

SYSTEMS BIOLOGY

Dynamic Pathway Modelling

Olaf Wolkenhauer

www.sbi.uni-rostock.de

December 6, 2006

This is a draft manuscript

While in the past it was more common to make preprints, algorithms and software freely available, authors have become more protective of their work. I do not know whether this is a reflection of the increased competition for recognition or the pressure to commercialize research results. By making these notes available I hope to receive information about errors, misconceptions, missing references or acknowledgment, and suggestions to improve the text. In the modern world of publishing one also feels a need to clarify that the author makes no warranty of any kind, expressed or implied, with regard to programs, algorithms, theoretical concepts, or the documentation contained in this manuscript. The author shall not be liable in any event for incidental or consequential damages in connection with, or arising out of, the use of the material provided here.

Contents

1	Preface	vii
2	Modelling Natural Systems	1
2.0.1	The Systems Biology Approach	3
2.0.2	Differential Equations	11
2.0.3	Dynamic Systems Theory	14
2.0.4	Dealing with Uncertainty	19
2.1	Cell Chemistry	23
2.2	Cell Signalling	26
2.3	Experimental Techniques	29
2.3.1	Gel Electrophoresis	29
2.3.2	Blotting	30
2.3.3	Scanning and Laser Densitometry of Western Blots	31
2.3.4	Quantification of Western blots - General Considerations	31
2.4	The Dynamic Proteome	33
2.5	Outlook	36
3	Biochemical Reactions	39
3.1	The Rate Equation Approach	39
3.1.1	Differential vs. Difference Equations	42
3.1.2	Numerical Simulation	43
3.2	Biochemical Reaction Modelling	44
3.3	Fundamental quantities and definitions	44
3.4	Basic Principles and Assumptions	46
3.5	Elementary Reactions	47
3.5.1	Monomolecular reactions	47
3.5.2	Bimolecular Reactions	49
3.5.3	Bimolecular reaction of identical species	51
3.5.4	Trimolecular reactions	52
3.5.5	Higher and rational reaction orders	54
3.6	Complex Reactions	55
3.6.1	Reversible Reactions	56
3.6.2	Parallel Reactions	59
3.6.3	Consecutive Reactions	62
3.6.4	Autocatalytic Reactions	68
3.7	Further Material to Enzyme Kinetics	70
3.7.1	Michaelis-Menten equation	71
3.7.2	The enzyme kinetic reaction with second intermediate	73

3.7.3	Constant substrate concentration	75
3.7.4	Interactions with other reaction mechanisms and reaction partners	75
3.7.5	Combination of catalyzed and uncatalyzed reaction	84
3.7.6	Reversible Michaelis-Menten equation	85
3.8	Activation and Deactivation of Proteins	87
4	Stochastic Modelling and Simulation	91
4.1	A matter of life and death	93
4.2	Mean values of Stochastic Processes	98
4.3	Discussion: ODEs and Stochastic Processes	100
4.4	Stochastic Simulation	102
4.4.1	Gillespie modelling	104
4.4.2	Stochastic rate constant versus rate constant	106
4.4.3	So are they, or are they not?	110
4.4.4	The Gillespie Algorithm	111
4.4.5	Examples	114
4.4.6	Molecules as individuals	119
4.5	An <i>ODE</i> to Differential Equations	120
4.6	A never ending story	127
4.6.1	Steady-state solution for the master equation	128
4.6.2	Temporal evolution of average and variance	135
4.6.3	Solution of the mass action model	137
4.6.4	Generating functions	139
4.6.5	Summary: The master-equation approach	149
5	Cell Communication	150
6	The Dynamic Systems Approach	151
6.1	Pathways as Dynamic Systems	151
6.2	The Role of Feedback	152
6.3	Tutorial Examples	154
6.4	Discussion	159
6.5	Phase-Plane Analysis	160
6.6	Nonlinear Dynamics	168
7	Receptor Modelling	175
8	Dynamic Modelling of Biochemical Networks	184
8.1	Simulation example	187
8.2	Michaelis-Menten modelling	188
8.3	Multinomial Systems	191
8.4	S-Systems	193

8.5	The Heinrich Model	194
8.6	The MAP Kinase (MAPK) Pathway	196
8.7	The Ras/Raf/MEK/ERK Pathway	198
8.8	Feedback and Oscillations in Signalling Pathways	202
9	Modules and Control Mechanisms	212
9.1	Linear Module	212
9.2	Hyperbolic Module	215
9.3	Sigmoidal Module	217
9.4	Robust or Adaptive Module	219
9.5	Feedback Systems	221
9.5.1	Positive Feedback/Feedforward - Switches	222
9.5.2	Negative Feedback - Oscillations	232
9.5.3	Mixed Control Mechanisms	238
10	Approximations and the Consequences	245
11	Cell Cycle	246
11.1	An Overview of the Cell Cycle	246
11.1.1	Mitosis - Key events	247
11.1.2	S-Phase	248
11.2	The cell-cycle Control System	248
11.2.1	Cell-Cycle Control System and Cyclically Activated Protein Kinase	249
11.2.2	Cell cycle and Cell growth	250
11.3	Model systems to study the cell-cycle control system	251
11.4	Modeling the cell cycle of yeast	252
11.4.1	Molecular components of the yeast cell cycle	252
11.4.2	Surveillance mechanisms, Checkpoints	254
11.4.3	Mathematical model	255
11.5	Numerical simulations	256
11.5.1	Wild-type cells	256
11.6	Mutations	258
11.6.1	<i>wee1⁻</i> - mutants	260
11.6.2	<i>wee1⁻ cdc25Δ</i> mutants	260
12	Metabolic Control Analysis	264
12.1	Fundamentals	265
12.1.1	Control and Regulation	265
12.1.2	Parameters and Variables	265
12.2	Control Coefficients	265
12.2.1	Flux Control Coefficient	267
12.2.2	Concentration Control Coefficient	268

12.2.3	The Summation Theorem	268
12.3	Metabolic Control Analysis and Rate-Determining Step	270
12.4	Elasticity Coefficients	271
12.4.1	Elasticity Coefficients for Metabolites	272
12.4.2	Elasticity Coefficients for Parameters	272
12.4.3	The Connectivity Theorem	273
12.5	Using the Summation and the Connectivity Theorem	273
12.6	Matrix Methods	275
12.6.1	Control Coefficients in an unbranched pathway	277
12.6.2	Control Coefficients in a branched pathway	279
12.7	Response Coefficients	280
12.8	Limits of Metabolic Control Analysis	280
12.9	Elasticity Coefficients for Specific RateLaws	281
12.9.1	Michaelis-Menten equation	281
12.9.2	Reversible Michaelis-Menten Equation	282
12.9.3	Hill equation	283
12.9.4	S-Systems	284
12.9.5	Elasticity Calculus	285
A	Glossary	287
B	Notation	295

1 Preface

A motivation not only for this text but for many if not most research projects in molecular biology, cell biology and bio-medicine is an improved understanding of disease mechanisms. We like to understand the causal basis of a disease that effects an entire organism. Taking a human body, it consists of organs, which in turn are build from cells. A biological cell is thus the basic building block for living systems and the focus of attention for an understanding of not only the cell but higher levels of structural and functional organization. Whether we are interested in cancer, neurodegenerative diseases like Parkinson's or Alzheimer's, metabolic diseases like diabetes or the malfunction of the immune system, modern life sciences tries to understand phenomena at the physiological level of an organism at the level of the cell. For example, a cancerous tumor is an uncontrolled proliferation of cells and it seems therefore logical to try and understand how a cell *functions*. The functions of a cell are its growth, differentiation (specialization), proliferation (division) and apoptosis (programmed cell death). In all of this the cell is interacting or communicating with its environment and other cells. Many drugs are designed to influence this communication, to suppress or stimulate the cell's behavior in a defined way. The generation of signals and transfer of information is achieved through biochemical reactions among molecules. The concept of a *pathway* is used to identify groups of molecules that interact in a specific way so as to realize the functions of cell. Intra- and intercellular interactions and cell functions are nonlinear dynamic processes. To understand how the cell functions requires therefore a theory of dynamic systems. We are going to present formal mathematical concepts that are the basis for an understanding of nonlinear dynamic interactions in cells.

The focus of this book is on systems biology, an emerging area of research that is a natural conclusion from the advances made in related areas, including genomics, molecular- and cell biology and bioinformatics. The areas of genomics and bioinformatics have identified and characterized many of the components that make up a living cell and maintain its function. In genomics the genetic information that is encoded in the genome is studied with respect to genes and the proteins they code for. A primary aim of bioinformatics has been to link genome sequences or genes with RNA products and proteins, i.e., to determine whether in a particular experimental context there exist a relationship between genes and proteins, amongst genes and proteins, and across genomes. The principal objective of modern life sciences is to describe the role of these components in developmental and disease mechanisms. While the developments in genomics and bioinformatics have brought tremendous advances in our understanding of molecular and cell biology, it is increasingly recognized that it is the temporal interaction amongst large numbers of molecules that determine phenomena observed at higher (metabolic, cellular, or physiological) levels. This dynamic or systems perspective and integrative approach (combining data from the genome, transcriptome, proteome, metabolome, ...) is considered in the area of research referred to as *Systems Biology*:

Systems biology investigates the functioning and function of inter- and intra-cellular dynamic networks, using signal- and systems-oriented approaches.

To understand the functioning and function of cells, systems biology addresses the following central questions:

How do the components within a cell interact to bring about its structure and function? (intra-cellular dynamics)

How do cells interact to bring about higher levels of structural and functional organization? (inter-cellular dynamics)

The functions of a cell do not reside in the molecules themselves but in their interactions, just as life is an emergent, rather than an immanent or inherent, property of matter. Although life, or the function of the cell arise from the material world, they cannot be reduced to it. Systems biology therefore signals a shift, away from molecular characterization and cataloguing of the components in the cell, towards an understanding of functional activity.

The term ‘systems’ in systems biology refers to systems theory, or more specifically, to dynamic systems theory. We here focus on dynamics and transient changes occurring within cells. These changes, which in most cases will be molecule concentrations, carry information and are at the root of cellular functions that sustain and develop an organism. The concept by which scientists organize these processes are *pathways*, i.e., *networks* of biochemical reactions. A pathway is an abstraction, a model, of an observed reality. The aim for us is to take the concept of pathways, from simple maps or graphs that name the components and indicate graphically and only roughly their relationship, towards a dynamic simulation of the interactions of proteins in a pathway. We are going to concentrate on signal transduction pathways for the examples given. However, it is important to emphasize that the methodologies used for modelling and simulation are generic, i.e., they are applicable to a wide range of processes related to intra- and inter-cellular dynamics. In fact, the mathematical concepts and techniques introduced here are widely used in various other areas, including engineering, physics, chemistry. Learning them as generic tools, has a number of advantages for the student who is interested in broad, interdisciplinary training.

The motivation for systems biology, and dynamic pathway modelling in particular, is that many neuro- and cancer-related diseases can be considered a failure of communication at molecular level. The area of *cell signaling* investigates the transmission of information from receptors at the cell surface to gene activation in the nucleus (intracellular dynamics) as well as the communication among cells (intercellular dynamics). Signals are relayed by means of biochemical reactions occurring in space and time and organized in pathways. We are going to investigate mathematical modelling and simulation of inter- and intracellular dynamics. Mihajlo Mesarović played an important role in defining the discipline systems biology. Already 1968 he wrote [Mes68]:

“In spite of the considerable interest and efforts, the application of systems theory in biology has not quite lived up to expectations. [...] one of the main reasons for the existing lag is that systems theory has not been directly concerned with some of the problems of vital importance in biology.” “The real advance in the application of systems theory to biology will come about only when the biologists start asking questions which are based on the system-theoretic concepts rather than using these concepts to represent in still another way the phenomena which are already explained in terms of biophysical or biochemical principles. [...] then we will not have the ‘application of engineering principles to biological problems’ but rather a field of *systems biology* with its own identity and in its own right.”

Since then there have been dramatic advances in technologies including, gene and protein expression assays, confocal microscopy, calcium imaging, and fluorescent tagging of proteins, which allow us to observe reactions in time and space. We should not ignore, the fact that as yet we have some way to go with regard to quantitative stimulus-response experiments that generate time series data suitable for conventional system identification

techniques. But even if the technologies are available possibly the greatest hurdle and certainly the reason why it is so attractive, is the human factor: advances in the life sciences will rely on experimentalists and theoreticians working closely together; they need each other.

One might argue that mathematical modelling in biology is anything but new. This is true although the emphasis may have been different, which could be summarized as follows:

1. Mathematical biology is most closely associated with:
 - (a) population modelling – modelling dynamics of infectious disease
 - (b) morphogenesis and spatial pattern formation
 - (c) evolutionary dynamics... a main reference is [\[Mur02\]](#).
2. Mathematical physiology is most closely associated with:
 - (a) cardiac rhythmicity
 - (b) modelling the circulatory system, respiration, blood, muscle
 - (c) hormone physiology
 - (d) renal physiology
 - (e) hearing and vision... a main reference is [\[KS01\]](#).
3. Systems biology is most closely associated with
 - (a) molecular and cell biology
 - (b) genetics and genomics (omics data)
 - (c) modelling metabolic and signal transduction pathways

The outline of the text is as follows. Chapter 2 provides an introduction to the subject area, including a discussion of the scientific approach and the role of modelling. The ‘novelty’ of systems biology is that it considers signal- and systems-oriented approaches to modelling and simulation of cell-biological and molecular systems. We are going to introduce the concept of a ‘system’ from a very general perspective which is then refined and adapted to fit the application under consideration. Systems biology considers dynamics, including transient changes of molecular concentrations and differential equations are therefore unavoidable. The first chapter provides a gentle introduction to the concept. For the theoretician it is essential to not only have a basic grasp of molecular and cell biology but also to appreciate the generation of data from experiments. Chapter 3 introduces the two basic modelling concepts for biochemical reaction networks: mass action models and the chemical master equation approach. We are going to provide a thorough discussion of the differences and similarities and on the way learn a number of important or useful mathematical techniques. With these tools at hand, in Chapter 6 this knowledge is applied to signal transduction pathways. The mixture of biology and mathematics, of basic and advanced material is deliberate. In interdisciplinary research it is important to be able to read a broad spectrum of literature and it is important to develop confidence for the experience that not everything can be understood after the first reading. The Appendix with its summary of mathematical notation used in the different chapters and a glossary of technical terms is an idea adopted from biological textbooks to help the reader in finding her/his way through the material. Throughout the text, the most important concepts and terms are indicated in the page margin at the place where they are introduced.

Rostock, December 6, 2006

Literature Review

Systems biology is an emerging area of research and which is truly an interdisciplinary area, combining various disciplines and areas of research. A consequence of this is that although there are already many relevant research journal publications, there are currently few suitable textbooks available. In trying to fill a gap with the present text, we should not suggest that it is possible to cover all aspects of systems biology in one book. Considering the large number of theoretical methodologies, experimental techniques and biological questions, it will be necessary to consult complementary literature. The targeted audience for the present text are graduate and postgraduate students and researchers from a range of disciplines. The aim is to make the text accessible to students and researchers who may be at different levels of their training/experience. Towards this end we are going to illustrate concepts with plots and line drawings wherever possible. Each Section will give numerous references to research publications and books. In addition, we here give a brief list of textbooks that could help the novice to complement the material presented here.

Although the edited volume, [BB01] was written as a textbook and provides a range of examples for models. It covers many methodologies and application areas, but is necessarily limited to brief introductions which do not allow a more comprehensive treatment of the mathematical basis of the models. Since it was written by practitioners it remains a valuable source book with motivating examples. The monograph by Davidson [Dav01] describes how embryonic development is the outcome of a vast spatial and temporal series of differential gene expressions, and how the control of these depends on a hardwired regulatory program built into the DNA sequence. Apart from few logical wiring diagrams, mathematical modelling and simulation does not play a role in this book. It does however provide a good example for theoreticians to understand the biological challenge related to regulatory systems that are involved in the development of an organism. The edited volume by Fall et al. [Fal02] comes closest to the present text, is well written with an interdisciplinary audience in mind and is broader in scope. Somewhat more advanced is the standard text in mathematical biology by Murray [Mur02]. It covers a vast range of mathematical techniques and biological examples. In fact, several older texts in the area of mathematical biology are ideal for studies in systems biology but unfortunately some of these texts are out of print.

[Kre93] is a standard reference in the engineering sciences and covers a large spectrum of basic mathematical techniques. There are excellent introductory treatments of differential equations available, including [BD01] to name only one. Those texts, written for an engineering undergraduate audience, have gone through various editions, are well illustrated and accessible to biologists. A more advanced but still introductory textbook is [HSD04], introductory texts focussing on nonlinear differential equations are [JS99] and [Str00b].

Mathematical modelling and simulation has been applied to metabolic pathways and a number of excellent texts are available, including, [CB04], [Voi00] for introductory material, whereas [Fel97] and [HS96] are more advanced texts, focussing on metabolic control analysis (MCA). The main difference between signalling and metabolic pathways is that for the latter we can concentrate on steady-states, which means that many problems are of algebraic nature and do not require the solution of differential equations. There are a large number of basic maths books aimed at the bio- or life science student. A good, short introduction to the mathematics that are required for any experimentalist are [Pho97] and [CB99], although they avoid differential equations and probability theory. For statistical techniques that are relevant for generating data, we refer to [QK02]. The books by Eason et al. [ECG80] and Batschelet [Bat79], although written for the life scientists, also introduce differential equations and other more advanced material. [MS99] is an introduction

to modelling of dynamic systems and is a good complementary text to the present one.

With regard to software tools, an important development for systems biology is the Systems Biology Markup Language (SBML). This standard provides a computer-readable format for representing models of biochemical reaction networks. SBML is applicable to metabolic networks, cell-signaling pathways, genomic regulatory networks, and many other areas in systems biology. It is an international effort to provide a free and open modelling language, supported by a large group of developers. The web-site www.sbml.org provides links to a number of software tools for modelling and simulation but also has a repository for SBML code of models published in the literature. These models are an excellent source for hands-on exercises.

For the theoretician or modeler, there are various excellent introductory textbooks for molecular and cell biology. Some of these books have gone through many editions and been translated into various languages¹. The comparison between mathematical and biological textbooks is striking. Biology textbooks are often heavy, large in size, rich in colorful illustration and images. A good mathematics textbook will have a couple of black & white line drawings but otherwise must appear rather dull and thin to the reader from the life science community. The complexity of systems dealt with and the level of abstraction used to describe such systems is in both areas very similar and yet there are very different means of representing information and generating knowledge.

A broad general introduction to modern life sciences is available, for example, through [P⁺01] and [Har01]. Focussing on the cell, the book by Alberts et al. [AJL⁺02] has become almost a standard text. For microorganisms, [MMP00] provides an excellent introduction and survey of microbiology. The book by Brown [Bro99] is an accessible introduction to the area of genomics. The biochemistry that underlies the reactions in pathways is covered by various books, including [SBT02] or [MVHA99]. The area of signal transduction is developing very rapidly, and there are few textbooks at introductory level available; on example is [Gom02]. For engineers and computer scientists the introductory text [TB04] provides a concise summary of the most important concepts and principles underlying modern life sciences research.

For the biologist who is interested in interdisciplinary research but whose school days instilled a dislike for mathematics, may find parts of the material presented here challenging. Throughout the text we are going to derive virtually all results in detail, rather than just presenting an equation. If the introductory maths texts, which we have described above, are not sufficient, we provide a very basic introduction to mathematical and statistical modelling as a complement to the present text, available from http://www.sbi.uni-rostock.de/data_handling.htm. Furthermore, we are going to encourage computational studies and simulations to ‘play’ with the ideas presented here. A collection of small programmes is available from www.sbi.uni-rostock.de.

¹The reader is advised NOT to consult a translation but always the original if it was published in English. The life sciences are not only an interdisciplinary but also international effort and it is important for the student to gain confidence in English for he will have to read the research literature, publish and present his work in English, whether one works in academia or in industry.

Acknowledgements

The research on the Ras-Raf-MEK-ERK signal transduction pathway has been conducted in collaboration with Walter Kolch from the Cancer Research UK, Beatson Laboratories in Glasgow, Scotland and Kwang-Hyun Cho from the School of Medicine, Seoul National University, South Korea. Mukhtar Ullah and Thomas Millat contributed material to several sections and produced the Matlab graphs. Wellstead contributed valuable, enjoyable discussions and pointed out errors. However, since they had not always all parts of the document available, all responsibility for expressed views and remaining errors is with me. Ulf Schmitz helped with \LaTeX and generated most of the PSTricks drawings.

This manuscript was produced using \LaTeX , specifically MiKTeX and WinEdt. Our thanks go to those who have contributed to these projects!

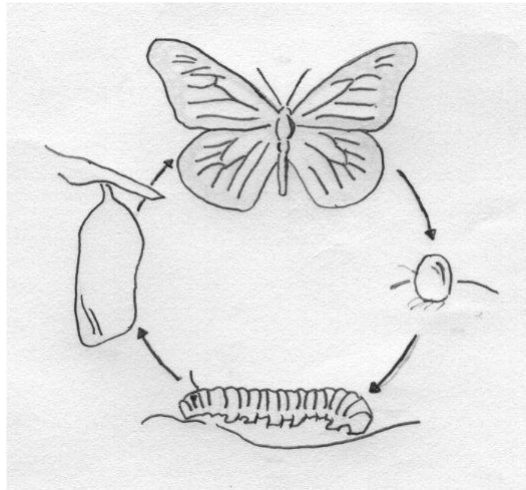


Fig. 2.1: The total metamorphosis of a butterfly is an example for one genomes realizing two proteomes. It is also a spectacular example for regulation, control and coordination in cell differentiation and the development of an organism.

2 Modelling and Understanding Natural Systems

One of the most spectacular example for the wonderful complexity and beauty in nature is the life cycle and total metamorphosis of the butterfly or dragonfly. During its morphological development the organism undergoes a dramatic transformation, in case of the dragonfly from an ugly predator, living under water, to a beautiful nectar eating insect. Nowadays it is argued that the information to enable this fascinating process is encoded in the *genome* of the organism; whereby the genome is understood as the entirety of the genetic information, encoded in a physical structure known as the DNA molecule. How is this process of the development of an organism realized? genome

What has been referred to as the *post-genomics era* of biology, is associated with areas of research that exploit the fact that we have now available the genome sequences for various organisms. The hope has been that using this information we should be able to understand observations at the cell-, phenotypic-, or physiological level. Zooming in from the entirety of an organism to the cell-level, we are covering an enormous scale of magnitude and quantity. While a human can reach heights of say 2.11m, a single cell has a diameter of only about ten micrometers. A human body may consist of some 10^{13} cells, where our largest organ, the liver consists of approximately 300 billion cells alone. The earth has only about six billion inhabitants, and without losing the enthusiasm for our research it is sometimes healthy to remind ourselves of the intellectual cosmos in which we are trying to travel.

Not surprisingly then, in many cases where there was an initial hope to discover a simple ‘gene/disease’ relationship, it was realized that what we are dealing here is a complex web of hierarchical, multi-leveled, regulated dynamic processes. These processes occur within cells and between cells and answering the question as to how a cell takes its place in higher levels of organization like tissues and organs, means we ought to ask questions about the communication and decision making in cells. The two central aspects of inter- and intra-cellular communication are therefore *signalling* and *gene expression*. signalling
gene expression

Studying intra- and inter-cellular communication requires sophisticated *technologies* to generate data. The complexity of the processes investigated and thus of the data, motivates the use of mathematics as an extension of common sense reasoning. In the words of Mike Mesarovic “Like you need a set of tools to do experiments, you need a set

of concepts to understand”.

The scientific approach, by which we wish to investigate questions like those mentioned above, is characterized by the attempt to *model natural systems*¹. An interesting aspect of interdisciplinary research is the diversity of perspectives and approaches individuals can contribute. The following story explains humorously differences among the modelers in systems biology. A University sends a philosopher, a biologist, a mathematician, a physicist, an engineer and a computer scientist to a hill walking trip in an attempt to stimulate interdisciplinary research. During a break, they rest on a bench, watching a cow in a field nearby. The philosopher asks “I wonder how one could decide on the size of a cow?”. Since the object under consideration is a biological species, the biologist responds first: “I have seen many cows in this area and know it is a *big* cow”. The mathematician is nervous about the lack of rigor and argues “The true volume is determined by integrating the mathematical function that describes the outer surface of the cow’s body.” The physicist realizes that this function is difficult to get and suggests an experiment: “You lift the cow into a completely filled swimming pool, and then measure the overflowing water, which corresponds directly to the volume of the cow, simple as that!” By now, the engineer had some time to think about the problem and suggests “Let’s assume the cow is a sphere...”. The computer scientist remained quite all along and is increasingly nervous: “Sorry mates, I thought my laptop wouldn’t work up here!”.

The underlying philosophy for the present text is to understand cell function through emphasizing transformations, processes over the objects. The technological developments in recent years have given us means to characterize the molecular components that make up a cell. For many researchers the function or biological role of a protein is largely defined by its three-dimensional structure. This obsession with molecular characterization has also led to the misconception of a gene as the causal agent for observations at the phenotype- or physiological level of an organism. The thrust of the present work is that it is systems dynamics that gives rise to biological function. A consequence of this is that the bioinformatics approach, i.e., mining large databases with information about the molecular characterization of components that make up the cell is necessarily limited. The physicist Erwin Schrödinger concluded that

“Life is an emergent, rather than an immanent or inherent, property of matter. Although it arises from the material world, it cannot be reduced to it.”

The relational character of our approach is also reflected in the following quotes attributed to the biologist Linus Pauling and universal genius Henri Poincaré, respectively:

“Life is a relationship among molecules and not a property of any molecule.”

“Science is built up of facts, as a house is with stones. But a collection of facts is no more a science than a heap of stones is a house. ”

In fact Poincaré apparently went as far as saying

“The aim of science is not things in themselves but the relations between things; outside these relations there is no reality knowable.”

¹A *natural system* is a system considered in the natural sciences, i.e., physics, chemistry, biology.

2.0.1 The Systems Biology Approach

The definition of a system as a set of objects with relations defined among those objects is not quite a proper mathematical expression and will require refinement. Instead of focussing on *things*, we are going to emphasize *processes*². The two things are not necessarily different: The existence or recognition of something as an object is at the same time a process. Appearances or phenomena are entities equivalent to the act by which they are apprehended. Likewise, a set is equivalent to the process or mapping that identifies elements of a set. A thing is identified by distinguishing it from something else.

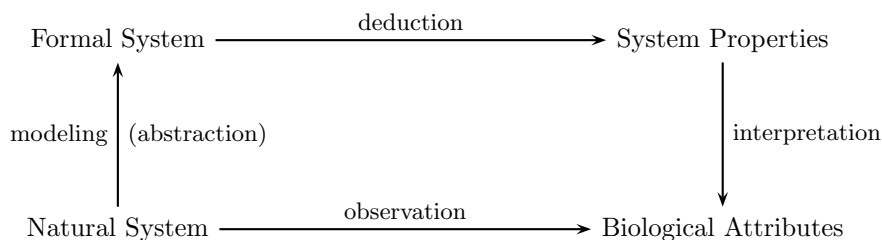


Fig. 2.2: Mathematical modelling is a process by which we establish an abstract representation of a natural system. For a model to be a valid representation, it is necessary to relate it with observable attributes of a biological system. Diagram adapted from [Mes68].

We are going to distinguish two kinds of systems: a natural system, which is an aspect of the phenomenal world under consideration and a formal system, which is the mathematical framework employed to represent the natural system (Figure 2.2). A formal system, theory or mathematical model is a collection of concepts. Establishing relationships between concepts is process of *modelling*. Mathematical modelling is the process by which we interpret a natural system using a mathematical model. Establishing the validity or truth of concepts requires the *a priori* concepts of space and time. This may not be obvious but if we, for example, consider the concept of a derivative, we find that any explanation of this using pen, paper or words, will use the ideas of movement, change, or rate of change. A mathematical model, or model for short, is thus a realization of the formal system employed. It should not escape our notice that the process of reasoning is a system itself. Since the world of objects is conditioned on the subject, science is not dealing with an independent reality. It is therefore desirable for us to look out for a conceptual framework in which not only to represent natural and formal systems but also the process of modelling itself. I believe this may be found in the context of category theory.

As simple as our definition of a system, as a set of objects with relations among these, may appear, since mathematics can be reduced to set-theoretic principles, our definition of a system is in fact as rich as the field of mathematics itself. Since we are going to derive every aspect of our understanding from this pair of a set and relation, we also realize the necessity or *a priori* nature of mathematics to establish truth in the phenomenal world.

That there is something else than the world of phenomena or “Wirklichkeit” which we can experience, can be seen from the fact that every argument we can establish has to have an absolute minimum of one premise and one rule of procedure, e.g. IF p , THEN q , before it can begin, and therefore begin to be an argument at all. So every argument has to rest on at least two undemonstrated assumptions, for no argument can establish either the truth of its own premises or the validity of the rules by which itself proceeds. We are therefore limited by the third class of objects that forms one of the roots for Schopenhauer’s principle of sufficient reason. The world as we know it, is our interpretation of the observable facts

²A process implies succession and so we are going to describe a dynamic system as a process, i.e., a sequence of events.

in the light of the theories we invent. The world of objects is thus conditioned by the subject: there is something that is grasped and something else that grasps it. In line with Kant and Schopenhauer, the entire world of phenomena or appearances is the world of representation in time and space, connected by causality. The world of phenomena is the self-objectification of the *noumena*. The noumena is what things are in themselves, or from our systems perspective, things in themselves are understood as things apart from *relation*. The existence of things independently of human representation or cognition, the unknowability of things, is the non-relational aspect of things. Schopenhauer described this as *will*. Before we further enter the slippery territory of philosophical arguments, we return to our definition of a system as a set of objects with relations, and refine it in mathematical terms.

With the definitions of sets and relations at hand, we rewrite our definition of a *general system*³, which is now considered as a relation on variables/indicators/items defined in set theoretic terms

$$\mathcal{S} \subset \times \{O_j\} , \quad (2.1)$$

or equivalently

$$\mathcal{S} \subset O_1 \times O_2 \times \dots ,$$

where the curly brackets denote a list or set, $j = 1, 2, \dots$ and \times denotes the so called Cartesian product, that is, a combination of sets. A *complex system* is a relation on systems/subsystems, i.e.,

$$\mathcal{S} \subset \times \{\mathcal{S}_j\} ,$$

such that there is a distinct behavior of the complex system while the integrity of the subsystem is preserved. The characteristics of a complex system include multilevelness and hierarchies. This definition of a complex system is more specific than the usual interpretation of molecular and cell-systems as being complex due to the fact that usually we are dealing with a large number of variables and nonlinear relationships among those. More generally we consider complexity

- a property of an encoding, i.e., the number of variables in a model.
- an attribute of the natural system under consideration, e.g., the connectivity, non-linearity of relationships.
- related to our ability to interact with the system, to observe it, to make measurement or generate experimental data.

A mathematical *model* is a realization of a formal system, an abstract representation of some aspect of the real world which is called the natural system under consideration. The process of modelling identifies and defines the variables and relationships among them. For our purposes, a mathematical model is subsequently identified with a set of parametric equations. The process of *simulation* is an ‘execution’ of, or ‘experiment’ with a model. For example, in case of differential equations models, a simulation refers to numerical integration, as the process of finding a solution to the set of equations. In case of mathematical models that use probabilities rather than direct values, a simulation run generates a *single* realization of the stochastic process.

The cell is a complex system and in studying the cell, we are considering *processes* and *components*. The components interact through or in processes to generate other components. We define the *organization* of a system as the pattern or configuration of processes. Pathways are therefore an example of describing organization in cells. The *structure* of

³The notion of a general system is due to Mihajlo Mesarovic, who developed the most general and most complete conceptual framework for general systems [MT75]. For a comprehensive overview of systems theory see [Kli91].

a system is the specific embodiment (implementation) of processes into material components. In modelling there is therefore a duality between abstract concepts and physical entities. Even the most basic concepts by which we make the world plausible, ‘space’ and ‘time’ have no material embodiment and the coexistence between the physical or material real and the mathematical or abstract should not be a real problem. In the present text we are trying to exercise this exciting aspect of modelling in the context of cell signalling, where the interactions of molecules lead to changes in protein concentrations, which define a signal that in turn carries or *transduces* information. We are going to consider two kinds of dynamics: *intracellular* dynamics and *intercellular dynamics*,. Intracellular dynamics are the processes by which the genome and proteome act, react and interact within the context of the cell so as to bring about its bio-physical structure and to realize its function (role) in a larger whole. Intercellular dynamics are the interactions of cells to realize higher levels of structural organization, including tissue or organs. It is obvious that both concepts are tightly coupled. Each cell has a copy of the entire genome and there are processes that run quite independently of other cells, cell-cell interactions are necessary for an individual cell to realize its role in a larger whole that is some tissue, an organ or the entire organism.

inter-/intra-cellular
dynamics

The key to successful modelling is that there has to be some correspondence between the causal structure of the natural system and the formal system. The decoding of a natural system into a mathematical model needs to be validated through an encoding that allows predictions about the natural system. In case of molecular systems modelling is hampered by *complexity*, and *observability*, i.e., the difficulties in making direct observations and measurements. Both, the complexity and observability of such systems lead to *uncertainty*.

With all these abstract formalisms, one may ask what the practical use of systems theory is. Although not simple, once it is mastered systems theory, mathematical modelling and simulation of dynamic systems provides a conceptual framework in which to discuss the following questions:

What is the influence inputs have on the system?

What do the outputs tell about the system?

What are the consequences of differences in the model structure?

More specifically, related to signal transduction pathways we wish to gain a better understanding of the following questions:

How does the physical interaction of molecules create signals, store and transmit information?

How are signals integrated and turned into decisions?

What is the role of space, location and translocation?

How are decisions, actions and reactions of cells integrated into higher levels of structure, organization and response pattern?

This discussion can proceed at different levels. An abstract but also most general discussion of properties of general systems, including issue such as coordination, hierarchies, multilevelness etc. can be conducted in the context of (2.1) [MT75]. We hereafter follow a different strategy. We first argue the case for differential equations as a suitable approach to represent signal transduction pathways and thereby concentrate on a special case of the

general system (2.1). Once we have shown that nonlinear state-space models are an appropriate modelling framework, we generalize our analysis of a particular pathway model to a class of pathways.

Let us first demonstrate how we get from the formal definition of a system as a relation of objects to some representation of dynamic systems. The purpose is not to frighten the reader with complicated mathematics but to show that there is a general systems theory which generalizes many of the models that have been published in systems biology. We are going to lower the level of abstraction immediately after this short excursion. Beginning with the general system

$$\mathcal{S} \subset O_1 \times O_2 \times \cdots ,$$

when (sub)systems interact they do this through defined interfaces, which may refer to as inputs and outputs. In cell signalling, membrane receptors suggest themselves as inputs that receive a stimulus in form of ligands binding to them. The expression level of a target gene may be defined as the response, respectively output of the pathway as a system. More formally, a signal transduction pathway may thus be described as the system

$$\mathcal{S} \subseteq \Omega \times \Gamma ,$$

where Ω and Γ related to the stimulus and response respectively. Most important for our philosophy of systems biology is that we understand a pathway not as a static graph but as a network of biochemical reactions, that is, a dynamic system which establishes a causal connection between stimulus and response. This means that we define a pathway formally as the mapping

$$\begin{aligned} \sigma : \Omega &\rightarrow \Gamma \\ \omega &\mapsto \gamma \end{aligned}$$

where a stimulus ω (respectively response γ) is defined as a temporal sequence of events

$$\Omega = \{\omega : I \rightarrow U\} , \quad \Gamma = \{\gamma : I \rightarrow Y\} .$$

At any point in time $t \in I$, our system \mathcal{S} receives a stimulus $u(t)$ to which it responds at time t with $y(t)$. We assume that stimuli and responses take their values from constant sets U respectively Y .

A key concept in the systems biological approach to cell signalling is that of a signal $\omega :]t_0, t_1] \rightarrow U$ acting on the system between time t_0 and t_1 , generating a response $\gamma :]t_0, t_1] \rightarrow Y$. In general, we denote the set of acceptable stimuli in terms of the mapping

$$\omega : I \rightarrow U ,$$

and for the response

$$\gamma : I \rightarrow Y ,$$

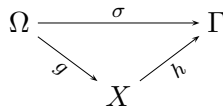
where for say concentrations we assume that data are described by positive real numbers such that $U \subseteq \mathbb{R}_+^m$ and $Y \subseteq \mathbb{R}_+^q$. For $m = 1$ and $q = 1$ the vector-valued notation reduces to a single signal or time series. If $I = \mathbb{Z}_+$ we have a discrete-time system model, which corresponds to the collection of experimental data, and in which case we could consider Ω and Γ as finite-dimensional vector spaces, encoding sequences of events

$$\begin{aligned} \omega &= (u(0), u(1), \dots) , \\ \gamma &= (y(1), y(2), \dots) . \end{aligned}$$

In modelling it often makes sense to assume a signal that is continuous in value and time. For $I = \mathbb{R}_+$ a continuous-time system, with $\omega :]t_1, t_2] \rightarrow U$ in Ω and $\gamma :]t_2, t_3] \rightarrow Y$ in Γ .

The entire sets of stimuli and responses that the cell can realize form the objects Ω and Γ of our definition of a stimulus-response system.

The description of a pathway σ as a mapping that takes an input sequence to a response pattern, is an external description, without consideration for the interactions that generate this input-output behavior. Extending our abstract model with a state-space X , we have the following state-space representation:



Dynamic pathway modelling is the process in which we identify and characterize the mathematical objects Ω , Γ , X as well as the mappings that put these objects in relation to another. At any point in time, the behavior of the system is thus encapsulated by the state $x \in X$. The temporal evolution of the state, $x(t)$, implies the existence of a state-transition map

$$\varphi : I \times I \times X \times \Omega \rightarrow X$$

whose value is the state $x(t) = \varphi(t; t_0, x, \omega)$, that is, an element of the state-space X . In this setting, the state x at time t arises from an initial state $x_0 = x(t_0) \in X$ at some initial time $t_0 \in I$ under the action of stimulus $\omega \in \Omega$. The graph of φ in $I \times X$, called the *trajectory* and describes the temporal evolution of the system. If we are to investigate a pathway or cell function in experiments we assume that, at least for the experiment, the system is *time invariant*, i.e.,

$$\varphi(t; t_0, x, \omega) = \varphi(t + s; t_0 + s, x, \omega')$$

for all $s \in I$. A dynamical system is *continuous-time* if I is a set of real numbers and *discrete-time* if I are integers. \mathcal{S} is *finite dimensional* if X is a finite-dimensional space and we speak of a *finite-state* system if X is a finite set. A finite system is more commonly known as a *automaton*.

Depending on the nature of the biological process under consideration but also motivated by personal preference and mathematical convenience, one can choose among a number formalisms by which to translate biological understanding into a mathematical model. To name but a few, we can distinguish between mass-action-, power-law or S-system-, and Michaelis-Menten models that are based on differential equations. The computer scientist tend to find the setting of automata theory, machines and formal languages (π -calculus, Petri-nets, process algebras) more preferable. For any formalism mentioned, one question is always the role or relevance of randomness. If the answer is that random fluctuations cannot be ignored, we are led to stochastic models, based on Markov-processes, Langevin-, or Chapman-Kolmogorov equations.

By far the most frequently employed approach is to represent temporal changes of protein concentrations as differential (rate) equations

$$\begin{aligned} \dot{x} &= V(x(t), u(t)) \\ y(t) &= h(x(t)) \end{aligned}$$

Here \dot{x} describes the rate of change in state vector $x = (x_1, \dots, x_n)$ at time t . Here V is again a map that in the context of dynamic systems theory is referred to as a vector field. The map h describes the observations we make on the states of the system. A special case of the above representation is a class of nonlinear systems models for which there exists some experience

$$\dot{x} = f(x(t)) + \sum_{i=1}^m g_i(x(t))u_i(t), \quad y_j = h_j(x(t)), \quad 1 \leq j \leq q.$$

Variable u is considered a state-independent external ‘control’-input to the system. Not surprisingly, most of the research around this formalism has emerged from the control engineering community.

We are going to consider a dynamic systems approach to understand inter- and intracellular processes. To identify a system’s behavior we require stimulus-response time course data. For most molecular and cell biological experiments it is not straightforward to generate sufficiently rich and quantitative data sets that satisfy the theoretician. The richest set of system-theoretic methods is available for *time-invariant linear* systems. Time invariance means that although the system variables change over time, the relationships among the variables do not. If we were to repeat an experiment the same mathematical relationships would be identified. The definition of linearity deserves attention as different scientific communities have different interpretations. For example, the standard model of a MAP kinase pathway is a *linear* cascade of three modules (cf. Figures 8.5 and 8.6). Linearity in this context refers to a *series* connection of modules. Any feedback loop branching of one of these modules and influencing a variable further up in the pathway is occasionally described as *nonlinear feedback*. This is rather unfortunate and should be avoided. Let $y_t(\theta, u_t)$ be the output of the model⁴ with parameters θ at time t and which is due to the input $u(\tau)$, which has been applied from initial conditions between time zero to t , $0 \leq \tau \leq t$, $t \in \mathbb{R}_+$. A model is said to be *linear in its inputs* (LI) if the outputs satisfy the superposition principle with respect to the inputs, i.e., if

$$\forall(\alpha, \beta) \in \mathbb{R}^2, y_t(\theta, \alpha u_1(t) + \beta u_2(t)) = \alpha y_t(\theta, u_1(t)) + \beta y_t(\theta, u_2(t)) . \quad (2.2)$$

A system is thus nonlinear if the output from the system is not proportional to the input. If we draw a graph of the output against the input on the abscissa, a linear system would define a straight line while a nonlinear system would diverge from the straight line. While this definition is common in engineering and applied mathematics, statisticians usually refer to a different kind of linearity: A model is said to be *linear in its parameters* (LP) if its outputs satisfy the following superposition principle with respect to its parameters:

$$\forall(\alpha, \beta) \in \mathbb{R}^2, y_t(\alpha \theta_1 + \beta \theta_2, u(t)) = \alpha y_t(\theta_1, u(t)) + \beta y_t(\theta_2, u(t)) . \quad (2.3)$$

For example, the simple straight line equation $y = \theta_1 x + \theta_2$ is LI and LP. We are going to return to a discussion of the difference between linear and nonlinear systems on page 168.

If \mathcal{S} is linear and time invariant, we can express the relationship between dependent and independent variables by the following equation:

$$y(t) = \sum_{k=0}^{t-1} \Theta_{t-k} u(k), \quad t \in T . \quad (2.4)$$

where $\Theta_t \in \mathbb{R}^{p \times m}$ denote the coefficient matrices which characterize the process and we have assumed a time-discrete system, i.e., $T \subset \mathbb{Z}_+$. For each t , (2.4) specifies a set of q equations in $q \cdot m$ unknowns of the matrix Θ_t . We find that for the linear system there exist a one-to-one correspondence:

$$\mathcal{S} \cong \{\Theta_1, \Theta_2, \Theta_3, \dots\} .$$

For the example of a linear discrete time-invariant system, the relationship between input $u(t)$ and output $y(t)$ is *linear*. Let $U \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^q$, $\Theta(t) \in \mathbb{R}^{p \times m}$, the system can then also be represented in a *canonical* form using matrices $F \in \mathbb{R}^{n \times n}$, $G \in \mathbb{R}^{n \times m}$ and $H \in \mathbb{R}^{n \times q}$

$$\Theta(t) = HF^{t-1}G \quad t = 1, 2, \dots$$

⁴Note that we slipped the time dependence from the brackets, $y(\theta, u(t), t)$ into the subscript y_t . This is to simplify the notation with no other meaning.

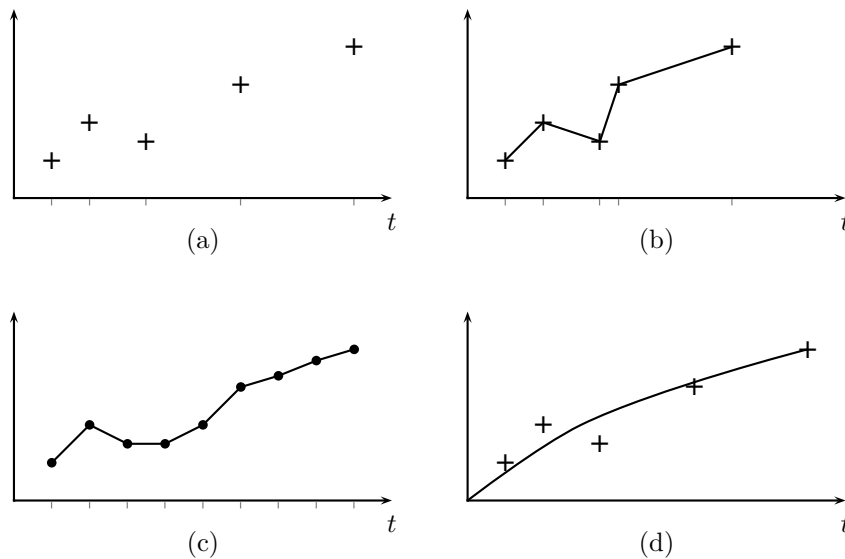


Fig. 2.3: Sampling and approximation of signals. (a) Experimental data define only points and in biological experiments measurements are frequently taken at equidistant time points. (b) The common representation of data as a line plot implies a model of linear interpolation, ignoring any possible random fluctuations. (c) For system identification and parameter estimation, it is usually assumed that measurements are sampled at equally spaced intervals. (d) In some cases it may be feasible to fit a curve through the points and model changes with a continuous-time model. It could however be the case that the dip at the third time point is not an outlier but rather an important biological aspect. Which assumption is correct depends on the context in which the data are generated.

The problem of modelling is then to define the dimension of X , for which the sequence $\Theta_1, \Theta_2, \dots$ is uniquely determined; leading to the discrete-time state-space model⁵:

$$\begin{aligned} x(t+1) &= Fx(t) + Gu(t) \\ y(t) &= Hx(t) \quad x(t_0) = x_0 . \end{aligned}$$

Given the present state $x \in X$ defined by $x(t)$ and input $u(t) \in U$ the map φ determines the next state and for every state x , the output map h determines an output $y(t)$. It is usually assumed⁶ that X is equal to or a subset of the Euclidean space of real numbers, $\mathbb{R}^n = \mathbb{R}_1 \times \dots \times \mathbb{R}_n$, and thereby any state can be represented as a point in X (see page 7). Note that the concept of state is a general notion, defining a set of n state-variables such that the knowledge of these variables at some initial point in time $t = t_0$ together with the knowledge of the input for $t \geq t_0$ completely determines the behavior of the system for any time $t \geq t_0$. State variables need not be physically measurable or observable quantities. The state-space representation is well established and forms the basis for automata theory and control theory. An *automaton* is a discrete-time system with finite input and output sets U and Y , respectively. We say the automata is finite if X is a finite set⁷. Automata theory has been used to model numerous systems including gene networks. However, we note that the finiteness of spaces, in which the inputs and outputs take their values, may require a quantization of measurements and discretization. With typically short time series and a lack of replicate measurements this may imply an unreasonable loss of information.

automaton

⁵Any control engineering textbook will provide further reading on properties of such systems, and how the matrices can be identified from experimental data.

⁶There are a number of mathematical requirements associated with the definitions and reasoning in this section. We leave these details for later chapters and refer to the extensive literature in mathematical systems and control theory, including for example [Son98], [Nv90], [Isi89] and [Bel90].

⁷The state of a linear dynamic system, continuous-time or discrete-time evolves in \mathbb{R}^n , whereas the state of an automaton resides in a finite set of *symbols*.

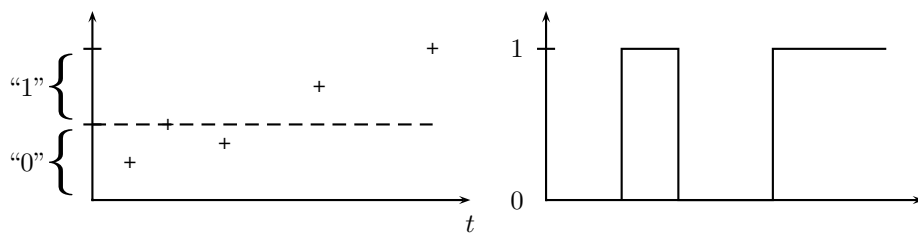


Fig. 2.4: Quantization of signals. Using a threshold on the amplitude of a signal, we can convert the data into a binary signal suitable for modelling boolean networks.

The state-space representation introduced above may look fairly general but there are more general cases to consider. For example, consider the state-space system

$$\begin{aligned} \dot{x}(t) &= V(x(t), u(t), w(t), t) , & x(t_0) &= x_0 \\ y(t) &= h(x(t), u(t), m(t), t) , \end{aligned} \quad (2.5)$$

where V can be changing with time and $\{m(t)\}$, $\{w(t)\}$ are stochastic processes. Representation (2.5) is usually too general for a detailed mathematical analysis of a specific model. The first step to a tractable model is by assuming that ϕ is not dependent on time, autonomous systems i.e., the system is *autonomous*⁸ or time-invariant

$$\dot{x}(t) = V(x(t), u(t)) .$$

Let us look at an example of the system above. In subsequent chapters, we are considering molecular populations that change as the result of chemical reactions. Under the hypotheses that all elementary reactions obey first-order kinetics and the compartment in which the reaction takes place has a constant temperature, the generalized mass action model is given by the set of coupled equations

$$\begin{aligned} \frac{dx_1}{dt} &= -\theta_1 x_1 + \theta_2 x_2 \\ \frac{dx_2}{dt} &= \theta_1 x_1 - (\theta_2 + \theta_3) x_2 \\ \frac{dx_3}{dt} &= \theta_3 x_2 . \end{aligned}$$

The structure of this mathematical model is given by prior knowledge or hypotheses about the system. The parameters θ_i are kinetic rate constants of the elementary reactions, and the state variables x_1 , x_2 , and x_3 are the concentrations of the reacting chemical species. All of them therefore have a precise meaning and interpretation⁹. These kind of models, describing *observed* changes, are therefore also referred to as *phenomenological* or *knowledge-based models*. Identifying the parameters of such a model from experimental data is called *parameter estimation*, and for nonlinear differential equations relies on sophisticated statistical tools. The decision of a model structure and parameter estimation, together are referred to as *system identification*.

One of the question that is going to arise is how we deal with stochasticity or randomness. When we are using differential equations we not arguing that the underlying physical

⁸The term *autonomous* is more frequently used in the context of differential equations, while *time-invariance* is more commonly used in the context of applications of differential equations to natural systems.

⁹In subsequent chapters we are going to use a different notation for biochemical reactions. The one based on x_i is commonly used in applied mathematics and in the context of differential equations. In biochemistry capital letters and square brackets, $[S_i]$, are used to denote the concentration of a molecular species S_i .

process is deterministic. One interpretation is that we describe the mean value of what is intrinsically a random process, arising from the interactions of molecules in the cell. There may however be situations in which we model a process involving only a relative small number of molecules in a larger volume. In this case we may prefer a stochastic model. Here again we have a selection of modelling approaches to choose from. Randomness of the states leads to *stochastic differential equations*:

stochastic differential equation

$$\dot{x} = V(x(t), w(t)) ,$$

If $w(t)$ is considered to a Gaussian process this representation is called *Langevin equation*:

$$\dot{x} = V(x(t)) + G(x(t))w(t) ,$$

where G is a matrix. We are not going to pursue this approach further but will look at Markov models and stochastic master equations in later chapters.

There are however also random fluctuations in measurements that are not related to the nature of molecular interactions but have to do the measurement itself. Adding an error model for such *measurement noise* $m(t)$ in the observation of the system gives us

measurement noise

$$\begin{aligned} \dot{x}(t) &= V(x(t), u(t)) , \\ y(t) &= h(x(t), u(t)) + m(t) , \end{aligned} \tag{2.6}$$

For measurement noise one usually assumes a Gaussian process $\{m(t)\}$, for which the joint probability distribution of $\{m(t)\}$ is multivariate *normal*. We are not going to consider such noise models. The reason is that at present it is rather difficult to validate such models with experimental data. This is however not to say that the consequences of perturbations to chosen parameter values are not an important part of the modelling process. To the contrary, an analysis of the sensitivity of the system's behavior on parameter variations is an important aspect of systems biology.

2.0.2 Differential Equations

While the physical object we focus on is the cell, the most fundamental process we consider is that of a biochemical reaction. In these reactions the concentration of a molecular species, referred to as the *substrate*, is changed. We are going to describe networks of coupled reactions with the help of equations. As an example let us consider the frequently occurring *Michaelis-Menten equation*:

$$V = \frac{V_{\max} \cdot S}{K_M + S} , \tag{2.7}$$

where the variable on the left-hand side of the equation is the *dependent variable* or 'output'. V_{\max} and K_M are fixed numbers, i.e., constants or parameters, while S denotes the concentrations of the substrate and is our *independent variable*. A graphical representation of equation (2.7) is shown in Figure 2.5. This equation, which is widely used in the biological sciences, is 'hiding' the fact that it is derived from a *differential equation*. Differential equations describe the rate of change of a variable over time, and are thus ideally suited to describe changes in concentrations in biochemical reactions. We are going to use differential equations extensively throughout and the present section serves as a gentle introduction to this most useful conceptual framework.

Pathways are the concept by which knowledge of interactions of proteins in cell functions is organized. A *pathway map* exhibits the names of the molecular components, whose interactions govern the basic cell functions. This leads us to a definition of pathways as

pathways

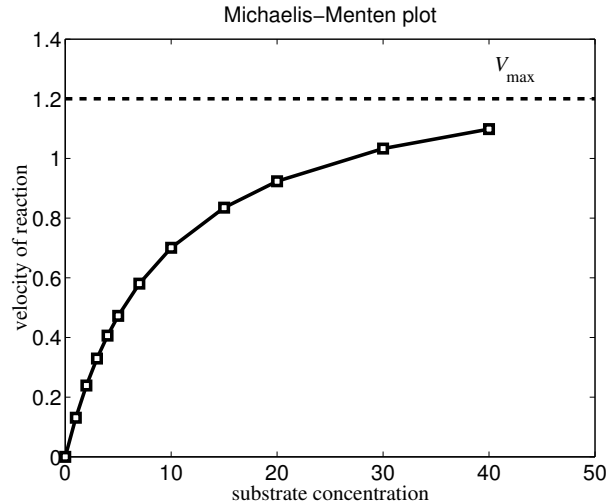


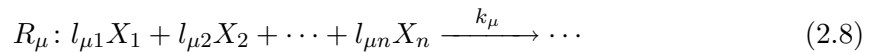
Fig. 2.5: Michaelis-Menten plot describing an enzyme kinetic reaction.

biochemical *networks*. One motivation for systems biology is to bring these static diagrams to life by modelling and simulating the biochemical reactions that underlie cell function, development, and disease.

Although a pathway or pathway map describes molecules, their physical state and interactions, it is an *abstraction* and has no physical reality. An *abstraction* is a *representation* of an aspect of the real world, so as to reason about it. A *model* is the consequence of this process, may it be by graphical means (e.g. the pathway map), by natural language, or through mathematical formalisms.

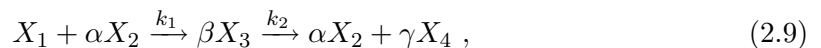
pathway map

A pathway map is thus a model, and the first step in the *art* of modelling is to identify which proteins need to be included in the model. One approach to support this process is to consider a pathway as a network of biochemical reactions. If we denote the chemical species and/or the modifications by capital letters, the following collection of biochemical equations would formalize a pathway:

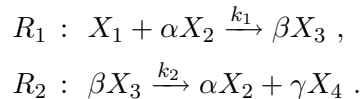


where X denotes a chemical species participating in reaction channel R_μ , the ‘+’ signs represents a combination, and the arrow a transformation. The coefficients $l_{\mu j} \geq 0$ indicate how many molecules of reactant species X_j are involved, and k_μ describes the rate at which the reaction proceeds. The reader unaccustomed to biochemical equations should not worry, we are going to introduce and discuss these representations in greater detail throughout the text.

While in the biochemical reaction equation participating components are denoted by X , in the mathematical equations we use small letters x to denote variables of the model. For example, the pathway



can be split into two *reaction channels*



When a reaction occurs, the changes to molecule populations is determined by the stoichiometric coefficients. For example, in reaction channel R_1 , X_2 loses α molecules, while in R_2 it gains α molecules.

The framework we are going to adopt to model biochemical reaction networks is that of nonlinear differential equations. I consider differential equations a *natural* choice for the following reasons. Causation is the principle of explanation of *change* in the realm of *matter*. For anything to be different from anything else, either *space* or *time*, or both have to be presupposed. Causation is a *relationship*, not between things, but between changes of *states* of things. As we going to demonstrate, differential equation models are an ideal means to realize this philosophy. Let consider the simplest of biological examples to demonstrate the view of causation as an explanation of change. Studying a protease cleaving peptide bonds in a substrate protein, we stipulate that

“The rate of proteolysis is somehow proportional to amount of substrate.”

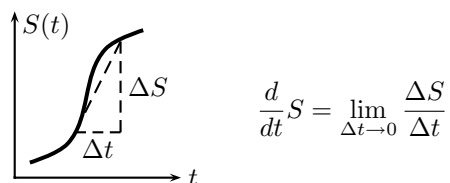
The purpose of mathematical modelling is to translate a hypothesis into a set of equations. These equations have parameter values, which we obtain from experimental data. In our example, a direct translation of the statement above is the following differential equation

$$\frac{d}{dt}S = -k_p S(t)$$

↳ variable (changes)

↳ parameter (fixed)

The operator d/dt is used to represent the *rate* of change of the substrate concentration $S(t)$ over time. The rate is thus also a velocity. As such it is related to the slope of the concentration profile, determined as the limit of Δt going towards zero. This is illustrated in the following diagram:



As indicated, we call a *parameter* a value of the model that does not change for the time interval of interest, while a *variable* changes. The former are typically rate constants while the latter are concentrations. Such *rate equations* are mathematically speaking differential equations with $dx/dt = \dots$ describing a rate of change over time. If the mathematical expression on the right-hand side is a “usual one”, not involving distributions like Dirac-distribution or the Heaviside step-function. These are then referred to as *ordinary differential equations* (ODEs). They consider changes only over time and not space. If diffusion across an area of the cell has to be considered we would end up with a description using *partial differential equations* (PDEs). Accounting for different regions of the cell, e.g., the cytosol and the nucleus, can be realized with ODEs by introducing different variables for the same protein, located in different regions. A more serious threat to the differential equation approach comes from the translocation of proteins, e.g., nucleocytoplasmic export. Time delays in a feedback loop more often than not have a significant effect on the dynamics of a system. An explicit representation of such phenomena leads to *delayed differential equations*. Needless to say that the theory for partial and delayed differential equations is more complicated than for ordinary differential equations.

rate equation
differential equations

In our example, the *state* of the system is fully determined by the equation that describes the value of the substrate S at any time t :

$$S(t) = S_0 \cdot e^{-k_p t}, \quad (2.10)$$

where S_0 denotes the initial concentration of S at time $t = 0$. This equations also called the *solution* to the differential equation above. Although there is also the “product”

concentration $P(t)$, the result of the proteolysis, its value can be directly determined from $S(t)$ using the conservation relation $S_0 = S(t) + P(t)$

$$P(t) = S_0(1 - e^{-k_p t}) . \quad (2.11)$$

The *simulation* of this mathematical model produces plots of (2.10) and (2.11) (Figure 2.6). As trivial as this example may seem, modelling arbitrary complex pathways is a straightforward extension of the approach demonstrate there.

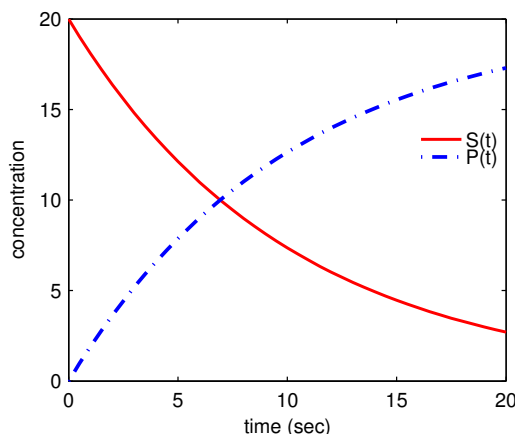


Fig. 2.6: Simulation of the differential equation model for the proteolysis.

Let us return to our pathway example (2.9) and consider the following set of ordinary differential equations with dynamic variables x_1, \dots, x_4 , corresponding to chemical species X_1, \dots, X_4 :

$$\begin{aligned} \frac{d}{dt}x_1 &= -k_1x_1(t)x_2^\alpha(t) \\ \frac{d}{dt}x_2 &= -\alpha k_1x_1(t)x_2^\alpha(t) + \alpha k_2x_3^\beta(t) \\ \frac{d}{dt}x_3 &= \beta k_1x_1(t)x_2^\alpha(t) - \beta k_2x_3^\beta(t) \\ \frac{d}{dt}x_4 &= \gamma k_2x_3^\beta(t) . \end{aligned} \quad (2.12)$$

Looking at the structure of these equations, we recognize the generalized representation for (2.8):

$$\frac{d}{dt}x_i(t) = \sum_{\mu=1}^M \nu_{\mu i} k_\mu \prod_{j=1}^n x_j^{l_{\mu j}}(t) \quad i = 1, 2, \dots, n \quad (2.13)$$

where the units of the concentrations x are mol per liter, $M=\text{mol/L}$. For simplicity, we omit the commonly used square brackets $[]$ to denote concentrations. The set of nonlinear differential equations (2.13) describes a large class of systems but is by no means the most general representation.

The aim of this section was to give a glimpse on how biological and experimental information is translated into mathematical expressions. We have omitted various questions but will return to these in due course.

2.0.3 Dynamic Systems Theory

The mathematical model (2.13), as general as it may seem, remains a particular choice for a conceptual framework in which to model biochemical reaction networks. We are

going to discuss various other approaches to represent intracellular processes. The choices involved in selecting one or the other approach are guided by practical or experimental considerations, mathematical convenience and personal preference. To guide a discussion of different approaches it is sometimes helpful to see whether modelling approaches can be generalized into a more abstract framework. An increased level of abstraction can mean that one loses the link of the model to a particular experimental setting but it can also help in the development and understanding of a mathematical model. Another motivation is that dynamic systems theory has developed various techniques and tools within other disciplines such as control engineering, applied mathematics and physics. By generalizing a biological model we can use these tools and techniques for our purposes. To illustrate this, let us consider a generalization of (2.13), which frees us from the discussion of how the right-hand side of the differential equation should be constructed. Towards this end we consider the ODE as a *mapping*

$$f : X \times \mathbb{P} \rightarrow \mathbb{R}$$

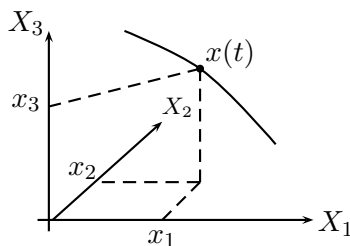
where X denotes the *state-space* and \mathbb{P} a parameter space. At any point in time $t \in I$, the concentrations are collected in a vector, called the *state* of the system and denoted

$$x(t) = (x_1(t), \dots, x_n(t)) .$$

The variables x_i are called state-variables and the space in which they take their values is the state-space. For a particular set of parameter values $\Theta \subset \mathbb{P}$, the current state $x(t)$ is associated with or mapped to a rate of change in the set of real numbers \mathbb{R} . Any particular parameter value is denoted as $\theta \in \Theta$. The model of ordinary differential equations (2.13) can then be generalized as

$$\dot{x} = f(x(t), \Theta) , \quad (2.14)$$

where \dot{x} is short for dx/dt and f is a n -valued mapping. The vector-valued mapping f determines the dynamics of the system¹⁰. If one imagines the state (vector) of concentrations as a *point* in state-space $X = \mathbb{R}_+^n$, the temporal evolution of the system describes a curve, called *trajectory*, in $X = X_1 \times \dots \times X_n$. The analysis of the dynamics of the system may thus be conducted in *geometric* terms, as illustrated in the following diagram:



The system (2.14) is said to be *unforced* (or *autonomous*) since there is no independent *input* $u(t)$ inside the bracket. We are also going to assume that rate coefficients are rate constants, i.e., the parameters of the system are not changing over time. One speaks of time-invariant systems, although their behavior is of course dynamic, changing over time. If there were an explicit dependence of the model on time, the brackets of the right-hand side of (2.14) should include a t .

Since the dynamics arising from (2.14) can be rather complex a common approach is to decompose a complex system into simpler subsystems or *modules* (see for example [TCN03, SJ04]). Under the headings of *bifurcation analysis*, *phase-space analysis*, *stability*

¹⁰In subsequent sections various variations of (2.14) will be discussed. The mapping f may alternatively be denoted V (for *velocity*). The parameter vector is often omitted despite the fact that it is always present.

analysis, reachability, observability, controllability, and realization theory, dynamic systems theory provides graphical and mathematical tools to predict the behavior of inter- and intra-cellular dynamics. The challenges for the modelling lie in the relatively large number of variables, in the inherent nonlinearity of interactions and the difficulties in getting quantitative time series data.

Graphical tools, such as phase plane and bifurcation analysis which help us to visualize the dynamic behavior of a system are restricted to systems in two or three dimensions. Even then, a mathematical analysis is focussing only on specific parts of the phase-plane and reduces in those areas the analysis of nonlinear dynamics to a linear systems theory. Although biological systems are for most case driven by nonlinear interactions we must acknowledge the fact that mathematical tools are for most cases only tractable if we revert to linear systems theory. This does not mean that we treat a nonlinear system as a linear one. Let us illustrate this approach.

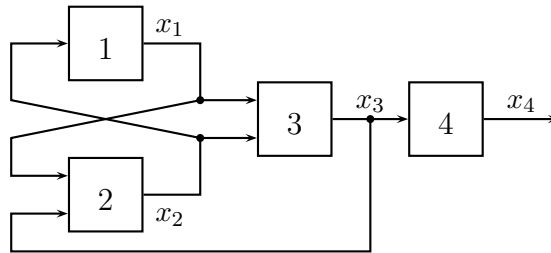


Fig. 2.7: Decomposition of the pathway model (2.12) into subsystem described by (2.15).

As can be seen from (2.13) the nonlinear properties of pathways are mostly determined from simple interactions among the system variables. One intuitive approach is therefore to decompose the network (2.14) into subsystems

$$\dot{x}_i = f_i(x_i(t), \bar{x}_i(t), \Theta) , \quad (2.15)$$

where x_i is now a scalar and \bar{x}_i is an input-vector consisting of the other state-variables x_k , $k \neq i$, of all other subsystems. Figure 2.7 illustrates the decomposition for the pathway model (2.12).

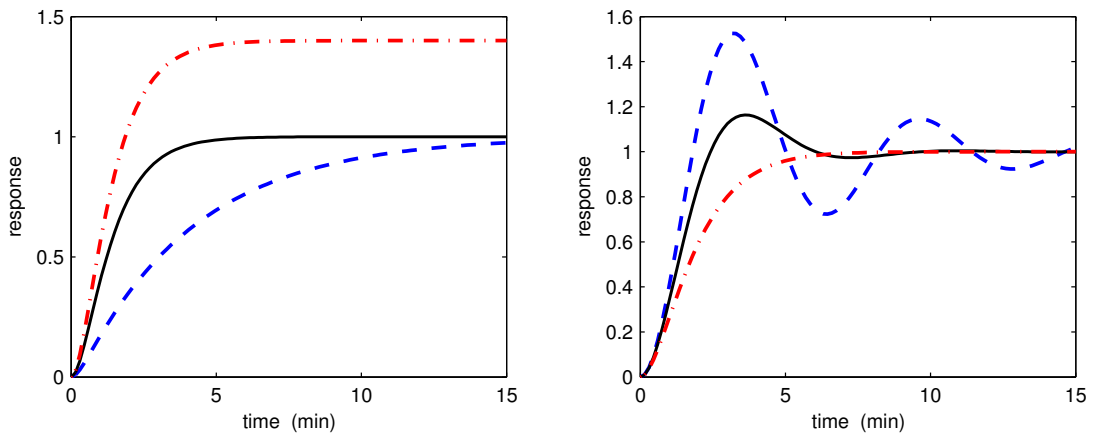


Fig. 2.8: The two plots illustrate the qualitative analysis that is a main purpose of dynamic pathway modelling. The aim is to establish the consequences or causes of parameter changes and/or the removal/introduction of feedback loops. The plots show typical changes in form of an acceleration/deceleration, amplification/suppression of signals.

Many techniques in systems theory focus on the long-term behavior of a system, as $t \rightarrow \infty$. For example, an important question is the stability of the system. Stability

theory helps us in this case to determine whether state variables remain within specified bounds as time proceeds. In dynamic pathway modelling we are however particularly interested in dynamic and short-term changes. For example, in cell signalling we wish to establish the response of a pathway to a stimulus of ligands binding to cell surface receptors. For signalling pathways, transient changes are of particular interest, while steady states are particularly relevant for metabolic pathways. The plots in Figure 2.8 show typical responses encountered in dynamic pathway modelling. The plot shows on the left monotonic responses, while on the right there are damped oscillations. Another important case are sustained oscillations as they are known from the cell cycle, calcium and glycolytic systems. More recently, sustained oscillations are also discussed in the context of cell signalling. In Figure 2.8 the systems depicted reach a *steady state* after about 15 minutes. A steady-state is thus reached when there are no observable changes in the concentrations¹¹. This observation corresponds to the mathematical condition where the rates of change on the left-hand side of the different equation are equal to zero:

$$0 = f(x(t), \Theta) . \quad (2.16)$$

For a given and fixed set of parameter values Θ the points in $X(t)$ for which this condition is met are called variously *steady-states*, *fixed points*, *critical points* or *equilibrium point*. This diversity in definitions is the result of various disciplines dealing with dynamic systems. For a new field like systems biology we thus have a choice of tools developed elsewhere but the different interpretations of the same mathematics may also be confusing to start with.

Stability analysis provides tools to characterize the dynamics of the system the transient behavior leading up to them and thereafter. It is then possible to predict whether the system will display cyclic changes of concentrations or whether they remain constant. Furthermore, it is possible to predict whether for small changes or perturbations the system remains stable with the critical point behavior or not. Bifurcation analysis is used to predict how system dynamics change as a function of parameter values or changes to them. For most practical purposes these analyzes are conducted by local linearization of the nonlinear system around critical points obtained from (2.16) and then use well established tools from linear systems theory. Ignoring the subscripts denoting a subsystem in (2.15), we linearize the system (2.14) around the critical point x^* (now including all x_i and x_k) by considering small perturbations to the steady-state

$$x(t) = x^*(t) + \Delta x(t)$$

where powers $(\Delta x)^p$, $p > 1$, are considered “very small” compared to Δx . This is indicated by the notation

$$(\Delta x)^p \doteq o(\Delta x) .$$

A classical technique to approximate a function around a point is by means of a Taylor series. Assuming $f(\cdot)$ is sufficiently smooth such that derivatives exist, the Taylor series expansion around x^* is given by

$$f(x(t), \Theta) = f(x^*(t)) + Df_{x^*} \Delta x(t) + o(\Delta x) ,$$

where

$$Df_{x^*} = \left(\frac{\partial f_i(x(t), \Theta)}{\partial x_j} \right) \Big|_{x^*} , \quad \text{with } i, j = 1, 2, \dots, n .$$

¹¹We are going to return to the question of steady-states and the assumptions involved throughout the text.

is the Jacobian matrix of first partial derivatives

$$Df_{x^*} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

of the mapping

$$f = (f_1(x_1, \dots, x_n), \dots, f_n(x_1, \dots, x_n)) .$$

The linearized system can now be written as

$$\Delta \dot{x} = Df_{x^*} \Delta x(t) , \quad (2.17)$$

where $\Delta x(t) = x(t) - x^*$ denotes deviations from the steady state. The Jacobian matrix is of particular interest. An off-diagonal element of the Jacobian describes the change of activation one variable, relative to another. Positive or negative entries correspond to *activation*, inhibition respectively *inhibition* of one variable by another¹². If we denote the Jacobian matrix by A and consider the diagonal matrix $\tilde{A} = \text{diag}(A)$, which contains the diagonal entries of A , then the decomposition of the linearized system (2.17) into one-component subsystems (2.15) is given by

$$\Delta \dot{x} = \tilde{A} \Delta x(t) + (A - \tilde{A}) \Delta \bar{x}(t) ,$$

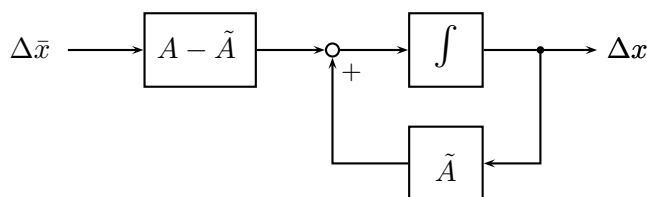
where the connections among linear subsystems are now denoted by

$$\Delta x(t) \doteq \Delta \bar{x}(t) .$$

In terms of individual subsystems this corresponds to

$$\Delta \dot{x}_i = a_{ii} \Delta x_i(t) + \sum_{j=1, j \neq i}^n a_{ij} \Delta x_j(t) ,$$

where the first term on the right-hand side corresponds to self-regulation and the last term corresponds feedback-type relations with other components. Interpreting $\Delta \bar{x}$ as an input and Δx as an output, we can study the behavior of the entire system, around the critical point x^* , using well established tools from systems and control theory (e.g. [SJ04]). In this setting the system can be represented by the following block-diagram:



Let us summarize what has been said so far. Beginning with the very simple biochemical reaction of proteolysis we introduced differential equations for modelling changes of molecular concentrations in the cell. We showed how sets of differential equations can be generalized and the dynamics analyzed using dynamic systems theory. From this brief introduction we can summarize the primary tasks in dynamic pathway modelling:

1. *Realization Theory*: Characterize model structures that could realize given stimulus-response data sets.

¹²The eigenvalues and eigenvectors of the Jacobian provide valuable information about the behavior of the dynamic system. There are many books available that describe this kind of analysis (e.g. [GH83, JS99, Str00b, HSD04]).

2. *System Identification*: Determine values for model parameters; using experimental data or simulation studies.
3. *Control Analysis*: Predict the consequence of changes to a pathway; in particular modifications to parameters, introduction and removal of feedback loops.

Figure 2.8 illustrated typical time course data and the kind of qualitative analysis the experimentalist is interested in. The main aim is to determine the causes of changes to parameters and the removal or introduction of feedback loops. The analysis is qualitative in the sense that exact values of the curves do not matter. More often we are interested in whether a response is ‘accelerated’ or ‘decelerated’, whether a signal is suppressed or amplified.

2.0.4 Dealing with Uncertainty

Uncertainty in modelling arises for various reasons, including randomness in the biological process and measurement noise from observations. For most practical situations we cannot expect the measured data to match the model perfectly. In other words, observations are usually subject to *random variations*. If we were to repeat the experiment, for a particular point in time t , we would obtain a set, called the *sample* of measurements. The purpose of statistics is to characterize this sample. The most intuitive approach to investigate a sample of data from a repeated random experiment is to visualize the *distribution* of the data in the *sample space*. Such a plot is called a *histogram*. In Figure 2.9 we show a frequency histogram. Dividing the heights of the bars by the total number of elements we obtain the *relative frequency histogram*. Dividing the relative frequency by the bar width, the total area sum of all areas the bars equals 1. This is then called *relative frequency density*. If the sample size is increased and the bar width reduced, the relative frequency density function approaches gradually a curve, called *probability density function*, denoted $p(x)$, where we used the letter x_i to denote an element of the sample space X . The area under the probability density function is equal to one, i.e., the probability that *any* value will occur is one.

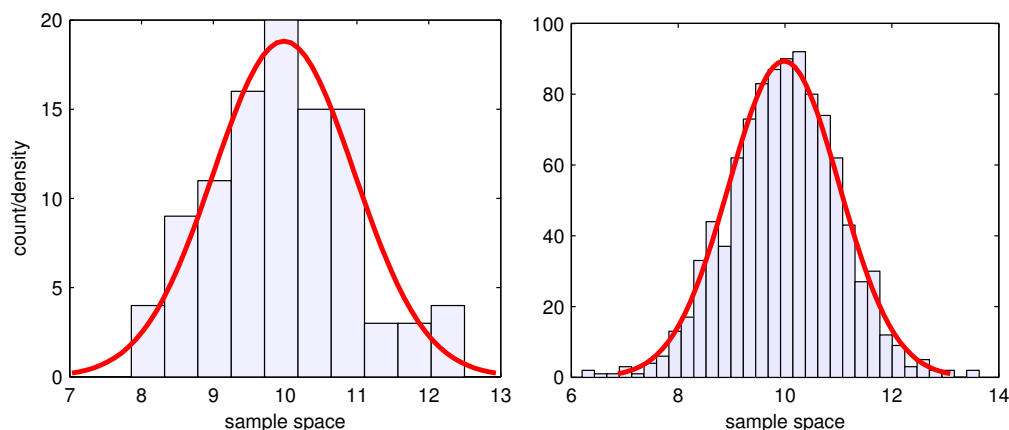


Fig. 2.9: As the sample size n , i.e., the number of repeated experiments increases, the shape of the histogram approaches a distribution which changes little and approximate the probability density function (solid line) from the sample values were drawn: On the left a sample of 100 values is used and 1000 for the plot on the right.

Note that a probability density function is a model of what we observe through statistics. We can therefore abstract from a sample to obtain a stochastic model of the process that is underlying the data. Instead of the histogram one could characterize the sample

(statistical model) or experiment (stochastic model) by some *characteristic*¹³ properties describing effectively the curve drawn by the histogram. For this we first consider the repeated data as referring to some real world variable (e.g. count of molecules at time t). If the measurement or observation is subject to random variations, it would make sense to speak of a random variable, say x . A description of a tendency for the data to cluster

mean value around a particular point, is called the *mean value*. From a statistical sample $\{x_i\}$ with sample average n elements, the mean is estimated as the *sample average*:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i .$$

variance Similar, a measure of variability around the mean value is obtained by the *variance*. The variance is an average distance to the mean value and given a sample, we can estimate the variance $\text{Var}[x]$ as

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

where the $\hat{\cdot}$ is to denote the fact that this is an estimate. The squaring is necessary to avoid an influence of the sign of the difference on the average. The problem is then that if the values are a physical measurement with a unit, the variance would not be in this unit. This problem can be solved to taking the square root of the variance, leading to what is

standard deviation known as the *standard deviation*

$$\text{Std}[x] \doteq \sigma .$$

We should remind ourselves of the difference of the mean value and the sample mean or sample average. One has to do with a statistical experiment and the other with a stochastic model of the process that generates such data. In this sense, statistics is the real-world interface for probability theory. The mean value may also be considered an *expected value*, if we are to repeat the random experiment many times we would expect this value, on average. If our random experiment can be modelled by a probability density $p(x)$, where each possible value x is effectively *weighted* by the distribution or density p , expectation we could define the mean value as the expectation, $E[x]$, of random variable x

$$E[x] = \sum_{i=1}^n xp(x_i) \quad \text{if } x \text{ is discrete,}$$

$$E[x] = \int xp(x)dx \quad \text{if } x \text{ is continuous.}$$

Similar, the variance can be defined as an expectation

$$E[(x - \bar{x})^2] = \int (x - \bar{x})^2 p(x) dx .$$

covariance For two random variables, x and y , the *covariance* is defined as

$$\sigma_{x,y} \doteq E[(x - \bar{x})(y - \bar{y})] .$$

If $\sigma_{x,y} = 0$, the two random variables are said to be *independent*. A bounded measure of correlation how two variables co-vary is the *correlation coefficient*

$$\rho_{x,y} \doteq \frac{\sigma_{x,y}}{\sigma_x \sigma_y} ,$$

such that $-1 \leq \rho \leq 1$. A positive correlation means that as one variable increases/decreases the other increases/decreases. In contrast, negative correlation means that as one variable increases/decreases, the other decreases/increases.

¹³In Section 4.6.4 we get to know this characteristic values of a distribution as *moments*.

Let us now instead of individual elements of X consider an *event* or set $A \subset X$ defined by its characteristic map. If we extend the concept of an expectation

$$E[1_A] = \int 1_A(x)p(x)dx .$$

This in effect determines the likelihood or probability of event A , and noting that taking the interval over X , the characteristic map 1_A in effect restricts the interval to A

$$\begin{aligned} 1_A : X &\rightarrow \{0, 1\} \\ x &\mapsto 1_A(x) = 1 \text{ if } x \in A, \text{ otherwise } = 0 . \end{aligned}$$

This leads us to an intuitive definition of *probability*:

probability

$$P(A) = E[1_A] = \int_A p(x)dx .$$

In terms of a sample of experimental data, it would seem plausible to define the probability then as the relative frequency:

$$P(A) = \frac{\text{number of outcomes in } X \text{ where } A \text{ occurs}}{\text{number of elements in } X} .$$

This is called the *relative frequency interpretation* for probabilities. There are however various other approaches to probability theory and it was the Russian mathematician Kolmogorov who put probability theory on a firm footing by linking probability theory to measure theory. The probability of an event A is then the measure of the area of the probability density function that overlaps with the subset $A \subset X$. For our purposes the mathematical details of these definitions are not of central importance and we refer to the vast literature on this. We do however note that a probability measure should satisfy the following *axioms*:

1. $0 \leq P(A) \leq 1$ for every event A .
2. $P(X) = 1$
3. $P(A \cup B) = P(A) + P(B)$, if A and B are mutually exclusive, i.e., $A \cap B = \emptyset$.

Remember that we started off by fixing a particular point in time and repeat an experiment. This generated a sample for which we now describe a stochastic model as illustrated in Figure 2.10.

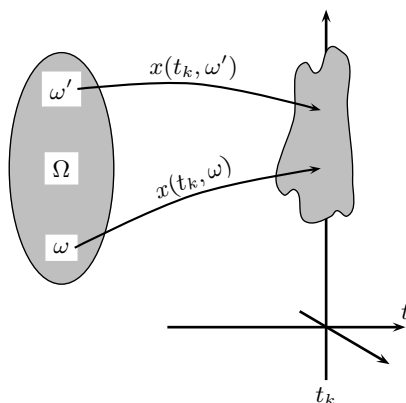


Fig. 2.10: Stochastic process $x(t, \omega)$ as a t -dependent random variable.

A further refinement of our random experiment is to distinguish between the random mechanism and an observation. Denote the sample space of an experiment with random sample space

outcomes as the set Ω . This consists of possible individual elementary outcomes $\omega \in \Omega$. These outcomes are mutually exclusive, i.e., only any one of the possible outcomes can occur. A collection of elements of Ω is called a *random event* and is denoted $A \subset \Omega$. We denote by $P(A)$ the probability that the event A will occur at each realization of the experiment. The collection of events or subsets of Ω is mathematically defined as a σ -algebra and denoted \mathbb{B} . The triple (Ω, \mathbb{B}, P) of a sample space, sigma algebra and probability measure is then referred to as a *probability space*. The variable which is associated with the random experiment, for example the measurement of a protein concentration at a particular point in time, is referred to as a *random variable*. If Ω is a continuous set, x is referred to as a continuous random variable and if Ω is a set of discrete outcomes, we speak of a discrete random variable and discrete probability distribution. A random variable x is a real-valued map defined on Ω such that for each real number α , $A_\alpha = \{\omega \in \Omega \mid x(\omega) \leq \alpha\} \in \mathbb{B}$. A_α is an event for which the probability is defined in terms of P . A random variable is neither random, nor variable, it is simply the mapping

$$x : \Omega \rightarrow X$$

$$\omega \mapsto x(\omega) .$$

Again we will drop the ω from $x(\omega)$ in most cases to simplify the notation, especially if x is a signal that is also a function of time. Since the experiment is associated with some random variable x , we write

$$p(\omega_i) = P(x = \omega_i) \text{ and } p(x) \text{ or } p_x \text{ for } P(x = \omega) ,$$

where p denotes the *probability distribution*, *probability mass function*, or *probability density function*. We use the term ‘mass function’ for discrete sample spaces and density function for continuous sample spaces.

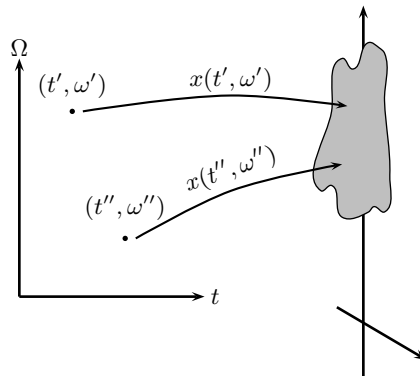


Fig. 2.11: Stochastic process $x(t, \omega)$ as a joint function of t and ω .

Talking of signals, considering a time set $I \subseteq \mathbb{Z}$, a time-varying process $x(t)$ is called a *random process* if for each t we cannot determine a precise value for $x(t)$, but instead have to consider a range of possible values with an associated probability distribution describing the relative likelihood of each possible value. More formally a *stochastic process* is a mathematical model of a random process, defined by the real-valued function

$$x : I \times \Omega \rightarrow X$$

such that for each fixed $t \in I$, $x(t, \omega)$ is a random variable. A stochastic process (Figure 2.11) is subsequently a sequence of t -dependent random variables

$$x : I \times \Omega \rightarrow X$$

$$(t, \omega) \mapsto x_t .$$

For each fixed $\omega \in \Omega$ the mapping from index set I into X describes a sequence of vectors $x_t(\omega)$, which is called a *realization* or *sample function* of the process. More commonly, we refer to the realization of a stochastic process as a *time series*, i.e., a sequence of observations and for which an observation at time t is modelled as the outcome of a random variable. The collection of all possible realizations is called the *ensemble*. All elements $x_t \equiv x(t)$ of a stochastic process $\{x(t)\}$ are defined on the same probability space. A stochastic process is in principle described by the joint distribution functions of all finite subcollections of $x(t)$'s but since these distributions will usually be unknown, most approaches will restrict themselves first and second order moments of the distributions, i.e., means, variances and co-variances.

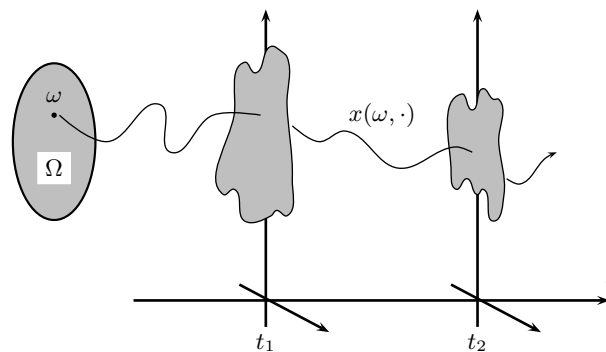


Fig. 2.12: Realization of the stochastic process $x(t, \omega)$, ω -dependent random variable.

We referred to a stochastic process as a model of a random process and should add that a stochastic model can take various forms. A dynamic model that determines or predicts for any t a precise value $x(t)$ is called *deterministic* and any model that accounts for random variations is called *stochastic*. A differential equation model such as equation (2.7) is thus deterministic. If we however add an additive noise term to this ODE model, we could refer to this as a stochastic model. In the application of dynamic systems theory to molecular and cell biology one generally has to make a decision whether to regard the process as a deterministic nonlinear system but with a negligible stochastic component or to assume that the nonlinearity is only a small perturbation of an essentially stochastic process. A theory of nonlinear stochastic processes has so far not found many applications to time-series analysis. We find that there are a large number of alternatives and combinations and the choice to whether account for uncertainty in the model or not will depend on various issues, including convenience and realism amongst others. This discussion how to translate a biological process into a mathematical model makes applied mathematics so interesting. The reader should avoid getting confused by the many alternatives and assumptions one can consider and see them as part of the *art* of modelling. The great artist Pablo Picasso once said “Art is a lie that makes us realize the truth”, to which we might add that science is the art that makes us realize reality.

The present section introduced the basic toolkit to describe systems as sets of objects with relationships defined on them. For the rest of this text all we do is to further refine and extend the concepts introduced here without actually adding more fundamental concepts than those introduced here.

2.1 Cell Chemistry

The cell is the basic building block of which higher organizational levels such as tissues and organs and entire organisms are composed. This chapter is to review some basic concepts from molecular- and cell biology. The text can however not even cover the bare

minimum of the information and a comprehensive book such as the standard text [AJL⁺02] is strongly recommended as a reference.

The cell is a rather complex environment, consisting of many different components. Because cells are about 70% water, life depends mostly on aqueous chemical reactions¹⁴. These reactions occur between *molecules*, where a molecule is a cluster of *atoms*, held together by so called *covalent bonds*. The weight of a molecule is its mass relative to that of an hydrogen atom. The mass of a molecule is specified in *Daltons*, 1 Da being an atomic mass unit approximately equal to the mass of a hydrogen atom.

$$\text{moles} = \frac{\text{weight}}{\text{molecular weight}} \quad (\text{a quantity})$$

One mole, 1 M, corresponds to $N_A \doteq 6.022 \cdot 10^{23}$ molecules of a given substance. N_A is referred to as the Avogadro's number. The molarity of a solution is defined by a concentration of 1 mole of the substance in 1 liter of solution:

$$1 \text{ molar} \equiv 1M \equiv 1 \frac{\text{mol}}{\text{L}} \quad (\text{a concentration})$$

For example, 1 moles of glucose weights 180 g; a molar solution, denoted 1 M, of glucose has 180 g/L. If we dissolve 1 mol in 0.5 liters, we have 2 M solution, although the amount of substance is the same.

If molecules are clusters of atoms, held together by bonds, these bonds can be broken by violent collisions amongst molecules. Average thermal motion at normal temperatures does not break these bonds and thus the deliberate breaking and making of bonds is the fundamental process that determines the concentrations of chemical species in a reaction and subsequently cell function. This process requires energy to take place and is carefully controlled by highly specific catalysts, which in biological systems are called enzymes.

kinetics How fast a reaction occurs is a matter of *kinetics*, defined by the rate of a reaction. In general, energy is the ability of a system to perform work, which is also why one speaks energetics of *energetics* in this context.

catabolic pathways There are two principle types of biochemical reactions: *catabolic pathways*, breaking down foodstuff and thereby generating energy and smaller building blocks. Secondly, anabolic pathways biosynthetic or *anabolic pathways* use energy to synthesize molecules. Both sets of reactions together constitute what is called the *metabolism* of the cell.

enzymes Apart from water, nearly all molecules in a cell are based on carbon. Carbon-based compounds are used in the cell to construct macromolecules, including the nucleic acids (DNA, RNA), and proteins. Proteins are particularly versatile, having various roles in maintaining the function of a cell and the organism as a whole. Many proteins serve as *enzymes* that are catalysts that control kinetic (bond-breaking and -making) reactions. Other proteins are used to build the structural components that make up the cell, or they act as motors and produce force and movement. Enzymes catalyze reactions by binding one or more ligands which are also called substrates, and converting them into one or more chemically modified products, without changing themselves. Enzyme-catalyzed reactions happen faster by a factor of a million or more than a non-catalyzed reaction. They are therefore an important mechanism by which the cell can respond to changes and regulate its functions. A typical enzyme will catalyze the reaction of a thousand substrate molecules every second. The enzyme therefore requires sufficient amounts of substrate around it. The motion caused by collisions and thus heat energy ensures that molecules are rapidly

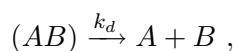
¹⁴There are alternative views that emphasize a gel-like character of the cell [Pol01]. The issue of what the inside of a cell is like should be important to us in modelling the interactions of molecules. In a somewhat brave act of modelling we later consider molecules as floating around as if they were in a gas.

moving about a confined area but can also move (diffuse) wider distances. The cell is a crowded environment and yet a small organic molecule can diffuse the entire distance across a cell in a fraction of a second.

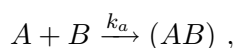
Enzymes move much more slowly than substrates, and the rate of encounter of each enzyme molecule with its substrate will depend on the concentration of the substrate molecule. For example, an abundant substrate may have a concentration of 0.5 mM and since water is 55 M, there is only about one such substrate molecule in the cell for every 10^5 water molecules. Nevertheless, an enzyme that could bind this substrate would collide with it about 500,000 times a second.

The biological properties or function of a protein is determined by its physical interaction with other molecules. The substance that is bound by a protein is referred to as a *ligand* for that protein. In cell signalling, ligands binding to membrane bound receptors provide the stimulus for intracellular reactions that transmit information to the genome of the cell. Antibodies, or immunoglobulins, are proteins produced by the immune system in response to foreign molecules. A specific antibody binds tightly to its particular target (called an antigen), and thereby inactivates it. Antibodies can therefore be used in experiments to select and quantitate proteins. For example, considering a population of antibody molecules which suddenly encounter a population of ligands, diffusing in the fluid surrounding them. The frequent encounters of ligands and antibody will increase the formation (association) of antibody-ligand complexes. The population of such complexes will initially increase but eventually complexes will also break apart (dissociate). Eventually, a chemical equilibrium is reached in which the number of association events per second is equal to the number of dissociation events. From the concentrations of the ligand, antibody and the complex at equilibrium, one can calculate the equilibrium constant K_{eq} of the strength of binding. The same principle described here for antibodies, applies to any binding of molecules.

We are going to use capital letters to denote molecular species, e.g., A , B , ERK, MEK. A complex formed from proteins A and B is denoted either AB , $A - B$ or A/B . If the molecules are not referring to particular names, like A and B , we usually write AB for the complex. For known proteins, e.g., Ras* and Raf we write Ras*/Raf. In some cases the protein complex gets a separate name, e.g., for the MAPK/ERK complex we write MEK. Considering a reversible reaction $A + B \leftrightarrow (AB)$, for dissociation the reaction diagram is



where the *dissociation rate* equals the product of k_d and the complex concentration (AB) . Note that in the reaction diagrams the symbols denote molecular species while in mathematical equations we use square brackets to distinguish concentrations from counts of molecules. For the association of molecules,



the *association rate* is the product of k_a , A and B . At equilibrium,

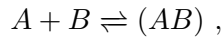
$$k_a A \cdot B = k_d (AB)$$

and which leads us to the definition of the *equilibrium constant*

$$K_{\text{eq}} = \frac{(AB)}{A \cdot B} = \frac{k_a}{k_d} . \quad (2.18)$$

The equilibrium constant of this particular reaction has a unit of liters per mole. The larger the equilibrium constant, the stronger the binding between A and B . Taking an example

which considers molecule numbers rather than concentrations [A_{JL}⁺02], considering 1000 molecules of A and 1000 molecules of B , with concentration 10^{-9} M. For $K_{\text{eq}} = 10^{-10}$ of the reversible reaction



there will be 270 A molecules, 270 B molecules and 730 (AB) molecules. For a reduction in binding energy of 2.8 kcal/M, reducing the equilibrium constant to $K_{\text{eq}} = 10^{-8}$, there will be 915 A molecules, 915 B molecules, and 85 AB molecules. For every, 1.4 kcal/M of free energy drop, the equilibrium constant increases by a factor of ten. Note that for the system to be in equilibrium, there is no flow of mass or material. In later sections we introduce the concept of *steady state*, for which changes in concentrations are zero. In dynamic systems theory, a steady state is sometimes also referred to as an equilibrium so that there is a risk for confusion. For the biochemist a biological system in equilibrium is dead.

2.2 Cell Signalling

For cells to combine into networks that realize higher levels of organization, including for example tissue and organs, it is necessary for them to communicate, exchange information. The basis for this *intercellular* signalling are the receptors in the cell membrane. The transmission of extracellular information to the genome is referred to *intracellular* signalling. Inter- and intra-cellular information effects the transcription of information from the genome and the synthesis of proteins. For more comprehensive discussions of cell signalling see [Dow01, Han97]. The glossary on page 287 provides help with unfamiliar terminology.

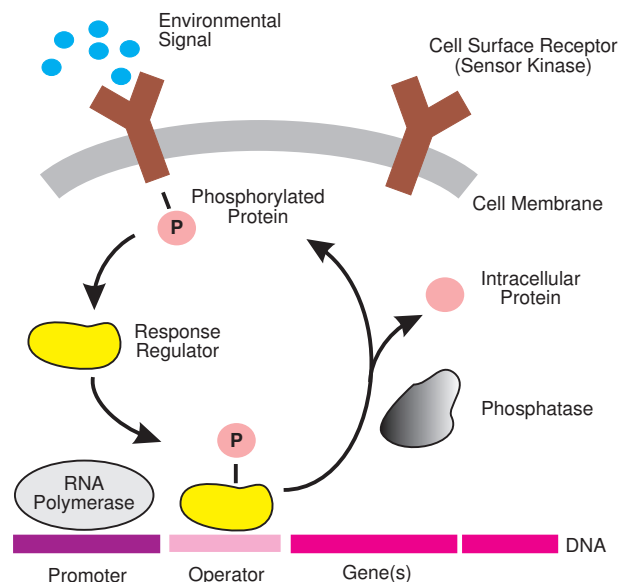


Fig. 2.13: A drastic simplification of intra-cellular signalling. Extracellular stimulation of the receptors is transduced into the cytoplasm. A series of biochemical reactions transmits the signal towards the genome, where the transcription of genes can be affected as a consequence of receptor stimulation.

The transmission of information is realized by chemical reaction networks, called pathways. Signals, passing these networks, are realized through changes in concentrations. The cell membrane and the nucleus in eucaryotic cells form physical barriers. There are principally two ways to pass these barriers - through active transport of molecules

passing through the cell surface (e.g. via pores or gap junctions) or nucleus or via signal transduction, i.e., receptor stimulation and phosphorylation as a means to transmit information without the movement of molecules. We may refer to these two modes of signalling as the “radio” versus “courier” mode of signal transmission. The location of a signalling molecule within the cell affects the interaction with other proteins and hence the movement of molecules to different cellular locations, called *translocation*, influences the dynamics of a signalling pathway.

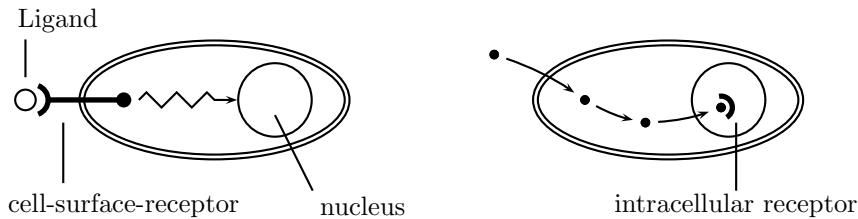
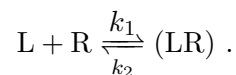


Fig. 2.14: Left: Most commonly receptors are bound to the transmembrane, where they bind an extracellular signal molecule (Ligand). Right: Small signaling molecules can enter the cell where they activate receptors inside the cell.

We are here going to focus on receptor-ligand signalling where extracellular molecules that bind to receptors in the cell membrane are referred to as *ligands*. Extracellular signalling molecules include hormones, cytokines and growth factors. Usually extracellular signals are found at very low concentrations, in the order of 10^{-8} mol/L [Han97]. Binding to receptors is highly specific to particular ligands. Not all ligands that bind to receptors result in the activation of that receptor. Ligands that bind to receptors and thereby prevent activation, are called *antagonists*. Denoting the ligand with a ‘L’ and the receptor with ‘R’, ligand binding to monovalent receptors with only one binding site can be described as a reversible reaction



The ratio of the concentrations, where 50% of the ligands are bound to receptors, is defined as via the *dissociation constant*

$$K_d = \frac{R \cdot L}{(LR)} = \frac{k_2}{k_1} .$$

The lower the K_d value, the higher the affinity of the receptor for its ligand. Generalizing¹⁵ the principle of *mass action* a mathematical model of monovalent receptor binding is given by the equation

$$\frac{d(LR)(t)}{dt} = k_1 L(t)R(t) - k_2 (LR)(t) ,$$

where k_1 ($M^{-1}sec^{-1}$) describes the rate constant of the receptor-ligand interaction and k_2 (sec^{-1}) describes the rate constant of the breakdown of the ligand/receptor complex LR . Solving this differential equation provides us with an equation for the temporal evolution of the ligand/receptor complex $(LR)(t)$. We got a bit ahead of ourselves here by jumping to this differential equation. In subsequent sections we are going to discuss ways to establish

¹⁵The term ‘generalized principle of mass action’ indicates the fact that we are not strictly using this balance principle. For various reasons, including the indirect measurements in cell signalling, the differential equation models described here are in most cases *phenomenological* models rather than exact representations of physical interactions among molecules.

such differential equation models and compare it with alternative formulations¹⁶. We return to receptor modelling in Section 7.

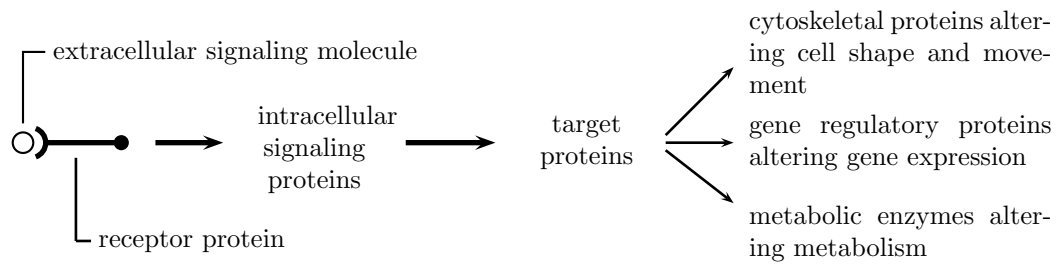


Fig. 2.15: Basic molecular components involved in intracellular signaling.

The consequence of signalling through ligand-binding is in most cases a modification of the activity of intracellular enzymes or activation factors (e.g. transcription factors that determine the reading or *transcription* of information encoded in the genome). A change in enzyme activity is achieved through a change in its conformation (three-dimensional structure). The altered spatial arrangement of the active site (amino acids) reduces or increases the protein's catalytic action and binding to substrate.

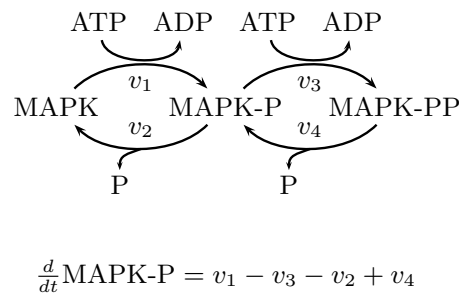


Fig. 2.16: Phosphorylation and dephosphorylation steps in the MAPK pathway.

One of the most common ways to alter the spatial arrangement and hence the properties of a protein is by adding of one or more phosphate groups, a process known as (de)phosphorylation *phosphorylation*. The enzymes that catalyze protein phosphorylation are known as *protein kinases* or kinases for short. The reverse process of dephosphorylation is catalyzed by phosphatases. Protein kinases and phosphatases are signalling molecules which catalyze the transfer of a phosphate group supplied by adenosine triphosphate (ATP) to and from target proteins respectively. The transfer ends with the release of adenosine diphosphate (ADP). The transfer of the phosphate group occurs only at specific *binding sites*, i.e., specific locations or amino acids of the target protein. The amino acids in the primary sequence of the polypeptide at which phosphorylation takes place are *serine*, *threonine*, and *tyrosine*. The kinases are grouped according to which amino acid they are specific for. Tyrosine protein kinases catalyze the phosphorylation of tyrosine residues, while serine/threonine protein kinases catalyze the phosphorylation of serine or threonine residues. Some protein kinases (such as MAPK) can act as both, tyrosine and serine/threonine kinases. Phosphorylated residues in a protein can act as binding sites for specific recognition domains in other proteins. A domain in a protein is a self-folding unit with a particular sequence and conformation. Certain domains allow proteins to recognize each other. Phosphorylation is thus a mechanism by which protein complexes can assemble. This results in

¹⁶Specifically for mathematical models of receptor binding the work of Lauffenburger and colleagues is notable [LL93].

a change of the localization or activity of enzymes. Phosphorylation/dephosphorylation is a good regulatory mechanism since it can occur in under one second [Han97]. Furthermore, the activation of a single kinase molecule results in the phosphorylation of many enzymes and therefore result in an *amplification* of the intracellular signal.

amplification

While there are a vast number of proteins involved in signalling, many of the proteins are similar in the sense that they consist of components (domains, modules, motifs), some of which are found in many otherwise different protein molecules. At the amino acid sequence level, this similarity is expressed as *homology*. It is therefore in some sense not just the protein as such but particular aspects of it which determines its role in signalling. Because the properties of proteins and hence their ability to interact depends on whether they are in a phosphorylated or unphosphorylated state, in mathematical modelling we are going to introduce two variables for each of the states.

homology

In response to signals important cell functions are influenced. These include

- cell death (apoptosis)
- cell growth (proliferation)
- specialization (differentiation)
- stress response
- cell cycle control

Cell signalling is therefore of relevance to the development of an organism and the onset of disease. For example, cancer is a disease of uncontrolled cell proliferation.

2.3 Experimental Techniques

This section is intended to provide a few notes on experimental aspects to investigate signal transduction¹⁷. For more comprehensive information on principles and techniques of practical biochemistry we refer to [WW00]. The motivation is that for mathematical modelling of signal transduction pathways it is important to appreciate the difficult and often indirect process by which information about the relationships and interactions among proteins are identified. A major challenge for the field of systems biology is to have available advanced measurement technologies to quantify protein concentrations. The purpose of this section is therefore to indicate the difficulties in obtaining quantitative data. Molecular and cell biology has been driven by the development of technologies to a large extent. With the current speed at which new tools become available, one can expect major changes to our understanding of how the cell functions¹⁸.

2.3.1 Gel Electrophoresis

Macromolecules, i.e., proteins or nucleic acids such as DNA and RNA, are commonly separated by gel electrophoresis. This method uses gels made of agarose for DNA and RNA, or polyacrylamide (PAA) for proteins. The sample is loaded on the top, or one end of the gel, and an electric field is used to pull the samples through the gel. As the gel is like a sieve the proteins or nucleic acids are separated by size. Big proteins move slower than small proteins. The resolution depends on the pore size and is only optimal for a

¹⁷The text is based on notes kindly provided by Walter Kolch.

¹⁸We cannot do justice to the technologies involved and the reader is advised to consult the literature for more information. Major technological breakthroughs are reported in journals such as *Nature* and *Science*.

certain size range. This is why agarose with wide pores is used for the large nucleic acids and PAA with smaller pores for the smaller proteins. Varying the concentration of PAA allows to adjust the size bracket for optimal resolution. For instance, proteins between 200 Da - 70 kDa are well resolved by 7.5% PAA gels; proteins between 120 Da -30 kDa are resolved on 10% PAA gels; and proteins between 50 Da - 10 kDa are resolved by 12.5% PAA gels.

2.3.2 Blotting

This is the transfer of macromolecules out from the gel onto a membrane, in order to make the separated macromolecules accessible, for instance, when other macromolecules are to be used as probes to specifically detect one of the separated macromolecules. Blotting can be done by various means. Traditionally, nucleic acids are blotted by capillary action as shown in Figure 2.17, where a stack of dry papertowels is used to draw the buffer from a tray at the bottom through the gel.

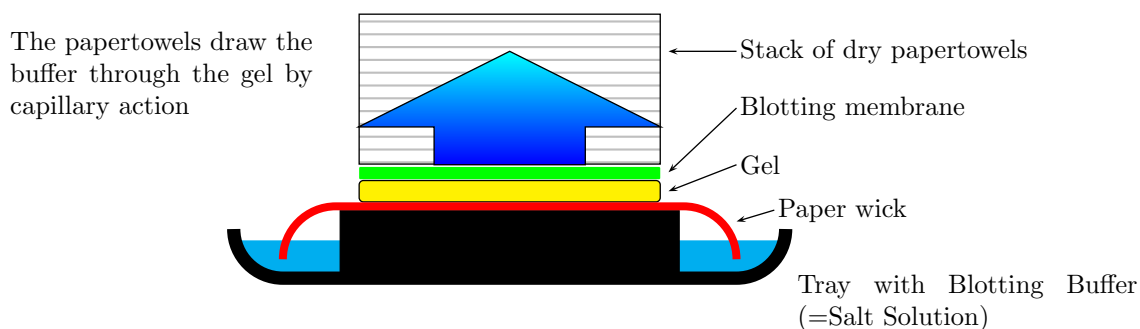


Fig. 2.17: Blotting is the transfer of macromolecules out from a gel onto a membrane.

The membrane with DNA on it is called Southern Blot¹⁹. If RNA is blotted, it is called a Northern Blot. If proteins are blotted, it is called Western blot.

Proteins are usually transferred by electroblotting (Figure 2.18), i.e., through an electrical field. The Western blot is then incubated with antibodies against the protein of interest, (e.g. Ras) and washed several times. The antibody against the protein of interest is called the primary antibody. The Ras antibody will bind to the Ras protein band on the blot, but will be washed off everywhere else. Then, the blot is incubated with a so-called secondary antibody and washed again several times afterwards. Then the blot is developed with chemiluminescence (see Figure 2.19). The secondary antibody will specifically bind to the primary antibody. For instance, if the primary antibody was raised in a rabbit, the secondary antibody is an anti-rabbit antibody, i.e., a Immunoglobulin, Ig; if the primary antibody was made in a mouse, the secondary antibody is an anti-mouse antibody etc. The secondary antibody is conjugated, i.e. covalently coupled, with an enzyme called horse-radish peroxidase, HRP. After the final wash the blot is overlaid with a thin layer of buffer containing hydrogen peroxide (H_2O_2), iodophenol and a chemiluminescent substrate, luminol, which will emit light when oxidized. The HRP conjugated to the secondary antibody will use the hydrogen peroxide to oxidize the luminol. The reaction is enhanced and prolonged by iodophenol. Therefore, light is produced at the place where the secondary antibody is bound. This will give a band on a film. The light emission lasts for about one hour. The buffer is wiped off, so that the blot stays damp and the blot is

¹⁹The procedure is called Southern blotting as it was invented by Ed Southern, Oxford.

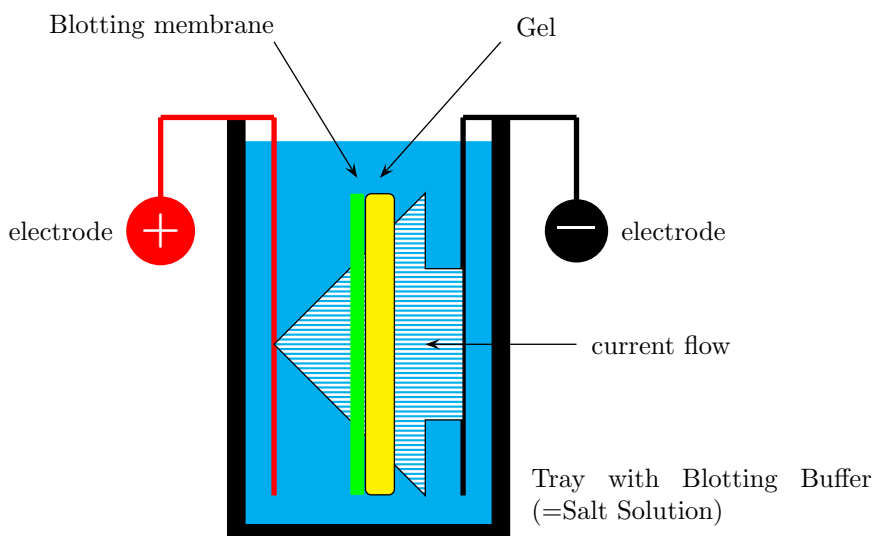


Fig. 2.18: In Western blotting proteins are usually transferred on to a membrane by electroblotting, i.e., through an electrical field.

quickly exposed to film to detect the light emission. This is the band under consideration. The detection procedure is called ECL (Enhanced Chemiluminescence)

2.3.3 Scanning and Laser Densitometry of Western Blots

The band on the film is scanned and then quantified by laser densitometry before image processing takes place. Care must be taken that the blots are not overexposed as one quickly is out of the linear range and goes into saturation. The linear range of film is not more than approximately 20, i.e., bands which differ by signal intensity of 20 fold or less are reliably quantified. This linear range is most of the times sufficient unless one has very strong and very weak bands on the same blot. Then one has to scan different exposures of the blot and extrapolate. This is rather accurate, but is also labor intensive. Our mentioning of the techniques is only to provide an appreciation of the difficulties in getting quantitative data. The technologies are rapidly improving with a wide range of supplies, which differ in the way they quantify measurements. The books [AJL⁺02] and [LC03] provide basic introductions to the various techniques and technologies.

2.3.4 Quantification of Western blots - General Considerations

Important considerations for modelling are:

1. A Western blot shows *relative* changes between the samples on the blot!
2. Different Western blots cannot be directly compared to each other!
3. A Western blot does not give you information about the concentration of a protein!

The reasons for that are many: The intensity of the bands, or signal intensity, depends on how long the ECL solution was incubated, how fast and how long the blot was exposed to film. For practical reasons these parameters are impossible to standardize. For instance, if the first exposure is under- or overexposed, then you have to put on another film - and everything has changed. Another reason is the affinity of the antibodies. Each antibody

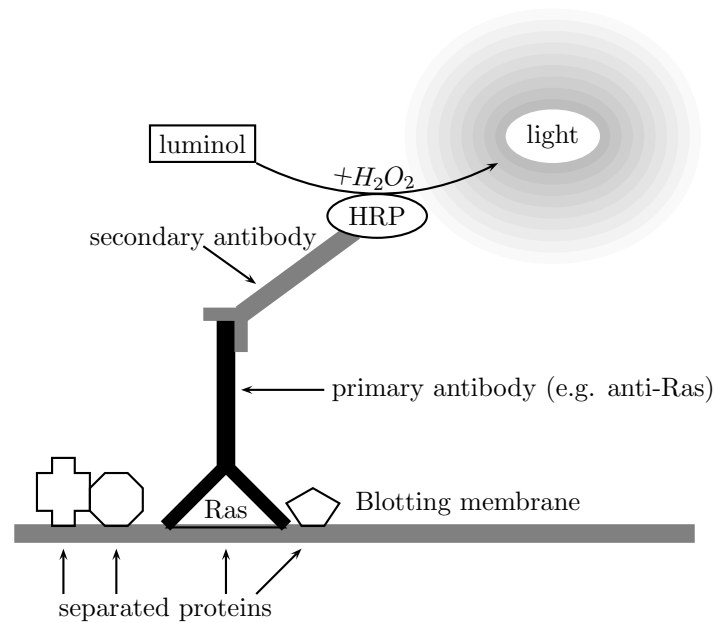


Fig. 2.19: Illustration of western blotting, the use of antibodies in separating proteins and how the presence of proteins is made visible.

has a different affinity for its protein antigen. Typical K_d values are between 10^6 and 10^9 . Thus, a good antibody will give a strong signal even with little protein present, and a poor antibody will give a weak signal even when lots of protein is present. Antibodies are proteins and therefore perishable molecules. There is batch to batch variation and also storage conditions can affect the affinity. These considerations apply to the primary and secondary antibody. Therefore, the observed signal intensity is a composite of the concentration of the protein antigen, the antibody affinities, the ECL conditions, and the exposure time of the film.

Thus, the only way to determine protein concentrations in a cell, is to compare the band intensity obtained from a cell lysate to that of a purified protein of known concentration. For instance, to determine the concentration of the ERK protein in different cell lines, lysates of these cells were run alongside a serial dilution of a purified, recombinant ERK. The densitometric scan values of this series is used to make a standard curve. “Standard” in biochemical terms means value for comparison. The standard curve should be a straight line, otherwise one is outside the linear range of the scanning. This curve relates protein concentration to densitometric scan units. These are the numbers from the scanner; they are arbitrary units. The concentrations of ERK in the different cell lines can be determined by mapping the scan units of the cell lysate to the standard curve. For the reasons mentioned above it is essential that the samples used to make the standard curve are highly pure, as the measured protein concentration reflects the quantity of the protein of interest plus any contaminating proteins. The purification of a protein to near homogeneity is extremely tedious. This makes the determination of absolute protein concentrations in cells so difficult and slow.

Improvements in the generation of quantitative and accurate data as well as standard operating procedures are an important aspect of systems biology. While in the early stages of molecular biology it was satisfactory to know whether a molecule is present, up- or downregulated, without much concern about measurement errors, this is a different story in systems biology. If experimental data are used for modelling it is paramount to have quantitative (rather than relative) and accurate long time series data.

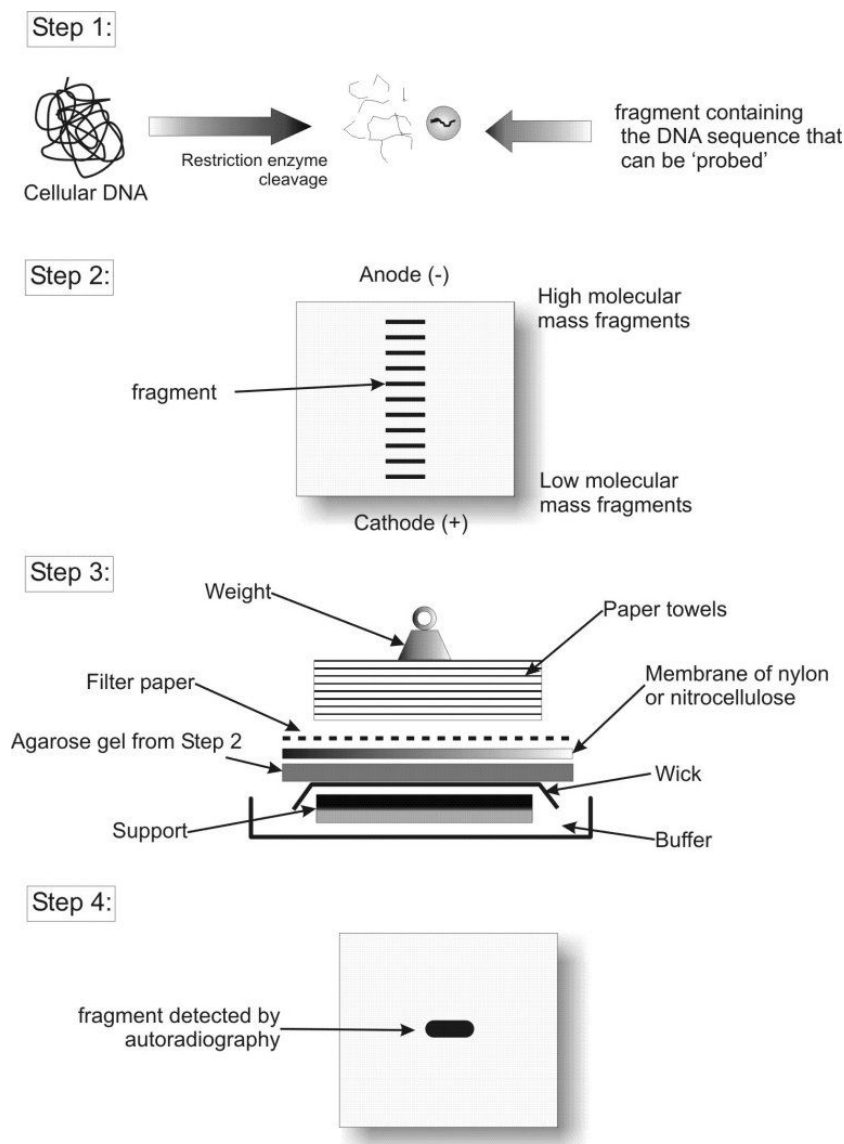


Fig. 2.20: Southern blotting. In Step 1, DNA is extracted and cleaved. In Step 2 the cleaved DNA fragments are separated according to their size using agarose gel electrophoresis. In Step 3 and after alkaline denaturation, the DNA fragments are replica-transferred to a nitrocellulose membrane, the blotting. In Step 4, hybridization with a radiolabelled probe takes place and is made visible with autoradiography.

2.4 The Dynamic Proteome

Before we continue with more technical chapters, we here briefly reflect upon the context of dynamic pathway modelling in which we are going to describe biochemical reaction networks using systems-theoretic concepts.

The area of cellular signalling investigates intracellular communication. The transmission of information within cells from receptors to gene activation by means of biochemical reaction networks (pathways) impinges on the development and disease of organisms. Our aim is to establish a mathematical/computational framework in which to investigate dynamic interactions within and between cells. In other words, we are concerned with *dynamic* pathway modelling since we do not simply map or list proteins in a pathway. Spatial-temporal sequences of reaction events in a biochemical network form a basis for *signals* that we observe as concentration changes in an experiment. We are not going to describe the physical basis of the reaction in every detail but choose a level of scale that

dynamic pathway
modelling

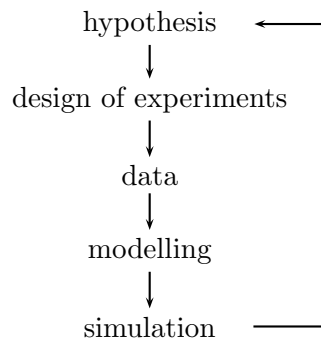


Fig. 2.21: Systems biology requires an iteration of the modelling loop illustrated here. The diagram shows the role of mathematical modelling and simulation in testing hypotheses but also in generating hypotheses through prediction. The purpose of modelling is to support experimental design, helping to identify which variables to measure and why.

is sufficiently high to allow a simplified description, which is however predictive.

Mathematical modelling and simulation of molecular or cellular biological systems is challenging. We consider such systems as ‘complex’ for the following reasons. A collection of cells, but also an individual cell consist of many interacting subsystems. For example, choosing any particular pathway there will be other pathways that “cross talk”. Due to the complexity of experiments to generate data and the sometimes complicated maths involved, we usually consider one pathway, or particular aspect of one pathway at a time. Since these systems (pathways) are interacting at different levels and in hierarchies, modelling is bound to be an *art* rather than an objective science. Although spatial aspects of the location of molecules in the cell, related diffusion or the transport of molecules, can in principle be encoded, for example, by partial differential equations, the available mathematical tools are often not easy to apply. Mathematical convenience is therefore one reason to make assumptions. Whether the underlying process is inherently random or deterministic may introduce further questions to how we represent this. For the kinetics of biochemical reactions, nonlinear ordinary differential equations are most commonly used for modelling while stochastic simulation is a popular avenue to avoid the complicated formal analysis of stochastic models.

In molecular biology experiments are typically expensive, time consuming undertakings, which in most cases deliver data sets which fall short of the expectations of statisticians or mathematicians. In contrast to the engineering sciences, the observation of molecular or cellular dynamics are indirect, i.e., it is as yet not possible to obtain a continuous stream of accurate, quantitative measurements of an intact living cell. Experiments are usually destructive with regard to the components of the cell, or in order to visualize effects it is difficult not to alter the state of the system, even if only slightly. Although there is a trend towards single cell measurements, to this day we are studying the processes within *a single* cell by using thousands or millions of cells in a biochemical experiment. While statisticians would usually argue the context in which the data are generated should be irrelevant, for the analysis of molecular- or cell biological data the context in which the data were generated is crucial information to allow any sensible conclusion. It can therefore not be avoided that our models are representations of observations that help us to speculate about the true nature of the physical processes which give rise to signals and communication within and between cells. For example, when we observe steady changes of protein concentrations in a signal transduction pathway, we may want to model this phenomena with differential equations, although the underlying reactions, due to collisions of molecules, may in fact be a random process. If we insist on a stochastic model, we immediately need to consider the question of how to validate the

model, i.e., how to estimate parameter values given only six to twelve time points of a nonlinear, non-stationary process.

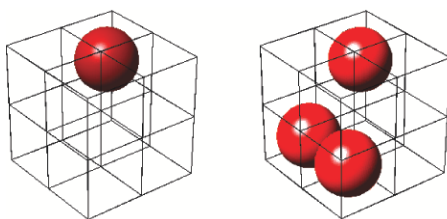
Modelling implies a process of abstraction and is often also a form of generalization. In this process we make numerous assumptions about the natural system under consideration and in order to simplify the mathematical approach, without losing the ability to make predictions. It is therefore possible to build predictive models without them being precise. The Lotka-Volterra predator-prey equations of two competing populations are an example of an unrealistic model that has nevertheless value in that it helps asking the right questions²⁰. Mathematical modelling and simulation should in this sense complement the biologist's reasoning, help him to generate and test hypotheses in conjunction with the design of experiments and experimental data.

In subsequent sections we investigate the mathematical foundations of the most commonly employed mathematical concepts in modelling pathways, discuss their properties, question the assumptions involved and compare their application with examples.

Differential equations describe rates of changes and thus appear to be a *natural* framework to describe the observations we make in experiments. Differences and changes are what we can observe and what provides us with information about a system. The state-transition map describes the changes of states. Causal entailment is the principal aim of scientific modelling and that causation is the principle of explanation of *change* in the realm of matter. However, in modelling natural systems, causation is a *relationship*, not between things, but between changes of states of systems. This view that changes in space and time are the essence of causal entailment has been well explained by the philosopher Arthur Schopenhauer who argued that the subjective correlative of matter or causality, for the two are one and the same, is the *understanding*:

“To know causality is the sole function of the understanding and its only power. Conversely, all causality, hence all matter, and consequently the whole of reality, is only for the understanding, through the understanding, in the understanding.” [Mag97].

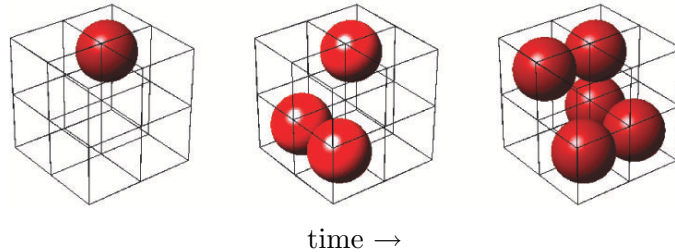
In experiments we usually look for *differences* but in our context of a dynamic systems perspective we are particularly interested in *change* over time. We can distinguish between a difference and a change, providing an example that also illustrates a difference between bioinformatics and systems biology. Let us consider the following picture as a toy model for two genomes: bioinformatics



Comparing the two, there are $2^8 = 256$ patterns to discern. For example, from comparative genomics we know that genome sequences can be very similar, while the organisms, their physiology, behavior and appearance are very different. One then wonders how this difference in complexity is possible if the genetic material appears to be so similar. Another example is the total metamorphosis of the butterfly, there is one genome but essentially two proteomes. An explanation is that the genome may provide the basic information for genomics

²⁰Murray, [Mur02] provides a discussion of the standard Lotka-Volterra system and how more realistic scenarios can be dealt with.

an organism to develop and function, but that it is the dynamic interactions of molecules and components in the cell that give rise to biological functions. If we therefore consider again our toy model, allowing the eight genes to be switched on or off in a temporal sequence, for only three time points, there are already 256^3 , i.e., more than 16 million information paths the system can describe:



systems dynamics We may conclude from this crude illustration that it is *system dynamics* that determines biological function. In bacterial systems it is known that the speed of transcription can influence the folding, structure and thus to an extent also the function of proteins. Another illustration that knowledge of components or symbols is not enough is the following set:

$$\{h, t, u, r, t, e, h, t, l, l, e, t, t, b, u, o, d, n, i, f, i\} .$$

Although not randomly arranged, looking at this list does not suffice to understand the meaning. Ordering them further does appear to make sense:

$$\{i, f, i, n, d, o, u, b, t, t, e, l, l, t, h, e, t, r, u, t, h\} ,$$

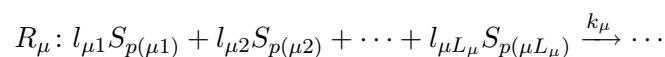
but a real understanding comes only if we read the symbols with a particular speed from left to right:

“If in doubt, tell the truth.” (Mark Twain)

Here again the temporal evolution is important: If we were given a letter a day, we would usually not be able to make sense of we are told, nor would make a foreign language and dialect make this easier. In line with the two quotations of Pauling and Poincare on page 2 we would argue that molecular and cell biology are built up from facts, as a cell is built from molecules. But a collection of facts is no more a science than a soup of molecules is a cell. Organisms and organs are complex structures of interdependent and subordinate components whose relationships and properties are largely determined by their function in the whole.

2.5 Outlook

There are various approaches to arrive at a mathematical model of intra- and inter-cellular dynamics. We are going to restrict ourselves to stochastic modelling and the use of differential equations. For differential equations there are again various perspectives one can take to motivate the set-up of the equations. To begin with, we are considering a reaction network or pathway involving N molecular species S_i . A network, which may include reversible reactions, is decomposed into M unidirectional basic reaction channels R_μ



where L_μ is the number of reactant species in channel R_μ , $l_{\mu j}$ is the stoichiometric coefficient of reactant species $S_{p(\mu j)}$, $K_\mu = \sum_{j=1}^{L_\mu} l_{\mu j}$ denotes the molecularity of reaction channel

R_μ , and the index $p(\mu j)$ selects those S_i participating in R_μ . Assuming a constant temperature and that diffusion in the cell is fast, such that we can assume a homogeneously distributed mixture in a fixed volume V and with a constant temperature and volume, one rate-equation model is the following

$$\frac{d}{dt}[S_i] = \sum_{\mu=1}^M \nu_{\mu i} k_\mu \prod_{j=1}^{L_\mu} [S_{p(\mu j)}]^{l_{\mu j}} \quad i = 1, 2, \dots, N \quad (2.19)$$

where the k_μ 's are rate constants and ν_μ denotes the change in molecules of S_i resulting from a single R_μ reaction. Representation (2.19) is a system of ordinary differential equations:

$$\frac{dx_i(t)}{dt} = V_i(x_1(t), \dots, x_N(t), \theta) , \quad (2.20)$$

where x_i denotes the n variables in question, θ denotes a parameter vector and V is a nonlinear function. The reader who is less accustomed to mathematical equations should not worry, we are going to approach the general formulations (2.19) and (2.21) below, with numerous examples of increasing complexity and generality.

Rate equation models have been widely used in describing biochemical reactions (e.g. [HS96, Fel97]). For modelling processes in living cells we are often not able to provide an exact description of the physical interactions of molecules and instead describe with our differential equations a mean of a population of interacting molecules, where the $[S_i]$ are most probable values. The fact that a differential equation is ‘deterministic’ does therefore not mean that it cannot describe a process that is inherently stochastic.

In a stochastic framework, we are looking at populations of molecules and wish to determine for each molecular species S_i the probability $\text{Prob}\{\#S_i(t) = n_i\}$ that at time t there are n_i molecules. For N molecular species, let n denote the N -dimensional state-vector, whose values are positive integers, $n \in \mathbb{Z}_+^N$. $\nu_\mu \in \mathbb{Z}^N$ are the step-changes occurring for elementary reaction indexed by μ . If S is a N -dimensional variable, we write $\text{Prob}\{\#S = n\} = P_n(t)$. Describing the changes in random variable S , we consider the following two state-transitions: First, from other states to state n , denoted

$$n - \nu_\mu \xrightarrow{a_\mu(n - \nu_\mu)} n ,$$

where $a_\mu(n - \nu_\mu)$ is the *propensity*, that is the probability per unit time, of a change ν_μ propensity occurring, given that we are in state $n - \nu_\mu$. Secondly, moving away from state n is given as

$$n \xrightarrow{a_\mu(n)} n + \nu_\mu .$$

From these definitions we arrive at an expression referred to as *master equation* or *chemical master equation* (CME)

chemical master equation (CME)

$$\frac{\partial P_n(t)}{\partial t} = \sum_{\mu=1}^M [a_\mu(n - \nu_\mu) P_{(n - \nu_\mu)}(t) - a_\mu(n) P_n(t)] . \quad (2.21)$$

The first term on the right-hand side describes the change to state n , while the second term describes the changes away from state n . M denotes the number of (unidirectional) reaction channels. The product of the propensity with the probability should be read as an ‘and’. The multiplication of a propensity and probability makes sense in light of the derivative against time on the left, in that a propensity, multiplied with dt gives a probability. In the above setting $a_\mu(n - \nu_\mu)$ and $a_\mu(n)$ are *transition probabilities*, transition probability while $P_n(t)$ and $P_{(n - \nu_\mu)}(t)$ are the probabilities that the system is at time t in state n , respectively $n - \nu_\mu$. A solution of the master equation describes the probability of

there being n molecules at time t . Chapter 4 is going to provide a series of examples for stochastic processes for which the methodologies become increasingly general until we arrive at (2.21).

A major difficulty with the CME is that the dimension of these sets of equations depends not only on the number of chemical species N but for any possible number of molecules of any species we have n differential-difference equations. Gillespie [Gil76, Gil77, Gil01, GB00] developed an algorithm to simulate or realize a CME model efficiently. The Gillespie approach to stochastic *simulation* has in recent years become popular in modelling intra-cellular dynamic processes [Kie02, RWA02, vGK01, MA97]. Some authors have unfortunately confused the simulation of a stochastic model with a stochastic model. The Gillespie algorithm does not provide a solution to the master equation but generates realizations. While a formal analysis of (2.21) is very difficult, it is possible to approximate the CME by a truncated Taylor expansion, leading to the Fokker-Planck equation, for which there exist some results [ELS01, Gar85, vK92]. Comparing (2.19) and (2.21), we note that while rate equations are deterministic in the sense that they employ differential equations, they are based on a probabilistic description of molecular kinetics. On the other hand, the CME is a stochastic formulation, but based on differential equations, with probabilities as variables. Although we are going to look at various stochastic models and their derivation, we eventually settle for model structures (2.20), to describe molecular principle from what we can observe in experiments.

The motto of this book is nicely captured in the following quotation by Ludwig von Bertalanffy, a founder of general systems theory and someone who laid the foundations for systems biology in the 1960s:

“Considering the inconceivable complexity of processes even in a simple cell, it is little sort of a miracle that the simplest possible model – namely, a linear equation between two variables – actually applies in quite a number of cases.

Thus even supposedly unadulterated facts of observation already are interfused with all sorts of conceptual pictures, model concepts, theories or whatever expression you choose. The choice is not whether to remain in the field of data or to theorize; the choice is only between models that are more or less abstract, generalized, near or more remote from direct observation, more or less suitable to represent observed phenomena.

On the other hand, one should not take scientific models too seriously. Kroeber (1952), the great American anthropologist, once made a learned study of ladies’ fashions. You know, sometimes skirts go down until they impede the lady in walking; again, up they go to the other possible extreme. Quantitative analysis revealed to Kroeber a secular trend as well as short-period fluctuations in the length of ladies’ skirts. This is a perfectly little law of nature. I believe a certain amount of intellectual humility, lack of dogmatism, and good humor may go a long way to facilitate otherwise embittered debates about scientific theories and models.”

(From Bertalanffy’s book *General Systems Theory*; the introductory part to the chapter ‘Some Aspects of Systems Theory in Biology’ [Ber68].)

The previous section took us from a relational world view to a general systems approach. The present section introduced the formal systems we are dealing with hereafter. The next section takes us on to the journey in which we apply the concept above to an understanding of the processes that make up a living system. Enjoy!

3 Biochemical Reactions

For anyone interested in modelling and simulation of biochemical reaction networks or pathways, there are plenty of tools to choose from. These packages are roughly divided into whether they deal primarily with deterministic or stochastic models. Here we are discussing the close relationship between both ideas, in the context of biochemical reactions. We begin with simple, particular examples and increase the generality of the equations before we arrive at the general chemical master equation (2.21) and have available a set of more advanced mathematical tools to investigate the properties of such stochastic representations.

3.1 The Rate Equation Approach

We are now going to go through a short example of the modelling process. The aim is to describe changes in a population of molecules. We first consider only one kind or *species* of molecules, which we call S . As a conceptual framework in which we formulate our models we initially consider Ludwig Boltzmann's *Kinetic Theory of Gases* from 1877. It begins with the assumption that for constant pressure, temperature, and volume V , the number of collisions between any two molecules should be constant. Let $\#S \doteq n$ denote the average number of molecules of species S . If the probability of a reaction to occur is independent of the details of that collision, then the change Δn in the number of molecules is proportional to n as well as to a time interval Δt :

$$\Delta n \propto n \cdot \Delta t$$

There are several kinds of biochemical reactions, which in turn can be combined into networks or pathways. We start with the simplest reaction which proceeds by itself, involving only one molecular species. The empirical rule we described above can now be turned into a mathematical equation

$$\Delta n = k \cdot n \cdot \Delta t, \quad (3.1)$$

where Δt is assumed to be a *relatively* small interval of time. Dividing both sides by Δt ,

$$\frac{\Delta n}{\Delta t} = k \cdot n,$$

we can now consider what happens as Δt becomes infinitesimal small, $\Delta t \rightarrow 0$, leading to the differential operator. For large n , changes in this very small time interval will be very small compared to the overall changes in the population. We could thus turn from discrete changes in the number of molecules n to continuous changes $S(t)$ in the population or concentration of S . This then leads us to the ordinary *differential equation*¹

differential equation

$$\frac{dS(t)}{dt} = k \cdot S(t), \quad (3.2)$$

as a model of the simple monomolecular or autocatalytic reaction. A differential equation describes the *rate of change* of variable S and is a translation of the observations and assumptions we make in modelling a natural system. To simulate the system and to make predictions for values of $S(t)$ for a range of time points t , we need to find a *solution* to the differential equation. For simple cases we may be able to find *analytical* solutions through mathematical analysis, while for more complex cases we need to resort to a *numerical*

analytical vs. numerical solutions

¹An *ordinary* differential equation, as opposed to a *partial* differential equation, does not consider spatial distribution of components or diffusion.

solution or simulation. The advantage of an analytical solution is that it is more general, we usually do not need to know an exact value for the parameters, here k . For our simple reaction the solution to the differential equation (3.2) is

$$S(t) = S(0) \cdot e^{kt} , \quad (3.3)$$

rate constant where k is the *rate constant* by which the conversion of reactant A proceeds and $S(t_0) = S(0)$ defines the initial condition. We frequently write S_0 for $S(0)$. The solution is thus dependent on a parameter value for k and the initial conditions. Since we could get the differential equation (3.2) from (3.3) by differentiating the equation, another way round to a solution of the differential equations is by integration. A simulation of a system of differential equation is therefore a numerical integration of (3.2). We are going to discuss the process by which we take the limit $\Delta t \rightarrow 0$ and integrate the differential equation in more detail further below in this section. Before that, we look at the application of our model.

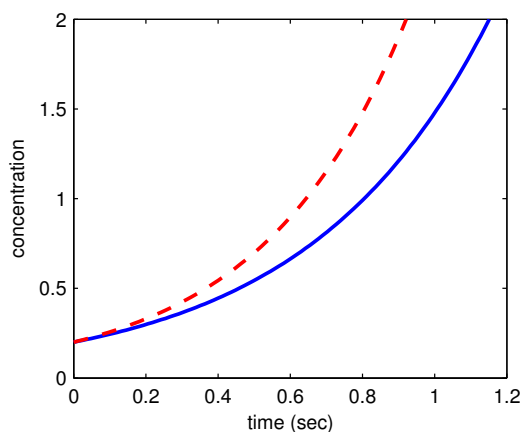
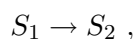


Fig. 3.1: Simple exponential growth of a population, $S(t) = S(0) \cdot e^{kt}$. The dashed line is for parameter $k = 2.5\text{sec}^{-1}$ and the solid line for $k = 2\text{sec}^{-1}$. The solution of the differential equation $dS(t)/dt = kS(t)$, depends on the initial condition (here $S(0) = 0.2$), that is, for each initial condition there is a different solution. The fact that this curve grows unbounded suggests that it is not a realistic growth model.

The model (3.2) describes an *irreversible* biochemical reaction, that is, the population or concentration of S can either only increase or only decrease as the result of changes to the molecules. As we shall see later, this irreversible reactions are somewhat unrealistic for living cells. The reaction is also *monomolecular* since it involves only one kind of species. Whatever happens to S , we could consider it to be a transformation of S_1 into S_2 . In other words, we have two kinds or forms of S , the biochemical notation for this is



reactant species where the chemical species on the left are referred to as the *substrate* or *reactant species*
product species while those on the right are called *product species*. As above, for a small time interval Δt , the changes ΔS_1 , ΔS_2 satisfy the proportionality relation

$$\Delta S_1 \propto -S_1 \Delta t ,$$

and if the number of molecules is conserved

$$\Delta S_2 = -\Delta S_1 .$$

For $\Delta t \rightarrow 0$ these relations turn into differential equations

$$\frac{d}{dt}S_1 = -kS_1 \quad \frac{d}{dt}S_2 = kS_1 .$$

We are now at a point where we need to discuss the units of the elements of our model. The molar concentration $[S]$ is given as mol per liter:

$$1 \frac{\text{mol}}{\text{L}} \equiv 1\text{M}$$

In general, in equations we denote the volume by V . Since a mole contains $6 \cdot 10^{23}$ molecules, in order to get a count of the actual number of molecules $\#S$ we would have to multiply the molar mass by the Avogadro constant

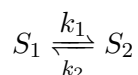
$$N_A = 6.02205 \cdot 10^{23} \text{mol}^{-1} .$$

With a slight abuse of terminology, we hereafter occasionally use S to denote a “count” of molecules in moles. Note, if it is clear from the context and explicitly stated, we may also use S to denote a concentration, leaving away the square brackets $[S]$ to simplify the notation. We are thus assuming that concentrations are taking real values and change smoothly. Using square brackets $[S]$ to denote molar concentrations (M) for our example, the unit of $d[S]/dt$ is M per second and we have

$$[S] = \frac{S}{V} \quad \text{and} \quad \#S = S \cdot N_A .$$

Although here the rate constant k is measured as ‘per second’, denoted sec^{-1} , in general the units of the rate constant will depend on the type of the reaction and whether we are dealing with concentrations or counts.

Most if not all biochemical reactions in the living cell are reversible and so the next step in making our approach to modelling biochemical reactions more comprehensive, is to consider a reversible reaction in which compound S_1 is transformed into compound S_2 , reversible reaction and *vice versa*:



where k_1 is called the *forward rate constant* and k_2 *reverse rate constant*. If the reversible reaction is in an *equilibrium*, the average concentrations remain constant² and the rates of changes are zero equilibrium

$$\frac{d}{dt}[S_1] = \frac{d}{dt}[S_2] = 0 ,$$

which is the same as to say

$$\frac{k_1}{k_2} = \frac{[S_2]}{[S_1]} .$$

This relation is what C.M. Guldberg and P. Waage in 1864 described as the *law of mass action* (LMA). The key to the differential equation model was the assumed proportionality law of mass action in (3.1). From our initial discussion, the following set of differential equations serves as a model for the reversible reaction above:

$$\frac{d}{dt}[S_1] = -k_1[S_1] + k_2[S_2] , \quad \frac{d}{dt}[S_2] = k_1[S_1] - k_2[S_2] .$$

If we consider the total concentration of S_1 and S_2 together as constant, $[S_1] + [S_2] = S_T$ and substitute $[S_2] = S_T - [S_1]$ for $[S_2]$ in the equation for S_1 , we obtain

$$\frac{d[S_1]}{dt} = -(k_1 + k_2)[S_1] + k_2 S_T .$$

²Note that this does not mean nothing happens. There is still a flux of material, although the reactions are balanced such that macroscopically we do not observe changes.

Simplifying the notation, let us denote $S_1 \doteq x$, $\alpha \doteq k_1 + k_2$ and $\beta \doteq k_2 S_T$ so that the differential equation looks tidier:

$$\frac{dx}{dt} = \beta - \alpha x . \quad (3.4)$$

This model is then an example of the system considered in the previous chapter:

$$\dot{x} = \phi(x, \theta) ,$$

where $\theta = (\alpha, \beta)$. This is a basic differential equation, for which there are various ways to solve it. Here we simply state the result in order to compare it with (3.3) as a model for population growth:

$$x(t) = \frac{\beta}{\alpha} + \left(x(0) - \frac{\beta}{\alpha} \right) e^{-\alpha t} .$$

Figure 3.2, shows simulation results for two different initial conditions. We observe that although the growth depends heavily on the initial conditions, the growth is limited to β/α .

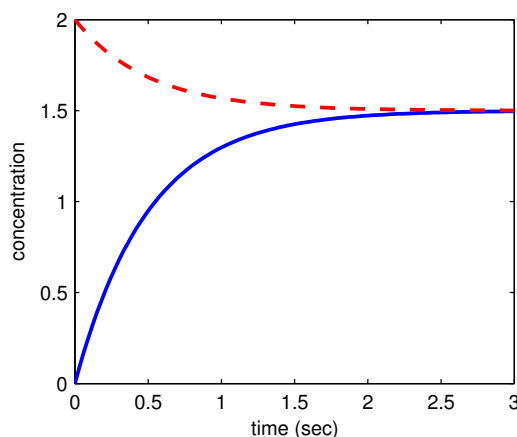


Fig. 3.2: Numerical solution for the differential equation (3.4), which was derived as a model for a reversible reaction. The solid line is for $x(0) = 0$, while the dashed line is for $x(0) = 2$ M. $\alpha = 2\text{sec}^{-1}$, $\beta = 3 \text{ M/sec}$.

3.1.1 Differential vs. Difference Equations

A justified criticism of the ODE model (3.2), as a representation of a biochemical reaction could be that we did not aim for *physical realism*, modelling the collisions of molecules. Instead we modelled what we observed: that at any time the change is proportional to Δt and the current number of molecules. A reasonable interpretation of (3.2) is then that $S(t)$ represents the average population level at time t . In this case we can view k as the difference between the formation rate k^+ of S and the decay rate k^- such that

$$\frac{dS(t)}{dt} = (k^+ - k^-)S(t) \quad \text{with solution} \quad S(t) = S(0) \cdot e^{(k^+ - k^-)t} . \quad (3.5)$$

In the derivation of the differential equation we made another assumption of a large population $\#S \doteq n$ such that discrete changes to the population are small enough to assume overall changes to the population are continuous (see Figure 3.3 for an illustration). This assumption in effect describes a discrete process with a continuous model. We can discuss this mathematically by approximating a difference equation by a differential equation. A simple intuitive example for this is the interest we receive for savings in a bank account. Say a bank offers on savings a rate of return r , in percent, compounded annually. If $S(0)$ is

the initial saving put in the bank, and $S(t)$ its future value after t years, then the following difference equation models the growth of our money:

$$S(t+1) = (1+r)S(t) ,$$

with initial condition $S(t_0) = S(0)$. If we are interested in growth on a monthly basis,

$$S\left(t + \frac{1}{12}\right) = \left(1 + \frac{r}{12}\right) S(t) .$$

Here r is divided by 12 because it is by definition compounded annually (not monthly). In general, if a year is divided into m equal intervals, the difference equation becomes

$$S\left(t + \frac{1}{m}\right) = \left(1 + \frac{r}{m}\right) S(t) ,$$

which can be written more conveniently

$$\frac{S\left(t + \frac{1}{m}\right) - S(t)}{\frac{1}{m}} = rS(t) .$$

As m goes to infinity, denoted $m \rightarrow \infty$, the above difference equation becomes a differential equation

$$\frac{dS(t)}{dt} = rS(t) , \quad (3.6)$$

where t is now a continuous time variable. One reason for choosing a continuous time representation with continuous changes is the rich set of analytical tools that is available to investigate the properties of such equations.

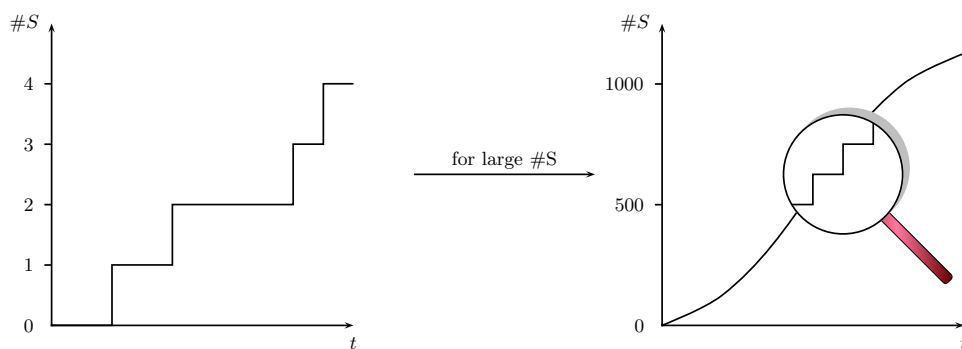


Fig. 3.3: Approximation of discrete steps as continuous changes. For large numbers of molecules changes to the population appear smooth and may be represented by a continuous model.

3.1.2 Numerical Simulation

Here we look at another assumption made in our model (3.2), that of a small interval of time Δt . This question is closely related to finding solutions to differential equations. For simple linear ordinary differential equations like the one above we can find exact or analytical solutions. For more complicated cases, in particular nonlinear equations, which we consider later, we can use numerical integration to obtain an approximation [AS72, PTVF93]. The simplest approach to obtain a numerical solution of a general ordinary differential is called the *forward Euler method*:

$$\frac{dS}{dt} \approx \frac{\Delta S}{\Delta t} = \frac{S(t + \Delta t) - S(t)}{\Delta t} , \quad (3.7)$$

Euler method

where ΔS and Δt are small, but not infinitesimal. If this approximation to the derivative is substituted into (3.5) and the equation is rearranged we get

$$\frac{S(t + \Delta t) - S(t)}{\Delta t} \approx k^+ S(t) - k^- S(t)$$

$$S(t + \Delta t) = S(t) + S(t)(k^+ - k^-)\Delta t .$$

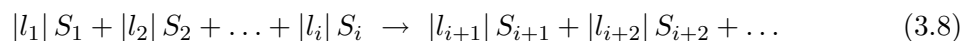
The justification for the differential d/dt may be considered a mathematical explanation. A physical argument is that in order to avoid surface effects, influencing the interactions among molecules, we consider an infinitely large system $\lim V \rightarrow \infty$. To avoid that the concentration goes to zero, the number of molecules must become very large in order to move from a discrete sum to a continuous integral.

3.2 Biochemical Reaction Modelling

Following the above introduction to differential equation modelling, the present section is to provide a more comprehensive survey of biochemical reaction modelling. The theoretical and experimental description of chemical reactions is related to the field of *chemical kinetics*. A primary objective in this area is to determine the rate of a chemical reaction, i.e., describing the velocity of conversion of reactants to products. Another task is the investigation of the influence of external factors, like temperature, pressure, and other chemical species on the chemical reaction under consideration. The determination of the reaction mechanism, the way the products are formed, which intermediates are created, is a further field of chemical kinetics.

3.3 Fundamental quantities and definitions

The pivotal quantity in the description of chemical reactions is the *reaction rate*. The general chemical reaction



can be summarized by

$$0 = \sum_i l_i S_i ,$$

where l_i denotes the *stoichiometric coefficient* for the i -th component of the chemical reaction defined in (3.8). The reaction rate of this reaction is defined as

$$r(t) = \frac{1}{l_i} \frac{d\#S_i(t)}{dt} , \quad (3.9)$$

where $\#S_i$ is the number of molecules of species S_i in the considered volume. The stoichiometric coefficients l_i are negative for reactants and positive for products. From the definition (3.9) it follows that the reaction rate is a positive definite quantity identical for all participating species S_i .

In wet-lab experiments and in modeling reactions in a cell we usually assume a constant volume, although in the cell this is surely not always a realistic assumption. For a system with constant volume we can transform (3.9) into an equation for the concentration $[S_i]$ of species S_i

$$r(t) = \frac{1}{l_i} \frac{1}{V} \frac{d\#S_i(t)}{dt} = \frac{1}{l_i} \frac{d\#S_i/V}{dt} = \frac{1}{l_i} \frac{d[S_i](t)}{dt} . \quad (3.10)$$

In general one can use any quantity which is proportional to the number of molecules, for instance the particle density or the partial pressure of gases but for practical considerations the measurement of concentrations is often easier than the count of molecules. But there is a more important difference between these quantities. The particle number is an *extensive property*, i.e., it depends on the size of the system. The concentration, particle density, ... are *intensive quantities* independent from the particular system under consideration. We hereafter assume a system with constant volume, in which case we get for (3.8) a reaction rate

$$\begin{aligned} v(t) &= -\frac{1}{|l_1|} \frac{dS_1(t)}{dt} = -\frac{1}{|l_2|} \frac{dS_2(t)}{dt} = \dots = -\frac{1}{|l_i|} \frac{dS_i(t)}{dt} \\ &= \frac{1}{|l_{i+1}|} \frac{dS_{i+1}(t)}{dt} = \frac{1}{|l_{i+2}|} \frac{dS_{i+2}(t)}{dt} = \dots, \end{aligned} \quad (3.11)$$

where $S_i(t)$ represents either a particle number or a proportional quantity. Degrading species are characterized by a minus sign while an increase is indicated by a plus sign. According to (3.11) the reaction rate is proportional to the change of concentration. Another possibility to investigate the reaction rate is given by the advancement or *extent of an reaction* $\epsilon(t)$. This quantity is a measure of the progress of the chemical reaction under consideration. It is defined as

$$\epsilon(t) = \begin{cases} \frac{1}{l_i} (S_i(t) - S_i(0)) & \text{for products,} \\ \frac{1}{l_i} (S_i(0) - S_i(t)) & \text{for reactants.} \end{cases} \quad (3.12)$$

The extent $\epsilon(t)$ relates the initial conditions $S_i(0)$ to the time dependent variables $S_i(t)$ and has the same value for all species. We are using it for the analytical integration of time laws of higher orders later in the text. With the extent the reaction rate is

$$v(t) = \frac{d\epsilon(t)}{dt} = \frac{1}{l_i} \frac{dS_i(t)}{dt} \quad (3.13)$$

and is interpreted as rate of change of the advancement of a reaction.

It is possible to formulate a conservation law which relates the reactants and the products in a closed system:

$$S_0 = S_1(t) + S_2(t), \quad (3.14)$$

where S_0 is the initial concentration of the reactant S_1 , where $S_1(t)$, $S_2(t)$ are time dependent concentrations. Here, we assume that the initial concentration of the product S_2 is zero. Relation (3.14) describes the conservation of the number of atoms in a *closed system* under the influence of the chemical process. One distinguishes between three types of systems according to their possibilities to exchange energy and matter with the environment. The most restricted system is an isolated or fully closed system, where no transfer of energy and matter is possible. If we allow the exchange of energy, but no exchange of matter, the system is called closed. Such a system is often used for chemical and biochemical experiments. In an *open system* also a transfer of matter is possible. Examples are flow reactors in the chemical industry and of course the cell in its natural environment. The creation and the disappearance of atoms is not a chemical process. Since the concentration is a function of the molecule number, we obtain a conservation of concentrations. Often, this law is also called *mass conservation*. From a physicist's perspective this is wrong. The mass is not a conserved quantity in chemical systems. As mentioned above, we can use any quantity proportional to the particle number. For each of them one can formulate a specific conservation law of the form (3.14). To avoid a restriction of systems in our treatment we will simply call it conservation law for this reason.

The advantage of conservation laws is, that they simplify the description of the system of interest and give conditions to narrow relevant solutions down. The simplification arises from the fact, that each conservation law eliminates a variable and reduces the order of a system of coupled differential equations. Further famous examples of conservation laws are energy conservation, momentum conservation and the angular momentum conservation.

3.4 Basic Principles and Assumptions

In the previous section we defined the reaction rate as a differential equation that depends on the change of participating species over time. In order to obtain the temporal behavior of molecular species, we have to specify the functional relation of change. The class of differential equations is not automatically restricted to ordinary differential equations. In ODEs the rate of change is described by a continuous function. The chemical conversion is however not a continuous process as it is postulated for the use of functions. If one wishes to consider the discreteness of the process without a change of the general framework, this is possible but requires the introduction of what are called the Dirac- δ -function and the Heaviside step-function θ . In a strict mathematical sense these are not functions but distributions and hence the differential equations are strictly speaking no longer ‘ordinary’ differential equations. The fact that it is possible to describe discrete changes in reactions with differential equations is worth mentioning. We are going to return to this question when we introduce stochastic models. In the literature stochastic models are frequently justified by stating that differential equations are not able to capture discrete changes. A formal theoretical description of discrete changes in reactions in terms of distributions is possible although non-trivial. One has to know the time of each reaction within the considered system. But since a prediction of the exact time is not possible, only statistical properties are known for chemical reactions. In particular the probability of a reaction can be calculated. Stochastic simulations work in this way. The simulations calculate the time of a reaction from probability distribution functions. The result of a simulation run is one possible realization of the temporal evolution of systems and corresponds to the formal treatment described above. The calculation of probabilities of a reaction requires a detailed information about the process of a chemical reaction. Collision theory is a successful approach to justify experimental rate laws. Two particles can only react if they interact with each other. In a classical view both particles must collide, like two balls on the billiard table.

The description of a molecule as a hard sphere is motivated by the subsequently feasible assumption of an interaction with a short contact time between the molecules. This means the interaction range is much smaller than the average distance between molecules. It is then possible to assume collisions as independent, an assumption useful in the context of stochastic Markov models. Because of the finite interaction range we also have a finite interaction time. This is the time, a molecule needs to move through the potential of its collision partner. Within this time old bonds break and new ones are established. Analogous to our assumption on the interaction length the interaction time is small and negligible in comparison to the time between two collisions. It follows from this that the number of reactions per time is related to the number of collisions within the considered time interval. Statistical physics is one possible tool for the solution of this problem. The main assumption made in statistical physics is, that the properties of the investigated system are well described by the ‘expected value’ or mean value. Therefore, all results obtained in due course have to be interpreted as averages. There are fluctuations around the expected value, but these are assumed small in comparison to the average. Because of the use of averages we change from a discrete description to a continuous formulation. Furthermore we assume an ideal gas in thermodynamic equilibrium. Then, the velocity

of the molecules is given by the Maxwell-Boltzmann distribution function. In order to avoid surface effects³ one assumes an infinitely expanded system. Because of the infinite system volume V one has to increase the particle number $\#S$ to infinity to keep the right particle density. All further treatments are in this thermodynamic limit, requiring that in the limit of $V, \#S \rightarrow \infty$ the ratio $\#S/V$ is constant and finite. Last but not least we have to make an assumption each chemical reaction is independent from the others. In addition to these assumptions we also restrict our treatment to a special class of systems called *isothermal* and *isochore* reaction systems. This means that during the reaction no temperature changes and no volume changes occur.

What follows is a short summary of chemical kinetics. From collision theory the reaction rate is equal to the number of molecules participating in the chemical reaction and the *rate coefficient* k . The rate coefficient⁴ k summarizes details of the reaction and is only fully known for very simple systems. This requires information about particle motion and is therefore temperature dependent. Furthermore, not all collisions lead to a chemical reaction. The collision energy can be too small to initiate the reaction. The molecules have a false direction, for instance, the molecules do not hit their reactive sites and cannot react. The molecules are too fast, it is not enough time to break the old and establish the new chemical bindings. These are only few of the possible reasons for a non-reactive collision. In summary, the theoretical calculation of the rate coefficient is complicated. For this reason we have to resort to experimental or estimated values. If identical species react, the rate coefficient also contains a symmetry factor avoiding a double counting⁵.

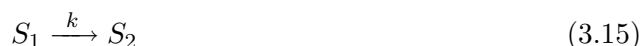
A further common assumption for pathway modeling is, that we can decompose more complicated chemical reactions into a sequence of elementary reactions, which we can describe using chemical kinetics. The most common classification is to distinguish the reaction by the number of participating molecules. The simplest reaction is the monomolecular reaction with only one molecule. In a bimolecular reaction two molecules or substrates form a product. This is the most frequently occurring reaction in nature. Trimolecular reactions are rare, because the probability for a collision of three particles within a tiny time interval is rather small. We will deal with these elementary reactions and more complicated reactions in the following sections.

3.5 Elementary Reactions

We assume that chemical reactions can be decomposed into a sequence of elementary steps, the elementary reactions. For simple reactions it is possible to derive time laws. In this section we will discuss the properties of the different elementary reactions in greater detail, finding solutions to differential equations.

3.5.1 Monomolecular reactions

The monomolecular reaction



is the simplest elementary reaction. It converts species S_1 to S_2 . The variable k above the arrow is the corresponding rate coefficient. For the monomolecular reaction the rate

³The presence of a reactive surface can dramatically change the properties of a (bio)chemical system. The catalyst of cars works in this way.

⁴Often k is called rate constant because of its time independence. But k is dependent on system parameters like the temperature or the pH-value and hence the term ‘coefficient’.

⁵For instance, for a reaction of two identical molecules one have to introduce a symmetry factor $1/2$ otherwise one counts each collision twice.

coefficient has the unit time^{-1} , independent from the units used for the species. The quantity $k dt$ is the probability that a reaction occurs within the time interval dt . According to the common chemical kinetics the reaction rate is

$$r(t) = -\frac{dS_1(t)}{dt} = k S_1(t) \quad (3.16)$$

which can be integrated by separation of variables

$$\frac{dS_1(t)}{S_1(t)} = -k dt . \quad (3.17)$$

The integration within the limits $S_1(0)$ to $S_1(t)$ and from 0 to t in the time domain

$$\int_{S_1(0)}^{S_1(t)} \frac{dS_1(t)}{S_1(t)} = -k \int_0^t dt \quad (3.18)$$

leads to

$$\ln \frac{S_1(t)}{S_1(0)} = -kt . \quad (3.19)$$

Solving this for $S_1(t)$ we obtain the familiar exponential law

$$S_1(t) = S_1(0) \exp\{-kt\} , \quad (3.20)$$

for the temporal evolution of S_1 . $S_1(0)$ is the initial condition at $t = 0$. The solution for the product S_2 is obtained from the conservation law

$$S_2(t) = S_1(0) - S_1(t) + S_2(0) \quad (3.21)$$

as

$$S_2(t) = S_1(0) [1 - \exp\{-kt\}] + S_2(0) , \quad (3.22)$$

with the initial value $S_2(0)$.

Characteristic times

The t_y -time is the time where the normalized quantity $S_1(t)/S_1(0)$ has the value y . Hence, possible values for y lie in the interval $[0, 1]$. From (3.20), it follows that

$$t_y = -\frac{\ln y}{k} . \quad (3.23)$$

half life The most commonly known t_y -time is the *half life* $t_{1/2}$. At this time point half of the initial amount $S_1(0)$ is transformed into product S_2 . It follows from the general definition (3.23)

$$t_{1/2} = \frac{\ln 2}{k} \approx \frac{0.69}{k} , \quad (3.24)$$

which is independent from S_1 . This means that it takes always $t_{1/2}$ to halve the amount of S_1 . This is illustrated in Figure 3.4, where $S_1(t)$, normalized by $S_1(0)$, is drawn as function of the half life $t_{1/2}$. At every multiple n of $t_{1/2}$ the ratio $S_1(t)/S_1(0)$ has the value $1/2^n$ as mentioned before.

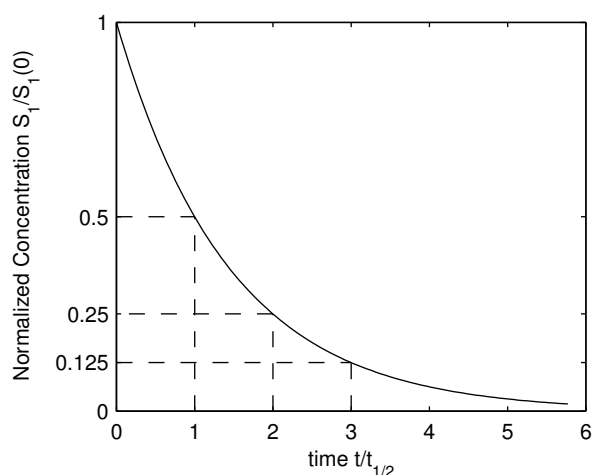


Fig. 3.4: The monomolecular reaction (3.15) as function of the half life $t_{1/2}$. The normalized concentration $S_1(t)/S_1(0)$ has the value 2^{-n} at the n -th multiple of the half life.

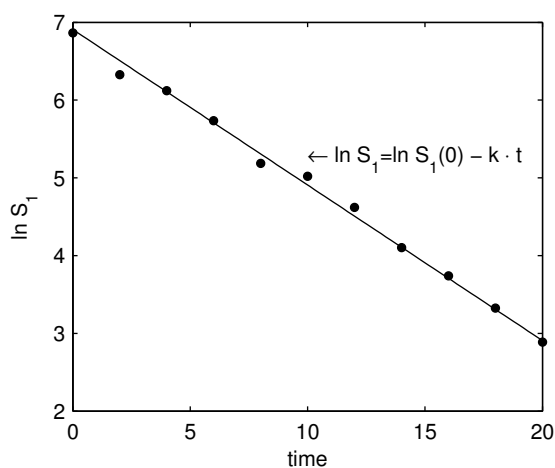


Fig. 3.5: The measurement of the rate coefficient k for a monomolecular reaction. The logarithm of the concentration S_1 as function of time t is a straight line with slope $-k$. The initial concentration $S_1(0)$ is given by the intersection with the y -axis. The parameters are obtained by a linear fit to measured data.

Measurement of the rate coefficient

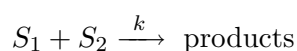
In case of a monomolecular reaction the rate coefficient can be measured in a simple way. From (3.19) the equation

$$\ln S_1(t) = \ln S_1(0) - k t \quad (3.25)$$

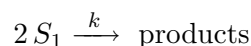
defines a straight line with slope $-k$ and an intersection with the y -axes $\ln S_1(0)$. If one applies the logarithm of measured data, as function of time, then the rate coefficient k is obtained from a linear fit. This is sketched in Figure 3.5.

3.5.2 Bimolecular Reactions

Probably the most common reaction that occurs in cells is the bimolecular reaction. There are two different ways by which two reactants combine to form one or more products. In the reaction



two different species are transformed into products, but the reaction



is also possible. We will discuss both types of bimolecular reactions in the following two sections.

Bimolecular reactions of two different molecules

The bimolecular reaction of two different molecular species S_1 and S_2 is defined by



where k is the rate coefficient for the bimolecular reaction. Its unit is now dependent on the units of S_1 and S_2 . If S_1 and S_2 are concentrations, the rate coefficient has the unit concentration per time. From collision theory we have for the reaction rate

$$r(t) = -\frac{dS_1(t)}{dt} = -\frac{dS_2(t)}{dt} = k S_1(t) S_2(t) . \quad (3.27)$$

With the help of the extent variable $\epsilon(t)$ (3.12) we can transform this equation to

$$r(t) = \frac{d\epsilon(t)}{dt} = k [S_1(0) - \epsilon(t)] [S_2(0) - \epsilon(t)] , \quad (3.28)$$

with initial concentrations $S_1(0)$ and $S_2(0)$. The differential equation is now solvable by separation of variables. The analytical solution of the integral

$$kt = \int_0^{\epsilon(t)} \frac{d\epsilon(t)}{[S_1(0) - \epsilon(t)] [S_2(0) - \epsilon(t)]} \quad (3.29)$$

is obtained, if one uses the following expansion into partial fractions

$$\frac{1}{[S_1(0) - \epsilon(t)] [S_2(0) - \epsilon(t)]} = \frac{1}{S_1(0) - S_2(0)} \left[\frac{1}{S_2(0) - \epsilon(t)} - \frac{1}{S_1(0) - \epsilon(t)} \right] .$$

The result

$$\begin{aligned} kt &= \frac{1}{S_1(0) - S_2(0)} \ln \frac{S_2(0) S_1(0) - \epsilon(t)}{S_1(0) S_2(0) - \epsilon(t)} \\ &= \frac{1}{S_1(0) - S_2(0)} \ln \frac{S_2(0) S_1(t)}{S_1(0) S_2(t)} \end{aligned} \quad (3.30)$$

is transposable to $S_1(t)$ and $S_2(t)$ using the relation $\epsilon = S_1(0) - S_1(t) = S_2(0) - S_2(t)$. One obtains

$$S_1(t) = \frac{S_2(0) - S_1(0)}{S_2(0)/S_1(0) \exp\{[S_2(0) - S_1(0)]kt\} - 1} \quad (3.31)$$

and

$$S_2(t) = \frac{S_1(0) - S_2(0)}{S_1(0)/S_2(0) \exp\{[S_1(0) - S_2(0)]kt\} - 1} \quad (3.32)$$

as the time law for bimolecular reaction (3.26). The reactants decrease exponentially. The component abounding at the beginning is left over at the end of the reaction. If one plots the logarithmic term in (3.30) versus the reaction time t , one obtain a straight line with slope $-k(S_1(0) - S_2(0))$ allowing the measurement of the rate coefficient k from experiment and a linear regression.

Equations (3.31) and (3.32) can be simplified in the case of stoichiometric concentrations $S_1(0) = S_2(0)$. Under this conditions the relation $S_1(t) = S_2(t)$ holds at every time. The reaction rate (3.27) can be rewritten as

$$r(t) = -\frac{dS_1(t)}{dt} = -\frac{dS_2(t)}{dt} = k [S_1(t)]^2 = k [S_2(t)]^2 , \quad (3.33)$$

which is again solvable by separation of variables. In this special case the time law is

$$S_1(t) = \frac{S_1(0)}{S_1(0)kt + 1} = \frac{S_2(0)}{S_2(0)kt + 1} = S_2(t) . \quad (3.34)$$

If one of the reacting species is in great excess, for instance $S_2(0) \gg S_1(0)$, the time law can be further simplified. The extent is controlled by the second component S_1 . For the abounding species S_2 holds

$$S_2(t) = S_2(0) - \epsilon(t) \approx S_2(0) \quad (3.35)$$

while (3.12) is valid for $S_1(t)$. With these approximations we obtain the differential equation

$$\begin{aligned} r(t) &= -\frac{dS_1(t)}{dt} = k S_1(t) S_2(t) \approx k S_1(t) S_2(0) \\ &\approx k' S_1(t) , \end{aligned} \quad (3.36)$$

where the new effective coefficient k' is the product of the original rate coefficient k and the concentration $S_2(0)$. It follows, that for this case the time law can be reduced to the monomolecular case treated in the previous section. Such reactions are referred to as *pseudo-monomolecular* or *kryptobimolecular*.

3.5.3 Bimolecular reaction of identical species

If both reactants are from the same species



we have a bimolecular reaction for two identical molecules. We write

$$r(t) = -\frac{1}{2} \frac{dS_1(t)}{dt} = -\frac{k [S_1(t)]^2}{2} \quad (3.38)$$

for the reaction rate. The prefactor 1/2 ensures that one gets the same rate for the reactants and products. If one uses the change of concentration, this condition is, in general, not fulfilled.

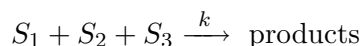
The differential equation can be solved by separation of variables. For the dynamic concentration $S_1(t)$ one obtains

$$S_1(t) = \frac{S_1(0)}{2S_1(0)kt + 1} , \quad (3.39)$$

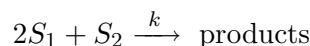
which is similar to (3.34). Both equations differ in a factor two in the denominator.

3.5.4 Trimolecular reactions

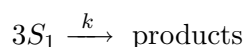
The reaction of three molecules to products is rare because the probability that three independent molecules collide at the same time or within a small time interval is very small. There are three possible ways for such a reaction. The first one is the reaction of three species



forming the products. In the second possibility



two molecules of species S_1 react with a third particle S_2 . Last but not least, three identical particles



can be transformed into the products.

Trimolecular reactions of different species

For the trimolecular reaction



of three different species the reaction rate is

$$r(t) = -\frac{dS_1(t)}{dt} = -\frac{dS_2(t)}{dt} = -\frac{dS_3(t)}{dt} = k S_1(t) S_2(t) S_3(t). \quad (3.41)$$

If one introduces the extent variable ϵ into the last equation

$$r(t) = \frac{d\epsilon(t)}{dt} = k [S_1(0) - \epsilon(t)] [S_2(0) - \epsilon(t)] [S_3(0) - \epsilon(t)] , \quad (3.42)$$

with initial concentrations $S_1(0)$, $S_2(0)$ and $S_3(0)$, it is possible to solve the differential equation by separation of variables. One obtains

$$\begin{aligned} kt = & \frac{1}{[S_1(0) - S_2(0)][S_3(0) - S_1(0)]} \ln \frac{S_1(t)}{S_1(0)} \\ & + \frac{1}{[S_1(0) - S_2(0)][S_2(0) - S_3(0)]} \ln \frac{S_2(t)}{S_2(0)} \\ & + \frac{1}{[S_2(0) - S_3(0)][S_3(0) - S_1(0)]} \ln \frac{S_3(t)}{S_3(0)} . \end{aligned} \quad (3.43)$$

If all participating species have stoichiometric concentrations⁶ we can simplify the approach ('ansatz'). Another known example is the so called 'product ansatz', where we assume that the solution of two parameter-dependent problem is separable into a product of two terms depending on one parameter only. In contrast, collision theory is an approach to the description of the temporal change of reacting species. For the reaction rate (3.41) with the relation $S_1(t) = S_2(t) = S_3(t)$ holding for all time. The resulting differential equation

$$r(t) = -\frac{dS_1(t)}{dt} = k [S_1(t)]^3 \quad (3.44)$$

⁶The ratio of concentrations is equal to the ratio of stoichiometric coefficients.

is easy to solve and one obtains for the concentrations

$$S_1(t) = \sqrt{\frac{S_1(0)^2}{1 + 2S_1(0)^2 kt}} \quad (3.45)$$

as function of time. The results for $S_2(t)$ and $S_3(t)$ follow, if we exchange S_1 for S_2 and S_3 in the same way.

Trimolecular reactions of two different species

The trimolecular reaction



describes the reaction of two molecules of species S_1 with one particle of species S_2 into products. For the reaction rate

$$r(t) = -\frac{1}{2} \frac{dS_1(t)}{dt} = -\frac{dS_2(t)}{dt} = k [S_1(t)]^2 S_2(t) . \quad (3.47)$$

Again, one substitutes the time dependent variables $S_1(t)$ and $S_2(t)$ by the corresponding relation between the initial values and the extent $\epsilon(t)$. The resulting differential equation

$$r(t) = \frac{d\epsilon(t)}{dt} = k [S_1(0) - \epsilon(t)]^2 [S_2(0) - \epsilon(t)] \quad (3.48)$$

can be solved by separation of variables. The result

$$kt = \frac{1}{2S_2(0) - S_1(0)} \left[\frac{S_1(0) - S_1(t)}{S_1(0)S_1(t)} + \frac{1}{2S_2(0) - S_1(0)} \right] \quad (3.49)$$

can be simplified, if one assumes stoichiometric concentrations. Then the relation $S_1(t) = 2S_2(t)$ has to be satisfied for all time. The differential equation is now

$$r(t) = -\frac{1}{2} \frac{dS_1(t)}{dt} = \frac{k [S_1(t)]^3}{2} . \quad (3.50)$$

Thus, the temporal evolution of the concentration of species S_1 obeys the same time law (3.45) as the species in the trimolecular reaction of three different particle for stoichiometric conditions. With the aim of the above relation between S_1 and S_2 one obtains

$$S_2(t) = \frac{1}{2} \sqrt{\frac{S_2(0)^2}{1 + S_2(0)^2 kt/2}} \quad (3.51)$$

for the second species S_2 .

Trimolecular reactions of three identical molecules

The reaction of three molecules of the same species



is the third possibility for a realization of a trimolecular reaction. The reaction rate is given as

$$r(t) = -\frac{1}{3} \frac{dS_1(t)}{dt} = k [S_1(t)]^3 . \quad (3.53)$$

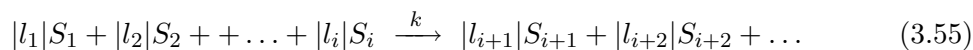
It follows

$$S_1(t) = \sqrt{\frac{S_1(0)^2}{1 + 6S_1(0)^2 kt}} \quad (3.54)$$

for the concentration of species S_1 as a function of time.

3.5.5 Higher and rational reaction orders

Up to now, we treated elementary reactions with one, two or three participating molecular species. The developed formalism shall expand to non-elementary reactions. For this purpose we revert to the general chemical reaction



and introduce the reaction rate as

$$\begin{aligned} r(t) &= k [S_1(t)]^{n_1} [S_2(t)]^{n_2} \dots [S_i(t)]^{n_i} [S_{i+1}(t)]^{n_{i+1}} [S_{i+2}(t)]^{n_{i+2}} \dots \\ &= k \prod_i [S_i(t)]^{n_i} . \end{aligned} \quad (3.56)$$

The total reaction order then is defined as sum over partial orders

$$n = \sum_i n_i , \quad (3.57)$$

which can take values greater than three, as well as rational values. Therefore, the so described reaction is in general not an elementary reaction. The partial reaction order n_i of each species, in general, do not coincide with the stoichiometric coefficients. The reaction rate obtained is an approximation and does not reflect the true reaction mechanism, but often it gives a first indication to the mechanism.

If one assumes stoichiometric concentrations, one is able to obtain some important quantities. The reaction rate is then

$$\frac{1}{l_1} \frac{dS_1(t)}{dt} = \frac{1}{l_2} \frac{dS_2(t)}{dt} = \dots = \frac{d\epsilon(t)}{dt} = k [S_1(0) - \epsilon(t)]^n , \quad (3.58)$$

where $\epsilon(t)$ is the extent variable. By separation of variables one obtains

$$\frac{S_1(t)}{S_1(0)} = \left(\frac{1}{1 + S_1(0)^{n-1} (n-1) kt} \right)^{1/(n-1)} , \quad (3.59)$$

for the normalized or dimensionless concentration $S_1(t)/S_1(0)$. The half life, defined as the time, where the normalized concentration is $S_1(t)/S_1(0) = 1/2$, can be determined from (3.59). After some algebraic transformations one obtains

$$t_{1/2} = \frac{2^{n-1} - 1}{(n-1)k} S_1(0)^{-n+1} . \quad (3.60)$$

The dependence from the initial concentration $S_1(0)$ in (3.60) allows the determination of the total reaction order n . In an experiment the half life can be measured as function of the initial concentration. A double logarithmic representation of the data gives a straight line with a slope $m = -n + 1$, which can be used to determine the total reaction order n . For such practical considerations and experimental techniques the reader is referred to [AdP02, Seg93, CB04].

Reactions of zeroth order

A special case of the generalized reaction is the reaction of “zeroth order”. For such a reaction the reaction rate is independent from the concentration. Choosing a zeroth order degradation of species S_1 as an example, the differential equation

$$- \frac{dS_1(t)}{dt} = k \quad (3.61)$$

reaction order	half life $t_{1/2}$
1	$t_{1/2} = \text{const.}$
2	$t_{1/2} \sim S_1(0)^{-1}$
3	$t_{1/2} \sim S_1(0)^{-2}$
1/2	$t_{1/2} \sim \sqrt{S_1(0)}$
3/2	$t_{1/2} \sim 1/\sqrt{S_1(0)}$

Table 3.1: The dependency of the half life $t_{1/2}$ from the initial concentration for the general chemical reaction (3.55) with stoichiometric conditions. $S_1(0)$ is the initial concentration of reactant S_1 .

is easy to integrate. The result is the linear function

$$S_1(t) = S_1(0) - kt, \quad (3.62)$$

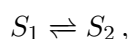
where the slope is given by the rate coefficient k . Reactions of zeroth order appear, if the rate is governed by a temporal constant non-chemical process. Examples include reactions on a surface, where the concentration of the reactant is constant by adsorption or a constant external flow of matter. The saturation of the enzyme complex in the enzyme kinetic reaction is a further example for such a behavior. Often, this approximation is used to simplify kinetic equations. The transient time, the reaction takes to reach this state and at the end of the reaction, is assumed as small compared to the saturated state.

At the end of this section we have to make an important comment on this treatment of (bio)chemical reactions. Within this framework one describes the dependency between the reactants and the products without a proper model of the reaction mechanism. Hence, the total order n cannot be interpreted as number of participating molecules in an elementary reaction. But it is an easy and fast way to find a functional relation between the reactants and the products.

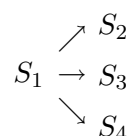
3.6 Complex Reactions

In the previous section we introduced the concept of elementary reactions and demonstrated their properties. As mentioned before, we assume that chemical reactions consist of a set of elementary reactions. In the following section we want to classify some basic complex reactions and describe their properties. Before we go into details, there are three basic kinds of complex reactions:

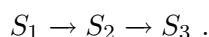
1. reversible reactions,



2. parallel reactions,



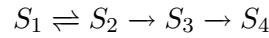
3. consecutive reactions,



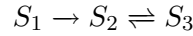
More complex reactions can be composed of these three basic classes. Some possible reaction schemes are:

- Consecutive reactions with reversible parts

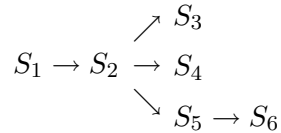
- pre-equilibrium



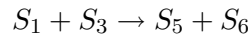
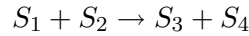
- downstream equilibrium



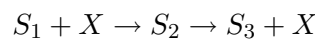
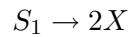
- Consecutive reactions combined with parallel reactions



- Competitive consecutive reactions



- Closed consecutive reactions (chain reactions)



We can continue this list with more complicated complex reactions, but in order to convey an idea of the complexity of reaction mechanisms this short overview is sufficient.

Each step in a complex reaction scheme is represented by an elementary reaction, which can be described by a differential equation. Because steps are not independent of each other, we now obtain a system of coupled differential equations. In general, there is no analytical solution for such systems. One has to use numerical methods to solve, simulate, the differential equations. In these notes we use some analytical solvable simple examples to demonstrate the properties of complex reactions.

3.6.1 Reversible Reactions

A reversible reaction consists of two elementary reactions, the forward reaction and the reverse reaction. Both are characterized by rate coefficients. The simplest example for such a reaction is the monomolecular reversible reaction



where the forward and the backward reaction are first-order reactions. The rate coefficient of the forward reaction is k_1 and the coefficient of the backward reaction is k_{-1} . If we assume the initial concentration $S_1(0) \doteq S_0$ and $S_2(0) = 0$ the conservation law

$$S_1(t) + S_2(t) = S_1(0) \quad (3.64)$$

must hold for all times. The corresponding reaction rate is obtained from the difference

$$r(t) = r_1(t) - r_{-1}(t) \quad (3.65)$$

of the forward and the backward reaction. With the use of the representation of elementary reactions we obtain the differential equations

$$\begin{aligned} r &= -\frac{dS_1(t)}{dt} = k_1 S_1(t) - k_{-1} S_2(t) \\ &= \frac{dS_2(t)}{dt} = k_1 S_1(t) - k_{-1} S_2(t) = -\frac{dS_1(t)}{dt} \\ &= \frac{d\epsilon(t)}{dt} = k_1 [S_1(0) - \epsilon(t)] - k_{-1} [S_2(0) - \epsilon(t)] \end{aligned} \quad (3.66)$$

for the reaction rate. The first term on the right-hand-side corresponds to the production of S_2 in a monomolecular reaction from S_1 and the second term to the analogue reverse reaction. From (3.66) it follows, that there is a point, where the reaction rate is zero. The forward and the backward reaction are balanced at this point, the same amount of S_2 is produced the same amount of S_2 as is transformed back to S_1 . From this we see that a reaction rate of zero does not mean, that nothing happens⁷. A net change of concentration is not measurable from a macroscopic point of view, but in a microscopic view the reactions are still going on. This special state is called *chemical equilibrium*. For our example of a reversible reaction, this state is defined as

$$0 = k_1 S_1(t) - k_{-1} S_2(t). \quad (3.67)$$

More general, the right-hand-side must be zero. This mathematical condition is a necessary but not a sufficient constraint, especially for complex systems. It has to fulfill some more physical conditions, by which we decide if it is a stable or an unstable state. Only the stable state is referred to as equilibrium state (see also Chapter 9). The solution of equation (3.67) is

$$K_{\text{eq}} = \frac{k_1}{k_{-1}} = \frac{S_{2,\text{eq}}}{S_{1,\text{eq}}} \quad (3.68)$$

defines the equilibrium constant K_{eq} as the ratio of the rate coefficient of the forward and the rate coefficient of the backward reaction. This quantity measures the affinity of S_1 to transform to S_2 . The equation (3.68) is the famous *law of mass action* for an monomolecular reaction. Equations (3.66) are kinetic rate equations. The *law of mass action* is a result of these equations and the assumptions behind them. To call this representations ‘law of mass action’ can therefore be misleading. The term ‘generalized law of mass action’ is sometimes used. For S_2 we define the dissociation constant K_d describing the process of the backward transformation into the reactants. It is the reciprocal of the equilibrium constant K_{eq} . From (3.68) it follows that in equilibrium the ratio of the concentrations match the equilibrium constant. The corresponding concentrations are called equilibrium concentrations $S_{1,\text{eq}}$ and $S_{2,\text{eq}}$. The species with a higher production rate has the higher equilibrium concentration.

The temporal evolution of the reversible reaction (3.63) can be solved analytically using the third differential equation in (3.66). The integration over the extent variable $\epsilon(t)$ can be carried out by separation of variables. We obtain the integrals

$$\int_0^{\epsilon(t)} d\epsilon(t) \left[(K_{\text{eq}} + 1) \left(\frac{K_{\text{eq}} S_1(0) - S_2(0)}{K_{\text{eq}} + 1} - \epsilon \right) \right]^{-1} = \int_0^t k_{-1} dt. \quad (3.69)$$

After integration and some manipulations we get the result

$$\begin{aligned} \epsilon(t) &= \frac{k_1 S_1(0) - k_{-1} S_2(0)}{k_1 + k_{-1}} \left[1 - \exp \{ - (k_1 + k_{-1}) t \} \right] \\ &= \epsilon_{\text{eq}} \left[1 - \exp \{ - (k_1 + k_{-1}) t \} \right] \end{aligned} \quad (3.70)$$

⁷Except at time $t = 0$, where the reaction was started. At this point, we postulate that nothing happened before.

for the extent variable. The extent increases exponentially to its equilibrium value. The relaxation time $\tau = (k_1 + k_{-1})^{-1}$ is a measure how fast the reaction tends to the equilibrium. The time evolution for the reactant S_1 follows as

$$S_1(t) = S_{1,\text{eq}} + (S_1(0) - S_{1,\text{eq}}) \exp \{ -(k_1 + k_{-1})t \} , \quad (3.71)$$

with initial concentration $S_1(0)$ and equilibrium concentration $S_{1,\text{eq}}$. From the conservation law

$$S_1(t) + S_2(t) = S_1(0) ,$$

where we assume that $S_2(0) = 0$, we obtain

$$S_2(t) = (S_1(0) - S_{1,\text{eq}}) \left[1 - \exp \{ -(k_1 + k_{-1})t \} \right] \quad (3.72)$$

for the product. Both, reactant and products, reach their equilibrium concentration, exponentially. This behavior is shown in Figure 3.6. Additionally, the equilibrium values are drawn as horizontal lines for comparison. We choose $k_1 > k_{-1}$ for this example, hence the equilibrium state of S_2 is higher than the equilibrium state of the reactant S_1 .

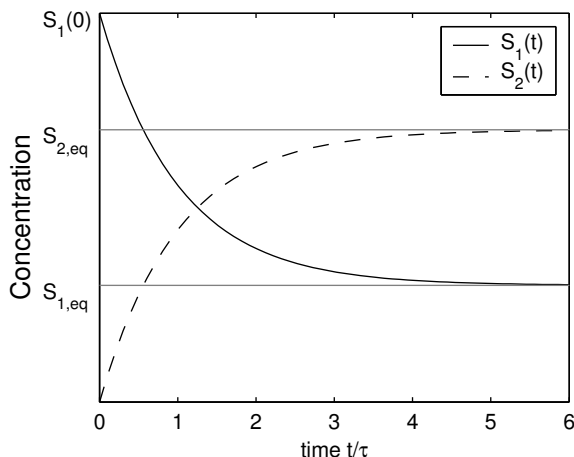


Fig. 3.6: The monomolecular reversible reaction as function of the relaxation time τ . The solid line represent the dynamic change of the reactant and the dashed line the change of the product to their equilibrium states shown as horizontal lines. The ratio of the equilibrium states is given by the law of mass action (3.68).

The rate coefficients of a monomolecular reversible reaction can be determined from a logarithmic representation of $(S_1(0) - S_{1,\text{eq}})/(S_1(t) - S_{1,\text{eq}})$ and the equilibrium with (3.68). The slope of the logarithmic plot is proportional to the inverse relaxation time $(k_1 + k_{-1})$. The determination of the rate coefficients k_1 and k_{-1} requires the measurement of the dynamic change of concentrations and the measurement of equilibrium data.

The treatment of reversible reactions can be generalized to higher reaction orders. As an example we choose a bimolecular forward- and backward reaction



We assume, that only the reactants are present in stoichiometric amounts at the beginning of the reaction. Furthermore we use the extent variable $\epsilon(t)$ to simplify the kinetic equation.

From this follows the conservation law

$$S_1(0) + S_2(0) = 2S_1(0) = 2S_{1,\text{eq}} + \epsilon_{\text{eq}} . \quad (3.74)$$

t	S_1	S_2	S_3	S_4
0	$S_1(0)$	$S_1(0)$	0	0
t	$S_1(0) - \epsilon(t)$	$S_1(0) - \epsilon(t)$	$\epsilon(t)$	$\epsilon(t)$
∞	$S_1(0) - \epsilon_{\text{eq}}$	$S_1(0) - \epsilon_{\text{eq}}$	ϵ_{eq}	ϵ_{eq}

Table 3.2: Relations between the reactants and products to the extent variable ϵ for the reversible reaction (3.73).

The time law is

$$\begin{aligned}
 r(t) &= -\frac{dS_1(t)}{dt} = -\frac{dS_2(t)}{dt} = \frac{d\epsilon(t)}{dt} \\
 &= k_2 S_1(t) S_2(t) - k_{-2} S_3(t) S_4(t) \\
 &= k_2 [S_1(0) - \epsilon(t)]^2 - k_{-2} [\epsilon(t)]^2
 \end{aligned} \tag{3.75}$$

is analytically solvable. The integration by expansion into partial fractions gives

$$\ln \frac{\epsilon(t) [S_1(0) - 2\epsilon_{\text{eq}}] + S_1(0)\epsilon_{\text{eq}}}{S_1(0) [\epsilon_{\text{eq}} - \epsilon(t)]} = \frac{2S_1(0) [S_1(0) - \epsilon_{\text{eq}}]}{\epsilon_{\text{eq}}} k_2 t \tag{3.76}$$

after numerous rearrangements. We can use it as an instruction for the determination of the rate coefficient k_2 . The second coefficient is obtained again from the equilibrium solution

$$K_{\text{eq}} = \frac{S_{3,\text{eq}} S_{4,\text{eq}}}{S_{1,\text{eq}} S_{2,\text{eq}}} = \frac{\epsilon_{\text{eq}}^2}{[S_1(0) - \epsilon_{\text{eq}}]^2} = \frac{k_2}{k_{-2}}, \tag{3.77}$$

as a generalization of the law of mass action for the bimolecular reversible reaction (3.73).

3.6.2 Parallel Reactions

In a parallel reaction, several reactions of the same reactants proceed side by side but produce different end products. Each reaction mechanism can consist of an elementary reaction, a reversible reaction, or a more complex mechanism. The reaction of one species with several partners in the reaction volume is a competitive reaction and no parallel reaction. We will discuss this towards the end of this section.

In order to illustrate some elementary properties of side reactions we consider the first-order reaction



composed of three irreversible monomolecular reactions. The species S_1 is converted in a monomolecular reaction into products S_2 , S_3 , and S_4 . The term $k_i dt$ is the probability that the i -th reaction occurs in the time interval dt . The resulting kinetic equation is

$$\begin{aligned}
 r(t) &= -\frac{dS_1(t)}{dt} = k_1 S_1(t) + k_2 S_1(t) + k_3 S_1(t) \\
 &= k S_1(t),
 \end{aligned} \tag{3.79}$$

where k is the total rate coefficient

$$k = \sum_i k_i \tag{3.80}$$

and $k dt$ the probability, that one of the reactions takes place. The solution of such a differential equation is known from the discussion of monomolecular reaction. In analogy, we obtain for the temporal evolution an exponential expression

$$S_1(t) = S_1(0) \exp\{-kt\}, \quad (3.81)$$

where $S_1(0)$ is the initial concentration. The temporal evolution of the reactant is indistinguishable from the evolution of a monomolecular reaction with the same rate coefficient. More generally, the dynamic change of the reactants is for the same coefficient k independent of the number of products.

The creation of products S_2 , S_3 , and S_4 is governed by the kinetic equations

$$\frac{dS_2(t)}{dt} = k_1 S_1(t), \quad \frac{dS_3(t)}{dt} = k_2 S_1(t), \quad \frac{dS_4(t)}{dt} = k_3 S_1(t), \quad (3.82)$$

which can be transformed with (3.81) into

$$\frac{dS_2(t)}{dt} = k_1 S_1(0) \exp\{-kt\}, \quad \dots \quad (3.83)$$

Integration leads the exponentials

$$S_2(t) = \frac{k_1}{k} [1 - \exp\{-kt\}] S_1(0) + S_2(0), \quad (3.84)$$

$$S_3(t) = \frac{k_2}{k} [1 - \exp\{-kt\}] S_1(0) + S_3(0), \quad (3.85)$$

$$S_4(t) = \frac{k_3}{k} [1 - \exp\{-kt\}] S_1(0) + S_4(0). \quad (3.86)$$

The comparison of the transient concentrations⁸ shows a further important property of parallel reactions. If we calculate the ratios of the dynamic concentrations

$$[S_2(t) - S_2(0)] : [S_3(t) - S_3(0)] : [S_4(t) - S_4(0)] = k_1 : k_2 : k_3 \quad (3.87)$$

we obtain the ‘Principle of Wegscheider’. It says, that the ratio is equal to the ratio of the rate coefficients k_i and constant.

The time evolution for the parallel reaction (3.78) is shown in Figure 3.7, for the special case of $S_2(0) = S_3(0) = S_4(0) = 0$. For these initial conditions we can expand the conservation law to

$$S_1(0) = S_1(t) + S_2(t) + S_3(t) + S_4(t). \quad (3.88)$$

At the end of the reaction, species S_1 is completely consumed and (3.88) is determined by the final concentration $S_{2,f}$, $S_{3,f}$, and $S_{4,f}$. We obtain with (3.84)-(3.86) for these concentrations

$$\begin{aligned} S_{2,f} &= \frac{k_1}{k} S_1(0), \\ S_{3,f} &= \frac{k_2}{k} S_1(0), \\ S_{4,f} &= \frac{k_3}{k} S_1(0). \end{aligned}$$

The concentrations are determined by the ratio of the individual rate coefficient k_i and the total rate coefficient k (3.80). The exponential decrease of S_1 and the exponential increase

⁸The prefix transient distinguishes between the produced concentration in the temporal process of the reaction and the initial concentrations. In this sense the full concentration is given by a transient and an initial part.

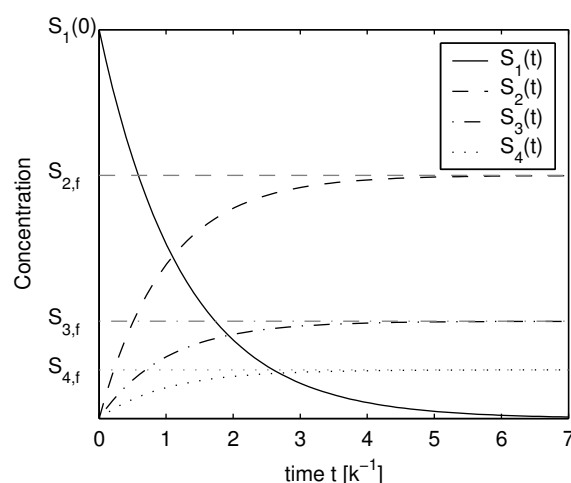


Fig. 3.7: The temporal evolution of the reactant S_1 and the products S_2, S_3, S_4 of the parallel reaction (3.78) in units of the initial concentration $S_1(0)$. The horizontal lines denote the final concentration $S_{2,f}, S_{3,f}$, and $S_{4,f}$ of the products.

of the product is clearly visible in the figure. For comparison, the product concentrations at the end of the reaction are also shown. The reaction $S_1 \rightarrow S_2$ is the fastest reaction, hence it is the main product of the side reaction (3.78).

These results are valid for parallel reactions in general. Recapitulating, parallel reactions have the following behavior

- The time evolution of the reactants is independent from the number of products and has the same order as each elementary reaction. It is determined by the total rate coefficient (3.80), only.
- The fastest side reaction is strongest participant on the extent. It controls the main product.
- The products are build in the ratio of their rate coefficient (Principle of Wegscheider).

Differentiation of parallel reactions and competitive reactions

In our discussion of parallel reactions we tried to show how a parallel reaction can be separated from a competitive reaction in an experiment. In a parallel reaction the same reactants can be transformed in different products. In a competitive reaction, for instance, the reactant S_1 participates in two independent reactions with different partners. For instance, let us consider two bimolecular reactions



with the common reactant S_1 . The corresponding system of coupled differential equations is

$$\begin{aligned}\frac{dS_1(t)}{dt} &= -(k_1 S_2(t) + k_2 S_3(t)) S_1(t), \\ \frac{dS_2(t)}{dt} &= -k_1 S_1(t) S_2(t), \\ \frac{dS_3(t)}{dt} &= -k_2 S_1(t) S_3(t), \\ \frac{dS_4(t)}{dt} &= k_1 S_1(t) S_2(t), \\ \frac{dS_5(t)}{dt} &= k_2 S_1(t) S_3(t).\end{aligned}\tag{3.90}$$

If we use Wegscheider's principle in differential form, we obtain with

$$\frac{dS_4(t)}{dS_5(t)} = \frac{k_1 S_1(t) S_2(t)}{k_2 S_1(t) S_3(t)} = \frac{k_1 S_2(t)}{k_2 S_3(t)}\tag{3.91}$$

an expression that depends on the present concentration of S_2 and S_3 . For a parallel reaction this ratio has to be a constant.

3.6.3 Consecutive Reactions

An important class of complex reactions are consecutive reactions. These reactions include one or more unstable intermediates. Some partial reactions can be reversible and there can be arborization as a result of parallel reaction.

To simplify matters we consider an unidirectional and monomolecular sequence of reactions



which is like a radioactive decay sequence, with initial condition $S_2(0) = S_3(0) = 0$. The decay of S_1 into the intermediate S_2 is governed by the differential equation

$$\frac{dS_1(t)}{dt} = -k_1 S_1(t),\tag{3.93}$$

with the known solution

$$S_1(t) = S_1(0) \exp\{-k_1 t\}.\tag{3.94}$$

The differential equation for S_2 is more complicated. It consists of two parts

$$\frac{dS_2(t)}{dt} = k_1 S_1(t) - k_2 S_2(t),\tag{3.95}$$

where the first term describes the production of S_2 from S_1 and the second term the decay into the final product S_3 . With (3.94) we transform this equation into a first-order linear differential equation that can be solved analytically:

$$\frac{dS_2(t)}{dt} = k_1 S_1(0) \exp\{-k_1 t\} - k_2 S_2(t).\tag{3.96}$$

We obtain for S_2

$$S_2(t) = \frac{k_1}{k_2 - k_1} S_1(0) [\exp\{-k_1 t\} - \exp\{-k_2 t\}].\tag{3.97}$$

The temporal evolution has a maximum at

$$t_{\max} = \frac{\ln(k_1/k_2)}{k_1 - k_2}.\tag{3.98}$$

This is a typical property for an unstable intermediate. In the beginning S_1 decays faster due to its higher concentration, whereas S_2 decays slowly. The result is an increase of the intermediate. In the course of the reaction the rate of decay of S_1 decreases. At t_{\max} the production rate and the decay rate of the intermediate are equal. After this time the decay rate is faster and S_2 decreases, too. The corresponding concentration is

$$S_{2,\max} = S_2(t_{\max}) = \left(\frac{k_1}{k_2}\right)^{\frac{k_2}{k_2-k_1}}. \quad (3.99)$$

If we use (3.97) the differential equation for the final product is

$$\begin{aligned} \frac{dS_3(t)}{dt} &= k_2 S_2(t) \\ &= \frac{k_1 k_2}{k_2 - k_1} S_1(0) [\exp\{-k_1 t\} - \exp\{-k_2 t\}]. \end{aligned} \quad (3.100)$$

It is proportional to the time law of S_2 . Hence, the reaction rate of S_3 has a maximum at t_{\max} . Remember we assumed $S_2(0) = S_3(0) = 0$, the conservation law

$$S_1(0) = S_1(t) + S_2(t) + S_3(t), \quad (3.101)$$

permits an easier way to solve the temporal evolution. The result

$$S_3(t) = S_1(0) \left[1 - \frac{k_2 \exp\{-k_1 t\} - k_1 \exp\{-k_2 t\}}{k_2 - k_1} \right] \quad (3.102)$$

changes the sign of its second derivative at t_{\max} . This behavior results in a typical sigmoidal shape of the time evolution of the final product, as shown in Figure 3.8. There the time evolution for all three participating species is compared. The starting substance S_1 decreases with the exponential decay law (3.94). As mentioned before, intermediate S_2 first increases, goes through a maximum and finally decreases. The time at which the maximum occurs and its value are plotted by thin dashed lines. Both are dependent on the ratio of the rate coefficient k_1 and k_2 . The time evolution of the final product S_3 is a monotonously increasing function with the predicted inflection point at t_{\max} .

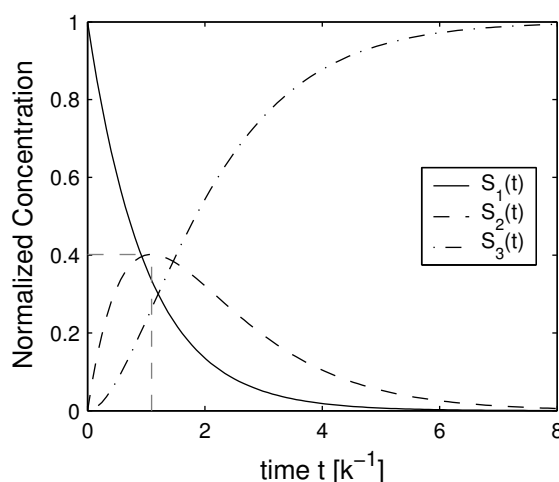


Fig. 3.8: The temporal evolution of consecutive reaction (3.92) as a function of the dimensionless time $k_1 t$. On the y-axis the normalized concentrations $S_1(t)/S_1(0)$, $S_2(t)/S_1(0)$, and $S_3(t)/S_1(0)$ are plotted. Additionally the time t_{\max} (3.98) and the maximum value of the intermediate $S_{2,\max}$ (3.99) are shown.

The radioactive decay sequence (3.92) is a simple example for a consecutive reaction. Reactions of higher order or/and higher complexity lead to more complicated kinetic equations and are often only numerically or approximately solvable.

Rate-determining steps

If one reaction is much slower than the others, we can simplify the formalism for a consecutive reaction. We consider again the radioactive decay sequence (3.92). First, let us assume the case where $k_1 \gg k_2$, i.e., the second reaction is slower than the first. Now we expand the prefactor⁹ in equation (3.97)

$$\frac{k_1}{k_2 - k_1} = -\frac{1}{1 - k_2/k_1} \approx \left(-1 - \frac{k_2}{k_1} - \mathcal{O}\left(\frac{k_2^2}{k_1^2}\right) \right) \approx -1, \quad (3.103)$$

where the symbol $\mathcal{O}(\dots)$ denotes the order in respect of the expansion parameter of first neglected term of the expansion. The sign corresponds to the trend of contribution. Furthermore we compare the exponential functions within the brackets. If $k_1 t$ is always much greater than $k_2 t$, we obtain

$$\exp\{-k_2 t\} \gg \exp\{-k_1 t\} \quad (3.104)$$

and neglect the k_1 dependent exponential function against the k_2 dependent exponential. With these approximations we obtain for the intermediate the new time law

$$S_{2,k_2}(t) = S_1(0) \exp\{-k_2 t\} \quad (3.105)$$

and for the final product

$$S_{3,k_2}(t) = S_1(0) [1 - \exp\{-k_2 t\}]. \quad (3.106)$$

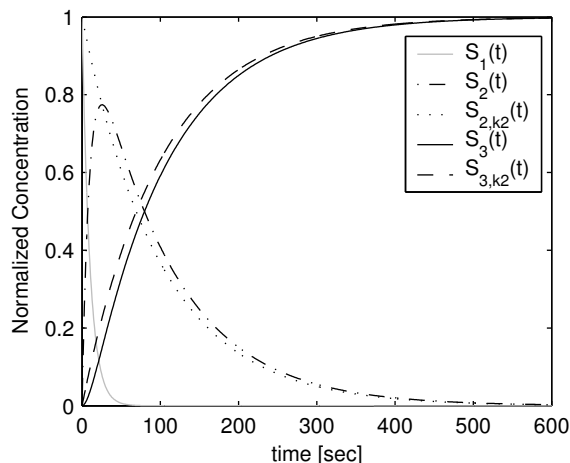


Fig. 3.9: Comparison of the full solution for the consecutive reaction (3.92) and the approximation of the rate-determining step for a slow second partial reaction. After a short starting time, the approximation shows good agreement with the full solution. We choose $k_1 = 0.1 \text{ s}^{-1}$ and $k_2 = 0.01 \text{ s}^{-1}$ for this example. The concentrations are normalized to the initial concentration $S_1(0)$.

Within this approximation, the temporal evolution is determined by the slow decay of the intermediate S_2 into the product S_3 . The subscript k_2 shall denote this property and distinguish between the full and the approximative solution. We compare both solutions in Figure 3.9, where our main focus is on the final product. The decay of S_1 is kept unchanged. It is almost completely transformed into the intermediate before the intermediate decays

⁹

$$\frac{1}{1-x} \approx 1 + x + x^2 + \dots \quad \text{for } x \ll 1$$

into S_3 . After a short starting time, the principle of the rate-limiting step is a good approximation.

In a second part we now assume, that the decay of S_1 is the slowest step of sequence reaction (3.92). This means $k_1 \ll k_2$. We again expand the prefactor

$$\frac{k_1}{k_2 - k_1} = \frac{k_1}{k_2} \frac{1}{1 - k_1/k_2} \approx \frac{k_1}{k_2} \left[1 + \frac{k_1}{k_2} + \mathcal{O}\left(\frac{k_1^2}{k_2^2}\right) \right] \approx \frac{k_1}{k_2} \quad (3.107)$$

and the comparison of the exponentials gets

$$\exp\{-k_1 t\} \gg \exp\{-k_2 t\}. \quad (3.108)$$

We can now simplify equations (3.97) and (3.102), whereas the equation for S_1 (3.94) remains unchanged. Within this approximation the evolution of the intermediate is governed by

$$S_{2,k_1}(t) = S_1(0) \frac{k_1}{k_2} \exp\{-k_1 t\} \quad (3.109)$$

and of the final product by

$$S_{3,k_1}(t) = S_1(0) [1 - \exp\{-k_1 t\}] = S_1(0) - S_1(t). \quad (3.110)$$

Again, the subscript denotes the rate-determining step. The comparison of the approximation and the full solution is shown in Figure 3.10. The production of the intermediate is much slower than its decay. A molecule from species S_2 is transformed practically immediately into the final product. The concentration of the intermediate is always small in comparison to S_1 and S_3 . Also in this case, the approximation gives a good description of the reaction except the very first time.

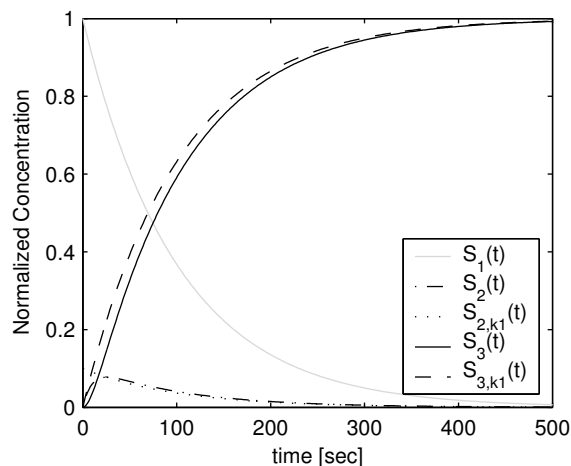
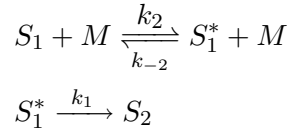


Fig. 3.10: Comparison of the full solution and the approximation for a fast second step. The rate coefficients are $k_1 = 0.01 \text{ s}^{-1}$ and $k_2 = 0.1 \text{ s}^{-1}$. The concentration of the intermediate is always small in comparison to other species. It is transformed practically immediately into S_3 .

We now give a more general discussion of the rate limiting step. If one step of a consecutive reaction is much slower than the others, it determines the total rate of reaction or more precisely, the rate of all following steps. This reaction is called the rate-determining step. The total reaction rate is the rate of the production of the final product of the sequence. This statement includes a further property of consecutive reactions. The reaction rate is not equal for each step. Each partial reaction before the rate-determining step has its own rate depending on the specific reaction and the steps before. The approximation

of the rate-determining step defines a limiting total reaction rate for the considered sequence. The reaction rate for the described system cannot be faster than this limit. This is a direct consequence from the principle of cause and effect. The effect cannot occur before the cause, you cannot create a new species without its components. If there are faster steps before the rate-determining step we obtain a congestion of its participating molecules. If the following steps are faster they have to wait on the slower reaction. However, the rate-determining step is not just the slowest step: it must be slow and be a crucial gateway for the formation of products. If a faster step also lead to products, the slowest step is irrelevant because the slow reaction can be sidestepped. To finish our discussion we have to give a criteria which step in a consecutive reaction is the slowest.

For this purpose we discuss a more sophisticated model for an monomolecular reaction. It was introduced by Lindemann and experimentally verified by Hinshelwood [AdP02]. Starting point for the model is the idea that a single molecule is excited by a collision. The excited molecule can loose its energy by a new collision or decays in a monomolecular step into the product.



The molecule M can be from species S_1, S_2 or an inert-gas. Here, an inert-gas is a gas of other species, which do not react with the considered molecules S_1, S_2 , and S_1^* . Collisions between two excited molecules S_1^* are negligible in comparison to the others. But a collision is a clear bimolecular process, how we can get a first-order kinetics from this mechanism? The rates for the components follow the system of coupled differential equations

$$\frac{dS_1(t)}{dt} = -k_2 S_1(t) M(t) + k_{-2} S_1^*(t) M(t), \quad (3.111)$$

$$\frac{dS_1^*(t)}{dt} = k_2 S_1(t) M(t) - k_{-2} S_1^*(t) M(t) - k_1 S_1^*(t) \quad (3.112)$$

$$\frac{dS_2(t)}{dt} = k_1 S_1^*(t). \quad (3.113)$$

A simplification arises, if we use a steady state approximation for the excited molecule S_1^* . From (3.112), it follows

$$S_1^*(t) = \frac{k_2 S_1(t) M(t)}{k_{-2} M(t) + k_1}, \quad (3.114)$$

which we can insert into (3.111) or (3.113). The result for the reaction rate is

$$-\frac{dS_1(t)}{dt} = \frac{dS_2(t)}{dt} = \frac{k_1 k_2 M(t) S_1(t)}{k_{-2} M(t) + k_1} \quad (3.115)$$

for which no reaction order can be defined. For a closed system is $M(t) \approx \text{const.}$ and the rate equation transforms to a pseudo-first order law

$$-\frac{dS_1(t)}{dt} = \frac{dS_2(t)}{dt} = k_{\text{eff}} S_1(t), \quad (3.116)$$

where the effective rate coefficient is

$$k_{\text{eff}} = \frac{k_1 k_2 M}{k_{-2} M + k_1}. \quad (3.117)$$

The effective constant depends on the collision partners M .

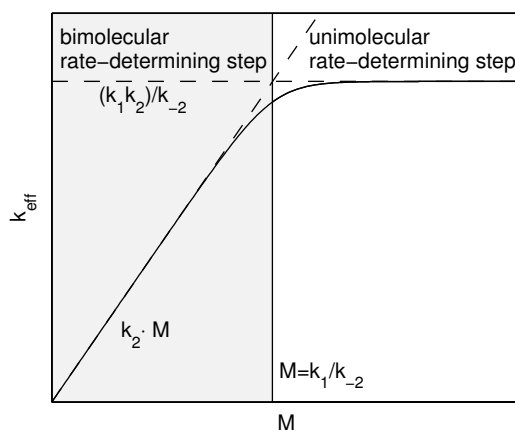


Fig. 3.11: Log-log plot for the effective rate constant k_{eff} as function of the collision partner M . For $k_2 M \ll k_1$, see (3.117), the reaction follows second-order kinetics determined by the bimolecular excitation of S_1 and for $k_2 M \gg k_1$ first-order kinetics with the monomolecular decay into the product S_2 as product. We use the intersection of the asymptotes of k_{eff} to separate the M - k_{eff} -plain into a bimolecular and monomolecular limited region.

Now, we go back to our original question and determine the rate-determining step. For that purpose we consider different amounts of M . First we assume $k_{-2} M \ll k_1$. The rate law simplifies to

$$-\frac{dS_1(t)}{dt} = \frac{dS_2(t)}{dt} = k_2 M S_1(t), \quad (3.118)$$

which obeys second-order kinetics. The rate-determining step is the bimolecular collision of S_1 and M . On the other hand, if $k_{-2} M \gg k_1$ we obtain a rate of first order

$$-\frac{dS_1(t)}{dt} = \frac{dS_2(t)}{dt} = \frac{k_1 k_2}{k_{-2}} S_1(t). \quad (3.119)$$

The rate-determining step is now the monomolecular decay of excited molecules into products.

From this example we are able to define the rate-determining step. The rate-determining step is the step with the smallest probability to occur within the time interval dt or the with the smallest rate, respectively. The rate coefficient alone is not a proper criteria. Only in special cases it is the process with the smallest rate coefficient. Remember, the rate for an elementary reaction is the product of the rate coefficient and the participating species. As we demonstrated the rate-determining step depends on the current conditions in the reaction volume. For complex reactions often one cannot define an unique rate-determining step. Instead one has to consider different regimes as we have done it for the sophisticated model of monomolecular reactions. Furthermore, the rate-determining step can be time dependent. Complex reactions with (auto)catalytic reactions are a typical example for such a behavior.

The (quasi-)steady state approximation

The second case of a preliminary rate-determining step can be expand to a more rigorous approximation. Let S_2 an unstable intermediate with a short lifetime and small concentration in comparison to the other participants on the sequence. Its consumption is approximatively simultaneously with its production. In this case we can assume for its rate

$$\frac{dS_2(t)}{dt} \approx 0. \quad (3.120)$$

It follows that the balance equation is given as

$$\frac{dS_2(t)}{dt} = k_1 S_1(t) - k_2 S_2(t) \approx 0, \quad (3.121)$$

leading to the relation

$$k_1 S_1(t) = k_2 S_2(t). \quad (3.122)$$

From this we obtain

$$S_2(t) = \frac{k_1}{k_2} S_1(t) \quad (3.123)$$

and finally from (3.100) the simplified time law

$$\frac{dS_3(t)}{dt} = k_1 S_1(t) \quad (3.124)$$

for the final product. An integration leads again to 3.110.

At the end of this section we generalize the steady state approximation. Within this approximation we assume, that during the major part of the reaction the rate of change of intermediates are negligible small¹⁰. Thereby we neglect an initial time period, where the intermediates rise from zero. In our previous examples, see Figure 3.9 and 3.10, the first 25 seconds cannot be described with the approximation of the rate-determining step. For the next minutes it is a good description. But a ratio of 10(0.1) is not a great difference between the rate coefficient. We chose this ratio for demonstration purposes only.

The amount of the intermediates do not need to be negligible in comparison to the reactants and products as we assume in our example. This more restrictive approximation is often called a quasi-steady state.

The steady state approximation is a very powerful tool in the analytic treatment of complex reactions. Because of the increasing mathematical complexity reaction schemes involving many steps is nearly always analytically unsolvable. One approach is a numerical solution of the differential equation. An alternative approach is to make an approximation. On the other hand, an approximation restricts the range of validity of the model. For instance, the steady state approximation in (3.112) assumes that the formation of the excited molecule and its decay back into the deactivated form are much faster than the formation of the product. This is only possible if $k_{-2} \gg k_1$, but not when $k_{-2} \ll k_1$.

Hence, the approximations of steady state and rate-determining step have to be used carefully. Their validity is limited and has to be rechecked for each specific reaction system and its parameters. An usage far away from the validity region leads to wrong results.

3.6.4 Autocatalytic Reactions

catalyst A particular class of reactions are catalytic reactions. A *catalyst* accelerates the reaction and is released unmodified as product. It occurs as reactant and as product. Note, a catalyst can affect only reactions happening from alone. In other words, the reactions have to be possible from thermodynamic reasons. It also does not change the equilibrium properties. In this section we want to discuss a special kind of catalytic reactions, the autocatalytic reaction. In these reactions a product accelerates its own production. In contrast to the other catalytic reactions we have not to add an additional substance. Known examples for such reactions are chain reactions¹¹. Thereby, one distinguishes between two types of (auto)catalytic reactions¹². The catalyst can increase the reaction

¹⁰Here negligible small means small in comparison to the rates of the other participating species.

¹¹In a chain reaction substances are involved recycling the reactands and starting a new reaction cycle.

¹²An acceleration is a change of velocity, not necessarily an increase.

rate, unfortunately, this behavior is called (auto)catalysis, too. The inverse effect of a decrease of reaction rate is called (auto)inhibition. The corresponding substance is an (auto)inhibitor.

For an introduction to autocatalysis we choose the simplest possible model, an unidirectional monomolecular reaction. The stoichiometric formula is



where the subscript S_2 denotes, that the product S_2 acts as autocatalyst. In a more detailed representation one often uses the formula



for an autocatalytic reaction. According to the second chemical formula the kinetic equation is

$$\begin{aligned} r(t) &= -\frac{dS_1(t)}{dt} = \frac{dS_2(t)}{dt} = \frac{d\epsilon(t)}{dt} \\ &= k S_1(t) S_2(t) \\ &= k [S_1(0) - \epsilon(t)] [S_2(0) + \epsilon(t)]. \end{aligned} \quad (3.127)$$

The autocatalytic effect of the product S_2 formally increase the reaction order. But be aware, this is only an ‘ansatz’ not a full description of the reaction mechanism. The mechanism of catalytic reaction is often complicated, hence we forbear to specify it. Instead, we discuss some basic properties in our model.

The autocatalytic reaction fulfills the conservation law

$$S_1(0) = S_1(t) + S_2(t) - S_2(0), \quad (3.128)$$

where we assume an initial concentration of the autocatalyst. Kinetic equations (3.127) are solvable by separation of variables. The integral

$$kt = \int_0^{\epsilon(t)} \frac{d\epsilon(t)}{[S_1(0) - \epsilon(t)][S_2(0) + \epsilon(t)]} \quad (3.129)$$

is solved by an expansion into partial fractions. We obtain

$$kt = \frac{1}{S_1(0) + S_2(0)} \ln \frac{S_1(0) S_2(t)}{S_2(0) S_1(t)} \quad (3.130)$$

from which we get

$$S_1(t) = \frac{S_1(0) + S_2(0)}{S_2(0)/S_1(0) \exp \left\{ [S_1(0) + S_2(0)] kt \right\} + 1} \quad (3.131)$$

and

$$S_2(t) = \frac{S_1(0) + S_2(0)}{S_1(0)/S_2(0) \exp \left\{ - [S_1(0) + S_2(0)] kt \right\} + 1} \quad (3.132)$$

for the reactant and the product as function of time. Both functions have an inflection point typically for autocatalytic reactions at

$$t_{\text{ip}} = \frac{1}{[S_1(0) + S_2(0)] k} \ln \frac{S_1(0)}{S_2(0)}. \quad (3.133)$$

The associated concentrations are

$$S_{1,\text{ip}} = \frac{S_1(0) + S_2(0)}{2} = S_{2,\text{ip}} . \quad (3.134)$$

Also, the reaction rate shows some interesting and typical properties. At the beginning the rate is small, because of the small amount of the catalyst. With increasing concentration of S_2 the rate increases and reaches a maximum

$$r_{\text{max}} = \frac{k}{4}(S_1(0) + S_2(0))^2 \quad (3.135)$$

at the same time point (3.133), where the concentrations have their inflection point. After this point the reaction gets slower because of the decreasing amount of the reactant S_1 .

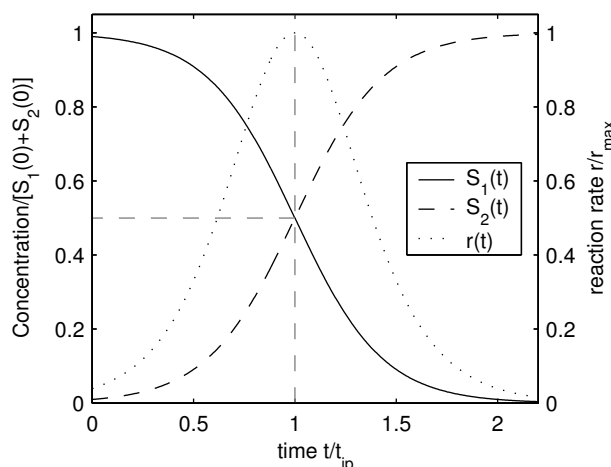


Fig. 3.12: The time evolution and the reaction as function of time for the autocatalytic reaction (3.125). We use normalized units for the concentrations and the time dependent rate. The time is plotted in units of the time of the inflection point (3.133). In contrast to uncatalyzed monomolecular reaction the reaction rate is an increasing function for $t < t_{\text{ip}}$.

In Figure 3.12 the time-dependent normalized concentrations $S_1(t)$, $S_2(t)$ and the normalized reaction rate are plotted. The time is in units of the inflection point. As mentioned before the concentrations have an inflection point where the reaction rate reaches its maximum.

The reaction starts with a small but finite initial concentration of the autocatalyst S_2 . Without this condition the ansatz for the reaction rate does not work. We avoid this non-physical behavior by invoking an extra uncatalyzed reaction converting S_1 directly to S_2 . How we mentioned before such a reaction must exist, but it can be very slow. The rate equation is now

$$r = -\frac{dS_1(t)}{dt} = \frac{dS_2(t)}{dt} = [k_0 + k S_2(t)] S_1(t) , \quad (3.136)$$

where k_0 is the rate coefficient of the uncatalyzed monomolecular reaction. After an initial time the second term within the brackets is dominant resulting in (3.127).

3.7 Further Material to Enzyme Kinetics

This section derives the well known Michaelis-Menten representation in detail and considers several distinct kinds of enzymatic reaction types, including competitive binding, inhibitors, reversible Michaelis-Menten equation. We discuss the consequences of these different formulations on the dynamics the models can capture.

3.7.1 Michaelis-Menten equation

The starting point of our discussion of enzyme kinetics is the well-known Michaelis-Menten equation. Within this framework we assume that the conversion of the substrate S into the product P is catalyzed by the enzyme E . The enzyme and the substrate form an intermediate enzyme-substrate complex C which can degrade into the reactants or into product and enzyme. The kinetic mechanism can be presented in the usual graphical form



where the k_i denote the rate coefficients for each elementary reaction. The kinetic rate equations for that are now a system of four coupled differential equations

$$\dot{S} = -k_1 E(t) S(t) + k_{-1} C(t), \quad (3.138)$$

$$\dot{E} = -k_1 E(t) S(t) + [k_{-1} + k_2] C(t), \quad (3.139)$$

$$\dot{C} = k_1 E(t) S(t) - [k_{-1} + k_2] C(t) = -\dot{E}, \quad (3.140)$$

$$\dot{P} = k_2 C(t). \quad (3.141)$$

There is a conservation law

$$E(0) = E(t) + C(t) \quad (3.142)$$

for the enzyme, where $E(0)$ is the initial concentration. A similar law can be found for the substrate

$$S(0) = S(t) + C(t) + P(t). \quad (3.143)$$

Considering these conservation laws the system of coupled differential equations (3.138)-(3.141) reduces from a fourth to a second-order system. Usually, the equations for the substrate (3.138) and for the enzyme-substrate complex C (3.140) are kept. In the further treatment, one obtains an expression for the rate of change of the substrate, namely is the Michaelis-Menten equation. This result is identified as the reaction rate. Formally, this is not correct, because of for a complex reaction scheme the reaction rate is the rate of change for the product only. As we show for consecutive reactions, the rate of change of the intermediate steps is different. Nevertheless, within the approach of Michaelis and Menten, the rates of substrate and product are equal. Therefore, the reaction rate can be derived from the substrate change.

To prove this we use only the conservation law (3.142) for the enzyme and obtain the system of differential equations (3.138), (3.140) and (3.141). This system is not solvable in terms of simple analytic functions [SM97], therefore we follow Michaelis and Menten [MM13] and Briggs and Haldane [BH25] and introduce a steady state assumption¹³ for the complex

$$\dot{C} \approx 0 = k_1 E(t) S(t) - [k_{-1} + k_2] C(t). \quad (3.144)$$

From this equation we obtain a constant concentration for the enzyme-substrate complex

$$C(t) = \frac{k_1}{k_{-1} + k_2} E(t) S(t) = \frac{E(t) S(t)}{K_M} \quad (3.145)$$

or for the enzyme

$$E(t) S(t) = K_M C(t). \quad (3.146)$$

¹³In their original work Michaelis and Menten consider only the reverse reaction into the complex and its decay into substrate and enzyme and neglect the conversion into the product. Briggs and Haldane generalized the steady state assumption into the given expression. The resulting relation for the reaction rate has in both cases the same form.

Insertion into equation (3.138) leads to

$$\dot{S} = -\dot{P} = -k_2 C(t) = -k_2 \frac{E(t) S(t)}{K_M}. \quad (3.147)$$

Within the steady state approximation the mechanism (3.137) formally reduces to a bimolecular reaction of the substrate and the enzyme. Furthermore, we introduce the Michaelis constant

$$K_M = \frac{k_{-1} + k_2}{k_1}. \quad (3.148)$$

Because the rate coefficients k_i depend on temperature, pH -value, etc. the Michaelis constant and the limiting rate are not constant quantities. Their values change with the environmental conditions of the considered system.

The reduction of the apparent reaction mechanism and the resulting mathematical simplification is typical for the steady state assumption. It makes this approximation a powerful tool in the analytical treatment of the complex biochemical reactions in metabolism and inter- and intracellular communications.

Combining the conservation law (3.142) and the steady state complex concentration (3.145), one gets

$$C(t) = \frac{E(0) S(t)}{K_M + S(t)} \quad (3.149)$$

for the enzyme. This result we insert into (3.147) and obtain

$$\begin{aligned} V = -\dot{S} = \dot{P} &= \frac{k_2 E(0) S(t)}{K_M + S(t)} \\ &= \frac{V_{\max} S(t)}{K_M + S(t)} \end{aligned} \quad (3.150)$$

the well-known Michaelis-Menten equation, where $V_{\max} = k_2 E(0)$ is the limiting rate for the enzyme kinetic reaction.

For $K_M \gg S(t)$ equation (3.150) provides a linear dependence of the reaction rate

$$-\dot{S} = \dot{P} \approx \frac{k_2 E(0)}{K_M} S(t) \quad (3.151)$$

from the substrate concentration. The rate-determining step for low substrate concentrations is the bimolecular formation of the enzyme-substrate complex. In this approximation the limiting rate is reached if the substrate concentration is equal to the Michaelis constant. For $S(t) \gg K_M$ the Michaelis-Menten equation (3.150) reduces to

$$-\dot{S} = \dot{P} = V_{\max} \quad (3.152)$$

a constant rate. The pre-equilibria of enzyme, substrate and enzyme-substrate complex rests completely with the side of the complex. All enzyme molecules are bound to a substrate molecule, and according to the conservation law (3.142), the complex concentration is equal to the initial enzyme concentration. The reaction rate is determined by the decay of the complex into product and enzyme. This process is described by the rate coefficient k_2 and independent from the substrate concentration. Hence, the reaction rate cannot be increased further by an increase of the substrate concentration. According to IUPAC¹⁴ it is called limiting rate and not maximum rate as in elder books, and the usual notation V_{\max} is retained for this quantity.

¹⁴IUPAC - International Union of Pure and Applied Chemistry

The term introduced limiting regimes which allow an independent measurement of the Michaelis constant and the limiting rate. For a known enzyme concentration we can calculate the rate coefficient k_2 . The coefficients k_1 and k_{-1} are not uniquely determined. Only if one assumes $k_{-1} \gg k_2$ the Michaelis constant reduces to the ratio

$$K_M \approx \frac{k_{-1}}{k_1}, \quad (3.153)$$

as did the constant originally introduced by Michaelis and Menten.

While the present section introduced the basic treatment of enzyme kinetic reactions we will use the following sections to consider more complex conversion schemes.

Dimensionless Representation – Activity Plot

The dimensionless representation of the Michaelis-Menten equation is a very useful tool for comparison of results for more complex enzyme kinetic reactions. It gives a very clear insight into the properties of the catalytic conversion of a substrate into the needed product. It allows the comparison of more complex enzyme kinetic reactions, we will introduce in the next chapters, with the Michaelis-Menten model. Furthermore its general form allows an easy discussion of general properties of the enzymatic conversion of a substrate into a product.

Therefore, we rearrange the Michaelis-Menten equation (3.150) into

$$A = \frac{V}{V_{\max}} = \frac{S(t)/K_M}{1 + S(t)/K_M}, \quad (3.154)$$

where the new defined ratio of the current reaction rate V and the limiting rate V_{\max} is called activity. The substrate concentration is now given in units of the Michaelis constant K_M . The important activity $A_{0.5}$, where the half limiting rate is reached, is related to this constant. Because of the chosen dimensionless concentration it is now obtained for a value $S/K_M = 1$. The activity has its limiting value of $A = 1$, if the reaction rate approaches to its limiting rate V_{\max} .

3.7.2 The enzyme kinetic reaction with second intermediate

Another possible reaction scheme for the catalytic conversion of a substrate is a mechanism with an additional intermediate complex C_2 . The stoichiometric formula for such a reaction is



where the transition between the intermediates is fully reversible. The corresponding system of coupled differential equations is

$$\dot{S} = -k_1 E(t) S(t) + k_{-1} C_1(t), \quad (3.156)$$

$$\dot{E} = -k_1 E(t) S(t) + k_{-1} C_1(t) + k_3 C_2(t), \quad (3.157)$$

$$\dot{C}_1 = k_1 E(t) S(t) - [k_{-1} + k_2] C_1(t) + k_{-2} C_2(t), \quad (3.158)$$

$$\dot{C}_2 = k_2 C_1(t) - [k_{-2} + k_3] C_2(t), \quad (3.159)$$

$$\dot{P} = k_3 C_2(t). \quad (3.160)$$

For this reaction scheme, all equations are linear independent, in contrast to the system of coupled differential equations (3.138)-(3.141) for the Michaelis-Menten model. For the enzyme we find now the conservation law

$$E(0) = E(t) + C_1(t) + C_2(t). \quad (3.161)$$

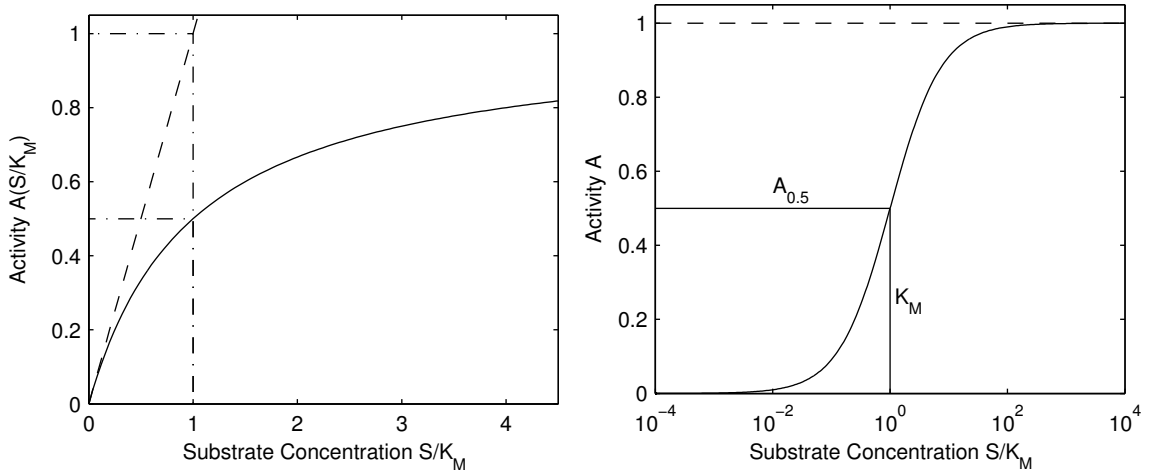


Fig. 3.13: Activity (solid line) as function of the substrate concentration in a linear (left) and a logarithmic (right) representation. In the linear representation, the small-concentration-approximation is drawn as a dashed line. Additionally, the activity value for the concentration $S/K_M = 1$ is indicated. In the logarithmic representation, the high-concentration-limit is drawn with a dashed line. The thin solid lines corresponds to the Michaelis constant and the half activity, respectively. The sigmoidal shape of the activity curve is a result of the logarithmic representation. It shows the change of the rate-determining step with increasing substrate concentration. For low concentrations the bimolecular formation of the enzyme-substrate complex, and for high concentrations the decomposition of the complex into product P and free enzyme E determines the reaction rate.

Again we assume the enzyme-substrate complexes C_1 and C_2 at steady states. Considering the conservation law (3.161) we obtain for the first complex C_1

$$C_1(t) = \frac{k_1 E(0) S(t) - [k_1 S - k_2] C(t)}{k_{-1} + k_2 + k_1 S(t)}. \quad (3.162)$$

This expression we insert into the equation for the second complex C_2 resulting in the steady state concentration

$$C_2(t) = \frac{k_1 k_2 E(0) S(t)}{k_{-1} k_{-2} + k_{-1} k_3 + k_2 k_3 + [k_1 k_2 + k_1 k_{-2} + k_1 k_3] S(t)}. \quad (3.163)$$

which we insert into the differential equation for the product (3.160). Finally, we obtain for the reaction rate of this type of enzyme kinetic reaction the expression

$$V(t) = \dot{P} = \frac{k_1 k_2 k_3 E(0) S(t)}{k_{-1} k_{-2} + k_{-1} k_3 + k_2 k_3 + [k_1 k_2 + k_1 k_{-2} + k_1 k_3] S(t)}. \quad (3.164)$$

Introducing an apparent Michaelis constant

$$K^{\text{app}} = \frac{k_{-1} k_{-2} + k_{-1} k_3 + k_2 k_3}{k_1 k_2 + k_1 k_{-2} + k_1 k_3} \quad (3.165)$$

and an apparent limiting rate

$$V^{\text{app}} = \frac{k_2 k_3}{k_2 + k_{-2} + k_3} E(0) \quad (3.166)$$

the expression (3.164) simplifies to

$$V(t) = \frac{V^{\text{app}} S(t)}{K^{\text{app}} + S(t)}, \quad (3.167)$$

a Michaelis-Menten-like form. Whereas the Michaelis constant (3.148) is interpretable as a dissociation constant, the apparent Michaelis constant prohibits such a physical interpretation.

Analogous to the simpler Michaelis-Menten scheme the mechanism (3.155) formally reduces to a bimolecular reverse reaction. To prove this, we transform Eq. (3.159) at steady state into an expression for

$$k_2 C_1(t) = [k_{-2} + k_3] C_2(t) \quad (3.168)$$

and insert it into (3.158). The result

$$k_{-1} C_1 = k_1 S(t)E(t) - k_3 C_2(t) \quad (3.169)$$

we use to simplify the rate of change of the substrate

$$\dot{S} = -\dot{P} = -k_3 C_2. \quad (3.170)$$

The reaction rate is

$$\begin{aligned} V(t) &= -\dot{S} = \dot{P} = k_3 C_2(t) \\ &= \frac{k_1 k_2 k_3 E(0) S(t)}{k_1 k_2 + k_2 k_{-2} + k_2 k_3 + (k_2 k_{-2} + k_1 k_3) S(t)}. \end{aligned} \quad (3.171)$$

3.7.3 Constant substrate concentration

If we assume a constant substrate concentration, for instance because of external flows or an external reservoir, we can further simplify the Michaelis-Menten equation (3.150). In this special case the right hand side is a constant. The reaction rate

$$V = -\dot{S} = \dot{P} = k_{\text{eff}} \quad (3.172)$$

follows zero-order kinetics. The introduced effective rate coefficient is defined as

$$k_{\text{eff}} = \frac{k_2 E(0) S}{K_M + S}. \quad (3.173)$$

For high substrate concentrations $S \gg K_M$ the rate (3.172) reduces to

$$V \approx k_2 E(0),$$

the decay of the enzyme-substrate complex into product and enzyme is the rate-determining step. On the other hand, for small substrate concentrations $S \ll K_M$ the rate

$$V \approx k_1 E(0) S(t)$$

is limited by the bimolecular formation of the enzyme-substrate complex C_1 .

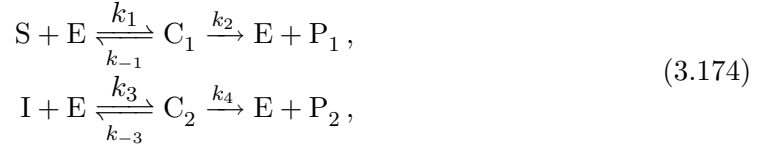
3.7.4 Interactions with other reaction mechanisms and reaction partners

In the previous part we discussed the conversion of a substrate into a product by a catalyzing enzyme for a single reaction channel and without the presence of further substances. The next sections focus on the interaction of different reaction channels and the influences of other substances to a specified conversion. We show some simple examples of inhibitory mechanisms to the enzyme kinetic reaction (3.137). The term inhibition is misleading here, because after the definition in chemistry a catalyst¹⁵ is an additional substance changing the reaction rate in a positive or negative way. But for some types of inhibition we have no additional substrates or enzymes.

¹⁵Every additional substance influencing the reaction rate is a catalyst. If it increases the reaction rate (positive catalysis) we call it catalyst. An inhibitor decreases the rate (negative catalysis).

Competitive substrates

A standard example for an inhibitory mechanism is the case of two competitive substrates. The two different substrates are catalyzed by the same enzyme E . In the stoichiometric representation



we assume that both processes follow the model of Michaelis and Menten. The first reaction in (3.174) is our desired conversion reaction from S to the product P_1 . It is suppressed by the reaction of an inhibitor I with the enzyme, where we allow the production of an unused product P_2 . This system of chemical reactions is described by a system of seven coupled differential equations. The system

$$\dot{S} = -k_1 S(t) E(t) + k_{-1} C_1(t) \quad (3.175)$$

$$\dot{I} = -k_3 I(t) E(t) + k_{-3} C_2(t) \quad (3.176)$$

$$\begin{aligned} \dot{E} &= -[k_1 S(t) + k_3 I(t)] E(t) \\ &\quad + (k_{-1} + k_2) C_1(t) + (k_{-3} + k_4) C_2(t) \end{aligned} \quad (3.177)$$

$$\dot{C}_1 = k_1 S(t) E(t) - (k_{-1} + k_2) C_1(t) \quad (3.178)$$

$$\dot{C}_2 = k_3 I(t) E(t) - (k_{-3} + k_4) C_2(t) \quad (3.179)$$

$$\dot{P}_1 = k_2 C_1(t) \quad (3.180)$$

$$\dot{P}_2 = k_4 C_2(t) \quad (3.181)$$

can be reduced in the same manner as for the single enzyme kinetic reaction. The conservation law for the enzyme

$$E(0) = E(t) + C_1(t) + C_2(t) \quad (3.182)$$

and the steady state assumption for both complexes cancel three differential equations. Furthermore we can show again, that the rate of degradation of substrate is equal to the production rate of product P_1 in this approximation. The same can be shown for the inhibitor reaction. Only the differential equations for the substrate and the inhibitor remain. It follows from (3.178) that

$$C_1(t) = \frac{E(0) S(t)}{K_{MS} [1 + I(t)/K_{MI}] + S(t)} \quad (3.183)$$

for the enzyme-substrate complex and

$$C_2(t) = \frac{E(0) I(t)}{K_{MI} [1 + S(t)/K_{MS}] + I(t)} \quad (3.184)$$

for the inhibitor-enzyme complex. The number of bound substrate molecules is reduced in comparison to the single enzyme kinetic reaction, see equation (3.149). Because this concentration determines the conversion rate into the product, a reduction of this concentration leads to a decrease of the reaction rate. The enzyme kinetic reaction (3.137) is suppressed by the presence of a second molecule reacting with the same enzyme.

For the reaction rate we obtain with (3.183)

$$V^S(t) = -\frac{V_{\max}^S S(t)}{K_{MS} [1 + I(t)/K_{MI}] S(t)} \quad (3.185)$$

and for the inhibitor

$$V^I(t) = -\frac{V_{\max}^I I(t)}{K_{\text{MI}} [1 + S(t)/K_{\text{MS}}] I(t)}, \quad (3.186)$$

where $V_{\max}^S = k_2 E(0)$ is the limiting rate for the substrate conversion and V_{\max}^I the corresponding quantity for the inhibitory reaction. In the limit of $S \gg K_{\text{MS}}$ the reaction rate reaches its limiting value and the inhibitor is negligible. In the contrary case, the reaction rate is again proportional to the substrate concentration but it is reduced by a factor $1 + I/K_{\text{MI}}$ in comparison to the result from our previous discussion. Finally we investigate the rate in dependence from substrate concentration and a given inhibitor concentration. Then we can treat the inhibitor amount as constant and combine

$$K^{\text{app}} = K_{\text{MS}} [1 + I/K_{\text{MI}}]$$

into the apparent Michaelis constant K^{app} , which is greater than the Michaelis constant K_{MS} of the single enzyme kinetic reaction. The result is a Michaelis-Menten like differential equation

$$V^S(t) = -\frac{V_{\max}^S S(t)}{K^{\text{app}} + S(t)}. \quad (3.187)$$

In Figure 3.14 the Michaelis-Menten equation (3.150) is compared with result in presence of an inhibitor I . Whereas for small substrate concentrations the inhibitor strongly affects the reactions rate, for high substrate concentrations the inhibitor is negligible, as mentioned before, and the reaction rate approaches to its limiting rate determined by the decay rate of the enzyme-substrate complex.

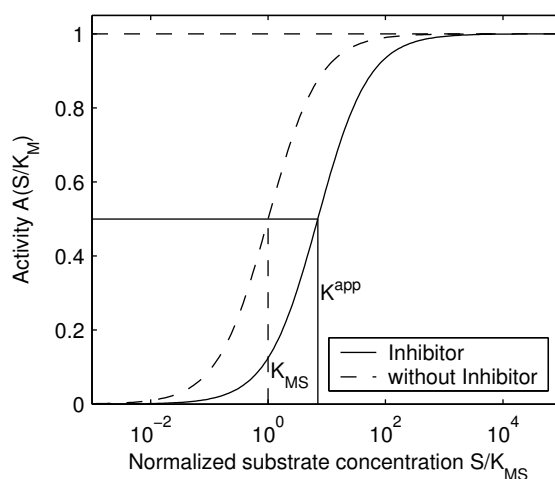
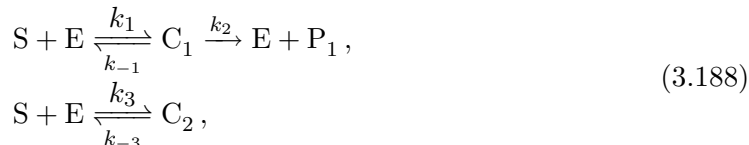


Fig. 3.14: The reaction rate of the enzyme kinetic reaction in presence of an inhibitor I reacting with the same enzyme to a second unused product as function of the substrate concentration in logarithmic representation. The corresponding concentration of the Michaelis constant K_{MS} and the apparent Michaelis constant K^{app} is shown by thin solid lines.

Non-productive binding

Another mechanism of suppression is the existence of other reaction channels. In a first example we want to discuss an enzyme kinetic reaction where the enzyme and the substrate can bind to a second enzyme-substrate complex that cannot convert into product and free enzyme. Because this second reaction channel does not form any product this scheme is

called non-productive binding¹⁶. The corresponding system of stoichiometric equations is given as



where the second reaction is a reversible reaction between substrate and enzyme and the enzyme-substrate complex. The rates of change are now given by the system of differential equations

$$\dot{S} = -(k_1 + k_3) E(t) S(t) + k_{-1} C_1(t) + k_{-3} C_2(t), \quad (3.189)$$

$$\dot{E} = -[k_1 + k_3] E(t) S(t) + [k_{-1} + k_2] C_1(t) + k_{-3} C_2(t), \quad (3.190)$$

$$\dot{C}_1 = k_1 E(t) S(t) - [k_{-1} + k_2] C_1(t), \quad (3.191)$$

$$\dot{C}_2 = k_3 E(t) S(t) - k_{-3} C_2(t), \quad (3.192)$$

$$\dot{P} = k_2 C_1(t) \quad (3.193)$$

with the conservation law

$$E(0) = E(t) + C_1(t) + C_2(t). \quad (3.194)$$

In a steady state approximation for the complexes we obtain

$$C_1(t) = \frac{E(t) S(t)}{K_{M1}} \quad (3.195)$$

$$C_2(t) = \frac{E(t) S(t)}{K_D} \quad (3.196)$$

for the complex concentration, where

$$K_{M1} = \frac{k_{-1} + k_2}{k_1} \quad (3.197)$$

is the usual Michaelis constant and

$$K_D = \frac{k_{-3}}{k_3} \quad (3.198)$$

the dissociation constant of the non-productive complex. An analysis of the remaining differential equations of substrate and product concentration shows again, that they are the same except for the sign. Taking into account the conservation law (3.194) we obtain for the enzyme concentration

$$E(t) = \frac{E(0)}{1 + S(t)/K_{M1} + S(t)/K_D}, \quad (3.199)$$

which leads to a reaction rate

$$\begin{aligned} V(t) &= \dot{P} = -\dot{S} \\ &= \frac{V_{\max} S(t)}{K_{M1} [1 + S(t)/K_{M1} + S(t)/K_D]}, \end{aligned} \quad (3.200)$$

where $V_{\max} = k_2 E(0)$ is the limiting rate known from previous sections. From (3.200) one can see, that the reaction rate is decreased by the second non-productive parallel reaction. The obtained expression for the rate we can transform into a Michaelis-Menten

¹⁶The enzyme-substrate complex is an unstable intermediate (transition state), it cannot exist alone.

like relation if we introduce effective constants. For the non-productive binding scheme (3.188) we have the apparent limiting rate

$$V_{\max}^{\text{app}} = \frac{V_{\max}}{1 + K_{M1}/K_D} \quad (3.201)$$

and an apparent Michaelis constant

$$K^{\text{app}} = \frac{K_{M1}}{1 + K_{M1}/K_D}. \quad (3.202)$$

Both coefficients are smaller than their pendants in the Michaelis-Menten equation (3.150). Again, we get a Michaelis-Menten like equation

$$V = \frac{V_{\max}^{\text{app}} S(t)}{K_M^{\text{app}} + S(t)}, \quad (3.203)$$

for the reaction rate. In comparison to a conversion according to (3.137) with the same rate coefficients the non-productive reaction channel decelerates the conversion into P , because a fraction of the enzyme is bound in non-productive complex C_2 . This is shown in Figure 3.15, where we compare the result of the Michaelis-Menten equation (3.150) and the Michaelis-Menten like equation (3.203). In both cases we use the same rate coefficients for the successful conversion reaction of the substrate.

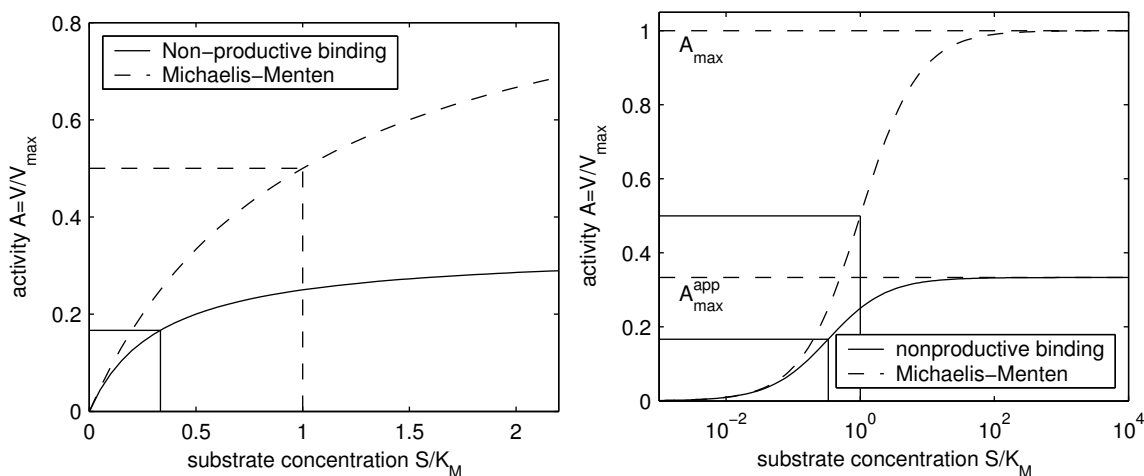


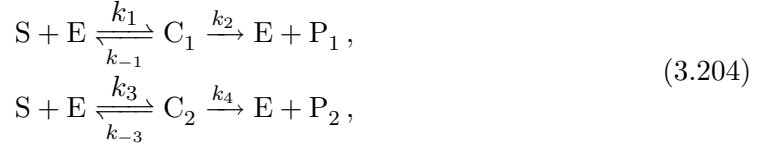
Fig. 3.15: Comparison of non-productive reaction scheme (3.188) with the Michaelis-Menten mechanism (3.137) in linear (left) and logarithmic (right) representation. In the presence of a second complex which cannot lead to a product the reaction rate is reduced. Note, that the apparent Michaelis constant and the apparent limiting rate are always smaller than the corresponding parameter of the original Michaelis-Menten equation (3.150).

An enzymatic reaction with a non-productive channel cannot be distinguished from a reaction without such a channel by the measurement of substrate or product concentration, because the apparent constants depends only on rate coefficient. For this purpose one has to identify the complexes.

Non-productive binding with an additional product

In this section we expand our previous treatment and introduce the second reaction as a parallel reaction, where the produced product P_2 is an unused by-product. From this we can expect that the second reaction channel decelerates down the reaction rate and

decrease production of the desired product P_1 . The stoichiometric representation of such a biochemical system is



from which follows the system of six coupled differential equations

$$\dot{S} = -(k_1 + k_3) E(t) S(t) + k_{-1} C_1(t) + k_{-3} C_2(t), \quad (3.205)$$

$$\dot{E} = -[k_1 + k_3] E(t) S(t) + [k_{-1} + k_2] C_1(t) + [k_{-3} + k_4] C_2(t), \quad (3.206)$$

$$\dot{C}_1 = k_1 E(t) S(t) - [k_{-1} + k_2] C_1(t), \quad (3.207)$$

$$\dot{C}_2 = k_3 E(t) S(t) - [k_{-3} + k_4] C_2(t), \quad (3.208)$$

$$\dot{P}_1 = k_2 C_1(t), \quad (3.209)$$

$$\dot{P}_2 = k_4 C_2(t). \quad (3.210)$$

Again, we have the conservation law

$$E(0) = E(t) + C_1(t) + C_2(t) \quad (3.211)$$

which cancels the differential equation for the enzyme. A further simplification arises from the steady state assumption for both complexes. Their steady state concentrations are

$$C_1(t) = \frac{k_1}{k_{-1} + k_2} E(t) S(t) = \frac{E(t) S(t)}{K_{M1}}, \quad (3.212)$$

$$C_2(t) = \frac{k_3}{k_{-3} + k_4} E(t) S(t) = \frac{E(t) S(t)}{K_{M2}}, \quad (3.213)$$

where we introduced the Michaelis constants K_{M1} and K_{M2} . With the conservation law (3.211) and $V_{\max1} = k_2 E(0)$ the reaction rate for P_1 is

$$V^{P_1}(t) = \dot{P}_1 = \frac{V_{\max1} S(t)}{K_{M1} + S(t) (1 + K_{M1}/K_{M2})}. \quad (3.214)$$

With the apparent Michaelis constant

$$K_1^{\text{app}} = \frac{K_{M1}}{1 + K_{M1}/K_{M2}} \quad (3.215)$$

and the apparent limiting rate is

$$V_{\max1}^{\text{app}} = \frac{V_{\max1}}{1 + K_{M1}/K_{M2}} \quad (3.216)$$

we transform (3.214) into the Michaelis-Menten like form

$$V^{P_1}(t) = \frac{V_{\max1}^{\text{app}} S(t)}{K_1^{\text{app}} + S(t)}. \quad (3.217)$$

In the same manner we obtain for by-product P_2 a similar relation, where the apparent Michaelis constant is given as

$$K_2^{\text{app}} = \frac{K_{M2}}{1 + K_{M2}/K_{M1}} \quad (3.218)$$

and the apparent limiting rate

$$V_{\max2}^{\text{app}} = \frac{V_{\max2}}{1 + K_{M2}/K_{M1}}, \quad (3.219)$$

where the limiting rate is $V_{\max 2} = k_4 E(0)$. Last, but not least we have to calculate the rate of substrate depletion. From (3.212) follows

$$k_1 E(t) S(t) = (k_{-1} + k_2) C_1(t)$$

and from (3.213)

$$k_3 E(t) S(t) = (k_{-3} + k_4) C_1(t).$$

If we insert both expressions into the rate of change of substrate S (3.205) the contributions of reverse reaction of complexes C_1 and C_2 are compensated and we obtain the rate

$$\dot{S} = -k_2 C_1(t) - k_4 C_2(t) \quad (3.220)$$

depending only on the decay into the products and free enzyme. On the other hand, the rate of production of P_1 and P_2 is determined by these two terms and we can rewrite the last equation into

$$\dot{S} = - \left[\dot{P}_1 + \dot{P}_2 \right]. \quad (3.221)$$

The rate of change of the substrate is the sum of the rates of change of the products. In contrast to our previous reaction schemes, the rates are not equal, except for the sign, any more. This is typical for a reaction with more than one product. The consumption of the substrate is separated into two independent channels. Therefore, the concentrations of substrate and product¹⁷ as a function of time should be measured in an experiment. A comparison of the sum of the reaction rates of the products and the rate of change of the substrate gives evidence if there are unknown reaction channels. Furthermore, the differences arising from unknown channels can be used to estimate the portion of these reactions on the consumption of the substrate. The insertion of the reaction rates for the products (3.217) and its analogon for P_2 leads to

$$\dot{S} = - \left(\frac{V_{\max 1}^{\text{app}}}{K_1^{\text{app}} + S(t)} + \frac{V_{\max 2}^{\text{app}}}{K_2^{\text{app}} + S(t)} \right) S(t) \quad (3.222)$$

a rate that is not presentable in a Michaelis-Menten-like form. Only in the special case $K_{M1}^{\text{app}} = K_{M2}^{\text{app}}$ the rate of substrate S can be transformed into such a form.

It is interesting to analyze some asymptotic properties of (3.222). First we consider the limit of low substrate concentrations $S \ll K_{M1}^{\text{app}}, K_{M2}^{\text{app}}$. Then the rate of degradation reduces to

$$\dot{S} = - \left(\frac{V_{\max 1}}{K_{M1}} + \frac{V_{\max 2}}{K_{M2}} \right) S(t), \quad (3.223)$$

which is a linear relation similar to the result for the Michaelis-Menten equation (3.151). In the other limit $S \gg K_{M1}^{\text{app}}, K_{M2}^{\text{app}}$ one obtains with

$$\dot{S} = - (V_{\max 1}^{\text{app}} + V_{\max 2}^{\text{app}}) = V_{\max}^{\text{T}} \quad (3.224)$$

a limiting rate, too. The formal differences appear only in an intermediate range of concentrations. In figure 3.16 we use these asymptotes to compute a Michaelis-Menten result for the substrate and compare it with (3.222).

Competitive Enzymes

Besides the inhibition by another substrate competing for the same enzyme, it is also possible that a second enzyme E_1 reacts with the substrate S . The reaction scheme then

¹⁷The known or desired product has to be measured.

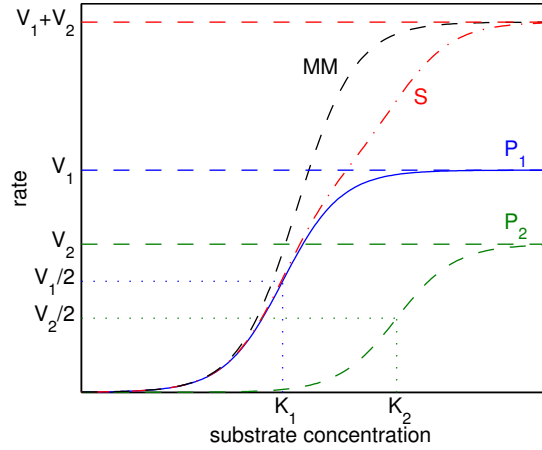
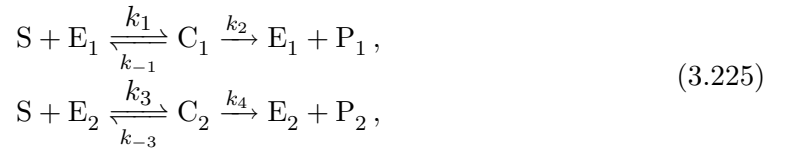


Fig. 3.16: The reaction rate for substrate, product P_1 , and unused product P_2 as function of the substrate concentration in comparison to a Michaelis-Menten curve with the same asymptotes. Whereas the products obey a Michaelis-Menten-like equation the substrate is not transformable in such a form. The apparent Michaelis constants of the products and the corresponding rate are mentioned by dotted lines in the same color.

is



where we assume, that the second reaction leads to a product P_2 used for instance in another pathway. The corresponding system of coupled differential equations consists of seven equations. The system

$$\dot{S} = -[k_1 E_1(t) + k_3 E_2(t)] S(t) + k_{-1} C_1(t) + k_{-3} C_2(t), \quad (3.226)$$

$$\dot{E}_1 = -k_1 E_1(t) S(t) + [k_{-1} + k_2] C_1(t), \quad (3.227)$$

$$\dot{E}_2 = -k_3 E_2(t) S(t) + [k_{-3} + k_4] C_2(t), \quad (3.228)$$

$$\dot{C}_1 = k_1 E_1(t) S(t) - [k_{-1} + k_2] C_1(t), \quad (3.229)$$

$$\dot{C}_2 = k_3 E_2(t) S(t) - [k_{-3} + k_4] C_2(t), \quad (3.230)$$

$$\dot{P}_1 = k_2 C_1(t), \quad (3.231)$$

$$\dot{P}_2 = k_4 C_2(t) \quad (3.232)$$

can be simplified if we use the conservation laws for the enzymes

$$E_1(0) = E_1(t) + C_1(t) \quad (3.233)$$

$$E_2(0) = E_2(t) + C_2(t) \quad (3.234)$$

to cancel the differential equations for both enzymes (3.227) and (3.228). If the complexes reside in a steady state the system reduces further. Analogous to the treatment of the Michaelis-Menten model (3.137) we obtain for the products

$$\dot{P}_1 = \frac{V_{\max 1} S(t)}{K_{M1} + S(t)} \quad (3.235)$$

and

$$\dot{P}_2 = \frac{V_{\max 2} S(t)}{K_{M2} + S(t)}, \quad (3.236)$$

where we define the limiting rate $V_{\max 1}$, $V_{\max 2}$, and the Michaelis constants K_{M1} , K_{M2} in the usual way. The reaction rates of the products remain unaffected by the presence of the other reaction. In steady state we have

$$\begin{aligned} k_1 E_1 S &= (k_{-1} + k_2) C_1(t) \\ k_3 E_2 S &= (k_{-3} + k_4) C_2(t) \end{aligned}$$

from which follows for the substrate

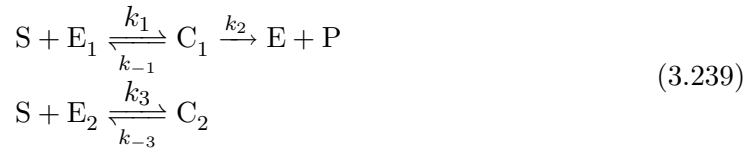
$$\begin{aligned} \dot{S} &= -[k_2 C_1(t) + k_4 C_2(t)] \\ &= -(\dot{P}_1 + \dot{P}_2). \end{aligned} \quad (3.237)$$

With (3.235) and (3.236) we obtain for the substrate

$$\dot{S} = -\left(\frac{V_{\max 1}}{K_{M1} + S(t)} + \frac{V_{\max 2}}{K_{M2} + S(t)}\right) S(t) \quad (3.238)$$

an equation which is no more Michaelis-Menten like. Only in the case of $K_{M1} = K_{M2}$ the equation (3.238) reduces to such a relation, but because we assume that E_1 and E_2 are different enzymes this is a hard restriction in the treatment of enzyme kinetic reactions.

Another possible reaction scheme is a competitive enzyme that cannot lead to a product P_2 and forms only a second complex C_2 . The stoichiometric equations than simplify to



and the corresponding system of differential equations is

$$\dot{S} = -k_1 S(t) E_1(t) - k_3 S(t) E_2(t) + k_{-1} C_1(t) + k_{-3} C_2(t), \quad (3.240)$$

$$\dot{E}_1 = -k_1 S(t) E_1(t) + (k_{-1} + k_2) C_1(t), \quad (3.241)$$

$$\dot{E}_2 = -k_3 S(t) E_2(t) + k_{-3} C_2(t), \quad (3.242)$$

$$\dot{C}_1 = k_1 S(t) E_1(t) - (k_{-1} + k_2) C_1(t), \quad (3.243)$$

$$\dot{C}_2 = k_3 S(t) E_2(t) - k_{-3} C_2(t), \quad (3.244)$$

$$\dot{P} = k_2 C_1(t). \quad (3.245)$$

Taking into account the conservation laws

$$E_1(0) = E_1(t) + C_1(t),$$

$$E_2(0) = E_2(t) + C_2(t),$$

and the steady state assumption for complexes we obtain again the result of a single enzyme kinetic reaction (3.150).

This result reflects some general properties in the discussed treatment of an enzyme kinetic reaction. The reaction rate of products obeys in this framework, conservation laws for enzymes and the steady state assumption for complexes, a Michaelis-Menten-like relation. The different mechanisms lead to new apparent Michaelis constants and limiting rates. Only if more than one product is converted, the absolute value of the rate of change of substrate is not equal to the reaction rate. The rate of change does not follow a Michaelis-Menten law in the case of more products.

3.7.5 Combination of enzyme kinetic reaction and an uncatalyzed conversion reaction

In our previous discussion of the conversion of a substrate into a product we considered only a catalyzed reaction. Because of thermodynamic reasons there is an uncatalyzed reaction too. This is a general property of catalysis. A catalyst cannot make possible a biochemical reaction, if the final state does not exist. But if such a state exists, there is a certain transition probability which might be very small but is not zero, between the initial state (substrate) and the final state (product). Furthermore, a catalyst only affects the rate of change not the equilibrium. The equilibrium is determined by thermodynamic quantities. For our investigation we choose an monomolecular reaction



as the uncatalyzed branch. In analogy to the treatment of Michaelis and Menten we neglect a possible reverse reaction into the substrate. With these assumptions we obtain the set of coupled differential equations

$$\dot{S} = -k_1 E(t) S(t) + k_{-1} C(t) - k_u S(t), \quad (3.247)$$

$$\dot{E} = -k_1 E(t) S(t) + [k_{-1} + k_2] C(t), \quad (3.248)$$

$$\dot{C} = k_1 E(t) S(t) - [k_{-1} + k_2] C(t) = -\dot{E}, \quad (3.249)$$

$$\dot{P} = k_2 C(t) + k_u S(t), \quad (3.250)$$

where the rate of change of the substrate and product have an additional term and the equations for the enzyme remains unchanged in comparison to the model of Michaelis and Menten. Taking into account the steady-state assumption for the complex C and the conservation law for the enzyme we are able to reduce the system of differential equations. We obtain that the degradation rate of the substrate and the conversion rate of the product are equal,

$$\dot{P} = -\dot{S} \quad (3.251)$$

$$= \left(\frac{V_{\max}}{K_M + S(t)} + k_u \right) S(t), \quad (3.252)$$

but now with an additional term within the reaction rate. There are two limiting cases. If

$$k_u \ll \frac{V_{\max}}{K_M + S(t)} \quad (3.253)$$

we can neglect the uncatalyzed reaction and the Michaelis-Menten equation is reproduced. The other asymptote we obtain in the limit

$$k_u \gg \frac{V_{\max}}{K_M + S(t)}. \quad (3.254)$$

In contrast to the treatment of Michaelis and Menten the reaction rate linearly increases with the substrate concentration. Because of the saturation of the enzyme kinetic reaction the uncatalyzed reaction is now the dominant reaction. The transition between both behaviors is characterized by the relation

$$k_u = \frac{V_{\max}}{K_M + S(t)}, \quad (3.255)$$

where both branches have a rate of the same magnitude. For biochemical reactions we can assume $S \gg K_M$ leading to the relation

$$k_u S = V_{\max}. \quad (3.256)$$

If the direct conversion rate is equal to the limiting rate V_{\max} the direct reaction becomes more important.

In Figure 3.17 the reaction rate is shown in an activity plot as a function of the substrate concentration. For demonstration purposes we vary the rate coefficient of the uncatalyzed conversion by a multiple of the rate coefficient k_2 of the Michaelis-Menten model.

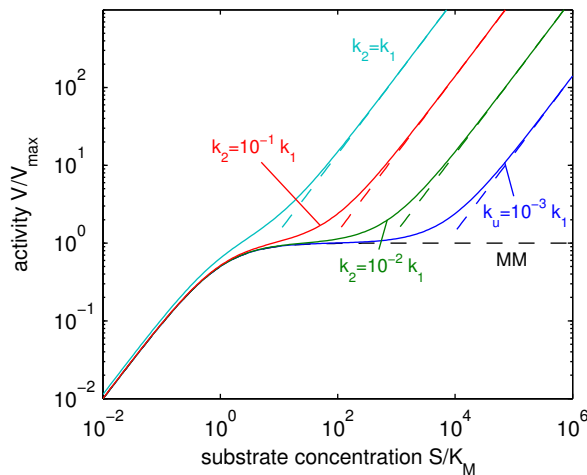
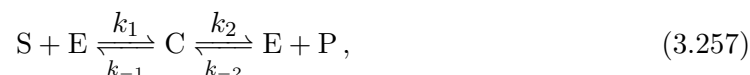


Fig. 3.17: Activity for a combination of an enzyme kinetic reaction with a parallel uncatalyzed reaction with the same product as a function of the substrate concentration. The rate coefficient of the direct reaction is a multiple of the rate coefficient k_2 of the Michaelis-Menten model. The usual enzyme kinetic reaction is the dominant part in the conversion for small concentrations fulfilling relation (3.253) and the direct conversion for high concentrations (relation (3.254)). The dashed lines are the corresponding asymptotes.

3.7.6 Reversible Michaelis-Menten equation

In the previous sections we neglected a possible reverse branch catalyzing the product into the substrate. Such an approximation assumes that the reverse reaction is very slow. Either the corresponding rate coefficient k_{-2} is nearly zero or the amount of the product is small in comparison to substrate. In the first case, the activation energy for the reverse way is very high. We can consider a small product concentration at the initial phase of the enzyme kinetic reaction and that the product is immediately transferred away. The simplest reaction mechanism of a reversible enzyme-kinetic reaction is



where we add an additional reverse reaction to mechanism of Michaelis and Menten (3.137). The corresponding system of coupled differential equations is

$$\dot{S} = -k_1 E(t) S(t) + k_{-1} C(t), \quad (3.258)$$

$$\dot{E} = -[k_1 S(t) + k_{-2} P(t)] E(t) + [k_{-1} + k_2] C(t), \quad (3.259)$$

$$\dot{C} = [k_1 S(t) + k_{-2} P(t)] E(t) - [k_{-1} + k_2] C(t) = -\dot{E}, \quad (3.260)$$

$$\dot{P} = k_2 C(t) - k_{-2} P(t) E(t). \quad (3.261)$$

The rate of change of enzyme, enzyme-substrate complex and product depends now on the product concentration, too. Again, there is a conservation for the enzyme

$$E(0) = E(t) + C(t)$$

relating the transient enzyme and complex concentration to a total enzyme concentration at $t = 0$. It enables us to reduce the system of differential equations. Considering the enzyme-substrate complex in steady state we obtain the balance equation

$$0 = (k_1 S(t) + k_{-2} P(t)) (E(0) - C(t)) - (k_{-1} + k_2) C(t), \quad (3.262)$$

where we apply the above conservation law to replace the transient enzyme concentration. After some straight forward transformation we obtain

$$C(t) = \frac{E(0) (S/K_M^F + P/K_M^R)}{1 + S/K_M^F + P/K_M^R} \quad (3.263)$$

for the steady state concentration of the complex. In this expression we introduce the abbreviations

$$K_M^F = \frac{k_{-1} + k_2}{k_1} \quad (3.264)$$

and

$$K_M^R = \frac{k_{-1} + k_2}{k_{-2}}. \quad (3.265)$$

With this result we get for the reaction rate

$$\begin{aligned} \dot{P} &= -\dot{S} \\ &= \frac{k_2 E(0) S(t)/K_M^F - k_{-1} E(0) P(t)/K_M^R}{1 + S/K_M^F + P/K_M^R} \\ &= \frac{V_{\max}^F S(t)/K_M^F - V_{\max}^R P(t)/K_M^R}{1 + S/K_M^F + P/K_M^R}, \end{aligned} \quad (3.266)$$

with the limiting rates $V_{\max}^F = k_2 E(0)$ and $V_{\max}^R = k_{-1} E(0)$. The reversible Michaelis-Menten equation can be interpreted as a combination of two irreversible enzyme kinetic reactions. The forward process converts the substrate to the product and the reverse reaction degrades the product. The additional product-dependent terms increase the value of the denominator and decrease the nominator. The sum of both changes decelerates the reaction rate in comparison to the irreversible version.

Because we consider a reversible reaction scheme, the substrate is not completely converted into the product. A chemical equilibrium is established after a certain time. It is defined by the balance equation

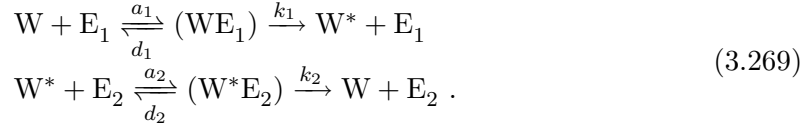
$$0 = \frac{V_{\max}^F S_{\text{eq}}}{K_M^F} - \frac{V_{\max}^R P_{\text{eq}}}{K_M^R} \quad (3.267)$$

from which follows the equilibrium constant

$$K_{\text{eq}} = \frac{P_{\text{eq}}}{S_{\text{eq}}} = \frac{V_{\max}^F K_M^R}{V_{\max}^R K_M^F} = \frac{k_2}{k_{-1}} \frac{k_1}{k_{-2}}. \quad (3.268)$$

3.8 Activation and Deactivation of Proteins by covalent Modifications

Signalling proteins often exist in an inactive form W and an active form W^* . The interconversion of the forms is catalyzed by two converter enzymes E_1 and E_2 . If we assume that the other substrates and products for modification and demodification are present at constant levels we obtain the reaction scheme [GKJ81]



In the signalling the modification system often consists of a phosphorylation and a dephosphorylation. In this case enzyme E_1 is the kinase and enzyme E_2 the phosphatase. The kinase modifies protein W by attaching a phosphate group to an amino acid that have a free hydroxyl group. The phosphate group is removed from ATP which is converted to ADP during the process of phosphorylation. The phosphatase changes the state of the protein due to hydrolysis taking away the phosphate group from the hydroxyl group. Hence, the rate constants in reaction scheme (3.269) depends on the concentrations of ATP and water.

Additionally to the usual conservation laws for enzymes

$$\begin{aligned} E_1^T &= E_1(t) + (WE_1)(t) \\ E_2^T &= E_2(t) + (W^*E_2)(t) , \end{aligned}$$

we furthermore assume a conservation of the protein itself

$$W^T = W(t) + W^*(t) + (WE_1)(t) + (W^*E_2)(t) . \quad (3.270)$$

Due these relations we are able to reduce the system of coupled ordinary differential equations from primary six equations to three equations. As usual, we cancel the equations for the enzymes and, additionally, the equation for the inactive protein W . The remaining differential equations are

$$\frac{dW^*(t)}{dt} = -a_2 W^*(t) E_2(t) + d_2 (W^*E_2)(t) + k_1 (WE_1)(t) \quad (3.271)$$

$$\frac{d(WE_1)(t)}{dt} = a_1 W(t) E_1(t) - [d_1 + k_1] (WE_1)(t) \quad (3.272)$$

$$\frac{d(W^*E_2)(t)}{dt} = a_2 W^*(t) E_2(t) - [d_2 + k_2] (W^*E_2)(t) . \quad (3.273)$$

A further simplification arises if we use the quasi-steady state approximation for the intermediate enzyme-substrate complexes and assume that the complex concentrations are negligible

$$W(t) + W^*(t) \gg (WE_1)(t) + (W^*E_2)(t) \quad (3.274)$$

in comparison to the protein concentration. The above assumption gives us the opportunity to relate directly the concentrations of the active and inactive protein to its overall concentration W^T . Alternatively, a transformation of rate equation (3.271) into expression in terms of the concentration of W^* is not possible.

From (3.272) follows the balance equation for the first complex

$$0 = a_1 [W^T - W^*(t)] [E_1^T - (WE_1)(t)] - [d_1 + k_1] (WE_1)(t)$$

which has the quasi-steady state solution

$$(WE_1)(t) = \frac{[W^T - W^*(t)] E_1^T}{K_{M1} + W^T - W^*(t)}. \quad (3.275)$$

In the same way we obtain with

$$(W^*E_2)(t) = \frac{W^*(t)E_2^T}{K_{M2} + W^*(t)}. \quad (3.276)$$

a similar expression for the second intermediate complex. As in the previous section the assumption of the quasi-steady state introduces the Michaelis constants

$$K_{M1} = \frac{d_1 + k_1}{a_1}, \quad K_{M2} = \frac{d_2 + k_2}{a_2}$$

into the mathematical representation. The rational form of solutions (3.275) and (3.276) results from the used enzyme conservation. We now insert the obtained expressions for both complexes into the differential equation (3.271) and obtain after some algebraic transformation¹⁸ the dynamic Goldbeter-Koshland function

$$\frac{dW^*(t)}{dt} = \frac{V_1 [W^T - W^*(t)]}{K_{M1} + W^T - W^*(t)} - \frac{V_2 W^*(t)}{K_{M2} + W^*(t)}, \quad (3.277)$$

where we introduce the abbreviations $V_1 = k_1 E_1^T$ and $V_2 = k_2 E_2^T$. This expression becomes slightly manageable, if we use dimensionless representations of the protein concentration and Michaelis constants

$$\overline{W^*}(t) = \frac{W^*(t)}{W^T}, \quad K_{1,2} = \frac{K_{M1,2}}{W^T}.$$

The mole fraction of active protein W^* is then determined by the differential equation

$$\frac{d\overline{W^*}(t)}{dt} = \frac{V_1 [1 - \overline{W^*}(t)]}{K_1 + 1 - \overline{W^*}(t)} - \frac{V_2 \overline{W^*}(t)}{K_2 + \overline{W^*}(t)}. \quad (3.278)$$

The corresponding steady state is defined by the balance equation

$$0 = (V_1 - V_2)\overline{W^*}^2 + [(V_1 + V_2)K_2 + (V_2 - V_1)]\overline{W^*} - V_1K_2 \quad (3.279)$$

which is a quadratic equation. An analysis of the solutions, see Figure 3.18, shows that only one fulfills all physical restrictions. The concentration of active protein W^* can lie only in the range $[0, W^T]$ and of course, it can not be negative. To distinguish between relevant and irrelevant solutions is a physical problem not a mathematical one.

The relevant solution

$$\begin{aligned} \frac{W_{SS}^*}{W^T} &= G(V_1, V_2, K_1, K_2) \\ &= \frac{\left(\frac{V_1}{V_2} - 1\right) - \left(\frac{V_1}{V_2} + \frac{K_1}{K_2}\right) K_2 + \sqrt{\left[\left(\frac{V_1}{V_2} - 1\right) - \left(\frac{V_1}{V_2} + \frac{K_1}{K_2}\right) K_2\right]^2 + 4\frac{V_1}{V_2} \left(\frac{V_1}{V_2} - 1\right) K_2}}{2\left(\frac{V_1}{V_2} - 1\right)} \end{aligned} \quad (3.280)$$

is the so-called Goldbeter-Koshland function [GKJ81] which is often abbreviated as $G(\dots)$. The steady state is primarily determined by the ratio of V_1 and V_2 . Because of the signal transduction in and between cells is carried out by changes of protein concentrations, this

¹⁸Use the relation $a_2 K_{M2} = d_2 + k_2$ to merge the degradation terms in (3.271).

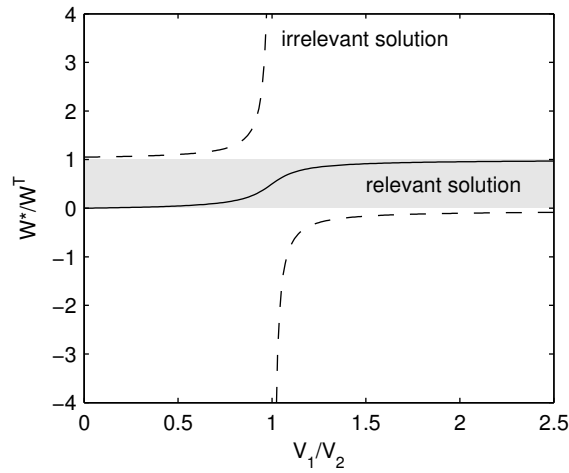


Fig. 3.18: The two solutions of balance equation (3.279) as function of the ratio of the limiting rates V_1 and V_2 . Only the solution inside the range $[0, W^T]$ is relevant for biological systems. The second one (dashed) breaks physical restrictions and is therefore irrelevant.

means that the ratio of the enzymes E_1 and E_2 or of kinase and phosphatase, respectively, plays a major role in the determination of the signaling state of a protein. The corresponding steady state (3.280) as function of the ratio of limiting rates is investigated in Figure 3.19. If we follow the previous assumption of negligible complex concentrations (3.274), the inactive protein can be calculated from the conservation relation (3.270) as

$$\frac{W}{W^T} = 1 - \frac{W^*}{W^T} .$$

Hence, both forms of protein W are directly linked as shown in Figure 3.19, where we use dashed lines for the inactive and solid lines for active state. The solution of the Goldbeter-

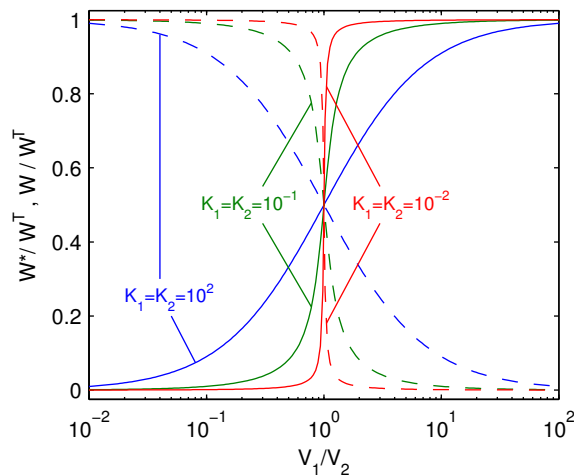


Fig. 3.19: Steady state of active (solid lines) and inactive (dashed lines) form of protein W as function of the ratio V_1/V_2 for different Michaelis constants.

Koshland function (3.280) has a typical sigmoidal shape. If the rate of deactivation is much greater than the activation corresponding to $V_1/V_2 \ll 1$ there is only a small fraction of activated protein W^* . On the other hand, in the limit of a much faster production rate almost all proteins are in the active form. The steady state as function of the ratio V_1/V_2 shows a switch-like behavior [FJ96]. Thereby the sharpness of the transition between a low level and a high level protein concentration or the ‘off’ and ‘on’-state, respectively, is determined by the values of the Michaelis constants K_1 and K_2 . As shown in Figure 3.19

small values lead to a sharp transition. Such a behavior is called ultrasensitivity. Because of the reason is a rate law following asymptotically a zero order kinetics¹⁹ this special feature of the covalent modification scheme is called zero-order ultrasensitivity. Apart from the discussed mechanism, there are further mechanisms leading to an ultrasensitive behavior, e.g. cooperativity and allosteric regulation (see for instance [CB04, Seg93]).

In Figure 3.19 we made with $K_1 = K_2$ a very special assumption. In fact, it is very improbable that two reactions have the same kinetic parameters. The steady state is a symmetric function of V_1/V_2 . The transition can be characterized by the inflection point located at $V_1/V_2 = 1$ and $\bar{W} = \bar{W}^* = 0.5$. The consequences of different constants are

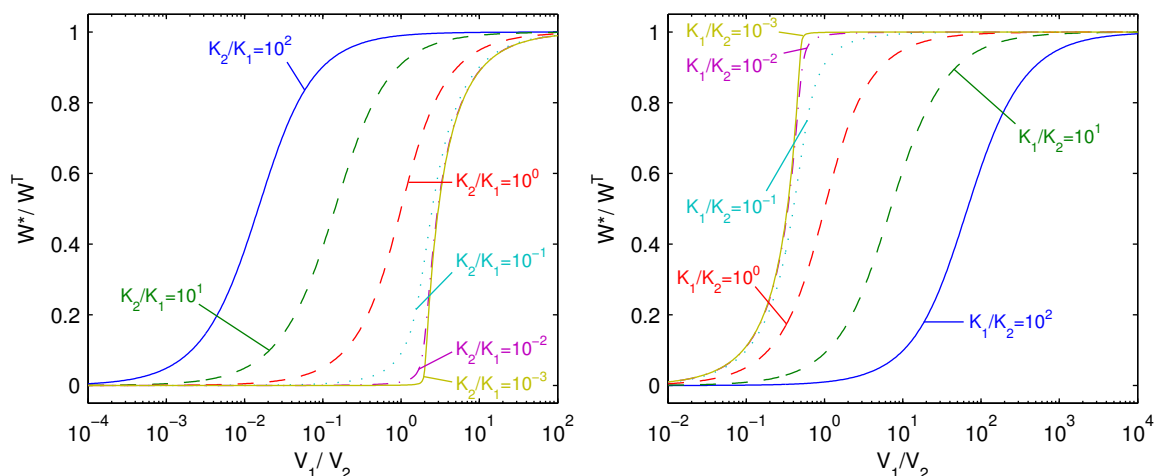


Fig. 3.20: Steady state of active protein W^* as function of the ratio V_1/V_2 . In contrast to Figure 3.19 we now assume different Michaelis constants K_1 and K_2 . In the left plot we vary K_2 and in the right plot K_1 . The red dashed line corresponds to a symmetric set of parameters as in Figure 3.19.

shown in Figure 3.20 where we vary the ratio K_1/K_2 . In the left plot we kept the parameter K_1 constant and in the right the parameter K_2 . Differences in the plots occur due to the fact that K_1 enters only through the ratio K_1/K_2 into the Goldbeter-Koshland function (3.280) whereas K_2 is also a multiplicative factor. The ratio of both parameters affects shape and position of the curves. The transition between a low level and a high level of activated protein is shifted as function of K_1 and K_2 . Furthermore the ultrasensitive properties are strongly influenced. Whereas some parameter values produce ultrasensitive behavior, this feature disappears for another combination of K_1 and K_2 .

¹⁹In the limits $W^* \ll W \approx W^T$ and $W \ll W^* \approx W^T$ we can reduce the dynamic Goldbeter-Koshland function (3.277) to a zero order kinetic expression whereby we made the assumption $W^T \gg K_{M1,2}$ and neglect the term corresponding to the degradation of the protein form with small concentration.

4 Stochastic Modelling and Simulation

The previous chapter derived differential equation models of biochemical reactions. We started off with the argument that molecules mix and collide, i.e., there is some randomness to the process. It makes therefore sense to describe changes in molecular populations in terms of stochastic processes.

Suppose the number of molecules of reactant S can be any positive integer n , $\#S \doteq n$ for ease of notation. The most elementary stochastic model would need to deal with $P_n(t)$, the probability that there are n molecules of S at time t . Considering the temporal evolution of this population, we regard P_n as a function of discrete time t and ask for the dynamic equations which govern its development. That is, we want to express $P_n(t + \Delta t)$ in terms of $P_m(t)$ for all m , that is, to express the probability of having n molecules at time $t + \Delta t$ in terms of the probabilities of all the possible values for the number, m of molecules at time t . The occurrence of *event* $\#S(t + \Delta t) = n$ can be thought of as the occurrence of either event the joint event $(\#S(t + \Delta t) = n, \#S(t) = 1)$ or event $(\#S(t + \Delta t) = n, \#S(t) = 2)$ or ... These events are mutually exclusive, and hence the addition law of probabilities implies that

$$P(\#S(t + \Delta t) = n) = \sum_m P(\#S(t + \Delta t) = n, \#S(t) = m) , \quad (4.1)$$

where we consider n as fixed for the moment, while m is varied. From the definition of conditional probability we have

$$P(\#S(t + \Delta t) = n, \#S(t) = m) = P(\#S(t) = m) \cdot P(\#S(t + \Delta t) = n \mid \#S(t) = m) .$$

Substituting this in (4.1), we get

$$P(\#S(t + \Delta t) = n) = \sum_m P(\#S(t + \Delta t) = n \mid \#S(t) = m) \cdot P(\#S(t) = m) .$$

Using a more compact notation where for each possible m we write $p_{m,n}$ for the *transition probability* of changing from m to n molecules in the interval Δt . This leads to the stochastic process model transition probability

$$P_n(t + \Delta t) = \sum_{m=1}^{\infty} p_{m,n} P_m(t) .$$

This stochastic representation, in terms of transition probabilities describes a class of stochastic processes called *Markov processes*. It is assumed that this transition probability does not itself explicitly depend upon the time at which the transition occurs, nor does it depend on the path on which the change occurred. In this case the Markov process, or *Markov chain*, is called *homogenous* with respect to time¹. The fact that the transitions depend only on the previous step/state is called the *Markov assumption*. One of the simplest cases of this is when the transition can only take place by an increase or a decrease of just one individual. Thus $p_{m,n} = 0$ unless $m = n - 1, n$ or $n + 1$. It then follows that Markov processes

$$p_{n,n} = 1 - p_{n,n+1} - p_{n,n-1} . \quad (4.2)$$

Thus the dynamic equation and stochastic equivalent of (3.5) is

$$P_n(t + \Delta t) = p_{n-1,n} \cdot P_{n-1}(t) + p_{n+1,n} \cdot P_{n+1}(t) + (1 - p_{n,n+1} - p_{n,n-1}) \cdot P_n(t) . \quad (4.3)$$

¹While for a homogenous Markov process the transition probabilities do not depend on time explicitly, they are through $n(t)$ implicitly a function of time.

The pattern of (4.3) will be a recurrent theme in subsequent sections. The probability of n molecules at time $(t + \Delta t)$ is composed of the sum of term which describe possible previous states, multiplied by their transition probabilities. The sum reflects a disjunction of statements reflecting events and the product reflects the conjunction (“and”) in describing the evolution of the system. One approach to define the transition probabilities is to assume they are proportional to n :

$$\begin{aligned} p_{n-1,n} &= k^+(n-1)\Delta t, & p_{n+1,n} &= k^-(n+1)\Delta t, \\ p_{n,n-1} &= k^-n\Delta t, & p_{n,n+1} &= k^+n\Delta t. \end{aligned}$$

Substituting this in (4.3), we get

$$\begin{aligned} P_n(t + \Delta t) &= k^+(n-1)\Delta t P_{n-1}(t) + k^-(n+1)\Delta t P_{n+1}(t) + (1 - k^+n\Delta t - k^-n\Delta t)P_n(t) \\ &= k^+(n-1)\Delta t P_{n-1}(t) + k^-(n+1)\Delta t P_{n+1}(t) - (k^+ + k^-)n\Delta t P_n(t) + P_n(t). \end{aligned}$$

Taking $P_n(t)$ to the left-hand side and dividing by Δt , we have

$$\frac{P_n(t + \Delta t) - P_n(t)}{\Delta t} = k^+(n-1)P_{n-1}(t) + k^-(n+1)P_{n+1}(t) - (k^+ + k^-)nP_n(t). \quad (4.4)$$

Taking the limit as $\Delta t \rightarrow 0$, gives us a differential equation $dP_n(t)/dt$ and the solution of which is an equation that gives us the probability of n molecules at time t , an example of the chemical master equation (2.21). Note that for simulation purposes, from (4.2), we have $p_{n,n+1} + p_{n,n-1} \leq 1$ and thus

$$\Delta t \leq \frac{1}{(k^- + k^+)n}.$$

The larger the number of molecules, the smaller the sampling intervals must be.

Suppose we would now like to derive from this a deterministic equation for the mean number of molecules at any time. This mean or ensemble average, denoted $\langle S(t) \rangle$, has the form

$$E[S(t)] = \sum_{n=1}^{\infty} n \cdot P_n(t). \quad (4.5)$$

We can therefore multiply the dynamic equation $P_n(t + \Delta t)$ by n and sum over n ; this automatically gives us on the left-hand-side the value of S at time $t + \Delta t$. The aim is to derive an expression in which the right-hand-side turns out to be expressible just in terms of S at time t ; that would lead to a deterministic dynamical equation for S . Let's see what happens if we insert (4.3) in (4.5). We get

$$\begin{aligned} E[S(t + \Delta t)] &= \sum_{n=1}^{\infty} n \left[p_{n-1,n} \cdot P_{n-1}(t) + p_{n+1,n} \cdot P_{n+1}(t) + (1 - p_{n,n+1} - p_{n,n-1}) \cdot P_n(t) \right] \\ &= \sum_{n=1}^{\infty} (n+1)p_{n,n+1}P_n(t) + \sum_{n=1}^{\infty} (n-1)p_{n,n-1}P_n(t) + \sum_{n=1}^{\infty} n(1 - p_{n,n+1} - p_{n,n-1})P_n(t) \\ &= \sum_{n=1}^{\infty} (n + p_{n,n+1} - p_{n,n-1})P_n(t). \end{aligned}$$

The most obvious way to make this depend upon $S(t)$ only is to make $p_{n,n+1}$, $p_{n,n-1}$ constant multiples of n , say $p_{n,n-1} = k^- \cdot n \cdot \Delta t$, $p_{n,n+1} = k^+ \cdot n \cdot \Delta t$ and let $\Delta t = 1$ be some unit time. In which case we get the deterministic equation

$$S(t + 1) = (1 + k^+ - k^-)S(t), \quad (4.6)$$

which is equivalent to (3.5) if we approximate dS/dt with (3.7). In general, this is an example of the linear first-order homogenous difference equation $S(t+1) = a(t)S(t)$, with solution

$$S(t) = \left[\prod_{i=1}^{t-1} a(i) \right] S(0) .$$

Note that the stochastic model we started off with, was *not* based on physical considerations of how molecules interact. The latter of which would involve a description of the velocities with which the molecules move.

4.1 A matter of life and death

In the present section we derive, like in the previous section, a stochastic equivalent for the simple reaction (3.2), but here we describe this process in the framework of *Poisson* and *birth-death processes*, which are special cases of Markov processes. The obvious conclusion from this section is then that there are many equivalent ways of deriving a chemical master equation. The different flavors of the derivations do however give us practice in thinking probabilistically. We find that there are different levels at which we can model a system: considering a population as a whole or starting off at the level of molecules and then generalizing towards a population of molecules. In either case, various assumptions have to be made as otherwise the model may be more realistic or too complicated. There may be assumption about the system or experimental set-up (e.g. constant temperature, constant volume), or assumptions for mathematical convenience (e.g. modelling single-step changes). Choosing a suitable modelling framework, whether stochastic or deterministic, and deriving the model is an *art*, which makes this process so interesting.

S is considered again a discrete variable, accounting for the actual number of molecules. For each model, we first consider probabilities describing individual molecules before aggregating these to describe changes in a population. This will be helped by the assumption that molecules are independent, i.e., whatever happens to one molecule is independent of other molecules (of the same species). This assumption may not always be realistic but is certainly convenient. Our strategy is to first consider a model for growth in a population and then a separate model for decay, before combining both into a stochastic model that is equivalent to (3.5).

Let us first look a process by which a molecule will ‘replicate’ itself, or is replicated by whatever means, and then consider the population. The probability for a molecule $S^{(j)}$ from a population with $j = 1, \dots, S(t)$ molecules to replicate itself during a small time interval Δt is

$$b_1(\Delta t) = k^+ \Delta t + o(\Delta t^2) , \quad (4.7)$$

where k^+ is the constant average rate (1/sec) at which the molecule replicates. Note that we use S , $\#S$ and n interchangeably to denote the number of molecules in the population. The term $o(\Delta t^2)$ accounts for other effects, which for small enough time intervals, are negligible. The transition probability function is thus approximated by a Taylor series and $o(\Delta t^2)$ describes terms of order Δt^2 and higher. The appearance of a term with Δt^2 can be explained as follows (see Figure 4.1). Consider the time interval $(t, t + \Delta t)$ being divided into two, $(t, t + \alpha\Delta t)$ and $(t + \alpha\Delta t, t + \Delta t)$, where $0 \leq \alpha \leq 1$. The probability for the events to occur in their respective time intervals is

$$b_1(\alpha\Delta t) = k^+ \alpha\Delta t \quad \text{and} \quad b_1((1 - \alpha)\Delta t) = k^+(1 - \alpha)\Delta t .$$

The probability for more than one event in $(t, t + \Delta t)$ is thus the probability for an event in $(t, t + \alpha\Delta t)$ and subsequently one in $(t + \alpha\Delta t, t + \Delta t)$, i.e., the product of the two probabilities above, leading to terms with Δt^2 .

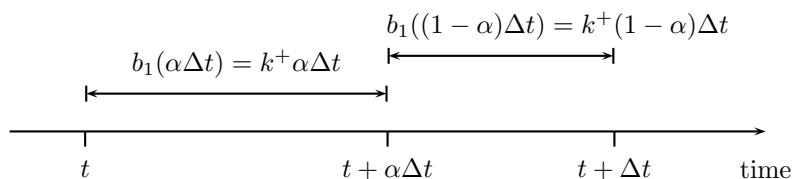


Fig. 4.1: The picture illustrates the occurrence of quadratic terms of Δt . See the text for an explanation.

The probability for $\#S$ not to change during the interval $(t, t + \Delta t)$ is

$$\begin{aligned} b_0(\Delta t) &= 1 - b_1(\Delta t) \\ &= 1 - k^+ \Delta t - o(\Delta t^2) . \end{aligned} \quad (4.8)$$

Equations (4.7) and (4.8) suggest that the probability could become negative. For this not to happen in simulations, Δt must be rather small in comparison to $1/k$. Finally the probability for any change m to be different to zero or one is negligible

$$b_m(\Delta t) = o(\Delta t^2) .$$

These considerations for one molecule should be extended to a population of $S(t)$ identical S molecules at time t . Define

$$\Delta S \doteq S(t + \Delta t) - S(t) .$$

Then the probability of no change in the entire population is obtained as

$$\begin{aligned} \text{Prob}\{\Delta S = 0\} &= \prod_{j=1}^{S(t)} b_0^{(j)} \\ &= \prod_{j=1}^{S(t)} (1 - k^+ \Delta t - o(\Delta t^2)) \\ &= 1 - S(t)k^+ \Delta t - o(\Delta t^2) . \end{aligned} \quad (4.9)$$

The probability for an increase of one in the entire population is obtained from the *union* (\sum) of single changes for a molecule j and (\prod) that the others do not replicate:

$$\begin{aligned} \text{Prob}\{\Delta S = +1\} &= \sum_{j=1}^{S(t)} \underbrace{b_1^{(j)}}_{(4.7)} \prod_{i \neq j} \underbrace{b_0^{(i)}}_{(4.8)} \\ &= \sum_{i=1}^{S(t)} k^+ \Delta t \underbrace{(1 - (S(t) - 1)k^+ \Delta t)}_{\text{from (4.9)}} - o(\Delta t^2) \\ &= S(t)k^+ \Delta t (1 - (S(t) - 1)k^+ \Delta t) - o(\Delta t^2) \\ &= S(t)k^+ \Delta t - \underbrace{S(t)(k^+)^2 \Delta t^2 (S(t) - 1)}_{\text{part of } o(\Delta t^2)} - o(\Delta t^2) \\ &= k^+ S(t) \Delta t - o(\Delta t^2) , \end{aligned} \quad (4.10)$$

and $\text{Prob}\{\Delta S = +m\}$ for $m \geq 2$ is $o(\Delta t^2)$, i.e., negligible. We have few more steps to take before we obtain the chemical master equation, so let us summarize where we are: If

at time t there are $n > 0$ molecules in the population, the probability that in the interval $(t, t + \Delta t)$ one molecule is added, equals (4.10), where $k^+ > 0$. The probability of two or more molecules to be added is vanishingly small, denoted $o(\Delta t^2)$. An expression for $P_n(t)$ is obtained by first considering the probability for n molecules at $t + \Delta t$, which can be the result of two previous situations: no change *or* an increase by one since time t . We consider a Δt such that other changes are negligible. If there is no change, we have for the first term the conjunction $\text{Prob}\{\Delta S = 0\}P_n(t)$, to represent the statement “ $P_n(t)$ and no change”. For the second term, a change of one implies that previously $S(t) = n - 1$ at t , leading to

$$P_n(t + \Delta t) = \text{Prob}\{\Delta S = 0\}P_n(t) + \text{Prob}\{\Delta S = +1, S(t) = n - 1\}P_{n-1}(t) .$$

Inserting (4.9) directly, and (4.10) with consideration for $S(t) - 1$, into this equation, we obtain

$$P_n(t + \Delta t) = (1 - nk^+\Delta t)P_n(t) + k^+(n - 1)\Delta tP_{n-1}(t) - o(\Delta t^2) . \quad (4.11)$$

The next step towards an equation for $P_n(t)$ is to divide (4.11) by Δt

$$\frac{P_n(t + \Delta t) - P_n(t)}{\Delta t} = -nk^+P_n(t) + k^+(n - 1)P_{n-1}(t) - \frac{o(\Delta t^2)}{\Delta t} .$$

Taking the limit $\Delta t \rightarrow 0$, leads us to a system of *differential-difference* equations

differential-difference
equation

$$\frac{dP_n(t)}{dt} = -k^+nP_n(t) + k^+(n - 1)P_{n-1}(t) \quad n = S(0), S(0) + 1, \dots \quad (4.12)$$

with initial conditions

$$\begin{aligned} P_n(0) &= 1 \quad \text{for } n = S(0) \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

Compare this with (4.4), noting that to this point we have only considered ‘birth’, that is, increases in populations. For equation (4.12) the first term on the right-hand side describes the source term for transitions from $n - 1$ to n and the right term is a sink term describing changes from n to $n - 1$. This set of equations is then the stochastic model or master equation for which the solution is $P_n(t)$. Although the differential-difference equations are linear, the set of equations is infinite and we can therefore not apply an matrix algebraic methods that can be used to solve linear differential equations. The solution requires the use of *moment generating functions*², leading to

$$P_n(t) = \binom{n - 1}{n - S(0)} e^{-k^+S(0)t} (1 - e^{-k^+t})^{n - S(0)}$$

for $n \geq S(0)$. The expected population size is obtained as

$$E[S(t)] = \sum_{n=0}^{\infty} nP_n(t) = S(0)e^{k^+t} , \quad (4.13)$$

which shows that the mean population size from the stochastic model, (4.13), is equal to the population size of the deterministic model (3.3). It is interesting that this relationship does not hold in general, i.e., the deterministic model does not always emerges as the mean of the stochastic model. We are going to consider a counter example in Section 4.2.

Before continuing with a similar model for a decreasing population, let us discuss the concept of treating a population of molecules as a birth-death process. Equation (4.7)

²Moment generating function feature prominently in Section 4.6.4.

considers a molecule as if it is an organism that can give birth. Looking at the human (female) population, the constant average rate k^+ would be determined by looking at the fact that on average a family has, say, two children and that these are born between the women's age of about 25-35. This means that to determine parameter k^+ we implicitly require the existence of a population from which to determine the value. We also note that in case of the human population, while for the individual woman it makes only sense to talk of changes approximately every 12 months, looking at the population, children can be born any hour, minute or second. In other words as the number of elements in a population increases we can approximate the discrete-time process by a continuous time model.

In analogy to the growth/birth-model (4.9) and (4.10), we now consider the decay/death of molecule populations:

$$\begin{aligned}d_0(\Delta t) &= 1 - k^- \Delta t - o(\Delta t^2) \\d_{-1}(\Delta t) &= k^- \Delta t + o(\Delta t^2) ,\end{aligned}$$

where k^- is the constant average decay rate. It follows

$$\begin{aligned}\text{Prob}\{\Delta S = 0\} &= 1 - S(t)k^- \Delta t - o(\Delta t^2) \\ \text{Prob}\{\Delta S = -1\} &= k^- S(t)\Delta t + o(\Delta t^2) .\end{aligned}\tag{4.14}$$

To combine formation and decay in one model we have two expressions for “no change”, which are combined as a product:

$$\text{Prob}\{\Delta S = 0\} = 1 - S(t)(k^+ + k^-)\Delta t - o(\Delta t^2) .\tag{4.15}$$

Together the transition probabilities (4.10), (4.14), and (4.15) describe the formation and decay of the population of molecules. For a full analysis of a general birth-death process see for example [JS01]. A notable result is that from an expression for the variance of the population, if $k^+ > k^-$, the behavior of the population is predicted with much less precision [Row94]. Due to the fact that we are dealing with a Poisson process, the variance of the population size increases without limit and the actual behavior is difficult to access.

To complete our collection of equations, for a death process, we would obtain a set of equations, analogous to (4.12):

$$\frac{dP_n(t)}{dt} = -k^- n P_n(t) + k^- (n+1) P_{n+1}(t) .$$

The differential-difference master equations for the combined birth-death process were derived above, see (4.4):

$$\begin{aligned}\frac{dP_0(t)}{dt} &= k^- P_1(t) \\ \frac{dP_n(t)}{dt} &= k^+(n-1)P_{n-1}(t) - (k^+ + k^-)nP_n(t) + k^-(n+1)P_{n+1}(t) .\end{aligned}$$

The expected population size for $k^+ \neq k^-$ is given by

$$E[S(t)] = S(0) \cdot e^{(k^+ - k^-)t} ,$$

as we set out to show, beginning with (3.5). We note that these birth-death processes assumed that the total rates k^+ and k^- are proportional to the population size. This can be generalized to consider possibly more accurate cases in which for example k^- increases for larger population sizes, a situation referred to as *crowding*.

Looking at (4.7) we might argue that with the assumption of a constant average rate, the crucial element of the random process considered, is *when* an event occurs. A standard model for stochastic processes where independent events, occur randomly in time, is the *Poisson process*. Let us take this perspective on the events occurring in molecular populations. Treating $S^{(j)}(t)$ as a time-varying random variable, representing the population size at time t , we consider the continuous-time probability distribution of the population that has arisen from $S^{(j)}$

$$\text{Prob} \left\{ S^{(j)}(t) = m \right\} \doteq p_m(t) = \frac{(\lambda t)^m e^{-\lambda t}}{m!}, \quad t \geq 0, \quad (4.16)$$

where $m = 0, 1, 2, \dots$ and λ is the rate of arrival/replication for a molecule indexed by j . Note that there we are looking at the population of molecules that has arisen from molecule $S^{(j)}$ (4.16) is thus a Poisson probability distribution function with parameter λ . For this to be a probability distribution, we must have

$$\sum_{m=0}^{\infty} p_m(t) = 1.$$

Deriving the derivative of (4.16), we note that

$$\frac{d}{dt} e^{-\lambda t} = -\lambda e^{-\lambda t} \quad \text{and} \quad \frac{d}{dt} \frac{(\lambda t)^m}{m!} = \frac{\lambda^m}{m!} m t^{m-1}.$$

Then using the product rule for differentiation, $d(u \cdot v)/dt = du/dt \cdot v + u \cdot dv/dt$,

$$\begin{aligned} \frac{dp_m}{dt} &= \frac{\lambda^m}{m!} m t^{m-1} \cdot e^{-\lambda t} - \frac{\lambda^m t^m}{m!} \cdot \lambda e^{-\lambda t} \\ &= \frac{(m t^{m-1} - \lambda t^m)}{m!} \lambda^m e^{-\lambda t} \\ &= \frac{(m - \lambda t)}{m!} \lambda^m t^{m-1} e^{-\lambda t}, \end{aligned} \quad (4.17)$$

where in the last step we made use of the fact that

$$m t^{m-1} - \lambda t^m = t^{m-1} \left(m - \lambda \frac{t^m}{t^{m-1}} \right) = t^{m-1} (m - \lambda t).$$

The maximum values of the probabilities for fixed m occur at $t = m/\lambda$, where $dp_m(t)/dt = 0$. Say we are looking for a change of one molecule for $S^{(j)}(t_0) = 0$, this is most likely to happen in the interval $\Delta t = 1/\lambda$. In simulating such a process, we would therefore require a sampling rate of about $1/\lambda$ or less to be able to capture the dynamics at a high enough resolution. For $m = 0$,

$$\frac{dp_0(t)}{dt} = -\lambda e^{-\lambda t}. \quad (4.18)$$

From (4.16), (4.17), (4.18) it follows that

$$\frac{dp_0(t)}{dt} = -\lambda p_0(t), \quad (4.19)$$

$$\frac{dp_m(t)}{dt} = \lambda (p_{m-1}(t) - p_m(t)), \quad m \geq 1. \quad (4.20)$$

We can consider these differential-difference equations to have arisen from a limiting process

$$\frac{dp_m(t)}{dt} = \lim_{\Delta t \rightarrow 0} \frac{p_m(t + \Delta t) - p_m(t)}{\Delta t},$$

such that for small Δt , we can write

$$\begin{aligned}\frac{p_0(t + \Delta t) - p_0(t)}{\Delta t} &\approx -\lambda p_0(t) , \\ \frac{p_m(t + \Delta t) - p_m(t)}{\Delta t} &\approx \lambda(p_{m-1}(t) - p_m(t)) ,\end{aligned}$$

so that

$$\begin{aligned}p_0(t + \Delta t) &\approx (1 - \lambda\Delta t)p_0(t) \\ p_m(t + \Delta t) &\approx p_{m-1}(t)\lambda\Delta t + p_m(t)(1 - \lambda\Delta t) .\end{aligned}\tag{4.21}$$

The probability of a single increase in the population can be obtained by setting $m = 1$ and assuming there were zero elements at t_0 , i.e., $p_0(t) = 1$ and $p_1(t) = 0$. Inserting this into (4.21), the probability for a single molecule increase in the short time interval Δt is equal to $\lambda\Delta t$ and that the probability for increases of two or more molecules is negligible. This conforms with what we said for (4.7) and we might as well have started with the Poisson model. (4.21) considers a molecule but for a population of n molecules, we must replace the rate λ with n multiples of k^+ with regard to P_n and with $k^+(n - 1)$ w.r.t. P_{n-1} , and where P_n replaces p_m such that we would obtain (4.11) as a representation for the entire molecular population. We have therefore n independent Poisson processes, each represent a molecule, occurring in parallel.

Note that a stochastic model is in mathematical terms a *stochastic process* - a mathematical abstraction, specifically a sequence of random variables, of an biochemical process whose development is governed by probabilistic laws. A stochastic process is thus a mathematical concept and should not be confused with the biological process itself.

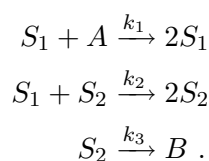
4.2 Mean values of Stochastic Processes

In the previous two sections we found for the monomolecular reaction, that a differential equation model was obtained from taking the expectation of the stochastic chemical master equation model. In other words, the ODE model determined the mean value of the stochastic model. This raises the question, whether this is true in general. In this section we look at an example where taking the expectation of the stochastic model does not recover the deterministic model of differential equations. We are going to conclude that chemical master equations are not suitable to describe average processes.

A frequently studied system of nonlinear differential equations was first investigated by Lotka and Volterra in 1925 and 1926. The *coupled* differential equations are given as

$$\begin{aligned}\frac{dS_1}{dt} &= k_1AS_1 - k_2S_1S_2 \\ \frac{dS_2}{dt} &= k_2S_1S_2 - k_3S_2 .\end{aligned}$$

In its original context, S_1 is to symbolize the prey, S_2 the predator, and A the food of the prey. As a reaction diagram we could write



The first equation describes how S_1 increases, facilitated by A . The second equation represents how S_2 increases, facilitated by S_1 . The last equation describes decay of S_2 by other means.

A probabilistic model is based on the assumption that in an infinitesimal time Δt , probability for changes in the number of molecules is given by

$$\begin{aligned}\text{Prob}\{\Delta S_1 = 1 \mid \Delta S_2 = 0\} &= k_1 A S_1 \Delta t \\ \text{Prob}\{\Delta S_1 = -1 \mid \Delta S_2 = 1\} &= k_2 S_1 S_2 \Delta t \\ \text{Prob}\{\Delta S_1 = 0 \mid \Delta S_2 = -1\} &= k_3 S_2 \Delta t \\ \text{Prob}\{\Delta S_1 = 0 \mid \Delta S_2 = 0\} &= 1 - (k_1 A S_1 + k_2 S_1 S_2 + k_3 S_2) \Delta t\end{aligned}$$

For the sake of an easier notation, we write S instead of $\#S$ to denote the number of molecules of S . Next, we use the Chapman-Kolmogorov equation to write the probability at $t + \Delta t$ as a sum of terms, each of which represents the probability of a previous state multiplied by the probability of a transition to state (S_1, S_2) . We get

$$\begin{aligned}\frac{P(S_1, S_2, t + \Delta t) - P(S_1, S_2, t)}{\Delta t} &= k_1 A (S_1 - 1) P(S_1 - 1, S_2, t) \\ &\quad + k_2 (S_1 + 1)(S_2 - 1) P(S_1 + 1, S_2 - 1, t) \\ &\quad + k_3 (S_2 + 1) P(S_1, S_2 + 1, t) \\ &\quad - (k_1 A S_1 + k_2 S_1 S_2 + k_3 S_2) P(S_1, S_2, t) .\end{aligned}$$

Suppose we now wish to get the equation for the expectation of S_1 :

$$E[S_1] = \sum_{S_1, S_2=0}^{\infty} S_1 P(S_1, S_2, t) .$$

We have

$$\begin{aligned}\dot{E}[S_1] &= \sum_{S_1, S_2=0}^{\infty} S_1 \dot{P}(S_1, S_2, t) \\ &= \sum_{S_1, S_2=0}^{\infty} S_1 \left[k_1 A (S_1 - 1) P(S_1 - 1, S_2, t) + k_2 (S_1 + 1)(S_2 - 1) P(S_1 + 1, S_2 - 1, t) \right. \\ &\quad \left. + k_3 (S_2 + 1) P(S_1, S_2 + 1, t) - (k_1 A S_1 + k_2 S_1 S_2 + k_3 S_2) P(S_1, S_2, t) \right] .\end{aligned}$$

Changing the sum indices to get every term to contain $P(S_1, S_2, t)$ we find, after some cancelations,

$$\dot{E}[S_1] = \sum_{S_1, S_2=0}^{\infty} k_1 A S_1 - k_2 S_1 S_2 = k_1 A E[S_1] - k_2 E[S_1 S_2] . \quad (4.22)$$

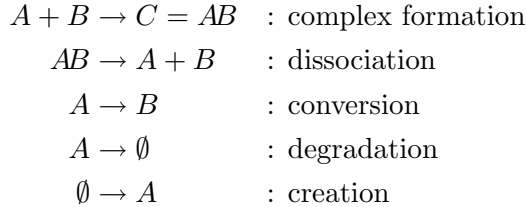
The next step would require

$$E[S_1(t) S_2(t)] = E[S_1(t)] \cdot E[S_2(t)] \quad \text{at any } t,$$

to make it an equation about expectations only: at present it is an equation about cross-correlations between the variables. Viewing $S_1(t)$ and $S_2(t)$ as (mean removed) time series, the equation above implies that they are *uncorrelated* sequences. Viewing S_1 and S_2 as random variables, the equation implies they are *independent* random variables. That last move to turn the stochastic model into a deterministic ‘‘average representation’’ would thus require the variables S_1, S_2 to be independent, but the very nature of the given *coupled* differential equations is a statement that this is not the case. Chemical master equations are therefore not suitable to model average processes. On the other hand, looking at the molecular interactions, collisions of usually larger numbers of molecules it seems that it would not make sense to have an accurate, physically realistic model. Instead in most cases the experimentalists obtain the rate constants k , which implicitly considers an ‘‘average process’’.

4.3 Discussion: ODEs and Stochastic Processes

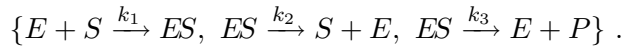
In considering biochemical reaction networks or pathways, we suppose a compartment with volume V , containing a spatially homogenous mixture of $\#S_i$ molecules of chemical species S_i , $i = 1, \dots, N$. These N species can interact through M reaction channels R_μ , $\mu = 1, \dots, M$. Examples for elementary reactions are:



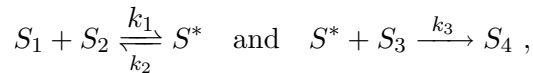
More complex reaction schemas can be constructed from those elementary irreversible ones. Note that each R_μ reaction is therefore uni-directional and any reversible reaction is decomposed into two separate *reaction channels*. For example, the common enzyme-kinetic reaction



can be decomposed into a set of three basic reactions:



Note that, depending of whether a reaction is monomolecular, bimolecular, or trimolecular (i.e., how many reactant molecules are involved on the left-hand side of the reaction diagram), the rate constant k will have different units. So for example, for $ES \xrightarrow{k} \dots$ we would obtain an ODE equivalent of the kind $\dot{S} = kS$ and k is in sec^{-1} , regardless of whether S is a concentration or a count of molecules. For the bimolecular reaction $E + S \xrightarrow{k} \dots$, the ODE equivalent is $\dot{S} = kS_1S_2$ such that k is in $\text{M}^{-1}\text{sec}^{-1}$ if we consider counts of molecules (in mol) and $\text{L}/(\text{mol sec})$ if we consider molar concentrations. For a trimolecular reactions $S_1 + S_2 + S_3 \rightarrow \dots$ and $\dot{S} = kS_1S_2S_3$, k is in $\text{M}^{-2}\text{sec}^{-1}$. For three molecules to simultaneously collide in an infinitesimal interval of time, dt , is proportional to dt^2 , and generally very unlikely³, and the probability of which is denoted $o(\Delta t^2)$. In fact, this was requirement for the Markovian framework of the stochastic models developed in Section 3. If we nevertheless want to consider trimolecular reactions in stochastic models, this can be done by introducing an unstable intermediate product with a short life time, S^* such that



where $1/k_2$ is the life time of the intermediate product, and $k_2 \gg k_1$ such that we can consider the trimolecular reaction in terms of two bimolecular and one monomolecular reactions.

The most important statement and fundamental assumptions for modelling are that molecules move independently, i.e., we are dealing with sufficiently low concentrations, that the volume and temperature are constant. This seems to contradict the common statement that for low concentrations a stochastic model is more appropriate, while an

³The get-to-together of three molecules is only unlikely if one subscribes to the collision theory that treats molecules as floating around in gas-phase. We should not forget that we are aiming at an understanding of processes within cells, which provides a very different environment to what is considered in the kinetic theory of gases.

mass action model is considered for larger fluxes of molecules. However, a large number of molecules does not imply a high concentration, if the action takes place in a large volume. Looking at the reversible reaction, $S_1 + S_2 \leftrightarrow S_3$, the complex $S_3 = S_1S_2$ is formed at a rate k_f , and dissociates at a rate k_d . We find that if we derive a mass action model, we implicitly make use of an stochastic interpretation. The change in S_3 has two components:

- (f) Increasing the number of S_1 or S_2 molecules, will proportionally increase the number of S_1 - S_2 encounters that give rise to S_3 .
- (d) A given S_3 molecule acts entirely independent of other S_3 molecules, which means that it has a certain probability per unit time of breaking apart into S_1 and S_2 . This is equivalent to saying the rate at which S_3 molecules break apart is proportional to the number of molecules present.

Putting this together in an equation, we have

$$\frac{d[S_1]}{dt} = -k_f[S_1][S_2] + k_d[S_3]. \quad (4.24)$$

The rate equations for S_1 and S_2 are simply equal to $-d[S_3]/dt$, i.e., when a molecule S_3 is formed, an S_1 molecule and a S_2 molecule disappear. Table 4.1 summarizes basic biochemical reactions and their ODE representation.

Table 4.1: Basic reactions and their mass action representation.

Reaction Syntax	ODEs
$S_1 \xrightarrow{k} S_2$	$[\dot{S}_1] = -k[S_1]$ $[\dot{S}_2] = k[S_1]$
$S_1 + S_2 \xrightarrow{k} S_3$	$[\dot{S}_1] = -k[S_1][S_2]$ $[\dot{S}_2] = -k[S_1][S_2]$ $[\dot{S}_3] = k[S_1][S_2]$
$S_1 + lS_2 \xrightarrow{k} S_3$	$[\dot{S}_1] = -k[S_1][S_2]^l$ $[\dot{S}_2] = -kl[S_1][S_2]^l$ $[\dot{S}_3] = k[S_1][S_2]^l$
$S + S \xrightarrow{k} \dots$	$[\dot{S}] = -2k[S]^2$
$S_1 + S_2 \xrightarrow{k} S_3 + S_3$	$[\dot{S}_3] = 2k[S_1][S_2]$
$S_1 \xrightleftharpoons[k_2]{k_1} S_2$	$[\dot{S}_1] = -k_1[S_1] + k_2[S_2]$ $[\dot{S}_2] = k_1[S_1] - k_2[S_2]$
$S_1 + S_2 \xrightleftharpoons[k_2]{k_1} S_3$	$[\dot{S}_1] = -k_1[S_1][S_2] + k_2[S_3]$ $[\dot{S}_2] = -k_1[S_1][S_2] + k_2[S_3]$ $[\dot{S}_3] = k_1[S_1][S_2] - k_2[S_3]$
$\emptyset \xrightarrow{k} S$	$[\dot{S}] = k$
$S \xrightarrow{k} \emptyset$	$[\dot{S}] = -k[S]$

For the enzyme kinetic reaction (4.23) we obtain the following set of nonlinear coupled differential equations as a special case of the mass action model (2.19)

$$\begin{aligned} \frac{d[S]}{dt} &= -k_1[S]([E]_0 - [ES]) + k_2[ES], \\ \frac{d[ES]}{dt} &= -(k_2 + k_3)[ES] + k_1[S]([E]_0 - [ES]), \end{aligned}$$

where we write $[S]$ instead of $[S](t)$ and $[S]_0$ instead of $[S](0)$, for short. The conservation law $[E]_0 = [E] - [ES]$ reflects the role of E as a catalyst for which the total enzyme (free and bound) remains constant. From (2.21), the chemical master equation is

$$\begin{aligned} \frac{dP(\#S, \#ES; t)}{dt} = & -[k_1\#S(\#E_0 - \#ES) + (k_2 + k_3)\#ES]P(\#S, \#ES; t) \\ & + k_1(\#S + 1)(\#E_0 - \#ES + 1)P(\#S + 1, \#ES - 1; t) \\ & + k_2(\#ES + 1)P(\#S - 1, \#ES + 1; t) \\ & + k_3(\#ES + 1)P(\#S, \#ES + 1; t) , \end{aligned} \quad (4.25)$$

subject to the assumptions and boundary conditions associated with the CME and which are discussed in [vK92]. Note that each equation (4.25) describes the probability of a state, for which there are many possibilities and hence many equations.

The previous sections introduced generalized mass action models and chemical master equations by specific examples. Since for virtually all cases the CMEs, like (4.25), cannot be solved analytically, a numerical simulation is required. The next section will look at one popular approach to stochastic simulation. This will be followed by further investigations into the differences between the mass action and CME approach.

4.4 Stochastic Simulation

In this section we introduce a generic framework for stochastic modelling based on chemical master equations and for which an elegant simulation algorithm is available. Here the simulation of a model is understood as an ‘execution of’ or ‘experiment with’ the model. For mass action models, simulation refers to numerical integration. The simulation of a stochastic model is the subject of Section 4.4.4.

state In the stochastic approach, the *state* of the system at time t , is defined by the number of molecules of each chemical species S_i :

$$\begin{aligned} S(t) & \doteq (\#S_1(t), \#S_2(t), \dots, \#S_i(t), \dots, \#S_N(t)) \\ & \equiv n , \end{aligned}$$

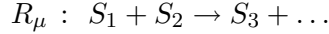
where $n \in \mathbb{Z}_+^N$ for each molecular species S_i we use the superscript $S^{(i)}$ to index individual molecules. In [Gil92b] Gillespie showed that for a chemically reacting system that is well stirred and in thermal equilibrium, the chemical master equations are a realistic microphysical representation, i.e., one that models the collision of physical objects and their velocities. The underlying assumptions that the reactants are in a fixed volume V and in gas-phase, well mixed, and in thermal equilibrium (at a constant temperature) may appear outrageous, considering the conditions in a cell, the models emerging from this framework may nevertheless be useful. We are reminded of Box’s dictum “All models are wrong but some are useful.”

The argument for the stochastic approach is that the evolution of the state (i.e., the variations of the reactant populations) cannot be a purely deterministic process. The basis for the stochastic approach is to set up and solve the chemical master equation. A system is then described by the evolution of the function (see also page 37)

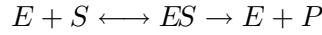
$$P(S_1, \dots, S_N; t) = P(S(t)) \equiv P_n(t)$$

which describes the probability that there will be $\#S_i$ molecules of species S_i at time t in volume V . The behavior of an entire system of reactions is thus described by a set of

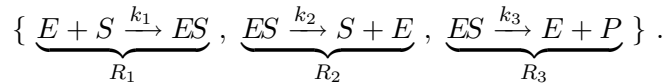
differential equations with probabilities as variables. Here we focus mostly on *bimolecular* reactions



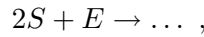
where the collision of *reactant species* S_1 and S_2 (one molecule each) leads to a conversion into one *product species* S_3 etc. S_1 and S_2 are allowed to be the same. Reversible reactions will be considered as two elementary reactions. For example, considering the enzyme kinetic reaction (4.23)



can be decomposed into a set of three basic reactions R_μ , $\mu = 1, \dots, M = 3$:



Let K_μ denote the number of participating reactant species in a basic uni-directional reaction channel R_μ , i.e., those on the left-hand side of the reaction diagram. K_μ is called *molecularity* of reaction R_μ . In our example $K_1 = 2$ (bimolecular), $K_2 = K_3 = 1$ molecularity (monomolecular). For a trimolecular reaction, for instance



we have a multiplicity of reactant molecules involved in R_μ , which we denote by $l_{\mu i}$, here $l_{\mu 1} = 2$ and $l_{\mu 2} = 1$. $l_{\mu i}$ is also called *stoichiometric coefficient* of reactant S_j in reaction R_μ . For the enzyme kinetic reaction $l_{11} = l_{12} = l_{21} = l_{31} = 1$. Finally, let L_μ denote the number of reactant species such that the number of participating reactant molecules in a reaction is

$$K_\mu = \sum_{j=1}^{L_\mu} l_{\mu j}.$$

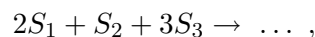
For the enzyme-kinetic reaction, $L_1 = 2$, $L_2 = L_3 = 1$. If there are $n_{p(\mu,j)}$ molecules of type $S_{p(\mu,j)}$, where index $p(\mu,j)$ selects those S_i participating in R_μ , then the total number of subsets, consisting of $l_{\mu j}$ reactants $S_{p(\mu,j)}$, in reaction channel R_μ , equals the binomial coefficient

$$\binom{n_{p(\mu,j)}}{l_{\mu j}} = \frac{(n_{p(\mu,j)})!}{l_{\mu j}!(n_{p(\mu,j)} - l_{\mu j})!} = \frac{1}{l_{\mu j}!} \left(n_{p(\mu,j)}(n_{p(\mu,j)} - 1) \times \dots \times [n_{p(\mu,j)} - (l_{\mu j} - 1)] \right).$$

Taking account for all reaction channel R_μ , we have the following expression for the number of distinct combinations of R_μ reactant molecules:

$$h_\mu = \begin{cases} \prod_{j=1}^{L_\mu} \binom{n_{p(\mu,j)}}{l_{\mu j}} & \text{for } n_{p(\mu,j)} > 0, \\ 0 & \text{otherwise.} \end{cases} \quad (4.26)$$

For example, let R_μ be defined as



then $L_\mu = 3$, $l_{\mu 1} = 2$, $l_{\mu 2} = 1$, and $l_{\mu 3} = 3$, such that

$$h_\mu = \prod_{j=1}^{L_\mu} \binom{n_{p(\mu,j)}}{l_{\mu j}} = \binom{n_1}{2} \cdot \binom{n_2}{1} \cdot \binom{n_3}{3}.$$

The expression for h_μ will be important in the following two sections, when we derive an expression for the probability that a reaction R_μ , which is part of a more complex reaction

schema, will take place in dt . If $n_{p(\mu,j)}$ in (4.26) is large and $l_{\mu j} > 1$, terms like $(n_{p(\mu,j)} - 1)$, \dots , $(n_{p(\mu,j)} - l_{\mu j} + 1)$ will not be much different from $n_{p(\mu,j)}$ and we may write

$$h_{\mu} \cong \prod_{j=1}^{L_{\mu}} \frac{(n_{p(\mu,j)})^{l_{\mu j}}}{l_{\mu j}!} = \frac{\prod_{j=1}^{L_{\mu}} (n_{p(\mu,j)})^{l_{\mu j}}}{\prod_{j=1}^{L_{\mu}} l_{\mu j}!}. \quad (4.27)$$

It should however be noted that this is an approximation, which can effect results in studies that compare mass action models with $l_{\mu j} > 1$ and stochastic simulations for small molecular populations. In fact, as we are going to show, it is misleading to compare a mass action model with stochastic simulation *as alternatives*.

4.4.1 Gillespie modelling

In most cases, the chemical master equation (2.21) is difficult or impossible to solve analytically. Gillespie developed subsequently in the 1970's a stochastic simulation algorithm [Gil77] and showed that the chemical master equation and the stochastic simulation are derived from the same set of theorems, and they are therefore logically equivalent to each other. Gillespie's stochastic model, the consequence of the physical model of reactive collisions (i.e., collisions which result in chemical alterations of the colliding molecules), is an alternative to the 'reaction rate' k of the mass-action type models we introduced in Section 3. In his approach, chemical reactions are characterized by a 'reaction probability per unit time'.

Asserting a *stochastic rate constant* c_{μ} for reaction channel R_{μ} , which only depends on the physical properties of the molecules and the temperature of the system,

$c_{\mu}dt$ is the probability that a particular selected combination of R_{μ} reactant molecules at time t will react in the next infinitesimal time interval $(t, t + dt)$.

$c_{\mu}dt$ is the probability that a particular selected combination of R_{μ} reactant molecules at time t will react in the next infinitesimal time interval $(t, t + dt)$. A reaction requires two separate phenomena: a collision to occur and for the collision to be reactive. In [Gil92b, Gil92a], for bimolecular reactions, Gillespie derived an expression for c_{μ} that contains a probability that a colliding pair of R_{μ} reactant molecules will chemically react. This probability is generally unknown. For trimolecular reactions the only relationship that can be derived from physical principles is the proportionality $c_{\mu} \propto V^{-K_{\mu}+1}$, where $K_{\mu} = 1, 2$, or 3 , and even this requires further unrealistic assumptions as Gillespie admits in [Gil92a]. Since a physical derivation for c_{μ} is in general not possible, implementations of c_{μ} in algorithms are relying on other arguments. Such a derivation will be introduced later on. In [Gil92a] he also showed how the linear relationship $c_{\mu}dt$ is justified on a mathematical basis. A consequence of this derivation is that c_{μ} must be analytical⁴. This can, for example, be achieved by keeping c_{μ} constant. If we multiply the probability $c_{\mu}dt$, which applies to a particular selected combination of reactant molecules, by the total number of distinct combinations of R_{μ} reactant molecules in V at time t , we obtain the probability that an R_{μ} will occur somewhere inside V in the next infinitesimal time interval $(t, t + dt)$. This leads us to

$$c_{\mu} \cdot h_{\mu} dt \equiv a_{\mu} dt$$

⁴A function is analytical if it has derivatives of all orders at any point.

or

$$a_\mu = c_\mu \cdot h_\mu \quad (4.28)$$

and

$$P_\mu(dt) = a_\mu dt . \quad (4.29)$$

as the probability that an R_μ reaction will occur in V in $(t, t + dt)$, given that the system is in state S at time t . a_μ is the *propensity* of the R_μ reaction. A propensity is thus understood as probability per unit time. Before deriving an expression for the propensity a_μ that deals with general reaction schemas, let us have a look at some simple cases we have dealt with before.

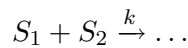
For the monomolecular reaction, $S \xrightarrow{k} \dots$, in which only one molecule of S is required for the reaction R_μ to take place, the *propensity* is simply

$$a_\mu = k_\mu(\#S) . \quad (4.30)$$

The probability of reaction R_μ to be realized in the infinitesimal time interval $(t, t + dt)$ is

$$P_\mu(dt) = k_\mu(\#S) dt .$$

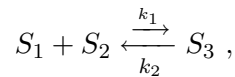
Considering the formation of a complex,



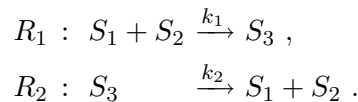
and labeling each molecule of $S_{p(\mu,j)}$ by $S_{p(\mu,j)}^{(i)}$, $i = 1, 2, \dots, \#S_{p(\mu,j)}$, we have $\#S_1 \cdot \#S_2$ distinct copies of the reaction that can occur:

$$a_\mu = k(\#S_1)(\#S_2)$$

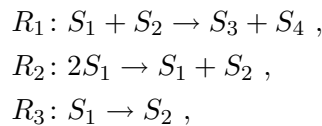
For the reversible reaction



we must split this into two separate reaction channels first:



Considering the following chemical reaction network with reaction channels



the number of distinct R_μ molecular reactant combinations h_μ for these channels are

$$\begin{aligned} h_1 : \#S_1\#S_2 , \\ h_2 : \#S_1(\#S_1 - 1)/2 , \\ h_3 : \#S_1 . \end{aligned}$$

Let $\nu_{\mu j}$ denote the change in the $S_{p(\mu,j)}$ molecular population caused by the occurrence of one R_μ reaction. The $\nu_{\mu j}$ values for the reaction channels above are

$$\begin{array}{llllll} \nu_{11} = -1 , & \nu_{12} = -1 , & \nu_{13} = +1 , & \nu_{14} = +1 , & \text{all other } \nu_{1j} = 0 \\ \nu_{21} = -1 , & \nu_{22} = +1 , & & & \text{all other } \nu_{2j} = 0 \\ \nu_{31} = -1 , & \nu_{32} = +1 , & & & \text{all other } \nu_{3j} = 0 \end{array}$$

In the next section we are going to derive a general expression for a_μ , c_μ , respectively, which will enable us to simulate realizations of more complex reaction pathway diagrams. Table 4.2 summarizes the stochastic parameters for basic reaction channels R_μ . Note that even if the relationship between the stochastic rate constant c_μ and the reaction rate k is only a constant factor for these simple examples, the conceptual difference between the stochastic and deterministic formulation is rather more complicated. The stochastic rate constant c_μ , multiplied by h_μ , is a propensity and thus referring to a stochastic model of a population of molecules. In contrast, the rate constants k_μ are indeed *rates* in the sense of a velocity. The fact that the c_μ , and hence any stochastic simulation, is dependent on knowledge of the rate constants of the mass action model is no coincidence. It has not been feasible to derive an expression for c_μ from physical principles without knowledge of either the rate constants or the probabilities that a colliding set of R_μ reactant molecules will chemically react. We are going to discuss these issues in the following sections in greater detail.

Table 4.2: Stochastic rate constant c_μ and number of distinct R_μ reactant combinations h_μ for elementary reaction channels R_μ .

R_μ	c_μ	h_μ
$S_1 + S_2 \xrightarrow{k} S_3$	k/V	$\#S_1 \cdot \#S_2$
$S_1 + S_2 \xrightarrow{k} 2S_1$	k/V	$\#S_1 \cdot \#S_2$
$S_1 + S_2 + S_3 \xrightarrow{k} S_4$	k/V^2	$\#S_1 \cdot \#S_2 \cdot \#S_3$
$S_1 \xrightarrow{k} \dots$	k	$\#S_1$
$2S_1 \xrightarrow{k} S_2$	$\frac{2! \cdot k}{V}$	$\frac{1}{2} \cdot \#S_1 \cdot (\#S_1 - 1) = \binom{\#S_1}{2}$
$2S_1 \xrightarrow{k} S_1 + S_2$	$\frac{2! \cdot k}{V}$	$\frac{1}{2} \cdot \#S_1 \cdot (\#S_1 - 1) = \binom{\#S_1}{2}$
$3S_1 \xrightarrow{k} S_2$	$\frac{3! \cdot k}{V}$	$\frac{1}{6} \cdot \#S_1 \cdot (\#S_1 - 1) \cdot (\#S_1 - 2) = \binom{\#S_1}{3}$
$l \cdot S_1 \xrightarrow{k} S_2$	$\frac{l! \cdot k}{V^{l-1}}$	$\binom{\#S_1}{l}$
$S_1 + 2S_2 \xrightarrow{k} S_3$	$\frac{2! \cdot k}{V^2}$	$\#S_1 \cdot \frac{1}{2} \cdot \#S_2 \cdot (\#S_2 - 1) = \#S_1 \cdot \binom{\#S_2}{2}$
$\emptyset \rightarrow S$	1	1

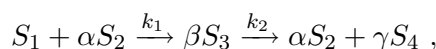
4.4.2 Stochastic rate constant versus rate constant

In this section we summarize the relationship between the stochastic rate constant c_μ and the rate constant k_μ in an ODE formulation of the mass-action law.

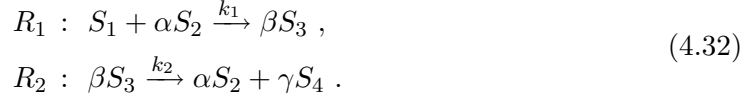
Let us start with a simple example. We are interpreting the rate constant, at the molecular level, as the probability that a particular combination of reactants is realized such that reaction R_μ will occur in the small unit time interval. For example, in Section 3, we were considering a reaction with only one reactant the number of ways a particular combination of molecules can occur is just the number of molecules $\#S$ of reactant S . In this monomolecular reaction ($K_\mu = 1$) the reaction does actually not require collisions and c_μ is thus independent of the volume. For a bimolecular reaction, increasing the volume (with the number of moles constant), reduces the likelihood of reactive collisions in a small interval of time and hence $c_\mu \propto V^{-1}$. In general we can therefore write

$$c_\mu \propto 1/V^{K_\mu-1} . \quad (4.31)$$

Using the following example for a chemical reaction



which is split into two reaction channels



When a reaction occurs, the changes to molecule populations are

$$\nu_1 = (-1, -\alpha, \beta, 0) , \quad \nu_2 = (0, \alpha, -\beta, \gamma) .$$

From (4.32) and from the definition of reaction velocity we have the following *reaction rates*

$$\begin{aligned} \left[\begin{array}{c} [\dot{S}_1] \\ -1 \\ [\dot{S}_2] \\ -\alpha \\ [\dot{S}_3] \\ \beta \end{array} \right] &= k_1 [S_1] [S_2]^\alpha , \\ \left[\begin{array}{c} [\dot{S}_3] \\ -\beta \\ [\dot{S}_2] \\ \alpha \\ [\dot{S}_4] \\ \gamma \end{array} \right] &= k_2 [S_3]^\beta . \end{aligned} \quad (4.33)$$

The rate equations are then easily derived as

$$\left. \begin{aligned} d[S_1]/dt &= -k_1 [S_1] [S_2]^\alpha \\ d[S_2]/dt &= -\alpha k_1 [S_1] [S_2]^\alpha + \alpha k_2 [S_3]^\beta \\ d[S_3]/dt &= \beta k_1 [S_1] [S_2]^\alpha - \beta k_2 [S_3]^\beta \\ d[S_4]/dt &= \gamma k_2 [S_3]^\beta . \end{aligned} \right\} \quad (4.34)$$

Looking at the structure of (4.34), we recognize in this set of equations the GMA representation (2.19). Substituting $[S] = S/V = \langle \#S \rangle / (N_A V)$ in (2.19), gives

$$\frac{d}{dt} \left(\frac{\langle \#S_i \rangle}{N_A V} \right) = \sum_{\mu=1}^M \nu_{\mu i} k_\mu \prod_{j=1}^{L_\mu} \left(\frac{\langle \#S_{p(\mu,j)} \rangle}{N_A V} \right)^{l_{\mu j}} ,$$

which can be rewritten as

$$\frac{d}{dt} \langle \#S_i \rangle = \sum_{\mu=1}^M \frac{\nu_{\mu i} k_\mu}{(N_A V)^{K_\mu - 1}} \prod_{j=1}^{L_\mu} \langle \#S_{p(\mu,j)} \rangle^{l_{\mu j}} , \quad (4.35)$$

where we made use of the fact that

$$\prod_{j=1}^{L_\mu} (N_A V)^{l_{\mu j}} = (N_A V)^{\sum_{j=1}^{L_\mu} l_{\mu j}} = (N_A V)^{K_\mu} .$$

The differential operator is justified with the assumption of large numbers of molecules involved, such that near continuous changes are observed⁵. Let us now assert for the temporal evolution of $\langle \#S_i \rangle$ a “particle-ODE”, :

$$\frac{d}{dt} \langle \#S_i \rangle = \sum_{\mu=1}^M \nu_{\mu i} k'_\mu \prod_{j=1}^{L_\mu} \langle \#S_{p(\mu,j)} \rangle^{l_{\mu j}} . \quad (4.36)$$

Comparing (4.36) with (4.35), we find

$$k'_\mu = \frac{k_\mu}{(N_A V)^{K_\mu - 1}} , \quad (4.37)$$

⁵Strictly speaking it is possible to take the differential, for a Heaviside step function the derivative is a δ -function. This leads however to some complications which we do not wish to go into here.

This equation then describes the interpretation of the rate constant, dependent on whether we consider concentrations or counts of molecules.

To arrive at a general expression for the propensity a_μ from (4.36) it follows that

$$\langle \#R_\mu \rangle = k'_\mu \cdot \prod_{j=1}^{L_\mu} \langle \#S_{p(\mu,j)} \rangle^{l_{\mu j}} dt \quad (4.38)$$

gives the average number of R_μ reactions occurring in $(t, t + dt)$. Note that ν_μ has been excluded above since we would otherwise have an expression for the number of molecules not the number of reactions. Considering $\#R_\mu$, the number of R_μ reactions, as a discrete random variable with probability mass function $p_{r_\mu} = \text{Prob}\{\#R_\mu = r_\mu\}$. The expectation value $\langle \#R_\mu \rangle$ is given by

$$\langle \#R_\mu \rangle = \sum_{r_\mu} r_\mu \langle p_{r_\mu} \rangle \quad r_\mu = 0, 1, 2, \dots \quad (4.39)$$

where

$$p_{r_\mu} = \begin{cases} a_\mu dt + o(dt) & : r_\mu = 1 \\ 1 - a_\mu dt + o(dt) & : r_\mu = 0 \\ o(dt) & : r_\mu > 1 \end{cases} \quad (4.40)$$

where $o(dt)$ denotes a negligible probability for more than one R_μ reaction to occur during dt . Since a_μ is a function of n , p_{r_μ} is randomly varying and hence the averaging $\langle p_{r_\mu} \rangle$ over the ensemble in (4.39). Equation (4.39) thus becomes

$$\langle \#R_\mu \rangle = 0 \cdot p_0 + 1 \cdot p_1 + \sum_{r_\mu > 1} r_\mu \langle p_{r_\mu} \rangle .$$

From (4.39) and (4.40) we then have

$$\langle \#R_\mu \rangle = \langle a_\mu dt \rangle + o(dt) . \quad (4.41)$$

Now, from (4.38) and (4.41) the propensity of R_μ reactions to occur in dt is given as

$$\langle a_\mu \rangle = k'_\mu \prod_{j=1}^{L_\mu} \langle \#S_{p(\mu,j)} \rangle^{l_{\mu j}} . \quad (4.42)$$

To proceed, we consider another alternative expression for a_μ , by substituting (4.27) into

$$a_\mu = h_\mu \cdot c_\mu , \quad \mu = 1, \dots, M \quad (4.43)$$

and considering the average

$$\langle a_\mu \rangle = c_\mu \cdot \left\langle \frac{\prod_{j=1}^{L_\mu} (\#S_{p(\mu,j)})^{l_{\mu j}}}{\prod_{j=1}^{L_\mu} (l_{\mu j}!)} \right\rangle , \quad (4.44)$$

where $\#S_{p(\mu,j)}$ denotes the random variable whose value is $n_{p(\mu,j)}$. Note that this implied the assumption of a large number of molecules for all species S_i and $l_{\mu j} > 1$. Comparing (4.44) and (4.42)

$$k'_\mu \prod_{j=1}^{L_\mu} \langle \#S_{p(\mu,j)} \rangle^{l_{\mu j}} = \frac{c_\mu \left\langle \prod_{j=1}^{L_\mu} (\#S_{p(\mu,j)})^{l_{\mu j}} \right\rangle}{\prod_{j=1}^{L_\mu} (l_{\mu j}!)} .$$

Making the same notorious assumption⁶ of zero covariance as in [Gil92a], gives

$$k'_\mu = \frac{c_\mu}{\prod_{j=1}^{L_\mu} (l_{\mu j}!)} , \quad (4.45)$$

which can be turned into an expression for c_μ :

$$c_\mu = k'_\mu \cdot \prod_{j=1}^{L_\mu} (l_{\mu j}!) . \quad (4.46)$$

Inserting (4.37) for k'_μ , we arrive at

$$c_\mu = \left(\frac{k_\mu}{(N_A V)^{K_\mu - 1}} \right) \cdot \prod_{j=1}^{L_\mu} (l_{\mu j}!) . \quad (4.47)$$

Equation (4.47) establishes a relationship between the stochastic constant c_μ and rate constant k_μ and is used in most implementations of Gillespie's algorithm. Note that if above we substitute S/V in (2.19) for $[S]$ instead of $\langle \#S \rangle / (N_A V)$, the only difference to (4.37) and (4.47) is that the N_A would not appear in these equations.

The difference of our derivation to the one given by Gillespie in [Gil92a] is that we introduced the average number of reactions (4.38) to move from the general mass action representation (2.19), which is independent of particular examples, to an expression that allows us to derive parameter c_μ of the stochastic simulation (4.47) without making reference to the temporal evolution of moments from the CME. In [Gil92a], the temporal evolution of the mean is derived for examples of bi- and tri-molecular reactions only. Taking the variance of $\#S(t)$ to be zero to make it a 'sure variable', this equation is compared to the mass action model to derive c_μ .

Equation (4.47) is at the heart of the Gillespie algorithm and its implementations. There are two conclusions from the derivation. First, using the approximation (4.27) for h_μ is valid for a large number of molecules with $l_{\mu j} > 1$. Although in most practical cases this will not lead to significant differences, this has been ignored by some authors. More important however is the fact that the derivation of (4.47) relies on the rate constants of the mass action model. In this respect, it does not make sense to compare a mass action model and a stochastic simulation *as alternatives* if the stochastic rate constant c_μ is derived from the rate constants of the mass action model.

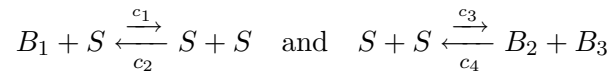
So how do we compare deterministic and stochastic models? First, we ought to compare models with models and simulations with simulations. The advantage of the mass action model (2.19) is that its terms and parameters have a precise meaning, they are a direct translation of the biochemical reaction diagrams that capture the biochemical relationships of the molecules involved. For a formal analysis of the model, as opposed to a simulation, rate equations are in virtually all cases simpler than the CME. One might argue that for any realistic pathway model a formal analysis is not feasible for either model and a simulation (numerical solution) is the way to go forward. In this case the Gillespie algorithm provides an efficient implementation to generate realizations of the CME. An advantage of simulations is furthermore that, in principle, it is possible to vary temperature and volume over time. One can imagine fluctuations in volume and temperature to be the

⁶The assumption of zero covariance such that $\langle \#S_i \#S_j \rangle = \langle \#S_i \rangle \langle \#S_j \rangle$ means for $i \neq j$ nullifying correlation, and for $i = j$ nullifying random fluctuations. The same assumption is required if one compares the temporal evolution of the mean of the CME model with the mass action model. This then demonstrates that a mass action model does *not* always arise as the mean of the CME model [Gil92a] (page 363). See also page 99 in Section 4.2.

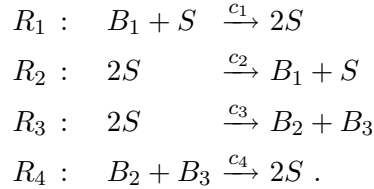
consequence of various physical processes in the cell. These variations may however also serve a ‘purpose’, in which case we ought to consider them explicitly in our mathematical models. If we are to consider random fluctuations in volume and temperature to matter, this would lead to random fluctuations of the rate constants.

4.4.3 So are they, or are they not?

We are now briefly returning to a question asked in Section 3. We would like to understand how the ODE representation and stochastic model compare and whether the ODE model does emerge as the mean of the stochastic process. To see how we can arrive at the law of mass action from stochastic modelling, consider the following set of reactions, also considered as an example by Gillespie in [Gil92a]



such that we get the following reaction channels R_μ :



We assume that B_1 , B_2 , and B_3 are buffered to allow for a univariate analysis of the process $S(t)$. We will for this part also simplify the notation and use $n_1 = \#B_1$, $n_2 = \#B_2$, $n_3 = \#B_3$, and $n_s = \#S$ to denote counts of molecules. The changes or ‘jumps’ in molecules in individual reaction channels R_1, \dots, R_4 are

$$\nu_1 = +1, \quad \nu_2 = -1, \quad \nu_3 = -2, \quad \nu_4 = +2$$

with propensities

$$a_1 = c_1 n_1 n_s, \quad a_2 = c_2 n_s (n_s - 1)/2, \quad a_3 = c_3 n_s (n_s - 1)/2, \quad a_4 = c_4 n_2 n_3 .$$

From the definition of $P_\mu(dt) = a_\mu dt$, we can write an expression for the mean change in dt :

$$\langle \#S(t + dt) - \#S(t) \rangle = \left\langle \underbrace{\sum_{\mu=1}^M \nu_\mu P_\mu(dt)}_{\text{average over channels}} \right\rangle = \left\langle \sum_{\mu=1}^M \nu_\mu a_\mu dt \right\rangle .$$

Note that in the right-hand term, a_μ is a random variable, which means it has a distribution and which therefore justifies the averaging $\langle \cdot \rangle$. This leads us to

$$\frac{d}{dt} \langle \#S(t) \rangle = \left\langle \sum_{\mu=1}^M \nu_\mu a_\mu \right\rangle \tag{4.48}$$

$$\begin{aligned} &= \langle (\nu_1 a_1 + \nu_2 a_2 + \nu_3 a_3 + \nu_4 a_4) \rangle \\ &= c_1 n_1 \langle \#S(t) \rangle - \left(\frac{c_2}{2} \right) \langle \#S(t)(\#S(t) - 1) \rangle - c_3 \langle \#S(t)(\#S(t) - 1) \rangle + 2c_4 n_2 n_3 . \end{aligned} \tag{4.49}$$

For this last equation, to be comparable to the ODE representation (4.36), it should only contain expression in $\langle \#S(t) \rangle$. For this to happen, we require two assumptions to be valid. The first one is

$$\langle \#S(t)(\#S(t) - 1) \rangle = \langle \#S^2(t) \rangle ,$$

which holds true for a reasonably large $\#S(t)$. The next step would require

$$\langle \#S^2(t) \rangle = \langle \#S(t) \rangle^2 .$$

Looking at the definition of the variance of $\#S(t)$

$$\text{Var}[\#S(t)] = \langle \#S^2(t) \rangle - \langle \#S(t) \rangle^2 ,$$

we can get there by assuming the variance of random variable $\#S(t)$ is zero. Compare this with the example in Section 4.2, where the same question, to whether the mass action representation is the mean or expectation of the stochastic model, was only sensible by assuming the dependent variables to be independent.

Gillespie's work, for which a wealth of details are provided in [Gil92a], provides us with a means to translate pathways, described in terms of mass-action laws, into a stochastic representation. However, it does also provide a simple alternative to the chemical master equations approach in stochastic modelling. In the following section we are going to discuss an algorithm that allows us to simulate fairly complex pathways in a quite effective manner using this framework.

4.4.4 The Gillespie Algorithm

Before Gillespie's work, the traditional method of calculating the time evolution of a system of chemically reacting molecules was to set up and solve the chemical master equation of a system under consideration. The behavior of an entire system of reactions is thus described by a set of differential equations with probabilities as variables. These equations are difficult to solve and instead Gillespie suggested a simulation of single trajectories of the system and then use statistics to estimate concentrations. The goal of stochastic simulation is then to describe the evolution of the state vector $S(t)$ from some given initial state $S(0)$. At each time step, the system is in exactly one state, and the two questions a simulation algorithm has to answer are:

- Which reaction will occur next?
- When does it occur?

Gillespie expresses the answer to questions above in form of the following 'reaction probability density function' $P(\mu, \tau)$ defined by

$$P(\mu, \tau)d\tau, \text{ the probability that reaction } R_\mu \text{ will occur in the infinitesimal time interval } (t + \tau, t + \tau + d\tau), \text{ given the system is in state } S(t) \text{ at time } t.$$

Note that the density function $P(\mu, \tau)$ is assumed to be proportional to $d\tau$ only. The probability $P(\mu, \tau)d\tau$ is calculated as

$$P(\mu, \tau)d\tau = P_0(\tau)P_\mu(d\tau) . \quad (4.50)$$

$P_\mu(d\tau) = a_\mu d\tau$ is the probability (4.29) that an R_μ reaction will occur in the time interval $(t + \tau, t + \tau + d\tau)$. $P_0(\tau)$ is the probability that, given state $S(t)$ at time t , no reaction will occur in the interval $(t, t + \tau)$. The probability that any of the M reactions occurs in the interval is

$$\sum_{\mu=1}^M a_\mu d\tau .$$

Let us simplify the notation by setting

$$a^* = \sum_{\mu=1}^M a_{\mu} .$$

The probability that no reactions occur in the interval $d\tau$ is thus $1 - a^*d\tau$ and therefore

$$P_0(\tau + d\tau) = P_0(\tau) \cdot (1 - a^*d\tau) ,$$

which can be rearranged into a differential equation

$$\frac{dP_0}{d\tau} = -a^*P_0 ,$$

for which the solution is given by

$$P_0(\tau) = e^{-a^*\tau} , \quad (4.51)$$

satisfying the initial condition $P_0(0) = 1$, i.e., the probability of nothing happening in zero time is one. Inserting (4.51) for $P_0(\tau)$ into (4.50)

$$P(\mu, \tau)d\tau = a_{\mu}e^{-a^*\tau}d\tau , \quad (4.52)$$

or equivalently

$$P(\mu, \tau) = a_{\mu}e^{-a^*\tau} .$$

where $\mu = 1, \dots, M$ and $0 \leq \tau < \infty$. This joint distribution can also be viewed as a μ -indexed family of continuous density functions

$$P_{\mu}(\tau) = h_{\mu}c_{\mu}e^{-a^*\tau} .$$

This is interpreted as the probability that at time t the reaction R_{μ} occurs in interval $(t+\tau, t+\tau+d\tau)$ and no reaction occurred in the previous interval. Fixing μ , the probability $P_{\mu}(\tau)$ follows an exponential distribution, beginning for $\tau = 0$ at height $h_{\mu}c_{\mu}$. Among all reactions, the most probable (or “first”) reaction μ is determined according to which distribution $P_{\mu}(\tau)$ has, in the interval $\tau + d\tau$, the largest area below the density function.

From the joint distribution one obtains the probability that the next reaction of any type will occur in the interval $(t + \tau, t + \tau + d\tau)$ by integrating (4.50) over all reactions:

$$P^{(1)}(\tau)d\tau = \sum_{\mu=1}^M P(\mu, \tau)d\tau = a^*e^{-a^*\tau}d\tau . \quad (4.53)$$

Note that equation (4.53) gives a probability, but for a time interval $(t + \tau, t + \tau + d\tau)$, where the exact point for the jumps in the molecule populations to occur is not specified and hence $d\tau$ is assumed to be negligible (see Figure 4.2). This will then give us a means to generate single-valued time point for the next reaction to occur. The probability that, given that a reaction occurs in τ , the reaction is of type μ is the conditional probability $P(\mu|\tau)$. From the definition of a conditional probability

$$P^{(2)}(\mu|\tau) = \frac{P(\mu, \tau)}{P(\tau)} = \frac{a_{\mu}e^{-a^*\tau}}{a^*e^{-a^*\tau}} = \frac{a_{\mu}}{a^*} .$$

We thus have two probabilities, $P^{(2)}$ and $P^{(1)}$, for *what* reaction is to occur next and *when* this it is to happen. For a simulation, this has to be turned into an algorithm. Let

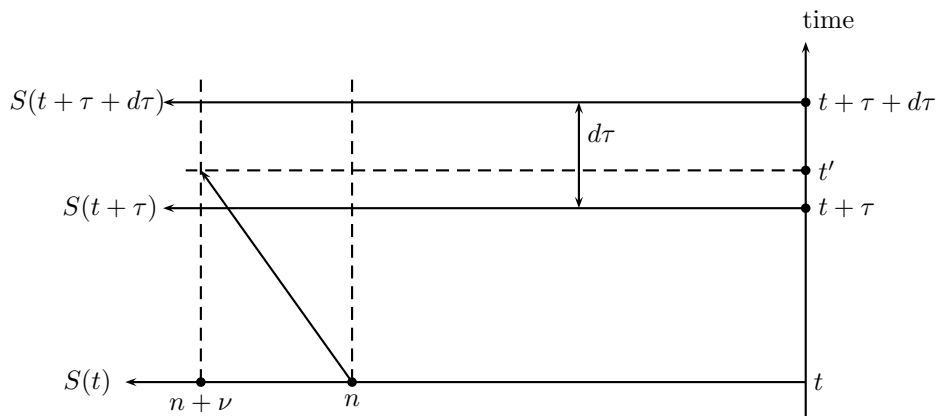


Fig. 4.2: Illustration for Equation (4.53).

us first determine the time to the next reaction. The probability density function (pdf) for this time was

$$P^{(1)}(\tau) = a^* e^{-a^* \tau}$$

with the corresponding cumulative distribution function defined by

$$F(t) = \int_{-\infty}^t P^{(1)}(\tau) d\tau = a^* \int_0^t e^{-a^* \tau} d\tau = 1 - e^{-a^* t} .$$

Suppose we use a random number generator to generate a number r_1 in the unit interval. If we choose a value t such that $F(t) = r_1$, the pdf of t will be the one for $P^{(1)}$. The random value of t can thus be obtained as

$$t = F^{-1}(r_1) = \frac{1}{a^*} \ln \left(\frac{1}{1 - r_1} \right) .$$

Since r_1 is uniformly distributed in the unit interval, so is $1 - r_1$ and we may replace $1 - r_1$ by r_1 to obtain a random value for the time τ to the next reaction:

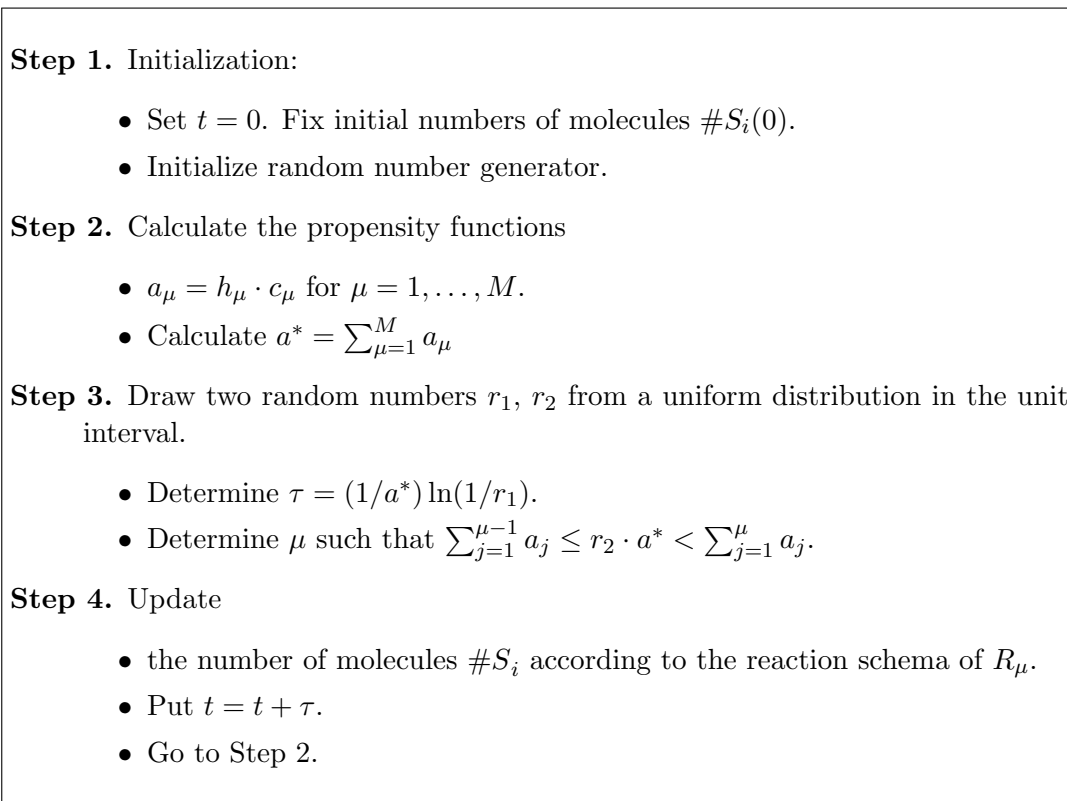
$$\tau = \frac{1}{a^*} \ln \left(\frac{1}{r_1} \right) = -\frac{1}{a^*} \ln r_1 .$$

Finally, the type of reaction is determined by generating a second random number r_2 in the unit interval. Then the type of reaction that occurs at time τ corresponds to that value of μ which satisfies the inequality

$$\sum_{j=1}^{\mu-1} \frac{a_j}{a^*} \leq r_2 < \sum_{j=1}^{\mu} \frac{a_j}{a^*} .$$

Gillespie introduced two equivalent simulation algorithms, called ‘Direct Method’ (see Fig.4.3) and ‘First Reaction Method’. These have been improved by others, including M.A.Gibson [GB00] who developed the ‘Next Reaction Method’. The differences between these algorithms have to do with the implementation and computational aspects, the basic idea how molecular interactions are represented is the same. Examples for stochastic simulations will be provided in subsequent sections.

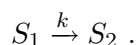
Fig. 4.3: Gillespie's Direct Methods for exact stochastic simulation.



4.4.5 Examples

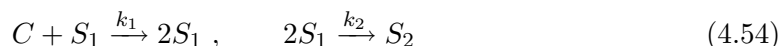
We here briefly recall a couple of examples for the Gillespie algorithm presented in [Gil77]. These are for illustration purposes only. We are going to discuss the roots of stochastic modelling and simulation further in the remaining sections of this chapter.

The trivial example of an irreversible isomerization, equivalent to reactive decay is the reaction



This has been solved analytically at the beginning of this chapter. Here the stochastic mean is indeed equal to the solution of the ODE model and as shown in Figure 4.5.

For the next example, consider a molecular species that remains constant, denoted by the letter C



The assumption of constancy can often be made for large pools, or at least for a certain period of time. The simulations in Figure 4.6 show that the system is stable for different initial conditions, i.e., beginning with only 10 molecules or beginning with 3000 molecules. In either case the average over two runs is near identical to the solution of the mass action model. Gillespie demonstrated with this example that stochastic simulation provides a rapid way of testing a hypotheses, which for analytical solutions can often be cumbersome. He disproved a claim from the literature that the stochastic model can become unstable. Interestingly Gillespie makes even an argument for saving money with his simulation technique. Thirty years later, for all of these examples it is not necessary to worry about computing time.


```

function [tt,nn,aa] = gillespie(pars)

% MATLAB (function) M-file implementing the Gillespie Algorithm.

n0 = pars{7}; % initial number of molecules.
nu = pars{9}; % changes in molecules.
tf = pars{10}; % end time for simulation.
rand('state',sum(100*clock)) % reset random number generator
t = 0; % initialize time to zero
n = n0; % initialize state
tt = 0; nn = n0; aa = propensity(n0,pars); while t < tf
    % Compute propensities for all reaction channels:
    a = propensity(n,pars);
    astr = sum(a); % sum propensities over reaction channels
    if ~astr
        t = tf;
    else
        tau = exprnd(inv(astr)); % time to next reaction
        % Next reaction Ru:
        mu = min( find( cumsum(a) > astr*rand ) );
        t = t + tau; % update t
        n = n + nu(mu,:); % update n
    end
    tt = [ tt; t ]; nn = [ nn; n ]; aa = [aa; a];
end tt(end) = tf; nn(end) = nn(end - 1); aa(end) = aa(end - 1);

function a = propensity(n,pars)

% Subfunction for the computation of propensities

c = pars{3}; M = pars{4}; p = pars{5}; l = pars{6}; for mu=1:M
    if n(p{mu})
        h(mu) = prod(matnchoosek(n(p{mu}),l{mu}));
    else
        h(mu) = 0;
    end
end
a = c.*h; % propensities

function c = matnchoosek(n,k)

% Subfunction that realizes a matrix version of nchoosek(.)

if any( [n k] <= 0 )
    error('matnchoosek takes positive inputs')
end m = n(:); k = k(:); M = length(m); for i = 1:M
    c(i) = nchoosek(m(i),k(i)) ;
end c = reshape(c,size(n));

```

Fig. 4.4: Matlab implementation of Gillespie's Direct Method for stochastic simulation of the chemical master equation.

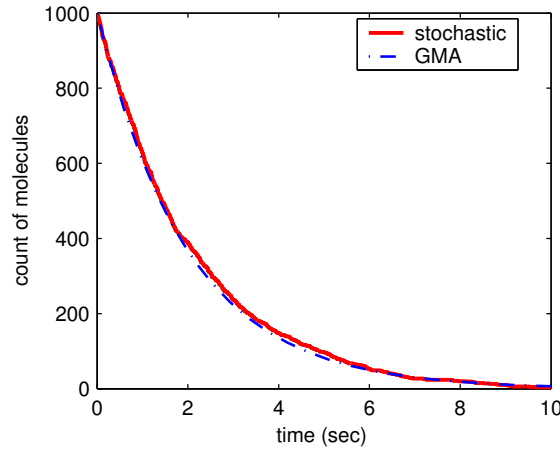


Fig. 4.5: Single run of a stochastic simulation (solid line) and solution of the mass action model for a simple irreversible isomerization reaction. The similarity is due to the fact that a large number of molecules are considered and because for a monomolecular reaction, the stochastic mean is exactly the same as the solution of the mass action model. Here $c = k = 0.5$.

The Volterra model of two nonlinear coupled differential equations,

$$\begin{aligned}\frac{dS_1}{dt} &= k_1 A S_1 - k_2 S_1 S_2 \\ \frac{dS_2}{dt} &= k_2 S_1 S_2 - k_3 S_2 ,\end{aligned}$$

is widely used for an introduction to nonlinear differential equations. In these equations S_1 is interpreted as the ‘prey’ and S_2 as the ‘predator’, A as the constant supply of food for the prey. This model can be interpreted as a set of coupled autocatalytic Lotka reactions



where the last reaction represents the ‘death rate’ of S_2 . The birth-death master equation for this system was already derived in Section 4.2. A steady state of the system is reached for

$$\frac{dS_1}{dt} = \frac{dS_2}{dt} = 0 ,$$

which is attained for

$$S_1 = \frac{k_3}{k_2} \quad \text{and} \quad S_2 = k_1 \frac{A}{k_2} .$$

For the simulation of these equations, we choose $k_1 A = 10$, $k_2 = 0.01$, $k_3 = 10$. For initial conditions, $S_1(0) = S_2(0) = A = 1000$, and an average of five runs, Figure 4.7 shows that for a stochastic simulation it may in principle be possible for the populations to get extinct. In the deterministic case, the system would describe well defined orbits around a fixed point and for small perturbations the system would recover and return to a stable oscillatory pattern. In the stochastic simulation, the amplitudes of the oscillations are quite variable, while the frequency and phase shift remain relatively stable.

While previous simulations assumed that food A changed insignificantly and thus could be assumed constant, we now consider the case in which $A(t)$ is depleted. For every time an R_1 reaction occurs, the population of A is reduced by one. As long as there is food, i.e., $A(t) > 1$, the prey can increase its population. However, as the food is depleted,

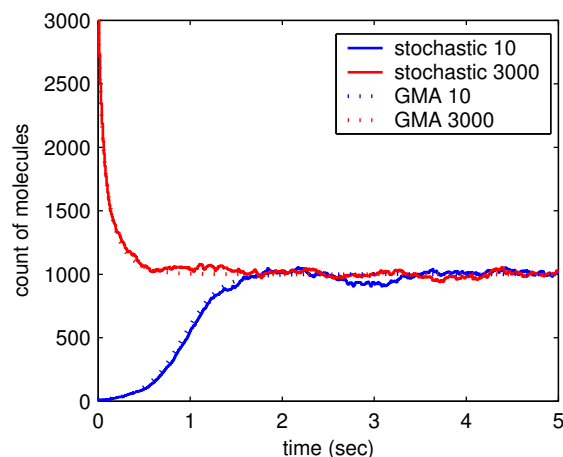


Fig. 4.6: The plot shows the average over two runs of the stochastic simulation for Example 4.54 with parameters $k_1 = 0.005, k_2 = 0.005$. The simulations show that the system is stable for different initial conditions, i.e., beginning with only 10 molecules or beginning with 3000 molecules. In either case the average over two runs is near identical to the solution of the mass action model.

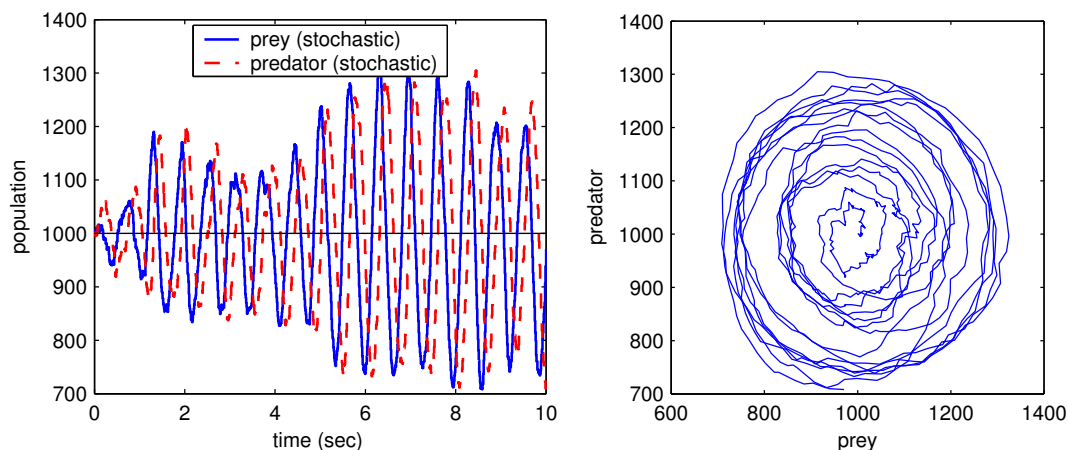
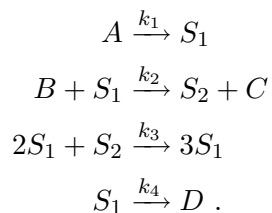


Fig. 4.7: Stochastic simulation for the Lotka-Volterra equations 4.55. Left: average of 5 runs for the temporal evolution of the populations of S_1 (“prey”) and S_2 (“predator”). Note how the predator lags behind the prey. The phase portrait is shown on the right.

even in the absence of predators the prey population will be diminished. In the absence of prey, the predators have the same fate. The previous situation in which there is an infinite supply of food is sometimes referred to an ‘open system’ in which there is an outside supply. To let A deplete would then correspond to a closure of the system. Figure 4.8 shows two simulations, in both cases the food is depleted, in the right plot the prey, S_1 , can also die of ‘natural causes’, which is realized through an additional reaction channel $R_4: S_1 \rightarrow B$.

The following set of reactions, demonstrate that it is possible to obtain a stable chemical oscillator, sometimes referred to as the *Brusselator*:



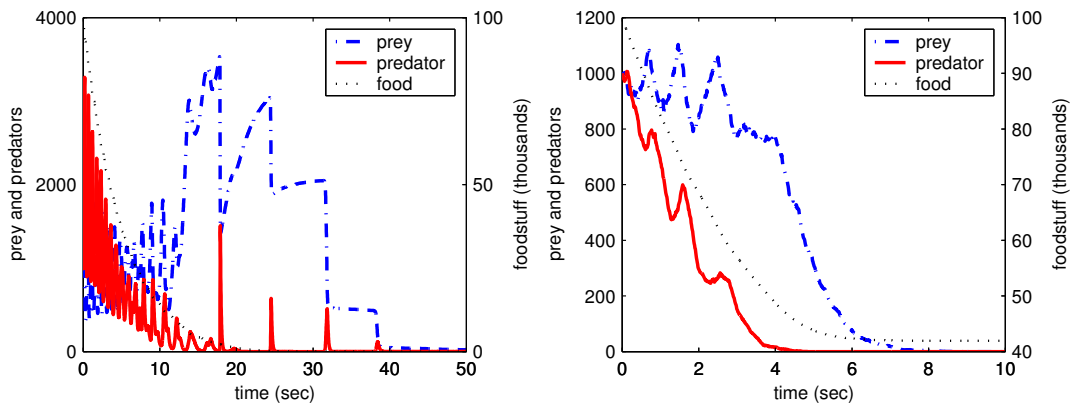


Fig. 4.8: Stochastic simulation of the Lotka-Volterra system in which the food supply is depleted. The initial condition for A is $A(0) = 100000$, $k_1 = 0.0002$, $k_2 = 0.01$, $k_3 = 10$. Right: the same system but now S_1 can also die of ‘natural causes’ at a rate of $k_4 = 10$. Shown are in all cases an average of five runs.

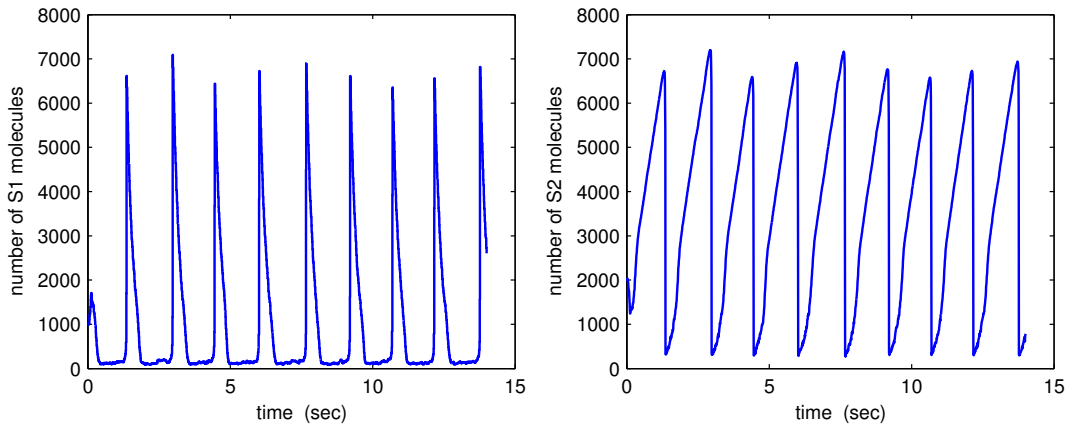


Fig. 4.9: Stochastic simulation of the Brusselator.

The corresponding rate equations are

$$\begin{aligned} \frac{d}{dt}S_1 &= k_1A - k_2BS_1 + \frac{k_3}{2}S_1^2S_2 - k_4S_1, \\ \frac{d}{dt}S_2 &= k_2BS_1 - \frac{k_3}{2}S_1^2S_2. \end{aligned}$$

The steady state values for $dS_1/dt = dS_2/dt = 0$ are k_1A/k_4 for S_1 and $2k_2Bk_4/k_1Ak_3$ for S_2 , respectively. Figure 4.9 and 4.10 show the stochastic simulation with $A = S_1(0)$, $B = S_2(0)$, $h_1 = S_1$, $h_2 = BS_1$, $h_3 = S_2S_1(S_1 - 1)/2$, and $h_4 = S_1$ and for $k_1A = 5000$, $k_2B = 50$, $k_3 = 0.00005$, $k_4 = 5$. Since for such large number of molecules a single realization is virtually identical to an average of realizations, the simulation was run only once. The initial conditions were set to $S_1(0) = 1000$ and $S_2(0) = 2000$. The two molecule populations clearly do not remain at their stationary values and instead immediately move into an oscillatory mode. Gillespie put the cost of these figures at 60 dollars. The phase diagram in Figure 4.10 shows that the system is unable to trace its route exactly in each cycle.

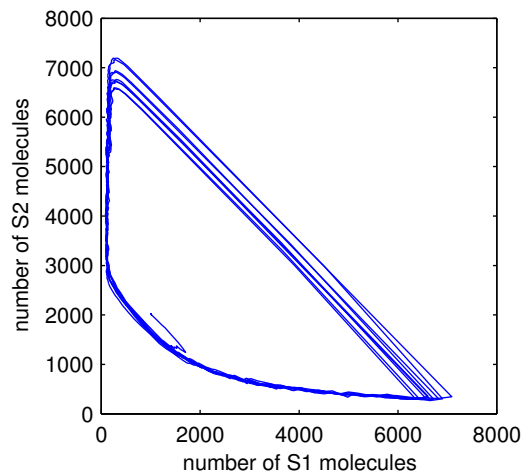


Fig. 4.10: Stochastic simulation of the Brusselator: phase diagram

The final example is referred to as the *Oregonator*, with the following set of reactions:

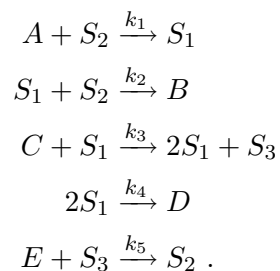


Figure 4.11 to 4.12 show the simulation results for $S_1(0) = 500$, $S_2(0) = 1000$, $S_3(0) = 2000$.

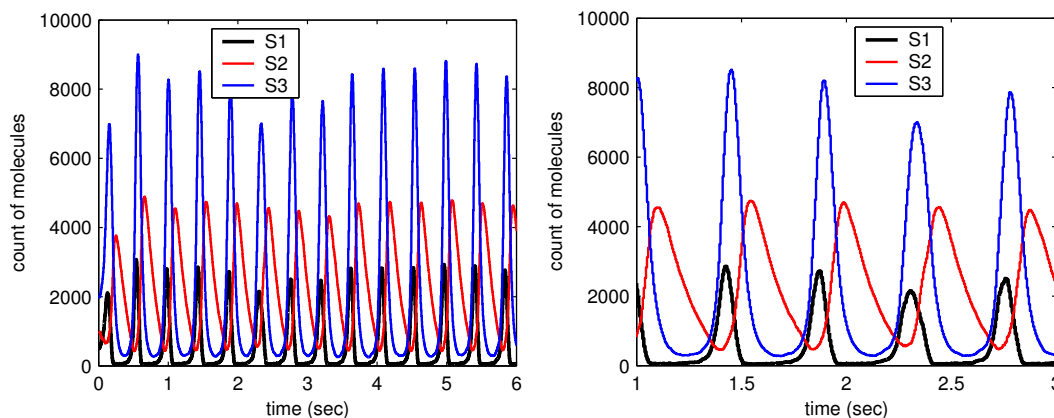


Fig. 4.11: Stochastic simulation of the Oregonator, showing the average of two runs for parameters $S_1(0) = 500$, $S_2(0) = 1000$, $S_3(0) = 2000$.

4.4.6 Molecules as individuals

In the previously discussed approaches, identical or similar molecules are grouped. Treating each molecule individually, it should be possible to have a more refined characterization of proteins (e.g. conformational changes, covalent modification, ligand binding, phosphorylation, methylation), allowing for a more detailed description of their catalytic activity,

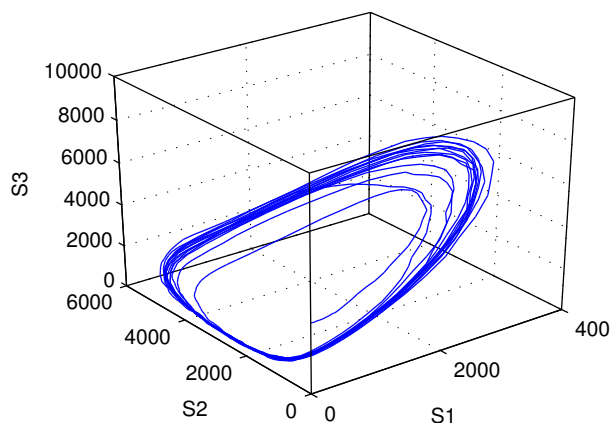


Fig. 4.12: Phase diagram for the stochastic simulation of the Oregonator.

binding affinity and so on. Each internal state of a protein may therefore lead to different reactive properties (and probabilities). Another advantage of a ‘multi-state molecular representation’ is that positional information could be added and the fate of an individual molecule could be traced.

In modelling groups of protein molecules we could introduce new variables for different internal states of the same protein.

In STOCHSIM [LS01] the enzyme kinetic reaction (4.23) would be simulated as follows. For a compartment containing four enzyme and four substrate molecules, and two product molecules, in the first step two molecules are chosen randomly. Say one S and one P , referring to a look-up table the probability for a reaction of these to molecules to occur is read off. If the probability exceeds a random number, the molecules will not react. If they do, a product is formed.

4.5 An ODE to Differential Equations

The stochastic models, derived in Section 3 were considering pools or populations of molecules and in deriving the model we made the assumption of a very small Δt . For decreasing numbers of molecules, and with a constant volume, fewer reactions will occur and a simulation of these stochastic models are bound to suffer accuracy. The basic assertion of the Gillespie stochastic simulation algorithm is that the probability for the reaction R_μ to occur within the small time interval $(t, t + dt)$ is equal to $a_\mu \cdot dt$ where a_μ is a factor that may depend on temperature, volume, current number of reactants. In principle, we could therefore vary a_μ during the simulation, something that would have been rather difficult for the stochastic models derived in Section 3. However, most commonly we take as the basis for the application of the Gillespie algorithm that within a compartment of fixed volume V we have a constant temperature, and moreover that all molecules have equal access to all others, they are well mixed and randomly distributed within the compartment space. A well-mixed system, means that molecules are homogeneously distributed in V . This allows us to determine the probability that the center of any randomly selected molecule will be found to lie inside a subregion of volume δV as equal to $\Delta V/V$. The position of a molecule is therefore uniformly distributed across the compartment of volume V . For this approach to make sense, one relies on the occurrence of many *non-reactive* molecular collisions to “stir” the system between successive *reactive* collisions. And because nonreactive collisions occur far more often, and thus the system can be described simply by the number of each kind of molecule. In other words, in the

Gillespie approach to stochastic simulation we do not describe or trace molecules individually but group molecular species S_i and describe them by their number of molecules $\#S_i$ in V . This has been occasionally used as an argument in favor of agent-based simulations. One should however note that for any realistic system there will always be pools of identical molecules or molecules in the same state and one way or another this will lead to aggregate descriptions.

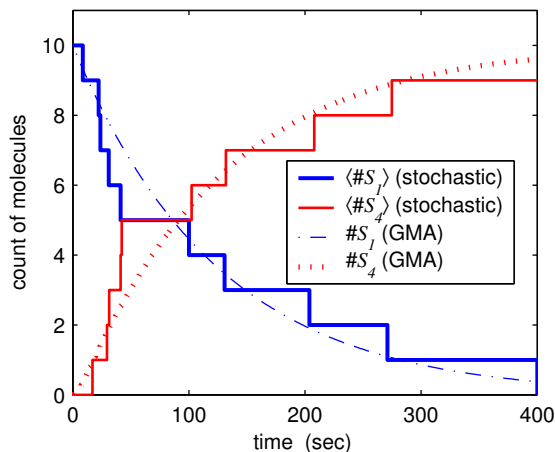


Fig. 4.13: Single run of a stochastic simulation using Gillespie’s Direct Method [Gil77] for Example (4.32) in the text. The parameters used are $V = 1$ pL, $k_1 = 0.5$ (nM · sec) $^{-1}$, $k_2 = 0.2$ sec $^{-1}$, $\alpha = 1$, $\beta = 1$, $\gamma = 1$, $\#S_4(0) = \#S_3(0) = 0$.

As the number of molecules becomes small, the variability of molecular populations in biochemical reaction networks increases. It is frequently argued that in this case differential equation models do not account for the observed variability and a stochastic approach should be preferred. To account for variability in chemical master equations (2.21) and rate equations (2.19), for *both* conceptual frameworks the identification of the model and its parameters requires a set of replicate experimental time series over which to average and estimate the moments of the distributions that account for the variability. While there are indeed good reasons to hypothesize stochastic mechanisms in intra-cellular dynamics (see [Pau04] for a recent overview and discussion), the arguments used for stochastic simulation and against differential equations are often misguided.

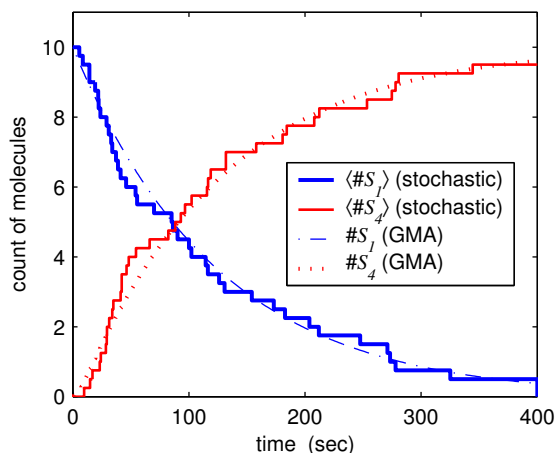


Fig. 4.14: Average over four realizations of a stochastic simulation of Example (4.32), using the same parameters as in Figure 4.13. The solutions of the mass action model have been multiplied by $N_A \cdot V \cdot 10^{-9}$ to convert concentrations into a count of molecules.

One ought to differentiate between a hypothesized principle or molecular mechanism

and the observations we make from experimental data. While rate equations are deterministic in the sense that they employ differential equations, they are based on a probabilistic description of molecular kinetics. On the other hand, the chemical master equation is a stochastic formulation, but based on differential equations, with probabilities as variables. The Gillespie algorithm realizes the chemical master equation but for most applications users of this algorithm make explicit use of the rate constants that define the generalized mass action model. In this situation it would then seem hypocritical to argue against mass action models as if it would their application be a great mistake. Note that this is not to mean that the CME approach relies on the mass action model, since to derive rather than postulate a rate equation, one must first postulate a stochastic mechanism from which the mass action model arises as a limit.

The Gillespie algorithm is an *exact* realization of a time-continuous Markov process. This is true regardless of how these processes are motivated physically, and one should not mistake this as an argument how well it can describe intracellular dynamics. Many of the assumptions regarding volume, temperature, homogenous culture etc., are not avoided by this fact. For many applications or examples in the literature the number of molecules is large enough to make a single stochastic realization appear to be an exact copy of the solution of the mass action model. In the context of a discussion of low molecular concentrations, one should not forget that, for fewer molecules, it is necessary to obtain a series of realizations over which we have to average before we get an ‘accurate’ of the properties of a stochastic system from either experimental data or simulations.

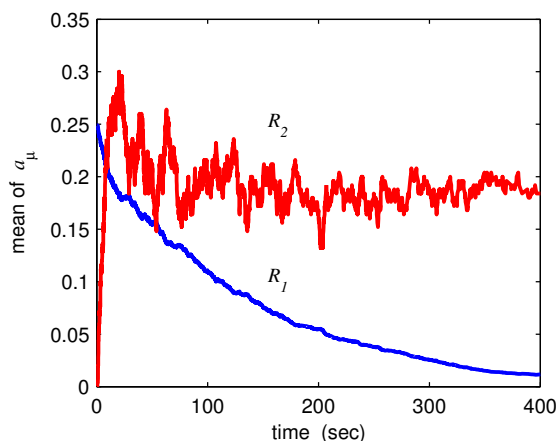


Fig. 4.15: Example (4.32). Temporal evolution of the a_μ . The parameters are the same as in Figure 4.13. What is shown is an average of the a_μ over realizations in order to illustrate the overall trend, free of the variability in single realizations.

A common argument is that if the concentration or the number of molecules of the chemical species involved in a biochemical reaction is low, a stochastic approach in form of the chemical master equation is a more accurate representation than rate equations [ELS01, vGK01, Kie02, XW03]. In case of [vGK01, Kie02] and [XW03] this discussion is not done on the basis of the chemical master equation but using the Gillespie algorithm for stochastic simulation. A question is what is meant by “low concentrations” or the consequences of small numbers of molecules? In [vGK01] a figure of the order of less than a few thousand is given. In [ELS01] the copy number of proteins is cited as less than a hundred. Since the number of molecules for most reactant species reduces either to very small values or increases steadily for others, we assume that authors, speaking of ‘numbers of molecules’ refer to initial numbers at time zero. Subject to approximation (4.27), Figures 4.13 and 4.14 compare realizations of the stochastic simulation of Example (4.32) with solutions of the rate equations. Figure 4.15 shows the temporal evolution of a_μ for a volume of 1 pL and initial numbers of 10 molecules. The simulations demonstrate that even for

very small numbers of molecules single realizations of stochastic simulations show steadily changing patterns that can be modelled well using a continuous representation. The close similarity between the numerical solution of the ODEs and the stochastic simulation is no surprise since the rate constants of the mass action model are also integral part of the stochastic simulation, as shown by equation (4.47).

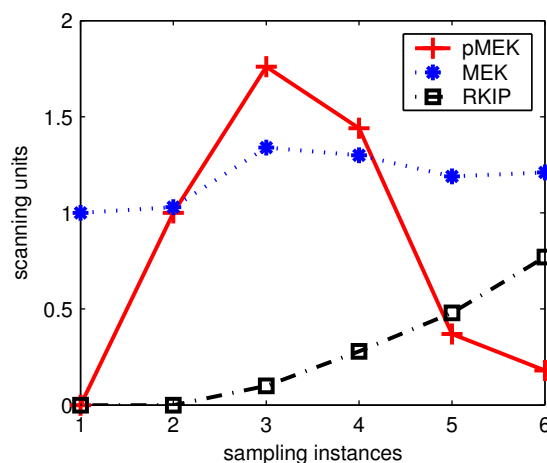


Fig. 4.16: Data from western blot time course experiments.

In fact, plots shown in those publications that argue for stochastic simulation in case of small molecule populations, are almost always displaying steady increases and decreases that are well approximated by ordinary differential equations. Figure 4.16 shows typical experimental data as obtained from western blot time course experiments to study proteins in signal transduction pathways. While there surely is measurement noise, it seems a reasonable assumption to believe the concentration profiles follow roughly the linear interpolations shown. If for the few time points that current experimental practice generates, we were not observing steadily increasing or decreasing pattern, and instead would argue for a truly random process, we would have a hard time to validate such a model from data like those shown in Figure 4.16. Figure 4.17 shows random simulations of time series with only six time points. How do we distinguish between random from deterministic pattern in data?

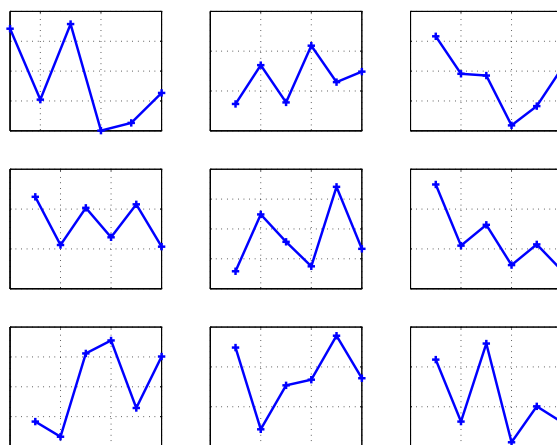


Fig. 4.17: Random simulations of time series. We can only start modelling by assuming that these curves are not random. If they are, we could not test this since there are not enough data points for a statistical test to be applicable. We call this the WYSIWYM principle: What You See Is What You Model!

Western-blot time series, like those shown in Figure 4.16, are generated from a pool

of about 10^7 cells although we are trying to understand what is happening in *a* cell. We could explain the deterministic pattern in experimental data as follows. Looking at the population of molecules of species S_i , from each reaction channel a change $\nu_{\mu i}$ arises for when the reaction channel R_μ is realized or active. The change $\nu_{\mu i}$ is a random variable, and the total change of S_i across all reaction channels is a sum of random variables

$$\Delta(\#S_i) = \sum_{\mu=1}^M \nu_{\mu i} .$$

For more than one reaction channel, from the central limit theorem, $\Delta(\#S_i)$ is approximately normal distributed, $\Delta(\#S_i) \sim \mathcal{N}(\cdot, \sigma_v^2)$. For any further averaging process with say m elements, e.g., using 10^7 cells in immunoblotting, the variance of measurements is of the order σ_v^2/m . This means that if we are not considering single-cell measurements we are likely to observe relatively smooth patterns. If we do consider single-cell measurements, we ought to have in any case replicates to average out random variations.

If we are to consider a stochastic simulation and wish to validate it with experimental data, we get the following requirements for the experimenters. In Gillespie's algorithm, the time interval for the next reaction to occur is calculated as

$$\tau = (1/a^*) \cdot \ln(1/r_1) ,$$

where r_1 is a random number in the unit interval and

$$a^* = \sum_{\mu}^M a_{\mu} . \quad (4.56)$$

Note that τ is a function of state n and thus implicitly also a function of time. As $\#S_i$ goes down, there are fewer reactive collisions and the propensity a_{μ} decreases. This means that for all relevant reaction channels, (4.56) will also decrease. This does however mean that the ratio a_{μ}/a^* changes little. Since the probability of the next reaction occurring is given by [Gil76]

$$P(\mu|\tau) = a_{\mu}/a^* , \quad (4.57)$$

this means that the resulting concentration levels are relatively similar. However, since τ , i.e., the time for the next reaction to occur is exponentially distributed,

$$P(\tau) = a^* \cdot \exp(-a^* \tau) , \quad (4.58)$$

with mean $1/a^*$ and standard deviation $1/a^*$, the variance of τ increases more substantially. This in turn means, that for a specified t the variance of the realizations will increase. Figure 4.18 illustrates the dependence of the variability on the initial number of molecules. A consequence is that for fewer molecules, more realizations are required to obtain an accurate picture through averaging across realizations. Also, the larger the number of reaction channels, M , the smaller is the average time to the next reaction τ , as shown by (4.58). However, at the same time, the number of possible transitions from state n will increase as can be seen from (4.57).

In considering mathematical modelling and simulation, most important are the context and the purpose of modelling. Do we wish to use a model to hypothesize a fundamental molecular mechanism, or are we trying to model the observed consequences of these molecular mechanisms? Is the phenomena we observe an aggregate of a group of dependent subcomponents (e.g. molecules or cells) that combine individual, discrete responses into graded response at higher levels of organization (e.g. in tissue and organs)?

In some cases, authors who wished to argue for their use of stochastic simulation, have unfortunately missed some of the subtleties of our foregoing discussion. Let us look at some examples. In [Kie02] it is argued that

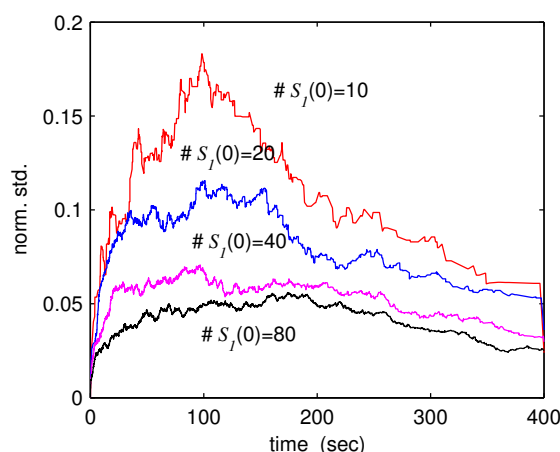


Fig. 4.18: Example (4.32). Temporal evolution of the normalized standard deviation $\sigma_{S_i}(t)/(\#S_1(0))$ over 50 realizations at t . $\alpha = \beta = \gamma = 1$, $k_1 = 0.6 \text{ (nM} \cdot \text{sec)}^{-1}$, $k_2 = 0.1 \text{ sec}^{-1}$. $\#S_4(0) = 10$, $\#S_2(0) = \#S_3(0) = 0$, $\#S_1(0) = 10, 20, 40, 80$. Note that the normalization is necessary to make the plots independent of the initial $\#S_i$ and thereby make them comparable.

“The availability of a huge amount of molecular data concerning various biochemical reactions provoked numerous attempts to study the dynamics of cellular processes by means of kinetic models and computer simulations.”

To take western-blot time course experiments as an example, the problem we face for modelling is anything but one of dealing with huge amounts of data. For a time series, usually only six to ten time points are available and replicates are the subject of hard fought negotiations between theoreticians and experimentalists. For realistic pathway models, because of the costs and time required, usually only a fraction of all pathway proteins can be covered by experiments. While the area of bioinformatics is often associated with a flood of data, in systems biology the lack of sufficiently rich quantitative stimulus-response time course data sets remains a problem.

The authors of [XW03] clearly missed the mark:

“There is also a problem of interpretation by users. Systems of differential equations have a number of parameters that must be fitted from experimental data. However, the parameters may have no meaning to the biologists, who are therefore unable to gauge whether the values are appropriate.”

Quite the opposite is true. The parameters of mass action model (2.19) have a very precise meaning, which can be fitted from experimental data. We would argue the fact that, for mass action models, we can identify parameters directly from experimental data is an advantage. Although this is not a trivial task, there are well established algorithms available for this purpose. Why would the authors of [XW03] think the CME (2.21) is more intuitive, and how would they validate their models?

Whether one starts with the mass action representation or the CME, it is often not possible to obtain all necessary parameter values from experimental data. For such practical reasons but also in order to simplify the mathematical analysis it is often desirable to make use of the quasi-steady-state assumption (QSSA) [SM97, Seg84]. The QSSA implies that for the time scale of interest the instantaneous rates of change of intermediate species are approximately equal to zero. Modelling signal transduction pathways, the consecutive activation of kinases is commonly described through phosphorylation and dephosphorylation steps, equivalent to enzyme kinetic reactions. Assuming the concentration of kinase-substrate complexes is small compared with the total concentration of the

reactants, phosphorylation is modelled as a bimolecular reaction and assuming that the concentration of active phosphatase is constant, dephosphorylation can be modelled as a first order reaction. Such assumptions allow a formal analysis of various important aspects of cell signalling rooted in mass action models. See [HNR02] for an outstanding example of such an analysis. These simplifications do of course also simplify the stochastic simulation since the k 's of the rate constants are implicitly used in the simulation. Alternatively, one considers the QSSA for the CME as discussed in [RA03].

A natural system may be inherently deterministic or random, whether we are going to use a stochastic model or a deterministic one is depending on various factors. To identify differential equations as a deterministic model is only partly true since a stochastic model (derived from the chemical master equation) is also a set of differential equations (with probabilities as variables). How easy it is to fall into a trap, is illustrated by the textbook example of a queue, e.g., the arrival of customers at a check-out desk. Every textbook introduces this as a prototypical example of a random process, modelled by a Poisson stochastic process. This is usually the right way to look at this problem and no-one would think of modelling this with differential equations. But like in studying biological systems, it depends on what (and how) we are looking at this process - under similar conditions a day at the check-out counter will show an expected (predictable) peak at lunch time and shortly after office hours, with longer intervals between arrivals. A Poisson process comes with an assumption about the constancy of the rate at which people arrive and it is the modeller's responsibility to ensure that this assumption is valid in the given context. Constant volume, temperature, rapid diffusion etc. were all assumptions we made in the previous section. Furthermore, in the modelling framework, the Markov assumption is justified for hard spheres colliding, i.e., collision time is minimal such that collisions are independent. None of these assumptions necessarily leads to a wrong model, and it would seem that systems biology is the art of making the right assumptions.

Systems of nonlinear ODEs can be numerically difficult to solve, for example, if there are widely ranging reaction rates in a pathway. It is also true that even for simple pathways we cannot expect to obtain a closed form solution⁷ for the set of ODEs that represent the system. With a wealth of knowledge about differential equations, there may at least be a chance for a study of general properties of the system without knowing its parameters exactly. There is scope for further research into phase plane and bifurcation analysis for pathway representations. Such analytic results are particularly important in light of the difficulties one has in obtaining sufficiently rich and accurate experimental data sets for a particular system under consideration.

The question of whether one should use ordinary differential equations or stochastic processes, raises interesting questions regarding the purpose of modelling. The principal aim of (mathematical) modelling is to represent available knowledge about system variables. The process of model building is then an essential part of the 'understanding' of the real world. The model is used to validate and generate hypotheses at the same time. This is done by 'prediction', in a sense that the model shows a behavior that either confirms available knowledge or the mismatch with observations suggests the need to revise the existing understanding.

In engineering, models that are composed of differential equations are often referred to 'physical models' since the equations are a direct representation of the physical laws that characterize the process under consideration. The structure of the model (the variables involved and how their are related) is then often directly interpretable in terms of the physical properties of the system. For complex technical systems, a common approach is to identify so called 'black-box models' from experimental data. System identification, a

⁷A *closed* solution of an ordinary differential equation is an expression that is composed of a finite number of "elementary" functions.

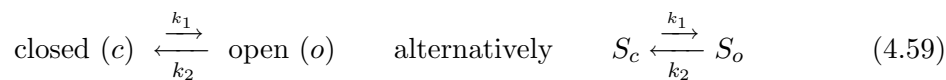
well established discipline within control engineering, and time series analysis provide a vast collection of tools for this purpose. In this context, a model is validated by means of prediction error, the aggregated difference between predicted and real output values. The internal structure of the model and its interpretation is usually of less interest. In systems biology we are aiming for a representation of ‘biological reality’. However, while aiming for physical or chemical realism in modelling, our choice for a mathematical representation should be guided by experimental considerations (what we can measure) and the need to limit the complexity of models, simulations or simply to reduce the effort in building a model.

Mathematical models deal with concepts, they are not direct reflections of reality. The model is constructed as to correspond in some useful sense to the real world, whatever that may mean. This correspondence is approximate, and there is no conflict between a deterministic model and a stochastic representation if we agree that mathematical models are not concerned with what is true but only with what can be observed.

4.6 A never ending story

In this section we continue the discussion of a suitable conceptual framework in which represent intracellular dynamics. We focus on a simple process that can have various practical interpretations. The development of the stochastic model and a study of its properties gives us an opportunity to learn more useful mathematical modelling techniques.

We are considering a molecule or molecular component like a channel or gate, which channeling has a switch-like function, i.e., changes between two states, say ‘open’ and ‘closed’. Alternatively, we may think of an isomerization reaction in which a molecule changes its isomerization reaction confirmation



Focussing initially on one channel or molecule, we write $P_o(t)$ for the probability that the channel is open at time t or the molecule is in state o , respectively. Let us assume the channel is closed at time t , the probability that it opens during the interval $(t, t + \Delta t)$ is defined as $P_o(t + \Delta t) = k_1 \Delta t$. Multiplying this probability with $P_c(t)$ gives us the probability that the transition $c \rightarrow o$ will actually occur:

$$\text{Prob}\{c \rightarrow o\} = k_1 \Delta t P_c(t) .$$

We can then summarize our knowledge of the behavior as follows

$$\begin{aligned} P_o(t + \Delta t) &= P_o(t) + k_1 \Delta t P_c(t) - k_2 \Delta t P_o(t) , \\ P_c(t + \Delta t) &= P_c(t) - k_1 \Delta t P_c(t) + k_2 \Delta t P_o(t) . \end{aligned}$$

Moving $P_o(t)$ and $P_c(t)$ to the left-hand-side and dividing by Δt we get

$$\begin{aligned} \frac{P_o(t + \Delta t) - P_o(t)}{\Delta t} &= k_1 P_c(t) - k_2 P_o(t) , \\ \frac{P_c(t + \Delta t) - P_c(t)}{\Delta t} &= -k_1 P_c(t) + k_2 P_o(t) . \end{aligned}$$

Since the system can only be in either state, $P_c(t) + P_o(t) = 1$ and we thus can describe the system with one equation

$$\frac{dP_o(t)}{dt} = k_1 P_c(t) - k_2 (1 - P_o(t)) .$$

Identifying the system as a two-state Markov process, *transition probability matrix* is given by

$$\Pi = \begin{bmatrix} \text{Prob}\{c \rightarrow c\} & \text{Prob}\{o \rightarrow c\} \\ \text{Prob}\{c \rightarrow o\} & \text{Prob}\{o \rightarrow o\} \end{bmatrix} = \begin{bmatrix} 1 - k_1\Delta t & k_2\Delta t \\ k_1\Delta t & 1 - k_2\Delta t \end{bmatrix} .$$

such that

$$P_S(t + \Delta t) = \Pi P_S(t) .$$

birth-death Markov process This birth-death process is a homogenous continuous-time, discrete-state Markov process. It is homogenous w.r.t. time, since the transition probabilities do not explicitly depend on time, although they are implicitly depending on time through the number of molecules.

Next we extend this model of one channel or molecule to a collection of n independent molecules and to simplify notation, let $P_c(t)$ and $P_o(t)$ denote the probability of $\#S_c \doteq c$ molecules in state ‘closed’ and $\#S_o \doteq o$ molecules in state ‘open’. Or, in terms of populations, $P_c(t)$, denotes is the probability of c molecules of species S_c and $P_o(t)$ gives the probability of o molecules of species S_o . Since n is given, we can focus on either o or c and derive the other from the fact that the total population is given by $n_T = c + o$. Considering the time interval $(t, t + \Delta t)$, where Δt is small enough to allow only one molecule to make the transition $o \rightarrow c$, there are four events that can determine, $P_c(t)$, the probability of c molecules of type S_c during $(t, t + \Delta t)$:

$$\begin{aligned} \text{Prob}\{c \rightarrow c - 1\} &= k_1 c \Delta t P_c(t) , \\ \text{Prob}\{c + 1 \rightarrow c\} &= k_1 (c + 1) \Delta t P_{c+1}(t) , \\ \text{Prob}\{c \rightarrow c + 1\} &= k_2 (n_T - c) \Delta t P_c(t) , \\ \text{Prob}\{c - 1 \rightarrow c\} &= k_2 (n_T - c + 1) \Delta t P_{c-1}(t) . \end{aligned}$$

Combining these equations into one we obtain

$$P_c(t + \Delta t) = P_c(t) + [k_2(n_T - c + 1)P_{c-1}(t) - k_2(n_T - c)P_c(t) + k_1(c + 1)P_{c+1}(t) - k_1cP_c(t)]\Delta t .$$

Moving $P_c(t)$ to the left-hand-side of the equation, dividing by Δt and taking the limit $\Delta t \rightarrow 0$, we obtain the master equation of S_c , for $0 \leq c \leq n_T$:

$$\frac{d}{dt}P_c(t) = k_2(n_T - c + 1)P_{c-1}(t) - k_2(n_T - c)P_c(t) + k_1(c + 1)P_{c+1}(t) - k_1cP_c(t) . \quad (4.60)$$

These are $n_T + 1$ differential equations and their solution would give us the temporal evolution of the probability of c molecules S_c . For S_o the master equation is

$$\frac{d}{dt}P_o(t) = k_1(n_T - o + 1)P_{o-1}(t) - k_1(n_T - o)P_o(t) + k_2(o + 1)P_{o+1}(t) - k_2oP_o(t) . \quad (4.61)$$

Although the master equation for this system are relatively simple, the explicit derivation of the solution to the differential equations (4.60) respectively (4.61) requires some effort⁸. We therefore first look at the steady state distributions, i.e., solve the equations for $\frac{d}{dt}P_c(t) = 0$. On the way we learn about the concept of probability generating functions.

4.6.1 Steady-state solution for the master equation

For $\frac{d}{dt}P_c(t) = 0$, we can simplify the notation by ignoring t and thus obtain a linear difference equation in one-variable for the isomerization reaction (4.59):

$$0 = k_1(c + 1)P(c + 1) - (k_1 - k_2)cP(c) - k_2n_T P(c) + k_2n_T P(c - 1) - k_2(c - 1)P(c - 1) . \quad (4.62)$$

⁸Well, actually using a computer and a mathematical software package, it is rather simple solve these equations numerically or analytically. We here prefer the ‘pen-and-paper’ approach as it will help us gain skills in manipulating equations.

Looking at the right-hand side we notice a structure with four cases

$$(c+1)P(c+1), \quad cP(c), \quad P(c), \quad (c-1)P(c-1).$$

To solve the difference equation (4.62) we introduce the *probability generating function* (pgf) probability generating function

$$\boxed{P(n) \xrightarrow{\text{pgf}} \mathfrak{P}(z) = \sum_{n=0}^{n_T} z^n P(n)} \quad (4.63)$$

where $z < 1$, and n_T denotes the total number of molecules,

$$\mathfrak{P}(1) = \sum_{n=0}^{n_T} P(n) = 1.$$

The rule (4.63) transforms $P(i)$ into $\mathfrak{P}(z)$ and the aim is to transform (4.62) entirely into the ‘ z -domain’ and thereby help solving it. In the engineering sciences similar transformations are a common tool to solve differential and difference equations. Next, to consider $P(n-1)$ let $n = i-1$ such that

$$\begin{aligned} \mathfrak{P}(z) &= \sum_{i=1}^{n_T+1} z^{i-1} P(i-1) = z^{-1} \left[\sum_{i=1}^{n_T+1} z^i P(i-1) \right] \\ &= z^{-1} \left[\sum_{i=0}^{n_T} z^i P(i-1) + z^{n_T+1} P(n_T) \right], \end{aligned}$$

where $P(-1)$, and the last term on the right corrects for the sum going only to n_T . We can rewrite this to obtain

$$z\mathfrak{P}(z) - z^{n_T+1}P(n_T) = \sum_{i=0}^{n_T} z^i P(i-1)$$

which implies the transformation rule

$$\boxed{P(i-1) \xrightarrow{\text{pgf}} z\mathfrak{P}(z) - z^{n_T+1}P(n_T)} \quad (4.64)$$

The next case to consider is $P(i+1)$, for which we let $n = i+1$ such that

$$\begin{aligned} \mathfrak{P}(z) &= \sum_{i=-1}^{n_T+1} z^{i+1} P(i+1) = P(0) + z \left[\sum_{i=1}^{n_T} z^i P(i+1) - z^{n_T} P(n_T+1) \right], \\ &= P(0) + z \left[\sum_{i=0}^{n_T} z^i P(i+1) \right], \quad \text{since } P(n_T+1) = 0 \end{aligned}$$

and from which we get the transformation rule

$$\boxed{P(n+1) \xrightarrow{\text{pgf}} \frac{1}{z} [\mathfrak{P}(z) - P(0)]} \quad (4.65)$$

Taking the first derivative of $\mathfrak{P}(z)$,

$$\frac{d\mathfrak{P}(z)}{dz} = \sum_{n=0}^{n_T} n z^{n-1} P(n) = z^{-1} \sum_{n=0}^{n_T} n P(n) z^n$$

from which get the rule

$$\boxed{nP(n) \xrightarrow{\text{pgf}} z \frac{d\mathfrak{P}(z)}{dz}} \quad (4.66)$$

From (4.64) and (4.66) we have

$$\boxed{(n-1)P(n-1) \xrightarrow{\text{pgf}} z^2 \frac{d\mathfrak{P}(z)}{dz} - z^{n_T+1} n_T P(n_T)} \quad (4.67)$$

From (4.65) and (4.67)

$$\boxed{(n+1)P(n+1) \xrightarrow{\text{pgf}} \frac{1}{z} \left[z \frac{d\mathfrak{P}(z)}{dz} - 0 \right] = \frac{d\mathfrak{P}(z)}{dz}} \quad (4.68)$$

If we now return to the steady-state equation (4.62) of our isomerization reaction example and apply the transformation rules we obtain

$$\begin{aligned} k_1 \frac{d\mathfrak{P}(z)}{dz} - (k_1 - k_2) z \frac{d\mathfrak{P}(z)}{dz} - k_2 n_T \mathfrak{P}(z) + k_2 n_T [z \mathfrak{P}(z) - z^{n_T+1} P(n)] \\ - k_2 \left[z^2 \frac{d\mathfrak{P}(z)}{dz} - z^{n_T+1} n_T P(n_T) \right] = 0 \end{aligned} \quad (4.69)$$

which reduces to

$$[k_1(1-z) + k_2 z(1-z)] \frac{d\mathfrak{P}(z)}{dz} - k_2 n_T (1-z) \mathfrak{P}(z) = 0$$

dividing both sides by $(1-z)$ and denoting $d\mathfrak{P}/dz = \mathfrak{P}'$, this leads to

$$\frac{\mathfrak{P}(z)'}{\mathfrak{P}(z)} = \frac{k_2 n_T}{k_2 z + k_1}.$$

From integration we know

$$\int \frac{f'(x)}{f(x)} dx = \ln |f(x)| + \alpha$$

where α denotes some constant. Applied to our expression this means that

$$\ln(\mathfrak{P}(z)) = n_T \ln(k_2 z + k_1) + \ln(\alpha) = \ln(\alpha (k_2 z + k_1)^{n_T}),$$

which implies

$$\mathfrak{P}(z) = \alpha (k_2 z + k_1)^{n_T}. \quad (4.70)$$

From (4.63), $\mathfrak{P}(z=1) = 1$, which gives us

$$\alpha = \frac{1}{(k_2 + k_1)^{n_T}},$$

inserted into (4.70) leads to the pgf transform of the steady-state equation of the isomerization reaction (4.59):

$$\mathfrak{P}(z) = \frac{1}{(k_2 + k_1)^{n_T}} \cdot (k_2 z + k_1)^{n_T} \quad (4.71)$$

$$\begin{aligned} &= \frac{1}{(k_2 + k_1)^{n_T}} \sum_{i=0}^{n_T} \binom{n_T}{i} (k_2 z)^{n_T-i} k_1^i \\ &= \frac{1}{(k_2 + k_1)^{n_T}} \sum_{i=0}^{n_T} \binom{n_T}{j} k_1^{n_T-i} (k_2 z)^i. \end{aligned} \quad (4.72)$$

Comparing this with the definition of the pgf, we obtain for the solution of (4.62) the steady state probability distribution of S_c as

$$P^\infty(c) = \frac{\binom{n_T}{c} k_1^{n_T-c} k_2^c}{(k_2 + k_1)^{n_T}} . \quad (4.73)$$

Likewise, the steady-state distribution for S_o is

$$P^\infty(o) = \frac{\binom{n_T}{o} k_2^{n_T-c} k_1^o}{(k_2 + k_1)^{n_T}} . \quad (4.74)$$

Looking at this discrete probability distribution, it looks very similar to a *binomial distribution*

binomial distribution

$$P(i) = \binom{n_T}{i} p^i (1-p)^{n_T-i} .$$

Indeed, we can show that (4.74), w.r.t. S_o , is indeed a binomial probability distribution with

$$p = \frac{k_1}{k_1 + k_2}$$

where p denotes the probability of the transition $o \rightarrow o + 1$:

$$\begin{aligned} \binom{n_T}{i} \frac{k_1^i}{(k_1 + k_2)^{n_T}} \cdot k_2^{n_T-i} &= \binom{n_T}{i} \frac{k_1^i}{(k_1 + k_2)^{n_T}} \cdot \frac{k_2^{n_T}}{k_2^i} \\ &= \binom{n_T}{i} \frac{k_1^i}{(k_1 + k_2)^{n_T}} \cdot \frac{k_2^{n_T}}{(k_1 + k_2)^i} \cdot \frac{(k_1 + k_2)^i}{k_2^i} \\ &= \binom{n_T}{i} \frac{k_1^i}{(k_1 + k_2)^{n_T}} \cdot \frac{k_2^{n_T}}{(k_1 + k_2)^i} \cdot \left(\frac{k_2}{k_1 + k_2} \right)^{-i} \\ &= \binom{n_T}{i} \left(\frac{k_1}{k_1 + k_2} \right)^i \cdot \left(\frac{k_2}{k_1 + k_2} \right)^{n_T} \cdot \left(\frac{k_2}{k_1 + k_2} \right)^{-i} \\ &= \binom{n_T}{i} \left(\frac{k_1}{k_1 + k_2} \right)^i \cdot \left(\frac{k_2}{k_1 + k_2} \right)^{n_T-i} \\ &= \binom{n_T}{i} \left(\frac{k_1}{k_1 + k_2} \right)^i \cdot \left(\frac{k_1 + k_2 - k_1}{k_1 + k_2} \right)^{n_T-i} \\ &= \binom{n_T}{i} \left(\frac{k_1}{k_1 + k_2} \right)^i \cdot \left(1 - \frac{k_1}{k_1 + k_2} \right)^{n_T-i} \\ &= \binom{n_T}{i} p^i (1-p)^{n_T-i} . \end{aligned}$$

For S_c the parameter p is equal to $k_2/(k_1 + k_2)$. The distribution depends on two parameters n_T and p .

Figure 4.19 illustrates the binomial probability distribution that is also the steady state probability distribution for our molecular system. The mean and variance of the binomial distribution give us the mean

$$\langle S_c(t) \rangle = n_T p ,$$

and standard deviation

$$\sqrt{n_T p (1-p)} .$$

To get a measure of variability that is independent of the mean, one usually considers the *coefficient of variation* (CV), which is the standard deviation, divided by the mean:

coefficient of variation

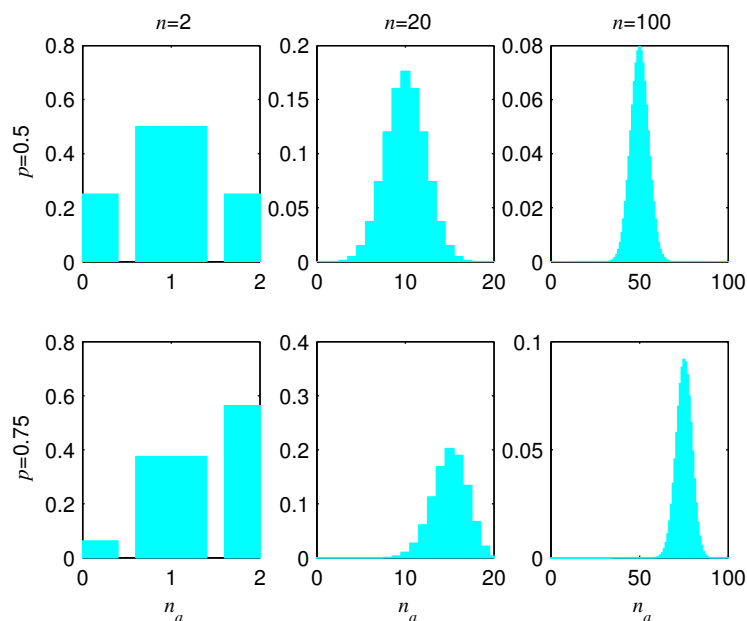


Fig. 4.19: Binomial distribution for different values of the parameters p and n_T . With regard to the isomerization reaction (4.59), the plot shows that the variance of the steady state distribution increases with a decreasing total number of molecules.

$$\text{CV}_o^\infty = \sqrt{\frac{1-p}{n_T p}} = \frac{1}{\sqrt{n_T}} \sqrt{\frac{k_2}{k_1}}, \quad (4.75)$$

$$\text{CV}_c^\infty = \frac{1}{\sqrt{n_T}} \sqrt{\frac{k_1}{k_2}}. \quad (4.76)$$

We can then say that the equilibrium coefficient of variation for S_o is inversely proportional to the square root of n_T . This is important for our discussion on the motivation for stochastic modelling. We find that if n_T increases by a factor of four, the relative variation halves. We remember that a small number of molecules suggested more randomness and hence the need for a stochastic model. However, so far we have only considered the steady state distribution and do not know how the relative variation changes beginning at time zero. In the next section we are going to show that the variation does actually increase with time, which means that if we are interested in the transient dynamic phase of the system's response and not in what happens as $t \rightarrow \infty$, a mass action model can be a good approximation even for small numbers of molecules. This is illustrated in Figure 4.20 for $n_T = 100$ and a single realization and in Figure 4.21 for $n_T = 10$ and an average over five simulation runs. What can be clearly seen is a dynamic or transient phase before the system enters a steady state. This transient phase obviously occurs only if the initial conditions are different to the steady-state value. Figure 4.22 shows a single run of the stochastic simulation and over a longer period of time. This plot has been used in [RWA02] but by choosing a longer time interval and starting with initial conditions near the steady-state, this plots hides the transient phase. Showing a single run or realization is in any case meaningless. If random fluctuations are important and they should not be ignored in modelling, then any analysis of observations or parameter estimation *must* be based on multiple repeated simulation runs or experiments. The simulations in Figure 4.20 and 4.21 also show that the variance of the stochastic simulation is smaller during the transient phase than at steady state. We will investigate this further in the next section.

In the present Section we introduced the concept of probability generating functions and conclude with properties of the pgf, which would have allowed us to derive the mean and variance of S_o directly from our knowledge of (4.71). Apart from the general definition

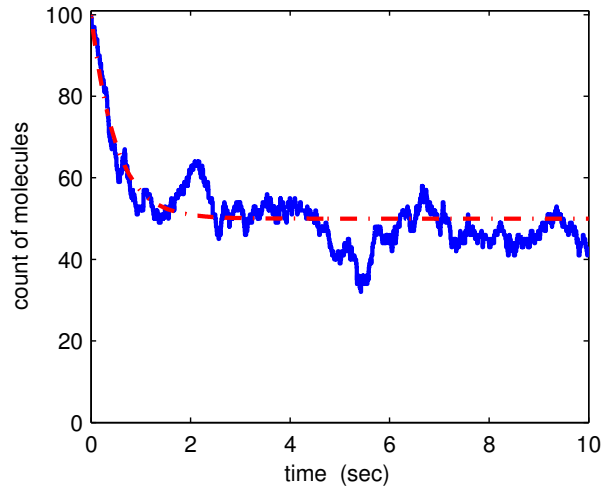


Fig. 4.20: Single realization of a stochastic simulation of an isomerization reaction with $n_T = 100$ molecules and $k_1 = k_2 = 1$. The smoother dash-dotted curve shows the solution of the mass action model. Both curves are for S_c of (4.59).

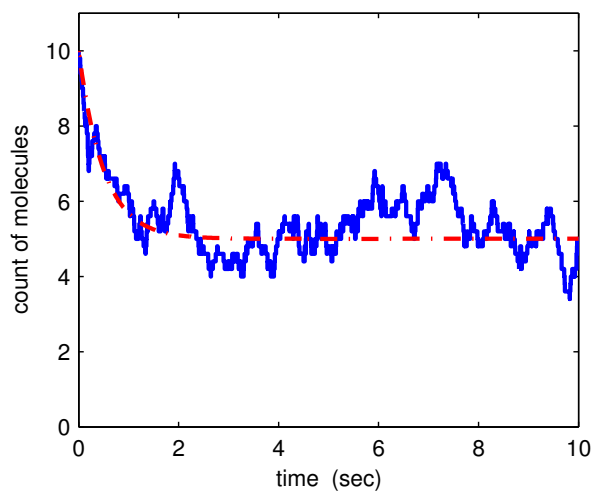


Fig. 4.21: Average over five runs of a stochastic simulation for an isomerization reaction (4.59), S_c , with $n_T = 10$ molecules and $k_1 = k_2 = 1$. The smoother dash-dotted curve shows the solution of the mass action model.

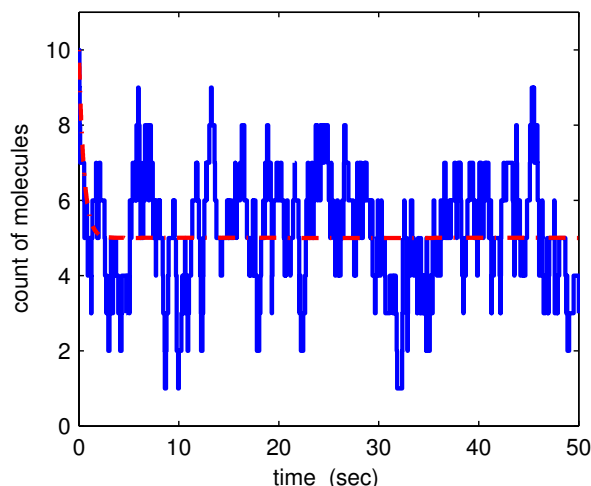


Fig. 4.22: Single run of the stochastic simulation using the parameters as in Figure 4.21. The point of this figure is that the time scale chosen matters, initial transients may be hidden for too long time scales on the plot and secondly that a single run is somewhat meaningless if the number of molecules are low and the variability is high. The more “stochastic” the data, the more realizations are required for averaging.

of the pgf $\mathfrak{P}(z)$ and its first derivative $\mathfrak{P}'(z)$ let us consider the second derivative $\mathfrak{P}''(z)$:

$$\begin{aligned}\mathfrak{P}(z) &\doteq \sum_{n=0}^{\infty} z^n P(n) , \\ \mathfrak{P}'(z) &= \sum_{n=1}^{\infty} nP(n)z^{n-1} , \\ \mathfrak{P}''(z) &= \sum_{n=2}^{\infty} n(n-1)P(n)z^{n-2} ,\end{aligned}$$

where it does not matter whether the limit of the sum is finite or not. We have

$$\mathfrak{P}(z=1) = \sum_{n=0}^{\infty} P(n) = 1$$

such that

$$E[S_c] = \mathfrak{P}'(1) = \sum_{n=0}^{\infty} nP(n)$$

and

$$\mathfrak{P}''(z=1) = \sum_{n=0}^{\infty} n(n-1)P(n) = E[S_c^2] - E[S_c]$$

such that the variance can be calculated from

$$\text{Var}[S_c] = \mathfrak{P}''(1) + \mathfrak{P}'(1) - \mathfrak{P}'(1)^2 = E[S_c^2] - E[S_c]^2 .$$

With (4.71) we have for the isomerization example (4.59)

$$E[S_c(t \rightarrow \infty)] = \frac{k_2 n_T}{k_1 + k_2}$$

and

$$\text{Var}[S_c(t \rightarrow \infty)] = \frac{k_1 k_2 n_T}{(k_1 + k_2)^2} .$$

4.6.2 Temporal evolution of average and variance

The previous section analyzed the steady-state properties of the isomerization process. The main reason to look at the steady-state distribution was mathematical convenience rather than biological reasoning. This analysis then showed that the smaller the total number of molecules n_T , the greater the variance of the steady-state distribution. This or similar arguments have been used to argue the case for stochastic modelling (e.g. [RWA02, RA03]). The simulations did however also show that for if the initial conditions of the system are away from the equilibrium, a transient period occurs during which the variance and mean of the stochastic process change and furthermore, that during the dynamic period of the system away from equilibrium the variance is smaller than it is at steady-state. This would then suggest that during this transient period, the discussion of whether a generalized mass action model or a chemical master equation (2.21) is more appropriate, the arguments against the supposedly “deterministic” model are less serious and even for low numbers of molecules the mass action model (2.19) may provide a reasonable approximation for the observations we make from experimental data. To investigate this question further, in the following three sections, we are going to derive the equations that describe the temporal evolution of the mean and variance of the chemical master equation (2.21). The solutions for the temporal evolution of mean and variance confirm and quantify the quality of approximation of the mass action model during transient and steady-state phase. An interesting finding is that the greater the differences between the rate constants of the process, the smaller is the coefficient of variation. This means that in order to provide a comprehensive comparison of the mass action and CME approach it is important to take account of the initial conditions as well as the relationship between the parameter values of the system.

To derive differential equations for the average and variance, we first introduce the *propagator function*

$$dS(dt; S(t), t) \doteq S(t + dt) - S(t) . \quad (4.77)$$

propagator function

Note that dS is a random variable as is $S(t)$. We can therefore as for

$$\text{Prob}\{ dS(dt; S(t), t) = \nu_\mu \} = a_\mu(S(t))dt ,$$

where ν_μ is then the realization of random variable dS and a_μ , defined on page 105, is the propensity (probability per unit time) of R_μ and $a_\mu dt$ is therefore the probability for an R_μ reaction to occur. The changes that in case of a reaction occur in the populations of the molecular species are described by ν_μ . The ν_μ are therefore realizations of random variable dS . For our particular example of the isomerization process (4.59) with n_T molecules we have

$$\nu_1 = 1 , \quad \nu_2 = -1 , \quad a_1 = k_1 S_c(t) , \quad a_2 = k_2 S_o(t) = k_2 (n - S_c(t)) .$$

$dS(dt; S(t), t)$ is therefore simply the average in $(t, t+dt)$, taken across the reaction channels

$$dS(dt; S(t), t) = \underbrace{\sum_{\mu=1}^M \nu_\mu a_\mu(S(t))}_{\text{average across channels}} dt .$$

We can generalize this propagator function to

$$dS^r(dt; S(t), t) \doteq W_r(S(t))dt = \sum_{\mu=1}^M \nu_\mu^r a_\mu(S(t))dt , \quad r = 1, 2, \dots$$

From (4.77) we have

$$S(t + dt) = S(t) + dS(dt; S(t), t)$$

Taking the average

$$\langle S(t+dt) \rangle = \langle S(t) \rangle + \langle W_1(S(t)) \rangle dt .$$

For the temporal evolution of the average we therefore have the differential equation

$$\boxed{\frac{d}{dt} \langle S(t) \rangle = \langle W_1(S(t)) \rangle} \quad (4.78)$$

The differential equation for the temporal evolution of the variance is slightly more complicated:

$$\begin{aligned} S^2(t+dt) &= S^2(t) + 2S(t)dS(dt; S(t), t) + dS^2(dt; S(t), t) \\ \langle S^2(t+dt) \rangle &= \langle S^2(t) \rangle + 2\langle S(t)W_1(S(t)) \rangle dt + \langle W_2(S(t)) \rangle dt \\ \frac{d}{dt} \langle S^2(t) \rangle &= 2\langle S(t)W_1(S(t)) \rangle + \langle W_2(S(t)) \rangle \end{aligned}$$

From

$$\text{Var}[S(t)] = \langle (S(t) - \langle S(t) \rangle)^2 \rangle = \langle S^2(t) \rangle - \langle S(t) \rangle^2$$

we have

$$\begin{aligned} \frac{d}{dt} \text{Var}[S(t)] &= \frac{d}{dt} \langle S^2(t) \rangle - 2\langle S(t) \rangle \frac{d}{dt} \langle S(t) \rangle \\ &= 2\langle S(t)W_1(S(t)) \rangle + \langle W_2(S(t)) \rangle - 2\langle S(t) \rangle \langle W_1(t) \rangle \end{aligned}$$

We therefore have for the ODE describing changes in the temporal evolution of the variance

$$\boxed{\frac{d}{dt} \text{Var}[S(t)] = 2[\langle S(t)W_1(S(t)) \rangle - \langle S(t) \rangle \langle W_1(S(t)) \rangle] + \langle W_2(S(t)) \rangle} \quad (4.79)$$

These equations are general, applying them to the isomerization reaction,

$$\begin{aligned} W_1(S_c) &= a_1\nu_1 + a_2\nu_2 = -k_1S_c + k_2(n_T - S_c) , \\ W_2(S_c) &= a_1\nu_1^2 + a_2\nu_2^2 = k_1S_c + k_2(n_T - S_c) , \end{aligned}$$

leading to

$$\frac{d}{dt} \langle S_c(t) \rangle = \langle W_1(S_c(t)) \rangle = k_2n_T - (k_1 + k_2)\langle S_c(t) \rangle \quad (4.80)$$

and finally for the isomerization reaction the changes of the average over time are described by

$$\frac{d}{dt} \langle S_c(t) \rangle = -(k_1 + k_2)\langle S_c(t) \rangle + k_2n_T . \quad (4.81)$$

Similarly, inserting all known values into (4.79),

$$\begin{aligned} \frac{d}{dt} \text{Var}[S_c(t)] &= 2[\langle S_c(t)W_2(S_c(t)) \rangle - \langle S_c(t) \rangle \langle W_1(S_c(t)) \rangle] + \langle W_2(S_c(t)) \rangle \\ &= -2(k_1 + k_2) [\langle S_c^2(t) \rangle - \langle S_c(t) \rangle^2] + k_2n_T + (k_1 - k_2)\langle S_c(t) \rangle , \end{aligned}$$

leading finally to

$$\frac{d}{dt} \text{Var}[S_c(t)] = -2(k_1 + k_2)\text{Var}[S_c(t)] + (k_1 - k_2)\langle S_c(t) \rangle + k_2n_T . \quad (4.82)$$

Since the differential equation for the variance does also depend on the mean it is somewhat more complicated, although not impossible to find a solution analytically. The easiest way is of course to use some mathematical software tool such as, for example, Matlab, Maple, or Mathematica.

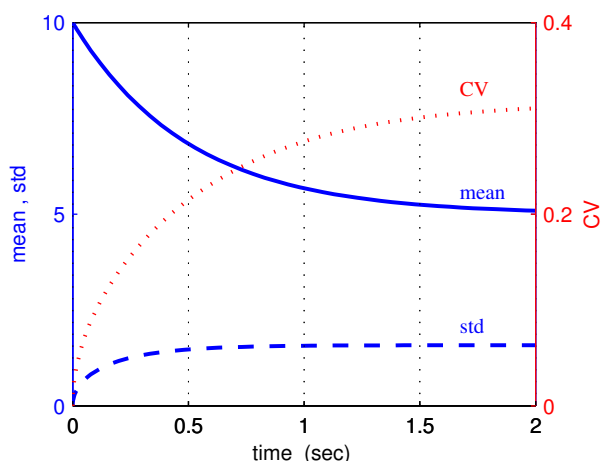


Fig. 4.23: Temporal evolution of the mean, standard deviation and coefficient of variation of S_c for the isomerization reaction (4.59), where $k_1 = 1$, $k_2 = 1$.

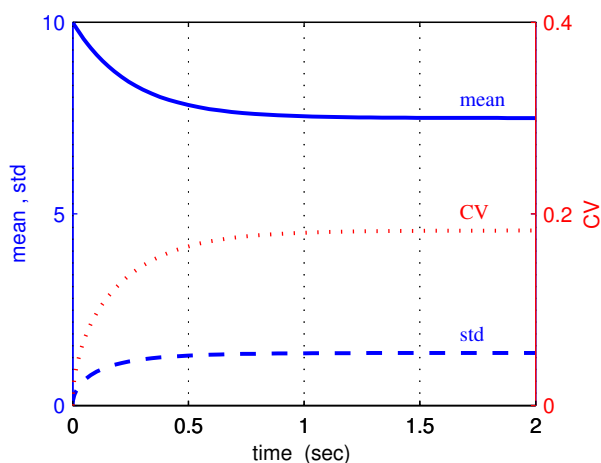
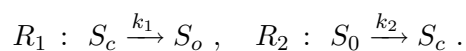


Fig. 4.24: Temporal evolution of the mean, standard deviation and coefficient of variation of S_c for the isomerization reaction (4.59), where $k_1 = 1$, $k_2 = 3$.

The numerical solutions of the differential equations (4.78) and (4.79) for our example of the isomerization reaction (4.59) are shown in Figures 4.23 and 4.24. In Figure 4.24 the rate constants are different, $k_2 = 3$, while $k_1 = 1$ such that $p = 3/4$. Figure 4.19 shows that for small n_T the steady state distribution is skewed. The steady state value of the coefficient of variation for S_c decreases as k_2 becomes larger than k_1 and this without any change to the total number of molecules.

4.6.3 Solution of the mass action model

For monomolecular reactions, the mean of the stochastic model should coincide with the solution of the mass action model. In this section we therefore derive the mass action model for (4.59)



The differential equations are

$$\begin{aligned} \frac{dS_c}{dt} &= -k_1 S_c + k_2 S_o \\ \frac{dS_o}{dt} &= k_1 S_c - k_2 S_o \end{aligned}$$

where $dS_c/dt = -dS_o/dt$. Focussing therefore on S_c , we can easily identify the steady-state part of the solution as

$$S_{c_{ss}} = \frac{k_2 n_T}{k_1 + k_2},$$

For the transient solution we expect an exponential function

$$S_{c_{tr}} = n^* \cdot e^{-(k_1+k_2)t},$$

where n^* is some unknown coefficient. The solution will be the sum of the transient and steady-state terms

$$S_c(t) = n^* e^{-(k_1+k_2)t} + \frac{k_2 n_T}{k_1 + k_2}.$$

At $t = 0$,

$$S_c(0) = n^* + \frac{k_2 n_T}{k_1 + k_2} \quad \text{from which} \quad n^* = S_c(0) - \frac{k_2 n_T}{k_1 + k_2}$$

such that

$$S_c(t) = \left(S_c(0) - \frac{k_2 n_T}{k_1 + k_2} \right) e^{-(k_1+k_2)t} + \frac{k_2 n_T}{k_1 + k_2}.$$

For $k_1 = k_2 = 1$, the solution of the mass action model is given by

$$S_c(t) = \left(S_c(0) - \frac{n_T}{2} \right) e^{-2t} + \frac{n_T}{2}$$

If $S_c(0) > n_T/2$ the amount of S_c exponentially decays to the value $n_T/2$.

We note in passing that numerical solutions for the mass action model (2.19) can very easily be obtained using some mathematical software package such as, for example, Matlab, Maple or Mathematica. There are also numerous free software packages, some of which are specifically developed to simulate biochemical reaction networks. The SBML web-site provides links to most of those packages: www.sbml.org. In Figure 4.25 this is illustrated for Matlab.

```
function [t,n] = gma(tspan,pars)
NAV = pars{1}; % product of Avogadro constant and Volume.
S0 = pars{8}; % initial conditions.
options = odeset('refine',16); % options for the ODE solver.
[t,S] = ode15s(@compdS,tspan,S0,options,pars); % ODE solver.
n = NAV*S*1e-9; % number of molecules predicted by gma

function dS = compdS(t,S,pars)
k = pars{2}; % rate constants
M = pars{4}; % number of reaction channels
p = pars{5}; % index selection for molecular species
l = pars{6}; % stoichiometric coefficient
nu = pars{9}; % changes that occur in each channel
for u=1:M
P(u) = prod(S(p{u}(:)).^l{u}(:));
end dS = (k.*P*nu)';
```

Fig. 4.25: Matlab functions to simulate a mass action model. The elements of these functions are a translation of Equation 2.19. The variable `tspan` defines the simulation period and `pars` collects all parameters that may be relevant.

This section has shown again that it is important to distinguish between stochastic mechanisms and modelling observations, regardless of whether they are ‘deterministic’ or ‘stochastic’. Here we distinguish between random and stochastic processes. A ‘random

process' or 'noise' is a description of a phenomena where observations of real or simulated data are subject to fluctuations that are either not predictable or appear unpredictable. A 'stochastic process' is a mathematical concept that can be employed to model random processes. Reducing the discussion to a choice of a 'deterministic' or 'stochastic' model is misleading. We need to distinguish between the hypothesis of a stochastic mechanism ("intrinsic noise"), random variations in observations ("observation or measurement noise") and how to represent these phenomena with the help of mathematical models. If intrinsic or extrinsic random fluctuations are significant, it is necessary to obtain repeated measurements and simulations to reduce uncertainty.

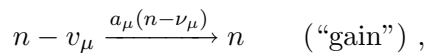
Noise can have a role in biological function or is simply a nuisance in understanding intracellular dynamics. The integration of many processes or components, each displaying random behavior, can lead to regularity. Tissue, and organs are examples of higher levels of organizations where a large number of cells are integrated into a larger whole.

4.6.4 Generating functions

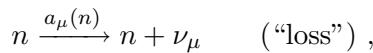
In Section (4.6.1) we made use of the probability generating function to derive the mean and variance of the steady-state probability distribution. Our derivation focussed on a transformation of the chemical master equation (CME) for our particular example. Here we introduce the concepts of *moment generating functions*⁹ and *cumulant generating functions*. These are convenient transformations to help generate equations for probabilities, mean and variance of the general CME (2.21):

$$\frac{\partial P_n(t)}{\partial t} = \sum_{\mu=1}^M \left[\underbrace{a_{\mu}(n - \nu_{\mu})P_{n-\nu_{\mu}}(t)}_{\text{gain term}} - \underbrace{a_{\mu}(n)P_n(t)}_{\text{loss term}} \right], \quad n = 0, 1, 2, \dots, n_T, \quad (4.83)$$

where $P(n, t) \equiv P_n(t)$ and state transitions, towards n :



with propensity $a_{\mu}(n - \nu_{\mu})$, and away from n :



with propensity $a_{\mu}(n)$. Let the propensity of reaction channel R_{μ} be generally defined as

$$a_{\mu}(n) = \sum_r c_{\mu r} n^r, \quad (4.84)$$

with some coefficient $c_{\mu r}$, unrelated to the stochastic reaction constant c_{μ} in stochastic simulation. This general expression includes as a special case the linear version (4.28) defined in Section 4.4.1.

Let us first consider the probability generating function (pgf)

$$\mathfrak{P}(z, t) \doteq \sum_n z^n P_n(t), \quad (4.85)$$

where z is some real number $0 \leq z \leq 1$ and $0 \leq n \leq \infty$. Note that if the total number of molecules is limited to n_T , terms higher than n_T will be zero. Multiplying both sides of

⁹For basic introduction to probability- and moment generating functions we refer to [JS01].

(4.83) with z^n , taking the sum over all possible states and inserting (4.84) for a_μ , gives us the following expression

$$\begin{aligned} \sum_n z^n \frac{\partial P_n(t)}{\partial t} &= \sum_\mu \left[\sum_n z^n \sum_r c_r (n - \nu_\mu)^r P_{n-\nu_\mu}(t) - \sum_n z^n \sum_r c_r n^r P_n(t) \right], \\ &= \sum_\mu \left[\sum_r c_r \sum_n z^n (n - \nu_\mu)^r P_{n-\nu_\mu}(t) - \sum_r c_r \sum_n z^n n^r P_n(t) \right]. \end{aligned} \quad (4.86)$$

From (4.85),

$$\sum_n z^n \frac{\partial P_n(t)}{\partial t} = \frac{\partial \mathfrak{P}(z, t)}{\partial t},$$

which translates us the left-hand side of (4.86). Furthermore, from (4.85)

$$z \frac{\partial \mathfrak{P}(z, t)}{\partial z} = \sum_n z^n n P_n(t), \quad (4.87)$$

$$z^2 \frac{\partial^2 \mathfrak{P}(z, t)}{\partial z^2} = \sum_n z^n n^2 P_n(t), \quad (4.88)$$

⋮

$$z^r \frac{\partial^r \mathfrak{P}(z, t)}{\partial z^r} = \sum_n z^n n^r P_n(t), \quad (4.89)$$

where the last rule gives us a transformation of the “loss-term” of (4.86). From the “gain-term” consider the sum

$$\sum_n z^n (n - \nu_\mu)^r P_{n-\nu_\mu}(t),$$

which by introducing a new index $m = n - \nu$ can be rewritten as

$$z^{\nu_\mu} \sum_m z^m m^r P_m(t),$$

which, from (4.89) gives us the transformation

$$\sum_n z^n (n - \nu_\mu)^r P_{n-\nu_\mu}(t) \rightarrow z^{\nu_\mu} z^r \frac{\partial^r \mathfrak{P}(z, t)}{\partial z^r}.$$

We are now in a position to transform (4.86) into

$$\begin{aligned} \frac{\partial \mathfrak{P}(z, t)}{\partial t} &= \sum_{\mu=1}^M \left[\sum_r c_r z^{\nu_\mu} z^r \frac{\partial^r \mathfrak{P}(z, t)}{\partial z^r} - \sum_r c_r z^r \frac{\partial^r \mathfrak{P}(z, t)}{\partial z^r} \right], \\ &= \sum_{\mu=1}^M (z^{\nu_\mu} - 1) \sum_r c_r z^r \frac{\partial^r \mathfrak{P}(z, t)}{\partial z^r}. \end{aligned} \quad (4.90)$$

With the definition for a_μ , (4.84), this leads us to the general transformation of the chemical master equation by means of a probability generating function:

$$\boxed{\frac{\partial \mathfrak{P}(z, t)}{\partial t} = \sum_\mu (z^{\nu_\mu} - 1) a_\mu \left(z \frac{\partial}{\partial z} \right) \mathfrak{P}(z, t)} \quad (4.91)$$

Solving and inverting this equation gives the temporal evolution of the probabilities.

For our example of an isomerization reaction, (4.59),



let $\#S_c = n$ and $\#S_o = n_T - n$,

$$\begin{aligned} a_1(n) &= k_1 n, & \nu_1 &= -1, \\ a_2(n) &= k_2(n_T - n), & \nu_2 &= +1. \end{aligned}$$

To apply (4.91), replace n by the operator

$$\begin{aligned} a_1 \left(z \frac{\partial}{\partial z} \right) &= k_1 \frac{\partial}{\partial z} \\ a_2 \left(z \frac{\partial}{\partial z} \right) &= k_2 n_T - k_2 z \frac{\partial}{\partial z}, \end{aligned}$$

such that

$$\begin{aligned} \frac{\partial \mathfrak{P}(z, t)}{\partial t} &= \sum_{\mu} (z^{\nu_{\mu}} - 1) a_{\mu} \left(\frac{\partial}{\partial z} \right) \mathfrak{P}(z, t) \\ &= ((z^{-1} - 1) k_1 z \frac{\partial \mathfrak{P}(z, t)}{\partial z} + (z - 1) [k_2 n_T - k_2 z \frac{\partial}{\partial z}]) \mathfrak{P}(z, t) \\ &= (z - 1) k_2 n_T \mathfrak{P}(z, t) - (z - 1) (k_1 + k_2 z) \frac{\partial \mathfrak{P}(z, t)}{\partial z}. \end{aligned}$$

Solving this equation is usually not trivial and for our particular example of an isomerization reaction, the solution gives the pgf

$$\begin{aligned} P(z, t) &= (k_1 + k_2)^{-n_T} \left[k_1 + k_2 z - k_2(z - 1)e^{-(k_1 + k_2)t} \right]^{n_T - \#S_c(0)} \\ &\quad \cdot \left[k_1 + k_2 z + k_1(z - 1)e^{-(k_1 + k_2)t} \right]^{\#S_c(0)}. \end{aligned} \quad (4.92)$$

The pgf can be used to obtain expressions for the mean and variance. For this we can facilitate the following nice properties of the pgf. Let us first summarize the pgf and its derivatives:

$$\begin{aligned} \mathfrak{P}(z, t) &\doteq \sum_{n=0}^{\infty} z^n P_n(t), \\ \mathfrak{P}'(z, t) &= \sum_{n=1}^{\infty} n P_n(t) z^{n-1}, \\ \mathfrak{P}''(z, t) &= \sum_{n=2}^{\infty} n(n-1) P_n(t) z^{n-2}, \end{aligned}$$

We have

$$\mathfrak{P}(z = 1, t) = \sum_{n=1}^{\infty} P_n(t) = 1$$

such that the mean is given by¹⁰

$$\begin{aligned} \langle S(t) \rangle &= \mathfrak{P}'(z = 1, t) \\ &= \sum_{n=0}^{\infty} n P_n(t) \end{aligned} \quad (4.93)$$

¹⁰We could also write for the mean $E[S(t)]$ but this may suggest a mean over time. We are here considering a mean of S at a particular, fixed t and hence write $\langle S(t) \rangle$.

and

$$\mathfrak{P}''(z = 1, t) = \sum_{n=0}^{\infty} n(n-1)P_n(t) = \langle S(t)^2 \rangle - \langle S(t) \rangle^2$$

such that the variance can be calculated from

$$\text{Var}[S(t)] = \mathfrak{P}''(1, t) + \mathfrak{P}'(1, t) - \mathfrak{P}'(1, t)^2 = \langle S(t)^2 \rangle - \langle S(t) \rangle^2. \quad (4.94)$$

Making all the substitutions required by hand is cumbersome. For the example of the isomerization reaction, in Figure 4.26 it is shown how, using Matlab, we can derive the equations symbolically.

```
clear all,clc
syms k1 k2 z t nT S0 positive % define symbolic variables.
P=(k1+k2)^-nT*(k1+k2*z-k2*(z-1)*exp(-(k1+k2)*t))^(nT-S0)...
*(k1+k2*z+k1*(z-1)*exp(-(k1+k2)*t))^S0;
mu = simple(subs(diff(P,z),z,1)); % mean as a function of time.
disp('mu = '); pretty(mu)
mu_init = simple(subs(mu,t,0)); % set the initial mean.
disp('mu_init = '); pretty(mu_init)
mu_ss = simple(subs(mu,t,inf)); % steady-state for mean.
disp('mu_ss = '); pretty(mu_ss)
sigma2 = simple(subs(diff(P,z,2),z,1)+mu-mu^2); % variance
disp('sigma2 = '); pretty(sigma2)
dsigma2 = simple(diff(sigma2,t)); % derivative of variance.
disp('dsigma2 = '); pretty(dsigma2)
sigma2_init = simple(subs(sigma2,t,0));
disp('sigma2_init = '); pretty(sigma2_init)
sigma2_ss = simple(subs(sigma2,t,inf)); % steady-state variance.
disp('sigma2_ss = '); pretty(sigma2_ss)
```

Fig. 4.26: Matlab code to obtain symbolic expressions for the temporal evolution as well as steady-state distributions for the isomerization reaction, using Eq. (4.92).

The mean and variance are extremely useful in characterizing a probability distribution, the mean is a measure of ‘central tendency’ while the variance describes the ‘spread’ of the distribution around this central value. In general, however these two measures cannot uniquely characterize a distribution. For this purpose, a sequence of constants, called **moments** *moments*, has been introduced to describe the shape of a distribution. At a particular t , for n being the scalar outcome of a random variable S , for discrete distributions the i th moment m_i around zero is defined¹¹ by

$$m_i = \sum_n n^i p_n, \quad i = 1, 2, \dots \quad (4.95)$$

The first two moments around the origin add nothing to what the mean and variance tell us since

$$m_1 = \sum_n n p_n = E[S]$$

$$m_2 = \sum_n n^2 p_n = \sum_n (n - E[S])^2 p_n + (E[S])^2 \left(\sum_n p_n \right),$$

such that

$$\text{Var}[S] = m_2 - (E[S])^2.$$

It then turns out that m_3 measures asymmetry or ‘skewness’ about the mean and m_4

¹¹In the definition of the moments we focus on a particular, fixed point in time and we may therefore leave out t from the equations to simplify the notation.

measures ‘flatness’. Skewness is defined by

$$\begin{aligned} E[(S - m_1)^3] &= E[S^3 - 3S^2m_1 + 3Sm_1^2 - m_1^3] \\ &= m_3 - 3m_1m_2 + 2m_1^3 . \end{aligned} \quad (4.96)$$

We notice that the definition of skewness is only indirectly defined by m_3 . Further below we find that *cumulants* will provide a more direct measure.

The *moment generating function* (mgf) \mathfrak{M} is defined as the Laplace transformation of the probability distribution p_n : moment generating function

$$\mathfrak{M}(\theta) = \int p_n e^{\theta n} dn , \quad (4.97)$$

which may also be seen as the expectation

$$\mathfrak{M}(\theta) = E \left[e^{\theta n} \right] .$$

For a discrete distribution, the mgf \mathfrak{M} is defined by

$$\mathfrak{M}(\theta) = \sum_n e^{\theta n} p_n .$$

Now allowing t to return into the equations, if we compare the probability generating function (4.85) with those of the moment generating function we have

$$\mathfrak{M}(\theta, t) = \mathfrak{P}(e^\theta, t) , \quad (4.98)$$

where we can arrive at the moment generating function simply by making the following substitution

$$z = e^\theta .$$

Making use of the chain rule,

$$\frac{\partial \mathfrak{M}}{\partial \theta} = \frac{\partial \mathfrak{P}}{\partial z} \cdot \frac{\partial z}{\partial \theta} = z \frac{\partial \mathfrak{P}}{\partial z}$$

therefore

$$z \frac{\partial}{\partial z} = \frac{\partial}{\partial \theta} .$$

Considering now the definition for propensity a_μ , (4.84), we can make the simple substitution

$$a_\mu \left(z \frac{\partial}{\partial z} \right) = a_\mu \left(\frac{\partial}{\partial \theta} \right)$$

in (4.91) to get the moment generating function of the general CME

$$\boxed{\frac{\partial \mathfrak{M}(\theta, t)}{\partial t} = \sum_\mu \left(e^{\theta \nu_\mu} - 1 \right) a_\mu \left(\frac{\partial}{\partial \theta} \right) \mathfrak{M}(\theta, t)} \quad (4.99)$$

To check that $\mathfrak{M}(\theta, t)$ is indeed a generating *function*, let us consider

$$\mathfrak{M}(\theta, t) = \mathfrak{P}(e^\theta, t) = \sum_n e^{\theta n} p_n(t)$$

with

$$e^{\theta n} = \sum_i \frac{(\theta n)^i}{i!}$$

this becomes

$$\mathfrak{M}(\theta, t) = \sum_n p_n(t) \sum_i \frac{\theta^i n^i}{i!} = \sum_i \frac{\theta^i}{i!} \sum_n n^i p_n(t) .$$

From the definition of moments

$$m_i(t) = \sum_n n^i p_n(t) , \quad i = 1, 2, \dots$$

and therefore

$$\mathfrak{M}(\theta, t) = \sum_i m_i(t) \frac{\theta^i}{i!} . \quad (4.100)$$

Taylor series Now, if we remember the Taylor series, which is a series expansion of a function, $f(t)$, about a point t_0 ,

$$f(t) = f(t_0) + (t - t_0)f'(t_0) + \frac{1}{2!}(t - t_0)^2 f''(t_0) + \dots \quad (4.101)$$

MacLaurin series where f' and f'' denote the first and second derivative. If $t = 0$, the expansion is known as a MacLaurin series. By letting $t - t_0 \equiv \theta$, and substituting $t = t_0 + \theta$ into (4.101) we get the alternative version

$$\begin{aligned} f(t_0 + \theta) &= f(t_0) + f'(t_0)\theta + \frac{1}{2!}f''(t_0)\theta^2 + \dots \\ &= \sum_i m_i \theta^i , \quad \text{where } m_i = \frac{f^{(i)}(t_0)}{i!} \\ &= \sum_i m_i \frac{\theta^i}{i!} , \quad \text{where } m_i = f^{(i)}(t_0) . \end{aligned} \quad (4.102)$$

where $f^{(i)}$ denotes the i th derivative of f . Comparing (4.102) with (4.100), we notice that the moment function $m_i(t)$ ought to be the i th derivative of $\mathfrak{M}(\theta, t)$ at $\theta = 0$:

$$m_i(t) = \left. \frac{d^i \mathfrak{M}(\theta, t)}{d\theta^i} \right|_{\theta=0} . \quad (4.103)$$

cumulant generating function

Next we introduce the cumulant generating function (cgf)

$$\mathfrak{C}(\theta, t) = \ln (\mathfrak{M}(\theta, t)) \quad (4.104)$$

and consider its expansion

$$\mathfrak{C}(\theta, t) = \sum_i \kappa_i(t) \frac{\theta^i}{i!}$$

so that $\kappa_i(t)$ is the *cumulant function* for which

$$\kappa_i(t) = \left. \frac{d^i \mathfrak{C}(\theta, t)}{d\theta^i} \right|_{\theta=0} = \mathfrak{C}^{(i)}(0, t) \quad (4.105)$$

It is useful to establish a relation between the cumulant and moment functions. Starting with

$$m_i(t) = \mathfrak{M}^{(i)}(0, t) ,$$

from (4.104),

$$\mathfrak{M}(\theta, t) = e^{\mathfrak{C}(\theta, t)}$$

such that

$$m_i(t) = \left(e^{\mathfrak{C}(\theta, t)} \right)^{(i)} \Big|_{\theta=0} .$$

Now

$$\left(e^{\mathfrak{C}(\theta,t)}\right)^{(i)} = \left(\mathfrak{C}'(\theta,t) \cdot \mathfrak{M}(\theta,t)\right)^{(i-1)}$$

Leibniz's theorem To proceed we require the *Leibniz's theorem* for differentiation¹² of a product of two functions f and g

$$\begin{aligned} (f \cdot g)' &= f \cdot g' + f' \cdot g \\ (f \cdot g)'' &= (f \cdot g'' + f' \cdot g') + (f' \cdot g' + f'' \cdot g) = f \cdot g'' + 2f' \cdot g' + f'' \cdot g \\ &\vdots \\ (f \cdot g)^{(i)} &= \sum_{r=0}^i \binom{i}{r} f^{(i-r)} \cdot g^{(r)}. \end{aligned}$$

Comparing this to

$$\left(\mathfrak{C}'(\theta,t) \cdot \mathfrak{M}(\theta,t)\right)^{(i-r)} = \sum_{r=0}^{i-1} \binom{i-1}{r} \mathfrak{C}^{(i-1)}(\theta,t) \cdot \mathfrak{M}^{(r)}(\theta,t).$$

Therefore

$$m_i(t) = \sum_{r=0}^{i-1} \binom{i-1}{r} \mathfrak{C}^{i-r}(0,t) \cdot \mathfrak{M}^{(r)}(0,t) \quad \text{at } \theta = 0.$$

However, we also know from (4.105) and (4.103)

$$\mathfrak{C}^{(i-r)}(0,t) = \kappa_{i-r}(t) \quad \text{and} \quad \mathfrak{M}^{(r)}(0,t) = m_r(t),$$

hence a general relationship between cumulants and moments is given by

$$m_i(t) = \sum_{r=0}^{i-1} \binom{i-1}{r} \kappa_{i-1}(t) m_r(t) \quad \text{where } m_0(t) = 1. \quad (4.106)$$

For $i = 1, 2, 3$, we have

$$m_1(t) = \kappa_1(t) = \langle S(t) \rangle \quad (4.107)$$

$$m_2(t) = \kappa_2(t) + \kappa_1(t)m_1(t) \quad (4.108)$$

$$m_3(t) = \kappa_3(t) + 2\kappa_2(t)m_1(t) + \kappa_1(t)m_2(t) \quad (4.109)$$

From (4.107) and (4.108),

$$\kappa_2(t) = m_2(t) - m_1^2(t) = \text{Var}[S(t)]. \quad (4.110)$$

Noting that $m_1 = E[S]$ and $m_2 = E[S^2]$ and $\text{Var}[S] = E[S^2] - (E[S])^2 = m_2 - m_1^2$, we see that κ_1 gives the mean and κ_2 defines the variance. Inserting (4.107), (4.110) into (4.109) gives

$$m_3(t) = \kappa_3(t) + 2m_1(t) (m_2(t) - m_1^2(t)) + m_1(t)m_2(t)$$

and thus

$$\kappa_3(t) = m_3(t) - 3m_1(t)m_2(t) + 2m_1^3(t) \quad (4.111)$$

gives a direct measure of skewness. To derive a differential equation for the temporal evolution of the cgf we first consider the differential equation for the mgf of a CME model, (4.99). Differentiating (4.104) we have

$$\frac{\partial \mathfrak{C}}{\partial \theta} = \frac{1}{\mathfrak{M}} \cdot \frac{\partial \mathfrak{M}}{\partial \theta}, \quad \frac{\partial \mathfrak{C}}{\partial t} = \frac{1}{\mathfrak{M}} \cdot \frac{\partial \mathfrak{M}}{\partial t}$$

¹²For ease of notation we use here f' and f'' to denote the first and second-order derivative of function f .

Dividing (4.99) by $\mathfrak{M}(\theta, t)$

$$\frac{1}{\mathfrak{M}} \frac{\partial \mathfrak{M}}{\partial t} = \sum_m \left(e^{\nu_\mu \theta} - 1 \right) \frac{1}{\mathfrak{M}} a_\mu \left(\frac{\partial}{\partial \theta} \right) \mathfrak{M}(\theta, t). \quad (4.112)$$

If we now restrict the general propensity $a_\mu(n)$, (4.84), to be of first order, i.e., following the assumption in Gillespie modelling, then

$$a_\mu(n) = c_{\mu 0} + c_{\mu 1} n$$

Such that

$$\begin{aligned} a_\mu \left(\frac{\partial}{\partial \theta} \right) \mathfrak{M}(\theta, t) &= c_{\mu 0} \mathfrak{M} + c_{\mu 1} \frac{\partial}{\partial \theta} \mathfrak{M} \\ \frac{1}{\mathfrak{M}} a_\mu \left(\frac{\partial}{\partial \theta} \right) \mathfrak{M}(\theta, t) &= c_{\mu 0} + c_{\mu 1} \frac{1}{\mathfrak{M}} \frac{\partial \mathfrak{M}}{\partial \theta} = c_{\mu 0} + c_{\mu 1} \frac{\partial \mathfrak{C}}{\partial \theta} \end{aligned}$$

Now (4.112) becomes for a first-order propensity,

$$\boxed{\frac{\partial \mathfrak{C}(\theta, t)}{\partial t} = \sum_\mu \left(e^{\nu_\mu \theta} - 1 \right) \left(c_{\mu 0} + c_{\mu 1} \frac{\partial \mathfrak{C}(\theta, t)}{\partial \theta} \right)} \quad (4.113)$$

as the cgf equivalent of (4.91) and (4.99). To derive the differential equations for cumulants, consider

$$\begin{aligned} e^{\nu_\mu \theta} - 1 &= \sum_{j>0} \frac{\nu_\mu^j \theta^j}{j!}, \quad \mathfrak{C}(\theta, t) = \frac{\kappa_i(t) \theta^i}{i!}, \\ \frac{\partial \mathfrak{C}}{\partial t} &= \sum_i \frac{\dot{\kappa}_i \theta^i}{i!}, \quad \frac{\partial \mathfrak{C}}{\partial \theta} = \sum_{i>0} \frac{\kappa_i \theta^{i-1}}{(i-1)!} = \sum_r \frac{\kappa_{r+1} \theta^r}{r!}. \end{aligned}$$

Inserting these expression into (4.113),

$$\begin{aligned} \sum_i \frac{\dot{\kappa}_i \theta^i}{i!} &= \sum_\mu \left[\left(\sum_{j>0} \frac{\nu_\mu^j \theta^j}{j!} \right) \left(c_{\mu 0} + c_{\mu 1} \sum_r \frac{\kappa_{r+1} \theta^r}{r!} \right) \right] \\ &= \sum_\mu c_{\mu 0} \sum_{j>0} \frac{\nu_\mu^j \theta^j}{j!} + \sum_\mu c_{\mu 1} \sum_{j>0} \sum_r \frac{\nu_\mu^j \kappa_{r+1} \theta^{j+r}}{j! r!}. \end{aligned}$$

Comparing coefficients with equal powers of θ , beginning with $i = 1$:

$$\dot{\kappa}_1 = \sum_\mu c_{\mu 0} \nu_\mu + \sum_\mu c_{\mu 1} \sum_{j+r=1} \frac{\nu_\mu^j \kappa_{r+1}}{j! r!}.$$

Since $j > 0$, and $j + r = 1$ implies $j = 1, r = 0$,

$$\frac{d\kappa_1(t)}{dt} = \left(\sum_\mu c_{\mu 0} \nu_\mu \right) + \left(\sum_\mu c_{\mu 1} \nu_\mu \right) \kappa_1 \quad (4.114)$$

For $i = 2$,

$$\frac{1}{2!} \dot{\kappa}_2 = \sum_\mu c_{\mu 0} \frac{\nu_\mu^2}{2!} + \sum_\mu c_{\mu 1} \sum_{j+r=2} \frac{\nu_\mu^j \kappa_{r+1}}{j! r!}.$$

Again, since $j > 0$, $j + r = 2$ implies $j = 2$, $r = 0$, or $j = 1$, $r = 1$,

$$\begin{aligned} \frac{d\kappa_2(t)}{dt} &= \sum_{\mu} c_{\mu 0} \nu_{\mu}^2 + 2 \sum_{\mu} c_{\mu 1} \left(\frac{\nu_{\mu}^2 \kappa_1}{2!} + \nu_{\mu} \kappa_2 \right) \\ &= 2 \left(\sum_{\mu} c_{\mu 1} \nu_{\mu} \right) \kappa_2 + \left(\sum_{\mu} c_{\mu 1} \nu_{\mu}^2 \right) \kappa_1 + \sum_{\mu} c_{\mu 0} \nu_{\mu}^2 . \end{aligned} \quad (4.115)$$

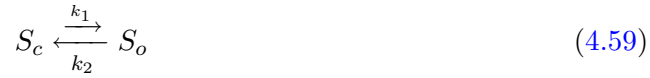
For $i = 3$,

$$\frac{1}{3!} \dot{\kappa}_3 = \sum_{\mu} c_{\mu 0} \frac{\nu_{\mu}^3}{3!} + \sum_{\mu} c_{\mu 1} \sum_{j+r=3} \frac{\nu_{\mu}^j \kappa_{r+1}}{j! r!} .$$

Again, $j + r = 3$ implies $(j, r) = (3, 0), (2, 1), (1, 2)$, and thus

$$\frac{d\kappa_3(t)}{dt} = 3 \left(\sum_{\mu} c_{\mu 1} \nu_{\mu} \right) \kappa_3 + 3 \left(\sum_{\mu} c_{\mu 1} \nu_{\mu}^2 \right) \kappa_2 + \left(\sum_{\mu} c_{\mu 1} \nu_{\mu}^3 \right) \kappa_1 + \sum_{\mu} c_{\mu 0} \nu_{\mu}^3 . \quad (4.116)$$

We therefore now have the differential equations for the first three cumulants. Let us apply this to our example of an isomerization reaction (4.59),



let $\#S_c = n$ and $\#S_o = n_T - n$,

$$\begin{aligned} a_1(n) &= k_1 n , & \nu_1 &= -1 , \\ a_2(n) &= k_2 (n_T - n) , & \nu_2 &= +1 . \end{aligned}$$

For $a_{\mu}(n) = c_{\mu 0} + c_{\mu 1} n$, $c_{10} = 0$, $c_{11} = k_1$, $c_{20} = k_2 n_T$, $c_{21} = -k_2$, such that

$$\begin{aligned} \sum_{\mu} c_{\mu 0} \nu_{\mu} &= k_2 n_T , & \sum_{\mu} c_{\mu 1} \nu_{\mu} &= -(k_1 + k_2) , \\ \sum_{\mu} c_{\mu 1} \nu_{\mu}^2 &= k_1 - k_2 , & \sum_{\mu} c_{\mu 0} \nu_{\mu}^2 &= k_2 n_T , \\ \sum_{\mu} c_{\mu 1} \nu_{\mu}^3 &= -(k_1 + k_2) , & \sum_{\mu} c_{\mu 0} \nu_{\mu}^3 &= k_2 n_T . \end{aligned}$$

For the isomerization reaction the differential equations, describing the temporal evolution of the cumulants are given as

$$\text{mean: } \frac{d\kappa_1}{dt} = -(k_1 + k_2) \kappa_1 + k_2 n_T , \quad (4.117)$$

$$\text{variance: } \frac{d\kappa_2}{dt} = -2(k_1 + k_2) \kappa_2 + (k_1 - k_2) \kappa_1 + k_2 n_T , \quad (4.118)$$

$$\text{skewness: } \frac{d\kappa_3}{dt} = -3(k_1 + k_2) \kappa_3 + 3(k_1 - k_2) \kappa_2 - (k_1 + k_2) \kappa_1 + k_2 n_T . \quad (4.119)$$

This then confirm the result of (4.80) and (4.82). Rewrite the first two equations in matrix form:

$$\begin{bmatrix} \dot{\kappa}_1 \\ \dot{\kappa}_2 \end{bmatrix} = \begin{bmatrix} -(k_1 + k_2) & 0 \\ k_1 - k_2 & -2(k_1 + k_2) \end{bmatrix} \begin{bmatrix} \kappa_1 \\ \kappa_2 \end{bmatrix} + \begin{bmatrix} k_2 n_T \\ k_2 n_T \end{bmatrix} \quad (4.120)$$

The lower left term of the matrix on the right-hand side, can be interpreted as a *coupling term* between mean and variance. For the special case $k_1 = k_2$, considered in [RWA02]

and Section 4.6.1, the mean and variance are uncoupled, the temporal evolution of the variance is independent of the mean. The solution for κ_2 is in this case given by

$$\begin{aligned}\kappa_2(t) &= e^{-2(k_1+k_2)t} \kappa_2(0) + k_2 n_T \int_0^t e^{-2(k_1+k_2)(t-\tau)} d\tau, \\ &= k_2 n_T \int_0^t e^{-2(k_1+k_2)(t-\tau)} d\tau \quad \text{since } \kappa_2(0) = 0, \\ &= k_2 n_T e^{-2(k_1+k_2)t} \int_0^t e^{2(k_1+k_2)\tau} d\tau, \\ &= \frac{n_T}{2} \cdot \frac{k_2}{(k_1+k_2)} \cdot \left(e^{-2(k_1+k_2)t} - 1 \right).\end{aligned}\tag{4.121}$$

In general, the solution for the mean of S_c in the isomerization reaction is,

$$\kappa_1(t) = \frac{n_T}{k_1+k_2} \left(k_2 + k_1 e^{-(k_1+k_2)t} \right),$$

and for the variance

$$\kappa_2(t) = \frac{k_1}{k_1+k_2} \left(1 - e^{-(k_1+k_2)t} \right) \kappa_1(t).$$

Such calculations are easy to do with a mathematical software package, capable of symbolic maths. For example, in Mathematica, the two equations are solved by the command:

```
DSolve[{kappa2'[t]==(k1-k2)*kappa1[t]-2*(k1+k2)*kappa2[t]+k2*nT,
kappa1'[t]==-(k1+k2)*kappa1[t]+k2*nT, kappa1[0]==nT, kappa2[0]==0}, {kappa1[t],
kappa2[t]}, t]
```

For $k_1 \neq k_2$ the variance is clearly a function of the mean. It is for this reason why one should consider the coefficient of variation (CV), i.e., the standard deviation divided by the mean to obtain a measure of variability that is independent of the mean. This then explains Figures 4.23 and 4.24. For S_c , the coefficient of variation is

$$CV[S_c(t)] = \sqrt{\frac{k_1}{k_1+k_2} (1 - e^{-(k_1+k_2)t})} \cdot \frac{1}{\sqrt{\kappa_1(t)}}.$$

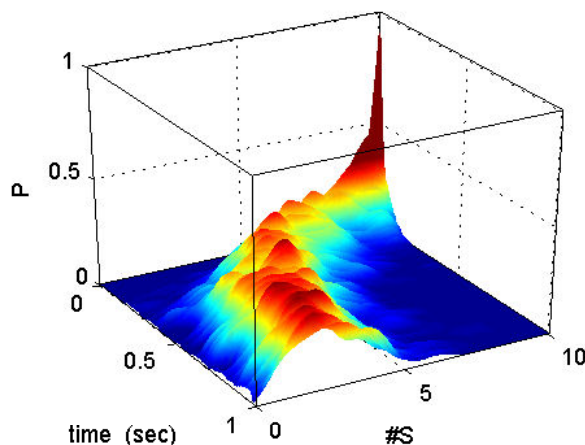


Fig. 4.27: The plots shows the probability distribution evolving with time for the isomerization reaction with $k_1 = 3, k_2 = 1$.

4.6.5 Summary: The master-equation approach

In conclusion, our analysis showed that in order to provide a comprehensive comparison of the mass action model, (2.19), and the CME, (2.21), it is important to take account of the initial conditions as well as the relationship between the parameter values of the system. If the initial conditions of the system are away from the equilibrium, a transient period occurs during which the variance and mean of the stochastic process change and furthermore, that during the dynamic period of the system away from equilibrium the variance is smaller than it is at steady-state. During this transient period, the arguments against the supposedly “deterministic” model are less serious and even for low numbers of molecules the mass action model can provide a reasonable approximation for the observations we make from experimental data. An interesting finding is that the greater the differences between the rate constants of the process, the smaller is the coefficient of variation.

5 Cell Communication

The life of multicellular animals begins with the fertilization of an oocyte (egg cell) by a sperm cell. Thereafter the fertilized egg undergoes a series of cell divisions. In the early stages of development, individual cells in the embryo are *totipotent*, i.e., each cell retains the capacity to differentiate into any one of the many different cell types in the body. As development proceeds cells become *pluripotent*, i.e., they become more restricted in their capacity to generate different types of descendent cells.

cell differentiation The processes of *cell differentiation* lead to individual cells acquiring specialized structures and functions. Some mature and terminally differentiated cells do not undergo cell division, while others (e.g. osteoblasts, chondroblasts, myoblasts,...) divide actively and thereby act as precursors of terminally differentiated cells. Those precursor cells that stem cells are also capable of self-renewal are known as *stem cells* (e.g. pluripotent hematopoietic stem cells in the bone marrow). The process of differentiation is closely related to morphogenesis *morphogenesis*, the process by which the structure of the cell is modified through regulated growth.

cell cycle While genes clearly have a role in these fundamental processes, by which cells grow, divide and differentiate, this role is primarily to provide information for the molecules whose dynamic interactions determine the *structure* and *function* of cells. The *cell cycle* mitosis is a sequence of events take the cell from division to division (*mitosis*). Progression proliferation through the cell cycle determines *proliferation* (the increase of the number of cells in a population). For example, the essence of cancer is that cells no longer act and react in a regulated fashion within the context of the organ that defines their environment.

pathway map The concept by which interactions of proteins in cell functions are organized are pathways. A *pathway map* exhibits the names of the molecular components, whose interactions govern the basic cell functions. This leads us to a definition of pathways as biochemical networks. One motivation for systems biology is to bring these static diagrams to life by modelling and simulating the biochemical reactions that underlie cell function, development, and disease.

To combine into networks that realize higher levels of organization, such as tissue and organs, cells must communicate. The physical interface between the inside and outside of a cell is comprised, amongst other things, of receptors, which can sense extracellular signals and transduce a signal to the genome where it can effect the transcription of genetic information. The biochemical reactions that relay signals are organized as signal transduction pathways in which regulatory feedback loops play a central role. Many cancer and neurodegenerative diseases are considered a failure of communication at molecular level.

... more to come.

6 The Dynamic Systems Approach

This chapter is to consider mathematical modelling and simulation of pathways, i.e., *networks* of biochemical reactions, focussing on dynamic or transient changes.

networks

We have so far encountered a range of representations including the biologist's graphical pathway map and the biochemist's formal reaction equations that describe the interactions of those components referred to in a pathway map. We hereafter seek a similar approach, devising a graphical representation in form of block diagrams as a representation of mathematical equations. The behavior of a formal system is then in turn visualized through simulation plots, phase planes, and bifurcation diagrams. In many ways both, the experimentalists and theoretician rely on visualizations to help an understanding. This does not come as a surprise if one accepts the philosophical arguments put forward in the first chapters of this text.



Fig. 6.1: Visualization plays an important role in all sciences. The biologist visualizes his understanding with a pathway map. Properties of mathematical models are visualized as simulation plots, phase plane and bifurcation diagrams.

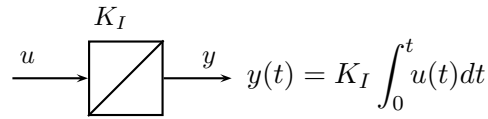
6.1 Pathways as Dynamic Systems

Systems theory and cell biology have enjoyed a long relationship that has received renewed interest in recent years in the context of *systems biology*. The term 'systems' in systems biology comes from *systems theory* or *dynamic systems theory*: Systems biology is defined through the application of systems- and signal-oriented approaches for an understanding of inter- and intra-cellular dynamic processes. The aim of the present text is to review the systems and control perspective of dynamic systems. The biologist's conceptual framework for representing the variables of a biochemical reaction network, and for describing their relationships, are pathway maps. A principal goal of systems biology is to turn these static maps into dynamic models which can provide insight into the temporal evolution of biochemical reaction networks. Towards this end we review the case for differential equation models as a 'natural' representation of causal entailment in pathways. Block-diagrams, commonly used in the engineering sciences, are introduced and compared to pathway maps. The stimulus-response representation of a molecular system is a necessary condition for an understanding of dynamic interactions among the components that make up a pathway. Using simple examples, we show how biochemical reactions are modelled in the dynamic systems framework and visualized using block-diagrams.

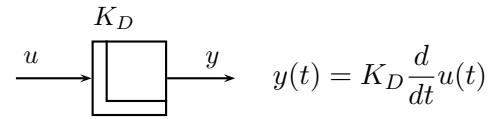
Pathway maps used are for most cases a graphical representation that lacks a standard and for which it is not clear which mathematical model should/could be used to simulate the system. We here introduce a *block diagram* representation of nonlinear dynamic systems, which is an unambiguous translation of the mathematical model. Admittedly it is therefore only suitable for differential equations. The biologist's conception of a pathway map is similar to block diagrams that are widely used in the physical- and engineering

sciences. Arbitrary complex systems can be built up from four basic building blocks:

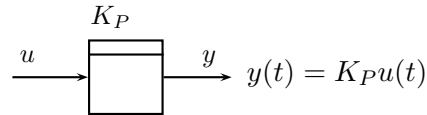
Integrator:



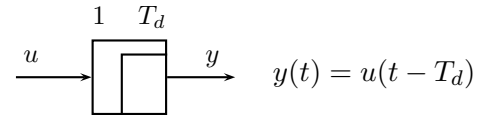
Differentiator:



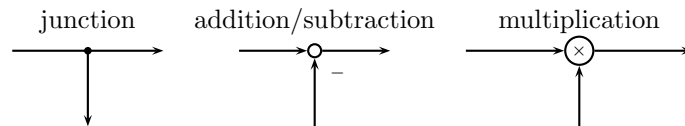
Gain:



Transport Delay:



The most important block we are going to focus on is that of an *integrator*, which describes an accumulation or growth process. The differentiator is simply the reverse operation to the integrator. As alluded to above, the *transport delay* block is of particular importance in simulating the effect of protein translocation, nucleocytoplasmic export and related spatial effects. Block diagrams differ to pathway maps in that they show the processing of *signals*. Block-diagrams are thus a *signal-oriented* approach, an arrow in these diagrams is associated with a variable that is changing over time. The arrows are thus not simply defining ‘associations’, plus/minus signs indicating amplification/inhibition but instead they are numbers that are added or subtracted. Towards this end, blocks or subsystems are connected through signals via the following nodes:



For the addition/subtraction node, if there is no sign, a “+” is assumed. These basic building blocks form a de facto standard for graphical modelling of control systems circuits. While the value and use of diagrammatic representations of pathway models and tools to visualize them are discussed, for example¹, in [Kit02, A⁺04, Laz02], there are no established standards for pathway maps. Given that we are discussing the value of control concepts in pathway modelling, we hereafter consider a couple of well studied examples of biochemical systems and investigate (a) how control block diagram representations might be used and (b) how a control analyst might incorporate feedback loops in pathway models. A discussion of how the more conventional pathway maps can serve as information organizers and simulation guides is discussed in [Koh01].

6.2 The Role of Feedback

Differential equations models are particularly suited to study the role of feedback loops. One of the first biologists who recognized the importance of biological feedback was René Thomas [TD90]. For any process that is to maintain, optimize or adapt a condition or value, information about the ‘is-state’ has to be fed back into the decision on any change that should occur. In other words, *feedback loops* are the basis for any form of regulation and/or control.

¹See also http://discover.nci.nih.gov/kohnk/interaction_maps.html

Control engineers distinguish between two principal kinds of control systems with different purposes: a) *reference tracking*, and b) *disturbance rejection*. We hereafter refer to the first case, where the system is *sensitive* to inputs, as the ability to make changes as required, e.g., to track or follow a reference signal, as *control*. On the other hand, we refer to *regulation* as the maintenance of a regular or desirable state, making the system *robust* against perturbations. Regulation that maintains the level of a variable is also referred to as *homeostasis*. Here we should distinguish two forms of robustness in a control system. The first is robustness against external disturbances (disturbance regulation). In a biochemical pathway, a disturbance might be caused by unwanted cross-talk from a neighboring signalling pathway. The second form of robustness, is one which tolerates parameter changes in a system, without significantly changing the system performance. Both forms of robustness are important properties in understanding pathways.

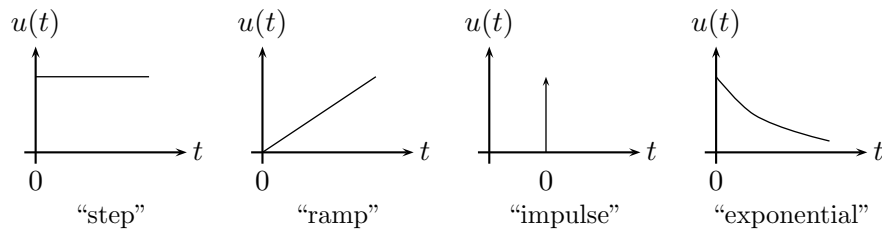


Fig. 6.2: Test signals that can be used to investigate the dynamic behavior of pathways.

A central objective of systems biology is to devise methods that allow the detection and description of feedback loops in pathways [K⁺02, SK04]. An important result from systems theory is that this is only possible through perturbation studies, where the system is stimulated with a well defined signal. Unfortunately, experiments in molecular and cell biology are difficult to set up in a way that suits systems-theoretic approaches. A major hurdle for the success of systems biology arises therefore from the need to conduct expensive, time consuming, complex perturbation experiments.

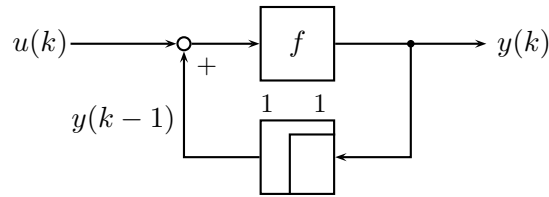
A superficial view of feedback would say that positive feedback is bad (destabilizing) and negative feedback is good (stabilizing). Indeed, the description of the role of feedback often implies that in the absence of negative feedback, a system is unbounded, unstable and not resistant to perturbations. In fact this is not the case, most dynamical systems exist in a stable manner without the need for feedback. A better way in which to describe the role of feedback is as a modifier of the dynamical behavior of a system. Depending upon the nature of the feedback, it can either stabilize, destabilize, sensitize or de-sensitize the behavior of a process. While positive feedback is conventionally associated with destabilization the truth is more complex, and in many circumstances negative feedback can have unwelcome effects. However, in the context of the special dynamical model forms found in pathway modelling, there are certain special dynamical features induced by feedback that are important to understand. The following simple models of accumulation or growth processes will illustrate some of these features as they manifest themselves within cells.

As an initial demonstration of the features associated with feedback, consider the simple model of growth (e.g. of a cell or of a population of molecules in the cell). Let $u(k)$ denote the stimulus of the system at time k and y the response. Let us take the view that the present depends not only on the current state but also on the past, leading to a discrete version of a differential equation, called *difference equation*:

$$y(k) = f(y(k-1), u(k)) . \quad (6.1)$$

where f describes the detailed functional relationship between the stimulus, the past of y

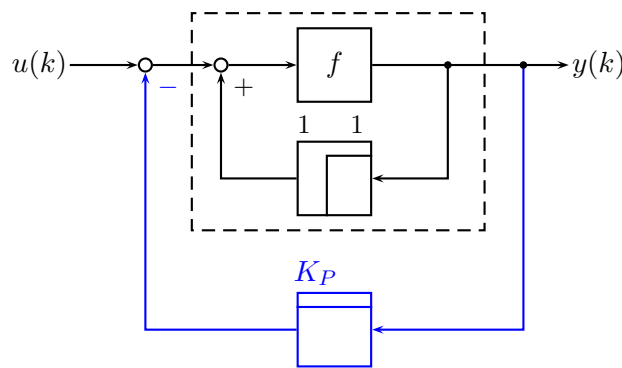
and the current response $y(k)$. One way to illustrate this is by the following block-diagram:



In the diagram the two numbers above the transport delay block denote an amplification of the signal, respectively the unit sampling time delay. For instance, let us look at a linear system, where f is realized by the following law

$$y(k) = u(k) + y(k - 1) .$$

For initial conditions $y_0 = 0$, $u_0 = 0$ if we stimulate the system with a step input, $u(k) = 1$ for $k \geq 1$, a simulation reveals a linearly increasing, unbounded signal (Figure 6.3, left). Whatever the initial conditions, $y_0 \geq 0$, the system is unstable and an unrealistic model for most purposes. Let us therefore see what happens if we add a *negative* feedback loop to the system:



The temporal evolution of the response signal is modelled by the following equation

$$y(k) = (u(k) - K_P \cdot y(k)) + y(k - 1) .$$

A simulation reveals a bounded signal (Figure 6.3, right).

6.3 Tutorial Examples

In the following we present very simple examples of biochemical reactions, which are subsequently translated into a set of mathematical (differential) equations. These in turn maybe related to a standard positive/negative feedback representation drawn from control engineering. In general, we say a component or variable of a system is subject to *negative feedback* when it inhibits its own level of activity. For example, a gene product that acts as a repressor for its own gene is applying negative feedback. Likewise, a component of a system is subject to *positive feedback* when it increases its own level of activity. Through these examples we are going to review the concepts of the biochemist's reaction equation, pathway maps, differential equations and block diagrams.

Returning to our proteolysis example from the introductory section, we generalize it in the context of the framework outlined above. Consider a simple monomolecular reaction where chemical species X is transformed. The change in concentration of X at time

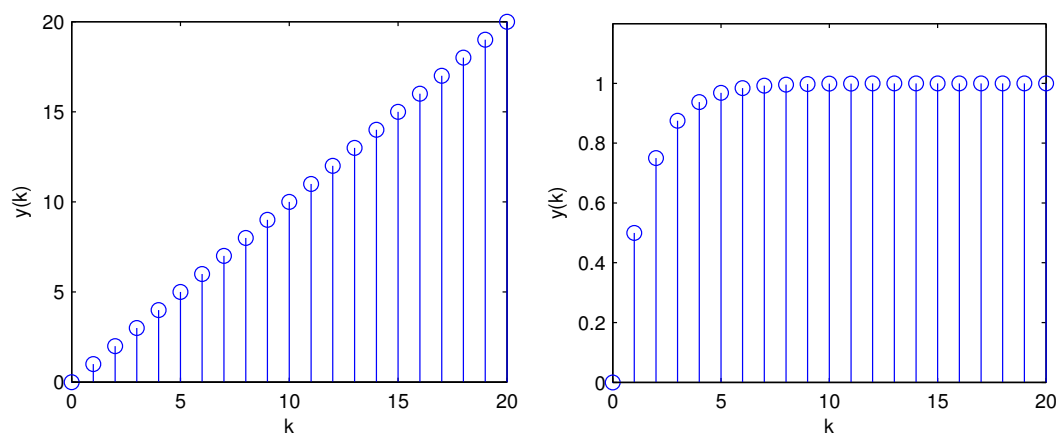
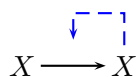


Fig. 6.3: Illustration of the stabilizing effect of a negative feedback loop in a discrete-time system. Left: Unstable system with a positive feedback loop. Right: Negative feedback loop with $K_P = 1$. In both cases the response to unit step input signal is shown.

t depends on the concentration of X at time t in that the rate by which the reaction proceeds is proportional to the concentration at each time instant,

$$\frac{dx(t)}{dt} \propto x(t)$$

with a certain positive *rate constant* k . A diagrammatic representation of this biochemical process illustrates the fact that chemical species X “*feeds back*” on itself:



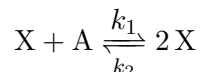
A linear mathematical ODE model of the process is given by

$$\frac{d}{dt}x(t) = k \cdot x(t) .$$

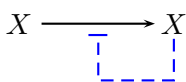
Here X acts as a *substrate* being converted and the *product*. There is *positive feedback* in that the larger the product X , the greater the rate of change by which substrate X is transformed. A simulation of this system reveals the expected unbounded growth in the concentration of X ,

$$x(t) = x_0 \cdot e^{kt} ,$$

where $x_0 = x(t = 0)$ denotes the initial condition. With increasing x , the growth rate dx/dt also increases in this system, leading to an unbounded growth. Next we look at the *autocatalytic* reaction



where for a given X molecule, A facilitates the doubling. A pathway map of this process would be



In pathway maps we use a bar at the end of the arrow to denote an inhibition or negative feedback loop. If A is considered to have a constant concentration, generalizing the law

of mass action, we arrive at the following differential equation model:

$$\begin{aligned}\frac{d}{dt}x(t) &= k_1ax(t) - k_2x^2(t) \\ &= k_1ax(t) \left(1 - \frac{k_2}{ak_1}x(t)\right) .\end{aligned}$$

Why we rewrote the equation in the form given in the second line will be clarified below. In this autocatalytic reaction the ‘product’ has a strong inhibitory effect on the rate at which X is transformed. In order to indicate the internal feedback mechanisms at work in this system, we will label the right-hand bracketed term $(1 - k_2x(t)/(ak_1))$ as a control input variable $u(t)$

$$\frac{d}{dt}x(t) = k_1au(t)x(t) .$$

Because of the product term on the right-hand side this equation is also referred to as a model of a *bilinear system*. If we consider variable x to represent the state of the system, and we write $dx(t)/dt = \dot{x}$ for short, this system becomes

$$\begin{aligned}\dot{x} &= f(x) + g(x)u , & x(t_0) &= x_0 , \\ y &= h(x) .\end{aligned}$$

We can alternatively write:

$$u(x) = \alpha - \beta x ,$$

where the constant α is called the *intrinsic growth rate* of the population and α/β corresponds to the maximum attainable population. The model we thus obtain is specified by the equation

$$\begin{aligned}\frac{dx}{dt} &= \alpha x \left(\frac{\alpha/\beta - x}{\alpha/\beta}\right) \\ &= \alpha x(t) \left(1 - \frac{\beta}{\alpha}x(t)\right) .\end{aligned}\tag{6.2}$$

This model form is called the *logistic growth model* and is equivalent to the autocatalytic reaction introduced above. The model describes the real growth rate as a proportion of the intrinsic growth rate. This proportion however decreases with an increase in the population, leading to a more realistic scenario of a system that remains within bounds (Figure 6.4). Both previous examples, echo the observations made in the discrete-time example of a simple growth process with added negative feedback.

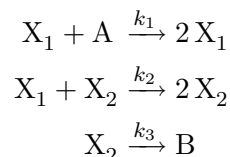
For two molecular species we can generalize the control of the system into

$$\begin{aligned}\dot{x}_1 &= u_1(x_1, x_2) x_1 , \\ \dot{x}_2 &= u_2(x_1, x_2) x_2 .\end{aligned}$$

If we specify for u_1 and u_2 ,

$$\begin{aligned}u_1(x_1, x_2) &= k_1a - k_2x_2 , \\ u_2(x_1, x_2) &= k_2x_1 - k_3 ,\end{aligned}$$

we obtain the well known *Lotka-Volterra model* of two competing populations. If variables x_1 and x_2 correspond to the chemical species X_1 and X_2 , the biochemical representation of this system is



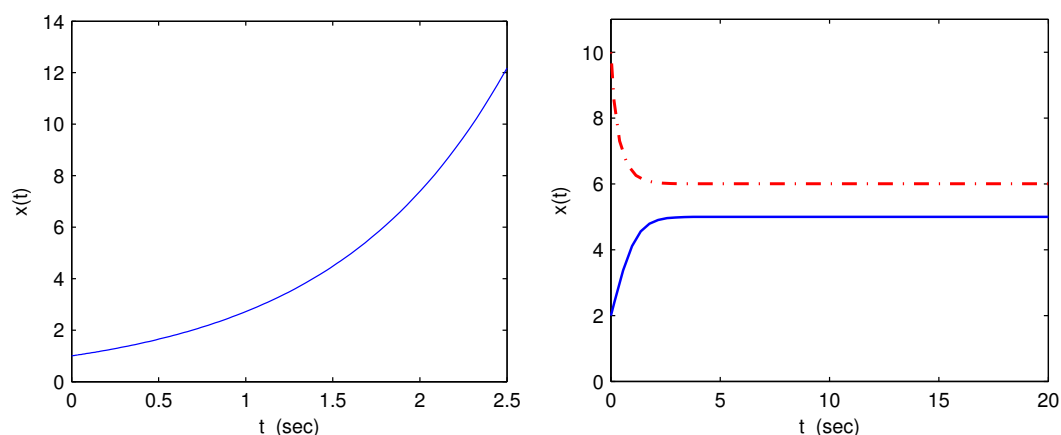
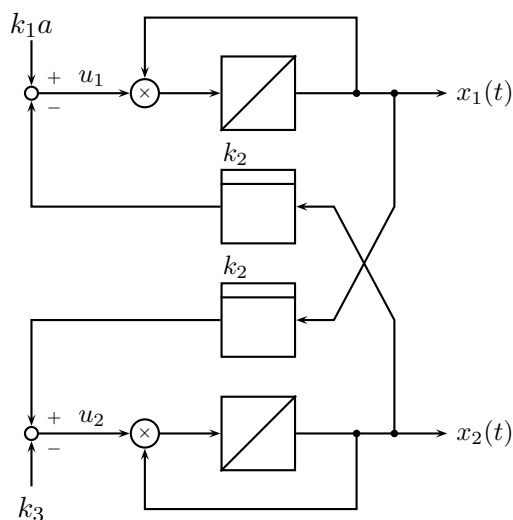
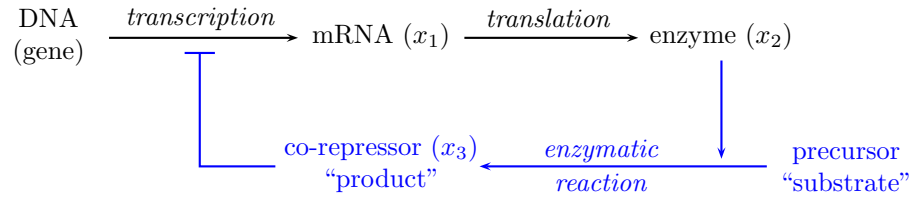


Fig. 6.4: Unbounded and limited growth. Left: Simulation of the monomolecular reaction with positive feedback. Right: Simulation of an autocatalytic reaction (logistic equation) with negative feedback. For the solid line $x_0 = 2$, $a = 2$, $b = 1/2.5$ and for the dashed line $x_0 = 10$, $a = 2$, $b = 1/3$.

where A is maintained at a constant concentration and B corresponds to the degradation of X_2 . The first two reactions are autocatalytic. Compared to the limited growth model from above, this system is capable of showing oscillatory behavior. The block diagram for the Lotka-Volterra model can be drawn directly from those equations:



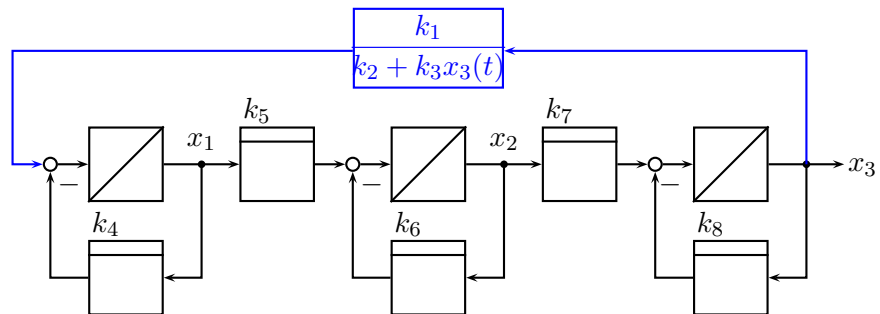
The Lotka-Volterra model of competing species gives an opportunity to discuss the purpose of mathematical models as a mechanism for illuminating basic principles, while not necessarily describing the details of a particular case. Specifically, the Lotka-Volterra model would nowadays be considered an *unrealistic* model for modelling animal population dynamics. However as an abstraction it has proven very *useful*, helping scientists to establish a conceptual approach and ask the right questions [Mur02]. It is in this spirit that models of intracellular dynamics are, or should be, developed in systems biology. The systems considered here are frequently used for an introduction to differential equations. The prototypical biological example of a regulatory system is the protein synthesis model of Jacob and Monod [JM61]. The conceptual model explains how the production of mRNA (x_1), is feedback controlled by a repressor (x_3). A simplified pathway map of this process is shown in the following diagram:



A differential equation model of this regulatory mechanism of protein synthesis is:

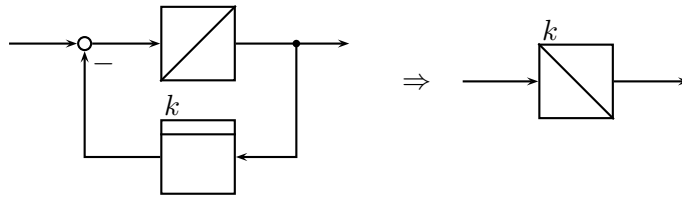
$$\begin{aligned}\frac{d}{dt}x_1 &= \frac{k_1}{k_2 + k_3x_3(t)} - k_4x_1(t) \\ \frac{d}{dt}x_2 &= k_5x_1(t) - k_6x_2(t) \\ \frac{d}{dt}x_3 &= k_7x_2(t) - k_8x_3(t) .\end{aligned}$$

For each of these equations, the last term describes degradation of the molecules. k_5 is the rate of synthesis for the protein that facilitates the production of the co-repressor. Note that there is no minus sign to indicate the negative feedback as in previous examples. The greater x_3 in the numerator of the first term of the rate equation for x_1 , the smaller its contribution towards the rate of change of x_1 . In contrast to the previous example where the feedback was linear, i.e., a simple additive or negative term, in this example the feedback is nonlinear. To illustrate the use of block-diagrams more clearly, let us consider the block-diagram for the Jacob-Monod model of protein synthesis.

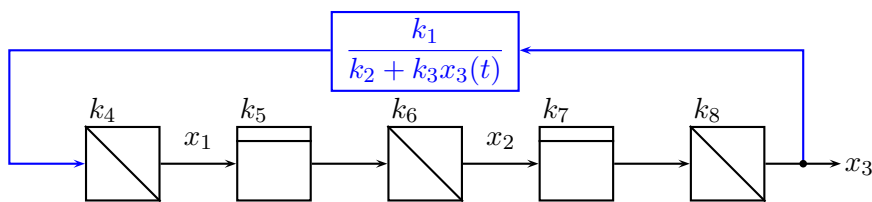


We are now alerted to the fact that *negative feedback* does not necessarily coincide with an explicit form of *negative feedback loop*. Specifically, we have in the block diagram arbitrarily chosen to arrange the figure such that $x_3(t)$ appears as the term fed-back to $x_1(t)$ and that because of the nonlinear form of the feedback it will in fact for small perturbations be negative. The arbitrary nature of the feedback variable is because there is no explicit control input. In such autonomous systems, it is the physical/biological structure that will determine what we (the analyst) chose to call the feedback signal. When the differential equation for x_1 is linearized by Taylor series expansion the $x_3(t)$ appears as a negative feedback term. Whether or not linearization is feasible depends on the system considered. A more comprehensive discussion of this system and further references can be found in [Mur02].

In the block diagram above we have also noticed that degradation is represented by an integrator with a negative feedback loop around it. This motif we can summarize into a single block:



Note that this is not just an arbitrary graphical simplification, the inner structure of the block remains unambiguously defined. That is, we do not lose information or accuracy in presentation by scaling the block diagram in this way. Finally, the protein synthesis model can be simplified to



Although the systems we have considered here are fairly simplistic, the consequences of feedback loops we have observed remain akin for more complex processes. For a related instructive discussion of the dynamic systems approach in biochemical pathway modelling we recommend [TO78, Fal02, TCN03].

6.4 Discussion

Although a pathway or pathway map describes molecules, their physical state and interactions, it is an *abstraction*, with no physical embodiment. A pathway map is thus a model; which proteins and what physical states of the molecules should be considered for experiments and the model is what we call the *art* of modelling.

Feedback loops are the essence of control and regulation, for only if information about the consequences of some output is fed back, the system can adjust itself or respond in an appropriate way. Using ordinary differential equations to model biochemical networks, we have shown that feedback loops can stabilize and destabilize a system, keep its variables and signals bounded, they can make the system robust against perturbations, they allow the system to adapt to changes, or track an input stimulus.

Another relevant feature of control systems is that they have specific intent, and control systems analysts have theories for understanding intent and methods for achieving a required intent or purpose [YHS00]. In a modelling framework, the causal structure of a control system provides a framework for the dynamical manipulation of information with a purposeful objective. This is topical and relevant in the light of recent discussion of the value of systems biology compared with mathematical biology [Lan04]. In this same spirit, feedback loops lie at the heart of the causal/purposeful mechanisms of control and regulation in dynamic systems. Specifically, it is only if information about the consequences or some output is fed back, can the system automatically adjust itself or respond in an appropriate way. Feedback is not always beneficial, for feedback loops can stabilize or destabilize a system. Feedback can keep a system's variables and signals bounded, or it can induce oscillations or unbounded growth. Likewise, feedback loops can make a system robust against perturbations, but at the same time they allow the system to adapt to changes, or track an input stimulus.

Apart from the role of feedback loops, we surveyed alternative and complementary representations and visualizations, including the biochemist's reaction equations, the mathematician's differential equation models, the control engineer's block diagrams and the biologist's pathway maps. Block diagrams are well established in the engineering sciences as a means of describing dynamic systems in general. Through the integrators used, these diagrams are inherently linked to differential equation models and are therefore less general than those molecular interaction maps [A⁺04], commonly used to visualize relationships in pathways. On the other hand, block diagrams are a direct and unambiguous visualization of the mathematical model. These diagrams also do not explicitly represent spatial aspects. While the transport of a protein from the nucleus to the cytosol can be modelled, compartments are realized by introducing more than one variable in the model for the same molecular species in different regions of the cell. For the analysis of the nonlinear differential equations models we only used time plots. Visualization is no less important to theoreticians than it is to biologists and so there are a range of tools available we have not mentioned here, including stimulus-response curves, phase-plane and bifurcation analysis (e.g. [GH83, KG95, Mur02, TCN03]). For an application of these mathematical tools applied to a model of the yeast cell cycle, we refer to the expositions of Novak and Tyson [Tys01, TN01, TCN03]. The building block approach to an understanding of systems, when associated with purpose, is very similar to the causality principles that are embedded in the dynamical system modelling methods of control engineering. One question we investigated here was whether the control engineer's proficiency with block diagram models and modular representations can contribute to systems biology by facilitating the translation of biological concepts into mathematical representations.

The cell is made up of molecules, like a car is made up from plastic and metal. But a soup of molecules is no more a cell than a heap of plastic and metal is a car. To understand the *functioning* and *function* of a cell we need to know the *relations* and *interactions* of the components that constitute it. If the central dogma of systems biology is that *it is dynamics that determines biological function*, we would refine this statement and argue that *the dynamical manifestation of feedback determines the development and maintenance of biological processes*.

6.5 Phase-Plane Analysis

Phase-plane analysis is an important technique in studying the behavior of linear and nonlinear dynamic systems. It is a graphical approach which allows the study of the behavior of the system for a large range of initial conditions. It is for this reason that this method is referred to as providing a *qualitative analysis* of the dynamic system. For linear systems the approach is truly global, while for nonlinear systems it is only locally applicable. The main purpose of this section is to illustrate the diversity of behavior nonlinear dynamics can generate and to introduce a tool for its analysis in the plane. For a more comprehensive description of phase-plane analysis see one of the many books available on differential equations (e.g. [BD01, EP01]).

The general form of systems considered for phase-plane analysis is

$$\frac{dx}{dt} = f(x, y) , \quad \frac{dy}{dt} = g(x, y) . \quad (6.3)$$

A system in which t does not explicitly occur in f and g is called an *autonomous system*. The two differential equations determine the velocity of two variables x and y moving in the xy -plane referred to as the *phase-plane*. As time increases, the system state moves along a curve in the xy -plane, called the trajectory. While for non-autonomous systems trajectories could cross in the plane, for autonomous systems the trajectories cannot cross.

The totality of all trajectories describes the *phase portrait*. Points (x, y) of the trajectory phase portrait for which

$$f(x, y) = g(x, y) = 0$$

are called *critical points* or *fixed points*, often also referred to as steady-states or equilibrium critical points². To see the appearance the phase portrait, we consider a *direction field* on a grid of points in the \mathbb{R}^2 plane and determine velocity vectors defined by

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{g(x, y)}{f(x, y)} .$$

As a first simple example let us consider the system

$$\frac{dx}{dt} = y \quad \text{and} \quad \frac{dy}{dt} = 4x .$$

The only fixed point for this system is the origin, $(0, 0)$, of the plane. We can solve the system by separation of variables:

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{4x}{y}$$

which implies

$$\int y \, dy = \int 4x \, dx \quad \text{i.e.,} \quad \frac{y^2}{2} = 2x^2 + c$$

leading to trajectories that are hyperbolas:

$$4x^2 - y^2 = c .$$

The phase-plane and some sample solutions are shown in Figure 6.5.

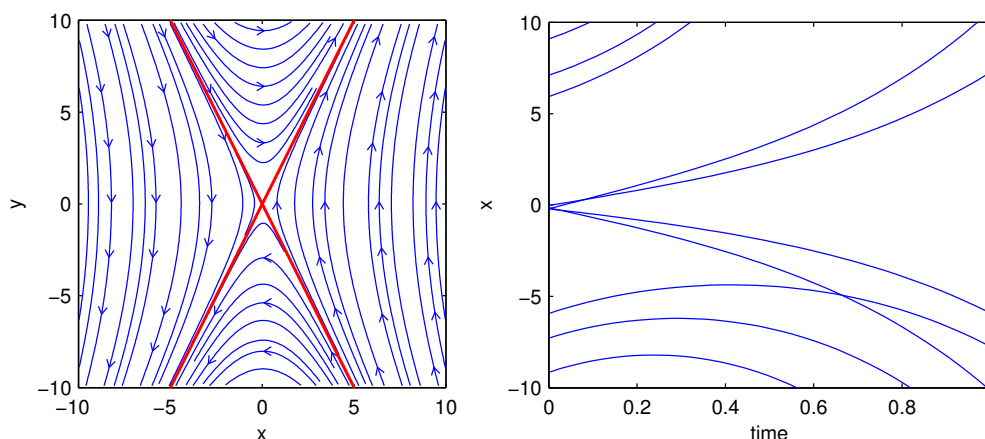


Fig. 6.5: Left: Phase portrait of the system $dx/dt = y$, $dy/dt = 4x$. The fixed point $(0, 0)$ is said to be *unstable* because trajectories close to the fixed point move away from it. From the appearance of graph, the fixed point is also referred to as a *saddle point*. Right: Sample solutions for $x(t)$ for different initial conditions.

A second motivating example, leading to trajectories that are circles is

$$\frac{dx}{dt} = y \quad \text{and} \quad \frac{dy}{dt} = -x ,$$

²In the engineering literature fixed points are also referred to as *equilibrium points*. In the context of biochemical networks in cells this can however lead to confusion.

The only fixed point is again $(0, 0)$. By separation of variables

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{-x}{y}$$

which implies

$$\int y \, dy = - \int x \, dx \quad \text{i.e.,} \quad \frac{y^2}{2} = -\frac{x^2}{2} + c$$

leading to circular trajectories:

$$x^2 + y^2 = c .$$

The phase-portraits is shown in Figure 6.6.

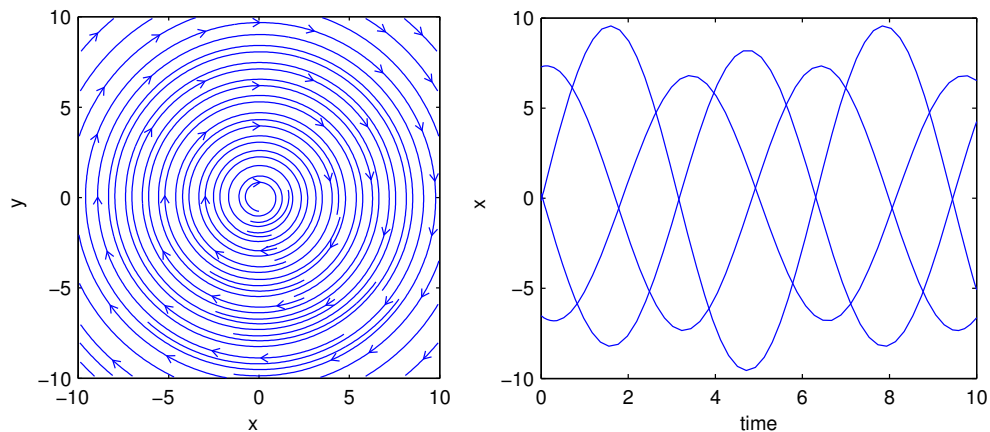


Fig. 6.6: Phase portrait of the system $dx/dt = y, dy/dt = -x$. The fixed point $(0, 0)$ of this system is *stable*, i.e., all trajectories close to it remain close for all t . This fixed point is called a *centerpoint*.

From these two examples we now consider a more comprehensive survey of linear dynamics, followed by nonlinear systems. If the derivatives are linear functions of the variables, we deal with a linear (autonomous) system:

$$\left. \begin{aligned} \frac{dx}{dt} &= a \cdot x + b \cdot y \\ \frac{dy}{dt} &= c \cdot x + d \cdot y \end{aligned} \right\} \quad (6.4)$$

In matrix form we can rewrite this system as

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} .$$

The matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

system matrix is called the *system matrix* of coefficients. If the determinant of A

$$\det(A) \equiv |A| = ad - bc ,$$

is nonzero, there is a unique solution to the equations. For a linear system, the origin of the phase-plane is this fixed point. If $\det(A) = 0$ there either aren't any solutions or there are infinitely many. In this case have to solve the system of linear algebraic equations

$$\begin{aligned} a \cdot x + b \cdot y &= 0 \\ c \cdot x + d \cdot y &= 0 . \end{aligned}$$

The solutions to the linear system differential equations are

$$\begin{bmatrix} x \\ y \end{bmatrix} = c_1 \vec{v}_1 e^{\lambda_1 t} + c_2 \vec{v}_2 e^{\lambda_2 t} ,$$

where \vec{v}_1 and \vec{v}_2 are the *eigenvectors* of matrix A corresponding to *eigenvalues* λ_1 and λ_2 of A . The eigenvectors and eigenvalues are found by asking whether there exists a nonzero vector \vec{v} such that the result $A\vec{v}$ is a simple scalar multiple of \vec{v} . With the eigenvalues on the diagonal of a diagonal matrix Λ and the corresponding eigenvectors \vec{v} forming the columns of a matrix L , we have

$$AL = \Lambda L$$

If L is nonsingular³, this becomes the *eigenvalue decomposition* of A :

$$A = L\Lambda L^{-1} .$$

The eigenvectors and values have a geometric interpretation. The length $|A\vec{v}|$ of vector $A\vec{v}$ is $\pm\lambda|\vec{v}|$. The multiplication of \vec{v} by the matrix A expands or contracts vector \vec{v} , while a positive eigenvalue preserves its direction, a negative value reverses it. In application of dynamic systems in the plane, the eigenvalue corresponds to the speed of response, while the eigenvector determines the principal direction. A line in the phase-plane that is not crossed by any trajectory is called a *separatrix*. Eigenvectors determine the separatrices.

The eigenvalues λ_1 and λ_2 of A are determined as solutions of the *characteristic equation*

$$\det(A - \lambda I) = \det \begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix} = 0$$

where I is the identity matrix with ones on the diagonal and zeros elsewhere. The eigenvalues are thus the roots of the *characteristic polynomial* $\det(A - \lambda I)$

$$(a - \lambda)(d - \lambda) - bc = \lambda^2 - (a + d)\lambda + (ad - bc) = 0 .$$

The constant term is equal to $\det(A)$ and the coefficient $(a + d)$ corresponds to the *trace* of A , denoted $\text{tr}(A)$. Let $A = [a_{ij}]$ be an $n \times n$ matrix, the trace of A is defined to be the sum of the diagonal entries $\text{tr}(A) = \sum_{i=1}^n a_{ii}$. The eigenvalues are then given by

$$\begin{aligned} \lambda_{1,2} &= \frac{1}{2} \left(\text{tr}(A) \pm \sqrt{[\text{tr}(A)]^2 - 4 \det(A)} \right) \\ &= \frac{1}{2} \left((a + d) \pm \sqrt{(a + d)^2 - 4(ad - bc)} \right) . \end{aligned}$$

Given the eigenvalues, the eigenvectors can be calculated by

$$\vec{v}_i = \frac{1}{\sqrt{1 + p_i^2}} \begin{bmatrix} 1 \\ p_i \end{bmatrix}$$

where

$$p_i = \frac{\lambda_i - a}{b}, \quad b \neq 0, \quad i = 1, 2 .$$

The sign of $[\text{tr}(A)]^2 - 4 \det(A)$ determines whether the eigenvalues λ are:

1. complex with nonzero imaginary part if $[\text{tr}(A)]^2 - 4 \det(A) < 0$

³A matrix is *singular* if its determinant is zero. It is *regular* if the determinant is nonzero, and in which case an inverse exist.

2. real and distinct if $[\text{tr}(A)]^2 - 4 \det(A) > 0$
3. real and repeated if $[\text{tr}(A)]^2 - 4 \det(A) = 0$.

If in the first case $[\text{tr}(A)]^2 - 4 \det(A) < 0$, then the real part of the eigenvalues is $\text{tr}(A/2) \equiv (a + d)/2$, determining a

- ... *spiral sink* if $\text{tr}(A) < 0$
- ... *spiral source* if $\text{tr}(A) > 0$
- ... *center* if $\text{tr}(A) = 0$.

If for the second case above $[(\text{tr}(A))^2 - 4 \det(A)] > 0$, if $\det(A) < 0$ we have a *saddle* and for $\text{tr}(A) > 0$ and $\det(A) > 0$ we have a *source*. In terms of the eigenvalues of matrix A we can distinguish four cases, discussed hereafter.

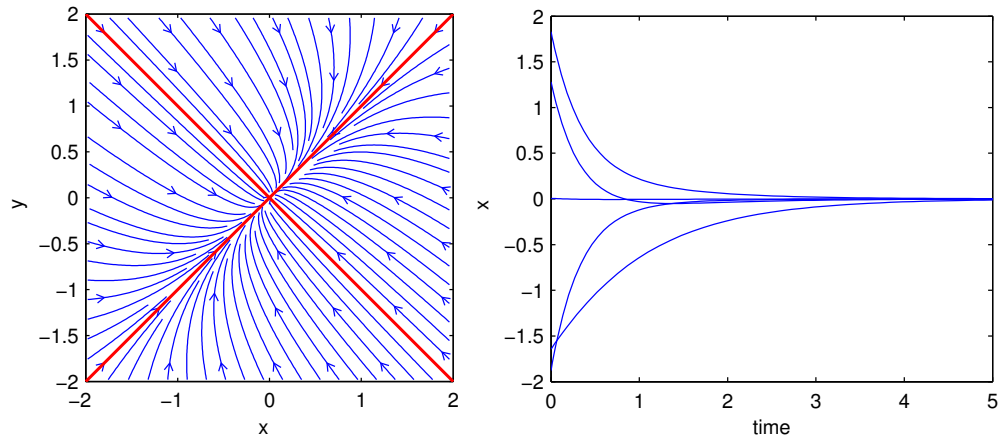


Fig. 6.7: Left: Phase portrait of the system $dx/dt = -2x + y$, $dy/dt = x - 2y$. Right: Sample solutions for x . Eigenvalues $\lambda_1 = -3, \lambda_2 = -1$. The critical point $(0,0)$ is called a nodal sink. It is asymptotically stable node.

Case 1: Unequal, real eigenvalues of the same sign: The general solution of (6.4) is

$$[x \ y]^T = c_1 \vec{v}_1 e^{\lambda_1 t} + c_2 \vec{v}_2 e^{\lambda_2 t} .$$

The eigenvalues can be either positive or negative. In Figure 6.7 the case for

$$A = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$$

with $\lambda_1 < \lambda_2 < 0$ is shown. From the general solution we see that both variables approach zero as time goes to infinity, regardless of the constants c_1 and c_2 . This means that all solutions approach the critical point at the origin as $t \rightarrow \infty$. The eigenvectors are in this case $\vec{v}_1 = [0.71 \ -0.71]^T$ and $\vec{v}_2 = [0.71 \ 0.71]^T$, forming a cross through the origin. Notice that all solutions approach the critical point tangent to \vec{v}_2 , except those solutions

nodal sink

that start on the line through \vec{v}_1 . This critical point is called a *node* or *nodal sink*. If λ_1 and λ_2 are both positive and $0 < \lambda_2 < \lambda_1$, the trajectories in the phase-plane have the same pattern as in Figure 6.7 but they are moving away from the critical point. $x(t)$ and $y(t)$ grow exponentially in this case. The critical point is, in this case, called *nodal source*.

nodal source

Case 2: Unequal, real eigenvalues of the opposite sign: The general solution of (6.4) is again

$$[x \ y]^T = c_1 \vec{v}_1 e^{\lambda_1 t} + c_2 \vec{v}_2 e^{\lambda_2 t} .$$

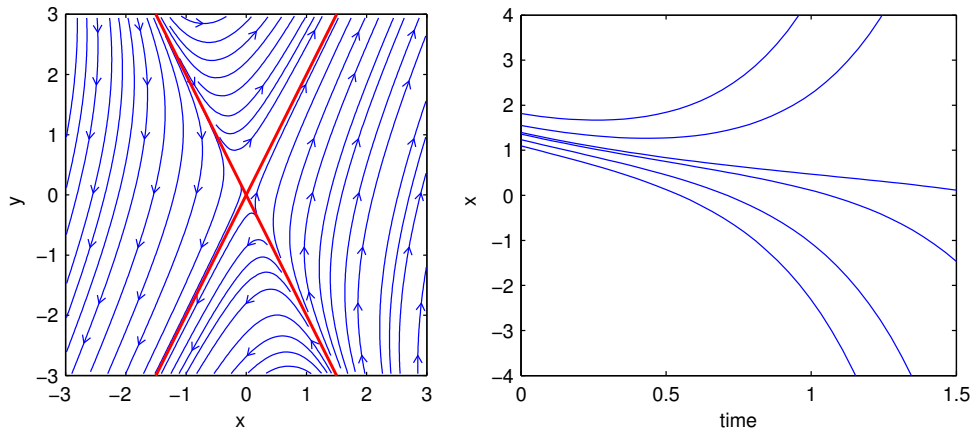


Fig. 6.8: Left: Phase portrait of the system $dx/dt = x + y$, $dy/dt = 4x + y$. Right: Sample solutions for x . Eigenvalues $\lambda_1 = 3, \lambda_2 = -1$. The critical point $(0,0)$ is called a saddle point.

In Figure 6.8 the system with matrix

$$A = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix}$$

is illustrated. The eigenvectors for this system are $\vec{v}_1 = [0.45 \ 0.89]^T$ and $\vec{v}_2 = [-0.45 \ 0.89]^T$. The eigenvectors are again forming a cross through the origin. If a solution starts on the line along \vec{v}_1 (which goes from the bottom left to top right corner of the plane), it will remain there for all time and $c_2 = 0$. The only solutions that approach the critical point in the origin are those that start on the line determined by \vec{v}_2 . For all other initial conditions the positive exponential term in the solution will eventually dominate. The origin is called a *saddle point*. The origin is also an unstable fixed point since no solution will remain saddle point there.

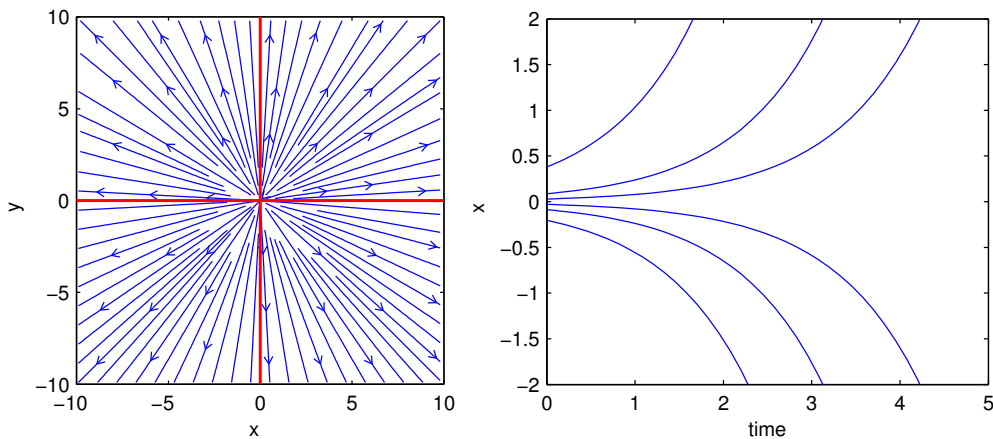


Fig. 6.9: Left: Phase portrait of the system $dx/dt = x$, $dy/dt = y$. Right: Sample solutions for x . Eigenvalue $\lambda = 1$. The critical point is called a proper node.

Case 3: Equal eigenvalues: In case $\lambda_1 = \lambda_2 = \lambda$, we have to distinguish two cases. **Two independent eigenvectors:** The general solution is

$$[x \ y]^T = c_1 \vec{v}_1 e^{\lambda t} + c_2 \vec{v}_2 e^{\lambda t} .$$

In this case the ratio y/x is only dependent on \vec{v}_1, \vec{v}_2 and independent of t . The trajectories generate a star-shaped pattern of the phase-plane. The fixed point is called a *proper*

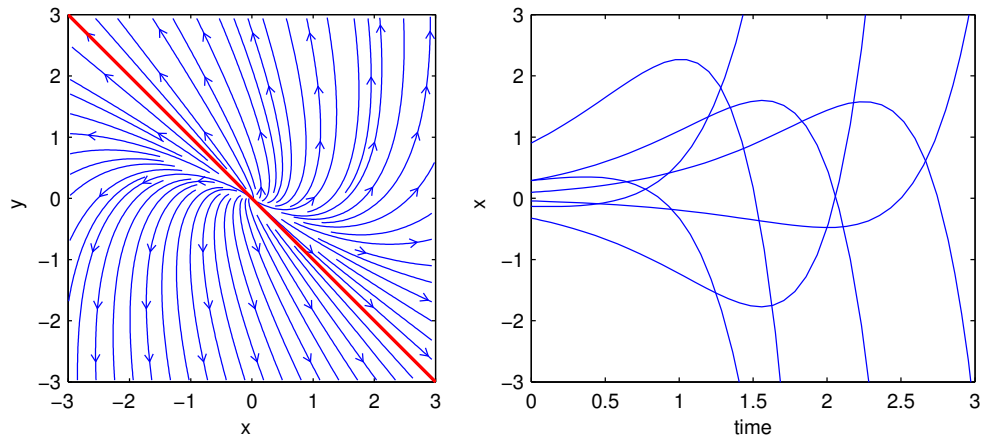


Fig. 6.10: Left: Phase portrait of the system $dx/dt = x - y$, $dy/dt = x + 3y$. Right: Sample solutions for x . Eigenvalue $\lambda = 2$. There is only one independent eigenvector. The critical point is called an improper node.

node or *star point*. Figure 6.9 illustrates a system with eigenvectors $\vec{v}_1 = [1 \ 0]^T$ and $\vec{v}_2 = [0 \ 1]^T$. The node is asymptotically stable or unstable, depending on whether the eigenvalue is negative or positive.

One independent eigenvector: The general solution is in this case

$$[x \ y]^T = c_1 \vec{v}_1 e^{\lambda t} + c_2 (\vec{v}_1 t e^{\lambda t} + \vec{v}_2 e^{\lambda t})$$

where \vec{v}_1 is the one independent eigenvector and \vec{v}_2 denotes the generalized eigenvector associated with the repeated eigenvalue. For a large t the dominant term is $c_2 \vec{v}_1 t e^{\lambda t}$, which means that for $t \rightarrow \infty$ all trajectories approach the origin tangent to the line through the eigenvector. The orientation of the trajectories depends on the relative positions of \vec{v}_1 and \vec{v}_2 . Figure 6.10 illustrates one situation for a system with eigenvectors $\vec{v}_1 = [-0.71 \ 0.71]^T$ and $\vec{v}_2 = [-0.71 \ 0.71]^T$. When a repeated eigenvalue has only one independent eigenvector, then the critical point is called an *improper node* or *degenerate node*. An improper node is asymptotically stable or unstable, depending on whether the eigenvalue are negative or positive.

improper node

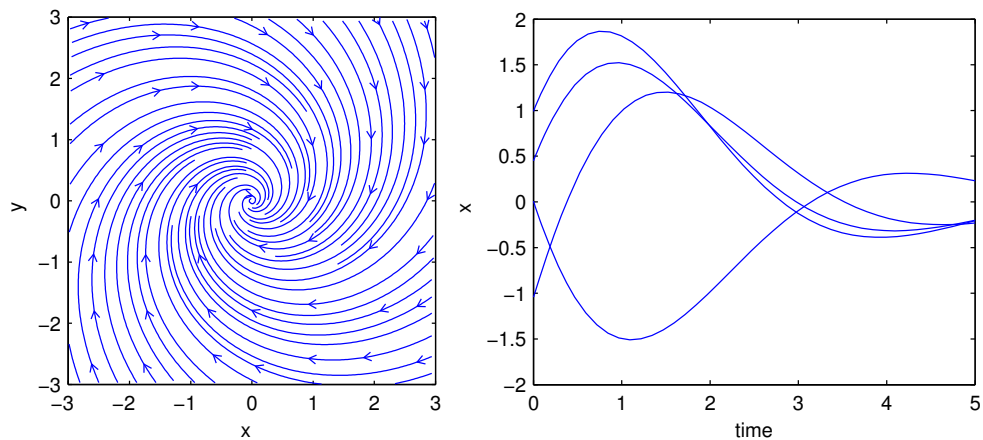


Fig. 6.11: Left: Phase portrait of the system $dx/dt = -0.5x + y$, $dy/dt = -x - 0.5y$. Right: Sample solutions for x . Eigenvalues $\lambda_1 = -0.5 + i$, $\lambda_2 = -0.5 - i$. Since the real part is negative the trajectories spiral inwards.

Case 4: Complex eigenvalues: In this case the eigenvalues are $a \pm ib$, where a is the real part and b denotes the imaginary part. The critical point is called a *spiral point*, *spiral sink* or *spiral source*. Whenever $a \neq 0$, the trajectories are spirals. They are directed

spiral point

inward or outward, depending on whether a is positive or negative. Figures 6.11 and 6.12 provides an illustration.

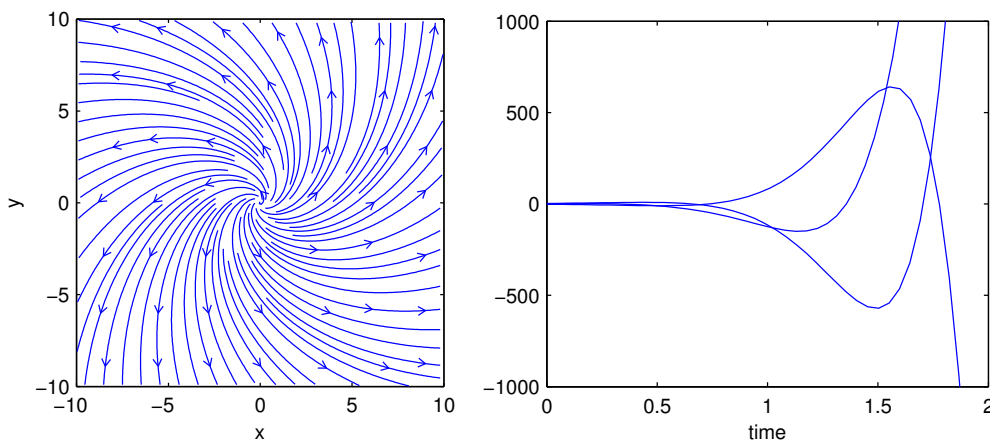


Fig. 6.12: Left: Phase portrait of the system $dx/dt = 4x - 3y$, $dy/dt = 3x + 4y$. Right: Sample solutions for x . Eigenvalues $\lambda_{1,2} = 4 \pm 3i$. Since the real part is positive the trajectories spiral outwards.

Case 5: Pure imaginary eigenvalues: In case $a = 0$ for the eigenvalues, the trajectories become circles around the origin, that are traversed clockwise if $b > 0$ and anticlockwise if $b < 0$. Figure 6.6 provides an illustration for the system $dx/dt = y$, $dy/dt = -x$, with eigenvalues $\lambda_{1,2} = \pm i$.

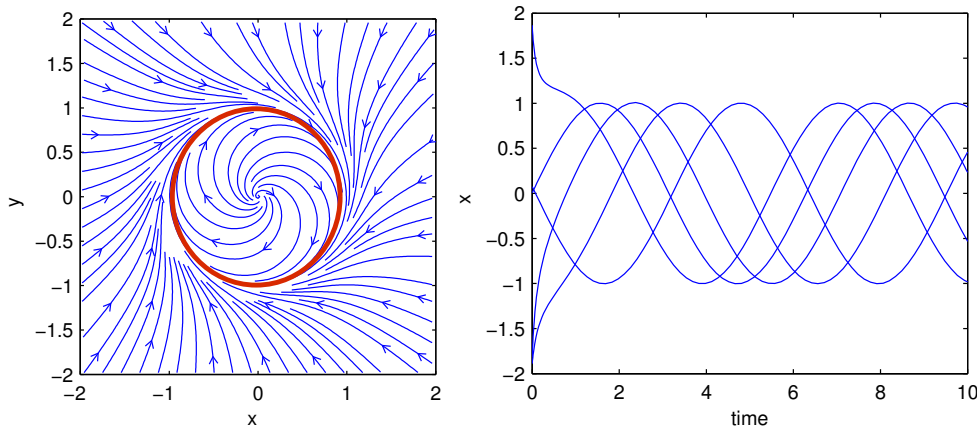


Fig. 6.13: Illustration of an autonomous nonlinear system, which is almost linear around the origin but displays a limit cycle.

We have summarized the dynamic properties or *stability* of the linear system $[x \ y]^T = A[x \ y]^T$ in Table 6.1. Before we continue with study nonlinear dynamics using phase-plane analysis, we look at a particular nonlinear autonomous system which is almost linear around the origin:

$$\begin{aligned}\frac{dx}{dt} &= y + x - x(x^2 + y^2) \\ \frac{dy}{dt} &= -x + y - y(x^2 + y^2) .\end{aligned}$$

The only critical point of this system is the origin $(0, 0)$. The corresponding linear system has the system matrix

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} .$$

and eigenvalues $\lambda_{1,2} = \pm i$, which suggests the origin is an unstable spiral point for the linear as well as the nonlinear system. However, rather than spiralling out completely, limit cycle as the linear analysis would suggest, the system exhibits what is known as a *limit cycle*. Figure 6.13 illustrates the limit cycle behavior.

Table 6.1: Stability of the linear system $[\dot{x} \ \dot{y}]^T = A[x \ y]^T$ with $\det(A) \neq 0$.

Eigenvalues	Type of Critical Point	Stability
$\lambda_1 > \lambda_2 > 0$	Node	Unstable
$\lambda_1 < \lambda_2 < 0$	Node	Asymptotically stable
$\lambda_2 < 0 < \lambda_1$	Saddle point	Unstable
$\lambda_1 = \lambda_2 > 0$	Proper or improper node	Unstable
$\lambda_1 = \lambda_2 < 0$	Proper or improper node	Asymptotically stable
$\lambda_{1,2} = \alpha \pm i\beta$	Spiral point	
$\alpha > 0$		Unstable
$\alpha < 0$		Asymptotically stable
$\lambda_{1,2} = \pm i\beta$	Center	Stable

6.6 Nonlinear Dynamics

As a motivating example for nonlinear systems let us find the trajectories of the following system with two coupled nonlinear equations:

$$\frac{dx}{dt} = 4 - 2y, \quad \frac{dy}{dt} = 12 - 3x^2.$$

To find critical points we set the derivatives to zero:

$$4 - 2y = 0, \quad 12 - 3x^2 = 0$$

and find that there are two critical points at $(-2, 2)$ and $(2, 2)$. For the trajectories we write

$$\frac{dy}{dx} = \frac{12 - 3x^2}{4 - 2y}.$$

Separation of variables and integration provides us with the solution

$$4y - y^2 - 12x + x^3 = c$$

where c is some arbitrary constant. Figure 6.14 illustrates the phase portrait.

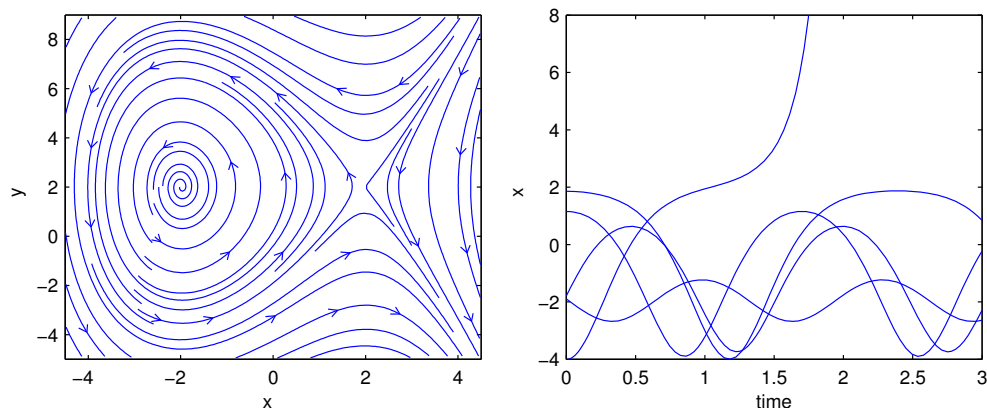


Fig. 6.14: Phase portrait of the nonlinear system $dx/dt = 4 - 2y$, $dy/dt = 12 - 3x^2$.

The phase-plane analysis introduced above does work for nonlinear systems by linearizing a system around a point of interest. The analysis in this case applies locally. Points of particular interest are critical- or fixed points. Let us denote such point of particular interest in the plane as (x^*, y^*) .

Linearizing a nonlinear system is done in the neighborhood of the fixed points using a Taylor expansion of $f(x, y)$ and $g(x, y)$. The Taylor expansion for a function of two variables $f(x, y)$ is given by Taylor expansion

$$\begin{aligned} f(x, y) = f(x^*, y^*) &+ \frac{\partial f}{\partial x} \Big|_{x^*, y^*} (x - x^*) + \frac{\partial^2 f}{\partial x^2} \Big|_{x^*, y^*} \frac{(x - x^*)^2}{2!} + \frac{\partial^3 f}{\partial x^3} \Big|_{x^*, y^*} \frac{(x - x^*)^3}{3!} \\ &+ \frac{\partial f}{\partial y} \Big|_{x^*, y^*} (y - y^*) + \frac{\partial^2 f}{\partial y^2} \Big|_{x^*, y^*} \frac{(y - y^*)^2}{2!} + \frac{\partial^3 f}{\partial y^3} \Big|_{x^*, y^*} \frac{(y - y^*)^3}{3!} + \dots \end{aligned} \quad (6.5)$$

If we neglect terms higher than first order the Taylor expansion is

$$f(x, y) \approx f(x^*, y^*) + \frac{\partial f}{\partial x} \Big|_{x^*, y^*} (x - x^*) + \frac{\partial f}{\partial y} \Big|_{x^*, y^*} (y - y^*) . \quad (6.6)$$

Introducing new variables u, v ,

$$u \doteq x - x^* , \quad v \doteq y - y^* ,$$

we can write for the expansion of (6.14),

$$\begin{aligned} \frac{du}{dt} &= au + bv + \dots , \\ \frac{dv}{dt} &= cu + dv + \dots , \end{aligned} \quad (6.7)$$

where

$$\begin{aligned} a &= \frac{\partial f}{\partial x} \Big|_{x^*, y^*} , & b &= \frac{\partial f}{\partial y} \Big|_{x^*, y^*} , \\ c &= \frac{\partial g}{\partial x} \Big|_{x^*, y^*} , & d &= \frac{\partial g}{\partial y} \Big|_{x^*, y^*} . \end{aligned} \quad (6.8)$$

The trick is then to assume that in the neighborhood of the fixed points, the higher-order terms in (6.7) are small enough to be neglected. If we collect the partial derivatives in matrix form this leads us to what is called the *Jacobian matrix*: Jacobian matrix

$$J^* = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{bmatrix} , \quad (6.9)$$

with the partial derivatives evaluated at (x^*, y^*) , treating in each case the other variable as a constant. Therefore in the neighborhood of a fixed point, the nonlinear system (6.14) can be approximated by the set of linear equations

$$\begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} . \quad (6.10)$$

The dynamics in the neighborhood of (x^*, y^*) are now determined by the eigenvalues of the system matrix, as illustrated above.

For more examples, we return to the population models referred to earlier in the text. On page 42 of Chapter 3 we considered a simple population model

$$\frac{dS(t)}{dt} = (k^+ - k^-)S(t) , \quad (3.5)$$

where $S(t)$ denotes the molecular population size at time t , k^+ the formation or birth rate and k^- the decay or death rate. In systems and control theory or the study of differential equations such model is more commonly written using the following notation

$$\frac{dx(t)}{dt} = u(t)x(t) , \quad (6.11)$$

where $u(t)$ may be considered a *control input variable*. Because of the product of the state variable and input variable, $u(t)x(t)$, such a system is also called a *bilinear system*, and which is an example of the more general system

$$\begin{aligned} \dot{x} &= \phi(t, x(t)) + u(t)g(t, x(t)) , & x(t_0) &= x_0 \\ y(t) &= h(t, x(t)) . \end{aligned} \quad (6.12)$$

In many realistic situations $u(t)$ may depend on $x(t)$, which in effect leads to some population constraints. A simple example of such a *feedback control loop* is the following definition for u

$$u(x(t)) = a - bx(t) ,$$

where for a population model, a may be considered as the *intrinsic* growth rate such that a/b corresponds to the maximum population level that can be reached [Ton90]. The input-output description and the description of feedback loops is central to the control engineering approach.

We can extend the idea of the bilinear formulation (6.11) to two molecular species x and y ,

$$\frac{dx}{dt} = u_1(x, y)x , \quad \frac{dy}{dt} = u_2(x, y)y , \quad (6.13)$$

where the dependence on t is suppressed to simplify the notation. If we define

$$\begin{aligned} u_1(x, y) &= \alpha - \beta y \\ u_2(x, y) &= \gamma x - \delta , \end{aligned}$$

we obtain the well known Lotka-Volterra model. This model has been rather useful, certainly for educational purposes. We encountered this system in Section 4.2, on page 116 and will further discuss it hereafter. But first let us consider the phase-portrait of a bilinear system with $u_1(x, y) = y + y/x$ and $u_2(x, y) = x + 3x/y$:

$$\begin{aligned} \frac{dx}{dt} &= xy + y \\ \frac{dy}{dt} &= xy + 3x . \end{aligned}$$

The system has two critical points, a trivial fixed point (x_1^*, y_1^*) at $(0, 0)$ and another at $(x_2^*, y_2^*) = (-1, -3)$. The Jacobian matrix is

$$J^* = \begin{bmatrix} y_{1,2}^* & x_{1,2}^* + 1 \\ y_{1,2}^* + 3 & x_{1,2}^* \end{bmatrix} .$$

Considering first the trivial fixed point. The eigenvalues are $\lambda_1 = -\sqrt{3}$, $\lambda_2 = \sqrt{3}$, with associated eigenvectors $\vec{v}_1 = [-0.5 \ 0.866]^T$ and $[0.5 \ 0.866]^T$. For the nontrivial fixed point the eigenvalues are $\lambda_1 = -3$, $\lambda_2 = -1$, with associated eigenvectors $\vec{v}_1 = [1 \ 0]^T$ and $[0 \ 1]^T$. Figure 6.15 illustrates the two phase-portraits of the system linearized around the critical points and Figure 6.16 shows the combined phase portrait of the nonlinear system.

What the Lotka-Volterra model is for the theoretician is the enzyme kinetic reaction for the experimentalist. The vast majority of biochemical reactions in a cell are catalyzed

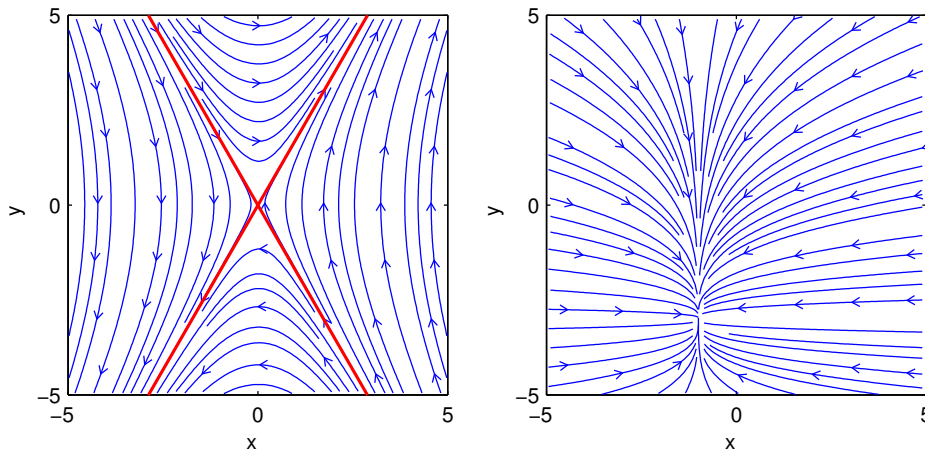


Fig. 6.15: Phase portrait of the nonlinear system $xy + y$, $dy/dt = xy + 3x$, around the critical points $(0, 0)$ on the left and $(-1, -3)$ on the right.

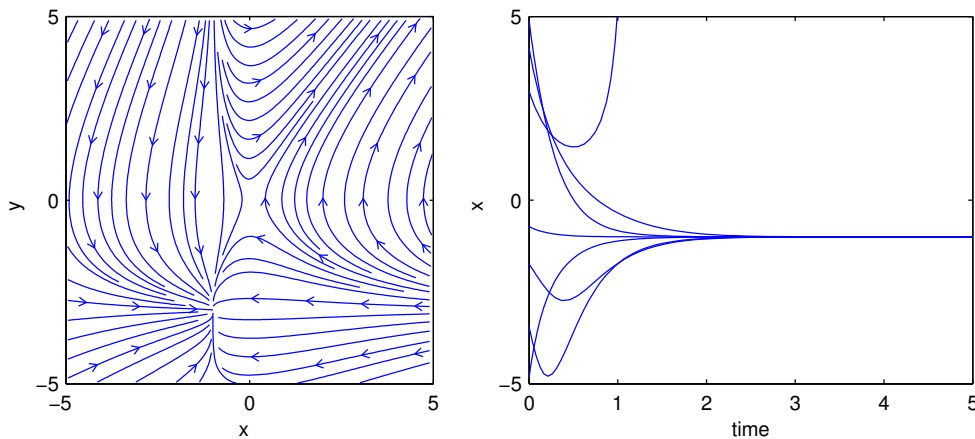
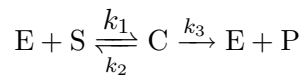


Fig. 6.16: Left: Phase portrait of the nonlinear system $dx/dt = xy + y$, $dy/dt = xy + 3x$. Right: sample solutions.

by enzymes. This type of reaction is therefore of particular interest in modelling and in experiments. We first introduced this system on page 100 and 185 and will have a detailed discussion of this reaction in Section 8.2. The biochemists diagrammatic representation is



where substrate S is under the action of enzyme E turned first into an intermediate complex C before further decomposed into a product P and the enzyme. The mass action kinetic equations for changes in the concentrations of the substrate and complex are

$$\begin{aligned} \frac{d[S]}{dt} &= -k_1[S][E] + k_2[C], \\ \frac{d[ES]}{dt} &= -(k_2 + k_3)[C] + k_1[S][E]. \end{aligned}$$

Since the enzyme is, in a sense, *controlling* the reaction, we may consider it as an input to the system. Rewriting these equations, using input variable u and state variables x_1 and x_2 for substrate and complex respectively, gives us the following compact representation that emphasizes an input-output representation and where state-variables and inputs are bilinearly related:

$$\begin{aligned} \dot{x}_1 &= -k_1x_1u + k_2x_2, \\ \dot{x}_2 &= -(k_2 + k_3)x_2 + k_1x_1u. \end{aligned}$$

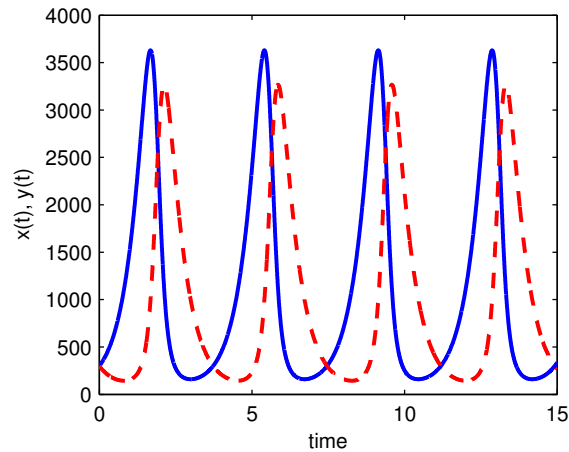


Fig. 6.17: Time series plot of the Lotka-Volterra system for $\alpha = 2$, $\beta = 0.002$, $\gamma = 0.0018$, $\delta = 2$ and initial conditions $x(0) = 300$, $y(0) = 300$. The solid line is for the prey population, $x(t)$, and the dashed line represents the predator population, $y(t)$.

Let us continue with a closer look at the predator-prey model for two competing or interacting populations, introduced by Lotka and Volterra:

$$\begin{aligned} \frac{dx}{dt} &= \alpha x - \beta xy, \\ \frac{dy}{dt} &= \gamma xy - \delta y, \end{aligned} \quad (6.14)$$

where $x, y \geq 0$, and α, β and δ are all positive constants. In (6.14) variable x is to represent the prey population and y the predator population. The structure of the Lotka-Volterra equations imply that in the absence of predators, the prey population will grow unbounded and in the absence of any prey, the predators will be extinguished. δ denotes the natural death rate of the predator and the term involving β describes the death of prey as being proportional to the encounters with predators.

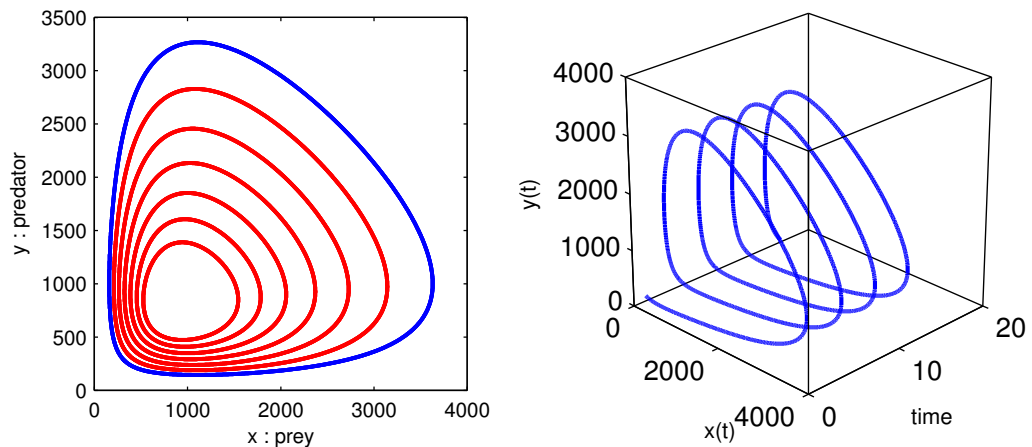


Fig. 6.18: Phase plane (left) and integral curve (right) of the Lotka-Volterra equations, (6.14).

Figure 6.18 shows two visualizations of the dynamic behavior to the Lotka-Volterra isocline system. The x -isoclines is described by those points in the phase-plane for which $dx/dt = 0$, i.e.,

$$f(x, y) = \alpha x - \beta xy = 0$$

which is true for

$$x = 0 \quad \text{and} \quad y = \frac{\alpha}{\beta}.$$

```

function lotkavolterra
delta = 2; gamma = 0.0018; alpha = 2; beta = 0.002;
x0 =300; y0 = 300; % Initial conditions. x: prey, y: predator
tspan = [0 15]; % Simulation time.
% ODE solver for numerical solution:
options = odeset('RelTol',1e-9,'AbsTol',1.e-9,'Refine',8);
[t,z] = ode45(@LV,tspan,[x0 y0],options,delta,gamma,alpha,beta);
plot(t,z(:,1),t,z(:,2)); % Plot time series.
plot(z(:,1),z(:,2)); % Trajectory in phase-plane.
plot3(z(:,1),t,z(:,2)); % Integral curve.
xlabel('prey'); ylabel('time');
% Subfunction for LV equations:
function dzdt = LV(t,z,delta,gamma,alpha,beta)
% z(1) : prey, z(2) : predator
dzdt = [alpha*z(1)-beta*z(1)*z(2); gamma*z(1)*z(2)-delta*z(2)];

```

Fig. 6.19: Matlab function to simulate the Lotka-Volterra equations.

These are two lines, equal to the y -axis going through the origin and a horizontal line at height α/β . The y -isoclines is defined in the same fashion,

$$g(x, y) = \gamma xy - \delta y$$

which is true for

$$y = 0 \quad \text{and} \quad x = \frac{\delta}{\gamma}.$$

These are again two perpendicular lines. There are two points of intersection between the x -isoclines and y -isoclines. There are therefore two fixed points for the system. The first fixed point is the origin of the plane, where $x = y = 0$. The second fixed point is given by $y = \alpha/\beta$, $x = \delta/\gamma$, which is a point at which the two populations are balanced so that there is no change to either population.

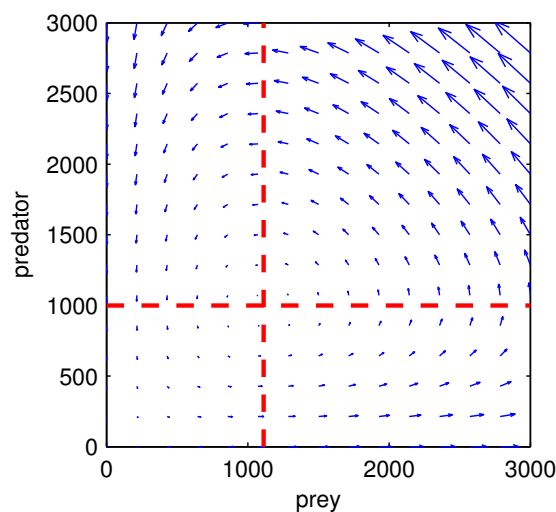


Fig. 6.20: Visualization of the flow of the Lotka-Volterra equations. The horizontal second x -isocline and vertical second y -isocline are shown as dotted lines. The fixed point lies where the isoclines meet. The first isoclines are the axis of the plot going through the origin.

Plotting trajectories in the phase-plane, as in Figure 6.18, requires a software tool for numerical integration (Figure 6.19). A quick way to visualize the flow of a two-dimensional nonlinear systems is to plot for a grid of (x, y) -values the gradient dy/dx as an arrow as

shown in Figure 6.20, plotted using the few lines of Matlab code in Figure 6.21. As an exercise, you may want to explain why the trajectories in Figure 6.18 are not circles or ellipsoids?

```
[X,Y] = meshgrid(linspace(0,3000,15));
dx = alpha.*X - beta.*X.*Y;
dy = gamma.*X.*Y - delta.*Y;
quiver(X,Y,dx,dy,1.5);
xlabel('prey'); ylabel('predator');
```

Fig. 6.21: Matlab function to visualize the flow of the Lotka-Volterra system.

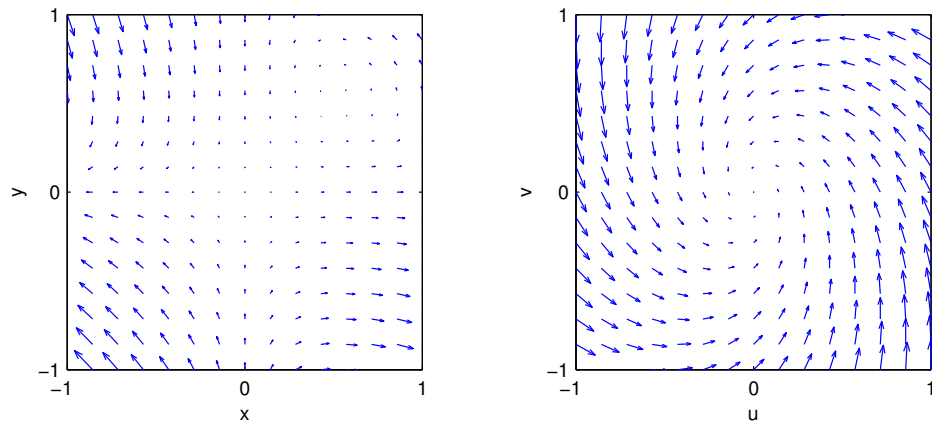


Fig. 6.22: Left: The flow of the Lotka-Volterra system, (6.14), with parameters $\beta = \gamma = 2$, $\alpha = \delta = 1$. Right: stable focus for $a = -1$, $b = -1$, $c = 1.9$, $d = -1$, which leads to eigenvalues $\lambda_1 = -1 + i1.378$, $\lambda_2 = -1 - i1.378$.

For the Lotka-Volterra system, (6.14), the fixed points occur at $x^* = 0$, $y^* = 0$ and at $x^* = \delta/\gamma$, $y^* = \alpha/\beta$. The constants of the system, linearized at (x^*, y^*) , are

$$\begin{aligned} a &= \left. \frac{\partial f}{\partial x} \right|_{x^*, y^*} = \alpha - \beta y^* , & b &= \left. \frac{\partial f}{\partial y} \right|_{x^*, y^*} = -\beta x^* , \\ c &= \left. \frac{\partial g}{\partial x} \right|_{x^*, y^*} = \gamma y^* , & d &= \left. \frac{\partial g}{\partial y} \right|_{x^*, y^*} = \gamma x^* - \delta . \end{aligned} \quad (6.15)$$

For the fixed point at the origin, $a = \alpha$, $b = 0$, $c = 0$, $d = -\delta$, such that the eigenvalues are

$$\lambda_1 = \alpha , \quad \lambda_2 = -\delta .$$

Since α and δ are positive, this fixed point is a saddle point, i.e., trajectories going towards it will drift off just before it. For the predator-prey model this means that even if the populations get near the extinction point, the populations will eventually grow again. For the second fixed point, we have $a = 0$, $b = -\beta\delta/\gamma$, $c = \gamma\alpha/\beta$, $d = 0$. The eigenvalues are

$$\lambda = \pm\sqrt{-\alpha\delta} .$$

Taking the square root of a negative number will lead to a complex number. The eigenvalues are therefore both imaginary numbers, which implies that the predator and prey populations oscillate around the fixed point, as can be seen well in Figure 6.20.

This section could only serve as a rudimentary introduction to nonlinear modelling. Other more comprehensive books at introductory level are [Str00b], and [JS99]. The book by Kaplan and Glass [KG95] is particularly suitable for biologists. Advanced texts are

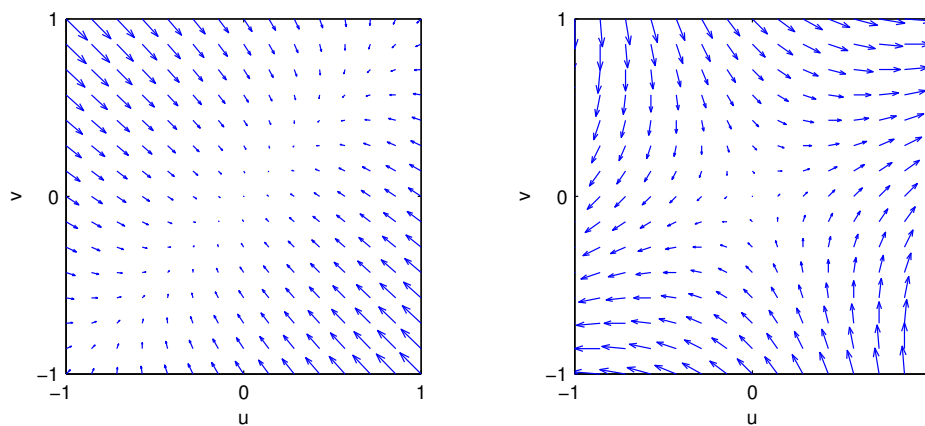


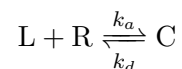
Fig. 6.23: Left: stable node of the linearized system, (6.10), with parameters taken from [KG95]: $a = -1.5$, $b = 1$, $c = 1$, $d = -1.5$, leading to $\lambda_1 = -2.5$, $\lambda_2 = -0.5$. Right: saddle point of the linearized system, (6.10), for $a = 1$, $b = 1$, $c = 1$, $d = -1$, leading to eigenvalues $\lambda_1 = -\sqrt{2}$, $\lambda_2 = \sqrt{2}$.

[Wig03] and [GH83]. The control engineering perspective of nonlinear systems is described in [Nv90], [Isi89], and [Sas99], all of which are advanced texts. The standard text on mathematical biology by Murray [Mur02] is an excellent source of examples for nonlinear modelling, applied to biological systems.

7 Receptor Modelling

In this section we pick up the thread of Section 2.2 and discuss in greater detail a mathematical model of cell surface receptor binding. A comprehensive study of receptors, models for binding, trafficking and signaling was first provided in [LL93]. Receptors are most commonly found at the cell surface, where extracellular signaling molecules, the ligand, can bind to them. Signaling proteins include cytokineses, insulin, hormones or growth factors, which could for example be transported through the blood stream. The binding process leads to a transmission of the signal into the cell where it can affect various processes, including the transcription of genes, which in turn can control various important cell functions.

We begin with a basic model for cell surface receptor binding, using the reversible bimolecular reaction



where R and L denote the free receptor of a cell and ligand molecules, respectively and C denotes the LR complex, i.e., receptors that are “occupied”. k_a is the rate constant at which ligands bind to receptors and k_d is the dissociation constant. We refer to receptor and ligand as *monovalent* to assume that at any time only one ligand molecule and one receptor form a complex. For a single cell, the kinetic model that describes temporal changes in the number of LR complexes is then

$$\frac{dC}{dt} = k_a R[L] - k_d C, \quad (7.1)$$

where $[L]$ gives the free ligand concentration in moles per liter of the medium; R is the number of free receptor per cell ($\#/cell$), C the number of receptor-ligand complexes per cell, k_d and k_a are in sec^{-1} and $M^{-1}sec^{-1}$, respectively. The number of receptors or

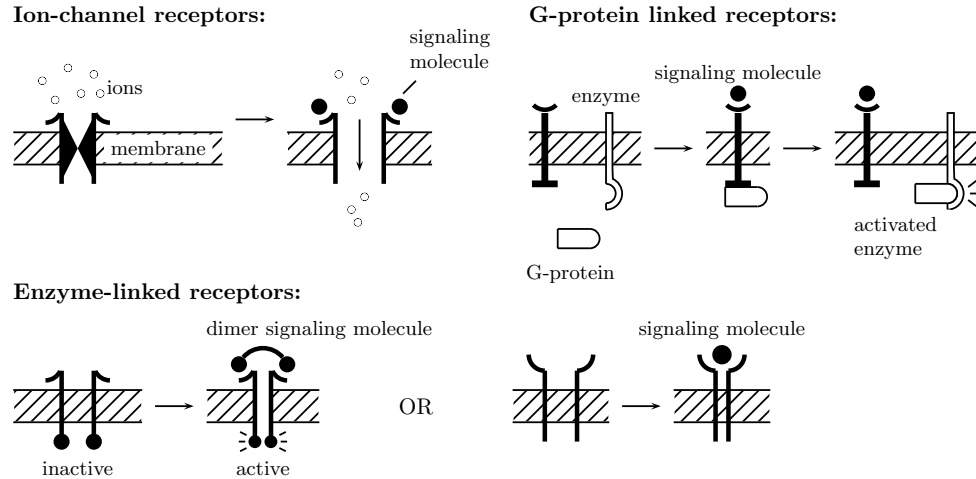


Fig. 7.1: The figure illustrate the three most common kinds of cell surface receptors and mechanisms for their activation.

complexes per cell can be converted into a concentration (moles per volume solution) or density ($\#/cell$ surface area) if necessary. To check the units of (7.1) we have

$$\frac{1}{\text{cell} \cdot \text{sec}} = \frac{1}{\text{M} \cdot \text{sec}} \cdot \frac{1}{\text{cell}} \cdot \text{M} - \frac{1}{\text{sec}} \cdot \frac{1}{\text{cell}} .$$

Equation (7.1) has three variables, C , L , and R . A reasonable assumption to simplify the analysis is that the total number of surface receptors, denoted R_T , is constant:

$$R_T = R + C , \quad (7.2)$$

leading to an equation in two variables:

$$\frac{dC}{dt} = k_a(R_T - C)[L] - k_dC . \quad (7.3)$$

The ligand concentration is determined by the initial concentration minus the ligands bound in complexes, $L = L_0 - C$. Furthermore, in most experimental set-ups we are going to have n cells in the medium to which also the concentration of the ligand refers. This means that we ought to multiply C (given in $\#/cell$) by n . Furthermore, since C is a count per cell, we turn the concentration of L , given in mol/liter, into a count of numbers by multiplying with N_A the Avogadro constant ($\#/mol$):

$$[L] = L_0 - \left(\frac{n}{N_A} \right) C , \quad (7.4)$$

where we write for the initial *concentration* of L , $[L](t = 0) = L_0$. The brackets are therefore left to simplify the notation. Inserting the two conservation assumptions (7.2) and (7.4) into (7.1) gives us a single differential equation to describe the receptor-ligand binding process:

$$\begin{aligned} \frac{dC}{dt} &= k_a(R_T - C) \left(L_0 - \frac{n}{N_A} C \right) - k_dC \\ &= k_a R_T L_0 - k_a R_T \frac{n}{N_A} C - k_a L_0 C + k_a \frac{n}{N_A} C^2 - k_d C . \end{aligned} \quad (7.5)$$

Rewriting this equation we recognize it as an inhomogeneous second-order differential equation:

$$\frac{dC}{dt} + \left(k_a R_T \frac{n}{N_A} + k_a L_0 + k_d \right) C - k_a \frac{n}{N_A} C^2 = k_a R_T L_0 . \quad (7.6)$$

Negligible Ligand Depletion

In order to obtain a simpler differential equation than (7.6) we make further assumptions. If we assume that ligand depletion can be neglected and therefore replace L in (7.3) with L_0 the differential equation becomes dependent on only one variable:

$$\frac{dC}{dt} = k_a(R_T - C)L_0 - k_dC . \quad (7.7)$$

If we imagine there are initially no ligands bound, the assumption of negligible ligand depletion implies that the initial concentration is much larger than the ligand bound in complexes, i.e., $(n/N_A)C \ll L_0$. Alternatively we could have replaced $L_0 - (n/N_A)C$ by L_0 in (7.5) to arrive at (7.7).

Let us now determine the solution to differential equation (7.7), by recognizing it as a linear, inhomogeneous, first-order ordinary differential equation¹

$$\frac{dC}{dt} + \underbrace{(k_aL_0 + k_d)}_{P(t)}C = \underbrace{k_aR_TL_0}_{Q(t)} . \quad (7.8)$$

Such equations are solved by defining an *integrating factor*

integrating factor

$$\begin{aligned} \rho(t) &= e^{\int P(t)dt} \\ &= \exp \left\{ \int (k_aL_0 + k_d) dt \right\} \\ &= \exp \{ (k_aL_0 + k_d)t \} . \end{aligned}$$

The next step is to multiply both sides of (7.8) with $\rho(t)$

$$\exp \{ (k_aL_0 + k_d)t \} \frac{dC}{dt} + \exp \{ (k_aL_0 + k_d)t \} (k_aL_0 + k_d)C = \exp \{ (k_aL_0 + k_d)t \} k_aR_TL_0 .$$

We notice that the left-hand side is the derivative of the product $C(t) \cdot \rho(t)$, which is in fact the whole idea behind the use of an integrating factor. We can thus write

$$\rho(t) \cdot C(t) = \int \rho(t) \cdot Q(t)dt + c ,$$

where c is some arbitrary constant. Insert the expressions for $\rho(t)$ and $Q(t)$,

$$\begin{aligned} \rho(t) \cdot C(t) &= \int \exp \{ (k_aL_0 + k_d)t \} \cdot k_aR_TL_0 dt + c \\ &= k_aR_TL_0 \int_0^t e^{(k_aL_0+k_d)t} dt + c \\ &= \frac{k_aR_TL_0}{k_aL_0 + k_d} \left(e^{(k_aL_0+k_d)t} - 1 \right) + c , \end{aligned}$$

leading to

$$C(t) = \frac{k_aR_TL_0}{k_aL_0 + k_d} - \frac{k_aR_TL_0}{k_aL_0 + k_d} e^{-(k_aL_0+k_d)t} + c \cdot e^{-(k_aL_0+k_d)t} , \quad (7.9)$$

¹This is an ordinary differential equation since there is only a single independent variable, t . The equation is linear since no terms such as C^2 appear and it is first-order since only the first derivative dC/dt appears. The homogenous version of (7.7) is obtained for $k_aR_TL_0 = 0$. This special case would correspond to the monomolecular reaction (3.2) on page 39.

where $1/\rho(t) = \exp\{-(k_a L_0 + k_d)t\}$. From the initial condition, $C(t = 0) = C_0$, we can determine the constant

$$C_0 = \frac{k_a R_T L_0}{k_a L_0 + k_d} - \frac{k_a R_T L_0}{k_a L_0 + k_d} + c ,$$

i.e., $c = C_0$, which inserted into (7.9), gives us the solution

$$C(t) = C_0 \exp\{-(k_a L_0 + k_d)t\} + \frac{k_a R_T L_0}{k_a L_0 + k_d} (1 - \exp\{-(k_a L_0 + k_d)t\}) . \quad (7.10)$$

Equilibrium and Steady State

equilibrium At *equilibrium* the reaction rates are equal

$$k_a [L] R = k_d C ,$$

dissociation constant where the *dissociation constant* is defined (cf. page 25) as

$$K_d = \frac{R[L]}{C} = \frac{k_d}{k_a} ,$$

where here we have for $R = R_T - C$ and $[L](t = 0) = L_0$, from our assumptions (7.2) and (7.4) above:

$$K_d = \frac{(R_T - C)L_0}{C} ,$$

with unit M (mol per liter). Let us denote by C_{eq} the number of ligand-receptor complexes at equilibrium,

$$C_{\text{eq}} = \frac{R_T L_0}{K_d + L_0} . \quad (7.11)$$

steady state At *steady state*, $dC/dt = 0$,

$$0 = k_a (R_T - C)L_0 - k_d C ,$$

leading to

$$C = \frac{k_a R_T L_0}{k_d + k_a L_0} = \frac{R_T L_0}{K_d + L_0} .$$

The steady state value is therefore in this case identical to the number of receptor-ligand complexes at equilibrium.

Dimensionless Representation

To visualize the solution (7.10) with a plot, we would have to specify the rate constants, the initial ligand concentration, the total number of receptors and the initial number of receptor-ligand complexes. The appearance of the plot may therefore vary considerably for different values of these parameters. This can be avoided by not looking at $C(t)$ but plotting the dimensionless fraction of occupied receptors,

$$y = \frac{C}{R_T} , \quad (7.12)$$

where $0 \leq y \leq 1$. For $y = 0$ all receptors are free (no complexes), and for $y = 1$ all receptors are occupied. Let us furthermore introduce a scaled time, τ ,

$$\tau = k_d t . \quad (7.13)$$

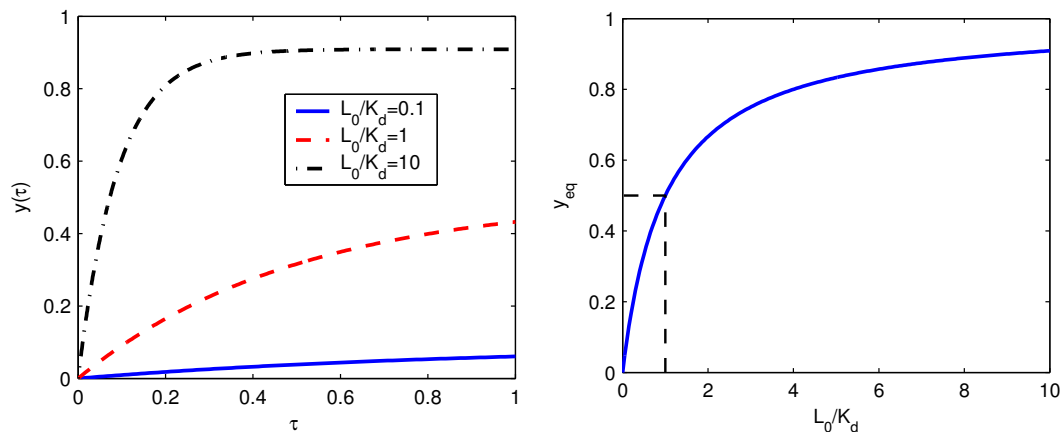


Fig. 7.2: Left: Transient binding of ligands to cell receptors, where a bimolecular reaction is used as a model, with ligand depletion assumed to be negligible and the total number of receptors remaining unchanged. At $t = 0$, it is assumed that all receptors are free. Right: The fractional occupancy of receptors at equilibrium.

Rewriting (7.7) first by dividing both sides by R_T

$$\begin{aligned} \frac{d}{dt} \left(\frac{C}{R_T} \right) &= k_a \left(1 - \frac{C}{R_T} \right) L_0 - k_d \frac{C}{R_T}, \\ \frac{dy}{dt} &= k_a(1-y)L_0 - k_d y, \end{aligned}$$

next taking account of (7.13) gives

$$\begin{aligned} \frac{dy}{d\tau} &= \frac{k_a}{k_d}(1-y)L_0 - y, \\ &= \frac{L_0}{K_d}(1-y) - y. \end{aligned} \quad (7.14)$$

For the fractional occupancy of receptors, the transient changes are therefore described by the following solution of (7.14)

$$y(\tau) = y_0 \exp \left\{ - \left(1 + \frac{L_0}{K_d} \right) \tau \right\} + \frac{L_0/K_d}{1 + (L_0/K_d)} \left(1 - \exp \left\{ - \left(1 + \frac{L_0}{K_d} \right) \tau \right\} \right), \quad (7.15)$$

and the equilibrium value is determined by the ratio L_0/K_d :

$$y_{\text{eq}} = \frac{L_0/K_d}{1 + (L_0/K_d)}. \quad (7.16)$$

See Figure 7.2 for an illustration of the transient binding of ligands to cell receptors. Note that $y_{\text{eq}} = C_{\text{eq}}/R_T$ and $y_{\text{eq}} = 0.5$ when $L_0/K_d = 1$, i.e., half the receptors are bound by ligands at equilibrium when the ligand concentration is equal to the value of K_d .

Half-Time!

Taking a break before we continue with our model of ligand binding, we complete the previous study by looking at the *half-time* of an experiment. Let us assume an experiment half time with initial condition $y_0 = C_0/R_T = 0$, i.e., initially no ligands are bound to the receptors.

One definition for a half-time, τ_h , is for the transient solution $y(\tau)$ to reach half of the change from $y_0 = 0$ to y_{eq} , $y = 0.5y_{\text{eq}}$:

$$\frac{1}{2}y_{\text{eq}} = y_{\text{eq}} \left[1 - \exp \left\{ - \left(1 + \frac{L_0}{K_d} \right) \tau_h \right\} \right], \quad (7.17)$$

$$\frac{1}{2} = - \exp \left\{ - \left(1 + \frac{L_0}{K_d} \right) \tau_h \right\}, \quad (7.18)$$

$$- \left(1 + \frac{L_0}{K_d} \right) \tau_h = \ln \frac{1}{2}, \quad (7.19)$$

such that

$$\tau_h = \frac{-\ln 1/2}{1 + L_0/K_d} = \frac{\ln 2}{1 + L_0/K_d} \approx \frac{0.69}{1 + L_0/K_d}. \quad (7.20)$$

This half time is shown in the right-hand plot of Figure 7.2.

Alternatively, we may ask for the left-hand plot of $y(\tau)$ in Figure 7.2 for when $y(\tau_h) = 1/2$. Considering again an initial value $y(t = 0) = 0$,

$$\begin{aligned} \frac{1}{2} &= y_{\text{eq}} \left[1 - \exp \left\{ - \left(1 + \frac{L_0}{K_d} \right) \tau_h \right\} \right], \\ \frac{1}{2y_{\text{eq}}} &= 1 - \exp \left\{ - \left(1 + \frac{L_0}{K_d} \right) \tau_h \right\}, \\ \exp \left\{ - \left(1 + \frac{L_0}{K_d} \right) \tau_h \right\} &= - \left(\frac{1 - 2y_{\text{eq}}}{2y_{\text{eq}}} \right), \\ - \left(1 + \frac{L_0}{K_d} \right) \tau_h &= \ln \left(\frac{2y_{\text{eq}} - 1}{2y_{\text{eq}}} \right). \end{aligned}$$

Since it must be possible for half of the receptors to be occupied, $y_{\text{eq}} \geq 0.5$.

$$\begin{aligned} \left(1 + \frac{L_0}{K_d} \right) \tau_h &= - \ln \left(\frac{2y_{\text{eq}} - 1}{2y_{\text{eq}}} \right) = \ln \left(\frac{2y_{\text{eq}}}{2y_{\text{eq}} - 1} \right), \\ \tau_h &= \left(\ln \frac{2y_{\text{eq}}}{2y_{\text{eq}} - 1} \right) \cdot \left(1 + \frac{L_0}{K_d} \right)^{-1} = \frac{\ln[2y_{\text{eq}}/(2y_{\text{eq}} - 1)]}{1 + L_0/K_d}. \end{aligned}$$

Stimulus-Response Analysis

The previous study of receptor-ligand binding, based on equation (7.7), assumed that the ligand concentration is more or less constant. The solution (7.10), respectively (7.15), can be thus be interpreted as the response to a step-change in the ligand concentration. We now return to equation (7.3) and consider different kinds of stimuli.

$$\begin{aligned} \frac{dC}{dt} &= k_a(R_T - C)[L] - k_dC \\ &= k_aR_T[L] - k_aC[L] - k_dC, \end{aligned} \quad (7.3)$$

where both C and the ligand concentration $[L](t)$ are a function of time. A check of the units is quickly done for the last equation

$$\frac{\#}{\text{cell} \cdot \text{sec}} = \frac{1}{\text{M} \cdot \text{sec}} \cdot \frac{\#}{\text{cell}} \cdot \text{M} - \frac{1}{\text{M} \cdot \text{sec}} \cdot \frac{1}{\text{cell}} \cdot \text{M} - \frac{1}{\text{sec}} \cdot \frac{\#}{\text{cell}}.$$

To investigate the response in receptor binding to different forms of ligand stimulation we again use the dimensionless variable $y = C/R_T$, which represents the fraction of occupied receptors

$$\frac{dy}{dt} = k_a[L] - k_a y[L] - k_d y.$$

Let us further scale time by introducing $\tau = t \cdot k_d$, which means that the right-hand side of the above equation is divided by k_d

$$\frac{dy}{d\tau} = \frac{k_a}{k_d}[L] - \frac{k_a}{k_d}y[L] - y .$$

For the reversible bimolecular reaction the equilibrium constant and dissociation constants are defined as (cf. page 25)

$$K_{\text{eq}} = \frac{k_a}{k_d} , \quad K_d = \frac{k_d}{k_a} ,$$

such that we have

$$\frac{dy}{d\tau} = \frac{1}{K_d}[L] - \frac{1}{K_d}y[L] - y . \quad (7.21)$$

In order to make this equation more appealing for the eye, we hereafter use x to denote the stimulus L and replace the equilibrium constant by θ :

$$\frac{dy}{d\tau} = \theta x - \theta xy - y . \quad (7.22)$$

This then is a nonlinear ordinary differential equations with one parameter θ and two temporal variables $x(t)$ and $y(t)$.

We begin with a downward step-change

$$x(\tau) = \frac{\alpha}{1 + \exp\left\{\frac{\tau - \beta}{\gamma}\right\}} , \quad (7.23)$$

where the parameters

- α : determines the initial height,
- β : defines the turning point of the curve,
- γ : determines sharpness of the transition.

For $\gamma \rightarrow 0$ we obtain the Heaviside step-function as illustrated in Figure 7.3. In Figure 7.4, on the left, the solution of (7.22) is shown for three different ratios of L_0/K_d , $\alpha = 2$, $\beta = 2.5$ and $\gamma = 0.2$. The right-hand plot of Figure 7.4 shows the response to a downward step with the slope of the change, γ , changing; $L_0/K_d = 10$ and α and β as before. The inverse of the “downward” or “negative” step, is the function

$$x(\tau) = \frac{\alpha}{1 + \exp\left\{-\frac{\beta}{\gamma}\right\}} - \frac{\alpha}{1 + \exp\left\{\frac{\tau - \beta}{\gamma}\right\}} .$$

The response pattern to this stimulus are shown in Figure 7.5.

Next we consider an ‘impulse’-like function, which we represent by a gaussian function

$$x(\tau) = \frac{\alpha}{\gamma} \cdot \exp\left\{-\frac{(\tau - \beta)^2}{\gamma^2}\right\} . \quad (7.24)$$

The pre-factor α/γ is chosen in this way as to ensure the integral of the right-hand side (i.e., the intensity) is constant. For $\alpha = 2$, $\beta = 2.5$ and changing width, γ , the function is shown in Figure 7.6 and the response pattern in Figure 7.7.

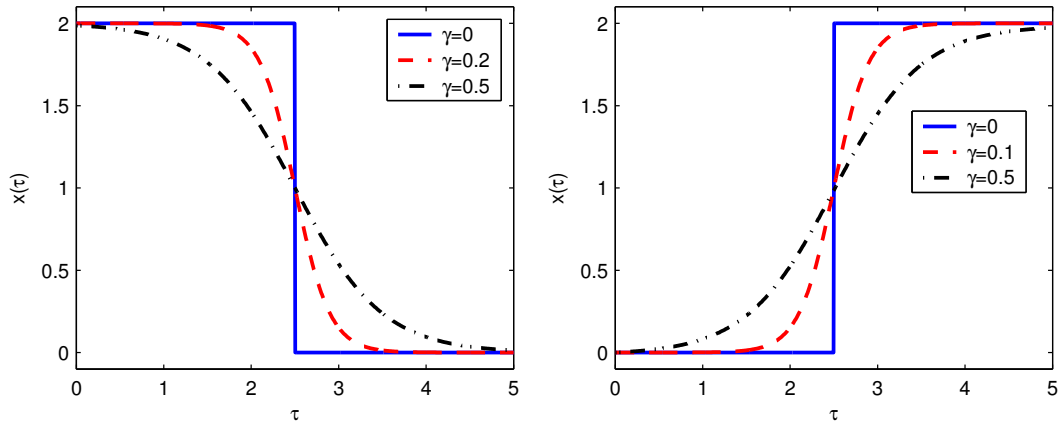


Fig. 7.3: Left: Downward step-change stimulus $x(\tau)$ for different parameter values of γ ; $\alpha = 2$, $\beta = 2.5$. Right: Positive step.

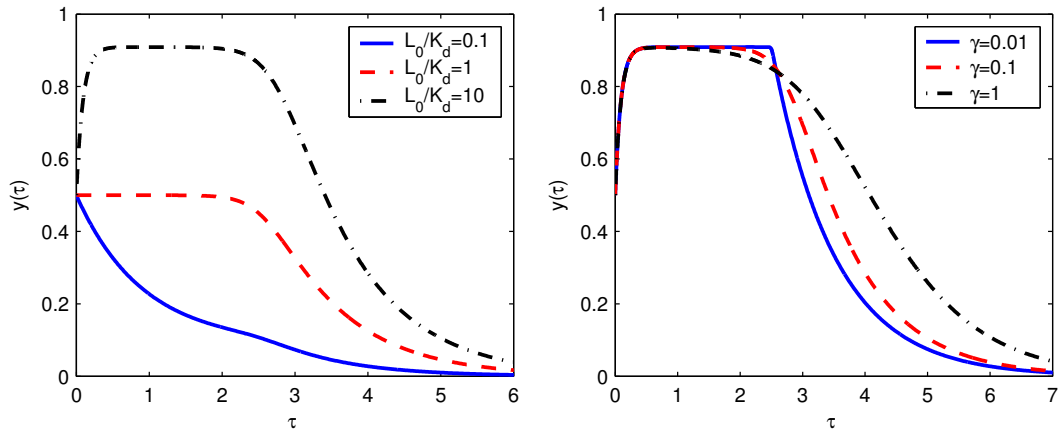


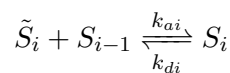
Fig. 7.4: Left: Response to downward step-changes for initial condition $y = C/R_T = 0.5$ and different values for L_0/K_d ; $\alpha = 2$, $\beta = 2.5$ and $\gamma = 0.2$. Right: Response to a downward step with the slope of the change, γ , changing; $L_0/K_d = 10$ and α and β as before.

Conclusions

For a full understanding of the dynamics of a system it is necessary to conduct a series stimulus-response experiments. For most experiments an initial concentration of ligands is depleted and it is not possible to control the exact shape of the stimulus. Here we have considered ‘typical’ input stimuli that may occur in a system.

We can easily extend the model of extracellular ligands binding to receptors, as discussed above, to *intracellular* signalling. A common mechanism for receptor regulated dimerization signalling is *dimerization*, a ligand induced monomer-to-dimer transition. As a monomer a single receptor is inactive, dimerization leads to an activation and intracellular autophosphorylation of the signaling domain as illustrated in Figure 7.8.

Let us denote by \tilde{S} the non-phosphorylated form of a molecular species or substrate S . The model for a signaling step suggested here is a phosphorylation



where analog to (7.1)

$$\frac{d[S_i]}{dt} = k_{ai}[\tilde{S}_i][S_{i-1}] - k_{di}[S_i].$$

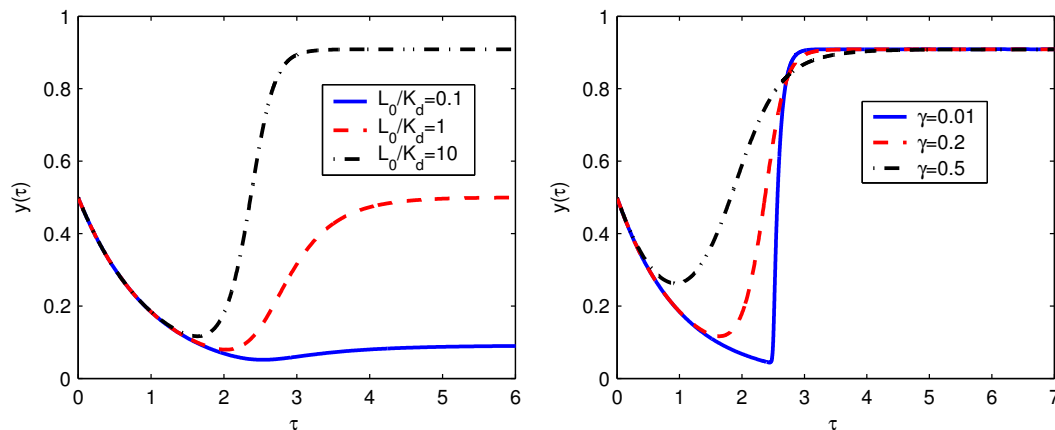


Fig. 7.5: Left: Response to downward step-changes for initial condition $y = C/R_T = 0.5$ and different values for L_0/K_d ; $\alpha = 2$, $\beta = 2.5$ and $\gamma = 0.2$. Right: Response to a downward step with the slope of the change, γ , changing; $L_0/K_d = 10$ and α and β as before.

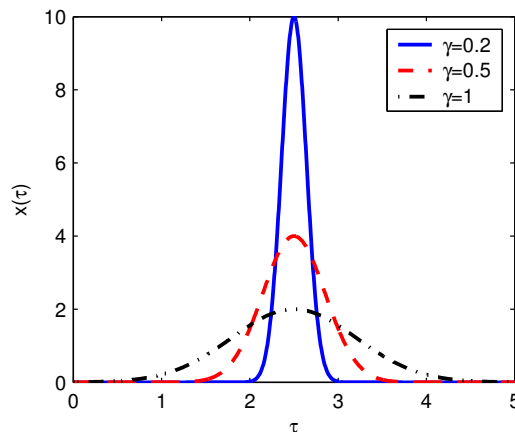


Fig. 7.6: Impulse-like stimulus $x(\tau)$, $\alpha = 2$, $\beta = 2.5$ for different width parameter.

We assume the total concentration of kinase S_i is constant

$$[S_{Ti}] \doteq [\tilde{S}_i] + [S_i] ,$$

which inserted into the previous equation gives

$$\begin{aligned} \frac{d[S_i]}{dt} &= k_{ai}[S_{i-1}]([S_{Ti}] - [S_i]) - k_{di}[S_i] , \\ &= k_{ai}[S_{Ti}][S_{i-1}] - k_{ai}[S_{i-1}][S_i] - k_{di}[S_i] . \end{aligned}$$

In support of our eyesight we introduce x_i to denote the phosphorylation of S_i and write $c_i \doteq [S_{Ti}]$, $\alpha_i \doteq k_{ai}c_i$, $\beta_i \doteq k_{di}$, leading to

$$\begin{aligned} \frac{dx_i}{dt} &= \alpha_i x_{i-1} - \frac{1}{c_i} x_{i-1} x_i - \beta_i x_i , \\ \frac{dx_i}{dt} &= \alpha_i x_{i-1} \left(1 - \frac{x_i}{\alpha_i c_i} \right) - \beta_i x_i . \end{aligned} \quad (7.25)$$

This last equation is the basis for an interesting study of properties of a signaling pathway, first introduced by Heinrich et al. [HNR02] and which we are going to discuss in greater detail in Section 8.5.

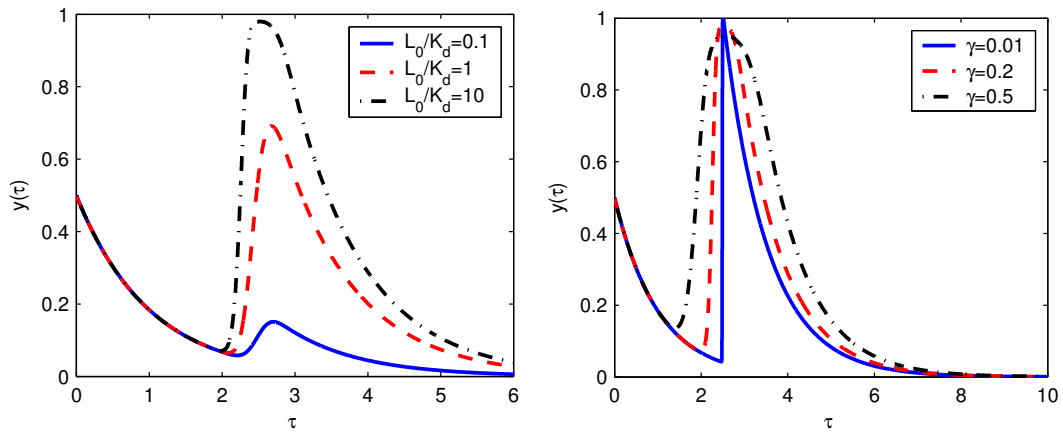


Fig. 7.7: Left: Response to impulse stimulus for initial condition $y = C/R_T = 0.5$ and different values for L_0/K_d ; $\alpha = 2$, $\beta = 2.5$ and $\gamma = 0.2$. Right: Response to an impulse with width, γ , changing; $L_0/K_d = 10$ and α and β as before.

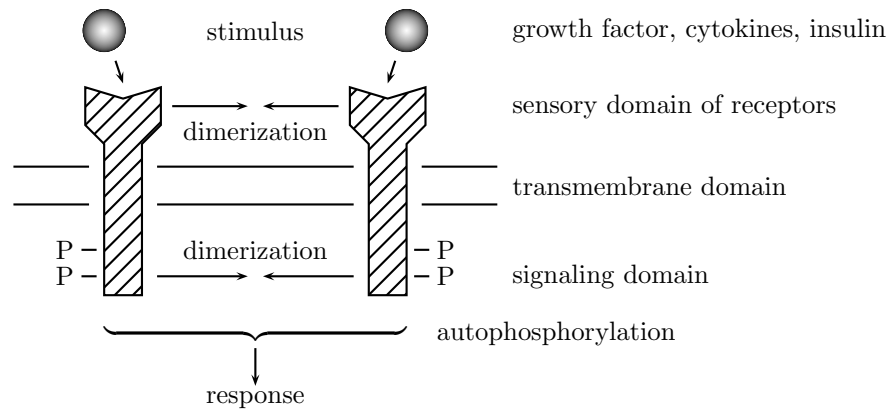


Fig. 7.8: Example of a common mechanism for receptor regulated signalling. Dimerization is a ligand induced monomer-to-dimer transition. As a monomer a single receptor is inactive, dimerization leads to an activation and intracellular autophosphorylation of the signaling domain.

8 Dynamic Modelling of Biochemical Networks

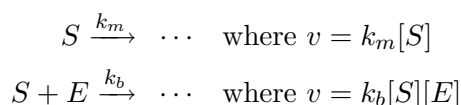
Pathways are networks of biochemical reactions, most of which are facilitated by highly specialized enzymes. The enzyme kinetic reaction can therefore serve as a template to construct more complex pathway models. In the present section we therefore first look at the equations that represent an enzyme kinetic reaction and hint at the use in dynamic modelling of signal transduction pathways.

We are going to consider a compartment or region of the cell with volume V for which we assume that diffusion is fast compared to the time scales of the reactions and hence concentrations within this volume are homogenous. In many cases it is possible to decompose more complex reaction networks into a set of uni- or monomolecular (first-order) reactions and bimolecular (second order reactions), depicted



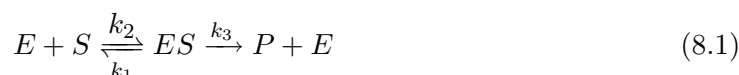
where the arrow denotes a conversion according to the law of mass action. Concentrations are specified in mol per liter (M). When it is clear from the context, the square brackets which denote concentrations, are often left away to have a simpler notation. The letters for

reaction rate the variables are chosen arbitrarily and depending on the context. The rate of the reaction or *reaction rate* v is, in case of the monomolecular reaction defined by the product k_m and $[S]$ and in case of the bimolecular reaction defined by the product of k_b with $[S]$ and $[E]$:

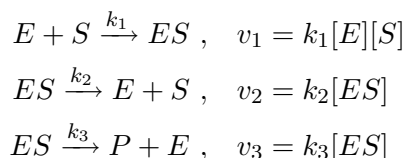


The linear (monomolecular reaction), respectively bilinear relationship (bimolecular reaction) of the reaction rate on the concentrations is in essence the law of mass action. Note that the units of the rate constant k is per second (sec^{-1}) for the monomolecular reaction and in moles per second ($\text{M}^{-1}\text{sec}^{-1}$) for bimolecular reactions.

One approach to model more complex signal transduction pathways is to model each step of the pathway on a template of an enzyme kinetic reaction (4.23)



with k_1 denoting the rate at which the complex ES is formed; k_2 at which ES dissociates into enzyme E and substrate S ; k_3 , the rate at which ES dissociates into product P and E . The reaction diagram (8.1) can be decomposed into a set of mono- and bi-molecular reactions



The ordinary differential equation model is directly derived from these reactions

$$\left. \begin{aligned} \frac{d[E]}{dt} &= -k_1[E][S] + k_2[ES] + k_3[ES] \\ \frac{d[S]}{dt} &= -k_1[E][S] + k_2[ES] \\ \frac{d[ES]}{dt} &= k_1[E][S] - k_2[ES] - k_3[ES] \\ \frac{d[P]}{dt} &= k_3[ES] . \end{aligned} \right\} \quad (8.2)$$

Figure 8.1 gives a graphical representation of these equations. If these graphical representation have a one-to-one mapping to the equations, they are an important communication tool in interdisciplinary collaborations. Biologists naturally draw cartoons to represent the relationship between variables in a pathway. There is however no standard about the meaning of symbols and it is usually not obvious how to translate this into equations.

Using the rate of reactions, an alternative compact representation of (8.2) is

$$\left. \begin{aligned} \frac{d[E]}{dt} &= -v_1 + v_2 + v_3 \\ \frac{d[S]}{dt} &= -v_1 + v_2 \\ \frac{d[ES]}{dt} &= v_1 - v_2 - v_3 \\ \frac{d[P]}{dt} &= v_3 . \end{aligned} \right\} \quad (8.3)$$

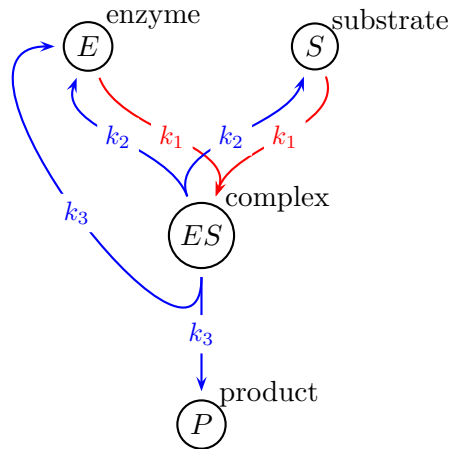


Fig. 8.1: Graphical representation of the enzyme kinetic reaction. More complex signal transduction pathways can be constructed using this template.

The enzyme is considered a catalyst, which facilitates the reaction without loss, i.e., the total enzyme concentration, i.e. the sum of free enzyme $[E]$ and enzyme in the complex $[ES]$, is constant. This is also apparent from adding the equations for $[\dot{E}]$ and $[\dot{ES}]$:

$$[\dot{E}] + [\dot{ES}] = 0 \quad \text{which implies} \quad [E](t) + [ES](t) = c_1 .$$

Assuming there is initially no complex, $[ES](0) = 0$, the constant equals the initial enzyme concentration $c_1 = [E](0)$. Inserting $[E](t) = [E](0) - [ES](t)$ into the equation for substrate $[S]$, and complex $[ES]$, the system of ODEs reduces to two equations:

$$\begin{aligned} [\dot{S}] &= -k_1[E](0)[S] + (k_1[S] + k_2)[ES] && \text{(substrate)} \\ [\dot{ES}] &= k_1[E](0)[S] - (k_1[S] + k_2 + k_3)[ES] && \text{(complex)} \end{aligned}$$

with initial conditions $[S](0)$ and $[ES](0) = 0$.

The structure of the equations (i.e., the signs, the number of terms, and the variables involved) are obtained as a direct translation of the biologist's knowledge of a pathway. This knowledge is usually not firm and the purpose of modelling is not only to fit experimental data to an assumed model, but to identify an appropriate model structure, validating or updating the knowledge we have of the proteins in a pathway and how they interact. To decide whether a model structure is realistic, a simulation could reveal whether the concentration profiles match experimental experience. For a simulation we do however need to know the values of the parameters. Ideally, we would like to derive some general properties of the system, without knowing the exact parameter values. For the set of equations (8.2) this can be done easily by looking at the equations. We know that the enzyme $[E]$ turns the substrate $[S]$ into the product $[P]$ and thus we would expect $[P]$ to increase steadily while $[S]$ decreases. The last equation $[\dot{P}] = k_3[ES]$ makes the product increase so long as $[ES]$ is positive. Since we deal with concentrations, all x_i can only be positive. Looking at the equation for the substrate, $[S]$ will decrease so long as the right-hand side is negative, i.e., $k_1[E][S] > k_2[ES]$. Thus from any initial condition $[E]$, $[S]$, and $[ES]$ would adjust themselves steadily until $k_1[E][S] > k_2[ES]$ and from then on $[S]$ would decrease steadily.

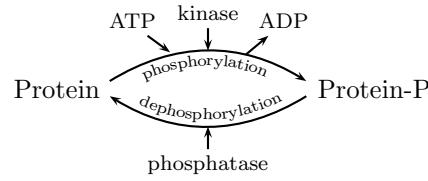
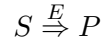
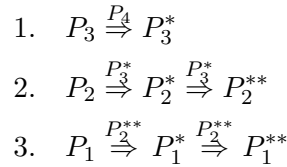


Fig. 8.2: A common means of signal transduction is through sequential phosphorylation and dephosphorylation. Phosphorylation is facilitated by means of a kinase and dephosphorylation is realized by a phosphatase. The phosphorylated state is denoted by adding -P to the name of the protein. More complex pathway diagrams can be built by connecting these basic components into cascades and networks.

Symbolically, we can summarize the enzyme kinetic reaction model (4.23), (8.1) as follows:



We read this as “the conversion of S into P , facilitated (or catalyzed) by E ”. For example, signal transduction pathways are commonly considered as a series or cascade of modules, each of which can be modelled using the enzyme kinetic reaction as a template. The signal in these pathways is transmitted through facilitated phosphorylation of proteins referred to as ‘kinases’:



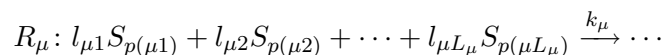
where the * and ** denote phosphorylation and double phosphorylation, respectively. Here P_4 facilitates the phosphorylation of P_3 and so forth. More generally we use the * to denote an activation, which can but must not be achieved by phosphorylation. A phosphorylation, e.g. of MEK, is often also denoted by adding -P to the name of the protein, MEK-P, or ERK-PP, for phosphorylation and double phosphorylation, respectively. Figure 8.2 shows another common way to illustrate signaling steps in diagrams.

8.1 Simulation example

As an illustration, we here describe the simulation of the enzyme-kinetic reaction (8.2). The four differential equations of (8.2) are an example for the mass action representation (2.19) consisting of N ordinary differential rate equations

$$\frac{d}{dt}[S_i] = \sum_{\mu=1}^M \nu_{\mu i} k_{\mu} \prod_{j=1}^{L_{\mu}} [S_{p(\mu j)}]^{l_{\mu j}} \quad i = 1, 2, \dots, N \quad (8.4)$$

where the k_{μ} 's are rate constants and ν_{μ} denotes the change in molecules of S_i resulting from a single R_{μ} reaction. For more complex reaction networks one first has to divide reversible reactions up into basic reaction channels



where L_{μ} is the number of reactant species in channel R_{μ} , $l_{\mu j}$ is the stoichiometric coefficient of reactant species $S_{p(\mu j)}$, $K_{\mu} = \sum_{j=1}^{L_{\mu}} l_{\mu j}$ denotes the molecularity of reaction

channel R_μ , and the index $p(\mu j)$ selects those S_i participating in R_μ . For the enzyme-kinetic reaction there are $M = 3$ reaction channels



For $i = 1, \dots, N = 4$, we translate the names ‘Enzyme’, ‘Substrate’, ‘Enzyme/Substrate complex’, and ‘Product’ into the notation of Chapter 3:

$$S_1 \doteq E \quad S_2 \doteq S, \quad S_3 \doteq ES, \quad S_4 \doteq P.$$

Subsequently, we have the matrix $\nu = [\nu_{\mu i}]$

$$\begin{array}{cccc} \nu_{11} = -1, & \nu_{12} = -1, & \nu_{13} = +1, & \nu_{14} = 0 \\ \nu_{21} = +1, & \nu_{22} = +1, & \nu_{23} = -1, & \nu_{24} = 0 \\ \nu_{31} = +1, & \nu_{32} = 0, & \nu_{33} = -1, & \nu_{34} = 1 \end{array}$$

The indices for participating species are collected in terms of vectors p_μ

$$p_1 = (1, 2), \quad p_2 = 3, \quad p_3 = 3$$

Similar, to facilitate the software implementation of the equations, the stoichiometry is defined by

$$l_1 = (1, 1), \quad l_2 = 1, \quad l_3 = 1$$

such that the molecularity $K_\mu = \sum_{j=1}^{L_\mu} l_{\mu j}$ is encoded as follows

$$\begin{array}{ccc} L_1 = 2, & L_2 = 1, & L_3 = 1 \\ K_1 = 2, & K_2 = 1, & K_3 = 1. \end{array}$$

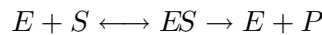
This leads us to a representation of the enzyme-kinetic reaction (4.23) in the form of (8.4)

$$\begin{aligned} \frac{dS_1}{dt} &= \nu_{11}k_1S_1^1S_2^1 + \nu_{21}k_2S_3^1 + \nu_{31}k_3S_3^1 \\ \frac{dS_2}{dt} &= \nu_{12}k_1S_1^1S_2^1 + \nu_{22}k_2S_3^1 + \nu_{32}k_3S_3^1 \\ \frac{dS_3}{dt} &= \nu_{13}k_1S_1^1S_2^1 + \nu_{23}k_2S_3^1 + \nu_{33}k_3S_3^1 \\ \frac{dS_4}{dt} &= \nu_{14}k_1S_1^1S_2^1 + \nu_{24}k_2S_3^1 + \nu_{34}k_3S_3^1 \end{aligned}$$

Figure 8.3 shows a Matlab coding for the enzyme kinetic reaction, and effectively realizing Equation (8.4). Figure 8.4 shows the inner and outer solution for the parameters in Figure 8.3. For most enzyme kinetic reactions, k_1 is usually magnitudes larger than k_2 , which means that there is initially a rapid drop in the enzyme and substrate concentrations. It is for this reason that we have split the solution into an early phase or inner solution and a later phase, called outer solution. For most practical cases, it is very difficult to take measurements at very short time intervals. See [Rub75] for a comprehensive discussion.

8.2 Michaelis-Menten modelling

Phosphorylation steps in signaling cascades are enzyme kinetic reactions, the kinase facilitating the phosphorylation of a substrate. However, even for the relatively simple system of an enzyme kinetic reaction



```

function enzyme_kinetic
p = {[1 2],3,3}; % species indicator
l = {[1 1],1,1}; % stoichiometric coefficients
k = 60*[10 2 0.02]; % rate constants [(nM*min)^-1 min^-1 min^-1]
e0 = 0.5; s0 = 1; c0 = 0; p0 = 0;
S0 = [e0 s0 c0 p0]; % initial concentrations (nM)
tf = 10; % simulation end time (min)
% Step changes in all species in all reactions
%      S1  S2  S3  S4
nu = [ -1  -1  1  0      % R1
       1   1 -1  0      % R2
       1   0 -1  1  ]; % R3
sol = ode15s(@comp_dSdt,[0 tf],S0,[],p,l,k,nu);
tau0 = 3/( k(1)*s0 + k(2) + k(3) ); % separating solutions at tau0
ti = linspace(0,tau0,1e3); to = linspace(tau0,tf,1e3);
Si = deval(sol,ti); plot(ti,Si) % inner solution
So = deval(sol,to); plot(to,So); ti = 60*ti; % outer solution
% Subfunction to calculate the GMA equation:
function dSdt = comp_dSdt(t,S,p,l,k,nu)
M = length(k); % Number of reaction channels
for u=1:M
    P(u) = prod(S{p{u}(:)}).^l{u}(:);
end
dSdt = ( k.*P*nu )';

```

Fig. 8.3: Matlab file to simulate an enzyme-kinetic reaction as an example of the mass action model. The solutions of the differential equations are split into two parts, which are shown in Figure 8.4.

it is already rather difficult to obtain an analytical solution¹ to the set of differential equations (8.3). Furthermore, not all variables may be observable, i.e., measurable, or identifiable from experimental data. Here we are going to discuss commonly used assumptions and simplifications.

Above we realized that the total enzyme is constant, $[E] + [ES] = c_1$. Ignoring degradation and reconstitution of the enzyme, the constant can be evaluated from the initial conditions:

$$[E](t) + [ES](t) = [E](0) + [ES](0) . \quad (8.5)$$

Using (8.5) to eliminate $[E]$, we obtain

$$\dot{[S]} = -k_1([E](0) + [ES](0) - [ES])[S] + k_2[ES] \quad (8.6)$$

$$\dot{[ES]} = k_1([E](0) + [ES](0) - [ES])[S] - (k_2 + k_3)[ES] . \quad (8.7)$$

which, together with initial conditions $[S](0)$ and $[ES](0)$ can be numerically solved. The solution for $[P]$ can be derived easily from the solution of $[ES]$. The solution of $[E]$ is given by (8.5).

From our discussion in the previous section, we know that the substrate concentration $[S]$ steadily decreases. However, if it is the case that the available amount of substrate is relatively large, i.e., we might consider it as unchanged for a suitable period of time, we would have $[S](t) = [S](0)$, so that (8.6) is not required. For this *steady state assumption* (w.r.t. $[S]$), we are left with only (8.7)

$$\dot{[ES]} = k_1([E](0) + [ES](0) - [ES](t))[S](0) - (k_2 + k_3)[ES](t) . \quad (8.8)$$

¹Schnell [SM97] describes a closed form solution employing the omega function.

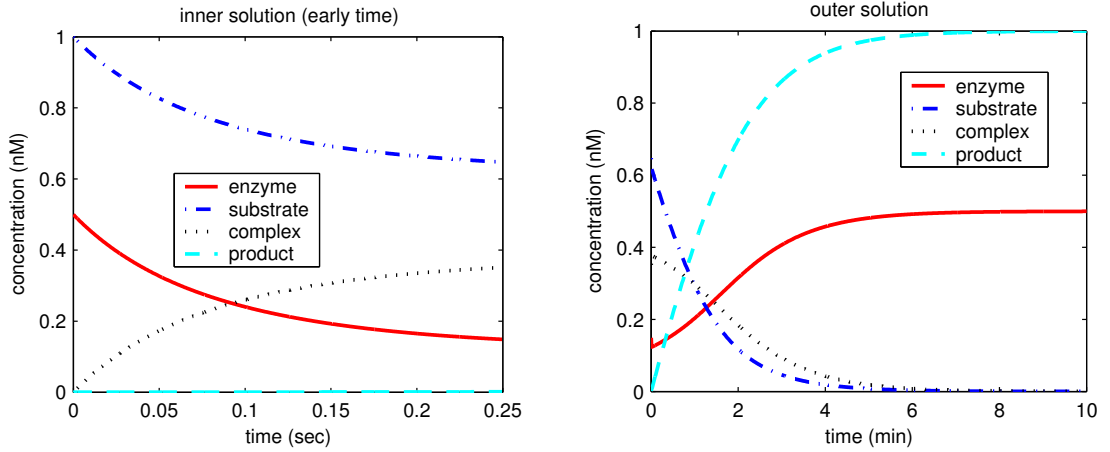


Fig. 8.4: Temporal evolution of substrate, enzyme, complex and product for the enzyme kinetic reaction. The plot on the left shows the inner solution

Denoting the steady state of $[ES]$ by the constant $[\tilde{E}S]$, and inserting this into (8.8)

$$k_1([E](0) + [ES](0) - [\tilde{E}S])[S](0) - (k_2 + k_3)[\tilde{E}S] = 0 ,$$

from which we obtain an expression for $[\tilde{E}S]$

$$[\tilde{E}S] = \frac{([E](0) + [ES](0))[S](0)}{K_M + [S](0)} ,$$

where

$$K_M = \frac{k_2 + k_3}{k_1} \quad (8.9)$$

is called the Michaelis- or Michaelis-Menten constant. Denoting the deviation of the complex from its steady state by $[ES]'$

$$[ES]' = [ES](t) - [\tilde{E}S]$$

and substituting this expression into (8.8)

$$[ES](t) = [ES]'(t) + [\tilde{E}S] , \quad (8.10)$$

and

$$\frac{d[ES]}{dt} = \frac{d[ES]'}{dt} .$$

It follows

$$\frac{d[ES]'}{dt} = -\lambda[ES]' \quad \text{where} \quad \lambda = k_1[S](0) + k_2 + k_3 ,$$

the solution of which is

$$[ES]'(t) = c \cdot e^{-\lambda t} ,$$

where c is a constant we obtain from initial conditions. From (8.10)

$$[ES](t) = c \cdot e^{-\lambda t} + [\tilde{E}S] .$$

The initial condition for $[ES]$ gives us then an expression for the constant c :

$$[ES](0) = c + [\tilde{E}S] , \quad \text{or} \quad c = [ES](0) - [\tilde{E}S] ,$$

leading to a solution for the temporal evolution of the complex concentration

$$[ES](t) = ([ES](0) - [\tilde{E}S]) \cdot e^{-\lambda t} + [\tilde{E}S] . \quad (8.11)$$

In conclusion, if substrate $[S]$ can be considered constant, then the complex concentration $[ES](t)$ approaches asymptotically the steady state $[\tilde{ES}]$, regardless of its initial conditions.

steady-state If we therefore let $[\dot{ES}] \approx 0$, this *quasi steady-state assumption* (w.r.t. $[ES]$) applied to (8.7), gives us the following expression

$$[ES](t) = \frac{(E[0] + [ES](0)) [S](t)}{[S](t) + K_M}.$$

Substitute this into (8.6) gives an expression that forms the basis for many experimental textbook settings:

$$[\dot{S}] = -\frac{k_3 (E[0] + [ES](0)) [S](t)}{[S](t) + K_M}$$

The value $V = |[\dot{S}]|$ is called the (initial) velocity of the reaction. Assuming $[ES](0) = 0$, the equation above is commonly written as

$$V = \frac{k_3[E](0)[S]}{K_M + [S]}.$$

Because $\partial V/\partial[S] > 0$, the reaction velocity is an increasing function of the substrate concentration. The maximum value of V , i.e., the maximum rate by which a product can be formed, is approached for very large values of $[S]$

$$V_{\max} = \lim_{[S] \rightarrow \infty} V = k_3[E](0)$$

and is called the *limiting rate*. Dividing this maximum rate by the enzyme concentration is called the *turnover number*. Typical turnover numbers are 1000 substrate molecules processed per second per enzyme molecule [AJL⁺02]. limiting rate

If the assumptions made above are realistic, the equation for V can be written as follows

$$V = \frac{V_{\max} \cdot [S]}{K_M + [S]}. \quad (8.12)$$

The Michaelis-Menten constant K_M gives the initial substrate concentration at which the reaction velocity is half maximal (since for $[S](0) = K_M$ substituted above gives $V = V_{\max}/2$). K_M is therefore an approximate measure of substrate affinity for the enzyme. A low K_M value means that the enzyme reaches its maximum catalytic rate at a low substrate concentration, which generally indices a tighter substrate binding.

Since V can be measured as a function of $[S](0)$, equation (8.12) allows us to estimate V_{\max} and K_M from curve fitting. To this end, we rearrange (8.12)

$$\frac{1}{V} = \frac{1}{V_{\max}} \left(1 + \frac{K_m}{[S]} \right).$$

This represents a straight line with slope K_M/V_{\max} and intercept $1/V_{\max}$. Plotting values of $1/V$ against $1/[S](0)$ is referred to the Lineweaver-Burk plot.

8.3 Multinomial Systems

In our discussion of stochastic modeling and simulation, equation (2.19) represented GMA models. The general structure of these differential equations is of the form²

$$\dot{x}_i = \sum_{k=1}^n \theta_{ik} \prod_{j=1}^n x_j^{l_{ijk}}. \quad (8.13)$$

²Dealing with differential equations we use n not as the state-vector but as an integer and limit to sums. Similar, rather than referring to molecular species S we denote all variables with x , a notation that is most frequently used in systems theory.

Including m independent variables, we write (see also page 37)

$$\dot{x}_i = \sum_{k=1}^{n+m} \theta_{ik} \prod_{j=1}^{n+m} x_j^{l_{ijk}} . \quad (8.14)$$

The mathematical structure was introduced by W.Peschel and W.Mende and is referred to as *multinomial systems*. Applied to biochemical reaction networks, these equations are the generalized mass action models (2.19). The set of equations for the enzyme kinetic reaction (4.23) is an example for a GMA system. We obtain the matrix of coefficients

$$[\theta_{ik}] = \begin{bmatrix} -k_1 & k_2 & k_3 & 0 \\ -k_1 & k_2 & 0 & 0 \\ k_1 & -k_2 & -k_3 & 0 \\ 0 & 0 & k_3 & 0 \end{bmatrix}$$

and the set of matrices $L_i = [l_{jk}]_i$ for the powers p_{jk} in equation i :

$$L_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad L_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad L_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad L_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

While the class of models, which is defined by (8.14) seems very general and it may seem difficult to conduct a formal analysis of these equations without turning to numerical solutions, we ought to remember that the values x_i cannot be negative and that although the coefficients can be positive or negative, which sign applies is predefined by the model structure or the reaction diagram respectively.

The ability to derive general properties of the dynamic behavior of a system, independent of specific parameter values, is the most attractive aspect of a formal mathematical analysis. If we consider the sparse data sets we obtain from experiments, such analysis would benefit parameter estimation and experimental design. If we are able to establish an order relationship between parameters, e.g. “ $k_1 \gg k_3$ ”, this would very useful in guiding parameter estimation, or providing confidence in using parameter values from literature. Due to experimental uncertainties, absolute values have little value, and an analysis in terms of basic temporal profiles (e.g. “ x_1 decays exponentially”, “ x_2 peaks before $[P]$ ”, “ x_1 is pulled down”, “ x_2 is delayed”) is at the heart of the biologists reasoning.

An example for a model of the ERK pathway, is the following set of equations:

$$\begin{aligned}
\dot{x}_1 &= -k_1x_1x_2 + k_2x_3 - k_{10}x_1x_{12} + k_{11}x_{13} + k_6x_7 + k_{14}x_{14} \\
\dot{x}_2 &= -k_1x_1x_2 + k_2x_3 \\
\dot{x}_3 &= k_1x_1x_2 - (k_2 + k_3)x_3 \\
\dot{x}_4 &= k_3x_3 \\
\dot{x}_5 &= k_3x_3 - k_4x_5x_6 + k_5x_7 \\
\dot{x}_6 &= -k_4x_5x_6 + k_5x_7 + k_9x_{10} \\
\dot{x}_7 &= k_4x_5x_6 - (k_5 + k_6)x_7 \\
\dot{x}_8 &= k_6x_7 - k_7x_8x_9 + k_8x_{10} \\
\dot{x}_9 &= k_{14}x_{14} - k_7x_8x_9 + k_8x_{10} \\
\dot{x}_{10} &= k_7x_8x_9 - (k_8 + k_9)x_{10} \\
\dot{x}_{11} &= k_9x_{10} - k_{12}x_{11}x_{13} + k_{13}x_{14} \\
\dot{x}_{12} &= -k_{10}x_1x_{12} + k_{11}x_{13} + k_{15}x_{15} \\
\dot{x}_{13} &= k_{10}x_1x_{12} - k_{11}x_{13} - k_{12}x_{11}x_{13} + k_{13}x_{14} \\
\dot{x}_{14} &= k_{12}x_{11}x_{13} - (k_{13} + k_{14})x_{14} \\
\dot{x}_{15} &= k_{14}x_{14} - k_{15}x_{15} .
\end{aligned}$$

The structure of these equations is determined from knowledge of the proteins (x_i) involved. For some of these proteins we can obtain experimental time course data but not for all. The question is then whether we could extract from this system of equations basic relationships between the k 's, considering that we are looking for basic temporal profiles (all of which either converge to zero or some steady state)?

Since the structure of these equations is fairly well defined (sums of simple products of variables, the signs of terms are given, the parameters are always positive), one might imagine some 'qualitative' analysis of the kind described above: Given time course data for some of the variables, we first of all wish to validate the model structure (e.g. test for feedback through an additional negative term in one of the equations). The model structure is in this sense more important than knowing exact parameter values. We elaborate on these issues further in Chapter 6.

8.4 S-Systems

The mathematical structure (8.13) suggests a form in which we separate positive terms (complex formation, production) from negative terms (dissociation, degradation, depletion):

$$\dot{x}_i = V_i^+(x_1, \dots, x_n) - V_i^-(x_1, \dots, x_n) .$$

If we are to include m independent variables we write

$$\dot{x}_i = V_i^+(x_1, \dots, x_n, x_{n+1}, \dots, x_{n+m}) - V_i^-(x_1, \dots, x_n, x_{n+1}, \dots, x_{n+m}) . \quad (8.15)$$

where $i = 1, \dots, n$. This general format allows for different classes of representations, one of which are S-systems.

The S-Systems approach, mostly developed by M.Savageau [Sav69a, Sav69b, Sav70] and E.O.Voit [Voi00], starts with the general description (8.15)

$$\dot{x}_i = V_i^+ - V_i^- \quad \text{for } i = 1, 2, \dots, n ,$$

where the general functions are reduced to simple products of the variables involved. Such power law representation has some attractive mathematical features, but is implicitly

based on a Taylor series approximation around a steady state value. This is alright for studying metabolic fluxes but does not work for transient phenomena in, for instance, signal transduction pathways.

For example, considering the conversion of x_1 into x_2 , catalyzed by x_3 . We assume a constant influx α to replenish x_1 . The degradation of x_1 depends on the concentration or pool size of x_1 itself and also on the enzyme x_3 :

$$\dot{x}_1 = \alpha - V_1^-(x_1, x_3) .$$

The production of x_2 is described in the same way as the degradation of x_1 , $V_2^+ = V_1^-$. Finally, the degradation of x_2 depends only on its current concentration or pool size:

$$\dot{x}_2 = V_1^-(x_1, x_3) - V_2^-(x_2) .$$

The S-systems approach would then choose the following power-law representations for V_1^- and V_2^- :

$$\begin{aligned} V_1^-(x_1, x_3) &= \beta x_1^a x_3^b , \\ V_2^-(x_2) &= \gamma x_2^c . \end{aligned}$$

For a general S-system we write

$$\dot{x}_i = \alpha_i \prod_{j=1}^{n+m} x_j^{g_{ij}} - \beta_i \prod_{j=1}^{n+m} x_j^{h_{ij}} \quad \text{for } i = 1, 2, \dots, n ,$$

and α denoting the rate constant for the production of each pool and β for its degradation or loss. These values can be positive or zero but not negative. There are numerous publications, exploring the theoretical properties of S-systems and applying them to metabolic pathways.

8.5 The Heinrich Model

In [HNR02], Heinrich and colleagues demonstrated that even without experimental data to estimate parameter values, mathematical modelling allows an interesting study of protein kinase signal transduction pathways. Amongst other things they concluded from their study that:

- The simplest model pathways allow amplified signalling only at the expense of slow signal transduction.
- Phosphatases have a more pronounced effect than kinases on the rate and duration of signalling, whereas signal amplification is controlled primarily by kinases.

Heinrich first considers a linear signalling cascade with the stimulation of a receptor, forming the upper part of the pathway and subsequent sequential activation of downstream proteins through phosphorylation via kinases. The output of the pathway is the phosphorylation of a protein, which is assumed to have further consequences, e.g. on the activation of transcription factors and thereby influencing transcription of genes. Deactivation of proteins is realized through dephosphorylation via phosphatases in case of intermediate pathway elements. Receptors can thought of being deactivated by receptor dephosphorylation, internalization of the receptor-ligand complex, and/or degradation of the receptor or ligand.

Considering the enzyme-kinetic reaction as a template for activation and deactivation, and assuming that the concentration of each kinase-substrate complex in the pathway is small compared to the total concentration of the reaction partners, and assuming that the concentration of active phosphatase is constant, each phosphorylation step or activation is described as a second-order or bi-molecular reaction, where the phosphorylated form x_{i-1} of protein $i - 1$ one step up in the pathway takes the role of the kinase which facilitates the activation of the unphosphorylated form \tilde{x}_i of the next downstream protein, referred to as the substrate of the reaction.



The phosphorylation rate is given by the expression

$$v_{p,i} = \tilde{\alpha}_i x_{i-1} \tilde{x}_i ,$$

where $\tilde{\alpha}_i$ is the phosphorylation second-order rate constant for phosphorylation of the i th kinase. Assuming that the concentration of active phosphatase is constant, dephosphorylation is modelled as a first order reaction with dephosphorylation rate

$$v_{d,i} = \beta_i x_i ,$$

where β_i is the rate constant for dephosphorylation by the i th phosphatase. Defining

$$c_i = \tilde{x}_i + x_i$$

as the total concentration of protein kinase i . The differential equation for the phosphorylation or activation of x_i is given by

$$\begin{aligned} \frac{dx_i}{dt} &= v_{p,i} - v_{d,i} \\ &= \tilde{\alpha}_i x_{i-1} \tilde{x}_i - \beta_i x_i \end{aligned}$$

Let

$$\alpha_i = \tilde{\alpha}_i c_i$$

be a pseudo-first-order rate constant, so that we can write

$$\frac{dx_i}{dt} = \alpha_i x_{i-1} \left(1 - \frac{x_i}{c_i} \right) - \beta_i x_i , \quad (8.16)$$

which we also introduced in Section 7. The first step of the pathway, receptor stimulation, is modelled as

$$\frac{dx_1}{dt} = \alpha_1 u(t) \left(1 - \frac{x_1}{c_1} \right) - \beta_1 x_1 , \quad (8.17)$$

where $u(t)$ is the concentration profile of the activated receptor. For example, the inactivation of the receptor may be modelled as $u(t) = \exp(-\lambda t)$, where $1/\alpha$ is the time constant of the receptor. For $\lambda \rightarrow 0$, the pathway is permanently activated. Heinrich *et al.* then introduce the *signalling time* as the average time to activate protein i

signaling time

$$\tau_i = \frac{T_i}{l_i} \quad \text{where} \quad l_i = \int_0^\infty x_i(t) dt \quad \text{and} \quad T_i = \int_0^\infty t x_i(t) dt . \quad (8.18)$$

l_i denotes the total amount of active protein i , generated during the signaling period. The *signal duration* is defined by

signal duration

$$\theta_i = \sqrt{\frac{Q_i}{l_i} - \tau_i^2}, \quad \text{where } Q_i = \int_0^\infty t^2 x_i(t) dt. \quad (8.19)$$

For a *weakly activated* pathway all of its proteins are phosphorylated to a low degree such that $x_i \ll c_i$. As a consequence, Equation (8.16) becomes

$$\frac{dx_i}{dt} = \alpha_i x_{i-1} - \beta_i x_i. \quad (8.20)$$

Heinrich *et al.* showed that the signaling time and signal duration can for this case be calculated explicitly:

$$\text{signalling time: } \tau = \frac{1}{\lambda} + \sum_{i=1}^n \frac{1}{\beta_i}, \quad (8.21)$$

$$\text{signal duration: } \theta = \sqrt{\frac{1}{\lambda^2} + \sum_{i=1}^n \frac{1}{\beta_i^2}}. \quad (8.22)$$

8.6 The MAP Kinase (MAPK) Pathway

This section is to introduce an important class of signaling pathways. For unfamiliar biochemical expressions the reader is referred to the glossary on page 287.



Fig. 8.5: Compact representation of the MAPK pathway.

The mitogen-activated protein kinase (MAPK) cascade, is part of the growth-factor/Ras pathway in eucaryotic cells. The cascade³ is highly *conserved*, which means that the same principles can be observed in a variety of organisms and cell types.

The core of this pathway is formed by a module which is defined by three protein kinases: MAPKKK (e.g. RAS/Raf), MAPKK (e.g. MEK) and MAPK. This module is activated by a collection of proteins, some of which have to occur in combination. The first element of the module to be activated is the MAPKKK. The activated MAPKKK phosphorylates MAPKK at two sites. This double phosphorylated MAPKK, denoted MAPKK**, acts as a threonine/tyrosine kinase and phosphorylates MAPK at two sites of the protein structure. MAPK can then act as a kinase for transcription factors, but may also have a feedback effect on the activity of kinases like the MAPKKK further *upstream*.

The ERK (extracellular-signal-regulated kinase) pathway is an example for a MAPK cascade, which features Ras as the G-protein, Raf as MAPKKK, MEK (MAPK/ERK kinase) as MAPKK and ERK as MAPK. Ras and Raf are proto-oncogenes which explains the interest in this pathway [Kol00]. Most stimuli to the receptor leads to an activation of the G-protein Ras by inducing the exchange of GDP with GTP. GDP and GTP are therefore also referred to as *exchange factors*. This exchange will convert the Ras molecule into its active conformation. Ras resides near the cell membrane and one says that the exchange factors are *recruited*. The protein SOS (son of sevenless) is another Ras exchange

³Biologists refer to an unbranched sequence of modules combined in a cascade as a *linear* cascade. The term ‘linear’ has not relationship with the question whether the biochemical reactions and their mathematical model are linear or nonlinear. See page 8 for a definition a linear model. The upper - membrane near parts of the pathway are also referred to ‘upstream regions’.

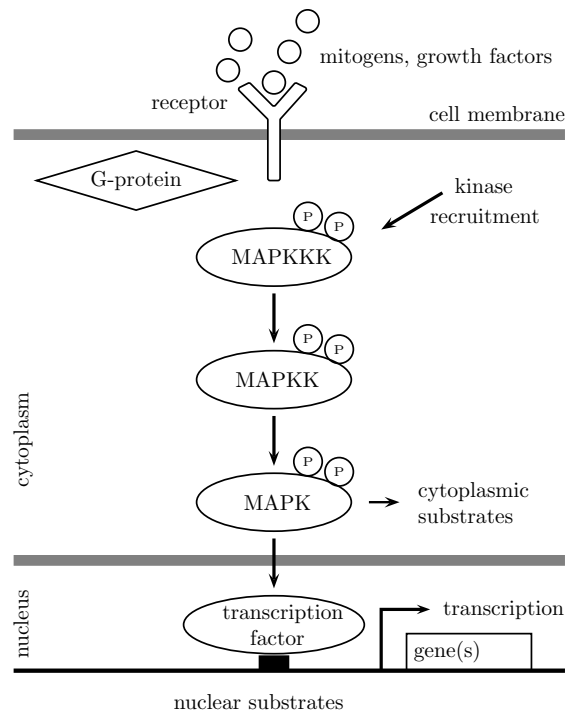


Fig. 8.6: The MAP kinase pathway. The core of this pathway is formed by a module which is defined by three protein kinases: MAPKKK, MAPKK and MAPK. Drawing adapted from [Kol00].

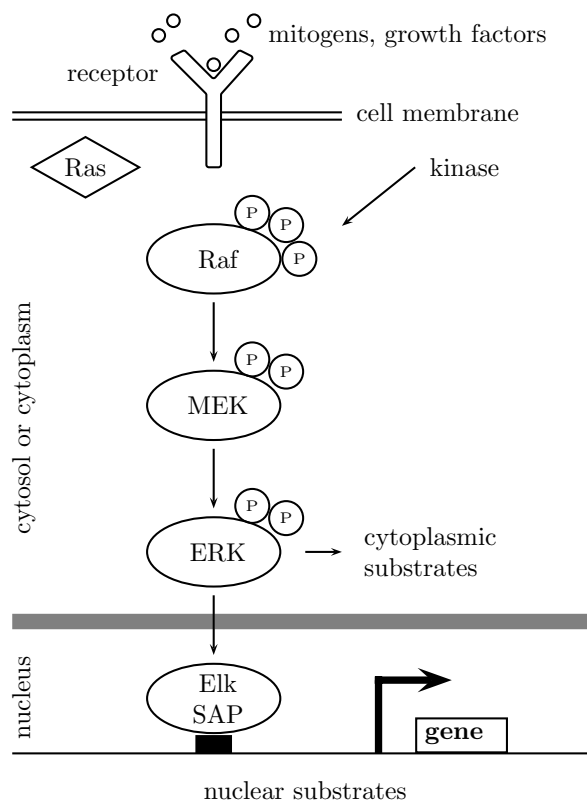


Fig. 8.7: Basic outline of the Ras/Raf/MEK/ERK pathway. Drawing adapted from [Kol00].

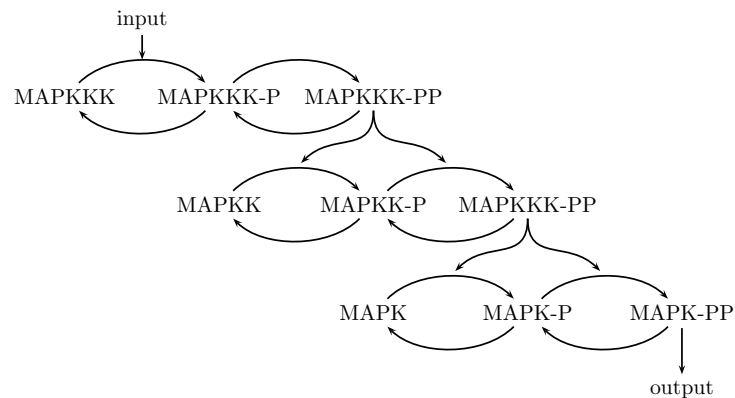


Fig. 8.8: General outline of the MAPK pathway.

factor, which can terminate Ras activation. This termination proceeds by phosphorylation of SOS, which leads to the disassembly of the complex. The phosphorylation of SOS is also feedback regulated by the activated ERK pathway [Kol00].

Activated Ras functions as an adaptor that binds to Raf kinases with high affinity and causes their translocation to the cell membrane, where Raf activation takes place.⁴ Mammals possess three Raf proteins which are also referred to as *isoforms*: Raf-2, A-Raf, and B-Raf. For all three, MEK acts as a substrate further ‘downstream’ of the pathway.

MEK is activated by phosphorylation of two serine residues in the activation loop. The most predominant activator of MEK in most cell types are Raf kinases. It is believed that Raf-1 can inhibit itself through some negative feedback. Raf seems to be suspended in a balance between activation and auto-inhibition [Kol00]. Raf can activate both MEK isoforms, MEK-1 and MEK-2, and both of them can activate the downstream ERK kinases. MEK is a ‘dual-specificity’ kinase which can phosphorylate both.

Finally, ERK is a serine/threonine kinase with more than 50 substrates [Kol00]. All components of the Ras/Raf/MEK/ERK pathway can interact with each other physically: Ras-GTP binds to Raf; Raf can bind to MEK; and MEK can bind to ERK.

8.7 The Ras/Raf/MEK/ERK Pathway

This section describes a mathematical model of the ERK module, as an example of a signal transduction pathway. The model is represented by a set of nonlinear differential equations. We show how this representation can be generalized to capture a large class of dynamic pathway models. In this framework, a pathway diagram corresponds to the state space of a dynamic system, while the entirety of dynamic processes that can occur in a particular pathway is defined by a one-parameter group of transformations in the manifold that is the state space. We are thus providing a conceptual framework in which to describe not only pathway diagrams but also the spatial-temporal interactions within and between cells.

Experimental data show that the inhibition of MEK phosphorylation by RKIP is not linear (Fig. 4.16). There is a threshold of RKIP expression that steeply reduces MEK phosphorylation. This is consistent with a positive feedback mechanism. If not all Raf is bound to RKIP, then there is Raf-1 available for interacting with and activating MEK.

⁴The spatial dimension, translocation of molecules, is an important aspect that is ignored by conventional models. One idea is to divide a space up into regions, model each region separately and allow for an exchange between them. We are thus in need for a multi-model concept.

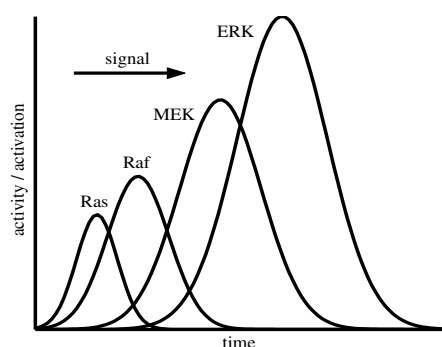


Fig. 8.9: Illustration of how a signal is relayed through sequential activation of proteins in the Ras/Raf/MEK/ERK pathway. Note that this picture is an idealized picture and actual amplitudes and base levels may differ significantly when considering experimental data. See also Figure 8.7.

MEK then activates ERK. Consequently, the positive feedback phosphorylation of RKIP by ERK will ensure that RKIP is phosphorylated and dissociates from Raf-1. In this situation there will be little or no inhibition of Raf-1 by RKIP. If the level of RKIP expression exceeds a certain threshold, all of the Raf-1 will be bound to RKIP. In this situation there is no phosphorylation of RKIP and no dissociation. Hence MEK phosphorylation remains inhibited.

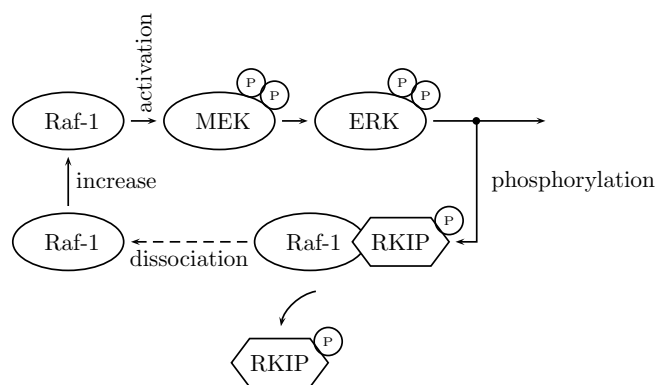


Fig. 8.10: Illustration of the interactions of RKIP with the Ras/Raf/MEK/ERK pathway.

To analyze the dynamics of the ERK signaling pathway, including the positive feedback mechanism, in both qualitative and quantitative manner, we constructed a mathematical model based on the mass action law and represented by nonlinear ordinary differential

equations (see Figure 8.11):

$$\left. \begin{aligned}
 \dot{x}_1 &= -k_1 x_1 x_2 - k_9 x_1 x_{12} + k_{10} x_{13} + k_5 x_7 + k_{13} x_{14} \\
 \dot{x}_2 &= -k_1 x_1 x_2 \\
 \dot{x}_3 &= k_1 x_1 x_2 - k_2 x_3 \\
 \dot{x}_4 &= k_2 x_3 \\
 \dot{x}_5 &= k_2 x_3 - k_3 x_5 x_6 + k_4 x_7 \\
 \dot{x}_6 &= -k_3 x_5 x_6 + k_4 x_7 + k_8 x_{10} \\
 \dot{x}_7 &= k_3 x_5 x_6 - (k_4 + k_5) x_7 \\
 \dot{x}_8 &= k_5 x_7 - k_6 x_8 x_9 + k_7 x_{10} \\
 \dot{x}_9 &= k_{13} x_{14} - k_6 x_8 x_9 + k_7 x_{10} \\
 \dot{x}_{10} &= k_6 x_8 x_9 - (k_7 + k_8) x_{10} \\
 \dot{x}_{11} &= k_8 x_{10} - k_{11} x_{11} x_{13} + k_{12} x_{14} \\
 \dot{x}_{12} &= -k_9 x_1 x_{12} + k_{10} x_{13} + k_{14} x_{15} \\
 \dot{x}_{13} &= k_9 x_1 x_{12} - k_{10} x_{13} - k_{11} x_{11} x_{13} + k_{12} x_{14} \\
 \dot{x}_{14} &= k_{11} x_{11} x_{13} - (k_{12} + k_{13}) x_{14} \\
 \dot{x}_{15} &= k_{13} x_{14} - k_{14} x_{15} .
 \end{aligned} \right\} \quad (8.23)$$

The mathematical structure is one of multinomial systems, concentrations x_i as well as the values for parameters k_i can only be positive. Each step in the pathway is modelled in analogy to an enzyme kinetic reaction (8.1). Possible algebraic simplifications of the model and parameter estimation are not discussed here as this is not the focus of the present section.

The Ras/Raf-1/MEK/ERK module is an ubiquitously expressed signaling pathway that conveys mitogenic and differential signals from the cell membrane to the nucleus [YJM+00]-[Kol00]. This kinase cascade appears to be spatially organized in a signaling complex nucleated by Ras proteins. The small G protein Ras is activated by many growth factor receptors and binds to the Raf-1 kinase with high affinity when activated. This induces the recruitment of Raf-1 from the cytosol to the cell membrane. Activated Raf-1 then phosphorylates and activates MAPK/ERK kinase (MEK), a kinase that in turn phosphorylates and activates Extracellular Kinase (ERK), the prototypic Mitogen-Activated Protein Kinase (MAPK). Activated ERKs can translocate to the nucleus and regulate gene expression by phosphorylation of transcription factors. This kinase cascade controls the proliferation and differentiation of different cell types. The specific biological effects of the kinase cascade are crucially regulated by the Raf-1 kinase inhibitor protein (RKIP) [YRD+01]. RKIP binds to Raf-1 thereby disrupting the interaction between Raf-1 and MEK. As a consequence RKIP inhibits the phosphorylation and activation of MEK by Raf-1. RKIP overexpression interferes with the activation of MEK and ERK, induction of AP-1 dependent receptor genes and transformation elicited by an oncogenically activated Raf-1 kinase [YSL+99].

Figure 8.12 shows a simulation of the ERK model for varying initial concentrations of the Raf-1 kinase inhibitor protein RKIP. The simulations show the rate of active Ras (x_2) binding to Raf-1 (x_1) linearly decreasing along with the initial value of RKIP. The plots demonstrate how the initial signal transduction through active Ras is interrupted by RKIP. The variation profile of active Raf (x_5) as a function of variations of the initial value for RKIP is similar to active MEK (x_8) and active ERK (x_{11}). The dynamics of these proteins in Figure 8.12 also exhibit the nonlinear relationships encapsulated by the model. At low initial concentration of RKIP all signal proteins are completely activated although with different time lags. At high concentrations of RKIP, the activation ratio

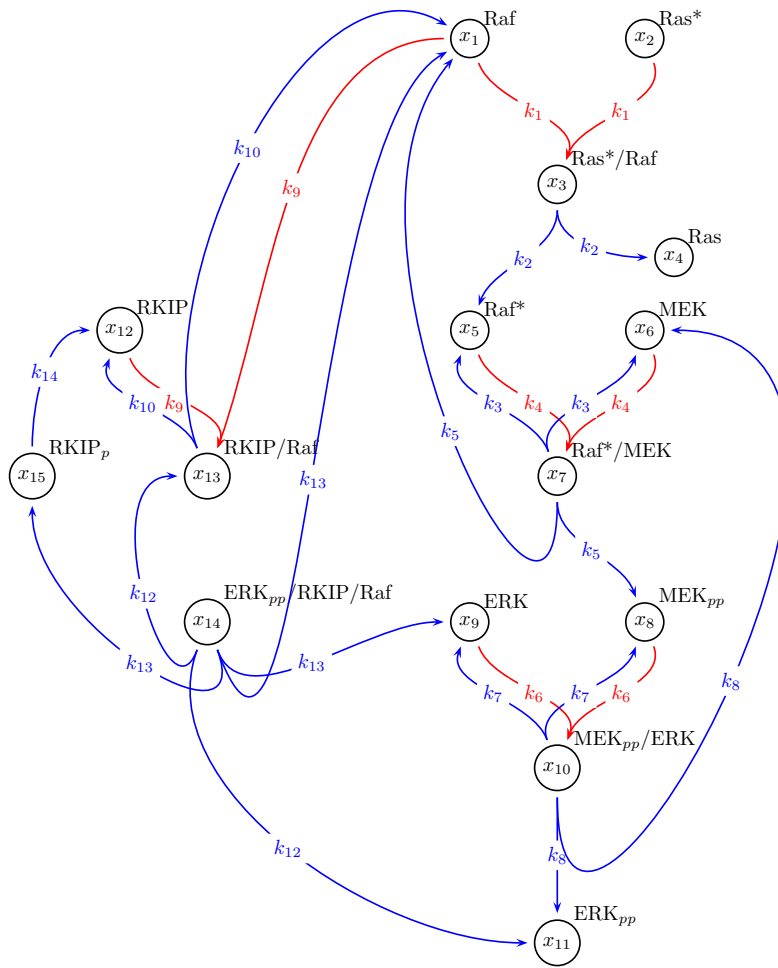


Fig. 8.11: Graphical representation of the model for Ras/Raf-1/MEK/ERK signaling module, describing a positive feedback loop between RKIP and double phosphorylated ERK (ERK_{pp}).

is about zero. These simulation results show that there is a threshold of concentration of RKIP that steeply reduces the phosphorylation of each protein, which is in accordance with the data in Figure 4.16.

The concept of a pathway is the framework in which a molecular- or cell biologist captures her/his knowledge of pathway variables, their states and relationships. A pathway model - whether a simple diagram or a mathematical representation like the one described here is an abstraction. In [RS02] the authors discuss the importance of abstractions in science and suggest an abstract computer language like π -calculus for pathway modelling. This language provides a means to formalize the knowledge of components and their interactions. In order to simulate a model in this framework, the process algebra requires the use of, for instance, the Gillespie algorithm in order to compile the model. Here we are seeking an abstract (and thus generally applicable) algebraic framework in which to discuss the dynamic properties of a pathway but also the entirety of dynamic processes the cell can realize as well as relationships between cells. Instead of a computer language, we begin with chemical rate equations and work ‘upwards’ through generalizations of the models that can be constructed to represent the dynamic interactions in pathways. The main argument for a π -calculus is its “linguistic structure” and “operational semantics” from which causal relationships can be derived. The motto of the approach presented in this section could be that we wish to realize what the biologist *could see* rather than what he *says*.

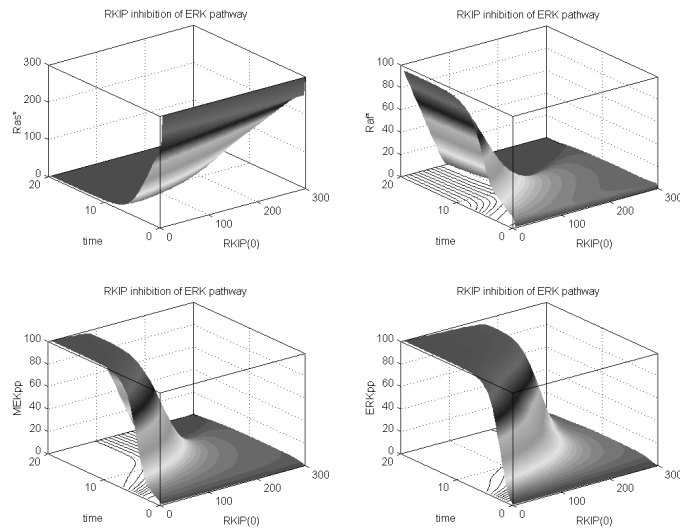
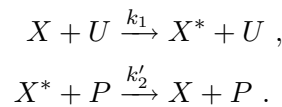


Fig. 8.12: Simulation of the ERK model for varying initial concentrations of RKIP.

8.8 Feedback and Oscillations in Signalling Pathways

In Figure 8.6 a prototypical mitogen-activated protein kinase (MAPK) cascade is shown [FJ96, Kol00]. MAP kinase pathways have been found in many organisms and cell types. They are also an important system for cancer research studies [Kol00, YSL⁺99, YJM⁺00] and have been considered in various modelling exercises (e.g. [FJ96, HFJ96, BI99, Kho00, BF00, AL01, K⁺02]).

Let us model phosphorylation and dephosphorylation as a reversible bimolecular reaction, where the phosphorylation of X into X^* is facilitated by the kinase U , and dephosphorylation by phosphatase P :



We write \tilde{x} for the unphosphorylated form of a molecular species or protein X , u for the kinase U , p for phosphatase P and x to denote the activated, i.e., phosphorylated protein X^* . Referring to the law of mass action, we obtain the following set of ordinary differential equations

$$\begin{aligned} \frac{d}{dt}x &= k_1u(t)\tilde{x}(t) - k'_2p(t)x(t) && : \text{phosphorylation,} \\ \frac{d}{dt}\tilde{x} &= -k_1u(t)\tilde{x}(t) + k'_2p(t)x(t) && : \text{dephosphorylation.} \end{aligned}$$

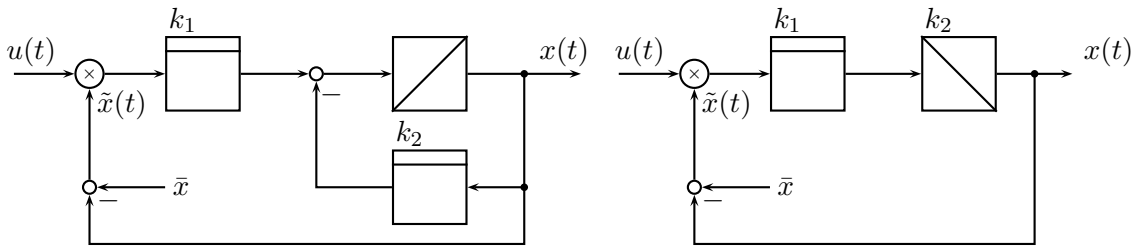
To simplify the mathematical model we assume that the phosphatase is constant. This means we can merge p and k'_2 into k_2 . Together with the conservation for a constant total of x ,

$$\bar{x} = x(t) + \tilde{x}(t),$$

we obtain the following differential equation, in one variable, to describe the phosphorylation of protein X :

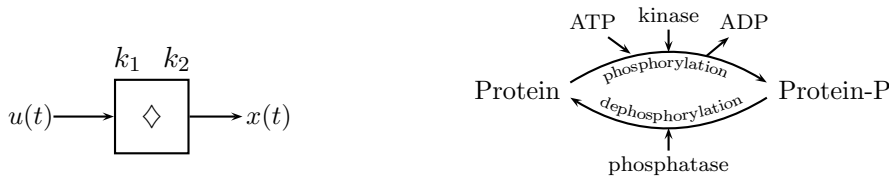
$$\frac{d}{dt}x = k_1u(t)(\bar{x} - x(t)) - k_2x(t).$$

The block diagram for a signaling step, (de-)phosphorylation is readily obtained:



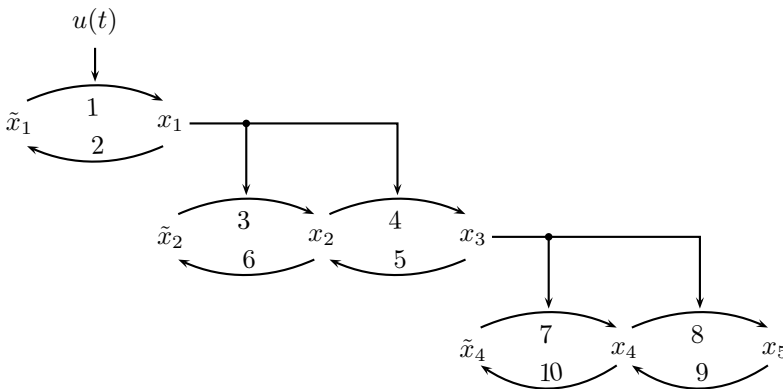
In this model for (de-)phosphorylation, it is assumed that kinase-substrate concentrations are low compared to the total concentration of the reactant species. Furthermore, for dephosphorylation to be considered a first-order reaction, it is assumed that the active phosphatase concentration is constant.

For more complex systems, let us now collapse the previous diagram into a single block to represent the (de-)phosphorylation of \tilde{x} through kinase u :



The diagram on the right illustrates the conventional, and more detailed, representation. Phosphorylation is facilitated by means of a kinase and dephosphorylation is realized by a phosphatase. The phosphorylated state is commonly denoted by adding -P to the name of the protein.

Given such a module, we are now in a position to construct more complex pathway diagrams by connecting these basic components into cascades and networks. For example, for the pathway in Figure 8.6 the map is [Kho00]:



In the diagram x_1 corresponds to activated MKKK-P, x_2 to MKK-P, x_3 to the double-phosphorylated MKK-PP, x_4 to MAPK-P, x_5 to MAPK-PP. Inactivated forms are denoted with a tilde \tilde{x} . In signalling activation/inactivation of proteins corresponds to phosphorylation/dephosphorylation, while in some cases one considers double-phosphorylations:

$$\begin{aligned}
 v_1 &= k_1 u(t) \tilde{x}_1(t), & v_6 &= k_6 x_2(t) \\
 v_2 &= k_2 x_1(t), & v_7 &= k_7 x_3(t) \tilde{x}_4(t) \\
 v_3 &= k_3 x_1(t) \tilde{x}_2(t), & v_8 &= k_8 x_3(t) x_4(t) \\
 v_4 &= k_4 x_1(t) x_2(t), & v_9 &= k_9 x_5(t) \\
 v_5 &= k_5 x_3(t), & v_{10} &= k_{10} x_4(t)
 \end{aligned}$$

The model is derived from the fact that

$$\begin{aligned} \dot{x}_1 &= v_1 - v_2 & \dot{x}_2 &= v_3 - v_6 - v_4 + v_5 \\ \dot{x}_3 &= v_4 - v_5 & \dot{x}_4 &= v_7 - v_{10} + v_9 - v_8 \\ \dot{x}_5 &= v_8 - v_9 \end{aligned}$$

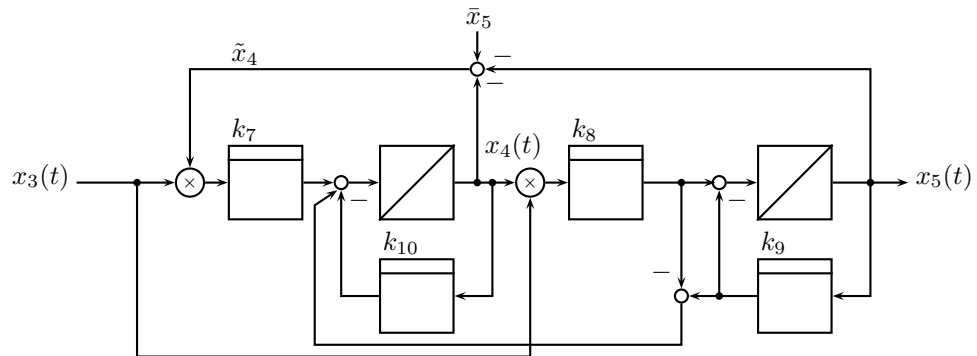
Inserting the reaction rates into these equations we obtain the following set of equations for the activated proteins in the pathway:

$$\begin{aligned} \dot{x}_1 &= k_1 u(t) \tilde{x}_1(t) - k_2 x_1(t) \\ \dot{x}_2 &= k_3 x_1(t) \tilde{x}_2(t) - k_6 x_2(t) \underbrace{-k_4 x_1(t) x_2(t) + k_5 x_3(t)}_{-\dot{x}_3} \\ \dot{x}_3 &= k_4 x_1(t) x_2(t) - k_5 x_3(t) \\ \dot{x}_4 &= k_7 x_3(t) \tilde{x}_4(t) + k_9 x_5(t) \underbrace{-k_8 x_3(t) x_4(t)}_{-\dot{x}_5} - k_{10} x_4(t) \\ \dot{x}_5 &= k_8 x_3(t) x_4(t) - k_9 x_5(t) . \end{aligned}$$

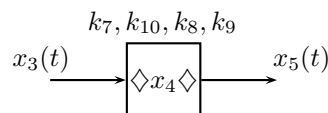
In addition, the following conservation relationships hold:

$$\begin{aligned} \bar{x}_1 &= \tilde{x}_1(t) + x_1(t) \\ \bar{x}_3 &= \tilde{x}_3(t) + x_2(t) + x_3(t) \\ \bar{x}_5 &= \tilde{x}_4(t) + x_4(t) + x_5(t) \end{aligned}$$

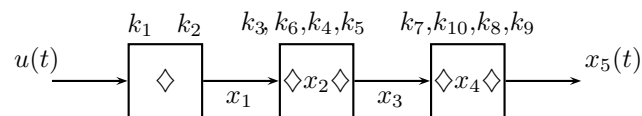
Using the mathematical model for (de-)phosphorylation from above, the double phosphorylation of x_4 into x_5 is described by the following block diagram:



This again we can collapse into a single block, without loss of information:



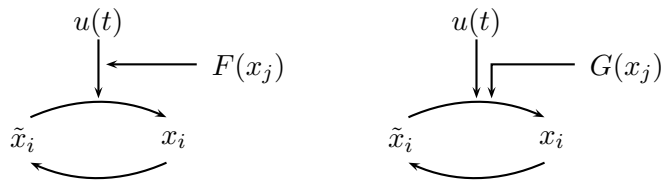
The MAP kinase pathway can then be represented by the compact block diagram:



This system is then another example of the state-space representation, where for $u(t)$ we might assume a negative exponential function $u(t) = e^{-\lambda t}$, where an initial concentration of ligands is depleted through binding to the receptors on the cell surface. The greater the value of λ , the faster the ligands bind with receptors to form complexes. In [HNR02], these series-connected submodels of (de-)phosphorylation, have been used to analyze pathways for their dynamic properties. In particular the authors derived expression for the *signaling time*, defined as the average time to activate a protein in the pathway and the *signal duration*, characterized by an integral of the concentration profile. In [Kho00] a very similar model as the one above is modified by introducing a negative feedback loop between the end product MAPK-PP and a Ras/MAPK complex at the top of the pathway. Kholodenko showed how ultrasensitivity, leading to switch-like behavior, combined with negative feedback can bring about sustained oscillations in this pathway. Considering populations of cells, this may be of particular interest in the context of the synchronization of coupled oscillators, which has been observed in a range of biological and physical systems [Str00a]. Approaches to test for feedback loops have been presented in [K+02, AFJS04]. In [BI99, BF00, AL01] computational studies of feedback effects on signal dynamics in a detailed MAPK pathway model are presented.

Feedback in Signalling Pathways

We continue the previous section, introducing feedback loops from a protein x_j further down the pathway, up to x_i . We have two options indicated in the following diagrams:



On the left-hand side, feedback manifests itself as a multiplicative term in the differential equation for x_i :

$$\frac{d}{dt}x_i(t) = k_1 u(t) F(x_j) (\bar{x}_i - x_i(t)) - k_2 x_i(t) ,$$

where for the function $F(x_j)$ we can choose from the following commonly used versions:

$$F(x_j) = \frac{1}{1 + \left(\frac{x_j}{K_I}\right)^n} , \quad F(x_j) = \frac{c}{d^n + x_j^n} , \quad F(x_j) = \frac{cx_j^n}{d^n + x_j^n} ,$$

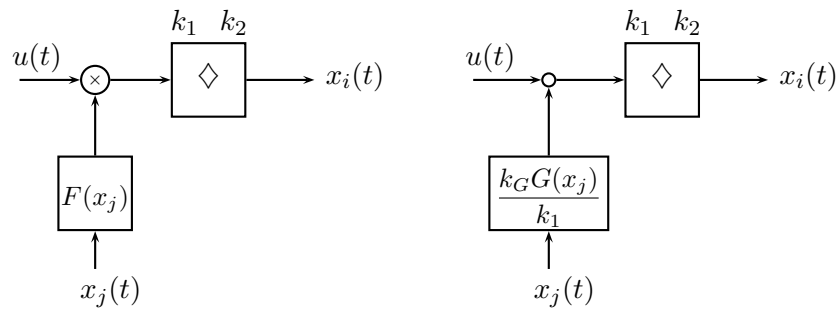
where $n \geq 1$ defines the steepness of the feedback function and the subscript I of K_I stands for ‘inhibition’. The main requirement for the choice of a function $F(x_j)$ is that at $x_j = 0$, we should have $F(x_j) = 1$. Mechanistic interpretations and experimental evidence for these functions are discussed in [Gol96, FJ96, HFJ96, LQ03]. Note the distinction between a *mechanistic* (or physical) and a *operational* (or phenomenological) definition for an interaction. An operational definition is based on observations, not necessarily requiring an interpretation/understanding of the physical interactions of the molecules involved, as would be the case for a mechanistic definition of kinetic behavior [CB04, pg. 116]. For the feedback indicated on the right-hand side, and represented by $G(x_j)$, there is an additional contribution to the activation of X :



leading to the following, modified ODE model:

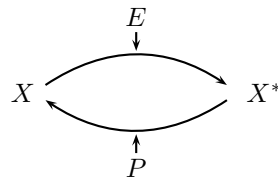
$$\begin{aligned} \frac{d}{dt}x(t) &= k_1 u(t)(\bar{x} - x(t)) + k_G G(x_j)(\bar{x} - x(t)) - k_2 x(t) \\ &= k_1(\bar{x} - x(t)) \left[u(t) + \frac{k_G G(x_j)}{k_1} \right] - k_2 x(t) . \end{aligned}$$

If $G(x_j)$ is monotonically increasing as x_j increases, we are dealing with positive feedback and, vice versa, if $G(x_j)$ monotonically decreasing with x_j , we are dealing with negative feedback. While for conventional pathway maps the kind of feedback employed remains unclear, if not explicitly stated. In the block diagram, however in our scheme we recognize the two situations as follows:



Michaelis-Menten Kinetics

The mass-action models for (de-)phosphorylation, as introduced above, could be criticized in that if the activation of a protein X is seen as an enzyme catalyzed reaction, the rate of activation in experiments is limited. Some authors might argue that a more realistic model would be to consider Michaelis-Menten kinetics. We here show how the Michaelis-Menten model can be derived from the mass-action model. In the context of signalling pathways an argument against Michaelis-Menten kinetics is the fact that in its derivation, the enzyme concentrations should be much smaller than the substrate concentration. In the context of cell signalling the relationships between kinase concentrations and those of inactivated proteins may however not satisfy this assumption. Let us consider the activation (phosphorylation) of protein X by means of an enzyme (kinase) E :



The dephosphorylation of the activated protein X^* is realized by some phosphatase P , which we here assume to have a constant concentration. The mass-action model of these two processes is



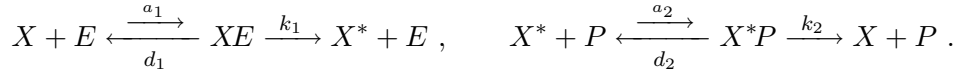
Using again the previously introduced notation, the corresponding differential equation for activation is

$$\frac{d}{dt}x = k_1 e(t)\tilde{x}(t) - k'_2 p(t)x(t) .$$

Assuming a constant phosphatase P , let us redefine $k_2 \doteq k'_2 p(t)$. With the conservation relation $\tilde{x}(t) + x(t) = \bar{x}$, we thus have

$$\frac{d}{dt}x = k_1 e(t)(\bar{x} - x(t)) - k_2 x(t) .$$

Note that only one differential equation is needed since changes in the non-phosphorylated form can always be derived from the conservation relation. A Michaelis-Menten model of a signalling module considers an intermediate complex in the enzyme catalyzed reaction



It is assumed that the first forward reaction is very fast, i.e., for phosphorylation $a_1 \gg d_1$. Focussing on steady-states, one finds for the Michaelis-Menten constants

$$K_{m1} = \frac{d_1 + k_1}{a_1} , \quad K_{m2} = \frac{d_2 + k_2}{a_2} .$$

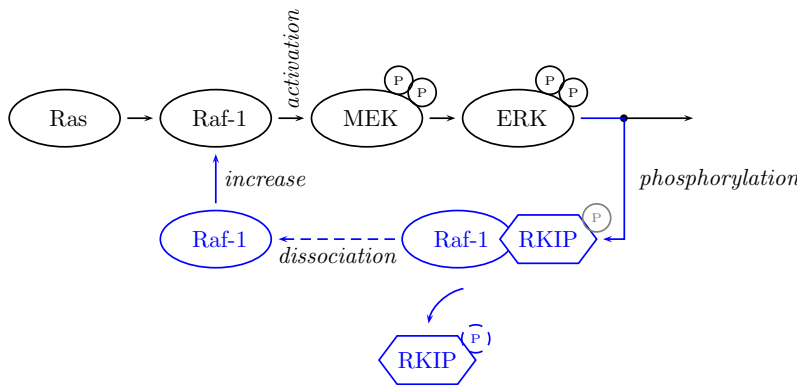
The equations for a Michaelis-Menten model are then

$$\frac{d}{dt}x = \frac{k_1 e(t)(\bar{x} - x(t))}{K_{m1} + \bar{x} - x(t)} - \frac{k_2 x}{K_{m2} + \bar{x} - x(t)} .$$

If E can be assumed to be constant, one would frequently define $V_{\max} \doteq k_1 E$ as the maximum velocity of the reaction, that is achieved for large concentrations of the non-activated form $x = \bar{x} - x$. Michaelis-Menten models of the MAP-kinase signaling cascade are considered in [HFJ96, Kho00, K+02].

The Ras/Raf/MEK/ERK Pathway

In this section we work out an example for the MAP-kinase pathway in Figure 8.6. In the Ras/Raf/MEK/ERK pathway Ras is the G-protein, Raf the MAPKKK, MEK the MAPKK and ERK the MAPK [Kol00]. While the linear cascade of Figure 8.6 is a textbook illustration, the research literature suggests the existence of various feedback loops such that we are dealing with a network rather than a linear cascade. As an example, for the Ras/Raf/MEK/ERK pathway a positive feedback mechanism can be illustrate with the following cartoon [YSL+99, YJM+00]:



The MAP kinase module is realized by the sequential activation of Raf-1, upstream near the cell membrane, followed by activation of the proteins MEK and ERK through structural modifications in the form of phosphorylations indicated by the P's. ERK translocates

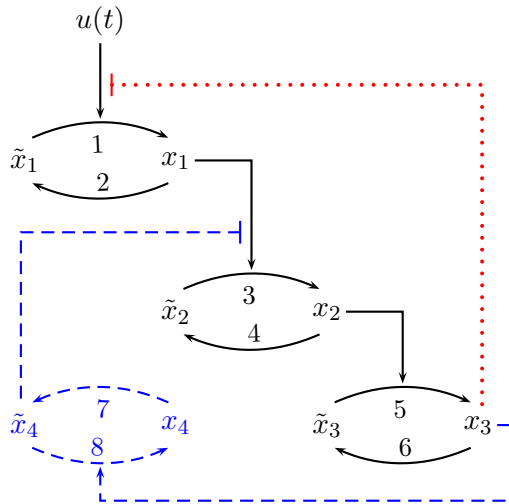


Fig. 8.13: Pathway map for the Ras/Raf/MEK/ERK signal transduction pathway with two speculated feedback loops. The dashed parts describe a positive feedback loop. The bar at the end of a line denotes an inhibition. Note that x_3 is acting on the phosphorylation of \tilde{x}_4 . The dotted line describes a negative feedback loop. The variables are $x_1 \doteq$ Raf, $x_2 \doteq$ MEK, $x_3 \doteq$ ERK, $x_4 \doteq$ RKIP. \tilde{x} denotes the non-activated or non-phosphorylated form.

into the nucleus of the cell, where it effects the transcription of genes. Double phosphorylated ERK-PP also phosphorylates RKIP and thereby releases Raf from the Raf-1/RKIP complex, and Raf in turn activates MEK. This positive feedback loop leads to switch-like behavior of the pathway.

We first translate the cartoon into a pathway map to reduce ambiguity. The pathway map is shown in Figure 8.13. The variables are $x_1 \doteq$ Raf, $x_2 \doteq$ MEK, $x_3 \doteq$ ERK, $x_4 \doteq$ RKIP. To simplify the example we ignored the double-phosphorylations of the previous Section. We first consider the pathway without any feedback loop. The mathematical model of this simple three module cascade is specified by the following set of equations.

$$\begin{aligned} \frac{d}{dt}x_1 &= \underbrace{\frac{k_1 u(t)(\bar{x}_1 - x_1(t))}{K_{m1} + (\bar{x}_1 - x_1(t))}}_{\text{phosphorylation}} - \underbrace{\frac{k_2 x_1(t)}{K_{m2} + x_1(t)}}_{\text{dephosphorylation}} \\ \frac{d}{dt}x_2 &= \frac{k_3 x_1(t)(\bar{x}_2 - x_2(t))}{K_{m3} + (\bar{x}_2 - x_2(t))} - \frac{k_4 x_2(t)}{K_{m4} + x_2(t)} \\ \frac{d}{dt}x_3 &= \frac{k_5 x_2(t)(\bar{x}_3 - x_3(t))}{K_{m5} + (\bar{x}_3 - x_3(t))} - \frac{k_6 x_3(t)}{K_{m6} + x_3(t)}, \end{aligned}$$

where for the conservation relations $\bar{x}_1 = \tilde{x}_1(t) + x_1(t)$, $\bar{x}_2 = \tilde{x}_2(t) + x_2(t)$, $\bar{x}_3 = \tilde{x}_3(t) + x_3(t)$ hold. Next we consider the positive feedback loop introduced by RKIP and which is denoted by x_4 . First phosphorylation and dephosphorylation are described as before,

$$\frac{d}{dt}x_4 = \frac{k_7 x_3(t)(\bar{x}_4 - x_4(t))}{K_{m7} + (\bar{x}_4 - x_4(t))} - \frac{k_8 x_4(t)}{K_{m8} + x_4(t)},$$

where $\bar{x}_4 = \tilde{x}_4(t) + x_4(t)$. Note that x_3 , activated ERK-PP, is acting on the phosphorylation of x_4 (RKIP). The inhibitory effect of RKIP on the phosphorylation of x_2 (MEK) is reflected by a change to the rate equation of x_2 :

$$\frac{d}{dt}x_2 = \frac{k_3 x_1(t) \left[1 / \left(1 + \left[\frac{\bar{x}_4 - x_4(t)}{K_P} \right]^p \right) \right] (\bar{x}_2 - x_2(t))}{K_{m3} + (\bar{x}_2 - x_2(t))} - \frac{k_4 x_2(t)}{K_{m4} + x_2(t)},$$

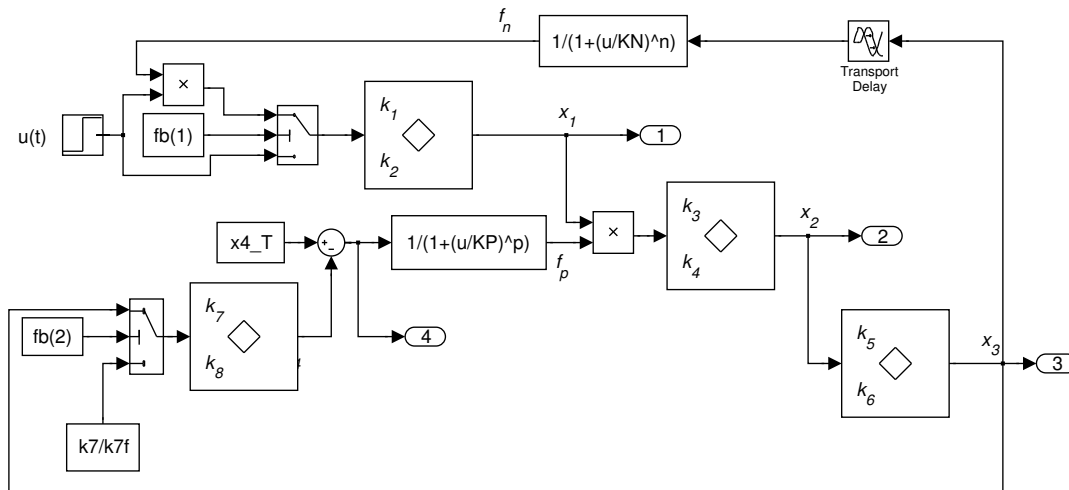


Fig. 8.14: Simulink block diagram of the Ras/Raf/MEK/ERK signal transduction pathway with RKIP regulation. Individual blocks can be “unmasked” to reveal their inside. The inside elements of these blocks are identical to those introduced before. The switches in the diagram are used to introduce or remove feedback loops in the simulation.

where K_P is a constant that defines the strength of the feedback and n defines the steepness of the response curve. The negative feedback from x_3 (ERK-PP) to x_1 (Raf) leads to an insertion in the equation for dx_1/dt :

$$\frac{d}{dt}x_1 = \frac{k_1 u(t) \left[1 / \left(1 + \left[\frac{x_3(t)}{K_N} \right]^n \right) \right] (\bar{x}_1 - x_1(t))}{K_{m1} + (\bar{x}_1 - x_1(t))} - \frac{k_2 x_1(t)}{K_{m2} + x_1(t)}$$

For all proteins involved, conservation relationships hold for a constant total of the activated and non-phosphorylated form.

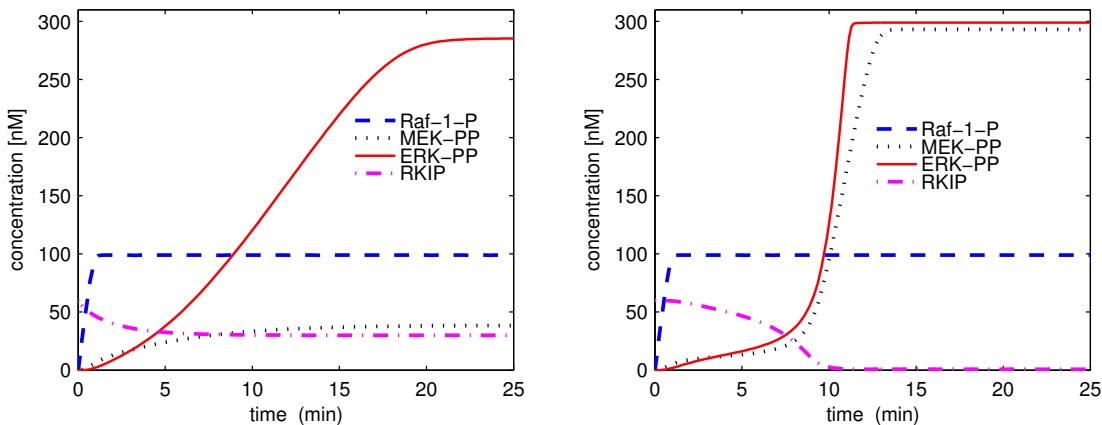


Fig. 8.15: Simulation of the Ras/Raf/MEK/ERK pathway in Figure 8.13. The plots show the response to a unit-step input signal. Left: without feedback loops. Right: as before but with a positive feedback loop that is realized by RKIP. Note the sharpening of the ERK response from the introduction of positive feedback loop.

Figure 8.14 shows a Simulink⁵ block diagram of the Ras/Raf/MEK/ERK signal transduction pathway with RKIP regulation. Simulink is a graphical simulation environment, using the mathematical programming language Matlab. Block diagrams such as those

⁵Matlab and Simulink files for all models and simulations shown here are available from www.sbi.uni-rostock.de.

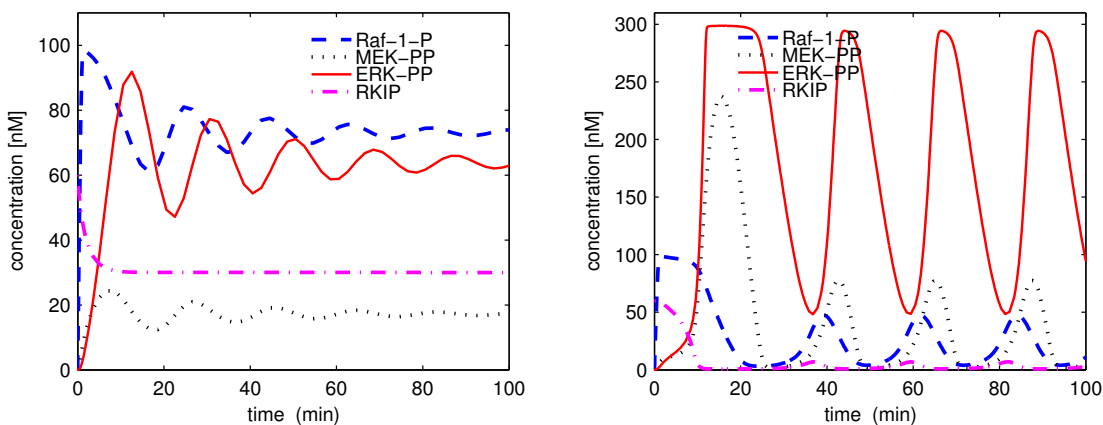


Fig. 8.16: Simulation of the Ras/Raf/MEK/ERK pathway in Figure 8.13. Left: with negative feedback loop. Right: with both, positive and negative feedback loops.

introduced in previous sections can be drawn and simulated directly from within an interactive graphical editor. The effect of changes to parameters, the removal or introduction of feedback loops can be simulated conveniently. As illustrated Figure 8.15, with only positive feedback added to the pathway and no transport delay, the pathway displays switch-like behavior. Switching dynamics have been found in various intracellular systems (e.g. [FJ96, TCN03]). Our previous observation that a positive feedback loop sharpens the response, making it *ultrasensitive*, remains true in this more complex system. Because the positive feedback loop effects only proteins from MEK downward, the Raf concentration profile has not been changed. Considering a negative feedback loop, no transport delay and without the positive feedback loop in the system, we observe that negative feedback can destabilize the response. What can also be observed are lower steady-state values for Raf and ERK.

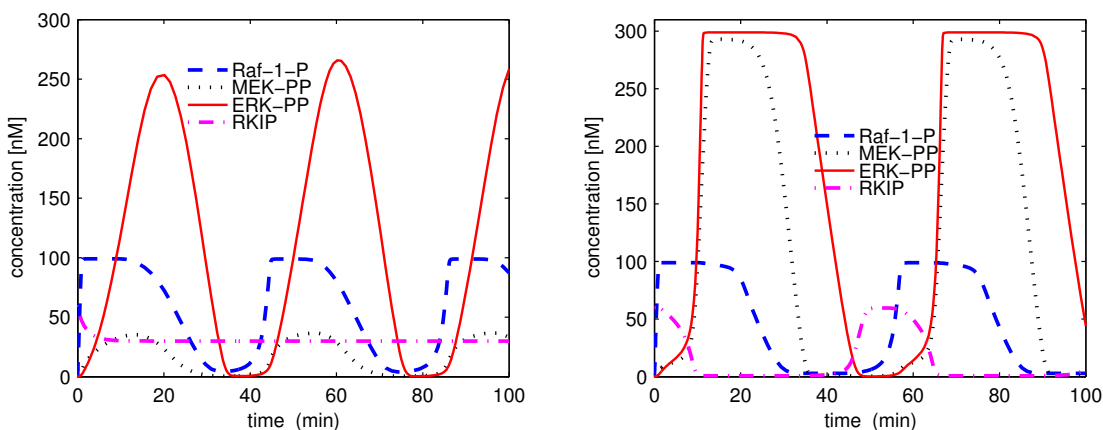


Fig. 8.17: Simulation of the Ras/Raf/MEK/ERK pathway in Figure 8.13. The plots show the response to a unit-step input signal. Left: Negative feedback loop with transport delay $T_d = 10\text{min}$. Right: As before but with additional positive feedback loop. In both cases $n = p = 1$.

Once a model is established, simulation allows quick studies of changes to the elements and parameters. For example, one way to make the model more realistic is to introduce a time delay between ERK near or inside the nucleus and its feedback effect on Raf further up the pathway. In Figure 8.17, we introduce a transport delay in the negative feedback loop with $T_d = 10\text{min}$. We observed that transport delays lead to increased oscillatory behavior,

turns the damped oscillations into sustained oscillations. In [SMT+03] this consideration for nucleocytoplasmic transport was crucial in obtaining a predictive mathematical model for the JAK-STAT pathway. Our next experiment is to change the feedback indices n and p , that were also introduced above and which define the sharpness or sensitivity of the feedback effect. In Figure 8.18 we find that without transport delay an increase from $n = 1$ to $n = 2$ in the negative feedback loop also leads to sustained oscillations.

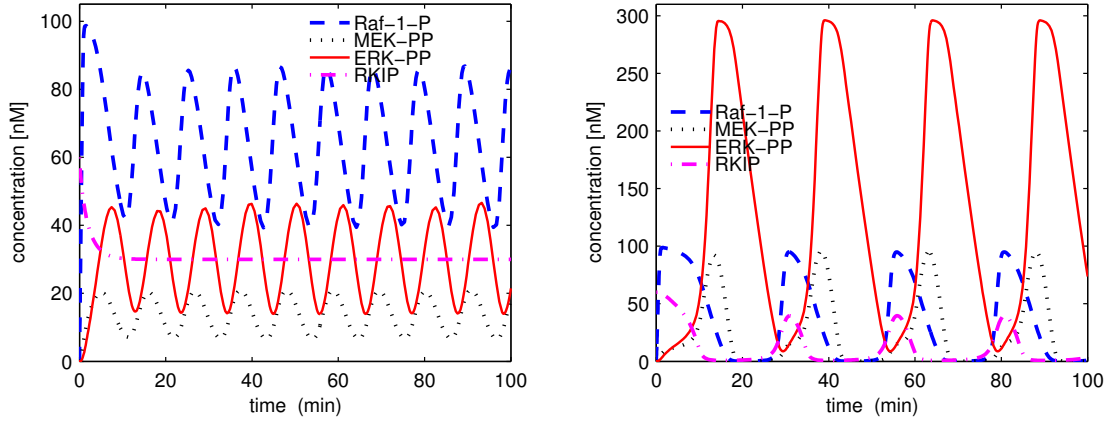


Fig. 8.18: Simulation of the Ras/Raf/MEK/ERK pathway in Figure 8.13. The plots show the response to a unit-step input signal. Left: Negative feedback loop with $T_d = 0$, $n = 2$, $p = 1$. Right: As before but with additional positive feedback loop.

Our study demonstrates various sources of sustained oscillations: negative feedback combined with ultrasensitivity, combined negative and positive feedback and transport delays in negative feedback loops. Oscillations have been investigated in various systems (e.g. [Gol96, WH00, Kho00]) and have been of interest in mathematical modelling for some time (e.g. [GH83, Str00a]). An interesting question is to ask whether our model applies to a single cell or a population of cells. If a single cell is an oscillator, one would have to consider a population of coupled oscillators. Oscillations are not easy to detect and require complex experimental set-ups. Simulation studies demonstrate the usefulness of mathematical model in generating hypotheses of phenomena that have yet to be observed in experiments. On the other hand, a simulation can also be used to support the design of experiments, helping to decide which variables to measure and why. The role of feedback in intracellular dynamics has been investigated for some time in the literature (e.g. [Gri68, Tys79, BPM82, Tys83, TD90]) and will, no doubt, play an important role in (re)emerging area of systems biology.

$k_1 = 2.5$	$K_{m1} = 10$	$k_2 = 0.25$	$K_{m2} = 8$	$[\text{Raf-1}]_T = 100$
$k_3 = 0.025$	$K_{m3} = 15$	$k_4 = 0.75$	$K_{m4} = 15$	$[\text{MEK}]_T = 300$
$k_5 = 0.025$	$K_{m5} = 15$	$k_6 = 0.5$	$K_{m6} = 15$	$[\text{ERK}]_T = 300$
$k_7 = 0.5$	$K_{m7} = 15$	$k_8 = 0.5$	$K_{m8} = 15$	$[\text{RKIP}]_T = 60$
$k_{7_f} = 0.025$	$K_N = 9$	$K_P = 9$	n, p, T_d	

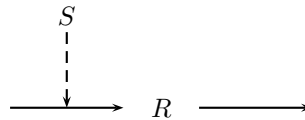
Table 8.1: Parameter values for the Ras/Raf/MEK/ERK pathway model. Concentrations are in nM; k_1, k_3, k_5, k_{7_f} are in s^{-1} ; k_2, k_4, k_6, k_7, k_8 in $\text{nM}\cdot\text{s}^{-1}$; K_{m1} to K_{m8} in nM. The right column specifies total concentrations. Note that the purpose of this model is to illustrate the role of feedback loops on protein concentration profiles only.

9 Modules and Control Mechanisms

In the present chapter we are going to investigate a class of *modules* with particular dynamics behaviors, such as switches, buzzers, oscillators etc. Rather than investigating dynamic properties in the time domain and with the help of phase-plane analysis, the present chapter introduces *rate curves* and *stimulus-response curves* for steady-states as a tool for the analysis of dynamic modules and control mechanisms. The presentation here is an extended version of the paper by Tyson et al. [TCN03]. The graphical representation of the modules follows the description in [TCN03].

9.1 Linear Module

To start with, we consider a linear system in which the synthesis of the response molecule R is facilitated by a stimulus S ; illustrated by the following diagram:



The mathematical model, the rate equation is given as

$$\frac{dR}{dt} = k_0 + k_1S - k_2R . \quad (9.1)$$

The term k_0 describes a constant base level synthesis of the *response component* R . What we describe here as a response could also be referred to as the *production* of R . Assuming monomolecular reactions for conversion and degradation, the second term k_1S is the conversion of the *stimulus component* S into the response component. The degradation of R is given by the last term $-k_2R$. The *rate of response* is then the sum of the base level flux and the conversion of S :

$$\text{rate of response} = k_0 + k_1S .$$

This allows us to write for the change of the response component

$$\frac{dR}{dt} = \text{rate of response} - \text{rate of degradation} .$$

rate curve We can then compare the degradation rate and the response rate by plotting the *rate curve* as function of R . This is shown in Figure 9.1. Note that hereafter we are going to suppress units in plots for easier viewing. Depending on the number of participating components the rate coefficient k_i has the unit ‘per second’ divided by the unit of the component to the power $(n - 1)$. Michaelis-Menten constants have the same units as the corresponding component. The solid line represents the degradation or removal of R , which is k_2R . The rate is a straight line with a slope equal to k_2 . The rate of response $k_0 + k_1S$ (dashed lines) is independent of R and thus a horizontal line. The points, where the rate of response and the rate of degradation are equal are of particular interest:

$$\text{rate of response} = \text{rate of degradation} . \quad (9.2)$$

At these points the system is in a steady-state such that no net change of concentrations is measurable. This means response and degradation are in balance, although this does not mean that no reaction occurs. This state is mathematically determined by the balance equation

$$\frac{dR}{dt} = \text{rate of response} - \text{rate of degradation} = 0 . \quad (9.3)$$

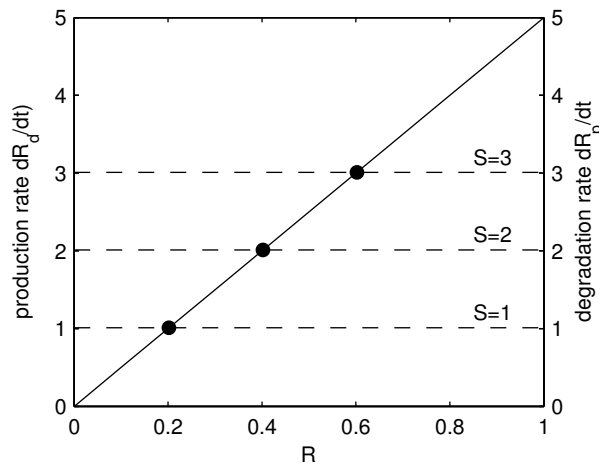


Fig. 9.1: Rate curve, comparing the response rate (dashed lines) and the degradation rate (solid line) for the linear system (9.1). The intersections of both rates (black dots) are the steady states for the given stimulus. The parameters for the system are: $k_0 = 0.01$, $k_1 = 1$, $k_2 = 5$.

Note that a system fulfilling condition (9.2) is either in equilibrium or in a steady-state. There is no difference between both states from the mathematical point of view. If the system is in one of these states, all sub-systems have to fulfil (9.2) too. The difference between both states lies in the considered type of system. An equilibrium is defined for a *closed system*, where there is no transport of matter and energy in and out of the system. A closed system will relax to the equilibrium state and will not leave it by itself. Following a small perturbation the system returns to the equilibrium. In an *open system*, such as a cell, the transport of energy is possible. If we further consider the flow of matter, we get a *fully open system*.

closed/open system

Here we want to describe the response of a system to an external stimulus, without assumptions about flow of energy or matter, and therefore assume a fully open system, that are characteristic of living systems. In such a system we usually reach a steady-state dependent from the external stimulus. The system state is held by the signal and can be far away from the equilibrium state. Only, if we close the system, for instance we choose $S = 0$, the system relaxes into its equilibrium state. In this sense, the equilibrium is a particular steady-state but if there is no flow of molecules in an out of the cell this could mean that they die. As we will see later for the sigmoidal module, the conditions (9.2) and (9.3) are not fully equivalent. Strictly speaking, if condition (9.2) is fulfilled, the system is in a steady-state. It can but must not be in a steady-state if (9.3) is fulfilled.

Note that the dR_r/dt of the ordinate in Figure 9.1 is not the *net-rate* dR/dt on the left-hand side of (9.1). Figure 9.2 shows the net-rate as a function of R . Is the net rate equal to zero, shown as horizontal dashed line in the plot, the production and degradation rates are balanced. The system is in a steady-state.

Next we investigate the steady-state as a function of the stimulus. According to (9.1) and (9.3) this state is obtained as a solution to the equation

$$0 = k_0 + k_1 S - k_2 R .$$

The steady-state value of the response component is an increasing straight line with slope k_1/k_2

$$R_{ss} = \frac{k_0}{k_2} + \frac{k_1}{k_2} S . \tag{9.4}$$

If the stimulus is increasing than the response is increasing too. For a finite constant external source of R the response is $R_{ss} = k_0/k_2$ at a signal strength $S = 0$. The repre-

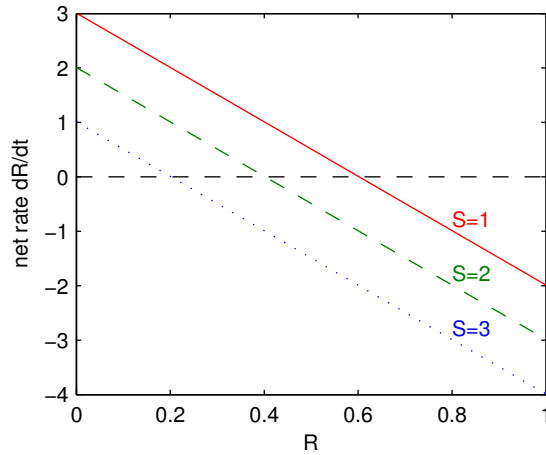


Fig. 9.2: The net rate for a linear system as function of the response component R for three values of the signal S . If the net rate is greater than zero, R increases and if the net rate is smaller than zero, R decreases. The value $dR/dt = 0$ (dashed horizontal line) corresponds to the steady-state. Parameters: $k_0 = 0.01$, $k_1 = 1$, $k_2 = 5$.

presentation of the steady-state R_{ss} as function of the signal S is shown in Figure 9.3. This kind of plot is called *stimulus-response curve*. Figure 9.3 shows the response of the linear module for different ratios of k_1/k_2 .

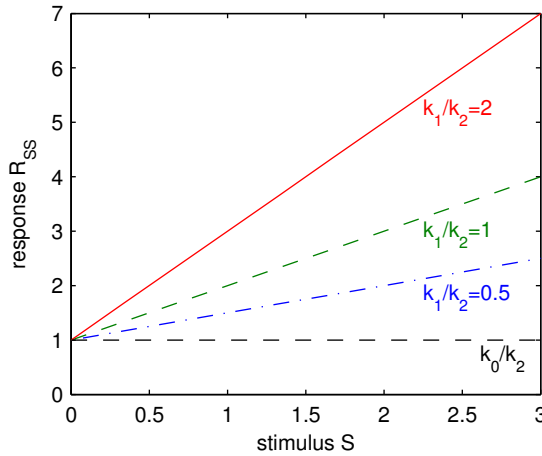


Fig. 9.3: Stimulus-response curve for a linear module. Three ratios of k_1/k_2 are compared, whereby the decay rate k_2 is constant. The response R_{ss} is the stronger the stronger the signal S is. The horizontal line is the limit for a vanishing stimulus. Parameters: $k_0 = 1$, $k_2 = 1$.

To this point, we have investigated static properties of the system. Investigating the temporal evolution of R , we now assume that the signal S does not change over time. This may also be interpreted as the response to an abrupt change of the stimulus, at $t_0 = 0$, to the given value, which is then constant (“step-response”). With this assumption, (9.1) is a linear first-order differential equation for which the solution can be found analytically:

$$R(t) = \frac{k_0 + k_1 S}{k_2} (1 - \exp\{-k_2 t\}) + R_0 \exp\{-k_2 t\} , \quad (9.5)$$

where R_0 is the initial value of R at t_0 . The first term on the right-hand-side describes the relaxation to the stationary state R_{ss} (9.4) and the second term is the degradation of the initial state R_0 . The relaxation time

$$\tau = \frac{1}{k_2}$$

is independent from external stimulus S . This quantity is an approximation to the time needed to relax into the steady state. It can also be interpreted as the time the system is able to react to perturbations. For $t \ll 1/k_2$ we can expand the exponential function

$$e^{-x} \approx 1 - x + o(x^2) ,$$

which for a short initial period of time leads to

$$R(t) \approx (k_0 + k_1 S - k_2 R_0) t + R_0 . \tag{9.6}$$

The response $R(t)$ as function of the dimensionless time $\tau = k_2 \cdot t$ is shown in Figure 9.4. We chose a fixed value for the decay rate coefficient k_2 and vary the rate coefficient k_1 . According to (9.5) the response component $R(t)$ exponentially relaxes to the steady-states (9.4). The solid straight lines are the asymptotic solutions for initial times (9.6). The horizontal dashed lines correspond to the steady state. Analogue to Figure 9.3 the steady-state depends on the ratio k_1/k_2 . If the production dominates, $k_1 > k_2$, the steady-state value is greater than the signal strength¹ S .

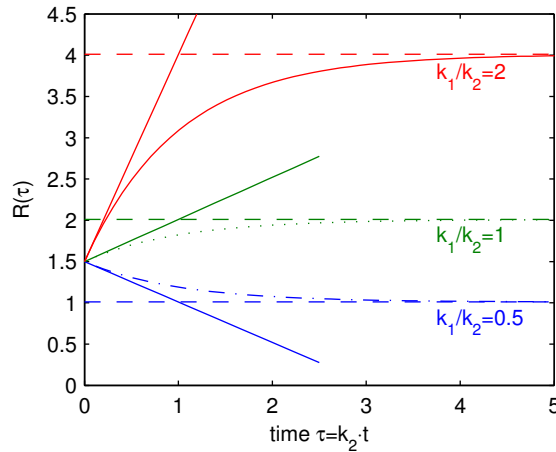
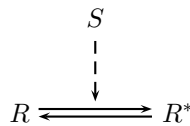


Fig. 9.4: The response of a linear module for a step change to $S = 2$ as function of the dimensionless time $\tau = k_2 \cdot t$. Three ratios of k_1/k_2 are shown. Parameters: $k_0 = 0.01$, $k_2 = 1$, $R_0 = 1.5$.

9.2 Hyperbolic Module

A hyperbolic module is slightly more complicated than the linear module discussed above. The response or activation is now understood as a bimolecular reaction while degradation is still considered monomolecular.



This formulation is a popular model for activation or phosphorylation in signalling pathways. In this situation we denote by R^* the activated or phosphorylated form and use R to denote the non-activated form. Such a reaction system we describe as

$$\frac{dR^*}{dt} = k_1 S R - k_2 R^* ,$$

¹The first term k_0/k_2 in (9.4) is negligible for $k_0 = 0.01$ and $S = 2$.

where in the case of a signalling pathway, stimulus S corresponds to a kinase that facilitates the phosphorylation of R . If we assume a constant total concentration,

$$R_T = R^* + R, \quad (9.7)$$

we can rewrite the differential equation of the hyperbolic module:

$$\frac{dR^*}{dt} = k_1 S (R_T - R^*) - k_2 R^*, \quad (9.8)$$

where the first term on the right-hand-side is the rate of response or activation and the second the rate of deactivation. The corresponding rate curve is shown in Figure 9.5. The rate of deactivation (solid line) has a slope of k_2 . According to the conservation law (9.7) the response component is restricted to the range $[0, R_T]$. If all R molecules are phosphorylated the response must be zero. The net rate of (9.8) is a straight line with a slope $-(k_1 S + k_2)$. The conservation law restricts the response to $[0, 1]$. If the net rate is greater than zero, activation dominates. The intersections with the zero line are the steady-states values of R^* for this system. From (9.8) we obtain for the steady-states a hyperbolic function, which gives this module its name:

$$R_{ss}^* = \frac{S R_T}{k_2/k_1 + S}. \quad (9.9)$$

For the limits $S \rightarrow 0$ and $S \rightarrow \infty$ we can expand (9.9). For a signal strength $S \ll k_2/k_1$ the contribution of S in the denominator is negligible and the steady-states govern the linear function

$$R_{ss}^* \approx \frac{k_1}{k_2} R_T S. \quad (9.10)$$

The ratio of the rate coefficients k_1/k_2 determines the slope of the asymptote. If for the signal strength $S \gg k_1/k_2$, the signal S dominates the denominator. In this case all proteins are phosphorylated and $R_{ss}^* \approx R_T$. The stimulus-response curve for a hyperbolic module is shown in Figure 9.6. The straight lines are the asymptotic expansion (9.10). If the rate coefficient of activation k_1 greater than the rate coefficient of deactivation k_2 , the hyperbolic properties of the system is in evidence. For small values of this ratio the hyperbolic system looks like a linear system within the presented signal range.

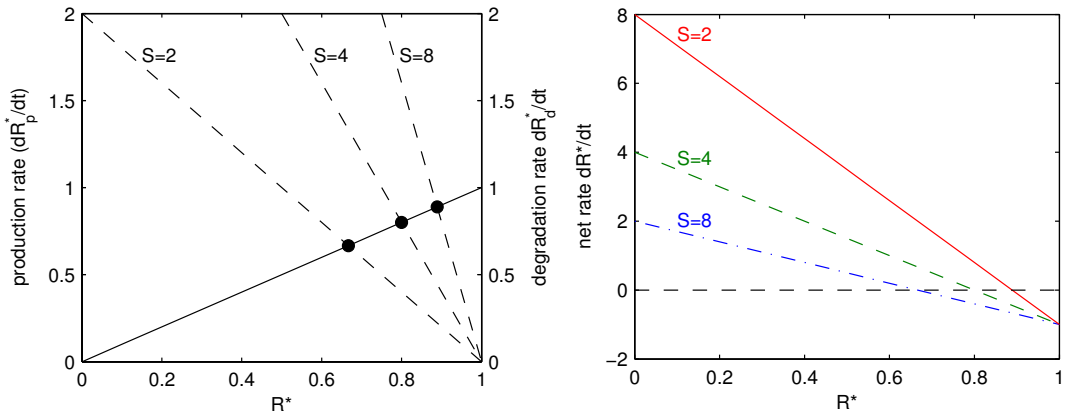


Fig. 9.5: Left: The rate curve for the hyperbolic module, assuming bimolecular activation and monomolecular deactivation, for different signal strengths. The deactivation rate is the solid line and the rate of response for three values of S are shown by dashed lines. The intersections marked by the black dots are the steady states for the shown parameters $k_1 = 1$, $k_2 = 1$, $R_T = 1$. Right: Net rate as a function of R^* for three different stimuli.

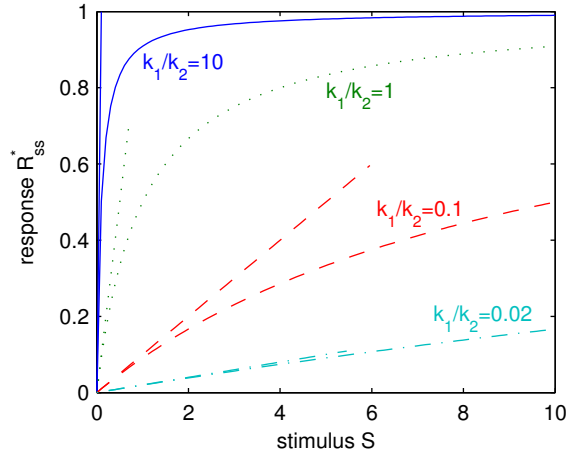


Fig. 9.6: The stimulus-response curve of a hyperbolic module (9.8) for different ratios of k_1/k_2 . The straight lines are the asymptotes for small signal strength (9.10). Parameters: $k_1 = 1$, $R_T = 1$.

The temporal evolution of the hyperbolic module is described by the differential equation (9.8). This is again a first-order linear differential equation that we can solve analytically:

$$R^*(t) = \frac{R_T S}{k_2/k_1 + S} \left(1 - \exp \{ -(k_2 + k_1 S)t \} \right) + R_0 \exp \{ -(k_2 + k_1 S)t \} . \quad (9.11)$$

The first term describes the relaxation to the steady-state while the second term corresponds to the degradation from the initial state. In contrast to the linear system the relaxation time

$$\tau = \frac{1}{k_2 + k_1 S}$$

is now dependent on the signal strength. The greater the signal strength, the faster the steady-state is reached. For times $t \ll \tau$ we can expand the exponential function and obtain for small initial times the asymptote

$$R^*(t) \approx \left[k_1(R_T - R^*)S - k_2 R_0 \right] t + R_0 . \quad (9.12)$$

The temporal evolution of the hyperbolic module is shown in Figure 9.7. We compare solutions for three different ratios of the rate coefficients. For comparison, the small times asymptotes (9.12) are also drawn.

9.3 Sigmoidal Module

The next module is an extension of the hyperbolic module. Both activation and deactivation are here considered as Michaelis-Menten type kinetic reactions, leading to the differential equation

$$\begin{aligned} \frac{dR^*}{dt} &= \frac{k_1 S R}{K_{M1} + R} - \frac{k_2 R^*}{K_{M2} + R^*} \\ &= \frac{k_1 S (R_T - R^*)}{K_{M1} + R_T - R^*} - \frac{k_2 R^*}{K_{M2} + R^*} , \end{aligned} \quad (9.13)$$

where on the second line, the conservation law (9.7) is used to relate the unphosphorylated form R of the response component to the phosphorylated form R^* . The parameter K_{M1}, K_{M2} are Michaelis-Menten constants. The first term on the right-hand-side describes

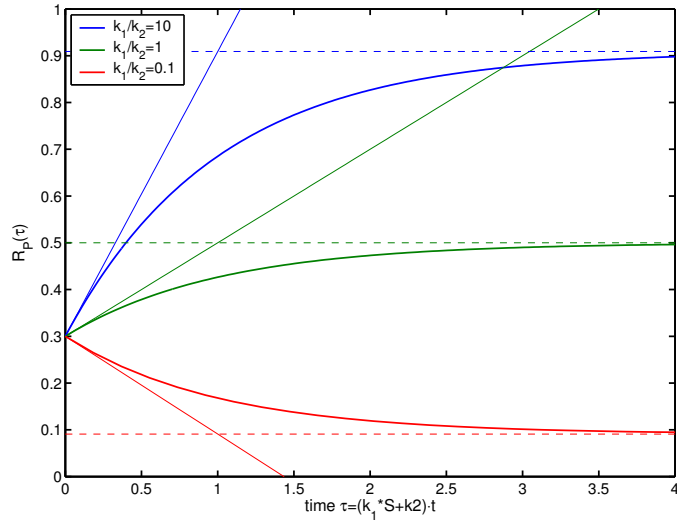


Fig. 9.7: Temporal evolution of the hyperbolic module (9.8). The full solution of the differential equation (9.11) is drawn for three different ratios of k_1/k_2 . Parameters: $k_1 = 1$, $S = 1$, $R_T = 1$, $R_0 = 0.3$.

the activation, while the second term corresponds to the deactivation or dephosphorylation. A comparison of response and deactivation rate is shown by the rate curve in Figure 9.8. The conservation law (9.7) limits the possible values of R^* to the interval $[0, R_T]$. If R^* reaches the value R_T , the activation rate has to be zero, independent of the signal strength, since there is no more unphosphorylated R available.

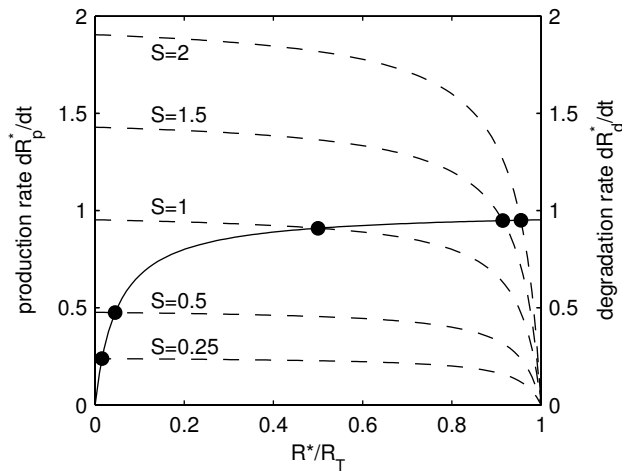


Fig. 9.8: The activation and deactivation rate of the sigmoidal module (9.13). The solid line is the deactivation rate. The dashed lines are the rate of activation for different signal strength. The black dots mark the points where both rates are in balance. Parameters: $k_1 = 1$, $k_2 = 1$, $K_{M1} = 0.05$, $K_{M2} = 0.05$, $R_T = 1$.

The steady-state of the sigmoidal module (9.13) is determined by the quadratic equation

$$0 = \frac{k_1 S (R_T - R^*)}{K_{M1} + R_T - R^*} - \frac{k_2}{K_{M2} + R^*}. \quad (9.14)$$

For $0 < R^* < R_T$, the solution is given by what is known as the *Goldbeter-Koshland*

function:

$$\frac{R_{ss}^*}{R_T} = \frac{k_2 \left(1 + \frac{K_{M1}}{R_T}\right) + k_1 S \left(\frac{K_{M2}}{R_T}\right)}{2(k_2 - k_1 S)} + \frac{\sqrt{\left[k_2 \left(1 + \frac{K_{M1}}{R_T}\right) + k_1 S \left(\frac{K_{M2}}{R_T}\right)\right]^2 + 4 \frac{K_{M2}}{R_T} (k_2 - k_1 S) k_1 S}}{2(k_2 - k_1 S)}. \quad (9.15)$$

Since this expression is somewhat complicated we replace it by the short form

$$\frac{R_{ss}^*}{R_T} = G(k_1 S, k_2, K_{M1}/R_T, K_{M2}/R_T). \quad (9.16)$$

The solutions are shown in Figure 9.9. Additionally, the allowed range of R^* is shown as a grey-shaded region. Only one solution of (9.14) lies inside this region. The others do not fulfil the physical restrictions on the response component and are therefore irrelevant for a biological system. This system is our first example, where a mathematical solution of (9.3) is not a steady-state. This demonstrates that the conditions (9.2) and (9.3) are not fully equivalent. Strictly speaking, if condition (9.2) is fulfilled, the system is in a steady-state. It can but must not be in a steady-state if (9.3) is fulfilled. In case of the sigmoidal module we cannot derive the solution $R^*(t)$ to differential equation (9.13) in an analytical form.

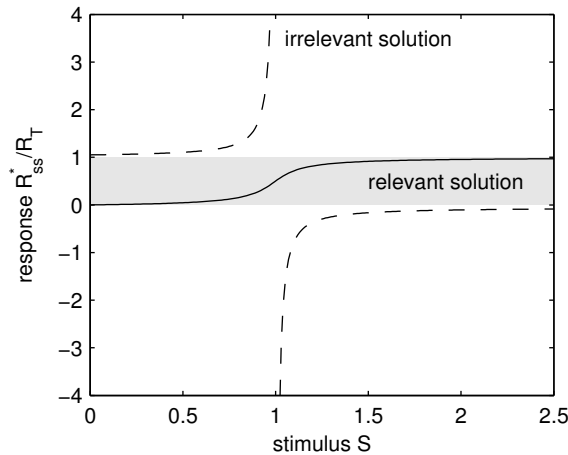


Fig. 9.9: Solutions of (9.14) as a function of the signal strength. Only the relevant solution ($+\sqrt{\dots}$) fulfils the conservation law (9.7). The irrelevant solution ($-\sqrt{\dots}$) lies in physically unreachable state-space. The grey shaded region shows the reachable states that follow from the conservation law.

The stimulus-response curve of the sigmoidal module is shown in Figure 9.10. The sigmoidal shape of the curve is determined by the ratio of the Michaelis-Menten constants. If the Michaelis constant of for the activation term is much smaller than the constant for the deactivation term, the typical sigmoidal shape vanishes. On the other hand, if the activation constant K_{M1} is much greater than K_{M2} , we get a switch-like behavior for the response function. It is for this reason that the system (9.13) is sometimes called Goldbeter-Koshland switch [GKJ81].

9.4 Robust or Adaptive Module

In the next dynamic module we consider, $R(t)$ is robust to changes in the stimulus. Looked at from the perspective of $X(t)$ the system is adaptive in that it tracks the stimulus. While

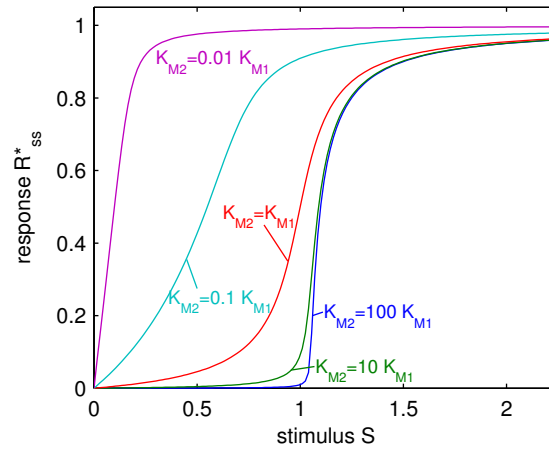
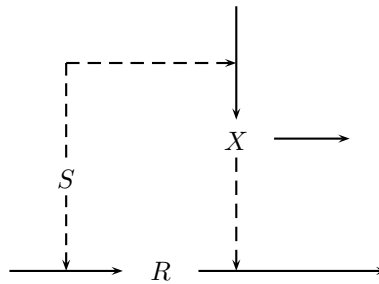


Fig. 9.10: The stimulus-response curve for the sigmoidal system (9.13) for different ratios of the Michaelis-Menten constants K_{M1} and K_{M2} . Parameters: $k_1 = 1$, $k_2 = 1$, $K_{M1} = 0.05$, $R_T = 1$.

there is a transient response of $R(t)$ to step-changes in $S(t)$, the response returns eventually to its steady state value R_{ss} . Such a behavior can be realized by two parallel reaction paths pushing the response in opposite directions:



For the simplest case we combine two linear systems

$$\begin{aligned} \frac{dR}{dt} &= k_1 S - k_2 X R, \\ \frac{dX}{dt} &= k_3 S - k_4 X, \end{aligned} \quad (9.17)$$

where both systems are coupled through the parallel stimulus by S . The degradation or deactivation of the response component R depends on the amount of X . Because an increasing signal will also increase the amount of X , this simultaneously increases the degradation rate of the response component. The rate curves for this module are shown in Figure 9.11. For the response component R we have:

$$\begin{aligned} \text{rate of response} &= k_1 S, \\ \text{rate of degradation} &= k_2 X R. \end{aligned}$$

The response rate is a linear function of the stimulus S , with slope k_1 . The degradation rate is independent of S but dependent on X . For this reason we vary the signal strength for the rate of production and the amount X in Figure 9.11. The intersections of corresponding curves are again the steady states of the system. In the representation of the net rate (Figure 9.12) this is more visible. According to the underlying linear system, we obtain again linear functions. The slope is given by $k_2 X$ and the ordinates by $k_1 S$.

For the steady-state R_{ss} we have to solve the system of balance equations

$$\begin{aligned} 0 &= k_1 S - k_2 X R, \\ 0 &= k_3 S - k_4 X, \end{aligned}$$

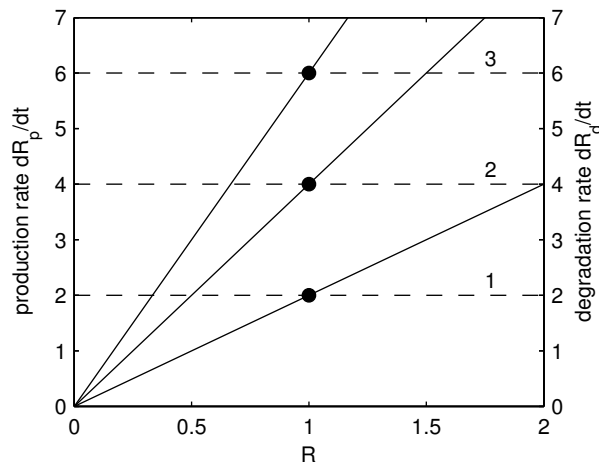


Fig. 9.11: Rate curve of a perfectly-adapted or robust system (9.17) for different external stimuli S and amounts of X . The rate of response (dashed lines) is proportional to the external signal and the rate coefficient k_1 , but constant with respect to R . The solid lines are the rates of degradation for three different X . These linear functions have the slope k_3X . The steady states, marked by dots, have a constant R -value. Parameters: $k_1 = k_2 = 2$, $k_3 = k_4 = 1$.

leading to a constant steady state response

$$R_{ss} = \frac{k_1 k_4}{k_2 k_3}.$$

The steady state response is therefore determined by the ratios of the rate coefficients for both channels. In Figure 9.13 we plot the temporal evolution of stimulus S , regulating component X , and response component R . We see that the response is only transient and returns eventually to its steady state. The height of the transient peak depends on X . The larger X , the smaller the peak. The duration for which S is kept constant is denoted by Δt and the height of the step-change by ΔS . The return of $R(t)$ to its steady state is due to the monomolecular decay that after some time is faster than the bimolecular degradation of R . Looking at R from S , the system may be considered *robust* to changes in S . Looked at from the perspective of X , the system shows perfect *adaptation* the changes.

9.5 Feedback Systems

In previous sections we considered systems without feedback. In the context of this chapter, feedback means, that the response component influences its own production or degradation. We refer to the feedback as positive or negative, depending on whether the response amplified or suppressed [MVHA99]. Positive feedback can be realized through:

1. Acceleration of the production, for instance in an autocatalytic process,
2. Inhibition of the degradation.

Feedback is negative, if it weakens the response signal through:

1. Inhibition of the production,
2. Acceleration of the degradation.

In subsequent sections we discuss different feedback mechanisms, starting with positive feedback, negative feedback followed by mixed negative and positive feedback.

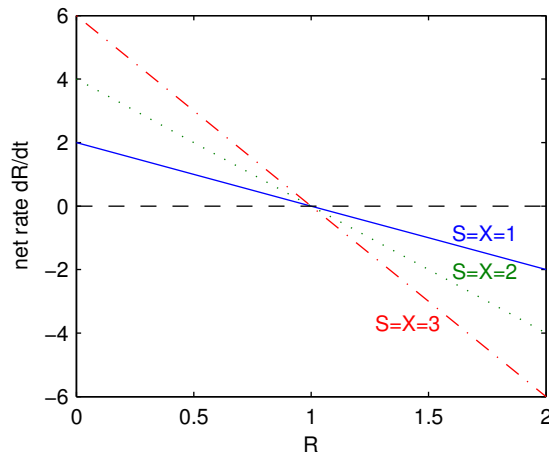


Fig. 9.12: Net rate for system (9.17) for various external signals S and amounts of X . According to the underlying linear system, the net rate is a linear function, too. The common intersection of all three curves approves the behavior of perfect adaptation. Parameters: $k_1 = k_2 = 2$, $k_3 = k_4 = 1$.

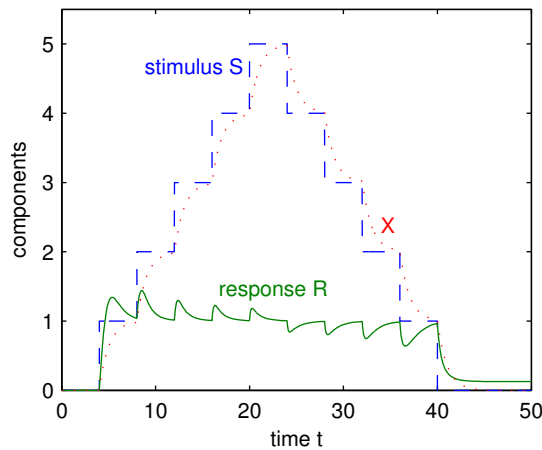


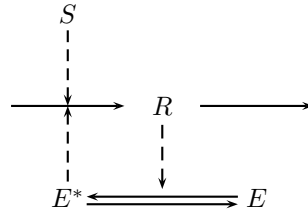
Fig. 9.13: Temporal evolution of robust-adaptive module (9.17) under the influence of a stepwise change in stimulus. Parameters: $k_1 = k_2 = 2$, $k_3 = k_4 = 1$, $\Delta t = 4$, $\Delta S = 1$, $S_0 = X_0 = R_0 = 0$.

9.5.1 Positive Feedback/Feedforward - Switches

The present section is to discuss the two ways by which positive feedback/feedforward control mechanisms can be realized. The acceleration of production of the response component R is related to *mutual activation*, while the inhibition of degradation is related to *mutual inhibition*. In both cases, the positive feedback/feedforward can create a switch-like behavior, in which a certain level of stimulus can lead to a sudden change of the response.

Mutual activation - positive feedback

We add to the linear system (9.1) a backward directed loop, i.e., feedback is realized through an intermediate enzyme E . The response component activates the enzyme E , for instance through phosphorylation. The activated enzyme E^* on the other hand enhances the synthesis of R . In this sense, both components S and R mutually activate R :



To study the properties of this system we assume that the activation of the enzyme E can be described by means of the Goldbeter-Koshland function (9.16) such that the differential equation model of the mutually activated system is

$$\frac{dR}{dt} = k_0 E^*(R) + k_1 S - k_2 R \tag{9.18}$$

where

$$E^*(R) = G(k_3 R, k_4, J_3, J_4) .$$

The rate of degradation is a linear function of the response component,

$$\text{rate of degradation} = k_2 R ,$$

and depends only on R . The proportionality coefficient is again the rate coefficient k_2 . In contrast to systems discussed in previous sections, the rate of production,

$$\text{rate of production} = k_0 E^*(R) + k_1 S ,$$

is now a function of the response component itself as well as signal S . The comparison of both rates, shown in Figure 9.14, illustrates new properties. Dependent on the signal strength S , the number of intersections between the curves varies between one and three. Let us, for the time being, consider these as steady states, although we have to refine the meaning of steady-states.

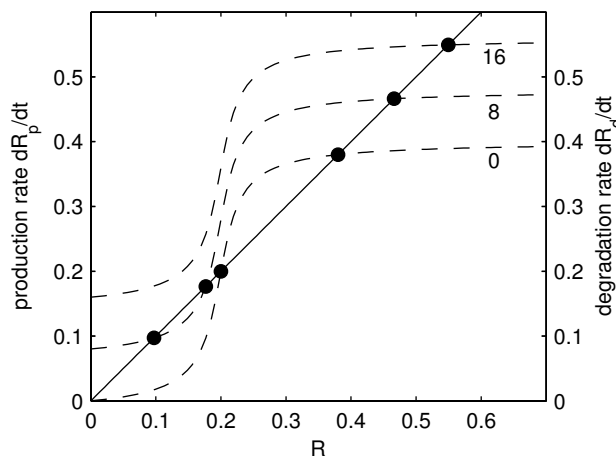


Fig. 9.14: Comparison of the rate of degradation (solid line) and the rate of production for different signal strengths (dashed lines), for a mutually activated system (9.18). Parameters: $k_0 = 0.4, k_1 = 0.01, k_2 = 1, k_3 = 1, k_4 = 0.2, J_3 = 0.05, J_4 = 0.05$.

Examining the net-rate of the mutually activated system, with the same parameters as in Figure 9.14, the resulting curves are shown in Figure 9.15. There is a signal strength, for which there are cases with a net rate equal to zero. We also notice that there is a region for which the net-rate increases with R . For this region the lines are dashed segments. Mathematically this corresponds to the condition

$$\frac{d}{dR} \frac{dR}{dt} > 0 ,$$

where the derivations are carried out in the given order from left to right. For the considered system (9.18) we have

$$\frac{d}{dR} \frac{dR}{dt} = k_0 \frac{dE^*}{dR} - k_2 . \quad (9.19)$$

This is an expression that is independent of the signal strength. The derivative of dE^*/dR can be derived analytically, although this is cumbersome and therefore avoided here. To see what happens, if a zero net-rate is located within this range, let us assume, that we are in steady state. If we perturb the system a little by increasing the response component, the rate of degradation will decrease. The system cannot return to his former state. On the other hand, if we decrease the response component the rate of production also decrease and the system cannot return. This state is unstable and hence not a steady or equilibrium state. The necessary condition for a stable (bio)chemical system is thus

$$\frac{d}{dR} \frac{dR}{dt} < 0 .$$

Only, if this condition is fulfilled, the system returns to the steady state for small perturbations to R . In Figure 9.16 we illustrate the differences between stable and unstable states, using the analogy of a ball in a landscape of changing altitudes. The only stable equilibrium² states are the minima of the height function. As we have seen above, the ball is moving back to the deepest point of the valley after a small perturbation. The position on the top of the mountain is very unstable. For a comprehensive discussion of these issues see [GS94]

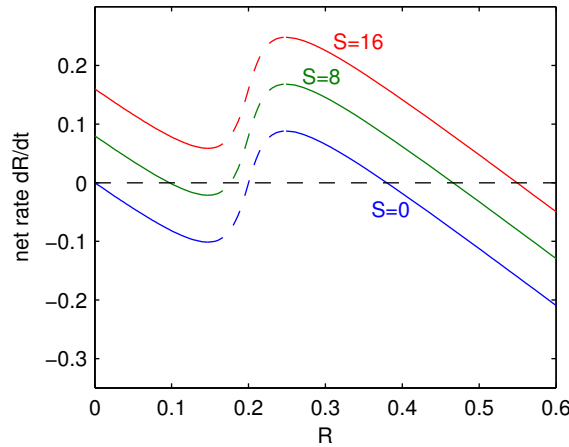


Fig. 9.15: The net-rate as function of the response component R for the system (9.18). The results for three different signal strengths are plotted. The curves are separated into two parts. The solid line represents stable solutions and the dashed lines unstable solutions. Parameters: $k_0 = 0.4$, $k_1 = 0.01$, $k_2 = 1$, $k_3 = 1$, $k_4 = 0.2$, $J_3 = 0.05$, $J_4 = 0.05$.

Returning to our mutually activated module (9.18), in Figure 9.15 a stimulus independent region occurred, for which no stable solution exists. For the chosen parameter values it lies within the interval $0.147 < R < 0.249$. On the other hand, for values of R outside this region we obtain stable solutions. Dependent on the signal strength, the net rate has now one or two stable zeros. Let us analyze the net-rate for a zero stimulus of Figure 9.15 in more detail. As S increases, the location of the zero moves to the right. If the value of the minima is positive, no steady state exists in the lower branch. The steady states of the mutually activated system are obtained from the balance equation

$$0 = k_0 E^*(R) + k_1 S - k_2 R . \quad (9.20)$$

²We assume, that the motion of the ball depends from the height only and that there are no further macroscopic forces. The system is then closed and can reach a state with minimal potential energy, an equilibrium state. Nevertheless, there can be more than one equilibrium states.

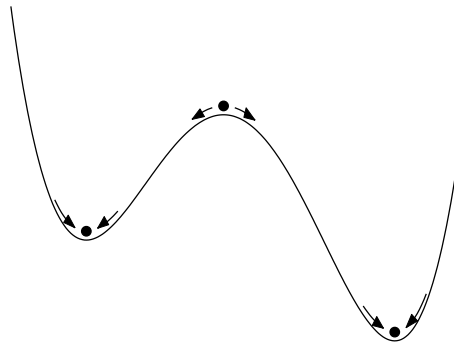


Fig. 9.16: Illustration of stable and unstable states for a ball in a landscape of changing altitudes. The maximum and the minima of this schematic mountain scenery fulfil the necessary condition $dh/dx = 0$, where h denotes the height and x the horizontal position. But only the minima are stable equilibrium states, because the ball will return back from alone to his position after a small perturbation. The maximum is a unstable state; a small perturbation means the ball drops down to one of the equilibrium states.

The analytic solution is lengthy and complicated so that we restrict our discussion to few important features. Equation (9.20) has three solutions. One is a full real function. The others are partly complex. Because the response component must be a real measurable quantity, such as a concentration, only the real parts are relevant for the analysis. All three solutions are shown in the signal-response curve of Figure 9.17. The mutually activated system (9.18) has two stable branches plotted as solid lines with an unstable branch, shown as a dashed line. Such a system is called *bistable*. As the stimulus S increases, the response is low until a critical level of stimulus is reached. At this point the module's behavior changes with a sudden increase of response R . If thereafter S decreases, the response remains high, the switch is irreversible. Note that the sigmoidal module can act as a reversible switch. The plot in Figure 9.17 as a (one-parameter) *bifurcation diagram*. The critical point is in this context called a *bifurcation point*.

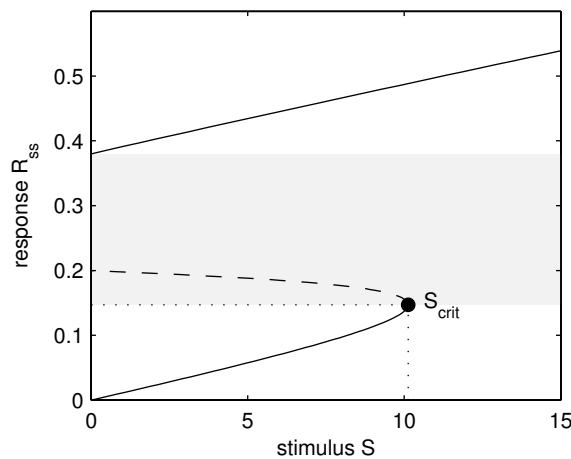


Fig. 9.17: Signal-response curve for the mutually activated system (9.18). The balance equation (9.20) has three branches. The upper one is the full real solution. The others are only in parts real. The steady-states (stable solutions) are shown as solid lines. The unstable solution is drawn as a dashed line. Because of the two stable branches, the system is called bistable. The grey region of the plots denotes the gap between the upper and the lower stable branches. Parameters: $k_0 = 0.4$, $k_1 = 0.01$, $k_2 = 1$, $k_3 = 1$, $k_4 = 0.2$, $J_3 = 0.05$, $J_4 = 0.05$.

Note that so far we have looked at steady values of S and R and whenever we spoke of a change in S , the switch-like behavior was discussed in the stimulus-response plane, not in the time domain. The transition that occurs with changes to S were not explicitly included

in our discussion. Figure 9.18 shows numerical solutions to the differential equations (9.18) for different initial values of R . Figure 9.18 displays solutions to the system with mutual activation through positive feedback. The range of the initial response component R_0 can be separated into two parts. The lower part relaxes to the steady state given by the lower branch of the bifurcation diagram, while for values higher than R_{crit} , the system moves to the upper branch in Figure 9.17. Both parts are separated by the unstable solution of the balance equation (9.20) not by the critical response R_{crit} . The unstable solution is a separatrix in the phase space. At this point it is a good opportunity to point out why stimulus-response curves are valuable. As can be seen from the time-plots, the behavior of the system displayed in those plots depends critically on the stimulus and initial conditions. The bifurcation plot on the other hand summarizes the behavior in an efficient way.

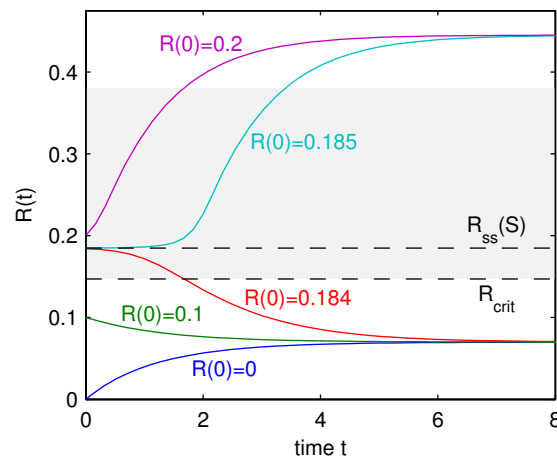


Fig. 9.18: Relaxation of the mutual-activation module into the steady state for different initial response signals R_0 . The unstable solution of the balance equation (9.20) acts as a separatrix, which separates trajectories. For $R_0 < R_{\text{ss}}(S)$ the system achieves a steady state on the lower branch. Outside this range the upper branch becomes the final state. For comparison, the unstable solution and the critical response signal R_{crit} are shown. Parameters: $k_0 = 0.4$, $k_1 = 0.01$, $k_2 = 1$, $k_3 = 1$, $k_4 = 0.2$, $J_3 = 0.05$, $J_4 = 0.05$, $S = 6$.

Figure 9.19 illustrates the irreversibility of the considered system. We compare the response of the system on a sub- and a supercritical signal step change at $t = 0$. The initial response signal is $R_0 = 0$. The signal remains constant until the system reaches a steady state. After a certain time, $t = 15$, the signal is switched and the response relaxes to a new steady state. The sub-critically stimulated system goes back to a zero response signal, while the supercritical stimulus shows the expected behavior of a continued high level. In the second case the nonlinearity of the system is visible. Looking at the response with the subcritical stimulus, one might interpret the temporal evolution $R(t)$ as the consequence of a linear system. Figure 9.20 illustrates the fact that in addition to a critical value of S , the stimulus must persist for a sufficient period of time if the full bistable behavior of the system is to be observed.

Mutual inhibition - positive feedforward

From the above definition of positive feedback/feedforward control, there is another possibility to increase the response signal. In the previous section we increased the rate of production via an intermediate enzyme. Now, we use a similar model to inhibit degradation. Here the response component is acting ‘forward’ via E :

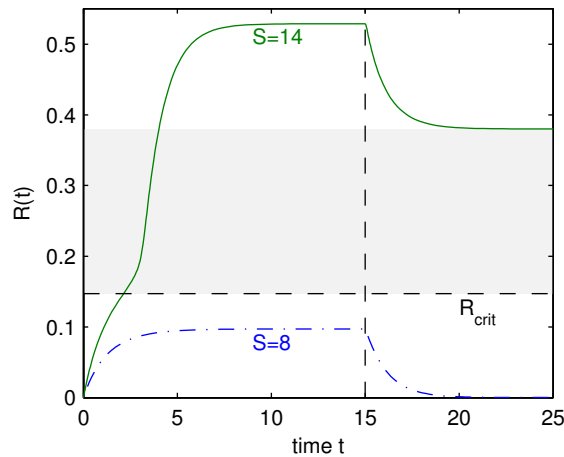


Fig. 9.19: The time evolution for two different stimuli, the first subcritical and the second supercritical. Once the response relaxes into the S -dependent steady state, the signal is switched off. The critical response signal R_{crit} is important for the change of the behavior of the system. At this point the activation/deactivation strongly increase from a lower to a high level. Parameters: $k_0 = 0.4$, $k_1 = 0.01$, $k_2 = 1$, $k_3 = 1$, $k_4 = 0.2$, $J_3 = 0.05$, $J_4 = 0.05$, $S = 14$, $R_0 = 0$.

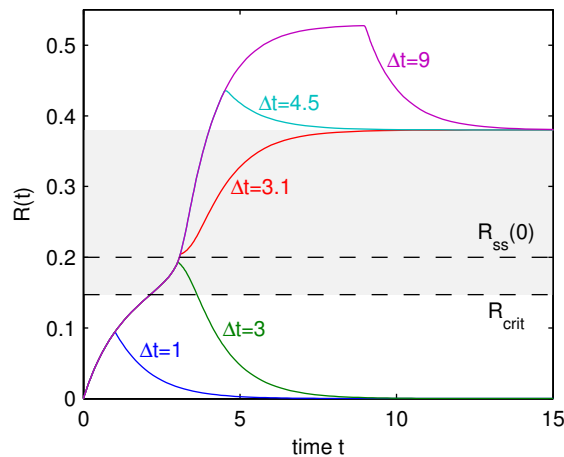
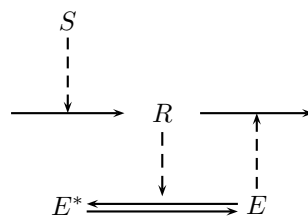


Fig. 9.20: The temporal evolution of the mutually activated system (9.18) for a supercritical signal $S = 14$ of different durations. The activated state is reached only and only if the separatrix $R_{ss}(0)$ is exceeded. Parameters: $k_0 = 0.4$, $k_1 = 0.01$, $k_2 = 1$, $k_3 = 1$, $k_4 = 0.2$, $J_3 = 0.05$, $J_4 = 0.05$, $S = 14$, $R_0 = 0$.



For mutual inhibition, the response component R facilitates the activation of enzyme E . The activated E^* in turn increases R . The corresponding system of differential equations is

$$\begin{aligned} \frac{dR}{dt} &= k_0 + k_1 S - [k_2 + k'_2 E(R)] R \\ E(R) &= G(k_3, k_4 R, J_3, J_4) , \end{aligned} \tag{9.21}$$

where again we assume that the enzyme reaction is much faster than the signalling reaction. Therefore we can use the steady state solution of this reaction given by the Goldbeter-

Koshland function. The term k'_2R describes the direct decay of R and $k'_2E(R)R$ the enzyme catalyzed degradation. The rate curve of this system is shown in Figure 9.21. The solid line is the rate of degradation and the dashed lines are the production rates for different signal strengths. For the mutually inhibited system, the rate of production is independent of the response component. Again, there are more than one intersection for some values of the stimulus. The corresponding balance equation is the necessary condition for a steady state:

$$0 = k_0 + k_1S - [k_2 + k'_2E(R)] R \quad (9.22)$$

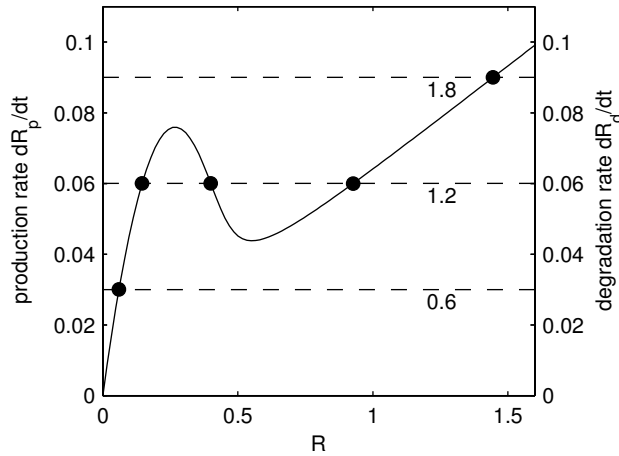


Fig. 9.21: The rate curve of a mutually inhibited system (9.21). The solid line is the rate of degradation. The dashed lines are the R -independent production rates for different signal strength S . Note, the change of the number of intersections (black dots) as a function of the signal.

A more detailed study the properties of the system are obtained from the net rate curve shown in Figure 9.22. We plotted (9.21) as function of the response component R . The stability criterion

$$\frac{d}{dR} \frac{dR}{dt} < 0$$

is fulfilled by the solid lines. For the dashed lines we have $d/dR dR/dt > 0$. Analog to the discussion in the previous section, the system exhibits instabilities. The derivative

$$\frac{d}{dR} \frac{dR}{dt} = - \left[k_2 + k'_2E(R) + k'_2R \frac{dE(R)}{dR} \right] \quad (9.23)$$

is independent of the stimulus S . As described above, the stability criterion has to be fulfilled for a stable steady state. Otherwise, the state is referred to as unstable.

Up to now, we only discussed the difference between the stable and the unstable steady state solutions. We pointed out, that the extremum,

$$\frac{d}{dR} \frac{dR}{dt} = 0 ,$$

of the net rate, limit the unstable range. But, for the critical point we gave no further conditions. We want to do this now but before that we repeat a restriction to our solution. It has to be physically relevant and thus the solution has to be positive, i.e., some quantity proportional to the molecule number (e.g. concentration, density, ...). From the net rate plot we obtain the critical point as zeros of the equation that describes the reaction rate as function of the signal strength. From the balance equation (9.22), it is this point that is limiting the real solution to the equation. The condition is fulfilled for two points of

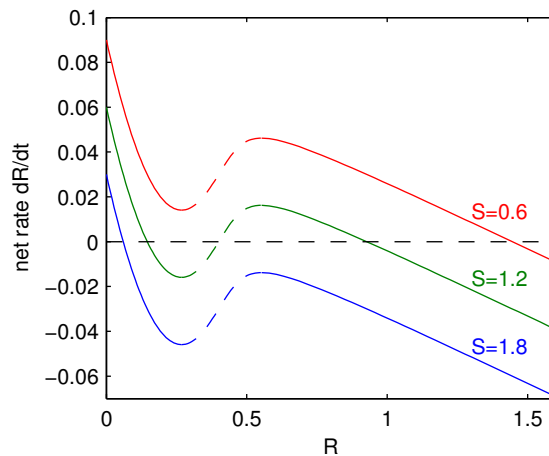


Fig. 9.22: The net rate curve of a mutually inhibited system (9.21) as a function of the response component for different signal strengths. The solid line are stable ranges of the net rate, the dashed lines correspond to unstable regions. The intersection between the net rate and the zero line are solutions of the balance equation (9.22).

the mutually inhibited system (9.21), one minimum and one maximum. The bifurcation points of the mutually inhibited system both satisfy the physical restrictions. We could expect, that the stimulus response curve for this system has two critical points. The stimulus-response curve is shown in Figure 9.23. The stable branches of the balance equation (9.22) are drawn as solid lines. The system is bistable for $S_{crit2} < S < S_{crit1}$ and monostable for all other stimuli. The dashed lines are the unstable solutions limits, described by the critical points S_{crit1} and S_{crit2} . If we increase the strength of the stimulus, starting from the upper level, the steady state response jumps at the first critical point to a high level output. If the stimulus decreases later, the response decreases accordingly. In contrast to the one-way switch in the previous section, the response now goes back to low level if the signal strength is smaller than the second critical point. It is for this reason that we call this system a *toggle switch*. The monostable solutions are reversible, toggle switch the response component is uniquely determined by the signal strength.

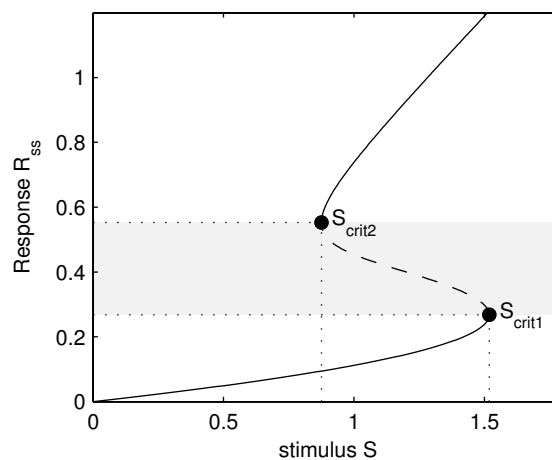


Fig. 9.23: The stimulus-response curve for a mutually inhibited system (9.21). The system is partial bistable. The stable branches are drawn as solid lines and the unstable as dashed lines. The critical points S_{crit1} and S_{crit2} limit the unstable solution. The horizontal and vertical dashed lines corresponds to the signal and response strength of the points. The forbidden range of the steady state response is denoted as grey box. Parameter: $k_0 = 0$, $k_1 = 0.05$, $k_2 = 0.06$, $k_2' = 0.5$, $k_3 = 1$, $k_4 = 0.2$, $J_3 = 10$, $J_4 = 0.01$.

Next, we discuss temporal properties of system (9.21). We assume that the activation of the enzyme is a much faster reaction than the conversion of the signal component S such that we can describe it with the Goldbeter-Koshland function. We first investigate the behavior of the response to a stimulus between the critical points as a function of the initial response component R_0 . For subcritical values $S < S_{\text{crit}2}$ and supercritical $S > S_{\text{crit}1}$ the system is monostable. The response signal moves to the unique steady state for a constant stimulus. The numerical results are plotted in Figure 9.24. Dependent on the initial value, the response signal evolves to the lower or upper steady state. The unstable stimulus dependent solution is again a separatrix. If $R_0 > R_{\text{ss}}(S)$ the upper branch is reached, otherwise the lower branch. Figure 9.25 shows the temporal evolution of the system for successive step-like stimuli of different strengths. Again, we let relax the system to the corresponding steady state and change the stimulus thereafter. In Figure 9.27 the influence of the duration of stimulus is investigated. We start with an initial value $R_0 = 0$ and an external signal $S = 0.6$. This is a subcritical stimulus and the response signal keep on the lower branch. We then increase the stimulus to a critical strength of $S = 1.2$. The response signal remains on the lower level. With the following supercritical signal we force the system to the upper branch of the steady state response. After some time we switch back to the critical signal. As expected, the system now settles to a steady state on the upper branch. The system remembers its previous state. The sub critical stimulus brings the system back on the lower branch. Finally, we switched off the stimulus and the response signal returns to zero.

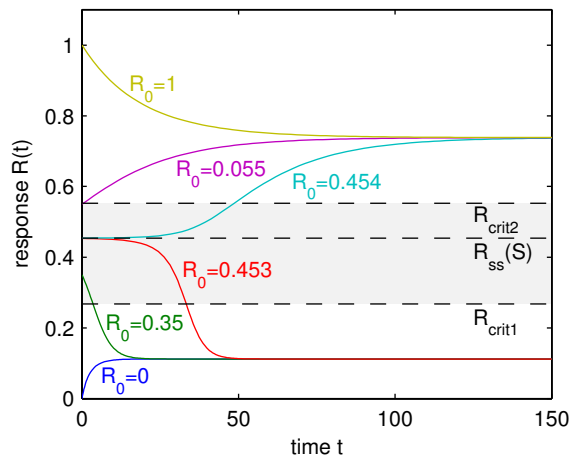


Fig. 9.24: The temporal evolution of the mutually inhibited module (9.21), dependent on the initial response signal R_0 . We choose a signal strength between the two critical points and hence the system is bistable. The unstable solution of the balance equation (9.22) is a separatrix (dashed horizontal line). For initial states greater than this value the system tends to the upper steady state. In the other case the lower steady state is reached. At the corresponding response value of the critical points (dashed horizontal lines) the response signal changes from unstable to stable behavior. Parameters: $k_0 = 0$, $k_1 = 0.05$, $k_2 = 0.06$, $k_2' = 0.5$, $k_3 = 1$, $k_4 = 0.2$, $J_3 = 10$, $J_4 = 0.01$ and $S = 1$.

The temporal evolution of the response component R for different signal strength is shown in Figure 9.26. We choose an example for each range of the stimulus strength. Further on we investigate the behavior for two initial values of the response signal $R_0 = 0$ (solid lines) and $R_0 = 0.7$ (dashed lines). The second initial value is above the unstable region of the steady state response shown as grey box. The limiting horizontal lines denotes $R_{\text{crit}1}$ and $R_{\text{crit}2}$. In both situations the response component settles to the same steady state if we apply a sub- and a supercritical signal. For a critical stimulus the steady state depends on the initial state. The response component changes its behavior within the response gap. In this region it is unstable and does never settle to a limiting value. The value of the inflection points is given by the critical points.

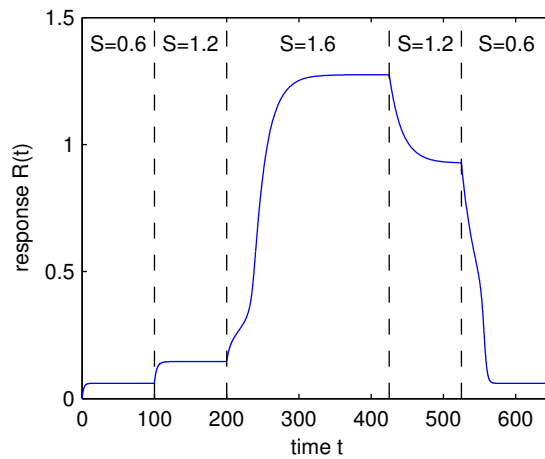


Fig. 9.25: Temporal response for the mutually inhibited system (9.21), given a step-like time dependent stimulus S . The steady state for critical signals (here $S = 1.2$) is dependent on the previous state. In the case of subcritical and supercritical stimuli the steady state is uniquely determined by the signal strength. Parameters: $k_0 = 0$, $k_1 = 0.05$, $k_2 = 0.06$, $k'_2 = 0.5$, $k_3 = 1$, $k_4 = 0.2$, $J_3 = 10$, $J_4 = 0.01$.

Finally, in Figure 9.27 we vary the duration of a step-like signal. We start with the high level response signal $R_0 = 1.1$ and the subcritical external stimulus $S = 0.5$. After the time Δt we switch back to the critical stimulus $S = 1.4$. It follows our previous discussion, that the response component R is decreasing to its lower level steady state if the stimulus is constant. This is clearly seen for long signal durations, for instance $\Delta t = 45$. If we switch to a critical stimulus the system will go to a steady state. Which one, depends on Δt . Only, if the response signal falls below the separatrix, the lower branch of the bistable system (Figure 9.23) defines the new steady state. As long as the response function has not enough time to do this, the system returns to the high level response. We found a similar behavior for the mutually activated system (9.18) in Figure 9.20. A stimulus greater than the critical stimulus value is not enough to change the properties of the system. The duration of the signal must be long enough. The separatrix works like a filter. Fluctuations are suppressed and do not lead to a change in the behavior of the system.

Dependency from parameters

The previous two sections discussed two systems with positive feedback/feedforward mechanisms, leading to bistability. In the present section we investigate the dependency of this special systems property on the rate coefficients k_i . As an example for the two systems, let us consider the mutually activated system (9.18) and vary the coefficient k_0 . The numerical simulations are shown in Figure 9.28. For a better comparison we extend the graphical representation to negative values of the external signal strength S . Remember, the signal strength is per definition a positive definite quantity. Especially, if we consider biochemical networks the external signal is a concentration of molecules.

For $k_0 = 0$ we have the system discussed in the previous section. The system behaves like an irreversible switch, once activated the system never return to the inactivated state. But if we extend our calculations to negative signals we obtain the same hysteresis-like behavior as for the mutually inhibited system. What happens, if we change the parameter k_0 ? In Figure 9.28 we show an example investigation.

Beginning with $k_0 = 0.4$, the situation is equivalent to Figure 9.17. For the second curve, k_0 is chosen such that the second critical point is reached at $S = 0$. For $k_0 = 0.2$,

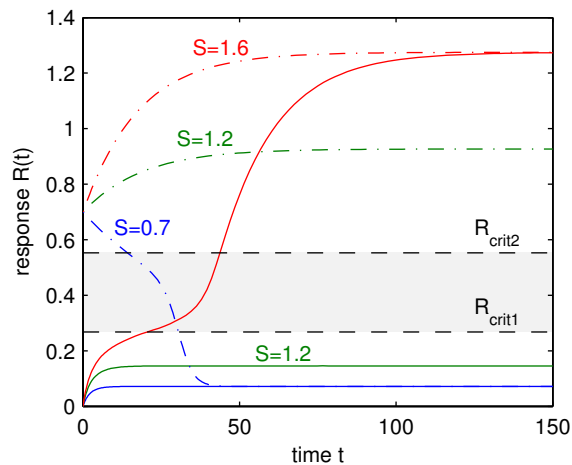


Fig. 9.26: Temporal evolution of the response signal for different signal strengths and initial values. The smallest signal is subcritical, the next is critical and the last is supercritical. Independent on the initial state, the system settles to the same steady state for the subcritical and analogue for the supercritical stimulus. For the critical signal the system shows again memory and reaches two different steady states. Parameters: $k_0 = 0$, $k_1 = 0.05$, $k_2 = 0.06$, $k'_2 = 0.5$, $k_3 = 1$, $k_4 = 0.2$, $J_3 = 10$, $J_4 = 0.01$.

the mutually activated system has a hysteresis-like behavior, equivalent to (9.21) (Figure 9.23). By changing one parameter we therefore alter the system from an irreversible switch to a reversible. If we further decrease k_0 the critical points coincide. The system is then continuous. For smaller values of the rate coefficient k_0 no more critical points exist. For $k_0 = 0$ the system (9.18) is linear:

$$\lim_{k_0 \rightarrow 0} R_{ss} = \frac{k_1}{k_2} S . \quad (9.24)$$

Equation (9.24) represents an approximation for small signal strengths S . In the limit $E(R)$ is a small quantity. The product of k_0 and $E(R)$ is negligible in comparison to the remaining terms in (9.18). The then linear system has a signal-response curve corresponding to (9.24).

The simple investigation of the properties of a mutually activated system illustrates a major problem in modeling biochemical networks: Often a behavior can be realized by more than one kinetic model. The falsification of these models is usually not possible with kinetic methods alone. The concentration of some hypothetic intermediates is not measurable with direct kinetic methods. As we have seen, by changing one rate coefficient, the behavior of the system change dramatically. In biochemical networks the coefficients depend on properties such as temperature, volume and pressure.

9.5.2 Negative Feedback - Oscillations

According to the definition of Section 9.5, negative feedback means the response counteracts the effect of the stimulus. There are two ways in which a negative influence can be exerted: through an acceleration of degradation and a deceleration of production of the response component.

Homoeostasis

homeostasis In *homeostasis*, the response on an external signal (stimulus) is approximately constant over a wide range of signal strength. This behavior may also be described as a kind

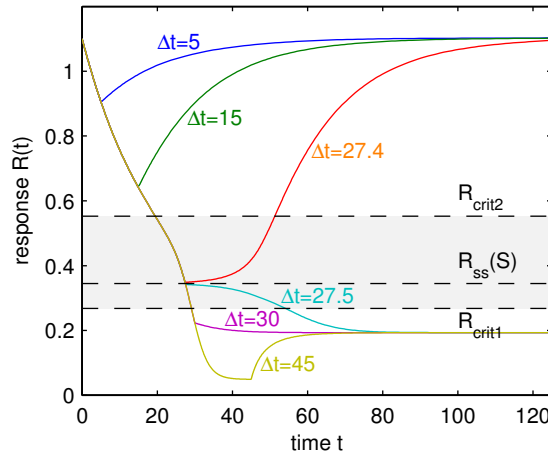


Fig. 9.27: Temporal evolution of the mutually inhibited module (9.21) for a time dependent subcritical signal $S = 0.5$ with varying duration. After the time Δt we apply the critical stimulus $S = 1.4$. Note, the final steady states depends on the duration of the subcritical stimulus. Only if Δt is long enough, the response signal falls below the separatrix given by the unstable solution $R_{ss}(1.4)$. After a short signal the system returns to its high level response. Parameters: $k_0 = 0$, $k_1 = 0.05$, $k_2 = 0.06$, $k_2' = 0.5$, $k_3 = 1$, $k_4 = 0.2$, $J_3 = 10$, $J_4 = 0.01$, $R_0 = 1.1$.

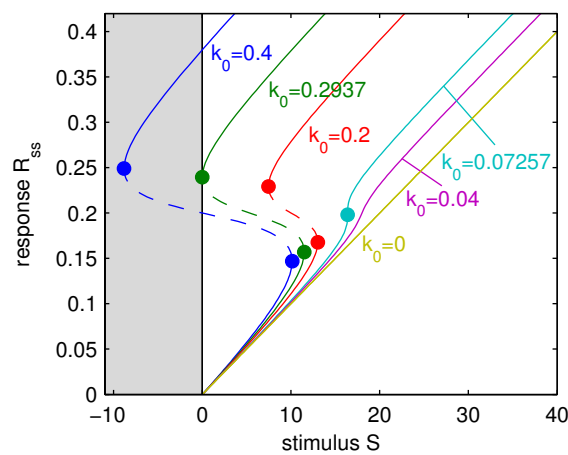
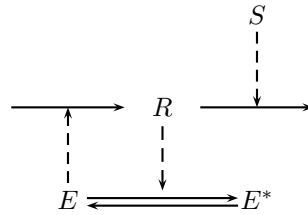


Fig. 9.28: Bistability of the mutually activated system, dependent on parameter k_0 . The stable branches are drawn as solid lines, unstable as dashed lines. The corresponding critical points are shown as circles. With decreasing k_0 the system changes from an irreversible to a reversible switch. The asymptotic system $k_0 = 0$ is linear. Parameters: $k_1 = 0.01$, $k_2 = 1$, $k_3 = 1$, $k_4 = 0.2$, $J_3 = 0.05$, $J_4 = 0.05$.

of imperfect adaption. In contrast to the perfectly-adapted system (9.17) the response component is not changed in response to step-change of the stimulus S (cf. Figure 9.13):



Such a system can be described by the coupled system of differential equations

$$\left. \begin{aligned}
 \frac{dR}{dt} &= k_0 E - k_2 S R \\
 \frac{dE}{dt} &= \frac{k_3 (1 - E)}{J_3 + 1 - E} - \frac{k_4 R E}{J_4 + E}
 \end{aligned} \right\} \quad (9.25)$$

where the response component R inhibits the enzyme catalyzing its synthesis. In (9.25) E is normalized to the total enzyme concentration E_T .

If one assumes, that the enzyme production reaches its steady state much faster than the whole system, we can simplify (9.25) using (9.16). The enzyme concentration is now

$$\left. \begin{aligned}
 \frac{dR}{dt} &= k_0 E(R) - k_2 S R \\
 \frac{dE(R)}{dt} &= G(k_3, k_4 R, J_3, J_4)
 \end{aligned} \right\} \quad (9.26)$$

A comparison of production and degradation rate is shown in Figure 9.29.

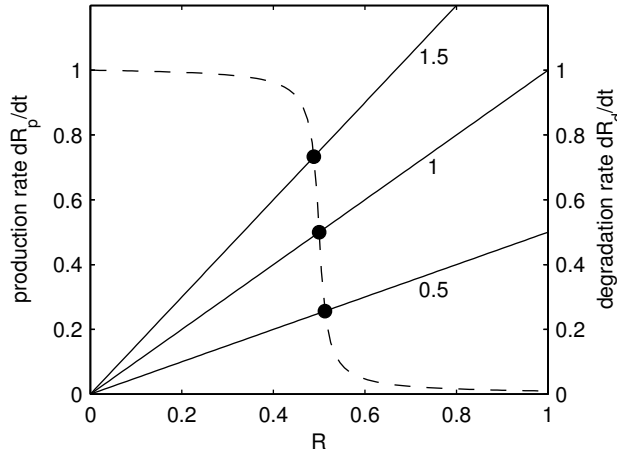


Fig. 9.29: Comparison of rate of response and degradation for the homoeostatic system (9.26) for stimulus S . The solid lines are the degradation rates, (9.26), i.e., linear functions with slope $k_2 \cdot S$. The rate of response depends on R and has the typical sigmoidal shape of (9.16). The intersections, denoted by dots, are again steady state solutions. Parameters: $k_0 = k_2 = 1$, $k_3 = 0.5$, $k_4 = 1$, $J_3 = J_4 = 0.01$, $E_T = 1$.

The rate of production

$$\text{rate of production} = k_0 E(R)$$

implicitly depends from the response component R . The result is a sigmoidal curve. The degradation rate

$$\text{rate of degradation} = k_2 S R$$

is a linear function with slope k_2S . For those chosen range, it is assumed that the steady state response is nearly independent from the external signal. The net rate, shown in Figure 9.30, displays an analogous behavior. Again, we compare the influence of different stimuli to the overall rate. The intersections with the dashed line, where the rate is zero, represent the steady states. The corresponding balance equation

$$0 = k_0E(R) - k_2SR \quad (9.27)$$

can be transformed into

$$\frac{k_2}{k_0}S = \frac{E(R)}{R} \quad (9.28)$$

A solution for the dependence on the response component can only be found numerically. On the other hand, (9.28) gives a rule to calculate the strength of the stimulus S for known R . This implies that the ratio of the rate coefficients k_0 and k_2 plays an important role for the behavior of the considered system. The corresponding signal-response curve is shown in Figure 9.31, where we the ratio k_2/k_0 is varied. The ratio determines the length of the expected plateau. The amplitude is weakly affected. For small signals the steady state response is singular and for strong signals it tends to zero.

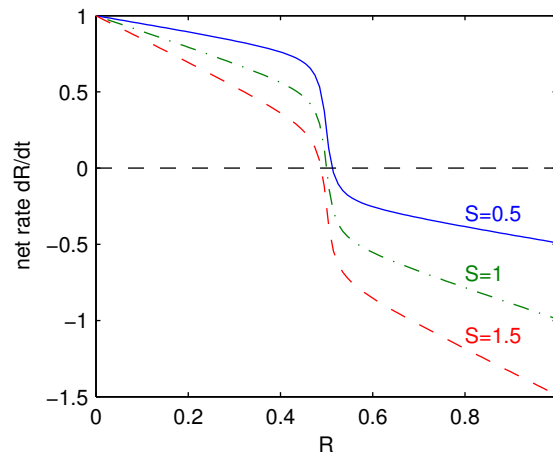


Fig. 9.30: Net rate for the homoeostatic system (9.26) for different external stimuli. The intersections with the horizontal dashed line are the corresponding steady states. The steady states are only weakly dependent on the stimulus S . Parameters: $k_0 = k_2 = 1$, $k_3 = 0.5$, $k_4 = 1$, $J_3 = J_4 = 0.01$, $E_T = 1$.

For a discussion of steady state properties for a homoeostatic system we neglected the time dependence of the enzyme activation/deactivation reaction in (9.25). For the study of the temporal behavior of such a system we do not make the assumption of a much faster reversible enzymatic reaction. Before we continue this, let us return to the steady state properties, derived from the coupled balance equations

$$0 = k_0E - k_2SR$$

$$0 = \frac{k_3(E_T - E)}{J_3 + E_T - E} - \frac{k_4RE}{J_4 + E}$$

The solution of the enzymatic equation is again the Goldbeter-Koshland function we used in (9.26). After insertion into the first equation we obtain again (9.28). The steady state solution of a coupled system of equations remains therefore unaffected. We thus expect, that the system will reach the same steady state in its temporal evolution. In Figure 9.32 we plot numerical solutions of (9.25) for the response component and compare it with solutions of (9.26). The differences in temporal behavior can be divided into three classes,

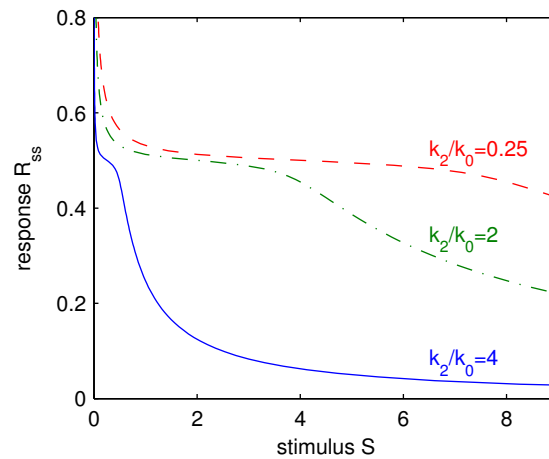


Fig. 9.31: Stimulus-response curve for system (9.26) as a function of different ratios of rate coefficients k_0 and k_2 . The steady state response shows, for small ratios, a plateau over a range of the signal S . The plateau decreases with increasing ratio and eventually disappears. Parameters: $k_0 = 1$, $k_3 = 0.5$, $k_4 = 1$, $J_3 = J_4 = 0.01$, $E_T = 1$.

corresponding to the three ranges in the stimulus-response curve. In the homeostatic range, the system displays damped oscillations around the steady state. Common for all three cases is a difference in the relaxation time. The system (9.25) takes longer than (9.26) to reach the steady state. The simplification made in (9.26) makes it easier to handle the differential equations and does not affect the steady state. On the other hand, no oscillations occur with this approximation. Nevertheless, the assumption of a much faster process is often used to simplify the treatment.

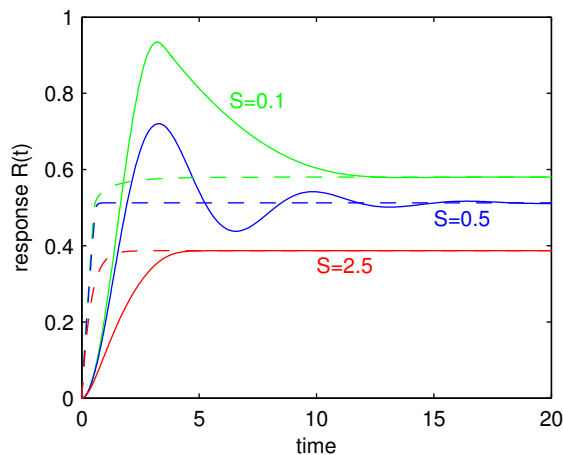
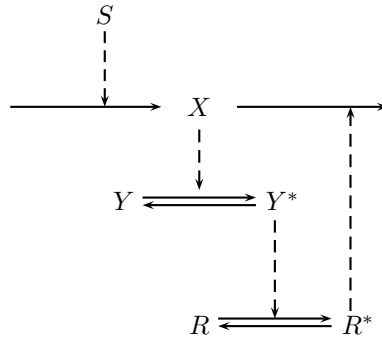


Fig. 9.32: Comparison of temporal evolution of homoeostatic system (9.25) and its simplified version (9.26) for different signal strengths S . The solutions of the full system are drawn as solid lines, while dashed lines are used for the simplified case. Parameters: $k_0 = k_2 = 1$, $k_3 = 0.5$, $k_4 = 1$, $J_3 = J_4 = 0.01$, $E_T = 1$.

Negative Feedback Oscillator

The previous section showed how negative feedback can introduce damped oscillations. We here show how negative feedback can lead to stable oscillations. Therefore we consider a system of three components



The mathematical model for this system is defined by the following equations

$$\left. \begin{aligned} \frac{dX}{dt} &= k_0 + k_1 S - (k_2 + k'_2 R^*) X \\ \frac{dY^*}{dt} &= \frac{k_2 X (Y_T - Y^*)}{K_{M3} + Y_T - Y^*} - \frac{k_4 Y^*}{K_{M4} + Y^*} \\ \frac{dR^*}{dt} &= \frac{k_5 Y^* (R_T - R^*)}{K_{M5} + R_T - R^*} - \frac{k_6 R^*}{K_{M6} + R^*} \end{aligned} \right\} \quad (9.29)$$

where X activates the protein Y . The activated protein Y^* activates the next protein R . Its activated form catalyzes the degradation of X . Another possible way to close the negative feedback loop is the inhibition of production from S . We focus on the first case. A numerical simulation of (9.29) is given in Figure 9.33. All three components show oscillations. The third component Y^* introduces a time delay in the feedback loop, causing the control system repeatedly to over- and undershoot its steady state. Within the shown interval the oscillations are damped. The system takes a certain time to establish stable oscillations, depending on the chosen set of parameters. For instance, if one increases the rate coefficient k'_2 to 20, the amplitude will initially increase until the limit cycle is reached. A variation of k_0 influences the strength of damping or amplifying and the amplitude of the limit cycle. In the phase-plane representation, Figure 9.34, it is shown, how the three components settle towards a limit cycle.

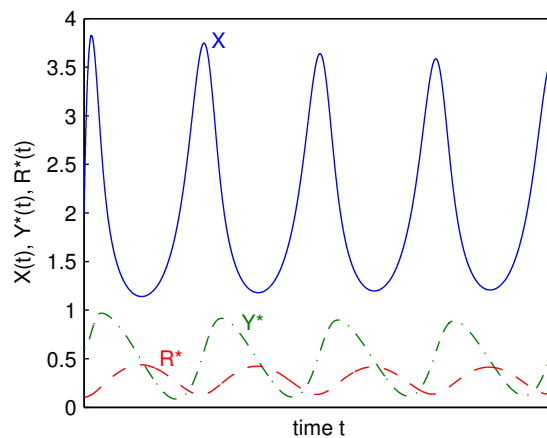


Fig. 9.33: The temporal evolution of the negative feedback oscillator (9.29). All three components X , Y^* , and R^* perform oscillations. Within the shown time interval these oscillations are damped. Parameters: $k_0 = 0$, $k_1 = 1$, $k_2 = 0.01$, $k'_2 = 10$, $k_3 = 0.1$, $k_4 = 0.2$, $k_5 = 0.1$, $k_6 = 0.05$, $K_{M3} = K_{M4} = K_{M5} = K_{M6} = 0.01$, $Y_T = R_T = 1$, $S = 5$, $X_0 = 2$, $Y^* = 0.6$, $R^* = 0.1$.

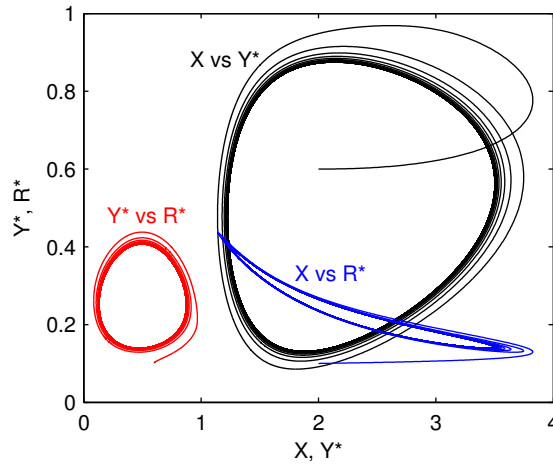


Fig. 9.34: Phase-plane representation of numerical solution of (9.29). We combine three possible combinations in this figure. All components perform oscillations and tend to the limit cycle within a certain number of oscillations. Parameters: $k_0 = 0, k_1 = 1, k_2 = 0.01, k_2' = 10, k_3 = 0.1, k_4 = 0.2, k_5 = 0.1, k_6 = 0.05, K_{M3} = K_{M4} = K_{M5} = K_{M6} = 0.01, Y_T = R_T = 1, S = 5, X_0 = 2, Y^* = 0.6, R^* = 0.1$.

Let us now focus on steady state properties. The state steady state are derived from the set of coupled balance equations

$$\begin{aligned}
 0 &= k_0 + k_1 S - (k_2 + k_2' R^*) X \\
 0 &= \frac{k_2 X (Y_T - Y^*)}{K_{M3} + Y_T - Y^*} - \frac{k_4 Y^*}{K_{M4} + Y^*} \\
 0 &= \frac{k_5 Y^* (R_T - R^*)}{K_{M5} + R_T - R^*} - \frac{k_6 R^*}{K_{M6} + R^*}
 \end{aligned} \tag{9.30}$$

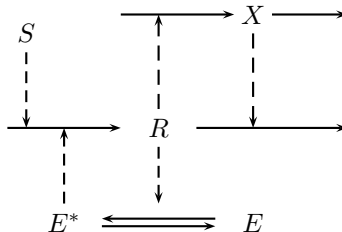
The balance equation for each subsystem itself is solvable but the steady state solution of the coupled system for the response component is lengthy and complicated. Therefore we prefer a numerical solution of (9.30). The corresponding stimulus-response curve is plotted in Figure 9.35. The solution is separated into two stable (solid line) and an unstable range (dashed line). In the unstable range the system performs stable oscillations. The amplitude of the oscillation depends on the stimulus S as shown by the dashed-dotted curve.

9.5.3 Mixed Control Mechanisms

In the present section, different feedback mechanisms are combined.

Activator-Inhibitor-Oscillator

Our first example combines the mutually activated system (9.18) and an autoinhibition of the response component.



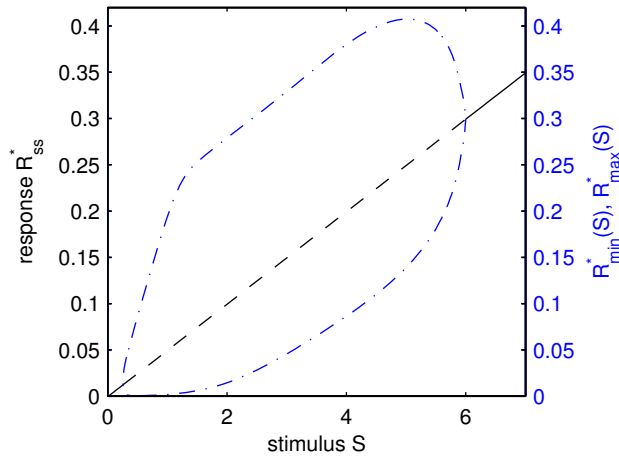


Fig. 9.35: Stimulus-response curve as function of the external signal strength S for the negative feedback oscillator (9.29). The straight line is the numerical steady state solution of (9.30). The solid parts denote stable ranges. In the interval $0.26 < S < 6$, the solution is unstable and the system performs oscillations, as in Figure 9.33. The maxima and minima as a function S are plotted as the dash-dotted curve. Parameters: $k_0 = 0$, $k_1 = 1$, $k_2 = 0.01$, $k'_2 = 10$, $k_3 = 0.1$, $k_4 = 0.2$, $k_5 = 0.1$, $k_6 = 0.05$, $K_{M3} = K_{M4} = K_{M5} = K_{M5} = 0.01$.

The response component R is produced in an autocatalytic process. It activates the enzyme E^* which accelerate the production of R . On the other hand, the response component promotes the production of the inhibitor X at the same time. The inhibitor speeds up the removal of R . Again, we assume that the enzyme is always in its steady state described by the relation (9.16), assuming the activation/deactivation process is much faster than the other reactions in the system. This assumption simplifies our further discussion and restrict the (mathematical) dimension of the corresponding system of differential equations

$$\begin{aligned} \frac{dR}{dt} &= k_0 E^*(R) + k_1 S - (k_2 + k'_2 X) R \\ \frac{dX}{dt} &= k_5 R - k_6 X \\ E^*(R) &= G(k_3 R, k_4, J_3, J_4) \end{aligned} \quad (9.31)$$

For this composed system an analysis in terms of rate of production/degradation and the net rate is only possible in three dimensions³. The assumption of the steady state for the enzyme avoids an additional fourth dimension. The temporal evolution of the response signal $R(t)$ and the inhibitor $X(t)$ is numerically solvable from the coupled differential equations (9.31). For the chosen set of parameters we obtain stable oscillations in Figure 9.36 for both components. If the amount of R small, the production of the response component is the main process. The degradation of the inhibitor is faster than its production. This results in an increase of the response component. With increasing R also the production of the inhibitor is increasing. The acceleration of the degradation of R leads to a decrease of the response component, returning us to where we started.

The phase plane representation of the oscillations is shown in Figure 9.37. Additionally to the limit cycle of the oscillation we plotted the steady states as a function of the response component R and X . The steady states are obtained from the balance equations

$$\frac{dR}{dt} = k_0 E^*(R) + k_1 S - (k_2 + k'_2 X) R = 0 \quad (9.32)$$

³One for the response component R , one for the rate of change of the response component, and one for the rate of change of the inhibitor X .

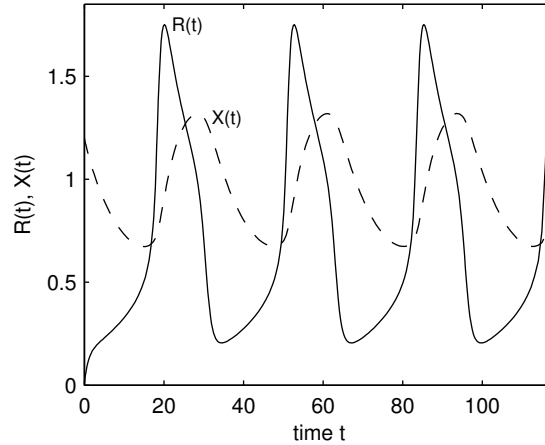


Fig. 9.36: The temporal evolution of system of an activator and an inhibitor (9.31) as function of time. The solid line is the response component and the dashed line the inhibitor. After a short starting time the system carries out stable oscillations. Parameters: $k_0 = 4$, $k_1 = k_2 = k'_2 = k_3 = k_4 = 1$, $k_5 = 0.1$, $k_6 = 0.075$, $J_3 = J_4 = 0.3$, $S = 0.2$, $R_0 = 0$, $X_0 = 1.2$.

and

$$\frac{dX}{dt} = k_5 R - k_6 X = 0 \quad (9.33)$$

Equation (9.33) for the inhibitor can be solved analytically leading to a straight line with slope k_6/k_5 :

$$R = \frac{k_6}{k_5} X . \quad (9.34)$$

In Figure 9.37, the closed curve is the phase-plane representation of the temporal evolution shown in Figure 9.36. After an initial time the system reaches a stable limit cycle. The straight line and the line that is in parts dashed, are the stimulus-response curves of the subsystems. The straight line is the analytic solution (9.34). The solution of (9.32) is numerically found. The stable solutions are shown as solid sections of the line and the unstable solution as a dashed line. The corresponding critical points are shown as filled dots. The solutions of the balance equations (9.32) and (9.33) have an intersection, which is the steady state solution shown as a dot on the straight line. For the chosen parameters the steady-state is unstable leading to an oscillating behavior. As one subsystem tries to reach one of its two stable states, the resulting production or degradation of X forces it back to the unstable state and the cycle starts again. The intersection depends on the signal strength S . By an increase or decrease of S we move the solution of (9.32) in Figure 9.37 until the intersection is stable. For the corresponding signal strengths the activator-inhibitor system has an stable steady state. No more stable oscillations occur. Such a situation is shown in Figure 9.38. At the intersection both solutions are stable, although near the critical point. The system (9.31) shows damped oscillations and reaches a stable steady state. The phase plane shows a typical spiralling curve with decreasing amplitude. For still higher signal strengths the intersection moves further away from the critical point and the damped oscillations will disappear.

As discussed above, the oscillatory behavior of the considered system (9.31) strongly depends on the strength of external signal S . Oscillations occur only if the intersection between the steady states of each subsystem is on the unstable branch of (9.32). The parameter we can change is the external signal S . The internal parameter we assume as inherent, uniquely determined by conditions like temperature, pH -value and so on. The stimulus-response curve shows a new qualitative property. There are intervals in the signal strength S , where the system tends to a steady state. For the chosen values of the rate coefficients, see Figure 9.39, this behavior is established in the intervals $0 < S < 0.066$ and

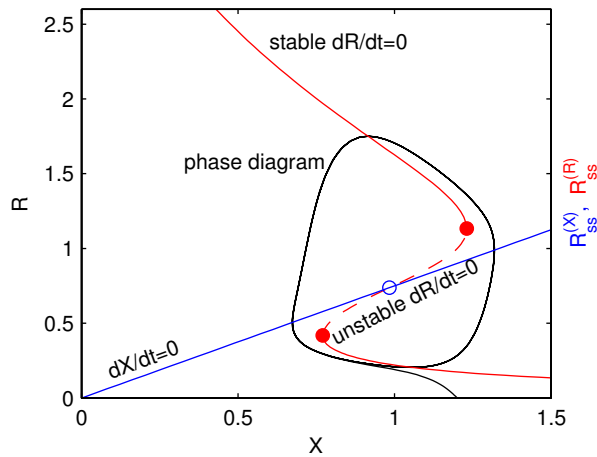


Fig. 9.37: Phase-plane representation and stimulus-response curve for the activator-inhibitor system (9.31) with a constant external signal strength. Parameters: $k_0 = 4, k_1 = k_2 = k'_2 = k_3 = k_4 = 1, k_5 = 0.1, k_6 = 0.075, J_3 = J_4 = 0.3, S = 0.2, X_0 = 1.2, R_0 = 0$.

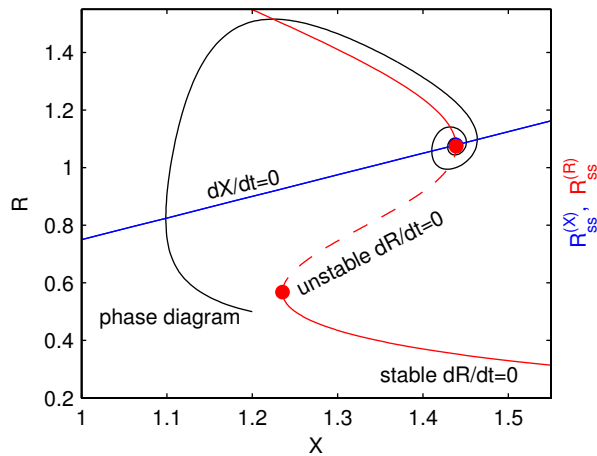
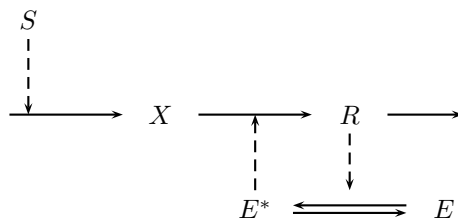


Fig. 9.38: Phase-plane representation and stimulus-response curves for the activator-inhibitor subsystems (9.32)-(9.33). In contrast to Figure 9.37 the system reaches a stable steady state. Starting with R_0 the system has damped oscillations around this stable state and ends finally in it. Parameters: $k_0 = 4, k_1 = k_2 = k'_2 = k_3 = k_4 = 1, k_5 = 0.1, k_6 = 0.075, J_3 = J_4 = 0.3, S = 0.43, X_0 = 1.2, R_0 = 0.5$.

$S > 0.41$. In the intermediate range the response signal R oscillates around the unstable steady state, shown as a dashed line in Figure 9.39. The amplitude and minimal (R_{\min}) and maximal (R_{\max}) values are functions of the stimulus signal S .

Substrate-Depletion-Oscillator

As second example of an oscillating mixed system we choose a substrate-depletion-oscillator.



The signalling component S is converted into the intermediate X . From the intermediate the response component R is produced. On the other hand the response component activates the enzyme E^* , increasing the conversion rate from X to R . For such a reaction

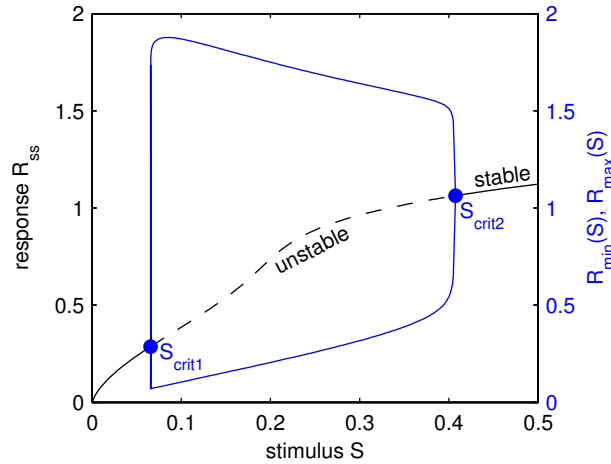


Fig. 9.39: The stimulus-response curve for the system (9.31). The steady state solutions of the coupled system of equation (9.32) and (9.33). The stable solutions are drawn as solid and the unstable solution as the dashed section of the line. The closed curve shows the maximal and minimal values of the response signal in the case of stable oscillations. On the critical points $S_1 \approx 0.066$ and $S_2 \approx 0.41$ the system changes its behavior abruptly. Parameters: $k_0 = 4$, $k_1 = k_2 = k_2' = k_3 = k_4 = 1$, $k_5 = 0.1$, $k_6 = 0.075$, $J_3 = J_4 = 0.3$.

scheme we obtain the following system of coupled differential equation

$$\begin{aligned} \frac{dX}{dt} &= k_1 S - (k_0' + k_0 E^*(R)) X \\ \frac{dR}{dt} &= (k_0' + k_0 E^*(R)) X - k_2 R \\ E^*(R) &= G(k_3 R, k_4, J_3, J_4) . \end{aligned} \quad (9.35)$$

We again assume that the activation/deactivation of the enzyme is much faster than the other reaction. Hence, the enzyme E^* is assumed to be always in a steady state. The corresponding steady state solution is then (9.16). In Figure 9.40 the numerical simulation of (9.35) is shown. For the chosen set of parameters the system displays stable oscillations. First, the amount of the intermediate X increases faster than it is converted into the response component R . But R promotes its own production via the enzyme E^* . According to the sigmoidal shape of (9.16), the positive feedback term $k_0 E^*$ is small in comparison to the direct transformation with k_0' . The response component will increase approximately linear. If R exceeds a critical value the enzyme concentration jumps to a high value. The conversion from X into R is now determined by the reaction of the activated enzyme and the intermediate. This reaction is faster than the production of X . The result is a strong increase of the response component until the intermediate is depleted. Then the fast production of R breaks down. On the other hand, the decay of the response component is now faster as its slow production and its concentration decrease. In the mean time the amount of X increase again.

Next we investigate the system (9.35) in more detail. We start with a computation of the steady state, solving the coupled system of balance equations

$$0 = k_1 S - (k_0' + k_0 E^*) X , \quad (9.36)$$

$$0 = (k_0' + k_0 E^*) X - k_2 R . \quad (9.37)$$

For the respective subsystems we derive

$$X = \frac{k_1 S}{k_0' + k_0 E^*(R)} \quad (9.38)$$

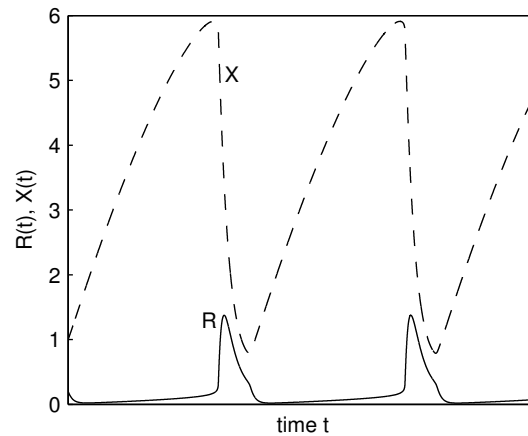


Fig. 9.40: The oscillatory behavior of the system (9.35). The response signal $R(t)$ is drawn as solid line and the intermediate $X(t)$ as a dashed line. Parameters: $k_0 = 0.4$, $k'_0 = 0.01$, $k_1 = k_2 = k_3 = 1, k_4 = 0.3$, $J_3 = J_4 = 0.05$, $S = 0.2$, $X_0 = 1$, $R_0 = 0.2$.

from (9.36) and

$$R = \frac{k'_0 + k_0 E^*(R)}{k_2} X \quad (9.39)$$

for the second equation (9.37). Both solutions are shown in Figure 9.41, together with a phase-plane representation of the temporal evolution in Figure 9.40. Whereas, (9.38) is a monostable function, the stimulus-response curve of the R -subsystem is bistable. This bistability triggers the occurrence of oscillations. Remember, the steady state of the substrate-depletion oscillator is given by the intersection of both curves. Mathematically we have to derive the balance equations (9.36) and (9.37) simultaneously. The result is the linear function

$$R_{\text{ss}} = \frac{k_1}{k_2} S. \quad (9.40)$$

But for a stable steady state all subsystems have to be in a stable steady state. This is not the case for the given set of parameters in Figure 9.41. The system performs oscillations around the steady state solution (9.40).

If we choose other rate coefficients the intersection of (9.38) and (9.39) change its position. Again, in analogy to the activator-inhibitor system (9.31) the state space of substrate-depletion oscillator (9.35) is separated into a region of stable oscillation and a non-oscillating part, where the system tends to a stable steady state. In Figure 9.42 this is illustrated with the stimulus-response representation. The straight line are the steady state solutions (9.40) following from (9.36) and (9.37) as a function of the external signal strength S . Solid parts of the line corresponds to a stable steady state. With the dashed part the system performs stable oscillations around the formal solution. The closed line corresponds to maximal and minimal values of the amplitude. This pictures shows a sharp transition between both behaviors. The value of minimal response component is nearly constant, where the maximum depends from the external signal.

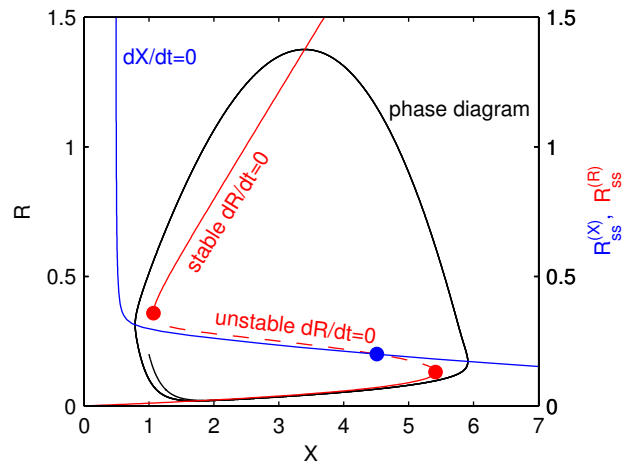


Fig. 9.41: Phase-plane representation of the substrate-depletion oscillator (9.35) combined with stimulus-response curves of the subsystems. The solid black line is the limit cycle of the stable oscillations of intermediate X and response component R . The balance equations (9.36) and (9.37) were numerically solved. The steady state solution, the intersection of the stimulus-response curves, is located in the unstable branch of the response component, the system therefore displays oscillations around this point. Parameters: $k_0 = 0.4$, $k'_0 = 0.01$, $k_1 = k_2 = k_3 = 1$, $k_4 = 0.3$, $J_3 = J_4 = 0.05$, $S = 0.2$, $X_0 = 1$, $R_0 = 0.2$.

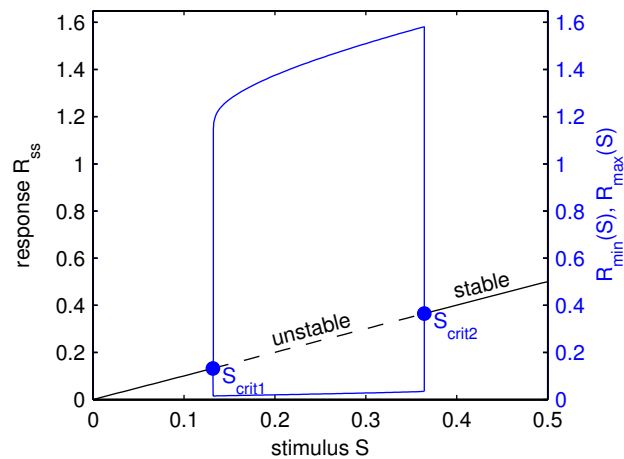


Fig. 9.42: Stimulus-response curve for the substrate-depletion oscillator (9.35). The straight line are the steady states. In the interval $0.132 < S < 0.365$ the system has no stable solution of the balance equations, shown as a dashed segment of the line. Under these conditions oscillations occur. The maximal and minimal values of the response component are shown as a closed curve. Again the amplitude of oscillations depends from the signal strength. Outside the unstable region the response component tends to the steady state value. Near the critical points damped oscillations occur. Parameters: $k_0 = 0.4$, $k'_0 = 0.01$, $k_1 = k_2 = k_3 = 1$, $k_4 = 0.3$, $J_3 = J_4 = 0.05$.

10 Approximations and the Consequences

The study of cellular dynamics is closely linked to the concept of a *pathway*. Here we consider a pathway as a *network* of biochemical reactions. We distinguish between *metabolic pathways* and *signalling pathways*. In metabolic analysis one usually considers a system in steady state, that is $dS/dt = 0$. In signal transduction or cell signalling, we care primarily for processes that transduce a signal. We are thus mostly interested in transient and oscillatory phenomena. In any case, it is of vital importance to be aware of the time scale in which we model observations. Processes in the cell occur across a wide range of time scales. For modelling this means that processes that are very slow in comparison to the one we model, can be assumed as constant (see Figure 10.1). In many cases this makes the model much simpler without losing too much realism.

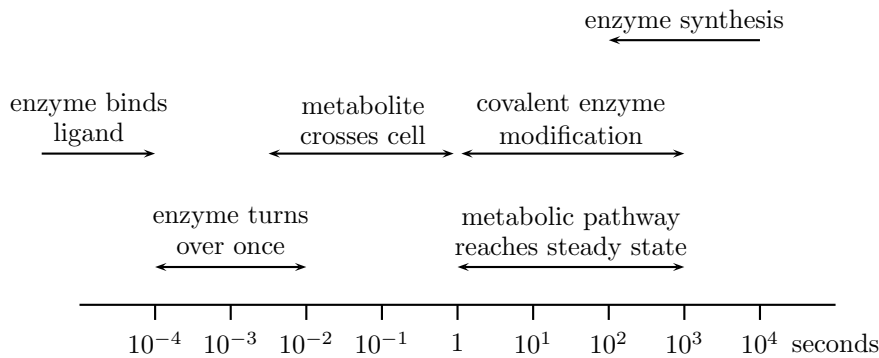


Fig. 10.1: Processes in the cell occur at widely differing time scales. If a process is relatively fast, relative to the one we are considering, we may assume that changes in this process are instantaneous. If on the other hand a process is relatively slow to the one we consider, we may assume the variable to be constant. Picture adopted from [Fel97].

...

11 Cell Cycle

Where a cell arises, there must be a previous cell, just as animals can only arise from animals and plants from plants. This cell doctrine, proposed by the German pathologist Rudolf Virchow in 1858, carried with it a profound message for the continuity of life. Cells are generated from cells, and the only way to make more cells is by division of those already exist. All living organisms, from the unicellular bacterium to the multicellular mammal, are products of repeated rounds of cell growth and division extending back in time to the beginning of life on Earth over three billion years ago.

A cell reproduces by performing an orderly sequence of events in which it duplicates its contents and then divides in two. This cycle of duplication and division, known as cell cycle, is the essential mechanism by which all living things reproduce. In unicellular species, such as bacteria and yeasts, each cell division produces a complete new organism. In multicellular species, long and complex sequences of cell divisions are required to produce a functioning organism. Even in the adult body, cell division is usually needed to replace cells that die. In fact, each of us must manufacture many million of cells every second simply to survive: if all cell division were stopped - by exposure to a very large dose of x-rays, for example - we would die in few days.

The minimum set of processes that a cell has to perform are those that allow it to accomplish its most fundamental task: the passing on of its genetic information to the next generations of cells. To produce two genetically identical daughter cells, the DNA in each chromosome must first be faithfully replicated to produce two complete copies, and the replicated chromosomes must then be accurately distributed (segregated) to the two daughter cells, so that each receives a copy of the entire genome. In addition to duplicating their genetic material, most cells also duplicate their other organelles and macromolecules; otherwise, they would get smaller with each division. To maintain their size, dividing cells must coordinate their growth (i.e., their increase in cell mass) with their division.

11.1 An Overview of the Cell Cycle

For this overview about the cell cycle we follow the representation of Alberts et al. [AJL⁺02]. The most basic function of the cell cycle is to duplicate accurately the vast amount of DNA in the chromosomes and then segregate the copies precisely into two genetically identical daughter cells. These processes defines the two major phases of the cell cycle. DNA duplication occurs during the S phase (S for synthesis), which requires 10-12 hours and occupies about half of the cell-cycle time in a typical mammalian cell. After S phase, chromosomes segregation and cell division occur in M phase (M for mitosis), which requires much less time (less than an hour in a mammalian cell).

Most cells requires much more time to grow and double their mass of proteins and organelles than they require to replicate their DNA and divide. Partly to allow more time for growth, extra gap phases are inserted in most cell-cycles — a G_1 phase between M phase and S phase and a G_2 phase between S phase and mitosis. Thus, the eucaryotic cell cycle is divided into four sequential phases: G_1 , S, G_2 , and M. G_1 , S, and G_2 together are called interphase, see also Fig. 11.1. In a typical human cell proliferating in culture, interphase might occupy 23 hours of a 24 hours cycle, with an 1 hour for M phase.

The two gap phase serve as more than simple time delays to allow cell growth. They also provide time for the cell to monitor the internal and external environment to ensure that conditions are suitable and preparations are complete before the cell commits itself to the major upheavals of S phase and mitosis. The G_1 phase is especially important in this

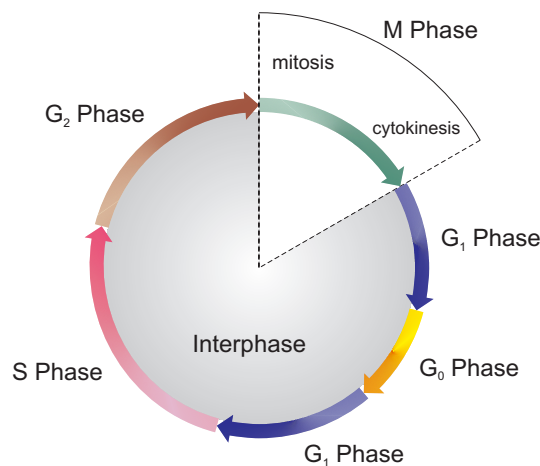


Fig. 11.1: The phases of the cell cycle. The cell grows continuously in interphase, which consists of three phases: S phase where the DNA is replicated; G₁ is the gap between M phase and S phase, which can be interrupted by a resting phase G₀. G₂ is the gap between S phase and M phase. In M phase, first the nucleus and then the cytoplasm divide.

respect. Its length can vary greatly depending on external conditions and extracellular signals from other cells. If extracellular conditions are unfavorable, for example, cells delay progress through G₁ and may even enter a specialized resting state known as G₀, in which they can remain for days, weeks, or even years before resuming proliferation. If extracellular conditions are favorable and signals to grow and divide are present, cells in early G₁ or G₀ progress through a commitment point near the end of G₁. After passing this point, cells are committed to DNA replication, even if the extracellular signals that stimulate cell growth and division are removed. An introduction to underlying modules and control mechanisms is given in Chapter 9.

The cell cycle is separated roughly into four phases (excluding an additional resting phase). But Mitosis and S-Phase play a special role in this complex sequence of cellular events. During these phases some remarkable changes in the contents and functionality of the cell are occurring.

11.1.1 Mitosis - Key events

Mitosis is the final step in the process of cell division. But before the cell can divide into two daughter cells some preparing steps are necessary. Mitosis can be separated into four key events as represented [MK89]. In the Prophase, the dispersed duplicated chromosomes are condensed into a state suitable for transport. Additionally, the metabolic activity is reduced. The condensed and inactive chromosomes are positioned at the equator of the nuclei in the Prometaphase. This phase is followed by the Anaphase, where the duplicated chromosomes are separated into two identical parts. Due to the progress in experimental technique and technology, this can be observed now in experiments. Subsequently, the chromosomes move towards opposite ends of the nuclei. In the Telophase the chromosomes decondense and become metabolically active. Simultaneously, there is a reformation of the nuclei.

Additionally to the processes inside the nuclei, there are further rearrangements in the cytoplasm providing the cell division. For example, the microtubuli and microfilaments of the cytoskeleton are rearranged, the Golgi apparatus is dispersed into a large number of small vesicles and is distributed throughout the cell. Additionally, the protein synthesis is slow down to about 25% of its normal rate.

11.1.2 S-Phase

As mentioned earlier, the chromosomes are doubled during the S-phase. But the synthesis not only replicates the chromosomes completely and precisely. In addition, the cell also has to duplicate the complex chromosome architecture [LFB89]. Especially, it copies the specific patterns of gene activity and inactivity. For this reason, a specialized cell divides into two daughter cells with the same functionality. A further sub-specialization of the daughter cell is possible during their life time. Additionally to functional information, also cell cycle specific data are stored and transferred during the S-phase. Eukaryotic cells tag the replicated DNA and distinguish between replicated and unreplicated DNA. Hence, the replication of DNA take place only once between cell divisions. This is a hallmark of eukaryotic cells distinguishing them from prokaryotes. Furthermore the most differentiate cells contains information about the number of replications. It prohibits the infinite division of a cell line. After a certain number of replications the cell cannot duplicate the chromosomes. The cell cycle cannot be finished and internal regulatory mechanisms drive the cell into Apoptosis.

11.2 The cell-cycle Control System

Eucaryotic cells have evolved a complex network of regulatory proteins, known as the cell-cycle control system, that governs progression through the cell cycle. The core of this system is an ordered series of biochemical switches that control the main events of the cycle, including DNA replication and the segregation of the replicated chromosomes. In most cells, additionally layers of regulation enhance the fidelity of cell division and allow the control system to respond to various signals from both inside and outside the cell. Inside the cell, the control system monitors progression through the cell cycle and delays later events until earlier events have been completed. Preparations for the segregation of replicated chromosomes, for example, are not permitted until DNA replication is complete. The control system also monitors the conditions outside the cell. In a multicellular organism, the system is highly responsive to signals from other cells, stimulating cell division when more cells are needed and blocking it when they are not. The cell-cycle control system therefore has a central role in regulation cell numbers in the tissues of the body. When the system malfunctions, excessive cell division can results in cancer.

In principle, one can imagine that the most basic control system should possess the following features:

- A clock, or timer, that turns on each event at a specific time, thus providing a fixed amount of time for the completion of each event.
- A mechanism for initiating events in the correct order; entry into mitosis, for example, must always come after DNA replication.
- A mechanism to ensure that each event is triggered only once per cycle.
- Binary (on/off) switches that trigger events in a complete, irreversible fashion. For example, once initiated the DNA synthesis must be completed.
- Adaptability, so that the system's behavior can be modified to suit specific cell types or environmental conditions.
- Robustness: backup mechanisms to ensure that the cycle can work properly even when parts of the system malfunction.

- **Abnormal termination:** control mechanisms to ensure that the cycle results in a perfect copy of the mother cell. For instance, if there are unrecoverable errors in the duplicated genome the cell cycle has to stop definitively and the cell death (apoptosis) must be initiated.

Especially, the required robustness and the ability of an abnormal termination are antagonists in the optimization process of the cell cycle. A very robust design might be insusceptible to malfunctions but crucial errors, e.g. in the DNA duplication, remains without consequences in the further course. On the other hand, an undersized error tolerance leads to a very fragile and inflexible cell cycle.

An adjustable cell-cycle control system is an very important feature for the adaptation, e.g. on environmental changes, of the cell cycle. The control system of simple embryonic cells is based on a clock. The timer is unaffected by the events it regulates and will progress through the whole sequence of events even if one of those events has not been successfully completed. In contrast, the control system of most cell cycles is responsive to information received back from the processes it is controlling. Sensors, for example, detect the completion of DNA synthesis, and, if some malfunction prevents the successful completion, signals are sent to the control system to delay or stop progression to the next phase. These delays provide time for the machinery to be repaired and also prevent the disaster that might result if the cycle progressed prematurely to the next stage.

In most cells there are several points in the cell cycle, called checkpoints, at which the cell cycle can be arrested if previous events have not been completed or extracellular conditions are unfavorable. For example, the progression through G_1 and G_2 is delayed by braking mechanisms if the DNA in the chromosomes is damaged by radiation or chemicals. The delays provide time for the damaged DNA to be repaired.

Furthermore, the cell cycle can be regulated by extracellular signals from other cells at the checkpoints. These signals can either promote or inhibit cell proliferation.

Although most checkpoints are not essential for normal cell-cycle progression under ideal conditions, populations with checkpoint defects often accumulate mutations due to occasional malfunctions in DNA replication, DNA repair, or spindle assembly. Some of these mutations can promote the development of cancer.

11.2.1 Cell-Cycle Control System and Cyclically Activated Protein Kinase

At the heart of the cell-cycle control system is a family of protein kinases known as cyclin-dependent kinases (Cdk). The activity of these kinases rises and falls as the cell progress through the cycle. The oscillations lead directly to cyclical changes in the phosphorylation of intracellular proteins that initiate or regulate the major events of the cell cycle — DNA replication, mitosis, and cytokinesis. An increase in Cdk activity at the beginning of mitosis, for example, leads to increased phosphorylation of proteins that control chromosome condensation, nuclear envelope breakdown, and spindle assembly.

Cyclical changes in Cdk activity are controlled by a complex array of enzymes and other proteins. The most important of these Cdk regulators are proteins known as cyclins. Cdks are dependent on cyclins for their activity: unless they are tightly bound to a cyclin, they have no protein kinase activity. Cyclins were originally named as such because they undergo a cycle of synthesis and degradation in each cell cycle. Cdk levels, by contrast, are constant, at least in the simplest cell cycles. Cyclical changes in cyclin levels result in the cyclic assembly and activation of the cyclin-Cdk complexes. This activation in turn triggers cell-cycle events, as it is outlined in Figure 11.2.

There are four classes of cyclins, each defined by the stage of the cell cycle at which

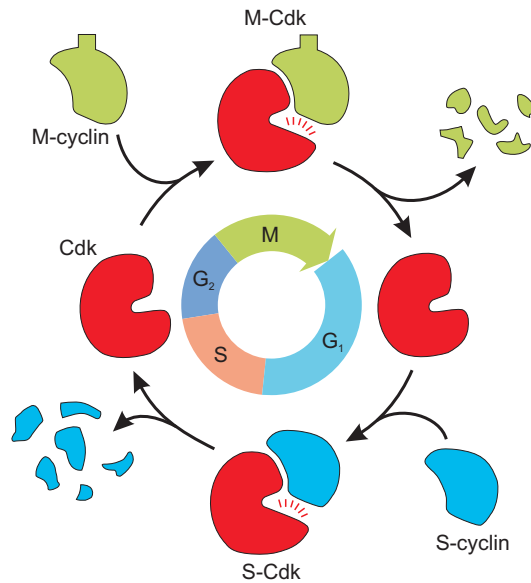


Fig. 11.2: Simplified view of the core of the cell-cycle control system [AJL⁺02]. For demonstration purposes, only the cyclins that act in S phase (S-cyclin (blue)) and M phase (M-cyclin (green)) are shown. The interaction between a single Cdk and a cyclin results in a cyclin-Cdk complex referred to as S-Cdk and M-Cdk, respectively.

they bind Cdks and function. Three of these classes are required in all eucaryotic cells:

1. G_1/S -cyclins bind Cdks at the end of G_1 and commit the cell to DNA replication.
2. S-cyclins bind Cdks during S-Phase and are required for the initiation of DNA replication.
3. M-cyclins promote the events of mitosis.

The fourth class, the G_1 -cyclins, helps promote passage through Start or the restriction point in late G_1 . Whereas in yeast a single Cdk protein binds all classes of cyclins and drive all cell-cycle events by changing cyclin partners at different stages of the cycle, in vertebrates cells there are four Cdks. Two interact with G_1 -cyclins, one with G_1/S - and S-cyclins, and one with M-cyclins.

The rise and the fall of cyclin concentrations is the primary determinant of Cdk activity during the cell cycle. Several additional mechanisms, however, are important for fine-tuning Cdk activity at specific stages in the cell cycle.

The activity of a cyclin-Cdk complex can be inhibited by phosphorylation at a pair of amino acids in the roof of the active site. Phosphorylation of these sites by a protein kinase known as Wee1 inhibits Cdk activity, while dephosphorylation of these sites by a phosphatase known as Cdc25 increase the activity, as shown in Figure 11.3[a]. This regulatory mechanism is particular important in the control of M-Cdk activity at the onset of mitosis.

A further regulation arises from the binding of Cdk inhibitor proteins (CKI), see Figure 11.3[b]. There are a variety of CKI proteins, and they are primarily employed in the control of G_1 and S phase.

11.2.2 Cell cycle and Cell growth

For proliferating cells to maintain a relatively constant size, the length of the cell cycle must match the time it takes the cell to double in size. If the cycle time is shorter

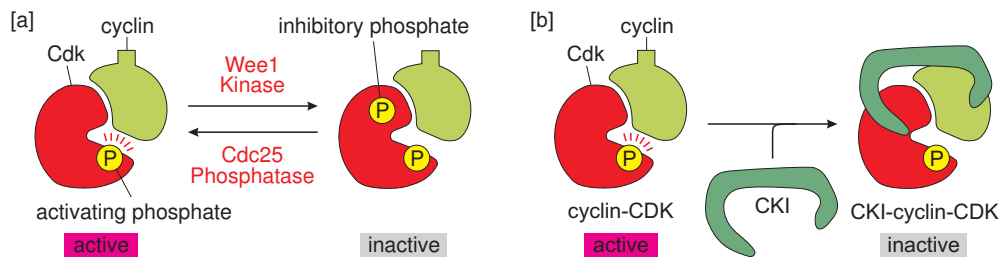


Fig. 11.3: Regulation of Cdk activity by inhibitory phosphorylation and a CKI [AJL⁺02]
 [a] Inhibitory phosphorylation: The active cyclin-CDK complex is deactivated when the kinase Wee1 phosphorylates two closely spaced sites above the active site. The phosphatase Cdc25 removes these phosphates and activates the complex.
 [b] Inhibition by CKI: The CKI binds to both the cyclin and Cdk in the complex, distorting the active site of Cdk. It also inserts into the ATP-binding site, further inhibiting the enzyme activity.

than this, the cells will get smaller with each division. The cells will get bigger with each division, if it is longer. Because cell growth depends on nutrients and growth signals in the environment, the length of the cell cycle has to be adjustable to varying environmental conditions. There must be a correlation between cell growth and cell cycle. However, it is not clear how proliferating cells coordinate their growth with the rate of cell-cycle progression to maintain their size.

There is evidence that cells coordinate their growth and cell-cycle progression by monitoring the total amount of a G₁ cyclin, which is synthesized in parallel with cell growth. But the mechanism, how the cell measures the total amount rather than its concentration, is still unclear.

Whereas cells of simple organisms grow and proliferate constitutively if nutrients are plentiful, animal cells generally grow and proliferate only when they are stimulated to do so by signals from other cells. The size at which an animal cell divides depends on, at least in part, on these extracellular signals, which can regulate cell growth and proliferation independently. Animal cells can also completely uncouple cell growth and division so as to grow without dividing or to divide without growing. The eggs of many animals, for example, grow to an extremely large size without dividing. After fertilization, this relation is reversed, and many rounds of division occur without growth.

Thus, cell division and cell growth are usually coordinated, but they can be regulated independently. Cell growth does not depend on cell-cycle progression. Neurons and muscle cells, for example, grow large after they are withdrawn permanently from the cell cycle.

11.3 Model systems to study the cell-cycle control system

The cell-cycle can be investigated experimentally on some 'simple' model systems. In studies of the genetics of the cell cycle a tiny, single-celled fungi, Yeast, is used. Two species are generally used in experiments. The fission yeast *Schizosaccharomyces pombe* is named after the African beer it used to produce. The budding yeast *Saccharomyces cerevisiae* is used by brewers, as well as by bakers. A microscopic view of both species of yeast is shown in Figure ???. These two species share a number of features that are extremely useful for genetic studies.

They reproduce almost as rapidly as bacteria and have a genome size less than 1% that of a mammal. They are amenable to molecular genetic manipulation, whereby genes can be deleted, replaced, or altered. Most importantly, they have the unusual ability to proliferate in a haploid state, in which only a single copy of each gene is present in the

cell. In such cells one avoids the complication of having a second copy of the gene in the cell and it is easy to isolate and study mutations that inactivate a gene. Many of these mutations cause cells to arrest at a specific point in the cell cycle, suggesting that the normal gene product is required to get the cell past this point.

While yeasts are ideal for studying the genetics of the cell cycle, the biochemistry of the cycle is most easily analyzed in the giant fertilized eggs of many animals, which carry large concentrations of the proteins needed for cell division. The egg of the frog *Xenopus*, for example, is over 1mm in diameter and carries about 10^5 times more cytoplasm than an average cell in the human body. Fertilization of the egg triggers an astonishingly rapid sequence of cell divisions, in which the single giant cell divides, without growing, to generate an embryo containing thousands of smaller cells. In this process, almost the only macromolecules synthesized are DNA — required to produce the thousands of new nuclei — and a small amount of proteins.

The cells in early embryos of *Xenopus*, as well as those of the clam *Spisula* and the fruit fly *Drosophila* are thus capable of exceedingly rapid division in the absence of either growth or many of the control mechanisms that operate in more complex cell cycles. These early embryonic cell cycles therefore reveal the working of the cell-cycle control system stripped down and simplified to the minimum needed to achieve the most fundamental requirements — the duplication of the genome and its segregation into two daughter cells. Another advantage is their large size. It is relatively easy to inject test substances into an egg to determine their effect on cell-cycle progression. It is also possible to prepare almost pure cytoplasm from *Xenopus* eggs and reconstitute many events of the cell cycle in a test tube.

11.4 Modeling the cell cycle of yeast

A often used model organism for studying the cell cycle, its components and its regulation is the single celled yeast. Because of its known gene sequence we can identify the encoding genes of proteins. With this knowledge knock-out experiments were made to identify the proteins participating on the cell cycle. Furthermore we obtain information of their function during the cycle, e.g. if a specific protein triggers the transition from G_2 to M phase. If we know the components and their function, we are able to build a mathematical model. The numerical simulation of this model provides us with information of the temporal operation of the cell cycle. This we demonstrate now using a model of the yeast cell cycle developed by Novak et al. [NPCT01]. But before we discuss the mathematical representation of the cell cycle engine in yeast we want to shortly introduce the molecular basis of the model.

11.4.1 Molecular components of the yeast cell cycle

Lower eukaryotes, like yeast, use only one essential Cdk subunit (generally called Cdk1). In yeast Cdk1 is often called Cdc2, in recognition of the gene (*cdc2*) that encodes this protein in fission yeast [Nur90]. In fission yeast complexes between Cdc2 and the B-type cyclin Cdc13 play the major roles in cell cycle regulation [FN95, Nas96a]. Deletion of the gene encoding this essential cyclin produces mutant cells that cannot enter mitosis [HFWN94].

DNA replication occurs once per cycle because Cdk activity not only triggers DNA replication but also inhibits replication of DNA [SN96]. To start DNA synthesis, Cdk activity causes properly licensed origins of replication to begin copying the chromosomes, and the same time it phosphorylates licensing factor molecules, making them more suscep-

tible to degradation. Disappearance of licensing factors from the nucleus prevent further rounds of replication. Cdk activity increases to higher level in late G_2 , thereby initiating M phase. As cells exit M phase, Cdk activity must be destroyed, to permit accumulation of licensing factors at replication origins in the next G_1 phase.

Destruction of Cdk activity as cells exit mitosis is the job of the anaphase promoting complex (APC). The APC is a large protein complex that attaches ubiquitin tags to target proteins, which are then rapidly degraded by proteasomes. The APC has two important functions at anaphase:

1. to initiate degradation of the cohesion proteins that hold two sister chromatids together, thereby initiating mitotic anaphase
2. to initiate degradation of B-type cyclins, thereby permitting cells to re-enter G_1 phase

To recognize the proper substrates for ubiquitination, the APC core requires specific “auxiliary” proteins. Slp1 targets the cohesion complex for disassembly, and both Slp1 and Ste9 present Cdc13 to the APC for ubiquitination [YON00, BSDdPM00]. Proper timing of these events is controlled by phosphorylation and dephosphorylation of Slp1 and Ste9.

MPF's enemies and friends

The activity of the Cdc2/Cdc13 complex (also called “M-phase promoting factor” MPF) is controlled by antagonistic interactions with its enemies. These enemies have negative effects on MPF, but MPF can down-regulate all of its enemies. Two of these enemies are active in G_1 phase, while a different group regulates the G_2/M transition.

The first G_1 enemy, Ste9 [YMO97, KMS98] targets Cdc13 to the APC core and promotes its degradation. On the other hand, phosphorylation of Ste9 by MPF inhibits its association with the APC core, rendering it inactive [YON00, BSDdPM00].

The other G_1 enemy of MPF is a stoichiometric inhibitor, called Rum1 [MN94], which can bind to Cdc2/Cdc13 complexes and inhibits their activity [CBN95, MCLM96]. On the other hand, a phosphorylation of Rum1 by MPF promotes its ubiquitination and rapid degradation [BMC98]. Hence, there is antagonism between MPF and Rum1, as well as between MPF and Ste9.

Because of these antagonistic relationships, MPF and its G_1 enemies cannot coexist. Either the enemies win and the cell is in G_1 phase corresponding to a low MPF activity or MPF wins and the cell is in S/ G_2/M phase of the cycle [NCNB⁺98]. The balance between MPF and its enemies is shifted by helper molecules in one direction or the other.

The Start transition ($G_1 \rightarrow S$) is supported by a “starter” kinase (SK), a group of Cdk/cyclin complexes. It helps the Cdc2/Cdc13 complex (MPF) to get the upper hand by phosphorylating Rum1 and Ste9. The helper molecule for the transition $M \rightarrow G_1$ is the Slp1(APC complex, which promotes the degradation of Cdc13 and activates Ste9. Slp1 can help the enemies because it is not inactivated by MPF phosphorylation, as is Ste9.

The duration of G_2 phase is regulated by different mechanism, namely enzymatic inhibition of MPF activity. The kinases Wee1 and Mik1 [RN87, LWB⁺91] can inactivate Cdc2. In return, MPF can also phosphorylate and inactivate them. So we have another case of mutual antagonism and alternative steady states: an S/ G_2 state and an M state.

The G_2/M transition is accelerated by a direct positive feedback loop. The inhibitory phosphate group of Cdc2 is removed by a specific phosphatase, called Cdc25 [MR92]. This phosphatase is phosphorylated by MPF, but the phosphorylated form is more active. Here, MPF helps its friend, Cdc25.

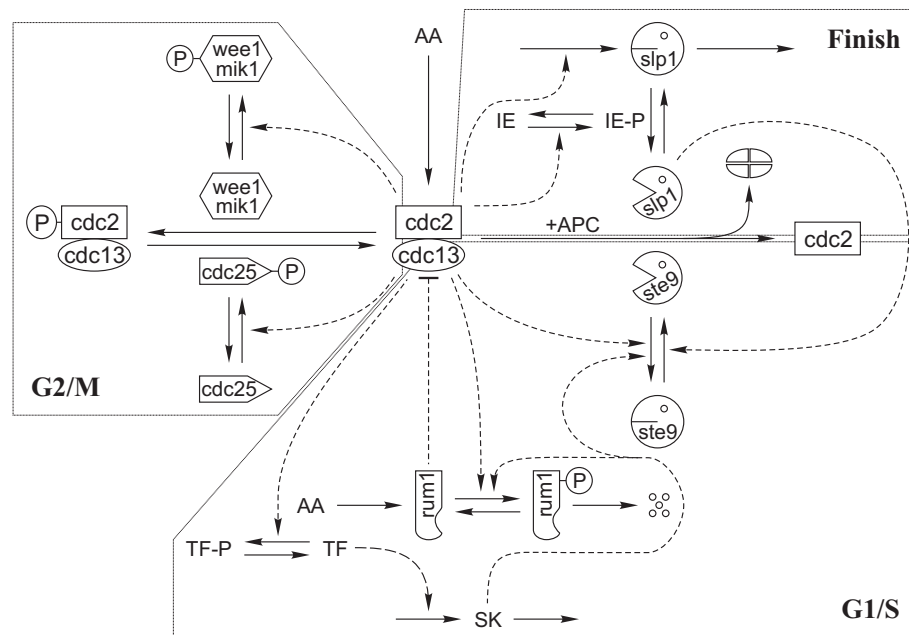


Fig. 11.4: The wiring diagram of the fission yeast cell cycle engine according to Novak et al. [NPCT01]. The core of the engine is the Cdc2/Cdc13 (MPF) complex, which is regulated by proteolysis of the Cdc13 component, phosphorylation of Cdc2 subunit, and stoichiometric inhibition of the complex. These processes are arranged according to the cell cycle transitions in which they are involved.

Feedback loops

For cells to proliferate, to make a repetitive sequence of properly controlled Start, G₂/M and Finish transitions, the helper molecules must be removed after they have done their jobs, because they are inhibitory for the next cell cycle. For instance, the starter kinase (SK) would inhibit the Finish transition, when the MPF enemies must come back. Therefore, MPF inhibits the synthesis of SK by phosphorylating its transcription factor (TF).

In analogy, Slp1 must disappear after Finish transition; otherwise, it would inhibit the next Start transition. The synthesis and activation of Slp1 depends on MPF creating a negative feedback loop [MR92]. It is essential that Slp1/APC complex is not directly activated by MPF, but rather through an intermediary enzyme (IE), which provides a time delay in the loop. This delay gives the chromosomes enough time to align before Slp1/APC breaks down their cohesion.

11.4.2 Surveillance mechanisms, Checkpoints

These helper molecules are regulated by surveillance mechanisms (also called checkpoints) [Nas96b]. Start is controlled by cell mass, Finish by the state of the cell's chromosomes, and the G₂/M transition is affected by both. The chromosome cycle, regulated by the cell cycle engine, must run in concert with overall cytoplasmic growth. Else the cells becoming hopeless small or enormously large. Without such a coordination mechanism, cells cannot be kept alive over the long term.

How cytoplasmic mass exerts its control over the cell cycle engine is not clear at present.

11.4.3 Mathematical model

After we identified the key proteins in the cell cycle of yeast and its functional relationships, we want to represent the cycle within a mathematical representation. From the graphical representation as wire diagram, Figure 11.4, we can establish a model of coupled ordinary differential equations (11.1)-(11.12)[NPCT01].

$$\frac{d[\text{Cdc13}_T]}{dt} = k_1 M - \left(k_2' + k_2'' [\text{Ste9}] + k_2''' [\text{Slp1}] \right) [\text{Cdc13}_T], \quad (11.1)$$

$$\begin{aligned} \frac{d[\text{preMPF}]}{dt} &= k_{\text{wee}} ([\text{Cdc13}_T] - [\text{preMPF}]) - k_{25} [\text{preMPF}] \\ &\quad - \left(k_2' + k_2'' [\text{Ste9}] + k_2''' [\text{Slp1}] \right) [\text{preMPF}], \end{aligned} \quad (11.2)$$

$$\begin{aligned} \frac{d[\text{Ste9}]}{dt} &= \left(k_3' + k_3'' [\text{Slp1}] \right) \frac{1 - [\text{Ste9}]}{J_3 + 1 - [\text{Ste9}]} \\ &\quad - \left(k_4' [\text{SK}] + k_4 [\text{MPF}] \right) \frac{[\text{Ste9}]}{J_4 + [\text{Ste9}]}, \end{aligned} \quad (11.3)$$

$$\frac{d[\text{Slp1}]}{dt} = k_5' + k_5'' \frac{[\text{MPF}]^4}{J_5^4 + [\text{MPF}]^4} - k_6 [\text{Slp1}_T], \quad (11.4)$$

$$\begin{aligned} \frac{d[\text{Slp1}^*]}{dt} &= k_7 [\text{IEP}] \frac{[\text{Slp1}_T] - [\text{Slp1}]}{J_7 + [\text{Slp1}_T] - [\text{Slp1}]} - k_8 \frac{[\text{Slp1}]}{J_8 + [\text{Slp1}]} \\ &\quad - k_6 [\text{Slp1}], \end{aligned} \quad (11.5)$$

$$\frac{d[\text{IEP}]}{dt} = k_9 [\text{MPF}] \frac{1 - [\text{IEP}]}{J_9 + 1 - [\text{IEP}]} - k_{10} \frac{[\text{IEP}]}{J_{10} + [\text{IEP}]}, \quad (11.6)$$

$$\frac{d[\text{Rum1}_T]}{dt} = k_{11} - \left(k_{12} + k_{12}' [\text{SK}] + k_{12}'' [\text{MPF}] \right) [\text{Rum1}_T], \quad (11.7)$$

$$\frac{d[\text{SK}]}{dt} = k_{13} [\text{TF}] - k_{14} [\text{SK}], \quad (11.8)$$

$$\frac{d[M]}{dt} = \mu M, \quad (11.9)$$

$$[\text{Trimer}] = \frac{2 [\text{Cdc13}_T] [\text{Rum1}_T]}{\Sigma + \sqrt{\Sigma^2 - 4 [\text{Cdc13}_T] [\text{Rum1}_T]}}, \quad (11.10)$$

$$[\text{MPF}] = \frac{([\text{Cdc13}_T] - [\text{preMPF}]) ([\text{Cdc13}_T] - [\text{Trimer}])}{[\text{Cdc13}_T]}, \quad (11.11)$$

$$[\text{TF}] = G \left(k_{15} M, k_{16}' + k_{16}'' [\text{MPF}], J_{15}, J_{16} \right) \quad (11.12)$$

where

$$k_{\text{wee}} = k_{\text{wee}}' + \left(k_{\text{wee}}'' - k_{\text{wee}}' \right) G \left(V_{\text{awee}}, V_{\text{iwee}} [\text{MPF}], J_{\text{awee}}, J_{\text{iwee}} \right), \quad (11.13)$$

$$k_{25} = k_{25}' + \left(k_{25}'' - k_{25}' \right) G \left(V_{\text{a25}} [\text{MPF}], V_{\text{i25}}, J_{\text{a25}}, J_{\text{i25}} \right), \quad (11.14)$$

$$\Sigma = [\text{Cdc13}_T] + [\text{Rum1}_T] + K_{\text{diss}}. \quad (11.15)$$

We use the Michaelis-Menten scheme for describing the activation and deactivation reactions of the participating proteins, whereby we assume the formed intermediate complexes in a quasi-steady state. As result we obtain Michaelis-Menten like expressions within our system of coupled differential equations, see for instance Eq. (11.3) for the protein ste9. A further simplification arises from the assumption of a quasi-stationary state for the following proteins

1. the transcription factor (TF) for the synthesis of the starter kinase (SK), Eq. (11.12),

2. the trimeric complexes of Cdc13/Cdc2 and Rum1 (Trimer), Eq. (11.10),
3. the enzymatic inhibitors Wee1, Eq. (11.13), and cdc25, Eq. (11.14).

Then their concentrations can be expressed in terms of the Goldbeter-Koshland function [GKJ81, TCN03]

$$G(a, b, c, d) = \frac{2ad}{b - a + bc + ad + \sqrt{(b - a + bc + ad)^2 - 4ad(b - a)}} \quad (11.16)$$

describing the steady state solution of an activation/deactivation cycle (see section 9.3).

Furthermore we used conservation laws for the proteins Wee1, cdc25, IE, and ste9 to reduce the model structure. We assume further, that Rum1 binds to unphosphorylated (MPF) and the phosphorylated (preMPF) form of the Cdc13/Cdc2 complex.

For the protein slp1 we have to distinguish between its deactivated form slp1 and its activated form slp1*, Eqs. (11.4) and (11.5).

As shown in the wiring diagram, Figure 11.4, the proteins rum1 and slp1 are also produced outside from the cell cycle engine. We assume a constant external production rate represented by zero-order rate constants k'_5 and k_{11} . The cell mass M takes in a special position in the presented model. Because of a lack of knowledge of the feedback regulation and to simplify the model we treat the cell growth as an independent dynamic variable, as shown in the corresponding differential equation (11.9). At the end of mitosis the cell mass is divided by two triggered by a decreasing MPF level through 0.1. In the model we assume an instantaneous cell division into two daughter cells of the same mass and volume, although daughter cells do not physically separate from another until 15-20 min after mitosis [NPCT01]. Due to this boundary condition we simulate a regulatory relationship between the cell cycle and the cell growth. Notice, that during the division the concentrations of the proteins remain unchanged¹. But the new cell mass influence the production rate of the protein Cdc13, see Eq. (11.1), which affects in succession all other components of the cell cycle engine. A possible resting phase G_0 or differences of growth behavior in G_1 , G_2 , M, and S phase are not included in this model [NPCT01]. Also, it does not describe the influence of changing external conditions to the cell cycle. We assume constant and ideal external conditions resulting in a exponential growth. The exponential function is characterized by the constant growth rate μ .

11.5 Numerical simulations

After we introduced the mathematical model of the cell cycle of yeast and its biological background, we now simulate the cell cycle. We compare the cell cycle of wild-type cells with mutant cells and investigate the differences in the dynamic change of key proteins [NPCT01]. Due to the mathematical complexity of the cell cycle model we can only numerically solve the system of coupled differential equations (11.1)-(11.14). The used parameter values of wild-type fission yeast are given in Table 11.1.

11.5.1 Wild-type cells

The temporal changes of the concentrations of components of the cell cycle of yeast are shown in Figure 11.5. We separate this graphical representation into three parts:

¹Concentrations are intensive quantities, whereas the mass is an extensive one. A doubling of the considered system doubles extensive quantities, but the intensive keep constant. In particular, the ratio of two extensive quantities, for instance number of molecules and volume, results in an intensive quantity, for our example in a concentration

Cdc13 synthesis and degradation:

$$k_1 = 0.03, k_2' = 0.03, k_2'' = 1, k_2''' = 0.1$$

Ste9 activation and deactivation:

$$k_3' = 1, k_3'' = 10, J_3 = 0.01, k_4' = 2, k_4'' = 35, J_4 = 0.01$$

Slp1 synthesis, degradation, activation, and deactivation:

$$k_5' = 0.005, k_5'' = 0.3, J_5 = 0.3, k_6 = 0.1$$

$$k_7 = 1, J_7 = 0.001, k_8 = 0.25, J_8 = 0.001$$

IE activation and deactivation:

$$k_9 = 0.1, J_9 = 0.01, k_{10} = 0.04, J_{10} = 0.01$$

Rum1 synthesis, degradation, and inhibition:

$$k_{11} = 0.1, k_{12} = 0.01, k_{12}' = 1, k_{12}'' = 3, K_{\text{diss}} = 0.001$$

SK synthesis and degradation:

$$k_{13} = k_{14} = 0.1$$

TF activation and deactivation:

$$k_{15} = 1.5, J_{15} = 0.01, k_{16}' = 1, k_{16}'' = 2, J_{16} = 0.01$$

Wee1 activation and deactivation:

$$V_{\text{awee}} = 0.25, J_{\text{awee}} = 0.01, V_{\text{iwee}} = 1, J_{\text{iwee}} = 0.01$$

Cdc25 activation and deactivation:

$$V_{a25} = 1, J_{a25} = 0.01, V_{i25} = 0.25, J_{i25} = 0.01$$

Rate of phosphorylation and dephosphorylation

$$k_{\text{wee}}' = 0.15, k_{\text{wee}}'' = 1.3, k_{25}' = 0.05, k_{25}'' = 5$$

Growth rate

$$\mu = 0.005$$

Table 11.1: Parameter values for the cell cycle model of Novak et al. [NPCT01], Eqs. (11.1)-(11.14). All constants have units min^{-1} , except the dimensionless Michaelis constants J_i and the dissociation constant of the trimer K_{diss} .

- [a] temporal evolution of the cell mass
- [b] temporal evolution MPF, preMPF and Cdc13_T
- [c] temporal evolution of the enemies of MPF (Ste9, Slp1, Rum1) and of the starter kinase (SK)

Additionally we draw the phases of the cell cycle on top of the figure and extend the separations with dashed lines into all three plots. The events determining the different stages are discussed in the further course.

The evolution of the cell mass is shown in Figure 11.5[a]. As mentioned before, it follows an exponential monotone increasing law until the Mitosis is finished. During the time between two divisions a common yeast cell double their mass. According to Figure 11.5[a] it takes about 140 minutes to complete the cell cycle of wild type yeast cells. If we further assume an ideal environment there are no fluctuations in the maximal cell mass and the cycle time. If the Mitosis finishes the cell mass is halved and the arisen daughter cells will growth with the same rate.

The duration of Mitosis is determined by the MPF level plotted in part [b] of Figure 11.5. In order to enter the Mitosis the MPF level has to change from low to high. A requirement for this is a previous production of Cdc13 which can combine with *cdc2* to the inactive form of the mitosis promoting factor, preMPF. As shown in the figure all the three components of yeast cell cycle among three levels. After a short time delay, Slp1/APC is activated by the high MPF activity in Mitosis, shown in part [c] of the figure, initiating the rapid degradation of Cdc13. As a consequence, MPF activity drops. Simultaneously, the increase of Slp1/APC activates Ste9. Because MPF inhibits the Ste9 activation, the low level of MPF has a further positive effect on the production of active Ste9. The period of high Ste9 level determines the G₁ phase. The low MPF concentration leads to an inactivation of Slp1 and relieves the inhibition on the transcription factor TF responsible for the synthesis of cyclin subunit of the SK. The SK level increases and inhibits the MPF enemies Ste9 and Rum1. Actually, G₁ is so short that Rum1 does not have time to come up [CBN95]. As soon as Ste9 gets inactivated, the Cdc13 level rises and the cell passes the G₁/S transition. However, SK does not inactivate the third enemy, Wee1, which phosphorylates Cdc2/Cdc13. The phosphorylated form has reduced protein-kinase activity, which seems to be enough to initiate S phase but not mitosis. When the cell reaches a critical size, the positive feedbacks for G₂/M transition turn on. Abrupt activation of MPF by Cdc25 drives the cell into mitosis. During this process virtually all Cdc13 is bound in MPF, the precursor form preMPF is completely converted into MPF. Hence, we come back to our starting point and a new cell cycle begins.

The SG₂ transition was arbitrarily chosen. A theoretical model of this transition is not included in the presented cell cycle model.

11.6 Mutations

Mutations play an important role in the development of organisms. Due to the increasing knowledge of proteins and their encoding genes and the arising new technologies to alter the genome of an organism in a well defined manner, it is possible to investigate the role of individual proteins and their influence on the cell cycle in experiments.

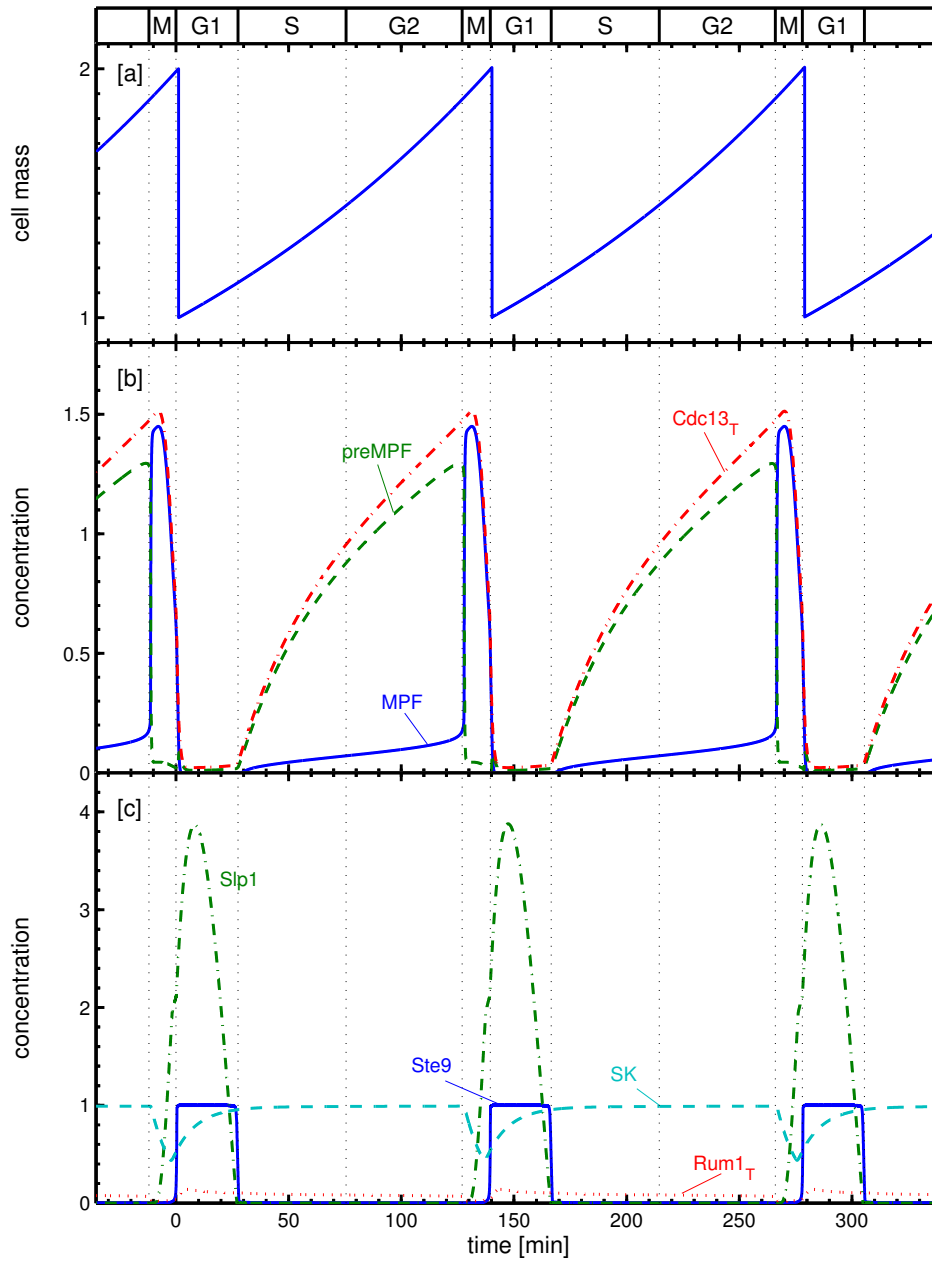


Fig. 11.5: Numerical solution of the cell cycle model of yeast of Novak et al. [NPCT01], Eqs. (11.1)-(11.14), for wild-type cells. The corresponding parameters values are given in Table 11.1. Two complete cell cycles are shown. The figure is divided into three parts:

- [a] The temporal evolution of the cell mass.
- [b] The temporal evolution of the Mitosis-promoting factor (MPF), its inactive form (preMPF) and the cyclin Cdc13, one of its main proteins.
- [c] The temporal evolution of enemies and friends of MPF.

11.6.1 *wee1*⁻ - mutants

The inhibition of MPF by Wee1 is an crucial step in the cell cycle of wild type cells. In order to inactivate Wee1, cells must grow to a critical size, which necessitates an extended S+G₂ phase (about 100 minutes). Hence, a change in the Wee1 activity strongly affects the duration of the cell cycle and the cell mass. The cell cycle of these *wee1*⁻ cells is investigated in Figure 11.6. Again, we use the same representation as described in the previous section. The activity of Wee1 is expressed in the rate constants for its inactivation k'_{wee} and for its activation k''_{wee} . A smaller activation constant of $k''_{wee} = 0.3$ (for wild-type cells see Table 11.1) shifts the concentration of activated Wee1 to a decreased concentration [NPCT01]. As consequence a smaller critical mass is needed to pass the S transition. As in Figure 11.6[a] is shown, the cell mass is smaller than the mass of a wild-type cell. This is the defining characteristic of 'wee mutants [Nur75].

The G₂/M transition is not size controlled. Consequently, the S+G₂ phase is much shorter (≈ 45 min) than in wild type. To adjust their cycle time to the mass doubling time (140 min), *wee1*⁻ cells have an extended G₁ phase, stabilized by up-regulation of Rum1 and Ste9 (Figure 11.6[c]).

Because of the extend M phase, ≈ 25 min, which is about the twice of the M phase of wild-type cells, and the stronger increase of MPF during the S+G₂ phase, the starter kinase shows a pronounced oscillatory behavior. The suppression of TF activation during the M phase leads to an almost complete degradation of SK. If the MPF concentration decreases at the end of mitosis the SK level rises until the comparable small cell size prefers the degradation again. First with increasing cell mass the SK production gets the upper hand until the negative feedback of MPF gets the dominant part in Eq. (11.12). The Rum1_T, on the other hand, benefits from the low level of starter kinase during the G₁ phase. Its antagonist SK is down-regulated and its inhibitor MPF can be neglected within this period. Hence, it can reach a pronounced maximum in G₁. If the MPF concentration and the SK concentration increase, Rum1_T is degraded to a similar low level as in wild-type cells, see Figure 11.5.

11.6.2 *wee1*⁻ *cdc25* Δ mutants

Wee1 and Cdc25 are the major tyrosine-modifying enzymes in fission yeast. If Wee1 (the inhibitor) is missing, then cells should not need Cdc25 (the activator). Indeed, *wee1*⁻ *cdc25* Δ double-mutant cells are viable, but they exhibit abnormal progression through the cell cycle.

For the simulation of these yeast mutants we use the rate constants $k'_{wee} = 0.15$ and $k''_{wee} = 0.3$ (for *wee1*⁻ mutation) and $k'_{25} = k''_{25} = 0.02$ (for *cdc25* Δ mutation) [NPCT01]. In comparison to wild-type cells the mutated rate constants are reduced (see Table 11.1).

The numerical results are shown in Figure 11.7. The cells alternate between short cycles (about 110min) and long cycles (about 170min). As observed experimentally [SNM99], the mutant cells always divide at size larger than wild type. Cells born at larger size are committed to the shorter cell cycle, and smaller newborns have longer cycles. Due to the larger cell size the mutants have a much shorter G₁ phase than wild-type cells.

The reason for the quantized cell cycle is the weak positive feedback in mitotic control due to lacking of activated Wee1 and Cdc25 [NPCT01]. The mitosis promoting factor is not activated abruptly, Figure 11.7[c], when cells are supposed to enter the M phase. As in Figures 11.5 and 11.6 we use the abrupt change in MPF behavior as the transition G₂/M. Then the *wee1*⁻ *cdc25* Δ double-mutant cells have an extended M phase compared to the previously considered fission yeast cells. Furthermore, only a small part of preMPF

is activated during this phase. During the M phase of wild-type cells and *wee1⁻* mutants the conversion to MPF was almost complete.

Because MPF rise is sluggish, it may not turn on the fully the exit-from-mitosis pathway. The negative feedback loop generates a rebound in MPF activity leading to a degradation and inactivation of the involved proteins. After a delay a 'second try' is started when the cell size is larger which leads to a successful mitosis.

Because of the comparable small MPF concentration in the cell cycle of *wee1⁻cdc25 Δ* double-mutants the inhibitory effect on the transcription factor TF is weak and the concentration of the starter kinase SK is almost constant and on a high level over the whole cycle. Consequently, the MPF enemy Rum1 is suppressed.

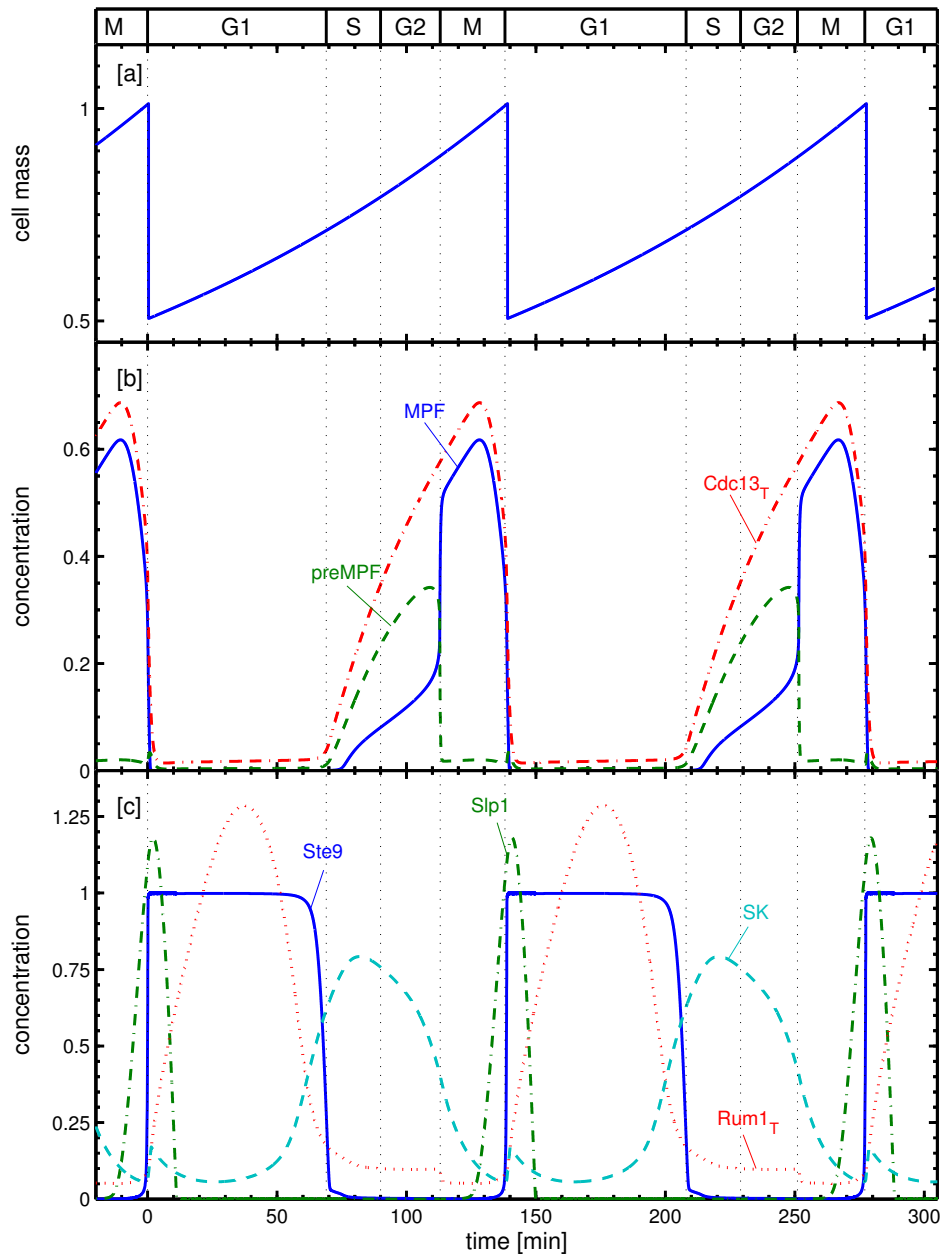


Fig. 11.6: Numerical solution of the cell cycle model of yeast of Novak et al. [NPCT01], Eqs. (11.1)-(11.14), for $wee1^-$ mutants. The corresponding parameters values are given in Table 11.1, except $k''_{wee} = 0.3$. Two complete cell cycles are shown. The figure is divided into three parts:

- [a] The temporal evolution of the cell mass.
- [b] The temporal evolution of the Mitosis-promoting factor (MPF), its inactive form (preMPF) and the cyclin Cdc13, one of its main proteins.
- [c] The temporal evolution of enemies and friends of MPF.

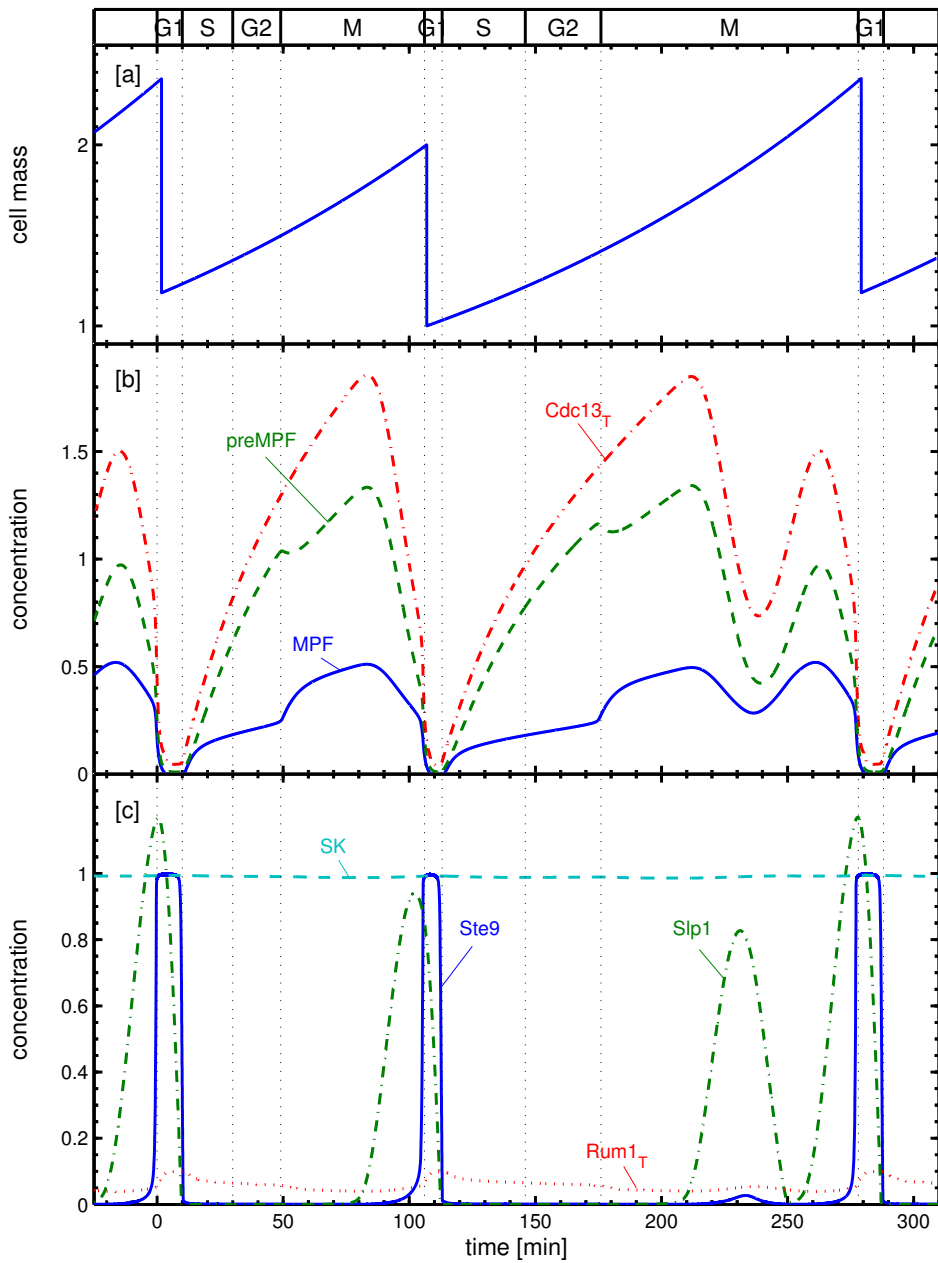


Fig. 11.7: Numerical solution of the cell cycle model of yeast of Novak et al. [NPCT01], Eqs. (11.1)-(11.14), for $wee1^- cdc25\Delta$ mutants. The corresponding parameters values are given in Table 11.1, except $k''_{wee} = 0.3$ and $k'_{25} = k''_{25} = 0.02$. Two complete cell cycles are shown. The figure is divided into three parts:

- [a] The temporal evolution of the cell mass.
- [b] The temporal evolution of the Mitosis-promoting factor (MPF), its inactive form (preMPF) and the cyclin Cdc13, one of its main proteins.
- [c] The temporal evolution of enemies and friends of MPF.

12 Metabolic Control Analysis

In the previous chapters we dealt with autonomous nonlinear systems and investigated, how the structure of a set of ordinary differential equations determines the properties and the structure of the corresponding solutions. There we focused on

- Structural properties of biochemical networks;
- Dependencies of the dynamic and transient behavior on initial states;
- Dependencies of the dynamic and transient behavior on rate coefficients.

Several methods for the analysis of nonlinear systems were introduced in Section 6.5 and, from a more system theoretic point of view, in subsequent sections. Simple examples were given that demonstrate how changes in parameters, initial conditions, and structure can change systems properties.

In this chapter, we introduce a general approach to investigate and quantify such changes in a systematic way. We are interested in the control exerted by the various reactions over the fluxes or rates and over the concentrations, especially at steady state. The main questions, which to be addressed during the course are:

- How much does the steady state change if one varies the concentration of participating species?
- How much does the steady state change if one varies the rate coefficients?
- How is the reaction network controlled and regulated?

This approach was developed for investigations of metabolic networks thus, it is called Metabolic Control Analysis (MCA). Nevertheless, it should be mentioned here that the mathematical framework is not restricted to metabolism. It can be applied to any problem that considers the transformations of elements or, more general, which considers the fluxes of some elements, e.g. economics, traffic and crowd control. The mathematical formalism, introduced in the next sections, provides us with a method to describe the control of metabolic systems. For this reason, we will use the usual terminology of MCA. Species at the beginning of a reaction network are referred to as substrates, intermediary components as metabolites and at the end of the metabolic pathway are products. Thereby, we follow the convention that the direction of the reaction goes from the substrates to the products [HS96]. This implies that there is a continuous conversion of substrate into products. The amount of converted material per unit of time is described as flux, for example as number of particles per second. In analogy to the definition of fluxes in Irreversible Thermodynamics, the fluxes in the Metabolic Control Analysis have a unique direction. However, a generalization for arbitrary direction is possible [Fel97]. Nevertheless, before we introduce the mathematical framework of MCA we will introduce some fundamental quantities in the next section.

Our notation follows the convention proposed by several leading authors of the MCA community [WGW84, BCBG⁺85]. Even if this notation differs from the one used in the original papers of Kacser and Burns [KB73] and Heinrich and Rapoport [HR73, HR74, HR75], it becomes more and more accepted in recent publications.

12.1 Fundamentals

12.1.1 Control and Regulation

Metabolic control analysis is a means to study the control of metabolic processes. It ascribes a specific meaning to the term control. The terms control and regulation are often used interchangeably, but here they have a distinct meaning. *Control* is the ability to make changes in the behavior of a system, whereas *regulation* is the maintenance (or the lack of it) of constant internal conditions when a system is exposed to external disturbances.

In this course, we consider control as the ability to influence something. For example, if an enzyme controls the flux of a pathway, a change in the concentration of that enzyme changes the flux of the pathway.

Regulation describes whether any control (by a controller), is actually affected or not. In metabolism, regulatory mechanisms often exist to maintain a degree of homeostasis. As well as maintaining constant conditions, regulatory mechanisms can also improve the performance of control.

When describing control we must be careful to define what is being 'controlled' and what is doing the 'controlling'. A controller is a component that controls some other component(s) of the system. In contrast to this, a component is controlled, if it is determined by another component. As we see later, these definitions are closely related to parameters and variables. However, it is important to note that, the definition of what is 'controlled' (a variable) and what is the controller (a parameter) is arbitrary and depends on the system under investigation.

12.1.2 Parameters and Variables

Parameters are quantities that can be changed independently, and they typically (but not always) remain constant during the evolution of the system toward its steady state. Examples include kinetic coefficients, enzyme concentrations, and external inhibitors, but also physical quantities like temperature and pH-value. Parameters can be classified as internal and external [HCB91]. Internal parameters have values that, although not invariant in the absolute sense, are fixed characteristics of a particular system, e.g. the thermodynamic and kinetic coefficients of the enzyme reactions. External parameters are those quantities through which the environment can affect the system. Most realistically external parameters are concentrations, e.g. of enzymes, terminal metabolites, and external effectors, such as inhibitors, activators, hormones or growth factors; their levels are determined by the environment of the system and do not depend in any way on what happens in the system. Consequently, changes in the environment can only be communicated to the system through changes of external parameters.

Variables are quantities determined by the system, and they are time-dependent before reaching their steady state. The most common variables are metabolite concentrations. Other important variables are the rates of reactions. These are functions of concentration variables and kinetic parameters. In most cases, the reaction rates will be described by enzyme kinetics. However, they may also represent other types of reactions or transport processes. A change in systems parameters will lead to a change of systems variables.

12.2 Control Coefficients

Control coefficients are fundamental to Metabolic Control Analysis. In general, a *control coefficient* is a relative measure of how much a perturbation of a systems parameter affects

a systems variable. In other words, a control coefficient describes how sensitive a system variable is towards changes of a system parameter. Hence, an approach that determines control coefficients is called a sensitivity analysis.

The above definition of control coefficients contains an implicit assumption – the measurement of control coefficients requires a reference point. In metabolic systems, a natural choice for the reference point is the steady state. Hence, unless noted otherwise, we measure the relative changes in the considered system against the steady state determined by the system parameters. Usually, this means for metabolism that we investigate the modification of the steady state by parameter variation.

However, before we can investigate metabolic pathways in more detail, we have to identify a proper set of variables and parameters that determines the considered metabolic system uniquely. As an example we use the linear pathway represented in Figure 12.1. Its steady state can be characterized by different quantities. First, we can consider the concentrations of substrates, intermediates, and products as in the investigation of transient phenomena in the previous sections. Second, the network can be characterized by the overall rate at steady state. From the definition of an open system and the convention that the net conversion of the investigated pathway is directed from the substrates to the products [HS96]¹, it follows that the flux is greater than zero, except at equilibrium which is uninteresting for living systems. The rate describing the amount of substrates that are converted into products per unit of time is usually called flux of the pathway in MCA.

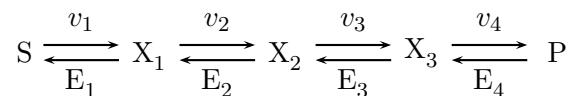


Fig. 12.1: Linear four step pathway. The substrate S is converted into the product P via four subsequent steps with the intermediates X₁, X₂, and X₃. Each reaction is facilitated by a specific enzyme E_i and is characterized by the reaction rate v_i.

At steady state, each enzymatic reaction of the metabolic pathway is operating at the same rate. No reaction is running slower than any other. Hence, we cannot deduce a controlling step from the fluxes. However this does not mean that not only few enzymes or even a single enzyme may control the steady state of the considered pathway. This leads us to the principle of rate-determining step introduced in Section 3.6.3. The relation of MCA and this principle shall be discussed later in Section 12.3.

Flux and concentrations are determined by kinetic parameters, such as rate coefficients, Michaelis constants or Hill coefficients, depending on the chosen level of mathematical description. In our example, at steady state, all fluxes v_i are equal. Nevertheless, we identify a flux for each intermediary step, because these fluxes are determined by the parameters of the corresponding step and can, therefore, be of different importance to the control of the considered pathway. Furthermore, the concentration of enzyme facilitating a step of the metabolic pathway is treated as a parameter. Finally, we assume that the substrate and product concentrations remain constant and can be treated as system parameters.

The fluxes v_i of the pathway and the concentrations of the intermediates are variables that are determined by the system parameters. The changes in the system variables are described by various control coefficients, which we introduce now.

¹Nevertheless, a generalization to arbitrary directions is discussed in [Fel97]

12.2.1 Flux Control Coefficient

First, we want to investigate, how a change of parameters changes the overall flux J of a pathway. Originally, Kacser and Burns [KB73] investigated how the concentration $[E_i]$ of the enzymes acting as catalysts of the individual reaction i affects the metabolic flux. Correspondingly, they defined a *flux control coefficient* as follows

flux control coefficient

$$C = \frac{\Delta J}{\Delta[E_i]}, \tag{12.1}$$

which relates the fractional change in the steady-state flux ΔJ to the fractional change in the total enzyme concentration $\Delta[E_i]$. At the present representation, the control coefficient depends on the units and the actual experimental conditions. Therefore, this absolute control coefficient will be different for other experimental setups. A comparison of different experiments and situations is not possible. This problem is resolved by normalization with the former steady-state flux J and the corresponding enzyme concentration $[E_i]$

$$C_{E_i} = \frac{[E_i]}{J} \frac{\Delta J}{\Delta[E_i]}. \tag{12.2}$$

We consider only infinitesimal changes in the enzyme concentration, and obtain the flux control coefficient [KB73]

$$C_{E_i}^J = \frac{[E_i]}{J} \frac{\partial J}{\partial [E_i]}, \tag{12.3}$$

where the subscript E_i denotes the considered enzyme of the metabolic pathway. The superscript J indicates the control coefficient as a flux control coefficient. For further simplification of (12.3), we use the mathematical properties of the derivative of the logarithmic function

$$\frac{\partial \ln f(x)}{\partial x} = \frac{f(x)}{x} \frac{\partial f(x)}{\partial x}$$

which leads to the logarithmic flux control coefficient

$$C_{E_i}^J = \frac{\partial \ln J}{\partial \ln [E_i]}. \tag{12.4}$$

As represented in Figure 12.2 the control coefficient is the tangent of the curve flux J against the enzyme concentration $[E_i]$ at the considered steady state.

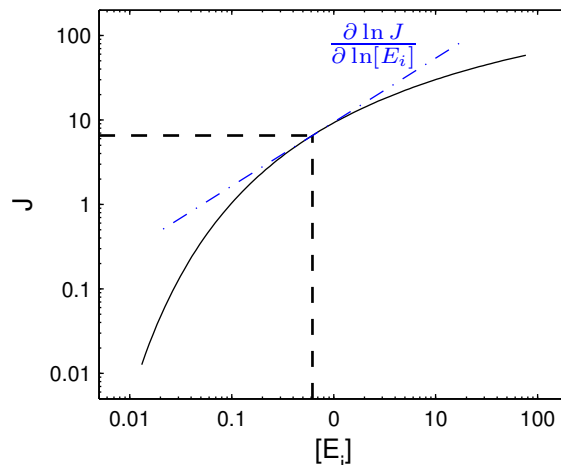


Fig. 12.2: Logarithmic representation of flux J of a metabolic pathway as function of the enzyme concentration $[E_i]$ (solid line). The flux control coefficient (12.4) defines the tangent (dashed-dotted line) of the curve at a considered steady state (dashed lines).

Taking into account that other parameters than enzyme concentrations may affect reaction rates v_i and, therefore, steady-state fluxes, Heinrich and Rapoport [HR73, HR74] defined the flux control coefficient in a more general way as

$$C_{v_i}^J = \frac{v_i}{J} \frac{\partial J / \partial p_k}{\partial v_i / \partial p_k} = \frac{v_i}{J} \frac{\partial J}{\partial v_i}, \quad (12.5a)$$

$$= \frac{\partial \ln J / \partial \ln p_k}{\partial \ln v_i / \partial \ln p_k} = \frac{\partial \ln J}{\partial \ln v_i}, \quad (12.5b)$$

where p_k is the parameter which is changing. Because, mathematically, the fluxes J cannot be directly expressed as function of the rates v_i , the simpler forms at the right are not strictly correct and have to be regarded as an abbreviated notation. But it is acceptable as long as it is remembered that there is always an implied external parameter p_k even if it is not shown explicitly.

From Eqs. (12.4) and (12.5) it follows that the flux control coefficient is state dependent. It varies according to the steady state (i.e. the value of J) at which it is measured. It should be noted, that the analysis assumes that only one parameter, e.g. the concentration of enzyme E_i , is changed and all other parameters remain constant. Hence, each parameter in the considered pathway is characterized by one flux control coefficient. The overall flux of a metabolic pathway is determined by a set of control coefficients, where the number of coefficients is equal to the number of parameters. Usually, the metabolic pathway is investigated in terms of its enzymes. Then, we have one control coefficient for each enzyme.

12.2.2 Concentration Control Coefficient

Apart from the flux, we can also use the concentrations of intermediary metabolites X_i to characterize the considered pathway. In analogy to the previous section substituting the system variable flux J for the concentration of the metabolites X_i , we obtain the concentration control coefficient [HR73, HR74] as

$$C_{v_i}^C = \frac{v_i}{[X_i]} \frac{\partial [X_i]}{\partial v_i}, \quad (12.6a)$$

$$= \frac{\partial \ln [X_i]}{\partial \ln v_i}, \quad (12.6b)$$

where we have to keep in mind that the rates depend on the external parameters p_k ². The concentration control coefficient measures the degree of control of an external parameter, as the total enzyme concentration, to the concentration of metabolites. Again, the concentration control coefficient is calculated at steady state.

12.2.3 The Summation Theorem

In the previous two sections we introduced control coefficients measuring the effect that a change of an enzyme concentration has on the flux or the metabolite concentration. From the definitions of flux control coefficient (Eqs. (12.3) and (12.4)) and concentration control coefficient (Eqs. (12.6a) and (12.6b)) it appears that the control coefficients for different enzymes are independent from each other. However, one can show within the framework of MCA that this is not the case. This dependency conclusion is one of most important results of this approach to biochemical networks. It concerns the *summation of all the control coefficients* of a pathway. By various procedures [KB73, HR75, Gie88, Red88] it

²See the comment to flux control coefficients below Eq. (12.5).

can be demonstrated that for a given flux the sum of its flux control coefficients of all steps in the system is constant.

The simplest way in which this property can be derived is by considering a simultaneous small relative increase Δv in all reaction rates of a metabolic system. Because for each metabolite the relative rates of its production increase exactly by the same amount Δv as the relative rates of its consumption, the metabolite concentrations remain unchanged. The flux of the system increases exactly by Δv . This means that the flux is a homogeneous function of first order and the metabolite concentrations homogeneous functions of zeroth order³. The summation theorems follow from this by applying a corollary of the Euler's homogeneous function theorem⁴ for homogeneous functions [Gie88]. Due to the normalization, Euler's theorem simplifies to the *summation theorem for fluxes* [KB73]

flux summation theorem

$$\sum_i C_{v_i}^J = 1 \quad (12.7)$$

and to the *summation theorem for concentrations* [HR75]

concentration summation theorem

$$\sum_i C_{v_i}^{X_j} = 0. \quad (12.8)$$

Note, that the sum over the control coefficients is equal to the order of homogeneity of the corresponding system's variable.

It follows clearly from both summation theorems that the control coefficients are not independent from each other. If, for example, one coefficient increases, one or more of the other coefficients have to decrease.

If and only if a flux control coefficient $C_j^J \approx 1$ and all other $C_{i \neq j} \approx 0$ then the j -th enzymatic reaction is rate-determining. In general, the control of the pathway may be distributed over more than one steps. This is summarized in the famous dictum – “Control can be shared!”

But the flux summation theorem does not restrict the flux control coefficients to the interval $[0; 1]$. Some coefficients may well be negative and some coefficients of the considered pathway can exceed unity.

For a concentration control coefficient $C_{E_i}^{X_i} \approx 0$, the enzymatic reaction is limited by other factors than the enzyme concentration, and a change of the enzyme concentration does not change the intermediate concentration.

Number of summation theorems

The number of summation theorems is determined by the structure of the investigated metabolic pathway. A summation theorem exists for each systems variable, i.e. flux. A short list is given in the following table

³A function $f(\alpha x_1, \alpha x_2, \dots)$ is homogeneous of order n if the equation

$$f(\alpha x_1, \alpha x_2, \dots, \alpha x_k) = \alpha^n f(x_1, x_2, \dots, x_k)$$

holds for all $\alpha, x_i \in \mathbb{R}$.

⁴Euler's homogeneous function theorem

$$n \cdot f(x_1, x_2, \dots, x_k) = \sum_{i=1}^k \frac{\partial f}{\partial x_i} x_i$$

relates a homogeneous function $f(x_i)$ to its partial derivatives, its coordinates x_i (i.e. enzyme concentrations), and its order of homogeneity n .

structure	# of fluxes	# of summation theorems
linear pathway	1	1
two branches	2	2
⋮

The number of summation theorems will increase with the complexity of metabolic pathways.

12.3 Metabolic Control Analysis and Rate-Determining Step

In previous sections we already mentioned the principle of a rate-determining step some times. In the metabolic control analysis this principle leads to a controversial discussion, see for instance [Fel97]. The origin is a traditional view of metabolic control which assumes that there has to be a rate-limiting step in a metabolic pathway. On the other hand, from the summation theorem (12.7) follows that the control can be distributed over several enzymes. This is also supported by some experiment, i.e. see the paper of Groen et al. [GWW⁺82].

But before we continue our discussion, we will cite two different definitions of the principle of rate-determining step. The first one is from Blackman (1905) [Bla05]:

When a process is conditioned as to its rapidity by a number of separate factors, the rate of the process is limited by the pace of the slowest factor.

and the second definition is a combination from the famous book of Atkins, “Atkins’ Physical Chemistry” [AdP02] and from the book of Schwetlik [SDP⁺73] about Chemical Kinetics :

If there is in a consecutive reaction a step much slower than all subsequent steps, see Figure 12.3[a] and [b], this step is rate-determining. The rate-determining step controls the rate of all subsequent steps. However, the rate-determining step is not just the slowest step: it must be slow and be a crucial gateway for the formation of products. If a faster reaction can also lead to products, then the slowest step is irrelevant because the slow reaction can then be side-stepped, as in the example in Figure 12.3[c].

In comparison to the modern definition of the rate-determining step in Chemical Kinetics, the definition of Blackman is vague. On one hand, the kind of process or biochemical reaction is not specified. On the other hand, the slowest factor is a very imprecise specification. As example, one factor may have a value of $f_1 = 10$ and another the value $f_2 = 10.5$. Then, both factors are similar. According to Blackman, the factor f_1 is rate-determining. But, as investigations of consecutive reactions show, see Section 3.6.3, similar reaction rates do not fulfill the principle of a rate-determining step.

In general, the principle of rate-determining step is defined only for consecutive reactions or processes⁵. These kind of processes have at least one irreversible step [SDP⁺73]. Then the principle is a direct consequence from the principle of cause and effect. The cause triggers the effect.

Whereas this condition is clearly fulfilled in industrial production processes, where it leads to the so-called bottleneck problem, it is rare in chemical processes. A further well-known example for a rate-determining step can be often observed at construction areas of

⁵For a more detailed discussion of complex reactions see Section 3.6 and for consecutive reactions Section 3.6.3.

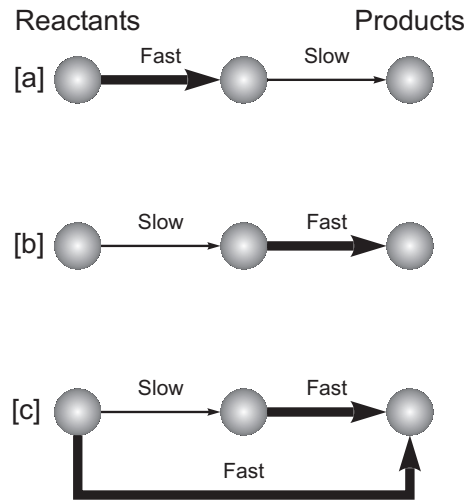


Fig. 12.3: Rate-determining steps in consecutive reactions. Heavy arrows represent fast steps and light arrows represent slow steps. [a] The second step is rate-determining; [b] the first step is rate determining; [c] although one step is slow, it is not rate-determining because there is a fast route that circumvents it.

highways. At the construction area the lanes are reduced, for instance from three lanes to a single lane. If the flux of cars exceed the passage limit of the bottleneck it determines the rate of all subsequent events. As consequence, there will be an increasing traffic jam before the construction area. In metabolic pathways such bottlenecks are usually the result of some limitations, for example in food, light, or necessary additional reactants. Only, in these limiting cases a metabolic pathway may have a rate-determining step. However, this is not the usual case. Interestingly, Blackman [Bla05] mentioned this already in 1905.

However, the modern definition characterizes only the necessary conditions for a rate-determining step, it does not predict the existence of such a step. An *a priori* assumption of a rate-determining step is a misinterpretation of the principle.

In the investigation of the dynamics of a biochemical system, the principle of a rate-determining step plays a major role, although it is often not obvious. Whereas in consecutive reactions the rate-determining step controls the overall reaction rate, the situation in more complex biochemical reaction networks is different. Here, there is usually not a global rate-determining step, similar to metabolic networks. But, if we model the network, we decompose it into a set of smaller subnetworks. Some of them may consist of consecutive reactions. For these reactions, local rate-determining steps can exist that enable a simplification of the mathematical representation. As example, the Michaelis-Menten equation 3.150, discussed in Section 3.7.1, is often used as a template to describe the conversion, the (de)activation or degradation of a protein. Here, the dissociation of the enzyme-substrate complex into product and enzyme has to be rate-determining, otherwise the quasi-steady-state assumption (see Section 3.6.3) is not applicable and the Michaelis-Menten equation is not an appropriate description of the dynamics. Nevertheless, the identification of a rate-determining step is crucial because it is dependent on the current situation. During the time course or through a change in the environmental conditions, the local rate-determining step may change or disappear.

12.4 Elasticity Coefficients

In Section 12.2 we investigated how a change in enzyme concentrations changes the flux and the metabolite concentrations. These are properties of the whole intact metabolic

pathway, referred to as global properties. In contrast, traditionally, biochemists have studied enzymes in isolation, as represented in Fig. 12.4. For this reason, we now investigate, how MCA can be used to describe the effect of changes in substrate, product or enzyme concentration on the individual reaction rates v_i . Furthermore, the rate v_i

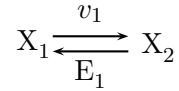


Fig. 12.4: Simple example of an isolated intermediary step of a metabolic pathway. The metabolite X_1 is converted to X_2 with the rate v_1 . The reaction is facilitated by the enzyme E_1 .

elasticity coefficient depends on external parameters, such as kinetic coefficients. The *elasticity coefficients* measure the changes to the individual reaction as a consequence of a change of these parameters. According to Heinrich and Schuster [HS96], we distinguish between two kinds of parameters: a) ε -elasticities describing the effects of changes in the concentrations of participating species, and b) π -elasticities describing the effects of varying kinetic coefficients to the rate of change. The elasticities are local or systemic coefficients because they consider only an isolated reaction.

12.4.1 Elasticity Coefficients for Metabolites

The rate of change of the conversion of a metabolite X_i to a metabolite X_{i+1} depends on the concentrations of the metabolites, of the enzymes facilitating the reaction, and of inhibitors or further influencing proteins. The elasticity coefficients are defined as [KB73, BCBG⁺85]:

$$\varepsilon_{v_i}^{X_j} = \frac{[X_i]}{v_j} \frac{\partial v_j}{\partial [X_i]}, \quad (12.9)$$

where the subscript i denotes the i -th metabolite X_i and the superscript j the rate v_j of the j -th reaction. The derivation of this expression is equivalent to the control coefficients represented in Section 12.2.1. The ε -elasticity quantifies how the reaction rate is modulated by the metabolite X_i . A positive coefficient corresponds to an increase of the rate and a negative coefficient to a decrease. A value of zero can have two different causes. First, the metabolite X_i is no part of the considered reaction and, therefore, the rate remains unchanged. Second, the metabolite X_i participating on the reaction but does not change the rate. This is common for reactions that follow a Michaelis-Menten regime. If the enzyme is saturated, a further increase of the metabolite does no longer increase the reaction rate. For a more detailed discussion of different reaction mechanisms and their related elasticities, see Section 12.9.

12.4.2 Elasticity Coefficients for Parameters

As mentioned earlier, the rates v_i also depend on external parameters. Kinetic parameters, such as rate coefficients and Michaelis-Menten constants, are functions of environmental conditions, e.g. temperature, pH-value or water balance. They vary over time and experiments. Hence, the rates v_i change if kinetic parameters change. Additionally, the reaction rates also depend on some model parameters. Depending on the used framework, these could be for example Hill parameters or kinetic orders. Because of the errors of experimental data, these model parameters can be determined only with a certain limited accuracy. Within the confidence interval the values may change. Therefore, it is important to investigate how these changes affect the rate of the considered reaction.

A relative measure of how a variation in external parameters modulate the rate is the π -elasticity. It is defined as [KSA90, SK90]

$$\pi_{p_k}^{v_i} = \frac{p_k}{v_i} \frac{\partial v_i}{\partial p_k}, \quad (12.10)$$

where p_k is the changing external parameter and v_i the considered rate. If the π -elasticity is positive, the rate v_i increases if the parameter p_k increases. For negative values it decreases if the parameter decreases. If the π -elasticity is zero, the parameter p_k does not affect the rate v_i . Either no step of the considered reaction is determined by this parameter or the rate is fully determined by other factors, e.g. a rate-determining step which is not affected by p_k .

12.4.3 The Connectivity Theorem

One of the major objectives of MCA is to reveal how the properties of a whole system, i.e. its control coefficients, can be explained in terms of the properties of its components, i.e. in terms of the elasticities. This is especially important, because the control coefficients are not usually easy to measure. But, many data are available for isolated (in vivo) enzymatic reactions. Thus, one is trying to assess the systemic behavior that is of fundamental interest but difficult to measure in terms of quantities that can be measured very easily. It turns out that for any pathway there are exact relationships between the control coefficients and the ϵ -elasticities. The first *connectivity theorem* relates the flux control coefficients and the ϵ -elasticities in a sum over products between both coefficients. It was found by Kacser and Burns as [KB73]:

$$\sum_i C_i^J \epsilon_{X_j}^i = 0. \quad (12.11)$$

As opposed to the summation theorem (12.7), the flux connectivity theorem depends on the kinetic expressions through the ϵ -elasticities.

An analogous theorem was derived for metabolite concentrations [WC84]:

$$\sum_i C_i^{X_j} \epsilon_{X_j}^i = -1. \quad (12.12)$$

It connects the concentration control coefficients $C_i^{X_j}$ to the elasticities of metabolites that participate in the i -th step of the pathway.

With summation theorems (12.7) for fluxes and (12.8) for concentrations and the above connectivity theorems, we have now a tool to investigate metabolic pathways from their global and their local properties. Moreover, we are able to determine control coefficients from elasticity coefficients using these relations. In the next section, we demonstrate how one can use summation and connectivity theorems on the example of a linear pathway. Later on, we generalize the treatment for arbitrary structures introducing the matrix representation of metabolic control analysis.

12.5 Using the Summation and the Connectivity Theorem

As mentioned earlier, the connectivity theorems (12.11) and (12.12) relate the control coefficients and the elasticities. For linear pathways, this fact can be used to express the control coefficients in terms of elasticity coefficients. As an example, we investigate a three step pathway which converts the substrate S into the product P via the intermediary metabolites X_1 and X_2 , see Fig. 12.5. We restrict the analysis to the flux control

connectivity theorem

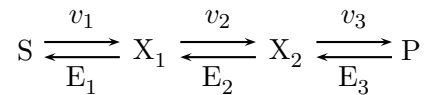


Fig. 12.5: Linear three step pathway converting substrate S into product P. v_i are the reaction rates of the conversion reactions, E_i the catalyzing enzymes and X_i intermediary metabolites.

coefficients, but an analogous treatment is possible for concentrations control coefficients.

Because of the linear nature of the considered pathway, there is only one flux and, therefore, we have only one summation theorem. Nevertheless, due to the two intermediates, we have to take into account two connectivity theorems, one for the metabolite X_1 and a second for the metabolite X_2 . The corresponding theorems are:

$$\text{Summation theorem (12.7):} \quad C_1^J + C_2^J + C_3^J = 1 \quad (12.13a)$$

$$\text{Connectivity theorem (12.11) for } X_1: \quad C_1^J \varepsilon_{X_1}^1 + C_2^J \varepsilon_{X_1}^2 = 0 \quad (12.13b)$$

$$\text{Connectivity theorem (12.11) for } X_2: \quad C_2^J \varepsilon_{X_2}^2 + C_3^J \varepsilon_{X_2}^3 = 0 \quad (12.13c)$$

where we can rearrange both connectivity relations with respect to the flux control coefficients:

$$C_1^J = -C_2^J \frac{\varepsilon_{X_1}^2}{\varepsilon_{X_1}^1}$$

and

$$C_3^J = -C_2^J \frac{\varepsilon_{X_2}^2}{\varepsilon_{X_2}^3}.$$

Substitution into the summation theorem (12.13a) leads to the equation

$$-C_2^J \frac{\varepsilon_{X_1}^2}{\varepsilon_{X_1}^1} + C_2^J - C_2^J \frac{\varepsilon_{X_2}^2}{\varepsilon_{X_2}^3} = 1,$$

holding only on flux control coefficient C_2^J and all elasticities. A transformation with respect to the flux control coefficient leads to the final result

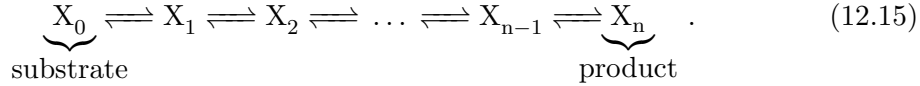
$$C_2^J = \frac{\varepsilon_{X_1}^1 \varepsilon_{X_2}^3}{\varepsilon_{X_1}^1 \varepsilon_{X_2}^3 - \varepsilon_{X_1}^2 \varepsilon_{X_2}^3 - \varepsilon_{X_1}^1 \varepsilon_{X_2}^2} \quad (12.14)$$

for the second control coefficient. It combines global properties (C_2^J) with local properties described by the elasticities ε_i^j . These local properties can be measured by in vitro enzyme kinetic experiments and can be applied around the considered steady state. Nevertheless, a support with in vivo data is strongly recommended to exclude differences for in vitro and in vivo experiments.

Similar expressions can be derived for the remaining flux control coefficients C_1^J and C_3^J . Note that this method is applicable only for some simple pedagogical examples. For real metabolic pathways, with many intermediary metabolites and branching points, it becomes intractable very fast. Moreover, it is not accessible for systematic numerical calculation. Hence, we want to generalize the method in the next section within a matrix approach that resolves the mentioned disadvantages.

12.6 Matrix Methods in Metabolic Control Analysis

As shown in the previous Section 12.5, using a simple linear pathway 12.5 as example, the combination of summation and connectivity theorem enables the determination of global flux control coefficients through the local elasticity coefficients. Also, if we consider all the theorems for an investigated metabolic pathway, we obtain a system of n^2 coupled equations. In this system, n is the number of all participating metabolites, where the substrate is the zeroth component and the product the n -th component, as in the generalized linear pathway



It is clear that with increasing complexity of metabolic pathways the treatment introduced in the last section becomes intractable. Hence, we use a more convenient and elegant method to represent the metabolic pathway in terms of the Metabolic Control Analysis. The matrix approach was originally introduced by [Red88] and developed further e.g. by [SSF87, SF89, HCBR93, HS96, Hof01].

Before we introduce the matrix approach in more detail, we summarize the theorems of MCA, that play the central role in this approach. First, there are the summation theorems

$$\sum_i C_i^J = 1 \quad \text{Flux Control Coefficients} \quad (12.16a)$$

$$\sum_i C_i^{X_j} = 0 \quad \text{Concentration Control Coefficients} \quad (12.16b)$$

for the flux and the concentration control coefficients. The connectivity theorems relate control coefficients to elasticity coefficients. For the combination of the flux related coefficients, they hold:

$$\sum_i C_i^J \epsilon_j^i = 0. \quad (12.17)$$

The connectivity theorem for concentrations (12.12) has to be generalized in order to distinguish between perturbed and unperturbed metabolites [WC84]. The theorem for the perturbed metabolite X_j was introduced in Section 12.4.3 as

$$\sum_i C_i^{X_j} \epsilon_{X_j}^i = -1.$$

However, we have to consider also the case where the metabolite X_j is not perturbed. Then, the connectivity theorem becomes:

$$\sum_i C_i^{X_j} \epsilon_{X_k}^i = 0.$$

Summarizing both equations, we obtain the generalized version of the connectivity theorem [WC84]

$$\sum_i C_i^{X_j} \epsilon_{X_k}^i = -\delta_{jk}, \quad (12.18)$$

where δ_{ij} is the Kronecker-Delta [AS72]

$$\delta_{jk} = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases} .$$

These information we now combine into in some matrices. This allows to use matrix algebra to derive control coefficients in terms of elasticity coefficients. Because of the

introductory nature of this chapter we follow here a more practical guide [SSF87, CB95a] instead of the theoretical approach based on structural properties of the pathway in [Red88, HCB93, HS96, Hof01].

Before we illustrate the practical use of the matrix approach we discuss the general treatment of metabolic pathways within this approach. The information about the metabolic pathway are represented in three quadratic matrices of order n . In the so-called E-matrix we collect elasticities, where the rows contains information about the fluxes J_i and the metabolites X_i . The rates are arranged over the columns of the matrix. The C-Matrix contains the control coefficients of the considered system. The arrangement with respect to the fluxes, metabolites and rates is now transposed in comparison to the E-matrix. Matrix multiplication of both matrices results in the M-matrix. In fact, this resulting matrix consists of a combination of the theorems which we summarized above. Finally, the metabolic pathway under investigation is described by the matrix equation

$$\begin{bmatrix} \varepsilon_1^1 & \varepsilon_1^2 & \cdots & \varepsilon_1^n \\ \varepsilon_2^1 & \varepsilon_2^2 & \cdots & \varepsilon_2^n \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_n^1 & \varepsilon_n^2 & \cdots & \varepsilon_n^n \end{bmatrix} \begin{bmatrix} C_1^1 & C_1^2 & \cdots & C_1^n \\ C_2^1 & C_2^2 & \cdots & C_2^n \\ \vdots & \vdots & \ddots & \vdots \\ C_n^1 & C_n^2 & \cdots & C_n^n \end{bmatrix} = \begin{bmatrix} M_1^1 & M_1^2 & \cdots & M_1^n \\ M_2^1 & M_2^2 & \cdots & M_2^n \\ \vdots & \vdots & \ddots & \vdots \\ M_n^1 & M_n^2 & \cdots & M_n^n \end{bmatrix} \quad (12.19)$$

or in its short representation

$$\mathbf{E} \cdot \mathbf{C} = \mathbf{M}. \quad (12.20)$$

This compact representation of the metabolic control analysis has some advantages over the classical treatment of the previous section: I.) The matrix representation makes it better amenable to numerical and symbolic approaches with suitable software packages or programming languages and to analytical investigation. II.) The matrices can be constructed directly from stoichiometric data as is shown in [Red88, HS96]. Hence, this approach can be connected to graphical tools.

Whereas the M-matrix is determined by the summation (12.16a) and connectivity theorems (12.18), and the E-matrix by local properties of the pathway, the C-matrix and its elements are unknown. Therefore, we are now seeking an expression for \mathbf{C} in terms of the other matrices. To this end, we calculate the inverse matrix of \mathbf{E} which is defined by the relation

$$\mathbf{E} \cdot \mathbf{E}^{-1} = \mathbf{1}.$$

The multiplication of a matrix with its inverse results in the identity matrix. The inverse matrix we multiply from left to matrix equation (12.20)

$$\mathbf{C} = \mathbf{E}^{-1} \cdot \mathbf{M} \quad (12.21)$$

and obtain the matrix of control coefficients as the product of the inverse of the E-matrix and the M-matrix⁶.

Before we demonstrate the practical use of the matrix approach with two examples, we have to give some general rules for the construction of the matrix. In the definition (12.19) we introduced only the elements of this matrix. Now, we relate them to the properties of the pathway. Here, we distinguish between two kinds of elements, flux related and elasticity related. If an element is part of a row describing a flux through the metabolic pathway, then its value is simply 1. In order to determine the elasticity related elements we have to consider the current metabolite X_i and the current rate v_j . If the metabolite X_i participate

⁶Note the order of multiplication of matrices! Generally, the matrix multiplication is not commutative. This means

$$\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}.$$

Hence, we distinguish between multiplication from left and multiplication from right.

on the j -th reaction, then the value of element ε_i^j follows from the definition of the elasticity coefficient (12.9), else it is zero.

After defining the structure and the values of the E-matrix, we can now calculate the M-matrix. The element-wise calculation of the elements M_i^j reproduces exactly the theorems (12.16a) and (12.18). Thus, we can represent the derivation of the matrix in a simple schematic way [CB95a].

$\begin{array}{c} E_1 \ E_2 \ E_3 \\ J \\ X_1 \\ X_2 \end{array} \cdot \begin{array}{c} J \ X_1 \ X_2 \\ J \\ X_1 \\ X_2 \end{array} = \begin{array}{c} v_1 \\ v_2 \\ v_3 \end{array}$	$\sum_{i=1}^n C_i^J = 1$	Flux control coefficient summation theorem
$\begin{array}{c} J \\ X_1 \\ X_2 \end{array} \cdot \begin{array}{c} J \ X_1 \ X_2 \\ J \\ X_1 \\ X_2 \end{array} = \begin{array}{c} v_1 \\ v_2 \\ v_3 \end{array}$	$\sum_{i=1}^n C_i^{X_j} = 0$	Concentration control coefficient summation theorem
$\begin{array}{c} J \\ X_1 \\ X_2 \end{array} \cdot \begin{array}{c} J \ X_1 \ X_2 \\ J \\ X_1 \\ X_2 \end{array} = \begin{array}{c} v_1 \\ v_2 \\ v_3 \end{array}$	$\sum_{i=1}^n C_i^J \varepsilon_{X_j}^i = 0$	Flux connectivity theorem
$\begin{array}{c} J \\ X_1 \\ X_2 \end{array} \cdot \begin{array}{c} J \ X_1 \ X_2 \\ J \\ X_1 \\ X_2 \end{array} = \begin{array}{c} v_1 \\ v_2 \\ v_3 \end{array}$	$\sum_{i=1}^n C_i^{X_j} \varepsilon_j^i = -1$	Concentration connectivity theorem (homogeneous)
$\begin{array}{c} J \\ X_1 \\ X_2 \end{array} \cdot \begin{array}{c} J \ X_1 \ X_2 \\ J \\ X_1 \\ X_2 \end{array} = \begin{array}{c} v_1 \\ v_2 \\ v_3 \end{array}$	$\sum_{i=1}^n C_i^{X_j} \varepsilon_k^i = 0$	Concentration connectivity theorem (heterogeneous)

Fig. 12.6: Schematic representation of the element-wise calculation of the M-matrix [CB95b]. As one can see, the matrix equation (12.20) contains all information about the investigated metabolic pathway which we derived in the previous sections. In general, for a system of n enzymes there are n^2 such relations which fall into five classes expressed by Eqs. (12.16 - 12.18).

Note, that in principle the arrangement of the elements is arbitrary, but most authors prefer the convenient rule that one starts with fluxes in the first rows followed by the metabolites. Also, some authors use the matrix equation

$$\mathbf{C} \cdot \mathbf{E} = \mathbf{M}$$

instead of (12.20). Due to the permuted order of the matrix multiplication one has to use the transposed versions of the E- and C-matrix which was introduced in (12.19). Furthermore, the multiplication with the inverse E-matrix have to carried from right.

In the next subsections, we apply the matrix method to the example of an unbranched and a branched pathway.

12.6.1 Control Coefficients in an unbranched pathway

The introduced matrix approach is used to investigate an unbranched pathway. We consider again a linear metabolic pathway with two intermediary metabolites (Figure 12.7). The corresponding matrix equation is

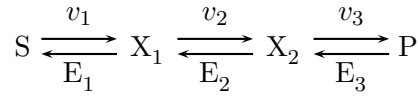


Fig. 12.7: Three step unbranched pathway. The v_i are the reaction rates of the conversion reactions, E_i are the catalyzing enzymes and X_i are intermediary metabolites.

$$\begin{array}{c} E_1 \quad E_2 \quad E_3 \\ J_1 \\ X_1 \\ X_2 \end{array} \begin{bmatrix} 1 & 1 & 1 \\ \varepsilon_1^1 & \varepsilon_1^2 & 0 \\ 0 & \varepsilon_2^2 & \varepsilon_2^3 \end{bmatrix} \begin{array}{c} J_1 \\ X_1 \\ X_2 \end{array} \begin{bmatrix} C_1^J & C_1^{X_1} & C_1^{X_2} \\ C_2^J & C_2^{X_1} & C_2^{X_2} \\ C_3^J & C_3^{X_1} & C_3^{X_2} \end{bmatrix} \begin{array}{c} v_1 \\ v_2 \\ v_3 \end{array} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \quad (12.22)$$

Due to the linear character of the pathway, there is only one flux which is determined by the three rates v_i . Furthermore three enzymes E_i affect these rates. Hence, the matrix equation consists of matrices of the order 3×3 . As noted in Eq. (12.22), the first row of the E-matrix is related to flux J and is, therefore, a unity matrix. The other two rows describe the elasticities of the reactions involved in the linear pathway. The elements ε_1^3 and ε_2^3 are zero because the enzyme E_3 does not affect the metabolite X_1 and the enzyme E_1 does not affect the metabolite X_3 . We conclude that in general a metabolite of a linear pathway has only two non-vanishing elasticities, namely one elasticity for the producing reaction and the second one for the degrading reaction. The control coefficients for the metabolic pathway are collected in the C-matrix. The elements of the first column are the flux control coefficients. The remaining columns consist of the concentration control coefficients with respect to the metabolites X_1 and X_2 . The M-matrix can be constructed according to the scheme represented in Figure 12.6. The first column of the M-matrix is given by the flux summation theorem (12.16a) and the second and third element by the connectivity theorem between flux control coefficients and elasticities. All non-diagonal elements of the next columns vanish as a consequence of the concentration summation theorem (12.16b) and the connectivity theorem (12.18) for the heterogenous case. The diagonal elements are determined by the control coefficients and the elasticities for the same metabolite (homogenous). Hence, these matrix elements have the value -1 . Note also the diagonal structure of the M-matrix. This is typical for linear pathways. Due to their special stoichiometric form, all non-diagonal elements have to equal zero.

After we constructed all three matrices, we now calculate the inverse E-matrix

$$E^{-1} = \frac{1}{\varepsilon_1^1 \varepsilon_2^2 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^2 \varepsilon_2^3} \begin{bmatrix} \varepsilon_1^2 \varepsilon_2^3 & \varepsilon_2^2 - \varepsilon_2^3 & -\varepsilon_1^2 \\ -\varepsilon_1^1 \varepsilon_2^3 & \varepsilon_2^3 & \varepsilon_1^1 \\ \varepsilon_1^1 \varepsilon_2^2 & \varepsilon_2^2 & \varepsilon_1^2 - \varepsilon_1^1 \end{bmatrix} \quad (12.23)$$

which we multiply from left with matrix equation (12.22). As result we obtain the matrix of control coefficients in terms of a new matrix

$$\begin{bmatrix} C_1^J & C_1^{X_1} & C_1^{X_2} \\ C_2^J & C_2^{X_1} & C_2^{X_2} \\ C_3^J & C_3^{X_1} & C_3^{X_2} \end{bmatrix} = \frac{1}{\varepsilon_1^1 \varepsilon_2^2 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^2 \varepsilon_2^3} \begin{bmatrix} \varepsilon_1^2 \varepsilon_2^3 & \varepsilon_2^2 - \varepsilon_2^3 & -\varepsilon_1^2 \\ -\varepsilon_1^1 \varepsilon_2^3 & \varepsilon_2^3 & \varepsilon_1^1 \\ \varepsilon_1^1 \varepsilon_2^2 & -\varepsilon_2^2 & \varepsilon_1^2 - \varepsilon_1^1 \end{bmatrix} \quad (12.24)$$

consisting of combinations of the elasticity coefficients. If analytical rate laws are known, the control coefficients can be obtained simply by calculating the partial derivatives of the definitions of the elasticity coefficient (12.9). Otherwise, they have to be measured experimentally.

12.6.2 Control Coefficients in a branched pathway

As a further example, we consider a branched pathway. The substrate S is converted via the intermediate X_1 into the products P_1 and P_2 (Figure 12.8). Because of the branching point

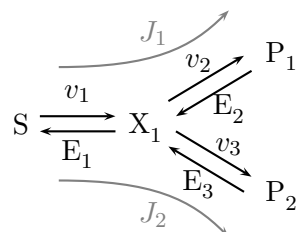


Fig. 12.8: An example of a branched pathway. The net rate of each step is denoted as v_i , whereas E_i is the facilitating enzyme. Note, due to the branching point, the flux is splitting into two parts, J_1 and J_2 .

we now have to consider two fluxes J_1 and J_2 . An important property of this branching point follows from the mass conservation in chemical reactions. The flux through the reaction v_1 has to be equal to the sum of the fluxes through the reactions v_2 and v_3 . This fact is a chemical equivalent to Kirchhoff's law for electrical circuits and is expressed mathematically as:

$$\sum_{i=1}^n J_i = 0. \quad (12.25)$$

Furthermore, we follow the usual convention that ingoing fluxes are positive and outgoing fluxes are negative. This law (12.25) can be used to search for missing branches or to estimate the crosstalk between different pathways. If one knows, e.g. all ingoing fluxes, and their sum is not equal to the sum of all considered outgoing fluxes, it is probable that there are further outgoing fluxes which one has to consider. On the other hand, one can use the normalized outgoing fluxes to estimate the crosstalk or the contribution of branches to the whole investigated system, whereby the normalization factor is the sum over all ingoing fluxes.

However, the matrix equation of the branched pathway of Figure 12.8 is:

$$\begin{array}{c} E_1 \quad E_2 \quad E_3 \\ J_1 \quad J_2 \quad X_1 \\ \left[\begin{array}{ccc} 1 & 1 & 0 \\ 1 & 0 & 1 \\ \varepsilon_1^{X_1} & \varepsilon_2^{X_1} & \varepsilon_3^{X_1} \end{array} \right] \left[\begin{array}{ccc} C_1^{J_1} & C_1^{J_2} & C_1^{X_1} \\ C_2^{J_1} & C_2^{J_2} & C_2^{X_1} \\ C_3^{J_1} & C_3^{J_2} & C_3^{X_1} \end{array} \right] \begin{array}{c} v_1 \\ v_2 \\ v_3 \end{array} = \left[\begin{array}{ccc} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -1 \end{array} \right], \end{array} \quad (12.26)$$

with consideration of two fluxes and the metabolite. In contrast to the E-matrix for an unbranched pathway, the flux related rows are not unity matrices. The element in the matrix corresponding to J_1 and E_3 has to be zero, because the reaction catalyzed by the enzyme E_3 have no contribution to the flux J_1 . In analogy, the element $\varepsilon_2^{X_1}$ has to be zero, too. The metabolite X_1 participates on all reactions, and thus it has three elasticities. The construction of the M-matrix follows again the scheme 12.6. The inverse of E-Matrix is then

$$E^{-1} = \frac{1}{\varepsilon_1^{X_1} - \varepsilon_2^{X_1} - \varepsilon_3^{X_1}} \left[\begin{array}{ccc} -\varepsilon_2^{X_1} & -\varepsilon_3^{X_1} & 1 \\ \varepsilon_1^{X_1} - \varepsilon_3^{X_1} & \varepsilon_3^{X_1} & -1 \\ \varepsilon_2^{X_1} & \varepsilon_1^{X_1} - \varepsilon_2^{X_1} & -1 \end{array} \right] \quad (12.27)$$

which leads finally to a C-Matrix of

$$\left[\begin{array}{ccc} C_1^{J_1} & C_1^{J_2} & C_1^{X_1} \\ C_2^{J_1} & C_2^{J_2} & C_2^{X_1} \\ C_3^{J_1} & C_3^{J_2} & C_3^{X_1} \end{array} \right] = -\frac{1}{\varepsilon_1^{X_1} - \varepsilon_2^{X_1} - \varepsilon_3^{X_1}} \left[\begin{array}{ccc} \varepsilon_2^{X_1} + \varepsilon_3^{X_1} & \varepsilon_2^{X_1} + \varepsilon_3^{X_1} & 1 \\ -\varepsilon_1^{X_1} & -\varepsilon_1^{X_1} & -1 \\ -\varepsilon_1^{X_1} & -\varepsilon_1^{X_1} & -1 \end{array} \right]. \quad (12.28)$$

The control coefficients of the branched pathway are expressed in terms of elasticities which can be measured experimentally.

12.7 Response Coefficients

Studies of metabolism have identified different mechanisms that regulate and coordinate the activity of metabolic pathways. Generally, these mechanisms involve ‘effectors’. These effectors may originate ‘internally’ from the pathway (e.g. a pathway intermediate) or ‘externally’ (e.g. a hormone, drug or change in the physical conditions).

Analogous to a control coefficient expressing the dependence of a system variable, such as flux, on an internal parameter, such as enzyme concentration, one can define a *response coefficient* R_z^J to express the dependence of a system variable on an external parameter, such as the concentration $[Z]$ of an external effector Z :

$$R_z^J = \frac{\partial \ln J}{\partial \ln [Z]}. \quad (12.29)$$

An external effector can only produce a local effect by acting on one or more enzymes in the system. Thus, it must have at least one non-zero effector elasticity

$$\varepsilon_z^{v_i} = \frac{\partial \ln v_i}{\partial \ln [Z]}, \quad (12.30)$$

which is defined in the same way as the elasticities in Section 12.4. In the case of an external effector that acts on a single enzyme, the degree to which it influences the pathway flux depends on the degree to which it affects the activity of the target enzyme, expressed by its effector elasticity (12.30), and the degree of control this enzyme has on the pathway flux, expressed by the flux control coefficient of the enzyme. This relationship was proven by Kacser and Burns [KB73] and formulated in the so-called *partitioned response*

$$R_z^J = C_i^J \varepsilon_z^{v_i}. \quad (12.31)$$

The partitioned response generalizes our previous definitions of control and elasticity coefficients. It follows from Eq. (12.31) that all the introduced coefficients can be treated as response coefficients. There is an implicit elasticity of unity connecting the response coefficient with the flux control coefficient. Analogous, an implicit control coefficient of unity relates the elasticity to the response coefficient.

12.8 Limits of Metabolic Control Analysis

After we introduced the building blocks of metabolic control analysis and illustrate them on some simple examples, closing, we want to discuss some problems and future developments of this approach to metabolic networks.

At present and also in these lecture notes, metabolic control analysis is largely concerned with steady states. Nevertheless, there are first approaches to extend it to transient phenomena [KC82, ASK89, MHTSK90, HR91, IS03, HBB⁺05].

Moreover, it assumes that a steady state exists and that it is stable and unique. As it is relatively easy to construct models, even for systems of only two enzymes, that have no steady state, one should not assume the existence of a steady state automatically.

Compartmentation might also seem to pose a problem but, in fact, MCA can readily accommodate multiple compartments as long as there are well defined connections between them and the components are distributed homogeneously within each compartment.

Some further restrictions follow from the mathematical framework of MCA. It deals with infinitesimal changes only, as it was shown in the schematic derivation in Section 12.2. Hence, it cannot predict the consequences of large changes in an experiment. Analytical considerations are more flexible but once a certain steady state was chosen as reference point, one is restricted to its neighborhood. This leads directly to the next point. As mentioned earlier, the coefficients are state dependent and cannot be assigned to other situations.

12.9 Elasticity Coefficients for Specific Rate Laws

As we demonstrated in the previous sections, the control coefficients are closely related to the elasticities of an isolated reaction in the metabolic pathway. Therefore, we want to discuss the elasticities of some well-known rate laws in this section [HS96].

12.9.1 Michaelis-Menten equation

A commonly used reaction motif of a conversion of a substrate into a product was proposed by Michaelis and Menten [MM13] as follows:



where the enzyme E catalyzes the conversion of the substrate S into the product P. One can derive the famous Michaelis-Menten equation

$$v = \frac{V_{\max} S}{K_M + S} \quad (12.33)$$

for the reaction rate if one assumes the intermediary complex C in a quasi-steady state (see Sec. 3.7.1 for a more detailed discussion). $V_{\max} = k_2 E^T$ is the limiting rate, E^T the total enzyme concentration, and $K_M = (k_{-1} + k_2)/k_1$ the Michaelis constant. According to the definition (12.9), the elasticity with respect to the substrate is

$$\begin{aligned} \varepsilon_S &= \frac{S}{v} \frac{\partial v}{\partial S}, \\ &= \frac{S}{v} \frac{\partial}{\partial S} \frac{V_{\max} S}{K_M + S}, \\ &= \frac{S}{v} \left[\frac{V_{\max}}{K_M + S} - \frac{V_{\max} S}{(K_M + S)^2} \right], \end{aligned} \quad (12.34)$$

where we use the quotient rule to get the derivative. Then we use the Michaelis-Menten equation (12.33) to substitute the terms inside the brackets by the reaction rate v

$$\varepsilon_S = \frac{S}{v} \left[\frac{v}{S} - \frac{v}{K_M + S} \right]$$

which enables us to cancel the reaction rate v from the left hand side. After a reduction to the common denominator we obtain the final result

$$\varepsilon_S = \frac{K_M}{K_M + S} \quad (12.35)$$

for the substrate elasticity of the Michaelis-Menten equation. Since the Michaelis constant K_M is positive, this is a monotonically decreasing function with respect to the substrate concentration S .

In the limiting case of a small substrate concentration $S \ll K_M$, we can neglect the substrate compared to the Michaelis constant, and the elasticity becomes $\varepsilon_S \approx 1$. A small change in the substrate concentration has a huge effect on the reaction rate. On the other hand, in the limit of a exceedingly substrate concentration $S \gg K_M$, the elasticity is vanishing. A change of the concentration does not affect the reaction rate. The enzyme is completely saturated. It follows from both limits that the substrate elasticity is in the range

$$0 \leq \varepsilon_S \leq 1.$$

Besides the substrate concentration, the reaction rate also depends on the Michaelis constant and the limiting rate. Both quantities are assumed to be parameters of the enzyme kinetic reaction. The effect of a change in their values is measured by the π -elasticities which we calculate now. From Eq. (12.10) follows for the elasticity with respect to the Michaelis constant K_M

$$\begin{aligned} \pi_{K_M} &= \frac{K_M}{v} \frac{\partial v}{\partial K_M} \\ &= \frac{K_M}{v} \left[-\frac{V_{\max} S}{(K_M + S)^2} \right]. \end{aligned} \quad (12.36)$$

Again we can substitute the rational expression in the brackets with the reaction rate. It follows for the elasticity

$$\pi_{K_M} = -\frac{K_M}{K_M + S} = -\varepsilon_S \quad (12.37)$$

which is a monotone decreasing function in the range $[-1; 0]$. The lower limit is valid in the case $K_M \gg S$ where a further increase of the Michaelis constant will decelerate the reaction. Additionally, the reaction rate is very sensitive to a change in the constant in that regime. In the limit of $K_M \ll S$ the elasticity is approximately zero. Here, the saturation of the enzyme is the dominating effect. A change of K_M has no or only a small effect on the reaction rate.

The limiting rate V_{\max} is the second parameter in the Michaelis-Menten equation (12.33). Its elasticity coefficient is

$$\pi_{V_{\max}} = 1. \quad (12.38)$$

The reaction rate is direct proportional to the limiting rate.

12.9.2 Reversible Michaelis-Menten Equation

The original Michaelis-Menten scheme (12.32) does not include a reverse reaction in the enzymatic reaction. For a well-founded investigation of biochemical pathways one has to consider such a reaction [HCB97]. Otherwise, the chosen description might be valid only in a restricted region of the parameter and variable space. Often these are very special limiting cases. Therefore we now include a reverse reaction into the enzyme kinetic reaction



which results in the rate equation (see Sec. 3.7.6)

$$v = \frac{V_{\max 1} S / K_{M1} - V_{\max 2} P / K_{M2}}{1 + S / K_{M1} + P / K_{M2}}. \quad (12.40)$$

The parameters are now

$$\begin{aligned} K_{M1} &= \frac{k_{-1} + k_2}{k_1} & V_{\max 1} &= k_2 E^T, \\ K_{M2} &= \frac{k_{-1} + k_2}{k_{-2}} & V_{\max 2} &= k_{-1} E^T. \end{aligned}$$

From Eq. (12.9) we obtain for the substrate elasticity

$$\begin{aligned}\varepsilon_S &= \frac{S}{v} \frac{\partial v}{\partial S}, \\ &= \frac{S}{v} \left[\frac{V_{\max 1}/K_{M1}}{1 + S/K_{M1} + P/K_{M2}} - \frac{1}{K_{M1}} \frac{V_{\max 1} S/K_{M1} - V_{\max 2} P/K_{M2}}{(1 + S/K_{M1} + P/K_{M2})^2} \right].\end{aligned}\quad (12.41)$$

We can substitute Eq. (12.40) in the second term and cancel the reaction rate from the prefactor. The first term we multiply with the inverse reaction rate. Then we reduce the resulting expression to the common denominator

$$\varepsilon_S = \frac{S}{K_{M1}} \left[\frac{V_{\max 1}}{V_{\max 1} S/K_{M1} - V_{\max 2} P/K_{M2}} - \frac{1}{1 + S/K_{M1} + P/K_{M2}} \right]$$

and obtain the substrate elasticity

$$\varepsilon_S = \frac{\left[V_{\max 1} + (V_{\max 1} + V_{\max 2}) \frac{P}{K_{M2}} \right] \frac{S}{K_{M1}}}{\left(\frac{V_{\max 1} S}{K_{M1}} - \frac{V_{\max 2} P}{K_{M2}} \right) \left(1 + \frac{S}{K_{M1}} + \frac{P}{K_{M2}} \right)} \quad (12.42)$$

which is positive if $v > 0$ corresponding to a flux from substrate to product. An increase of substrate concentration will increase the flux and the production rate of the product, respectively.

In contrast to the irreversible Michaelis-Menten equation (12.33) the reaction rate depends now also on the product. The corresponding product elasticity is

$$\begin{aligned}\varepsilon_P &= \frac{P}{v} \frac{\partial v}{\partial P}, \\ &= \frac{P}{v} \left[\frac{V_{\max 2}/K_{M2}}{1 + S/K_{M1} + P/K_{M2}} - \frac{1}{K_{M2}} \frac{V_{\max 1} S/K_{M1} - V_{\max 2} P/K_{M2}}{(1 + S/K_{M1} + P/K_{M2})^2} \right], \\ &= \frac{P}{K_{M2}} \left[\frac{V_{\max 2}}{V_{\max 1} S/K_{M1} - V_{\max 2} P/K_{M2}} - \frac{1}{1 + S/K_{M1} + P/K_{M2}} \right], \\ &= \frac{\left[V_{\max 2} + (V_{\max 1} + V_{\max 2}) \frac{S}{K_{M1}} \right] \frac{P}{K_{M2}}}{\left(\frac{V_{\max 1} S}{K_{M1}} - \frac{V_{\max 2} P}{K_{M2}} \right) \left(1 + \frac{S}{K_{M1}} + \frac{P}{K_{M2}} \right)},\end{aligned}\quad (12.43)$$

which one obtains in the same way as described for the substrate elasticity above. The product elasticity is negative for a reaction rate $v > 0$. Hence, an increase of the product will decrease the reaction rate.

12.9.3 Hill equation

Many enzymes show a much more complicated behavior than considered in the derivation of the Michaelis-Menten equation. The corresponding reaction rates can be described by the phenomenological Hill equation

$$v = \frac{V_{\max} (S/K)^n}{1 + (S/K)^n}, \quad (12.44)$$

where V_{\max} is a limiting rate, K an enzyme specific constant, n the Hill coefficient and S the substrate concentration. The substrate elasticity for the Hill equation is

$$\begin{aligned}\varepsilon &= \frac{S}{v} \frac{\partial v}{\partial S} \\ &= \left[\frac{V_{\max} (S/K)^{n-1}}{1 + (S/K)^n} - \frac{V_{\max} (S/K)^n (S/K)^{n-1}}{[1 + (S/K)^n]^2} \right] \frac{S}{K} \frac{n}{v},\end{aligned}\quad (12.45)$$

which we can transform with Eq. (12.44) into the expression

$$\varepsilon_S = \left[v - v \frac{(S/K)^n}{1 + (S/K)^n} \right] \frac{n}{v}$$

leading to the final result

$$\varepsilon_S = \frac{n}{1 + (S/K)^n}. \quad (12.46)$$

The substrate elasticity is a decreasing function with respect to the substrate concentration.

In the limit of small concentrations $S \ll K_M$ we neglect the substrate dependent term in the denominator and obtain the asymptotic elasticity

$$\varepsilon_S \approx n.$$

Again, the reaction rate is very sensitive to changes in the substrate concentration. The sensitivity increases if the Hill coefficient increases. In the limit of high substrate concentrations $S \gg K$ the elasticity becomes

$$\varepsilon_S \approx \frac{n}{(S/K)^n}$$

which simplifies further if we assume $(S/K)^n \gg n$ to

$$\varepsilon_S \approx 0.$$

The Hill equation is insensitive to changes in the substrate concentration. Here, the enzyme saturation is dominant and determines the reaction rate.

Similar to the substrate elasticity we obtain the π -elasticity with respect to the enzyme-specific constant K

$$\begin{aligned} \pi_K &= \frac{K}{v} \frac{\partial v}{\partial K} \\ &= -\frac{n}{1 + (S/K)^n} = -\varepsilon_S. \end{aligned} \quad (12.47)$$

In analogy to the treatment of the Michaelis-Menten equation the reaction rate decreases with increasing constant K . Since the limiting rate V_{\max} acts only as constant factor in the Hill equation (12.44), its π -elasticity is again

$$\pi_{V_{\max}} = 1.$$

An increase in the limiting rate will increase the reaction rate in the same manner.

12.9.4 S-Systems

S-Systems [Sav69a] can be used to approximate the reaction rate near the steady state. Within this approach we aggregate all contributions to the formation of X_i in one single term and all contributions to the degradation of X_i in a second one. The rate law for the j -th reaction is now

$$v_j = k_{+j} \underbrace{\prod_i X_i^{n_{ij}^-}}_{\text{production}} - k_{-j} \underbrace{\prod_i X_i^{n_{ij}^+}}_{\text{degradation}} \quad (12.48)$$

with the forward rate constant k_{+j} and the reverse rate constant k_{-j} and the stoichiometric coefficients n_{ij}^- for reactants and n_{ij}^+ for products. In contrast to the kinetic orders of

elementary reactions introduced in Section 3.5, the kinetic orders n_{ij}^{\pm} can have real values. Note that a stoichiometric coefficient of zero has two possible interpretations: i.) the metabolite does not take part at the considered reaction and ii.) the rate with respect to the considered species is limited by physical processes (see also Section 3.5.5). The steady state is then determined by a generalized law of mass action:

$$K_{\text{eq}} = \frac{k_{+j}}{k_{-j}} = \frac{\prod_i X_i^{n_{ij}^+}}{\prod_i X_i^{n_{ij}^-}}. \quad (12.49)$$

The elasticity for the j -th reaction and with respect to the metabolite X_i is in the S-Systems approach

$$\begin{aligned} \varepsilon_i^j &= \frac{X_i}{v_j} \frac{\partial v_j}{\partial X_i} \\ &= \frac{X_i}{v_j} \left[k_{+j} \frac{\partial}{\partial X_i} \prod_l X_l^{n_{lj}^-} - k_{-j} \frac{\partial}{\partial X_i} \prod_l X_l^{n_{lj}^+} \right]. \end{aligned}$$

The partial derivatives of products in the above equation is

$$\frac{\partial}{\partial X_i} \prod_l X_l^{n_{lj}^{\pm}} = n_{ij}^{\pm} X_i^{n_{ij}^{\pm}-1} \prod_{l \neq i} X_l^{n_{lj}^{\pm}},$$

where the metabolite X_i is excluded from the product. Furthermore, its kinetic order is reduced by one. But a multiplication with the normalization factor restores the original order, and we can group again all participating metabolites in a common product. If we apply the definition of the reaction in the S-System approach (12.48), the elasticity becomes

$$\varepsilon_i^j = \frac{k_{+j} n_{ij}^- \prod_l X_l^{n_{lj}^-} - k_{-j} n_{ij}^+ \prod_l X_l^{n_{lj}^+}}{k_{+j} \prod_l X_l^{n_{lj}^-} - k_{-j} \prod_l X_l^{n_{lj}^+}}. \quad (12.50)$$

The π -elasticities of the rate constants are easy to derive. We obtain an elasticity of

$$\pi_{k_{\pm i}}^j = \pm \delta_{ij} \frac{k_{\pm i} \prod_l X_l^{n_{lj}^{\mp}}}{k_{+j} \prod_l X_l^{n_{lj}^-} - k_{-j} \prod_l X_l^{n_{lj}^+}}, \quad (12.51)$$

where

$$\delta_{ij} = \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{for } i = j \end{cases}$$

is the Kronecker-Delta [AS72]. The π -elasticity is non-vanishing only if the rate constant determines the considered reaction. Otherwise, the rate does not depend on the investigated rate constant and the elasticity is zero. If $v_j > 0$, then the π -elasticity with respect to the forward rate constant is always positive. Its increase further increases the reaction rate. In contrast, the π -elasticity of the reverse rate constant is negative, corresponding to a decreasing reaction rate with increasing rate constant.

12.9.5 Elasticity Calculus

As demonstrated in the previous sections the derivation of elasticity coefficients becomes more and more complicated if the complexity of the considered reactions increases. Due to the resulting partial derivatives the expressions are very complex. Formally, due to

the normalization often a simplification arises because of cancelations between numerator and denominator terms with the normalization factor. Hence, one can use an elasticity calculus [Fel97] to avoid the generation of terms that will subsequently cancel.

Often the reaction rate is a rational form of

$$v = \frac{N}{D} \quad (12.52)$$

where both function N and D depend on the metabolite concentration X_i . The corresponding elasticity is

$$\begin{aligned} \varepsilon_i &= \frac{X_i}{v} \frac{\partial v}{\partial X_i} \\ &= X_i \frac{D}{N} \left[\frac{\partial N / \partial X_i}{D} - N \frac{\partial D / \partial X_i}{D^2} \right] \end{aligned} \quad (12.53)$$

which simplifies if both terms in the brackets are multiplied with the inverse reaction rate of the normalization factor. Finally we obtain with

$$\varepsilon_i = X_i \left[\frac{1}{N} \frac{\partial N}{\partial X_i} - \frac{1}{D} \frac{\partial D}{\partial X_i} \right] \quad (12.54)$$

a simple rule for the calculation of the elasticity coefficients. Nevertheless, we must emphasize, that the introduced calculus is applicable only if the reaction rate has a rational form as proposed in Eq. (12.52). Otherwise, one has to use the general definition (12.9). In this case a step by step derivation, as shown on the examples in the previous section, cannot be avoided.

A Glossary

Most biological textbooks have a glossary, for mathematical expressions we refer to Eric Weisstein's MathWorld web-site <http://mathworld.wolfram.com/> or a mathematics dictionary (e.g. [BB89]).

abscissa The horizontal axis or x -coordinate of a point in the two-dimensional plane. See also *ordinate*.

absolute value The positive value of a number, disregarding its sign and written $|x|$.

activation loop A segment of the amino acid sequence that contains phosphorylation sites usually at the surface of a protein and accessible by protein kinases.

active site Region of an enzyme surface to which a substrate molecule binds in order to undergo a catalyzed reaction.

active transport Movement of a molecule across a membrane or other barrier driven by energy other than that stored in the electrochemical gradient of the transported molecule.

adaptors Adaptor proteins typically do not possess a catalytic function but bind to other proteins. Adaptors serve to physically connect proteins with each other. See also *exchange factors*.

algebra A branch of mathematics that generalizes arithmetic operations with numbers to operations with variables, matrices etc.

amino acid Class of biochemical compounds from which proteins are composed. Around 20 amino acids are present in proteins.

analysis A branch of mathematics concerned primarily with limits of functions, sequences, and series.

analytic function A function possessing derivatives of all orders and agreeing with its Taylor series locally.

antibody A protein molecule produced in response to an antigen.

antigen Molecules that is able to provoke an immune response.

argument The argument of a function is the element to which a function applies. Usually the independent variable of the function.

associative A law or operation is called associative if the placement of brackets does not matter: $(a \cdot b) \cdot c \equiv a \cdot (b \cdot c)$.

ATP The principal carrier of chemical energy in cells.

autocatalysis Reaction that is catalyzed by one of its products, creating a positive feedback (self-amplifying) effect on the reaction rate.

autoinhibition Mechanism for inhibiting own activity; e.g., Raf contains an autoregulatory domain that inhibits its own activity by binding to its catalytic domain. The autoregulatory domain is relieved from the catalytic domain by phosphorylation of characteristic residues.

- autonomous** A system (of differential equations) is said to be autonomous if it does not *explicitly* depend on time.
- apoptosis** Controlled cell death.
- attractor** A region of the space describing the temporal solution of a dynamic system towards which trajectories nearby converge, are attracted to. An attractor can be an equilibrium point or a circle. An attractive region that has no individual equilibrium point or cycle is referred to as a chaotic or strange attractor.
- bifurcation point** An instability point in which a single equilibrium condition is split into two. At a bifurcation point the dynamics of a system changes structurally.
- bioinformatics** The management and analysis of genomic data, most commonly using tools and techniques from computer science.
- calculus** A branch of mathematics concerned with the rate of change of a dependent variable in a function.
- category theory** A branch of mathematics that considers mappings and their effect on sets. A category is a structure consisting of a set of objects and a class of maps, which satisfy specific properties.
- chain rule** A rule used in the context of differential equations and which states that $dy/dx = dy/dt \times dt/dx$.
- class** Another name for set, especially a finite set.
- closed form** An expression or solution in terms of well understood quantities.
- coefficient** A numerical or constant multiplier of a variable in an algebraic term.
- continuous function** A function for which the value changes gradually.
- control** Target or set-point tracking, making the system sensitive to changes in the input. See also *regulation* and *homeostasis*.
- control coefficient** Relative measure of how much a perturbation of a systems parameter affects a systems variable. The control coefficient is a global property and one of three building blocks in Metabolic Control Analysis.
- cytokine** Extracellular signal protein or peptide that acts as a local short distance mediator in cell-cell communication. Cytokines are called lymphokines if produced by lymphocytes, interleukines if produced by leucocytes, and monokines if produced by monocytes and macrophages.
- cytoplasm** Contents of a cell that are contained within its plasma membrane but, in the case of eucaryotic cells, outside the nucleus.
- damped oscillations** An oscillation in which the amplitude decreases over time.
- differentiation** process by which the cell acquires specialized functional properties.
- differentiable A system (usually a process described by differential equations) is called differentiable if its *phase space* has the structure of a differentiable *manifold*, and the change of *state* is described by differentiable functions.
- dimer** A protein molecule which consist of two subunits separated polypeptide chains); homodimer: the subunits are identical; heterodimer: the subunits are different; heterotrimer: three subunits, some different.

dimerization The process by which two molecules of the same chemical composition form a condensation product or polymer.

discretization An approximation of a continuous object.

dynamic system A system that changes with time.

EGF Epidermal Growth Factor. EGF is expressed by many cells and stimulates the proliferation of many cell types via Ras and the Raf/MEK/ERK pathway.

EGFR EGF Receptor, a prototypical receptor tyrosine kinase.

electrophoresis An experimental technique to separate DNA fragments or proteins from a mixture. The molecules are separated by their mass, size or rate of travel through a medium (typically agarose or gels) and their electrical charge.

enzyme Protein that catalyzes a specific chemical reaction.

elasticity coefficient Relative measure of the dependence of the reaction rate of an isolated step in a metabolic pathway on systems variables and internal parameters. The elasticity coefficient is a local or system property and one of three building block of Metabolic Control Analysis.

epithelial A epithelial is a coherent cell sheet formed from one or more layers of (epithelial) cells covering an external surface or lining a cavity. For example, the epidermis is the epithelial layer covering the outer surface of the body.

equilibrium State where there is no net change in a system. E.g. in a chemical reaction the equilibrium is defined by the state at which the forward and reverse rates are equal.

equilibrium point Point such that the derivatives of a system of differential equations are zero. An equilibrium point may be stable (then called an *attractor*) or unstable (*repellor*).

exchange factors Bind to the activated receptor, i.e., act as an adaptor; facilitate the exchange of bound GDP for GTP on small G-proteins, which are several steps away from the receptor, and thus activate them.

expression Production of a protein which has directly observable consequences.

extended phase space See *phase space*.

feedback inhibition Regulatory mechanism in metabolic pathways - an enzyme further up in the pathway is inhibited by a product further down in that pathway.

finite-dimensional A process is called finite-dimensional if its *phase space* is finite dimensional, i.e., if the number of parameters needed to describe its *states* is finite.

fixed point See *steady state*.

formal system A mathematical framework in which to represent natural systems.

fun What we experience doing mathematics.

function A relation between two sets that describes unique associations among the elements of the two sets. A function is sometimes called a mapping or transformation.

GAP GTPase Activating Protein. Ras proteins possess intrinsic GTPase activity which hydrolyzes the bound GTP to GDP, i.e., cleaves off a phosphate from GTP. This hydrolysis is a dephosphorylation and as such a phosphatase reaction. A dephosphorylation or phosphatase reaction is a special case of a hydrolysis reaction. Hydrolysis reactions are all reactions where water, H_2O , is used to break a chemical bond. The intrinsic GTPase activity of Ras is weak. However, GAPs can accelerate this activity almost 1000fold. GAPs do not hydrolyze GTP, they bind to Ras and make Ras a more efficient GTPase.

gene product The macromolecules, RNA or proteins, that are the result of *gene expression*.

gene expression The process by which the information, coded in the genome, is transcribed into RNA. Expressed genes include those for which the RNA is *not* translated into proteins.

genome The entirety of genetic material (DNA) of a cell or an organism.

gradient The slope of a line measured as the ratio of its vertical change to its horizontal change.

Grb-2 Growth-factor Receptor Binding protein-2. Grb-2 is an adaptor protein.

group A mathematical group is a set, together with a binary operation on the group elements.

growth factor Extracellular signalling molecule that can stimulate a cell to grow or proliferate.

G-proteins Small monomeric GTP-binding proteins (e.g. Ras), molecular switches that modulate the connectivity of a signalling cascade: resting G-proteins are loaded with GDP and inactive, replacement of GDP with GTP by exchange factors means activation.

GTP/GDP Guanosine triphosphate (GTP) refers to three phosphate molecules attached to the sugar, guanosine diphosphate for two (GDP). See also *GAP*.

homeostasis Regulation to maintain the level of a variable. See also *regulation*.

homologues proteins/genes Have descended from a common ancestor; genes are either homologous or non-homologous, not in between; though, due to multiple genomic rearrangements, the evolutionary history of individual components (domains = evolutionary units) of a gene/protein might be difficult to trace.

hydrolysis See *GAP*.

immunoglobulin General expression for antibody molecules.

infinitesimal Infinitely small. Infinitesimal quantities are used to define integrals and derivatives, and are studied in the branch of maths called analysis.

integral curve A trajectory in extended phase space.

in vitro Experimental procedures taking place in an isolated cell-free extract. Cells growing in culture, as opposed to an organism.

in vivo In an intact cell or organism.

in silico In a computer, simulation.

isoforms Closely homologous proteins (from different genes) that perform similar or only slightly different functions, e.g., under tissue-specific control. Two or more RNAs that are produced from the same gene by different transcription and/or differential RNA splicing are referred to as isoforms.

kinase Enzyme which catalyzes the phosphorylation of a protein.

ligand Molecule that bind to a specific site on a protein or other molecule.

linear equation An equation $y = ax + b$ is linear because the graph of y against x is a straight line (with slope a and intercept b). A linear equation should not be confused with a linear system. See also *nonlinearity*.

linear system A system is nonlinear if changes in the output are not proportional to changes in the input.

linearization Taylor expansion of a dynamical system in the dependent variable about a specific solution, discarding all but the terms linear in the dependent variable.

locus The position of a gene on a chromosome, the DNA of that position; usually restricted to the main regions of DNA that are expressed.

lysis Rupture of a cell's plasma membrane, leading to the release of cytoplasm and the death of the cell.

manifold A mathematical space in which the local geometry around a point in that space is equivalent to the Euclidean space.

MAP-kinase Mitogen-activated protein kinase that performs a crucial step in transmitting signals from the plasma membrane to the nucleus.

Metabolic Control Analysis (MCA) Method for analyzing variation in fluxes and intermediate concentrations in a metabolic pathway relating to the effects of the different enzymes that constitute the pathway and external parameters. The building blocks of MCA are: control coefficients, elasticity coefficients, and response coefficients.

metabolism The entirety of chemical processes in the cell.

mitogen Substance that stimulates the mitosis of certain cells.

mitosis Process in cell division by which the nucleus divides.

monomer A protein molecule which consist of one subunits separated polypeptide chains; homodimer: the subunits are identical; heterodimer: the subunits are different; heterotrimer: three subunits, some different.

morphism Generalization of the concepts of relation and function. Often synonymously used with mapping.

multimer A protein molecule which consist more than four subunits separated polypeptide chains); homodimer: the subunits are identical; heterodimer: the subunits are different; heterotrimer: three subunits, some different.

natural system An aspect of the phenomenal world, studied in the natural sciences.

noise A description of real or simulated data for which the behavior is or appears unpredictable.

- nonlinearity** Linearity is defined in terms of functions that have the property $f(x+y) = f(x)+f(y)$ and $f(ax) = af(x)$. This means that the result f may not be proportional to the input x or y .
- oncogene** An altered gene whose product which takes a dominant role in creating a cancerous cell.
- ordinate** The vertical or y -axis of the coordinate system in the plane.
- orbit** The set of points in phase space through which a trajectory passes.
- organization** Pattern or configuration of processes.
- peptide** A small chain of amino acids linked by peptide bonds.
- percepts** The consequence of cognitive processes or observations.
- phase space** Phase space is the collection of possible states of a dynamical system, i.e., the mathematical space formed by the dependent variables of a system. An extended phase space is the cartesian product of the phase space with the independent variable, which is often time.
- phenomenon** A collection of percepts to which relationships are assigned.
- phosphatase** Enzyme that removes phosphate groups from a molecule.
- phosphorylation** Important regulatory process, one third of mammalian proteins are regulated by reversible phosphorylation; phosphate groups P from ATP molecules are transferred to the -OH groups of serine, threonine or tyrosine residues by protein kinases; phosphate groups are two times negatively charged, their addition will change the protein's local conformational characteristics and can thus activate a protein. See also *GAP* and *protein phosphorylation*.
- polymer** Large molecule made by linking monomers together.
- protein** A linear polymer of linked amino acids, referred to as a macromolecule and major constituent component of the cell.
- protease, proteinase** Enzymes that are degrading proteins by splitting internal peptide bonds to produce peptides.
- proteinase inhibitor** small proteins that inhibit various proteinase enzymes. An example is antitrypsin.
- protein kinase** Enzyme that transfers the terminal phosphate group of ATP to a specific amino acid of a target protein.
- protein phosphorylation** The covalent addition of a phosphate group to a side chain of a protein catalyzed by a protein kinase.
- random process** A description of real or simulated data for which the behavior is or appears unpredictable.
- RAS protein** Member of a large family of GTP-binding proteins that helps transmit signals from cell-surface receptors to the nucleus. Ras-GDP is the inactive form of Ras, which is bound to Guanosine-Di-Phosphate. Ras-GTP is the active form of Ras, which is bound to Guanosine-Tri-Phosphate. This form of Ras undergoes a conformational change that enables it to bind with high affinity to other proteins such as Raf.

receptor tyrosine kinase Receptor tyrosine kinases play an important role in the regulation of cell proliferation, survival and differentiation. The binding of the ligand (including growth factors, hormones etc.) to the extracellular portion of the receptor typically activates the kinase activity of the intracellular portion of the receptor, resulting in autophosphorylation on several tyrosine residues. the phosphorylated tyrosines serve as docking sites for adaptor proteins such as Grb-2 resulting in the assembly of a multiprotein complex at the receptor. This complex is a platform that typically mediates the specific biological responses by activating several intracellular signalling pathways.

regulation The maintenance of a regular or desirable state, making a system robust against perturbations. See also *homeostasis* and *control*.

repressor Protein that binds to a specific region of DNA to prevent transcription of an adjacent gene.

residue Proteins are built of amino acids by forming peptide bonds under removal of water; what remains of the amino acids are the amino acid residues.

response coefficient Relative Measure of the dependence of a system variable of a pathway on an external parameter. The response coefficient is one of three building blocks in Metabolic Control Analysis.

sample space The set of possible outcomes in a statistical experiments.

scaffold protein Protein that organizes groups of interacting intracellular signalling proteins into signalling complexes.

signalling, signal transduction A process by which signals are relayed through biochemical reactions.

sigma algebra A σ -algebra is a collection of subsets of a set that contains the set itself, the empty set, the complements in the set of all members of the collection, and all countable unions of members.

steady state A system state in which the system remains. A steady state is associated with a *fixed point*, i.e., the point in the state-space in which the system remains.

stochastic process A mathematical concepts defined as a sequence of random variables.

SOS Son of Sevenless. SOS is the prototypic GDP/GTP Exchange Factor, GEF. There are many GEFs, but SOS is ubiquitously expressed. GEFs cause Ras to release GDP. Since the cell contains much higher concentrations of GTP than GDP, per default a GTP molecule will bind to Ras in place of the released GDP. Oncogenic Ras mutants cannot release GDP. Therefore, they are always in the active (GTP bound) form.

system A collection of objects and a relation among these objects.

tangent bundle The set of tangent vectors to a manifold.

terminal domain N-terminal domain, C-terminal domain chain of amino acid residues leaves an amino group free at one end, and a carboxyl group at the other end; by convention a protein chain starts at the N-terminus, i.e., the N-terminal domain is the first domain near the amino terminus; the C-terminal domain the last near the carboxyl terminus.

tetramer A protein molecule which consist of four subunits separated polypeptide chains; homodimer: the subunits are identical; heterodimer: the subunits are different; heterotrimer: three subunits, some different.

TNF Tumor necrosis factor, protein produced by macrophages in the presence of an endotoxin.

trajectory The solution of a set of differential equations, synonymous with the phrase phase curve.

tyrosine kinase See *receptor tyrosine kinase*.

vector A mathematical vector is an ordered set of elements, e.g., (a, c, b) . An unordered list is denoted $\{a, b, c\}$, where the position of the elements in the list does not matter.

B Notation

The notation used in this text was one of the biggest challenges. Since we are dealing with various aspects of mathematics and different application areas, there are conflicting customary uses of symbols. For example, in stochastic modelling we use n to denote the state vector, i.e., the number of molecules at any particular time. In modelling with differential equations, n is a constant used to denote the number of equations, \dot{x}_i , $i = 1, \dots, n$. The letter x refers to a variable, random variable $x(t)$, vector $x = (x_1, \dots, x_n)$, An effort is made to introduce notation and symbols where they appear first. According to convention in biological textbooks, acronyms printed in lower case indicate genes (e.g. ras), capitalized acronyms indicate their protein products (Ras or RAS).

Units

L	liter.
Da	Dalton.
mol	moles, molar mass.
M	molarity, molar concentration.
sec	seconds.
min	minutes.
g	grams.

Mathematical Symbols

\rightarrow	mapping, function, morphism, arrow.
\mapsto	“maps to”.
$:$	“for which”, “such that”.
$ $	“conditional on”.
\forall	“for all”.
\in	“element of”.
\doteq	“by definition”.
\exists	“there exists”.
\equiv	“equivalent”, “identical”.
\propto	“proportional to”.
\approx	“approximately”.
\Rightarrow	“implies”, material implication.
\Leftrightarrow	“if and only if” (iff).
\therefore	“therefore”.

N_A	Avogadro number.
$\{ \}$	set, list.
$()$	ordered set, sequence, vector.
\mathbb{Z}	set of integers $\{\dots, -2, -1, 0, 1, 2, \dots\}$.
\mathbb{Z}_+	set of nonnegative integers $\{0, 1, 2, \dots\}$.
\mathbb{N}	set of natural numbers $\{1, 2, \dots\}$.
\mathbb{R}	set of real numbers.
\mathbb{Q}	set of rational numbers.
\mathbb{C}	set of complex numbers.
$\mathbb{R}^{p \times m}$	set of real $p \times m$ matrices.

\mathbb{B}	σ -algebra.
\emptyset	empty set.
\subseteq	subset.
\subset	proper subset.
\cap	intersection.
\cup	union.
\preceq	partial or semi-ordering.
\vee	disjunction, “or”.
\wedge	conjunction, “and”.
\circ	composition.
$1(\cdot)$	identity map.
d/dt	differential operator in an ODE.
\dot{x}	short form of the differential dx/dt .
$\partial/\partial t$	partial differential operator.
$\mathcal{N}(\bar{x}, \sigma_x^2)$	normal or Gaussian probability distribution/density function.
\bar{x}	mean value.
σ^2	variance.
ρ	Euclidean distance.
$n!$	factorial, $n! = 1 \times 2 \times 3 \times \cdots \times n$.
∞	infinity.

Abbreviations

ADP	adenosine diphosphate.
ATP	adenosine triphosphate.
CME	chemical master equation.
EGF	epidermal growth factor.
ERK	extracellular signal-regulated kinase.
GDP	guanosine diphosphate.
GTP	guanosine triphosphate.
GMA	generalized mass action.
JAK	janus kinase.
LMA	law of mass action.
MAP	mitogen-activated protein.
MAPK	mitogen-activated protein kinase.
MAPKK	mitogen-activated protein kinase kinase.
MEK	MAPK/ERK kinase.
MEKK	MEK kinase.
ODE	ordinary differential equation.
pgf	probability generating function.
mgf	moment generating function.
cgf	cumulant generating function.
CV	coefficient of variation.
Var	variance.
Std	standard deviation.
lim	limes, in the limit.
det	determinant.
w.r.t.	with respect to.
iff	if and only if.
SOS	son of sevenless.
STAT	signal transducers and activators of transcription.

TNF	tumor necrosis factor.
TPA	12-O-tetradecanoyl-phorbol-12-acetate.

Chapter 2

\mathcal{S}	system.
O	object(s).
R	relation.
$A \times B$	Cartesian product.
T	time set.
I	index set.
U, Y	input, output objects/spaces.
ϕ	state mapping.
g, h	input, output mapping.
u, y, x	input-, output-, and state-variable/vector.
Ω	sample space of a random variable.
\mathbb{B}	σ -algebra.
$P(\cdot)$	probability measure/function.
$\text{Prob}\{A\}$	probability of event A .
$\omega \in \Omega$	elementary event.
$w(\omega)$	random variable.
$w_t(\omega)$	stochastic process.
n	number of state variables/ODEs.
m	number of dependent variables.
K_{eq}	equilibrium constant.
K_d	dissociation constant.

Chapter 3

n	number of molecules, state-vector.
Δ	small but not infinitesimal change.
$\#S$	number of molecules.
n_T	total number of molecules.
k	rate constant.
R_μ	reaction channel (irreversible reaction).
M	number of reaction channels.
\tilde{x}	steady state.
$\langle S(t) \rangle$	mean or average of the process $S(t)$.
K_m	Michaelis-Menten constant.
V	volume, or velocity.
V_{max}	limiting rate in a kinetic reaction.
S_j	chemical species.
N	number of chemical species.
$[S]$	concentration of S .
S	state (vector) of the system.
R_μ	reaction channel.
c_μ	stochastic reaction constant (stochastic simulation).
a_μ	propensity of reaction R_μ .
a^*	propensity for any of the R_μ to occur.
h_μ	number of distinct combinations of R_μ reactant molecules.

K_μ	molecularity of reaction R_μ .
$l_{\mu j}$	stoichiometric coefficient.
L_μ	number of reactant species.
$\nu_{\mu j}$	change in the population of molecular species S_j in reaction R_μ .
$P(\cdot)$	probability measure.
$F(\cdot)$	cumulative distribution function.
$p_{m,n}$	transition probability.
Π	probability transition matrix.
v_μ	rate of reaction.
\mathfrak{P}	probability generating function (pgf).
\mathfrak{P}'	derivative of the pgf.
\mathfrak{M}	moment generating function (mgf).
\mathfrak{C}	cumulant generating function (cgf).

Chapter 6

θ	parameter(s).
n, m	number of dependent, independent variables.
x, X	state variable, state space or fiber.
u, y	input and output variable.
ϕ	state mapping.
h	output mapping.
J	Jacobian matrix.
$H(A, B)$	set of all mappings from A to B .
φ	flow.
\mathcal{G}	group.
\mathcal{C}	category.
$T_x X$	tangent space to domain X .
TX	tangent bundle.
\mathbb{M}	family of models.
(\mathbb{P}, π)	parametrization, $\pi: \mathbb{P} \rightarrow \mathbb{M}$.
\mathbb{P}	parameter space, base space.
(\mathbb{P}, X)	fiber bundle.
B^A	exponential of maps from A to B .
e_f	evaluation map.

Bibliography

- [A⁺04] M.I. Aladjem et al. Molecular interaction maps—a diagrammatic graphical language for bioregulatory networks. *Sci. STKE*, 222:pe8, 2004. <http://www.stke.org>.
- [AdP02] P.W. Atkins and J. de Paula. *Atkins' Physical Chemistry*. Oxford University Press, seventh edition, 2002.
- [AFJS04] D. Angeli, J.E. Ferrell Jr., and E.D. Sontag. Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proc. Natl. Acad. Sci. USA*, 101(7):1822–1827, February 2004.
- [AJL⁺02] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Molecular Biology of the Cell*. Garland Science, fourth edition, 2002.
- [AL01] A.R. Asthagiri and D.A. Lauffenburger. A computational study of feedback effects on signal dynamics in a mitogen-activated protein kinase (MAPK) pathway model. *Biotechnol. Prog.*, 17:227–239, 2001.
- [AS72] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions*. Number 55 in Applied Mathematics Series. National Bureau of Standards, tenth edition, 1972.
- [ASK89] L. Acerenza, H.M. Sauro, and H. Kacser. Control analysis of time-dependent metabolic systems. *J. Theor. Biol.*, 137:423–444, 1989.
- [Bat79] E. Batschelet. *Introduction to Mathematics for Life Scientists*. Springer., third edition, 1979.
- [BB89] E.J. Borowski and J.M. Borwein. *Collins Dictionary of Mathematics*. Harper Collins Publishers, 1989.
- [BB01] J.M. Bower and H. Bolouri. *Computational Modeling of Genetic and Biochemical Networks*. MIT Press, 2001.
- [BCBG⁺85] J.A. Burns, A. Cornish-Bowden, A.K. Groen, R. Heinrich, H. Kacser, J.W. Porteus, S.M. Rapoport, T.A. Rapoport, J.W. Stucki, J.M. Tager, R.J.A. Wanders, and H.V. Westerhoff. Control analysis of metabolic systems. *Trends Biochem. Sci.*, 10(1):16, January 1985.
- [BD01] W.E. Boyce and R.C. DiPrima. *Elementary differential equations and boundary value problems*. John Wiley and Sons, seventh edition, 2001.
- [Bel90] D. Bell. *Mathematics of Linear and Nonlinear Systems*. Oxford University Press, 1990.
- [Ber68] L. von Bertalanffy. *General Systems Theory*. George Braziller, 1968.
- [BF00] F.A. Brightman and D.A. Fell. Differential feedback regulation of the MAPK cascade underlies the quantitative differences in EGF and NGF signalling in PC12 cells. *FEBS Lett.*, 482(3):169–174, October 2000.
- [BH25] G.E. Briggs and J.B.S. Haldane. A note on the kinetics of enzyme action. *Biochem. J.*, 19:338–339, 1925.
- [BI99] U.S. Bhalla and R. Iyengar. Emergent properties of networks of biological signaling pathways. *Science*, 283:381–387, 1999.

- [Bla05] F.F. Blackman. Optima and limiting factors. *Ann. Bot.*, 19(2):281–296, April 1905.
- [BMCM98] J. Benito, C. Martín-Castellanos, and S. Moreno. Regulation of the g_1 phase of the cell cycle by periodic stabilization and degradation of the $p25^{rum1}$ cdk inhibitor. *EMBO J.*, 17(2):4882–497, January 1998.
- [BPM82] R.D. Bliss, P.R. Painter, and A.G. Marr. Role of feedback inhibition in stabilizing the classical operon. *J. Theor. Biol.*, 97:177–193, 1982.
- [Bro99] T.A. Brown. *Genomes*. βIOS Scientific Publishers, 1999.
- [BSDdPM00] M.A. Blanco, A. Sánchez-D´az, J.M. de Prada, and S. Moreno. $Apc^{ste9/srw1}$ promotes degradation of mitotic cyclins in g_1 and is inhibited by $cdc2$ phosphorylation. *EMBO J.*, 19(15):3945–3955, August 2000.
- [CB95a] A. Cornish-Bowden. *Fundamentals of Enzyme Kinetics*. Portland Press, 1995.
- [CB95b] A. Cornish-Bowden. Metabolic control analysis in theory and practice. *Advances in Molecular and Cell Biology*, 11:21–64, 1995.
- [CB99] A. Cornish-Bowden. *Basic Mathematics for Biochemists*. Oxford University Press, second edition, 1999.
- [CB04] A. Cornish-Bowden. *Fundamentals of Enzyme Kinetics*. Portland Press, third edition, 2004.
- [CBN95] J. Correa-Bordes and P. Nurse. $p25^{rum1}$ orders s phase and mitosis by acting as an inhibitor of the $p34^{cdc2}$ mitotic kinase. *Cell*, 83(6):1001–1009, December 1995.
- [Dav01] E.H. Davidson. *Genomic Regulatory Systems*. Academic Press, 2001.
- [Dow01] J. Downward. The ins and outs of signalling. *Nature*, 411(6839):759–762, June 2001.
- [ECG80] G. Eason, C.W. Coles, and G. Gettinby. *Mathematics and Statistics for the Bio-Sciences*. Pearson Education, 1980.
- [ELS01] J. Elf, P. Lötstedt, and P. Sjöberg. Problems of high dimension in molecular biology. In *17th GAMM-Seminar*, pages 1–10, Leipzig, 2001.
- [EP01] C.H. Edwards and D.E. Penney. *Differential Equations and Linear Algebra*. Prentice Hall, 2001.
- [Fal02] C.P. Fall. *Computational Cell Biology*. Springer Verlag, 2002.
- [Fel97] D. Fell. *Understanding the Control of Metabolism*. Portland Press, 1997.
- [FJ96] J.E. Ferrell Jr. Tripping the switch fantastic: how a protein kinase cascade can convert graded inputs into switch-like outputs. *TIBS*, 21(12):460–466, December 1996.
- [FN95] D. Fisher and P. Nurse. Cyclins of the fission yeast *Schizosaccharomyces pombe*. *Seminars in Cell Biology*, 6(2):73–78, April 1995.
- [Gar85] C.W. Gardiner. *Handbook of Stochastic Models*. Springer, second edition, 1985.

- [GB00] M.A. Gibson and J. Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *J. Phys. Chem.*, A104:1876–1888, 2000.
- [GH83] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer, 1983.
- [Gie88] C. Giersch. Control analysis of metabolic networks. 1. homogeneous functions and the summation theorems for control coefficients. *Eur. J. Biochem.*, 174:509–513, 1988.
- [Gil76] D.T. Gillespie. General method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22:403–434, 1976.
- [Gil77] D.T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977.
- [Gil92a] D. T. Gillespie. *Markov Processes*. Academic Press., 1992.
- [Gil92b] D.T. Gillespie. A rigorous derivation of the chemical master equation. *Physica A*, 188:404–425, 1992.
- [Gil01] D.T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting system. *J. Chem. Phys.*, 115(4):1716–1733, July 2001.
- [GKJ81] A. Goldbeter and D.E. Koshland Jr. An amplified sensitivity arising from covalent modification in biological systems. *Proc. Natl. Acad. Sci. USA*, 78(11):6840–6844, November 1981.
- [Gol96] A. Goldbeter. *Biochemical Oscillations and Cellular Rythms*. Cambridge University Press, 1996.
- [Gom02] B. Gompertz. *Signal Transductions*. Academic Press, 2002.
- [Gri68] J.S. Griffith. Mathematics of cellular control processes. *J. Theor. Biol.*, 20:202–208, 1968.
- [GS94] P. Gray and S. Scott. *Chemical Oscillations and Instabilities: Nonlinear Chemical Kinetics*. Oxford University Press, 1994.
- [GWW⁺82] A.K. Groen, R.J.A. Wanders, H.V. Westerhoff, R. van der Meer, and J.M. Tager. Quantification of the contribution of various steps to the control of mitochondrial respiration. *J. Biol. Chem.*, 257(6):2754–2757, March 1982.
- [Han97] J.T. Hancock. *Cell Signalling*. Prentice Hall, 1997.
- [Har01] F.M. Harold. *The Way of the Cell*. Oxford University Press, 2001.
- [HBB⁺05] J.J. Hornberg, F.J. Bruggeman, B. Binder, C.R. Geest, A.J.M.B. de Vaate, J. Lankelma, R. Heinrich, and H.V. Westerhoff. Principles behind the multifarious control of signal transduction. ERK phosphorylation and kinase/phosphatase control. *FEBS J.*, 272(1):244–258, January 2005.
- [HCB91] J.-H. S. Hofmeyr and A. Cornish-Bowden. Quantitative assessment of regulation in metabolic systems. *Eur. J. Biochem.*, 200:223–236, 1991.

- [HCB97] J.-H. S. Hofmeyr and A. Cornish-Bowden. The reversible hill equation: how to incorporate cooperative enzymes into metabolic models. *Comput. Appl. Biosci.*, 13(4):377–385, 1997.
- [HCBR93] J.-H.S. Hofmeyr, A. Cornish-Bowden, and J.M. Rohwer. Taking enzyme kinetics out of control; putting control into regulation. *Eur. J. Biochem.*, 212:833–837, 1993.
- [HFJ96] Chi-Ying F. Huang and J. E. Ferrell Jr. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proc. Natl. Acad. Sci.*, 93(19):10078–10083, September 1996.
- [HFWN94] J. Hayles, D. Fisher, A. Woollard, and P. Nurse. Temporal order of s phase and mitosis in fission yeast is determined by the state of the p34cdc2-mitotic b cyclin complex. *Cell*, 78(5):813–822, September 1994.
- [HNR02] R. Heinrich, B.G. Neel, and T.A. Rapoport. Mathematical models of protein kinase signal transduction. *Mol. Cell*, 9:957–970, May 2002.
- [Hof01] J.-H. S. Hofmeyr. Metabolic control analysis in a nutshell. In T.-M. Yi, M. Hucka, M. Morohashi, and H. Kitano, editors, *Proceedings of the 2nd International Conference on systems Biology*, pages 291–300. Omnipress, Wisconsin, 2001.
- [HR73] R. Heinrich and T.A. Rapoport. Linear theory of enzymatic chains; its application for the analysis of the crossover theorem and of the glycolysis of human erythrocytes. *Acta Biol. Med. Germ.*, 31:479–494, 1973.
- [HR74] R. Heinrich and T.A. Rapoport. A linear steady-state treatment of enzymatic chains. general properties, control and effector strength. *Eur. J. Biochem.*, 42:89–95, 1974.
- [HR75] R. Heinrich and T.A. Rapoport. Mathematical analysis of multienzyme systems. II. steady-state and transient control. *Biosystems*, 7:130–136, 1975.
- [HR91] R. Heinrich and C. Reder. Metabolic control analysis of relaxation processes. *J. Theor. Biol.*, 151:343–350, 1991.
- [HS96] R. Heinrich and S. Schuster. *The Regulation of Cellular Systems*. Chapman and Hall, 1996.
- [HSD04] M.W. Hirsch, S. Smale, and R.L. Devaney. *Differential Equations, Dynamical Systems and an Introduction to Chaos*. Elsevier - Academic Press, 2004.
- [IS03] B.P. Ingalls and H.M. Sauro. Sensitivity analysis of stoichiometric networks: an extension of metabolic control analysis to non-steady state trajectories. *J. Theor. Biol.*, 222:23–36, 2003.
- [Isi89] A. Isidori. *Nonlinear Control Systems*. Springer-Verlag, 1989.
- [JM61] F. Jacob and F. Monod. Genetic regulatory mechanisms in the synthesis of protein. *J. Mol. Biol.*, 3:318, 1961.
- [JS99] D.W. Jordan and P. Smith. *Nonlinear Ordinary Differential Equations*. Oxford University Press, third edition, 1999.
- [JS01] P.W. Jones and P. Smith. *Stochastic Processes*. Arnold Publishers, 2001.

- [K⁺02] B.N. Kholodenko et al. Untangling the wires: A strategy to trace functional interactions in signaling and gene networks. *Proc. Natl. Acad. Sci. USA*, 99(20):12841–12846, October 2002.
- [KB73] H. Kacser and J.A. Burns. The control of flux. *Symp. Soc. Exp. Biol.*, 27:65–104, 1973.
- [KC82] M.C. Kohn and E. Chiang. Metabolic network sensitivity analysis. *J. Theor. Biol.*, 98:109–126, 1982.
- [KG95] D. Kaplan and L. Glass. *Understanding Nonlinear Dynamics*. Springer Verlag, 1995.
- [Kho00] B.N. Kholodenko. Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. *Eur. J. Biochem.*, 267(6):1583–1588, March 2000.
- [Kie02] A.M. Kierzek. STOCKS STOchastic Kinetic Simulations of biochemical systems with Gillespie algorithm. *Bioinformatics*, 18:470–481, 2002.
- [Kit02] H. Kitano. Computational systems biology. *Nature*, 420:206–210, November 2002.
- [Kli91] G.J. Klir. *Facets of Systems Theory*. Plenum Press, 1991.
- [KMS98] K. Kitamura, H. Maekawa, and C. Shimoda. Fission yeast *ste9*, a homolog of *hct1/cdh1* and *fizzy*-related, is a novel negative regulator of cell cycle progression during G_1 -phase. *Mol. Biol. Cell*, 9(5):1065–1080, May 1998.
- [Koh01] K.W. Kohn. Molecular interaction maps as information organizers and simulation guides. *Chaos*, 11(1):84–97, 2001.
- [Kol00] W. Kolch. Meaningful relationships: the regulation of the Ras/Raf/MEK/ERK pathway by protein interactions. *Biochem. J.*, 351:289–305, 2000.
- [Kre93] E. Kreyszig. *Advanced Engineering Mathematics*. John Wiley & Sons, 1993.
- [KS01] J. Keener and J. Sneyd. *Mathematical Physiology*. Springer Verlag, 2001.
- [KSA90] H. Kacser, H.M. Sauro, and L. Acerenza. Enzyme-enzyme interactions and control analysis. 1. The case of non-additivity: Monomer-oligomer associations. *Eur. J. Biochem.*, 187(3):481–491, February 1990.
- [Lan04] A.D. Lander. A calculus of purpose. *PLoS Biol.*, 2:0712–0714, June 2004.
- [Laz02] Y. Lazebnik. Can a biologist fix a radio? - or what I learned while studying apoptosis. *Cancer Cell*, 2:179–182, September 2002.
- [LC03] S. Lorkowski and P. Cullen. *Analysing Gene Expression*. Wiley-VCH, 2003.
- [LFB89] R.A. Laskey, M.P. Fairman, and J.J. Blow. S phase of the cell cycle. *Science*, 246(4930):609–614, November 1989.
- [LL93] D.A. Lauffenburger and J.J. Linderman. *Receptors: Models for binding, trafficking, and signalling*. Oxford University Press, 1993.
- [LQ03] G. Li and H. Qian. Sensitivity and specificity amplification in signal transduction. *Cell Biochem. Biophys.*, 39:45–59, 2003.

- [LS01] N. Le Novere and T.S. Shimizu. STOCHSIM: modelling of stochastic biomolecular processes. *Bioinformatics*, 17(6):575, June 2001.
- [LWB⁺91] K. Lundgren, N. Walworth, R. Booher, M. Dembski, M. Kirschner, and D. Beach. mik1 and wee1 cooperate in the inhibitory tyrosine phosphorylation of cdc2. *Cell*, 64(6):1111–1122, March 1991.
- [MA97] H.H. McAdams and A.P. Arkin. Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. USA*, 94:814–819, 1997.
- [Mag97] B. Magee. *The Philosophy of Schopenhauer*. Clarendon Press, 1997.
- [MCLM96] C. Martin-Castellanos, K. Labib, and S. Moreno. B-type cyclins regulate g1 progression in fission yeast in opposition to the p25^rum1 cdk inhibitor. *EMBO J.*, 15(4):839–849, February 1996.
- [Mes68] M.D. Mesarović. Systems theory and biology - view of a theoretician. In M.D. Mesarović et al., editors, *System Theory and Biology*, volume 351, pages 59–87. Springer-Verlag, 1968.
- [MHTSK90] E. Melendez-Hevia, N.V. Torres, J. Silica, and H. Kacser. Control analysis of transition times in metabolic systems. *Biochem. J.*, 265(1):195–202, January 1990.
- [MK89] J.R. McIntosh and M.P. Koonce. Mitosis. *Science*, 246(4930):622–628, November 1989.
- [MM13] L. Michaelis and M.L. Menten. Die Kinetik der Invertinwirkung. *Biochemische Zeitschrift*, 19:333–369, 1913.
- [MMP00] M.T. Madigan, J.M. Martinko, and J. Parker. *Brock: Biology of Microorganisms*. Prentice Hall, ninth edition, 2000.
- [MN94] S. Moreno and P. Nurse. Regulation of progression through the g1 phase of the cell cycle by the rum1⁺ gene. *Nature*, 367:236–242, 1994.
- [MR92] J.B.A. Millar and P. Russell. The cdc25 m-phase inducer: An unconventional protein phosphatase. *Cell*, 68(3):407–410, February 1992.
- [MS99] D.D. Mooney and R.J. Swift. *A Course in Mathematical Modeling*. The Mathematical Association of America, 1999.
- [MT75] M. D. Mesarović and Y. Takahara. *General Systems Theory: Mathematical Foundations*. Academic Press., 1975.
- [Mur02] J.D. Murray. *Mathematical Biology*. Springer Verlag, third edition, 2002.
- [MVHA99] C.K. Mathews, K.E. Van Holde, and K.G. Ahern. *Biochemistry*. Benjamin-Cummings, third edition, 1999.
- [Nas96a] K. Nasmyth. At the heart of the budding yeast cell cycle. *Trends Genet.*, 12(10):405–412, October 1996.
- [Nas96b] K. Nasmyth. Putting the cell cycle in order. *Science*, 274(5293):1643–1645, October 1996.
- [NCNB⁺98] B. Novak, A. Csikasz-Nagy, Györfy B., K. Nasmyth, and J.J. Tyson. Model scenarios for evolution of the eukaryotic cell cycle. *Philosophical Transactions of the Royal Society London B*, 353:2063–2076, 1998.

- [NPCT01] B. Novak, Z. Pataki, A. Ciliberto, and J.J. Tyson. Mathematical model of the cell division cycle of fission yeast. *Chaos*, 11(1):No, March 2001.
- [Nur75] P. Nurse. Genetic control of cell size at cell division in yeast. *Nature*, 256(5518):547–551, August 1975.
- [Nur90] P. Nurse. Universal control mechanism regulating onset of m-phase. *Nature*, 344(6266):503–508, April 1990.
- [Nv90] H. Nijmeijer and A.J. van der Schaft. *Nonlinear Dynamical Control Systems*. Springer, 1990.
- [P⁺01] W.K. Purves et al. *Life: The Science of Biology*. Sinauer Associates, sixth edition, 2001.
- [Pau04] J. Paulsson. Summing up the noise. *Nature*, 427:415–418, January 2004.
- [Pho97] D. Phoenix. *Introductory Mathematics for the Life Sciences*. Taylor and Francis, 1997.
- [Pol01] G.H. Pollack. *Cells, Gels and the Engines of Life*. Ebner and Sons Pub, 2001.
- [PTVF93] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C*. Cambridge University Press, second edition, 1993.
- [QK02] G.P. Quinn and M.J. Keough. *Experimental Design and Data Analysis for Biologists*. Cambridge University Press, 2002.
- [RA03] C.V. Rao and A.P. Arkin. Stochastic chemical kinetics and the quasi-steady-state assumption: Applications to the gillespie algorithm. *J. Chem. Phys.*, 118:4999–5010, March 2003.
- [Red88] C. Reder. Metabolic control theory: A structural approach. *J. Theor. Biol.*, 135:175–201, 1988.
- [RN87] P. Russell and P. Nurse. Negative regulation of mitosis by *wee1+*, a gene encoding a protein kinase homolog. *Cell*, 49(4):559–567, May 1987.
- [Row94] G.W. Rowe. *Theoretical Models in Biology*. Oxford University Press, 1994.
- [RS02] A. Regev and E. Shapiro. Cells as computation. *Nature*, 419:343, 2002.
- [Rub75] S.I. Rubinow. *Introduction to Mathematical Biology*. John Wiley, 1975.
- [RWA02] C.V. Rao, D.M. Wolf, and A.P. Arkin. Control, exploitation and tolerance of intracellular noise. *Nature*, 420:231–237, 2002.
- [Sas99] S. Sastry. *Nonlinear Systems*. Springer Verlag, 1999.
- [Sav69a] Michael A. Savageau. Biochemical systems analysis: I. some mathematical properties of the rate law for the component enzymatic reactions. *J. Theor. Biol.*, 25:365–369, 1969.
- [Sav69b] Michael A. Savageau. Biochemical systems analysis: II. steady state solutions for an n- poll system using a power-law approximation. *J. Theor. Biol.*, 25:370–379, 1969.
- [Sav70] M. A. Savageau. Biochemical systems analysis: III. dynamic solutions using a power-law approximation . *J. Theor. Biol.*, 26:215–226, 1970.

- [SBT02] L. Stryer, J.M. Berg, and J.L. Tymoczko. *Biochemistry*. W H Freeman, 2002.
- [SDP⁺73] K. Schwetlik, H. Dunken, G. Pretzschner, K. Scherzer, and H.-J. Tiller. *Chemische Kinetik*, volume LB 6 of *Lehrwerk Chemie*. VEB Deutscher Verlag für Grundstoffindustrie, Leipzig, 1973. in German.
- [Seg84] L.A. Segel. *Modeling dynamic phenomena in molecular and cellular biology*. Cambridge University Press, 1984.
- [Seg93] I.H. Segel. *Enzyme Kinetics*. John Wiley and Sons, 1993.
- [SF89] J.R. Small and D.A. Fell. The matrix method of metabolic control analysis: its validity for complex pathway structures. *J. Theor. Biol.*, 136:181–197, 1989.
- [SJ04] H. Schmidt and E.W. Jacobsen. Linear systems approach to analysis of complex dynamic behaviours in biochemical networks. *IEE Proc. Syst. Biol.*, 1(1):149–158, June 2004.
- [SK90] H.M. Sauro and H. Kacser. Enzyme-enzyme interactions and control analysis. 2. The case of non-independence: heterologous associations. *Eur. J. Biochem.*, 187(3), February 1990.
- [SK04] E. Sontag, E. Kiyatkin and B.N. Kholodenko. Inferring dynamic architecture of cellular networks using time series of gene expression, protein and metabolite data. *Bioinformatics*, 20:1877–1886, 2004.
- [SM97] S. Schnell and C. Mendoza. Closed form solution for time-dependent enzyme kinetics. *J. Theor. Biol.*, 187(2):207–212, July 1997.
- [SMT⁺03] I. Swameye, T.G. Müller, J. Timmer, O. Sandra, and U. Klingmüller. Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by data-based modeling. *Proc. Natl. Acad. Sci. USA*, 100(3):1028–1033, February 2003.
- [SN96] B. Stern and P. Nurse. A quantitative model for the cdc2 control of s phase and mitosis in fission yeast. *Trends Genet.*, 12(9):345–350, September 1996.
- [SNM99] A. Sveiczler, B. Novak, and J.M. Mitchison. Mitotic control in the absence of cdc25 mitotic inducer in fission yeast. *J. Cell Sci.*, 112(7):1085–1092, April 1999.
- [Son98] E.D. Sontag. *Mathematical Control Theory*. Springer, 1998.
- [SSF87] H.M. Sauro, J.R. Small, and D.A. Fell. Metabolic control and its analysis - extensions to the theory and matrix method. *Eur. J. Biochem.*, 165:215–221, 1987.
- [Str00a] S.H. Strogatz. From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D*, 143:1–20, 2000.
- [Str00b] S.H. Strogatz. *Nonlinear Dynamics and Chaos*. Westview Press, 2000.
- [TB04] A. Tözeren and S.W. Byers. *New Biology for Engineers and Computer Scientists*. Pearson Prentice Hall, 2004.

- [TCN03] J.J. Tyson, K.C. Chen, and B. Novak. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr. Opin. Cell Biol.*, 15(2):221–231, April 2003.
- [TD90] R. Thomas and R. D’Ari. *Biological Feedback*. CRC Press, 1990.
- [TN01] J.J. Tyson and B. Novak. Regulation of the eucaryotic cell cycle: Molecular antagonism, hysteresis, and irreversible transitions. *J. Theor. Biol.*, 210:249–263, 2001.
- [TO78] J.J. Tyson and H.G. Othmer. The dynamics of feedback control circuits in biochemical pathways. *Progress in Theoretical Biology*, 5:1–62, 1978.
- [Ton90] H. Tong. *Non-linear Time Series*. Oxford University Press, 1990.
- [Tys79] J.J. Tyson. Periodic enzyme synthesis: Reconsideration of the theory of oscillatory repression. *J. Theor. Biol.*, 80:27–38, 1979.
- [Tys83] J.J. Tyson. Periodic enzyme synthesis and oscillatory repression: Why is the period of oscillation close to the cell cycle time? *J. Theor. Biol.*, 103:313–328, 1983.
- [Tys01] J.J. Tyson. Network dynamics and cell physiology. *Nat. Rev. Mol. Cell Biol.*, 2:908–916, 2001.
- [vGK01] C. van Gend and U. Kummer. STODE - automatic stochastic simulation of systems described by differential equations. In Yi and Hucka, editors, *Proceedings of the 2nd International Conference on Systems Biology*, pages 326–333, Pasadena, 2001. Omnipress.
- [vK92] N.G. van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland, 1992.
- [Voi00] E.O. Voit. *Computational Analysis of Biochemical Systems*. Cambridge University Press, 2000.
- [WC84] H.V. Westerhoff and Y.-D. Chen. How do enzyme activities control metabolite concentrations? an additional theorem in the theory of metabolic control. *Eur. J. Biochem.*, 142:425–430, 1984.
- [WGW84] H.V. Westerhoff, A.K. Groen, and R.J.A. Wanders. Modern theories of metabolic control and their application. *Biosci. Rep.*, 4(1):1–22, January 1984.
- [WH00] J. Wolf and R. Heinrich. Effect of cellular interaction on glycolytic oscillations in yeast: a theoretical investigation. *Biochem. J.*, 345:321–334, 2000.
- [Wig03] S. Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. Springer Verlag, 2003.
- [WW00] K. Wilson and J. Walker. *Principles and Techniques of Practical Biochemistry*. Cambridge University Press, fifth edition, 2000.
- [XW03] X.-Q. Xia and M.J. Wise. DiMSim: A discrete-event simulator of metabolic networks. *J. Chem. Inf. Comput. Sci.*, 43:1011–1019, 2003.
- [YHS00] T.M. Yi, Y. Huang, and J. Simon, M.I. Dolye. Robust perfect adaption in bacterial chemotaxis through integral feedback control. *Proc. Natl. Acad. Sci. USA*, 97:4649–4653, 2000.

- [YJM⁺00] K. Yeung, P. Janosch, B. McFerran, D. W. Rose, H. Mischak, J. M. Sedivy, and W. Kolch. Mechanism of suppression of the Raf/MEK/extracellular signal-regulated kinase pathway by the Raf kinase inhibitor protein. *Mol. Cell Biol.*, 20:3079–3085, May 2000.
- [YMO97] S. Yamaguchi, H. Murakami, and H. Okayama. A wd repeat protein controls the cell cycle and differentiation by negatively regulating cdc2/b-type cyclin complexes. *Mol. Biol. Cell*, 8(12):2475–2486, December 1997.
- [YON00] S. Yamaguchi, H. Okayama, and P. Nurse. Fission yeast fizzy-related protein srw1p is a g₁-specific promoter of mitotic cyclin b degradation. *EMBO J.*, 19(15):3968–3997, August 2000.
- [YRD⁺01] K. C. Yeung, D. W. Rose, A. S. Dhillon, D. Yaros, M. Gustafsson, D. Chatterjee, B. McFerran, J. Wyche, W. Kolch, and J. M. Sedivy. Raf kinase inhibitor protein interacts with NF- κ B-inducing kinase and TAK1 and inhibits NF- κ B activation. *Mol. Cell Biol.*, 21:7207–7217, November 2001.
- [YSL⁺99] K. Yeung, T. Seitz, S. Li, P. Janosch, B. McFerran, C. Kaiser, F. Fee, K. D. Katsanakis, D. W. Rose, H. Mischak, J. M. Sedivy, and W. Kolch. Suppression of Raf-1 kinase activity and MAP kinase signalling by RKIP. *Nature*, 401:173–177, September 1999.