Springer Series in Synergetics

Springer : COMPLEXITY

Péter Érdi
Gábor Lente

Stochastic Chemical Kinetics

Theory and (Mostly) Systems Biological Applications

Springer

Springer Complexity

Springer Complexity is an interdisciplinary program publishing the best research and academic-level teaching on both fundamental and applied aspects of complex systems – cutting across all traditional disciplines of the natural and life sciences, engineering, economics, medicine, neuroscience, social and computer science.

Complex Systems are systems that comprise many interacting parts with the ability to generate a new quality of macroscopic collective behavior the manifestations of which are the spontaneous formation of distinctive temporal, spatial or functional structures. Models of such systems can be successfully mapped onto quite diverse "real-life" situations like the climate, the coherent emission of light from lasers, chemical reaction-diffusion systems, biological cellular networks, the dynamics of stock markets and of the internet, earthquake statistics and prediction, freeway traffic, the human brain, or the formation of opinions in social systems, to name just some of the popular applications.

Although their scope and methodologies overlap somewhat, one can distinguish the following main concepts and tools: self-organization, nonlinear dynamics, synergetics, turbulence, dynamical systems, catastrophes, instabilities, stochastic processes, chaos, graphs and networks, cellular automata, adaptive systems, genetic algorithms and computational intelligence.

The three major book publication platforms of the Springer Complexity program are the monograph series "Understanding Complex Systems" focusing on the various applications of complexity, the "Springer Series in Synergetics", which is devoted to the quantitative theoretical and methodological foundations, and the "SpringerBriefs in Complexity" which are concise and topical working reports, case-studies, surveys, essays and lecture notes of relevance to the field. In addition to the books in these two core series, the program also incorporates individual titles ranging from textbooks to major reference works.

Springer Series in Synergetics

Founding Editor: H. Haken

The Springer Series in Synergetics was founded by Herman Haken in 1977. Since then, the series has evolved into a substantial reference library for the quantitative, theoretical and methodological foundations of the science of complex systems.

Through many enduring classic texts, such as Haken's *Synergetics and Information and Self-Organization*, Gardiner's *Handbook of Stochastic Methods*, Risken's *The Fokker Planck-Equation* or Haake's *Quantum Signatures of Chaos*, the series has made, and continues to make, important contributions to shaping the foundations of the field.

The series publishes monographs and graduate-level textbooks of broad and general interest, with a pronounced emphasis on the physico-mathematical approach.

For further volumes:
http://www.springer.com/series/712

Péter Érdi • Gábor Lente

Stochastic Chemical Kinetics

Theory and (Mostly) Systems Biological Applications

Springer

Péter Érdi
Center for Complex Systems Studies
Kalamazoo College
Kalamazoo, MI, USA

Wigner Research Centre for Physics
Hungarian Academy of Sciences
Budapest, Hungary

Gábor Lente
Department of Inorganic and Analytical Chemistry
University of Debrecen
Debrecen, Hungary

ISSN 0172-7389
ISBN 978-1-4939-0386-3 ISBN 978-1-4939-0387-0 (eBook)
DOI 10.1007/978-1-4939-0387-0
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2014936407

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To our teacher, coworker and friend,
János Tóth

Preface

This book is, of course, about stochastic chemical kinetics. Chemical kinetics is a prototype of nonlinear science, since the rate of reaction is generally a nonlinear function of the quantities of the reacting chemical components.

Stochastic models are able to describe fluctuations around some deterministic values, but also internally random processes without having any deterministic trend (as in small chemical systems). In general, stochastic kinetic models are Markovian jump processes, and the temporal evolution can be described by a Kolmogorov equation, also called the master equation:

$$\frac{dP_n}{dt} = \mathbf{A}P_n(t). \tag{1}$$

Equation (1) is a linear differential-difference equation, and the special structure of $\mathbf{A}$ is governed by the stoichiometry. In spite of the fact that Eq. (1) is linear, it can give rise to *nonlinear phenomena*. The master equation can be converted into nonlinear equations for some macroscopic variables, e.g. by calculating the expectations of concentrations or amounts of substance.

The book deals with spatially homogeneous systems (or a system of spatially homogeneous systems, such as compartmental systems), so reaction-diffusion systems and the related stochastic models based on stochastic partial differential equations and/or on Markov fields are neglected.

There is a generation gap between the two authors. The senior author worked on stochastic kinetics with his mathematician friend János Tóth, mostly in the 1970s, and the cooperation led to a book that is a quarter of a century old now (P.É. and J.T.: Mathematical Models of Chemical Reactions. Theory and applications of deterministic and stochastic models. Manchester Univ. Press., Princeton Univ. Press. 1989). János should have been a natural coauthor of the present book, too, but he has been busy writing another one (Tóth, J., Nagy, A. L., & Papp, D.: Reaction Kinetics: Exercises, Programs and Theorems. Mathematical and Computational Chemistry. New York: Springer Verlag. In preparation.) Many ideas and techniques presented

in this book reflect (we hope) his spirit, too, as we deliberately adopted parts from János's works, and we thank him for his permission to do it.

The more junior author began his involvement in this field a decade ago by re-discovering stochastic kinetics without any knowledge of the previous literature – as a hobby for himself. The primary question for him was the interpretation of the *Soai reaction*, which is connected to efforts to understand the origins of homochirality and, ultimately, life on Earth. His wife, who was his single-member audience at that time, convinced him to try to publish his thoughts and results. The response to his first article from other scientists was sufficiently enthusiastic to keep him busy thinking about stochastic kinetics and writing scientific papers – although his main professional field is still experimental.

The development of the experimental techniques implied the much more extensive application of stochastic models, and we felt we should write a book now, which tries to be a bridge between the theory-induced pioneer period and the present and future somewhat more application-oriented times. Obviously, stochastic kinetics has an increasing popularity, mostly due to the renaissance of systems biology, as it is reflected in the subtitle of the book.

Chapter 1 is a light introduction to the fluctuation phenomena and to the most frequently used concepts of stochastic processes and stochastic kinetics. The scope and limits of the applicability of the deterministic model is discussed. Stochastic modeling grew up from the studies of fluctuation phenomena, particularly of the Brownian motion, famously studied by Einstein. His studies led to the first formulation of the fluctuation-dissipation theorem. Continuous time, discrete state space stochastic models are now often used to describe chemical fluctuations. Systems biology combines new experimental techniques and theoretical/computational methods containing a strong component related to the measurement, analysis and modeling of noise processes.

Chapter 2 is a more formal description of the topic. The mathematical framework of the most often used stochastic models of chemical reactions are discussed. First, a brief overview on and some classification of the stochastic (mostly Markovian) processes is given. The standard stochastic model of homogeneous reaction kinetics is defined, and the construction leads to the most extensively used master equations. The analogies between the deterministic and stochastic models are analyzed, among others, with the concept of the stochastic map. The different methods of obtaining transient and stationary solutions, and then the simulation techniques are reviewed. The deterministic continuation and the continuous state approximation are considered, and finally, a brief hint on the non-Markovian approximation is given.

Chapter 3 reviews the most important applications of stochastic kinetic models. Fluctuations particularly cannot be neglected in small systems and around unstable stationary points. Compartmental systems and enzyme kinetics are popular fields of stochastic kinetics, autocatalytic systems are somewhat neglected despite their historical role. Other fields of systems biology (and related areas), as signal processing, gene expression and chiral symmetry, also convincingly show the necessity of applying stochastic models. After two technical subsections (parameter

estimation and stochastic resonance), the application of stochastic kinetics in the theory of computation is reviewed. Finally, Chapter 4 gives a subjective summary of what is written in the previous three chapters.

Our book is hopefully an organic sprout on the verdant existing scientific literature. Crispin Gardiner's *Stochastic Methods: A Handbook for the Natural and Social Sciences* is an excellent resource for learning (and teaching) concepts, methods and applications of stochastic processes. Peter Schuster has an extremely good textbook on the net (*Stochasticity in Chemistry and Biology*. *When Small Population Sizes Matter and Environments Fluctuate*), Darren Wilkinson's *Stochastic Modelling for Systems Biology* has its second edition, and it is very well usable to learn simulation methods and statistical inference techniques. The literature is now rapidly growing, and we might have overlooked very important items. It is not necessarily reflected in the references, but many papers of Hong Qian were read, and our way of thinking is hopefully not too far from the spirit we spelled out from them. We deliberately adopted text, figures, and ideas from the scientific works of other colleagues. We think precise citation/credit was given.

We benefited from having a wonderful working environment. Kalamazoo College was awarded by a Henry R. Luce Professorship and one of us (P.É.) has had the privilege to serve here to build a program about Complex Systems. The Wigner Research Centre for Physics of the Hungarian Academy of Sciences in Budapest also provides a supportive environment when he spends the summers there. He also benefited from spending the Michaelmas term of 2012 as a fellow of the Institute of Advanced Studies at Durham University. G.L. is expecting to be a full professor at the Department of Inorganic and Analytical Chemistry of the University of Debrecen in Hungary soon. This environment constantly reminds him of the fact that the primary role of scientific theories is to interpret experimental data.

Budapest, Hungary/Debrecen, Hungary/Kalamazoo, USA — Péter Érdi
December 2013 — Gábor Lente

Contents

Chapter 1
Stochastic Kinetics: Why and How?

1.1 Chemical Kinetics: A Prototype of Nonlinear Science

The kinetic behavior of a chemical reaction is traditionally described by a system of (generally) nonlinear differential equations. The unknowns in these equations are the concentrations of the species present in the system. Let A_1, A_2, ..., A_n be the chemical species (usually molecules or ions) in an n-component system, $[A_1], [A_2], \ldots, [A_n]$ their concentrations given as a continuous function of time, and $\mathbf{c}(t)$ is the n-dimensional vector (the **state** of the system) composed of scalars $[A_1], [A_2], \ldots, [A_n]$. In this case, the differential equation describing the concentration changes, commonly called the **rate equation** or **rate law**, takes the following general form:

$$\frac{d\mathbf{c}(t)}{dt} = \mathbf{f}(\mathbf{c}(t); \mathbf{k}); \quad \mathbf{c}(0) = \mathbf{c}_0 \tag{1.1}$$

where $\mathbf{f}$ is the function which governs the temporal evolution of the system, $\mathbf{k}$ is the vector of the parameters (**rate constants** or **rate coefficients**) and $\mathbf{c}_0$ (with elements $[A_1]_0, [A_2]_0, \ldots, [A_n]_0$) is the initial value vector of the component concentrations.

It is to be noted that rate equation (1.1) assumes a **homogeneous** system, which means that the intensive physical properties within the reactor do not depend on the spatial coordinates. The description of such a system is independent of the values of extensive physical properties, most significantly the volume of the reactor. As a rule, an initially homogeneous system will conserve homogeneity unless it is under a special, direction-dependent external influence.[1]

[1] Famously, Turing constructed a model of a reaction – diffusion system [70] in which there exists a stable homogeneous stationary state losing its stability as a result of inhomogeneous perturbations. It was shown a few decade later [68, 69] that the presence of cross-inhibition (i.e. $(\partial f_i / \partial x_j)(c) < 0$) is a necessary condition of Turing instability. This result implies that the presence of higher than first order reactions is a necessary condition of Turing instability.

P. Érdi and G. Lente, *Stochastic Chemical Kinetics: Theory and (Mostly) Systems Biological Applications*, Springer Series in Synergetics, DOI 10.1007/978-1-4939-0387-0_1,

Deterministic models of chemical reactions might be identified with Eq. (1.1). Function **f** typically has a very special structure related to the individual chemical reaction steps possible in the system. Therefore, not all kinds of systems of differential equations (not even all those with a polynomial right-hand side) can be considered as reaction kinetic equations. Trivially, the term $-kc_2(t)c_3(t)$ cannot occur in a rate equations referring to the rate of c_1 since the quantity of a component cannot be reduced in a reaction in which the component in question does not take part. Putting it another way, the **negative cross-effect** is excluded. Chemical kinetic equations always have unique solutions, which sometimes can be deduced analytically, but more often is only available using approximate numerical methods of integration.

An overall chemical reaction is understood to consist of a finite number of individual reaction steps. Chemical reaction steps possible in the system are represented by stoichiometric equations, which have the following form:

$$0 = \sum_{i=1}^{n} \nu_{j,i} \mathrm{A}_i \quad (j = 1, 2, \ldots, m) \tag{1.2}$$

where m is the number of different reactions. The value $\nu_{j,i}$, called the stoichiometric coefficient of component A_i in reaction step j, is positive for species that are produced, negative for species that are consumed, and 0 for species that do not appear in the reaction step j. It is a long established custom to use integers only as stoichiometric coefficients, with 1 as their greatest common factor in any reaction (a given value of j in Eq. (1.2)). This convention is especially important in mass action type kinetics, and also in the stochastic equivalent of deterministic kinetics. The matrix composed of the stoichiometric coefficients is called the **stoichiometric matrix** of the system:

$$\nu = \begin{pmatrix} \nu_{1,1} & \nu_{1,2} & \cdots & \nu_{1,n} \\ \nu_{2,1} & \nu_{2,2} & \cdots & \nu_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \nu_{m,1} & \nu_{m,2} & \cdots & \nu_{m,n} \end{pmatrix} \tag{1.3}$$

The stoichiometric equations can very often be used to determine the number of independent components and the number of independent elementary reactions. It is possible to show that certain linear combinations of concentrations $[\mathrm{A}_1]$, $[\mathrm{A}_2]$, ..., $[\mathrm{A}_n]$ are constant (i.e. do not depend on time, only on the initial conditions), which leads to a reduction of the number of the differential equations. A systematic approach to determine the minimal number of independent reactions was given in the now classical papers of Rutherford Aris [1–3], see also [7].[2]

[2]Chapter 3: stoichiometry: the algebraic structure of complex chemical reactions of [19] still seems to be a good overview.

It is also quite customary to represent stoichiometric equations using a reaction arrow, left of which are given the substances with negative stoichiometric coefficients (reactants), and substances with positive stoichiometric coefficients (products) are displayed on the right. For example, the stoichiometric equation $0 = -\mathrm{A}_1 - 3\mathrm{A}_2 + 2\mathrm{A}_3$ is given as $\mathrm{A}_1 + 3\mathrm{A}_2 \longrightarrow 2\mathrm{A}_3$ in this formalism.

At this point, an interesting and often neglected point should also be made. Despite the fact that textbooks define the concept of the rate of reaction based on the rates of concentration changes, this definition should be limited to chemical reactions that can be represented by a single reaction step (a careful reading of IUPAC recommendations also reveals this restriction). In other words, the rate of reactions should generally be a vector whose dimension is determined by the number of reactions steps. The rate of an individual reaction step (v_j) can be defined without problems. The constituent functions of **f** can be given as the sum of the rates of the individual steps, taking the stoichiometric coefficients into account:

$$f_i(\mathbf{c}) = \sum_{j=1}^{m} \nu_{j,i} v_j \quad (i = 1, 2, \ldots, n) \tag{1.4}$$

The rates of individual reaction steps (v_j) depend on some of the elements of the **k** vector as parameters. Typically, different rates have different parameters, but symmetry laws may results in different reaction steps having identical rate constants.

1.1.1 The Power Law and Mass Action Type Deterministic Model of Homogeneous Reaction Kinetics

Many real systems adhere to **power law kinetics**, which means that rates v_j can be obtained by multiplying the concentrations raised to a suitable power:

$$v_j = k_j \prod_{k=1}^{n} [\mathrm{A}_k]^{\alpha_{j,k}} \tag{1.5}$$

In this case, the constituent functions of **f** can be given as:

$$f_i(\mathbf{c}) = \sum_{j=1}^{m} \nu_{j,i} k_j \prod_{k=1}^{n} [\mathrm{A}_k]^{\alpha_{j,k}} \quad (i = 1, 2, \ldots, n) \tag{1.6}$$

The $\alpha_{j,k}$ values are often, but not necessarily, integer[3] and are called the order of reaction step j with respect to substance A_k. The sum $\sum_{k=1}^{n} \alpha_{j,k}$ is called the overall order of reaction step j. This overall order has special significance in determining the physical unit of the scalar k_j, which is referred to as the rate constant of reaction step j, and is an element of the parameter vector **k** in Eq. (1.1). In general, values of $\alpha_{j,k}$ cannot be deduced from stoichiometric coefficients and can be built into a separate **order matrix**:

$$\alpha = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m,1} & \alpha_{m,2} & \cdots & \alpha_{m,n} \end{pmatrix} \tag{1.7}$$

At this point, it should be pointed out that experiments in chemical kinetics are typically carried out in a way that the dependence of parameter **k** on external conditions (most prominently on temperature) can be neglected during a single experiment, so scalars k_j are constants as far as the solution of Eq. (1.1) is concerned. Should this not be the case, it is still very common to keep the notation of power law kinetics and introduce additional differential equations describing the time dependence of the parameters. It should be stated that such equations typically preserve the autonomous property of differential equation (1.1): time never appears as an explicit variable, only through the concentrations or the values of external parameters. The practically rare case of a system with non-constant volume can be handled in an analogous manner, a special difficulty arises here because concentration change not only in chemical reactions, but also as a result of the volume change.

Within power law kinetics, **mass action type kinetics**[4] has very special importance. It is characterized by the fact that the order matrix can be determined from the stoichiometric matrix using the following simple rule:

$$\begin{aligned} \alpha_{j,i} &= -\nu_{j,i} \quad \text{if} \quad \nu_{j,i} < 0 \\ \alpha_{j,i} &= 0 \quad \text{if} \quad \nu_{j,i} \geq 0 \end{aligned} \tag{1.8}$$

[3] Savageau proposed the power law approximation for systems with non-ideal kinetics [62–64]. A highly cited paper for the breakdown of the mass action, and of the use of fractal kinetics (and also of kinetics with time-dependent rate constants) is [35]; see also [65]. In the power-law approach, rather than introducing a time dependence to the rate constants of second- and higher-order reactions, the reactant concentrations are raised to non-integer powers.

[4] The theory of *formal reaction kinetics*, called also as *chemical reaction network theory*, found beautiful relationships between the structure of the reaction network and its dynamic behavior. We can point here to the pioneer papers only [21, 31].

It is a very basic tenet of chemical kinetics that observations in all chemical reactions can be described by special mass action type kinetics called a series of elementary reactions; this property will be referred to as **reducibility** in this text. Unfortunately, the concept of an elementary reaction is not very clearly defined in a mathematical sense. A necessary but not sufficient condition for a mass action type system to qualify as a series of elementary reactions is that all stoichiometric coefficients are 0, ± 1, or ± 2 (possibly ± 3 in very exceptional cases) and none of the reaction steps have an overall reaction order greater than 2 (3 is possible again as a rare exception). If observations can be described by a rate law that cannot represent a series of elementary reactions, this fact is usually taken to imply that not all components or reactions have been correctly identified, and a more complete (and doubtlessly more complicated) description is possible by taking into account more elementary reactions. However, experimental data often do not allow such identifications or the determination of all rate constants. In these cases, it is entirely up to the judgment of the experimenter to decide whether finding a suitable series of elementary reactions is necessary for research purposes or the simpler, but theoretically incomplete description serves the objectives better.

A very common notation used to condense the stoichiometric and kinetic information of a reaction step with a power rate law is to write the following chemical equation:

$$\sum_{i=1}^{n} \alpha_{j,i} \mathrm{A}_i \longrightarrow \sum_{i=1}^{n} (\alpha_{j,i} + \nu_{j,i}) \mathrm{A}_i \tag{1.9}$$

In a mathematical sense, this sort of notation is limited to cases when all $\alpha_{j,i}$ and $(\alpha_{j,i} + \nu_{j,i})$ are non-negative. However, from a practical point of view, this is not much of a limitation as most known processes (e.g. all reactions with mass action type kinetics) satisfy this criterion. Because of its brevity, this notation is more popular than giving separate stoichiometric and order matrices.

Further limitations from physical and chemical laws apply for differential equation (1.1). Because of the property of reducibility and the difficulty in defining elementary reactions, the mathematical consequences of these limitations are most practically stated in terms of power law kinetics. One obvious limitation is that concentrations should remain non-negative at any reaction time. A sufficient but not necessary condition for the non-negativity of concentrations is that $\alpha_{j,i} > 0$ should hold for any pair of (i, j) values for which $\nu_{j,i} < 0$. Mass action kinetics not only satisfies this necessary condition, but also guarantees that all component concentrations remain positive (i.e. they cannot reach the value of zero at finite times).

Another set of limitations are imposed by the law of mass conservation. These can often be deduced from the stoichiometric equations in closed systems (i.e. those which cannot exchange particles with the surroundings). In an open system, the effect of in- and outflow is often most conveniently described as

reactions which have no reactant (for inflow) or product (for outflow). The notation $\varnothing$ is often used in these cases. It should not be left without notice that certain conservation laws apply for open systems as well, but their mathematical formulation may be much more difficult than in closed systems.

A more special set of limitations is given by the principle of detailed balance. These posit that for each stoichiometric reaction (1.9), the model must also contain the exact reverse reaction as well ("*microscopic reversibility*"):

$$0 = \sum_{i=1}^{n} -\nu_{r(j),i} \mathrm{A}_i \tag{1.10}$$

where $r(j)$ is a function giving the number of the stoichiometric equation corresponding to the reverse of stoichiometric equation j. Furthermore, the principle of detailed balance also requires a very specific relationship between $\alpha_{j,i}$ and $\alpha_{r(j),i}$ values:

$$\nu_{j,i} = \alpha_{r(j),i} - \alpha_{j,i} \tag{1.11}$$

Finally, the ratio of the values of the rate constants of the forward and reverse steps, $k_j / k_{r(j)}$, should be equal to the equilibrium constant of the process, that can be obtained from measurements independently of the kinetic studies. However, the required values of reverse rate constants are often so low so that they have no experimentally detectable consequences. It is, therefore, very common to deal with rate laws that do not adhere explicitly to the principle of detailed balance.[5]

Different types of physical limitations apply to the values of rate constants. All rate constants have a lower limit of 0. Reaction steps with exactly 0 rate constants can be deleted from the system without any change in the results. The upper limits of the rate constant values are set by the time scale of intramolecular motion or the velocity molecules move relative to each other depending on the overall order of the reaction step.

The introduced concepts are illustrated by a few examples in the next paragraphs.

Example 1.1. Consider the simple reaction representing inflow: $\emptyset \longrightarrow \mathrm{A}_1$. Here $n = m = 1$, and $\nu_{1,1} = 1$. The mass action type induced kinetic differential equation of this reaction (with $\alpha_{1,1} = 0$) is:

$$\frac{d[\mathrm{A}_1]}{dt} = k_1 \tag{1.12}$$

[5] At the beginning of the twentieth century, Wegscheider [74] gave an example to show that in some cases, the existence of a positive equilibrium state alone does not imply the equality of all the individual forward and backward reaction rates in equilibrium, and to ensure it, some relationship should be among the rate constants. More generally, but rather vaguely, the principle of detailed balance was formulated by Fowler and Milne [24]. Necessary and sufficient conditions for detailed balancing in mass action systems was given by a champion of formal chemical kinetics, Martin Feinberg [22], whereas the relationship between detailed balance and the second law of thermodynamics was also investigated later [42]. For some applications related to biophysical kinetics see [12, 53, 54].

Taken literally, this model would predict that the concentration can grow infinitely high. This means that the validity of this rate equation must be limited to a finite time period.

Example 1.2. Consider the irreversible second order dimerization reaction: $2A_1 \longrightarrow A_2$. Here $n = 2$, $m = 1$, and $\nu = (-2, 1)$. The mass action type induced kinetic differential equation of this reaction assumes $\alpha = (2, 0)$:

$$\begin{aligned} \frac{d[A_1]}{dt} &= -2k_1[A_1]^2 \\ \frac{d[A_2]}{dt} &= k_1[A_1]^2 \end{aligned} \tag{1.13}$$

The fact that the usual definition of rates requires a coefficient 2 in the differential equation defining the concentration change of A_1 should be noted here. Including or excluding this stoichiometric coefficient is often a source of ambiguity in the literature. A time-independent linear combination of concentrations here is $[A]_1 + 2[A_2]$.

Example 1.3. Consider the irreversible half-order reaction: $A_1 \longrightarrow A_2$. Here $n = 2$, $m = 1$, $\nu = (-1, 1)$, $\alpha = (0.5, 0)$. The kinetic differential equation of this reaction is:

$$\begin{aligned} \frac{d[A_1]}{dt} &= -k_1\sqrt{[A_1]} \\ \frac{d[A_2]}{dt} &= k_1\sqrt{[A_1]} \end{aligned} \tag{1.14}$$

Whilst concentrations seldom actually reach the value of zero[6] in deterministic kinetics, this reaction features $[A_1] = 0$ after a finite time of $t = 2\sqrt{[A_1]_0}/k$, independently of the initial concentration of A_2.

Example 1.4. Consider the irreversible fist order catalytic reaction: $A_1 \longrightarrow A_2$. A_3 is assumed to be a significant substance in this system that does not appear in the stoichiometric equation, but has an effect on the rate. Here $n = 3$, $m = 1$, and $\nu = (-1, 1, 0)$. The kinetic differential equation of this reaction with $\alpha = (1, 0, 1)$ is:

$$\begin{aligned} \frac{d[A_1]}{dt} &= -k_1[A_1][A_3] \\ \frac{d[A_2]}{dt} &= k_1[A_1][A_3] \\ \frac{d[A_3]}{dt} &= 0 \end{aligned} \tag{1.15}$$

A_3 is called a **catalyst** in this reaction, as its stoichiometric coefficient is zero, but the order of reaction with respect to it is positive.

[6] Such kind of behavior, i.e. convergence to **terminal attractors** may emerge if the $|\frac{\partial f_i}{\partial x_j}| < \infty$ Lipschitz conditions are violated [76].

Example 1.5. Consider the first order irreversible autocatalytic reaction with the stoichiometric equation $\emptyset \longrightarrow A_1$. Here $n = m = 1$, $\nu_{1,1} = 1$, $\alpha_{1,1} = 1$. The kinetic differential equation of this reaction is:

$$\frac{d[A_1]}{dt} = k_1[A_1]. \tag{1.16}$$

A_1 is called an **autocatalyst** in this reaction, as the stoichiometric coefficient and the order of reaction are both positive for this component. This is an example of power law kinetics, which is not mass action type. Some confusion may arise here, though. The first order irreversible autocatalytic reaction is often represented by the equation $A_1 \longrightarrow 2A_1$ and is said to have mass action type kinetics because this form show one as the coefficient before A_1 on the left. However, the equation itself does not adhere to the formalism of stoichiometric equations, in which it is not possible to have the same substance both as a reactant and a product. The term 'mass action' itself originates from equilibrium thermodynamics, where the use of stoichiometric equations is exclusive. Therefore, it is probably more precise to adopt the usage established earlier. However, to avoid superfluous confusions, we will use the conventional notation.

Example 1.6. Consider the second order (or: quadratic) autocatalytic reaction with the stoichiometric equation $\emptyset \longrightarrow A_1$. Here $n = m = 1$, $\nu_{1,1} = 1$, $\alpha_{1,1} = 2$. The kinetic differential equation of this reaction is:

$$\frac{d[A_1]}{dt} = k_1[A_1]^2. \tag{1.17}$$

Similarly to the previous example, the equation is also often given in the $2A_1 \longrightarrow 3A_1$ form.[7] This system has the property that the concentration approaches infinity at a finite time [13].

Example 1.7. Consider the reversible reaction between adduct formation between reactants A_1 and A_2 to give product A_3. The two reaction in this scheme are $A_1 + A_2 \longrightarrow A_3$ and its reverse $A_3 \longrightarrow A_1 + A_2$. Here $n = 3$, $m = 2$, and the matrices ν, α are given as:

$$\nu = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix} \quad \alpha = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.18}$$

[7] The equation itself implements large, "higher-than-linear" positive feedback (or "hyperbolic growth"), which seems to be a general mechanism behind finite time singularities. It ensures that the instantaneous doubling time tends to zero after a finite period. **Finite time singularity** roughly speaking means that a dynamical variable gets an infinite value at a finite time. This phenomenon is qualitatively different from the exponential growth, when infinite value can be attained during infinite time only.

The kinetic differential equations of this reaction scheme are:

$$\begin{aligned}\frac{d[\mathrm{A}_1]}{dt} &= -k_1[\mathrm{A}_1][\mathrm{A}_2] + k_2[\mathrm{A}_3] \\ \frac{d[\mathrm{A}_2]}{dt} &= -k_1[\mathrm{A}_1][\mathrm{A}_2] + k_2[\mathrm{A}_3] \\ \frac{d[\mathrm{A}_3]}{dt} &= k_1[\mathrm{A}_1][\mathrm{A}_2] - k_2[\mathrm{A}_3]\end{aligned} \tag{1.19}$$

This is an example of mass action type kinetics that also conforms to the principle of detailed balance. Time-independent linear combinations of concentrations are $[\mathrm{A}_1] + [\mathrm{A}_3]$ and $[\mathrm{A}_2] + [\mathrm{A}_3]$.

Example 1.8. Consider the well-known Michaelis–Menten reaction scheme composed of three processes $\mathrm{A}_1 + \mathrm{A}_2 \longrightarrow \mathrm{A}_4$, $\mathrm{A}_4 \longrightarrow \mathrm{A}_1 + \mathrm{A}_2$, and $\mathrm{A}_4 \longrightarrow \mathrm{A}_1 + \mathrm{A}_3$. The conventional meaning of the components are: A_1: enzyme, A_2: substrate, A_3: product and A_4: enzyme-substrate complex. Here $n = 4$, $m = 3$, and the matrices ν, α are given as:

$$\nu = \begin{pmatrix} -1 & -1 & 0 & 1 \\ 1 & 1 & 0 & -1 \\ 1 & 0 & 1 & -1 \end{pmatrix} \quad \alpha = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.20}$$

The kinetic differential equations of this reaction scheme are:

$$\begin{aligned}\frac{d[\mathrm{A}_1]}{dt} &= -k_1[\mathrm{A}_1][\mathrm{A}_2] + k_2[\mathrm{A}_4] + k_3[\mathrm{A}_4] \\ \frac{d[\mathrm{A}_2]}{dt} &= -k_1[\mathrm{A}_1][\mathrm{A}_2] + k_2[\mathrm{A}_4] \\ \frac{d[\mathrm{A}_3]}{dt} &= k_3[\mathrm{A}_4] \\ \frac{d[\mathrm{A}_4]}{dt} &= k_1[\mathrm{A}_1][\mathrm{A}_2] - k_2[\mathrm{A}_4] - k_3[\mathrm{A}_4]\end{aligned} \tag{1.21}$$

This is again mass action type kinetics. It should be noted that two time-independent linear combinations of concentrations can be derived in this system: $[\mathrm{A}_1] + [\mathrm{A}_4]$, and $[\mathrm{A}_2] + [\mathrm{A}_3] + [\mathrm{A}_4]$.

Example 1.9. Consider the steady state Michaelis–Menten reaction scheme described stoichiometrically as $\mathrm{A}_1 \longrightarrow \mathrm{A}_3$, and A_2 playing the role of the catalyst. It is very typical in the Michaelis–Menten scheme presented in the previous example that experimental methods cannot distinguish A_1 and A_4, therefore only one of them is included in the steady state scheme. In addition, it is also very common that the time resolution of the experimental method is not sufficient to follow the first,

adduct formation step in the full scheme given in the previous example. In this description, $n = 3$, $m = 1$, and $\nu = (-1, 1, 0)$. Typically, non-power law kinetics is used in this case as defined by the following differential equations:

$$\begin{aligned} \frac{d[\mathrm{A}_1]}{dt} &= -\frac{k_1[\mathrm{A}_1][\mathrm{A}_3]}{k_2 + [\mathrm{A}_1]} \\ \frac{d[\mathrm{A}_2]}{dt} &= \frac{k_1[\mathrm{A}_1][\mathrm{A}_3]}{k_2 + [\mathrm{A}_1]} \\ \frac{d[\mathrm{A}_3]}{dt} &= 0 \end{aligned} \tag{1.22}$$

This rate equation itself is called the Michaelis–Menten equation. Constant k_2, which has a dimension of a concentration, is called the Michaelis constant. The concept of orders of reaction cannot be defined here, as this is not power-law kinetics.

Example 1.10. Consider the Gray model [47] with second order autocatalysis: $\mathrm{A}_1 \longrightarrow \mathrm{A}_2$, $\mathrm{A}_1 \longrightarrow \emptyset$, $\emptyset \longrightarrow \mathrm{A}_1$, $\mathrm{A}_2 \longrightarrow \emptyset$, $\emptyset \longrightarrow \mathrm{A}_2$. Here $n = 2$, $m = 5$, and the matrices ν, α are given as:

$$\nu = \begin{pmatrix} -1 & 1 \\ -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{pmatrix} \quad \alpha = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{1.23}$$

The kinetic differential equations of this reaction scheme are:

$$\begin{aligned} \frac{d[\mathrm{A}_1]}{dt} &= -k_1[\mathrm{A}_1][\mathrm{A}_2] - k_2[\mathrm{A}_1] + k_3 \\ \frac{d[\mathrm{A}_2]}{dt} &= k_1[\mathrm{A}_1][\mathrm{A}_2] - k_4[\mathrm{A}_2] + k_5 \end{aligned} \tag{1.24}$$

Usually, $k_5 = 0$ is set by assuming that A_2 is not present in the feed, and the difference $k_3 - k_4$ is kept positive by assuming that there is a first-order chemical reaction consuming A_2 in addition to the outflow.

With the exception of Example 1.9, all of the ten listed examples show power law kinetics, but only Examples 1.1, 1.2, 1.7 and 1.8 adhere to mass action kinetics in the sense defined here. In addition, only Examples 1.2, 1.7 and 1.8 satisfy the necessary conditions to be a sequence of elementary reactions.

1.1.2 Stationary States and Their Stability

Returning now to the general rate equation given in Eq. (1.1), stationary conditions are said to prevail if concentrations do not change. The concentration vector $\mathbf{c}_{st}$ is a **stationary point** in the state space if it is a solution of the following equation:

$$0 = \mathbf{f}(\mathbf{c}_{st}; \mathbf{k}) \tag{1.25}$$

In effect, a stationary point is a single-point trajectory in the concentration space. If the limiting value $\lim_{t\to+\infty} \mathbf{c}(t)$ of a trajectory exists, it is a stationary point. As the solution of the differential equation (1.1) is unique, each initial condition can give rise to no more than one stationary state. The existence and uniqueness of the solutions are ensured by the Picard-Lindelöf theorem.

Stationary states are classified based on their stability. Intuitively, the stability of a stationary point implies the property that the system returns to this state after a small perturbation. The absence of this property is referred to as instability. A stationary point $\mathbf{c}_{st}$ is said to be **Lyapunov stable** if for any U neighborhood of $\mathbf{c}_{st}$, there exists another U_1 neighborhood so that all trajectories originating from U_1 remain entirely in U. A stable stationary point $\mathbf{c}_{st}$ is **asymptotically Lyapunov stable** if a suitable U_1 can be given so that all trajectories originating from U_1 tend to the stationary point as time approaches infinity, $\lim_{t\to+\infty} \mathbf{c}(t) = \mathbf{c}_{st}$. A stationary point is **Lyapunov unstable** if it is not stable. A system is **globally stable** if $\lim_{t\to+\infty} \mathbf{c}(t) = \mathbf{c}_{st}$ is true for every trajectory.

It may happen that the solutions tend to a periodic trajectory, and this behavior is related to the important class of the oscillatory reactions. In more than two dimensions, it may also happen that the trajectories remain in a bounded set, but they neither tend to an equilibrium point, nor to an oscillatory solution. Loosely speaking this behavior is called chaotic.

If a system of differential equations has multiple attractors, the phenomenon of **multistability** occurs. Specifically, if the attractors are equilibrium points, it is called **multistatonarity**.

The stability properties of deterministic stationary points have special significance for stochastic kinetics. In general, a system in an unstable stationary state is expected to show phenomena that can only be interpreted by theory containing stochastic elements, and the nature of fluctuations near instability points should be carefully studied.

1.2 Applicability of the Deterministic Model

The line of thought presented in the previous section is said to represent the deterministic approach to chemical kinetics. This means that the rate law shown in Eq. (1.1) always yields a unique $\mathbf{c}(t)$ function, which can be also called a

trajectory in the concentration space of the system. Each trajectory in this space is unambiguously identified by any of its points: the value of the concentration vector $\mathbf{c}$ at a single time point is sufficient to calculate the entire function $\mathbf{c}(t)$. If two trajectories ($\mathbf{c}_1$ and $\mathbf{c}_2$) have a common point in the concentration space at different time values so that $\mathbf{c}_1(t_1) = \mathbf{c}_2(t_2)$, then a shift in time must convert the two trajectories into each other, $\mathbf{c}_2(t) = \mathbf{c}_1(t + t_2 - t_1)$ must hold for any value of t.

If the initial state is not a single vector but known as a probability distribution of a range of states,[8] the probability distribution of states at time t can be obtained using the trajectories starting from the possible initial states. If state $\mathbf{c}(0)$ occurs with probability P at the initial time, the state $\mathbf{c}(t)$ from the state trajectory must occur with probability P after time t. This property is sometimes a useful tool that can simplify calculations in stochastic systems. Despite the appearance of probability distributions, this case should still be considered deterministic kinetics as the stochastic element is introduced by the initial conditions and not by the rate law.

A number of time-tested computational algorithms are known to solve the differential equation (1.1) even in quite general cases. These numerical methods are implemented in several commercially or even freely available software packages. Some of these programs also contain minimizing components as well, which make it possible to find the vector of parameters ($\mathbf{k}$) that best interpret an experimental data set of arbitrary size. This needs to be stated to show that determining rate constants is in fact not a central problem in modern chemical kinetics, despite some beliefs to the contrary. The fundamental problem in twenty-first century kinetics is establishing the rate laws based on experimental observations and interpret them on a molecular level.[9]

From a theoretical point of view, the major insufficiency in the deterministic model is the assumption that concentrations are continuous functions of time. Dalton's atomic theory, postulated in the beginning of the nineteenth century and generally accepted up to date, states that matter is composed of particles. Therefore, concentration values should be discrete rather than assuming any values from a continuous range. This seriously contradicts the assumptions made before setting up rate law (1.1). However, the continuous state space deterministic approach is still

[8] Starting from deterministic models, randomness can be incorporated with different implementation. First, the weakest way is to assume that the only source of the randomness is due to the lack of precise knowledge of the initial values. Second, the parameters might contain some uncertainty, so the constants can be replaced by values taken from some stationary stochastic process. Third, the governing law, i.e. the forcing function itself may contain randomness.

[9] The derivation of mass action kinetic laws from elementary principles (at least for reversible bimolecular gas-phase reactions) was given by a seminal paper of Ross and Mazur in 1961 [61]. Starting from a Boltzmann equation containing also a term due to reactive collisions, the mass action kinetic equations were derived by using the Chapman-Enskog approximation method. In this special case, macroscopic kinetics equations can be considered as the 'zeroth order approach' of the equations of non-equilibrium statistical mechanics.

a very useful one in chemical kinetics. Chemically detectable amounts of substance are usually macroscopic and contain more (often much more) than 10^{10} molecules. At these high molecule numbers, the errors caused by assuming a continuous range of possible concentrations are usually much smaller than experimental errors from a multitude of other sources. As illustrated in this book, stochastic approaches to chemical kinetics are usually much more difficult to implement computationally. It was already mentioned that standard algorithms are available for solving the rate law (1.1) no matter what its actual form is. For experimenters, it usually makes more sense to use this routine method because its systematic error will by no means limit its applicability to interpret observed data.

Yet it must be recognized that deterministic kinetics is based on the assumption of continuously changing concentrations, which leaves open the possibility that Eq. (1.1) may not give a physically acceptable model of reality well in certain exceptional cases. An obvious exceptional case is when the reactor is small and the number of molecules in it is low (e.g. $<10^4$). This is not uncommon in biological systems, and the improvement of detection techniques increasingly enables experimental studies in such systems. Furthermore, some rate laws have the inherent property that changes at very low concentrations are of profound consequence even when the molecule numbers grow to macroscopic values. A systematic study of the applicability limits of deterministic kinetics will be one of the recurring themes in this book.

Another phenomenon where deterministic kinetics is not usually useful is describing inherent fluctuations, which are known to exist independently of any other fluctuations caused by changing external conditions.

1.3 Fluctuation Phenomena

1.3.1 Brownian Motion

The Scottish botanist Robert Brown discovered the existence of fluctuations when he studied microscopic living phenomena. However, the physical nature of the motion, which was named after its discoverer, was not known for a long time. As Darwin [14] wrote in 1876: “I called on him [Brown] two or three time before the voyage of the Beagle (1831), and on the occasion he asked me to look through a microscope and describe what I saw. This I did, and believe now that it was the marvelous currents of protoplasm in some vegetable cell. I then asked him what I had seen; but he answered me, ‘This is my little secret.’ ”

Brownian motion could be well detected in colloidal solutions. The mass of the – literally microscopic – Brownian particle is much greater than the mass of the solvent molecules, and the observable motion is the result of the individual motions

of the small molecules. The theory of Brownian motion[10] was given by Einstein [17] and Smoluchowski [72]. The theory[11]

- Gave a relationship for the time-dependence of the average of the square of the displacement $X(t)$ of the Brownian particle,

$$< X(t)^2 >:=< [x(t) - x(0)]^2 >= Dt, \tag{1.26}$$

  where $x(t)$ is the actual and $x(o)$ is the initial coordinate of the Brownian particle;
- Found the connection between the mobility of the particle and the – macroscopic – diffusion constant:

$$D = \mu k_B T. \tag{1.27}$$

  Since the diffusion constant D measures the deviation from the average, and $\mu := \gamma m$, the mobility is the measure of the dissipation (γ is a damping constant, m is the mass of the Brownian particle) Eq. (1.27) is the first version of the fluctuation-dissipation relation (k_B is the Boltzmann constant, T is the temperature).
- Considered the motion of the Brownian particles as *memoryless* and non-differentiable trajectories, and helped prepared the pathway to the formulation of the theory of stochastic (actually Markov) processes.
- Offered a new method to determine the Avogadro constant.

1.3.1.1 Diffusion

The theory of Brownian motion led to the construction of a new types of equations, **stochastic differential equations**, first formulated by Paul Langevin. It implements a principle, which assumes that the forcing function has a *systematic* and *deterministic* part, and a term due to the *rapidly varying, highly irregular* random effects. In a general form it is written as

$$dx/dt = a(x,t) + b(x,t)\xi(t) \tag{1.28}$$

[10]For the early history, see Chapter 15 of the seminal book on the history of the kinetic theory of gases [8]. An excellent website for historical items and surveys related to Brownian motion is found at http://www.physik.uni-augsburg.de/theo1/hanggi/, [28].

[11]Historically, it is interesting that Brownian motion had earlier application in finance than in physics. Louis Bachelier (1870–1946) in a paper in 1900 [5] defined Brownian motion and applied it as a model for **asset price** movements. While the paper (actually a dissertation) did not gain a very high reputation after its preparation, it is qualified now as the starting point of financial mathematics.

It is well-known now that Eq. (1.28) is not precise, since $\xi(t)$ is often non-differentiable, and therefore $x(t)$ is also non-differentiable.[12]

Retrospectively, Einstein's theory considers $a(x,t) = 0$ and $b(x,t) = \sqrt{2D}$ and assuming Gaussian white noise[13] (1.28) becomes

$$dx/dt = \sqrt{2D}\xi(t) \tag{1.29}$$

or

$$dx/dt = \sqrt{2D}dW(t) \tag{1.30}$$

where $W(t)$ is the Wiener process.[14] After the integration of (1.30) the mean position and the mean square displacement can be calculated, so one finds $< x >=< x_0 >$, and $< x^2 > - < x >^2 = 2Dt$.

Langevin also took into account the friction effects on the motion of particles, and his description is for the velocity v is

$$dv/dt = -\gamma v + \sqrt{2D}dW(t). \tag{1.31}$$

This model leads to a time-dependent mean position, later called the Uhlenbeck–Ornstein process [71, 73] (see Sect. 2.8) as opposed to Einstein's model (1.30).

Both from theoretical and practical points of view, the connection between the stochastic differential equation and the evolution equation for the P probability density function (called the **Fokker–Planck equation** in physics) is very important. The solution of (1.28) under a rather general condition is a diffusion process (a special case of Markovian stochastic processes) defined by the infinitesimal generator A (for scalar case):

$$A = a\frac{\partial}{\partial x} + \frac{1}{2}b\frac{\partial^2}{\partial x^2} \tag{1.32}$$

Here $a(x,t)$ is the velocity of the conditional expectation (called "drift"), and $b(x,t)$ is a the velocity of the conditional variance (called a "diffusion constant"). The general form of a Fokker–Planck equation is:

$$\frac{dP}{dt} = AP \tag{1.33}$$

---

[12] One much discussed problem of **stochastic calculus** is the interpretation of stochastic integrals, "Ito versus Stratonovich", and we refer to Sect 4.2 of the now classical book [25].

[13] White noise is considered as a stationary Gaussian process with $E[\xi_t] = 0$ and $E[\xi_t \xi_{t'}] = \delta_{tt'}|t - t'|$, where δ is the Dirac delta function.

[14] A Wiener process $W(t)$ is a process with independent increments $(t_2) - W(t_1)$ that follow Gaussian distribution. See Sect. 2.8 for details.

Specifically, the Wiener process is defined by $a(x,t) := 0$, and $b(x,t) := \sqrt{2D}$. This equations is the well-known diffusion equation. The probability density function of finding a particle at position x at time t is governed by the diffusion equation when the microscopic motion of the particle is described by the Wiener process.

Anomalous diffusion is characterized by the deviation from the linear time dependence of the mean squared displacement. Instead, anomalous diffusion (related to the breakdown of the Gaussian and Markovian assumptions, short-range correlations) is described by $< X^2(t) > \sim t^n$, where $n < 1$ and $n > 1$ are called subdiffusion and superdiffusion, respectively, and the theory grew up from the studies of transport process (for a review, see [52]).

1.3.1.2 Fluctuation-Dissipation Theorem

The spirit of the Einstein relationship was applied to characterize noise occurring in electric circuits being in thermal equilibrium by Nyquist (theory) and Johnson (experiments) in the Bell Labs in 1928 [32, 55].[15] The internal voltage fluctuation is proportional to the resistance R, temperature T and the b bandwidth of the measurement:

$$< V^2 >= 4RK_BT \tag{1.34}$$

The fact that the same forces that cause the fluctuations also result in their dissipation was formulated as the fluctuation-dissipation theorem in seminal papers [10, 11, 38] of non-equilibrium thermodynamics.

The 'spirit' of the fluctuation dissipation theorem can be utilized in chemical kinetics, and individual rate constants can be estimated form equilibrium concentration fluctuations when from the equilibrium concentration, only the equilibrium constant (the ratio of rate constants) can be calculated. Direct measurements of concentration fluctuations started in the 1970s by using electric conductance measurements [20] and fluorescence correlation spectroscopy [45]. In the first case, the kinetic parameters of the dissociation reaction of beryllium sulfate were obtained from the analysis of the frequency spectrum of the fluctuations in the concentrations of the reactants. The latter led to the discovery of new optical methods, see e.g. [37, 57].

Fluctuation or noise phenomena have representation in the time domain and the frequency domain. Loosely speaking, at least for stationary processes (where some statistical characteristics are time-independent), the **two point autocorrelation function** can be converted by Fourier transformation to **power spectral density**.

[15] Excuse us for a little off-topic remark. The Nyquist-Johnston noise is different form the *shot noise* occurring due to the flow of the discrete nature of the flowing objects (electrons) introduced by Schottky in 1918, also in the Bell lab related to the emergence of the semiconductor industry. Off-off topic: for a very well written history of Bell Labs, see [26].

$$C(t) := E[\xi(t)\xi(t-\tau)] \tag{1.35}$$

$$S(\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} C(t)cos\omega d\tau = \frac{1}{2\pi} < \xi^2 >^{eq} \frac{\gamma(\omega)}{\gamma(\omega)+\omega^2} \tag{1.36}$$

$\gamma(\omega)$ is the Fourier transform of the dissipation constant, and $< \xi^2 >^{eq}$ characterized the measure of the equilibrium fluctuations. The relationship is famously called the Wiener–Khinchine theorem.

We will return to the topic of the fluctuation-dissipation theorem in chemical kinetics at Sect. 3.9.

1.3.1.3 Towards the Theory of Stochastic Processes

Followed (and sometimes preceded) by studies of Einstein, Langevin and Schmoluchowski, the formal definition of Brownian motion was given by Wiener [75]; Uhlenbeck and coworkers gave a generalization. Non-differentiable Brownian trajectories (fractals, using the modern language) are statistically self-similar on all scales, and were studied famously by Paul Levy [44] and his student, Benoit Mandelbrot [46]. Equations for the discrete state space jump process[16] referred as Kolmogorov equations, going back to the work [34].

1.3.1.4 Experimental Determination of the Avogadro Constant

The experimental verification of the heterogeneous nature of colloid solutions, studies on disperse and discontinuous structure of matter, and of the Einstein–Schmoluchowski theory resulted in three Nobel-prizes in chemistry in 1925–1926 (Zsigmondy, Svedberg) and physics (Perrin). Perrin improved very much the procedure to estimate the Avogadro constant by using the relation between colloidal osmotic pressure and concentration.[17]

1.4 Stochastic Chemical Kinetics

1.4.1 Model Framework: Preliminary Remarks

To describe fluctuation phenomena a continuous time, discrete state space stochastic model has to be defined. Introducing a stochastic description, let ξ be a stochastic vector process, the dimension of which is equal to the dimension of the concentration vector.

[16] "in a small time interval there is an overwhelming probability that the state will remain unchanged; however, if it changes, the change may be radical", [23].

[17] For the short history of colloidal suspensions and Brownian motion, see e.g. [30].

$$P_n(t) := P(\xi(t) = \mathbf{n}) \tag{1.37}$$

is the probability that the vector of the numbers of components is **n**. $P_n(t)$ is the time-dependent distribution function. The temporal evolution of the distribution is determined by the assumption that the chemical reaction is considered as a Markovian jump process, and the temporal evolution can be described by a Kolmogorov equation, also called the master equation:

$$\frac{dP_n}{dt} = \mathbf{A}P_n(t). \tag{1.38}$$

Equation (1.38) is a linear differential-difference equation, the special structure of **A** given by the stoichiometry. Introducing $a_{nn'}$ as the infinitesimal transition probability, which gives the probability (per unit time) of the jump from n' to n, the master equation can be interpreted as a gain-loss equation for the probability of each state **n**:

$$\frac{dP_n}{dt} = \sum n'[a_{nn'}P_n(t) - a_{n'n}P_n(t)]. \tag{1.39}$$

The first term of the right-hand side is the gain due to the transition form all the other states n', and the second term is the loss due to the jump to other states. For a more formal description of the model framework see Sect. 2.1.

1.4.2 Historical Remarks

As we know, probably Leontovich [43] was the first to investigate stochastic models of chemical reactions. Delbrück [15] studied the autocatalytic reaction (see Sect. 3.4) $A + X \longrightarrow 2X$ which is used to describe the formation of tripsin from tripsinogen. The deterministic and the standard stochastic models were formulated and the binomial distribution of the latter is approximated by a normal distribution. The distribution of the time at which a given number of particles is attained is also determined. In the same year, in 1940, Kramers [36] provided a general approximation of CDS models with CCS models. General compartmental systems have been solved by Siegert [66]. The Hungarian mathematician Alfréd Rényi [59] was the first to provide a detailed analysis of the standard stochastic model of a higher than first order reaction. He has shown that the expectation of the numbers of molecules in a stochastic model is close to the corresponding quantities in the deterministic model, and the difference is proportional to the reciprocal of the square root of the number of molecules (see also [6]). Therefore, the relative error is small, if this number is large, otherwise it is large. He also made a series of approximations of the distributions of molecule numbers under different conditions. A direct continuation of his work can be found in [40]. The author carried out the

calculations on the standard stochastic model of $A + B \longleftrightarrow C$ in detail using Laplace transforms. His results correspond to the results of [39] (see Sect. 2.3.3). For the early history of stochastic kinetics consult [48–51].

To make the long story short, theoretical calculations showed why and when the application of stochastic models is **necessary** in chemical kinetics, and the emergence of methods of measuring concentration fluctuations give the evidence that it is possible to compare theoretical predictions to experimental data.

1.4.3 On the Solutions of the Stochastic Kinetic Models: Analytical Calculations Versus Simulations

While the solution of the master equation contains all the information about the system that is required in the practice, closed form solutions can be obtained for a very restricted class of systems. In Sect. 2.4 the methods of obtaining exact or approximate solutions will be reviewed. Not being able to solve the master equation, we are often satisfied by the determination of the first and second moments. The approximation of the jump processes by continuous (such as diffusion) processes also helps, equations for the latter processes are more easily solvable. Instead of obtaining transient solutions we might be satisfied with having stationary solutions, which might be unimodal or multimodal.

Stochastic simulation is an often used method to approximately generate a **realization** of the stochastic process. By generating a large number of realizations, an approximate distribution function can be constructed. About the different algorithms, also from historical perspective, see Sect. 2.6.

1.4.4 The Renaissance of Stochastic Kinetics: Systems Biology

Systems biology combines (i) the collection and analysis of large data sets of experimental data and (ii) mathematical modeling and (iii) statistical analysis to interpret these data and to make predictions for the result of new experiments. Its main tenets are, among others, to predict phenotype from genotype, understand metabolism, cell-cell communication, cellular networks etc. The whole approach grew up from the studies of traditional biological systems (autocatalysis, Sect. 3.4, and classical enzyme kinetics Sect. 3.5).

From the perspective of their function, chemical systems can be considered as signal processing devices, and (stochastic) biochemical networks convert time-dependent inputs to time-dependent outputs. The relationship between the structure of the reaction network and the efficiency of information transfer is discussed in Sect. 3.6.2.

Many biochemical and biophysical processes are involved in gene regulation. While traditional biochemistry adopted a rather rigid deterministic scenario considering the execution of instructions encoded in the DNA, chemical reactions taking place at single cell level are now admittedly better described by stochastic models than with deterministic ones. Reactions in gene expression, such as promoter activity and inactivity, transcription, translation, and decaying of mRNA and proteins are the most important chemical steps. Measurements on stochastic gene expression in single cells with single molecule sensitivity [9, 18] implied the necessity of stochastic description [56]. Models of stochastic gene expression [58, 60] will be reviewed in Sect. 3.7.3.

Section 3.8 is dealing with symmetry breaking in chiral systems, in particular with the Frank model and the Soai reaction. Chirality, after all, is as excellent model for demonstrating the importance of fluctuations.

While a cell is a highly heterogeneous spatial system, nevertheless, modeling of spatially homogeneous systems brought reasonable results in the past, and we decided to restrict ourselves to such situations. Both theoretical frameworks [4] and stochastic simulation of reactions diffusion systems were offered quite early [29], and the latter developed further decades later (e.g. [16, 27, 33]). We expect, however, the development and applications of methods (analogous to partial differential equations as deterministic models) of analyzing spatially inhomogeneous and heterogeneous systems such as [41, 67].

References

1. Aris R (1965) Prolegomena to the rational analysis of systems of chemical reactions. Arch Ration Mech Anal 19:81–99
2. Aris R (1968) Prolegomena to the rational analysis of systems of chemical reactions. Arch Ration Mech Anal II. Some addenda 27:356–364
3. Aris R, Mah RHS (1963) Independence of chemical reactions. Ind Eng Chem Fundam 2:90–94
4. Arnold L, Theodosopulu M (1980) Deterministic limit of the stochastic model of chemical reactions with diffusion. Adv Appl Probab 12:367–379
5. Bachelier L (1900) Theorie de la Speculation. Annales Scientifiques de l'École Normale Superieure III 17:21–86
6. Barabás B, Tóth J, Pályi G (2010) Stochastic aspects of asymmetric autocatalysis and absolute asymmetric synthesis. J Math Chem 48:457–489
7. Bowen R (1968) On the stoichiometry of chemically reactiong materials. Arch Ration Mech Anal 34:97–127
8. Brush SG (1976) The kind of motion we call heat. A history of the kinetic theory of gases in the 19th century. North-Holland, Amsterdam/New York/Oxford. xxxix + 769pp (2 books)
9. Cai L, Friedman N, Xie X (2006) Stochastic protein expression in individual cells at the single molecule level. Nature 440:358–362
10. Callen HB, Greene RF (1952) On a theorem of irreversible thermodynamics I-II. Phys Rev 86:702–710; 88:1387–1391
11. Callen HB, Welton TA (1951) Irreversibility and generalized noise. Phys Rev 83:34
12. Colquhoun D, Dowsland KA, Beato M, Plested AJR (2004) How to impose microscopic reversibility in complex reaction mechanisms. Biophys J 86:3510–3518

13. Csikja R, Tóth J (2007) Blow up in polynomial differential equations, Enformatika. Int J Appl Math Comput Sci 4:728–733
14. Darwin C (1859) On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. John Murray, London
15. Delbrück M (1940) Statistical fluctuation in autocatalytic reactions. J Chem Phys 8:120–124
16. Drawert B, Lawson MJ, Petzold L, Khammash M (2010) The diffusive finite state projection algorithm for efficient simulation of the stochastic reaction-diffusion master equation. J Chem Phys 132:074101
17. Einstein A (1905) Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. Ann Phys (Leipzig) 17:549–555
18. Elowitz MB, Levine AJ, Siggia ED, Swain PS (2002) Stochastic gene expression in a single cell. Science 297:1183–1186
19. Érdi P, Tóth J (1989) Mathematical models of chemical reactions. Theory and applications of deterministic and stochastic models. Manchester University Press/Princeton University Press, Princeton
20. Feher G, Weissman M (1973) Fluctuation spectroscopy: determination of chemical reaction kinetics from the frequency spectrum of fluctuations. Proc Natl Acad Sci USA 70:870–875
21. Feinberg M (1972) On chemical kinetics of a certain class. Arch Ration Mech Anal 46:1–41
22. Feinberg M (1989) Necessary and sufficient conditions for detailed balancing in mass action systems of arbitrary complexity. Chem Eng Sci 44:1819–1827
23. Feller W (1949) On the theory of stochastic processes, with particular reference to applications. In: Proceedings of the (first) Berkeley symposium on mathematical statistics and probability, Berkeley, pp 403–432
24. Fowler RH, Milne EA (1925) A note on the principle of detailed balancing. Proc Natl Acad Sci USA 11:400–401
25. Gardiner CW (2009) Handbook of stochastic methods for physics, chemistry and the natural sciences, 4th edn. Springer, Berlin
26. Gertner J (2012) The idea factory: Bell labs and the great age of American innovation. Pinguine Press, New York
27. Gillespie DT, Hellander A, Petzold LR (2013) Perspective: stochastic algorithms for chemical kinetics. J Chem Phys 138:270901
28. Hänggi P, Marchesoni F (2005) Introduction: 100 years of Brownian motion. Chaos 15:026101
29. Hanusse P, Blanché A (1981) A Monte Carlo method for large reaction-diffusion systems. J Chem Phys 74:6148–6153
30. Haw MD (2002) Colloidal suspensions, Brownian motion, molecular reality: a short history. J Phys Condens Matter 14:7769–7779
31. Horn F, Jackson R (1972) General mass action kinetics. Arch Rat Mech Anal 47:81–116
32. Johnson J (1928) Thermal agitation of electricity in conductors. Phys Rev 32:97–109
33. Koh W, Blackwell KT (2011) An accelerated algorithm for discrete stochastic simulation of reaction? Diffusion systems using gradient-based diffusion and tau-leaping. J Chem Phys 134:154103
34. Kolmogorov A (1931) Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung (On analytical methods in the theory of probability). Math Ann 104:415–458
35. Kopelman R (1988) Fractal reaction kinetics. Science 241:1620–1626
36. Kramers HA (1940) Brownian motion in a field of force and the diffusion model of chemical reactions. Physica 7:284–304
37. Krichevsky O, Bonnet G (2002) Fluorescence correlation spectroscopy: the technique and its applications. Rep Prog Phys 65:251–297
38. Kubo R (1966) The fluctuation-dissipation theorem. Rep Prog Phys 29:255–284
39. Kurtz TG (1972) The relationship between stochastic and deterministic models for chemical reactions. J Chem Phys 57:2976–2978
40. Laurenzi IJ (2000) An analytical solution of the stochastic master equation for reversible bimolecular reaction kinetics. J Chem Phys 113:3315–3322

41. Lawson MJ, Drawert B, Khammash M, Petzold L, Yi TM (2013) Spatial stochastic dynamics enable robust cell polarization. PLoS Comp Biol 9:e1003139
42. Lente G (2010) The connection between the second law of thermodynamics and the principle of microscopic reversibility. J Math Chem 47:1106–1111
43. Leontovich MA (1935) Basic equations of kinetic gas theory from the viewpoint of the theory of random processes. J Exp Theor Phys 5:211–231
44. Lévy P (1948) Processus stochastiques et mouvement brownien. Gauthier-Villars, Paris
45. Magde D, Elson EL, Webb WW (1972) Thermodynamic fluctuations in a reacting system – measurement by fluorescence correlation spectroscopy. Phys Rev Lett 29:705–708
46. Mandelbrot BB (1982) The fractal geometry of nature. W. H. Freeman, New York
47. Marion G, Mao X, Renshaw E, Liu J (2002) Spatial heterogeneity and the stability of reaction states in autocatalysis. Phys Rev E 66:051915
48. McQuarrie DA (1963) Kinetics of small systems. I. J Chem Phys 38:433–436
49. McQuarrie DA (1967) Stochastic approach to chemical kinetics. J Appl Probab 4:413–478
50. McQuarrie DA (1968) Stochastic approach to chemical kinetics. Corrigendum. J Appl Probab 5:484–485
51. McQuarrie DA (1969) Stochastic theory of chemical rate processes. In: Shuler KE (ed) Stochastic processes in chemical physics. Wiley, New York
52. Metzler R, Klafter J (2000) The random walk's guide to anomalous diffusion: a fractional dynamics approach. Phys Rep 339:1–77
53. Nagy I, Kovács B, Tóth J (2009) Detailed balance in ion channels: applications of Feinberg's theorem. React Kinet Catal Lett 96:263–267
54. Nagy I, Tóth J (2012) Microscopic reversibility or detailed balance in ion channel models. J Math Chem 50:1179–1199
55. Nyquist, H (1928) Thermal agitation of electric charge in conductors. Phys Rev 32:110–113
56. Qian H (2013) Stochastic physics, complex systems and biology. Quant Biol 1:50–53
57. Qian H, Elsom EL (1999) Quantitative study of polymer conformation and dynamics by single-particle tracking. Biophys J 76:1598–1605
58. Paulsson J (2005) Models of stochastic gene expression. Phys Life Rev 2:157–175
59. Rényi A (1953) Kémiai reakciók tárgyalása a sztochasztikus folyamatok elmélete segítségével (in Hungarian). (Treating chemical reactions using the theory of stochastic process.) MTA Alk Mat Int Közl 2:83–101
60. Ribeiro AS (2010) Stochastic and delayed stochastic models of gene expression and regulation. Math Biosci 223:1–11
61. Ross J, Mazur P (1961) Some deductions from a formal statistical mechanical theory of chemical kinetics. J Chem Phys 35:19–28
62. Savageau MA (1969) Biochemical systems analysis. I. Some mathematical properties of the rate law for the component enzymatic reactions. J Theor Biol 25:365–369
63. Savageau MA (1976) Biochemical systems analysis: a study of function and design in molecular biology. Addison-Wesley, Reading
64. Savageau MA (1995) Michaelis-Menten mechanism reconsidered: implications of fractal kinetics. J Theor Biol 176:115–124
65. Schnell S, Turner TE (2004) Reaction kinetics in intracellular environments with macromolecular crowding: simulations and rate laws. Prog Biophys Mol Biol 85:235–260
66. Siegert JF (1949) On the approach to statistical equilibrium. Phys Rev 76:1708–1714
67. Sturrock M, Hellander A, Matzavinos A and Chaplain MAJ (2013) Spatial stochastic modelling of the Hes1 gene regulatory network: intrinsic noise can explain heterogeneity in embryonic stem cell differentiation. J R Soc Interface 10:20120988
68. Szili L, Tóth J (1993) Necessary condition of the turing instability. Phys Rev E 48:183–186
69. Szili L, Tóth J (1997) On the origin of turing instability. J Math Chemistry 22:39–53
70. Turing AM (1952) The chemical basis of morphogenesis. Philo Trans R Soc Lond Ser B 237:5–72
71. Uhlenbeck GE, Ornstein LS (1930) On the theory of Brownian motion. Phys Rev 36:823–841

72. von Smoluchowski M (1906) Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen. Ann Phys (Leipzig) 21:756–778
73. Wang MC, Uhlenbeck GE (1945) On the theory of Brownian motion, II. Rev Mod Phys 17:323–342
74. Wegscheider R (1901–1902) Über simultane Gleichgewichte und die Beziehungen zwischen Thermodynamik und Reaktionskinetik homogener Systeme. Zeitsch Phys Chem 39:257–303
75. Wiener N (1923) Differential space. J Math Phys 2:131–174
76. Zak M (1993) Introduction to terminal dynamics. Complex Syst 7:59–87

Chapter 2
Continuous Time Discrete State Stochastic Models

2.1 Model Frameworks

The deterministic models of classical kinetics are appropriate only when the system can be considered as macroscopic, and even then the deviations from the "average values" remain negligible. Chemical equilibrium, the study of which usually precedes chemical kinetics within physical chemistry, is well understood to be modeled better by a particle-based view with some stochastic elements developed in statistical thermodynamics than classical thermodynamics based entirely on macroscopic functions. It follows quite logically that deterministic chemical kinetics should have a more generally applicable particle-based counterpart as well that should give a better approach to observed physical reality. A number of situations are known when assuming continuous concentration-time functions is a very crude approximation. Often times, **natural fluctuations** (i.e. not caused by any external influences) are important in the system. Some of the possible reasons are as follows:

- The size of the chemical system is small. In this case, the state space is discrete, the continuous approximation is very bad. Discrete state space, but deterministic models are also out of question, since fluctuations cannot be neglected even in the "zeroth" approximation, because they are not superimposed upon the phenomenon, but they represent the phenomenon itself.
- The system operates near an instability point of a deterministic model. In this case, small fluctuations may be amplified and produce observable, even macroscopic effects.
- Fluctuations can be a source of information. The fluctuation-dissipation theorem connects the spontaneous fluctuations around an equilibrium state, and the dissipative process leading to equilibrium. Using this theorem applied to chemical kinetics, rate constants can be calculated from equilibrium fluctuations.

P. Érdi and G. Lente, *Stochastic Chemical Kinetics: Theory and (Mostly) Systems Biological Applications*, Springer Series in Synergetics, DOI 10.1007/978-1-4939-0387-0_2,

- A specific importance is found for a concentration that falls to zero at a finite time value. As pointed out earlier, this is not possible within the mathematical framework of deterministic mass action type kinetics. However, the stochastic approach may associate this event with a non-zero probability.

The alternative to deterministic kinetics used in most of this book is called the continuous time discrete state stochastic model. Its relationship to deterministic kinetics is analogous to the connection between classical and statistical thermodynamics. Statistical thermodynamics can give some sort of a theoretical prediction about equilibrium fluctuations, stochastic kinetics can describe the time evolution of chemical fluctuation phenomena. The concentration, (a continuous variable) should be transformed into a function of the number of components (discrete variable): $\mathbf{c}(t) \rightarrow N_{\mathrm{A}}^{-1} V^{-1} \mathbf{N}(t)$, where N_{A} is Avogadro's constant (necessary to account for the fact that concentration is defined using the amount of substance rather than particle numbers), V is the volume of the system, $\mathbf{c}(t)$ is a concentration vector and $\mathbf{N}(t)$ is the vector of the particle numbers at a fixed time t.

Stochastic kinetics and its applications will be described in this book, but not before some more general definitions and theorems of stochastic processes are introduced.

2.2 Stochastic Processes

A brief summary of the mathematical foundations of stochastic processes will be given in the next sections. A more detailed, and relatively easy-to-read classic summary can be found in the literature [44]. Our description mainly follows the approach of Barabás et al. [4], which was partially motivated by our old wording (section 5.1.3 in [24]).

2.2.1 Definition

Given the **independent variable set** T, whose elements usually represent time, and the set of possible outcomes S, called the **state space**, a **stochastic process** is defined as a collection of random variables $\{X(t),\ t \in T\}$ on the same probability space.

The mapping $t \mapsto X(t, \cdot)$ defined on set T is a **realization** or a **sample path** of the process. (The name **temporal trajectory** is also used, but it may cause misunderstanding because this term is used differently in the case of deterministic models.)

2.2.2 Time and State Space

Stochastic processes can be divided into four categories depending on the continuous or discrete nature of the time variable and the random variables.

Discrete time, discrete state stochastic processes (DDS) feature a time variable and a state space that are both discrete (i.e. they belong to a finite or countable set). If the assumed values of time, t_i $(i = 1, 2, \ldots)$, form an increasing sequence, $X(t_i)$ is a **random sequence**. If the process is Markovian (see Sect. 2.2.4), then it is called a **Markov chain**. Simulation techniques in stochastic chemical kinetics usually belong to this class.

In the case of **continuous time, discrete state stochastic processes (CDS)**, the state space is still discrete, but the time variable assumes a continuous range of values in $(-\infty, +\infty)$. A Markovian process with these features is called a **continuous time Markov chain**. The standard stochastic model of chemical reactions is a CDS approach, with $\mathbb{R}$ as the set of times, and $\mathbb{N}^n$ as the state space, where n is the number of chemical species.

For **discrete time, continuous state stochastic processes (DCS)**, $X(t)$ assumes a continuous range of values, but time variable t is discrete. If the process is Markovian, it is called a **discrete time Markov process**.

Finally, both $X(t)$ and t assume continuous ranges of values in a **continuous time, continuous state stochastic processes (CCS)**. A **continuous time Markov process** is a Markovian CCS process.

2.2.3 Important Types of Stochastic Processes

A **stationary process** is a stochastic process whose joint probability distributions do not change when shifted in time. A stochastic process $X(t)$ with the cumulative distribution function $F_{X_{t_1},\ldots,X_{t_k}}(x_{t_1}, \ldots, x_{t_k})$ at times $t_1, \ldots, t_k$ is stationary if the following equation holds for all k, τ, and $t_1, \ldots, t_k$ values:

$$F_{X_{t_1},\ldots,X_{t_k}}(x_{t_1}, \ldots, x_{t_k}) = F_{X_{t_1+\tau},\ldots,X_{t_k+\tau}}(x_{t_1}, \ldots, x_{t_k}). \tag{2.1}$$

Weak sense stationarity (wide-sense stationarity or covariance stationarity) only requires that the first and second moments do not vary with respect to time. Any strictly stationary process that has a mean and a covariance is weakly stationary.

A DDS stochastic process or a random sequence $\{X_n\}$ is an **independent process** if the joint density function can be written as the product of the density functions of the variables:

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n; t_1, \ldots, t_n) = f_{X_1}(x_1; t_1) \cdots f_{X_n}(x_n; t_n). \tag{2.2}$$

A random vector $\mathbf{X}$ is a **white random vector** if its mean vector is zero and its autocorrelation matrix is a multiple of the identity matrix:

$$\langle \mathbf{X} \rangle = \mathbf{0} \quad R_X = \left\langle \mathbf{X}\mathbf{X}^T \right\rangle = \sigma^2 \mathbf{I}. \tag{2.3}$$

A continuous time random process $\mathbf{X}(t)$ ($t \in \mathbb{R}$) is a **white noise process** if its mean function and autocorrelation satisfy the following equations:

$$\langle \mathbf{X}(t) \rangle = 0 \quad R_{\mathbf{X}}(t_1, t_2) = \left\langle \mathbf{X}(t_1)\mathbf{X}(t_2)^T \right\rangle = \frac{N_0}{2}\delta(t_1 - t_2). \tag{2.4}$$

In Eq. (2.4), δ stands for the **Dirac delta function**, which returns a zero value for every real number except 0, but its definite integral over any interval around 0 is exactly 1. A possible way to define this function is as follows:

$$\delta(\mu_i) = \lim_{x \to 0} \frac{1}{x\sqrt{\pi}} e^{-\mu_i^2/x^2} \tag{2.5}$$

The most important property of the Dirac delta function for the present purposes is that the integral of its product with any function $h(\mu)$ can be given very simply:

$$\int_{-\infty}^{+\infty} h(\mu_i)\delta(\mu_i - x)d\mu_i = h(x) \tag{2.6}$$

The **increments** of a continuous time stochastic process by $X(t)$ are the differences $X(s) - X(t)$ between its values at different times $t < s$. The increments of the process are **independent** if increments $X(s) - X(t)$ and $X(u) - X(v)$ are independent random variables whenever the two time intervals $[t, s]$ and $[v, u]$ do not overlap and more generally, any finite number of increments assigned to pairwise non-overlapping time intervals are mutually (not just pairwise) independent.

2.2.4 Markov Chains

A highly important class of stochastic processes was named after Russian scientist Andrey Andreyevich Markov (1856–1922), who was not the only notable mathematician in his family. A **Markov chain** is a sequence of random variables $X_1, X_2, \ldots X_n, \ldots$ that satisfy the following equation for every n ($n = 1, 2, \ldots$):

$$\begin{aligned} &P(X_n = j | X_1 = i_1, X_2 = i_2, \ldots, X_{n-1} = i_{n-1}) \\ &\quad = P(X_n = j | X_{n-1} = i_{n-1}) \end{aligned} \tag{2.7}$$

The Markovian property means that the probability of the state at time n depends only on the state at time $n - 1$ and does not depend directly on the earlier states: the

future depends on the past only through present. This is often summarized briefly by stating that a Markov chain has no memory.

The probabilities $p_{ij} = P(X_n = j | X_{n-1} = i)$ are called **single-step transition probabilities**, which give the probability of the Markov chain going from state i into state j in one step. In the special case when the single-step transition probabilities do not depend on n, the Markov chain is called **(time-)homogeneous** (the word 'stationary' is also used occasionally in this sense, which may be the source of some confusion as this stationarity is different form the one defined in Sect. 2.2.3).

In the homogeneous case, **m-step transition probabilities** can be defined as follows:

$$ {}^{n}p_{ij}^{m} = P(X_{n+m} = j | X_n = i). \tag{2.8} $$

From now on, this text will deal with homogeneous Markov chains exclusively, and the m-step transition probabilities will be denoted simply p_{ij}^{m}.

From Eq. (2.7), the following Markov property, which is a special case of the Chapman–Kolmogorov equation, is easily proven for any integer l $(0 < l < m)$:

$$ p_{ij}^{m} = \sum_{k \in S} p_{ik}^{m-l} p_{kj}^{l} \quad (m = 2, 3, \dots) \tag{2.9} $$

If the state space of a homogeneous Markov chain is finite (or infinite but countable), the transition probabilities can be collected into a matrix, which is called the **transition probability matrix**:

$$ \mathbf{P} = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots \\ p_{2,1} & p_{2,2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \tag{2.10} $$

The sum of each row in $\mathbf{P}$ is one and all elements are non-negative. Therefore, $\mathbf{P}$ is a **(right) stochastic matrix**. The k-step transition probabilities for a time-homogeneous Markov chain are given by the kth power of the transition probability matrix, $\mathbf{P}^k$.

A probability distribution $\boldsymbol{\pi}$ is called a **stationary distribution** if it satisfies the following equation.

$$ \boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P} \tag{2.11} $$

A stationary distribution $\boldsymbol{\pi}$ is a normalized (i.e. the sum of its elements is 1) left eigenvector of $\mathbf{P}$ associated with the eigenvalue 1. A stationary distribution always exists, but it is not necessarily unique. If the Markov chain is irreducible

and aperiodic (see the next few paragraphs for the definitions of these properties), a unique stationary distribution exists. In this case, $\mathbf{P}^k$ converges to a rank-one matrix in which each row is the stationary distribution:

$$\lim_{k \to +\infty} \mathbf{P}^k = \mathbf{1}\boldsymbol{\pi} \tag{2.12}$$

In Eq. (2.12), $\mathbf{1}$ is the column vector with all entries equal to 1.

A careful analysis of the structure of the state space can often give useful hints for obtaining the state probabilities, or prove notable symmetry properties. Therefore, some definitions relevant to the properties of individual states or a group of states will be given in the next paragraphs.

State j is **accessible** from state i (written as $i \to j$) if there is non-zero probability of passing into state j from state i (there exists an integer $n > 0$ such that $P(X_n = j | X_0 = i) = p_{ij}^n > 0$). A Markov chain is **irreducible** if all pairs of states are mutually accessible from one another.

State i has period k if any return to state i must occur in multiples of k time steps ($k = \gcd\{n : P(X_n = i | X_0 = i) > 0\}$, where gcd is the greatest common divisor). If $k = 1$, the state is **aperiodic**. Otherwise ($k > 1$), the state is said to be **periodic with period** k. A Markov chain is **aperiodic** if all its states are aperiodic.

State i is **transient** if the probability of never returning to this state is larger than zero. If the random variable T_i represents the the **hitting time**, which is the first (earliest) return time to state i: $T_i = \inf\{n \geq 1 : X_n = i | X_0 = i\}$, state i is transient if and only if $P(T_i = +\infty) > 0$. A **recurrent** or **persistent** state is not transient (it has a finite hitting time with probability 1). Even if the hitting time is finite, it does not necessarily have a finite expectation. If the expected return time, $M_i := \langle T_i \rangle$, is finite, state i is **positive recurrent**. Otherwise, the state is **null recurrent** (the terminology non-null persistent or null persistent is also used).

It can be shown that a state is recurrent if and only if the following equation holds:

$$\sum_{n=0}^{+\infty} p_{ii}^n = +\infty \tag{2.13}$$

An irreducible Markov chain has a stationary distribution if and only if all of its states are positive recurrent. In that case, $\boldsymbol{\pi}$ is unique, and the chain converges to the stationary distribution from any initial state. Such a $\boldsymbol{\pi}$ is called the **equilibrium distribution** of the chain.

A state is **absorbing** if it is impossible to leave this state. Therefore, the state i is absorbing if and only if $p_{ii} = 1$ and $p_{ij} = 0$ for $i \neq j$. State i is **ergodic** if it is aperiodic and positive recurrent. If all states in a Markov chain are ergodic, then the chain is said to be ergodic. It can be shown that a finite state irreducible Markov chain is ergodic if its states are aperiodic.

2.2.5 *Continuous Time Discrete State Markov Process*

In analogy with DDS Markov chains, a stochastic process $\{X(t)\}_{t\in(R)}$ is a **continuous time Markov process** if the following equation holds for every $t_1 < t_2 < \cdots < t_{n+1}$ (n is a positive integer):

$$\begin{aligned} P\left(X(t_{n+1}) = j \,|\, X(t_1) = i_1, X(t_2) = i_2, \ldots X(t_n) = i_n\right) \\ = P\left(X(t_{n+1}) = j \,|\, X(t_n) = i_n\right) \end{aligned} \tag{2.14}$$

Equation (2.14) is fully analogous to Eq. (2.7). Thus, most of the properties of the continuous time Markov process are similar to those of the Markov chain.

Let τ_i denote the time that the process spent in state i. According to the Markov property in Eq. (2.14), τ_i does not depend on the past of the process, so the following equation holds:

$$P\left(\tau_i > s + t \,|\, \tau_i > s\right) = h(t) \tag{2.15}$$

Function $h(t)$ in Eq. (2.15) only depends on the remaining time t, and not on the past time s. The only continuous probability distribution which satisfies Eq. (2.15) is the exponential distribution. Equation (2.15) is often quoted as the **memorylessness property** of the Markov process. In the discrete time case, requirement (2.15) leads to the geometric distribution.

The transition probability for a Markov process is defined as:

$$p_{ij}(s,t) = P\left(X(t) = j \,|\, X(s) = i\right). \tag{2.16}$$

Obviously, the following equation holds for the transition probabilities for all possible i values:

$$\sum_{k\in S} p_{ik}(s,t) = 1 \tag{2.17}$$

Furthermore, it follows from Eq. (2.15) that $p_{ik}(s,t)$ only depends on the difference $(t-s)$ in a Markov process, but not on the values of t and s individually. The definition of the transition probabilities gives rise to the **Chapman–Kolmogorov equation**:

$$p_{ij}(s,t) = \sum_k p_{ik}(s,u)\, p_{kj}(u,t) \quad (i,j = 0,1,2,\ldots) \tag{2.18}$$

The **transition probability matrix** $\mathbf{P}(s,t)$ can be constructed from the individual transition probabilities:

$$\mathbf{P}(s,t) = \begin{pmatrix} p_{1,1}(s,t) & p_{1,2}(s,t) & \cdots \\ p_{2,1}(s,t) & p_{2,2}(s,t) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \tag{2.19}$$

For internal consistence, let $\mathbf{P}(t,t)$ be the identity matrix ($\mathbf{I}$). Using this notation, the Chapman–Kolmogorov equation given in Eq. (2.18) can be stated in a matrix form:

$$\mathbf{P}(s,t) = \mathbf{P}(s,u)\mathbf{P}(u,t) \quad s \leqq u \leqq t. \tag{2.20}$$

If the transition probabilities satisfy some continuity conditions, a system of differential equations can be written that describes the process. The **absolute state probabilities** $P_i(t) := P(X(t) = i)$ of a CDS Markov process, i.e. the probabilities for the system to be in the state i, satisfy a relatively simple recursive equation with the transition probabilities:

$$P_i(t) = \sum_j p_{ji}(s,t) P_j(s) \tag{2.21}$$

$P_i(t)$ functions are very often the most preferred characteristics of CDS Markov processes in physical and chemical applications. From the Chapman–Kolmogorov equation, the **master equation** of the CDS Markov process can be derived:

$$\frac{dP_i(t)}{dt} = \sum_j (q_{ji} P_j(t) - q_{ij} P_i(t)), \tag{2.22}$$

The q_{ij} values are called **infinitesimal transition probabilities** (transition rates), and can be obtained from the transition probabilities as follows:

$$\begin{aligned} q_{ii} &= 0 \\ q_{ij} &= \lim_{s \to t} \frac{p_{ij}(s,t)}{t-s} \quad (i \neq j) \end{aligned} \tag{2.23}$$

In a Markov process, $p_{ij}(s,t)$ only depends on the difference $(t-s)$, so q_{ij} is independent of t as already implied in Eq. (2.23). Time-dependent transition rates may only appear in non-Markovian processes. The master equation can also be stated in a matrix form for the vector of absolute probabilities, $\mathbf{P}(t)$, this is often most suitable for carrying out numerical calculations and will be discussed in more detail later. A few examples of the Markov process will be introduced in the next few paragraphs.

Example 2.1. A Markov process $X(t)$ $(t \geqq 0)$ is a **birth-and-death process** with parameters $\lambda_0, \lambda_1, \ldots$ and $\mu_0, \mu_1, \ldots$ if the transition probabilities satisfy the following equations:

$$\begin{aligned} p_{i,i+1}(t, t+h) &= \lambda_i h + o(h) \\ p_{i,i-1}(t, t+h) &= \mu_i h + o(h) \\ p_{i,i}(t, t+h,) &= 1 - (\lambda_i + \mu_i)h + o(h) \\ p_{i,j}(t, t+h) &= o(h) \quad \text{if} \quad j \neq i \quad \text{and} \quad j \neq i \pm 1 \quad h \to 0. \end{aligned} \tag{2.24}$$

Because the birth-and-death process is a special CDS Markov process, the master equation can be re-arranged into a special form. If $P_n(t)$ is the probability of the process being in state n at time t, the following equations hold:

$$\begin{aligned} \frac{dP_0(t)}{dt} &= -\lambda_0 P_0(t) + \mu_1 P_1(t) \\ \frac{dP_n(t)}{dt} &= -(\lambda_n+\mu_n) P_n(t)+\lambda_{n-1} P_{n-1}(t)+\mu_{n+1} P_{n+1}(t) \quad (n \geqq 1) \end{aligned} \tag{2.25}$$

A chemical reaction scheme that gives rise to a birth-and-death process is as follows:

$$\begin{aligned} \emptyset &\xrightarrow{k_1} \mathrm{A}_1 \\ \mathrm{A}_1 &\xrightarrow{k_2} 2\mathrm{A}_1 \\ \mathrm{A}_1 &\xrightarrow{k_3} \emptyset \end{aligned} \tag{2.26}$$

Example 2.2. A CDS stochastic process $N(t)$ is a **Poisson process** if it starts at zero ($N(0) = 0$), it has independent, stationary increments, and, for every $t > 0$, $N(t)$ is a Poisson random variable with parameter λt described by the following equation:

$$P\,(N(t) = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \quad n = 0, 1, 2, \ldots \tag{2.27}$$

A Poisson process is a birth-and-death process, in which only birth occurs with $\lambda_i = \lambda$ and $\mu_i = 0$ for $i = 0, 1, 2, \ldots$. In this case, Eq. (2.25) assumes a simpler form:

$$\frac{dP_n(t)}{dt} = -\lambda(P_n(t) - P_{n-1}(t)). \tag{2.28}$$

The solution of Eq. (2.28) is exactly given by Eq. (2.27). A chemical example of the Poisson process is:

$$\emptyset \xrightarrow{k_1} \mathrm{A}_1 \tag{2.29}$$

Example 2.3. Let G^d denote the points of a d-dimensional lattice. The position of a point at time $t = n$ is given by S_n^d. The point changes its position such that one of its coordinates changes by ± 1 with probability $\frac{1}{2d}$ and all the other $d - 1$ coordinates remain unchanged. If X_k^d stands for the shift in the time interval $(k-1, k)$, then S_n^d can be given as:

$$S_n^d = S_0^d + \sum_{k=1}^{n} X_k^d \tag{2.30}$$

X_k^d are independent, identically distributed random variables, hence S_n^d is a Markov process. Let $P(d)$ be the probability that a random walk on a d-dimensional lattice returns to the initial position. It is easily proven that $P(1) = P(2) = 1$, but $P(d) < 1$ for any $d \geqq 3$. Moreover, the probability that a random walk on a d-dimensional lattice returns to the initial position infinitely many times equals 1 for $d = 1$ and $d = 2$ but equals 0 for $d \geqq 3$. Some approximate values of $P(d)$ are as follows: $P(3) = 0.340537$, $P(4) = 0.193206$, and $P(5) = 0.135178$.

Example 2.4. Let $T_i > 0$ for $i = 1, 2, \ldots$ be a sequence of independent, identically distributed random variables. The random variable T_i is referred to as the ith **holding time**. The nth **jump time** $(J_n, n > 0)$ is defined by the following equation:

$$J_n = \sum_{i=1}^{n} T_i \text{ and } J_0 = 0 \tag{2.31}$$

The process $N_t = \sup\{n : J_n \leqq t\}$ is called a **renewal (counting) process**. If the random variables T_i are exponentially distributed with parameter λ, then N_t is a Poisson process with parameter λ, and the following equation holds:

$$\lim_{t \to +\infty} \frac{N_t}{t} = \frac{1}{\mu_{T_1}} \tag{2.32}$$

If $X_1, X_2 \ldots$ is a sequence of independent, identically distributed random variables satisfying $\mu_{|X_i|} < +\infty$, the random sum $S_t = \sum_{i=1}^{N_t} X_i$ is called a **renewal reward process**, for which the following equation holds:

$$\lim_{t \to +\infty} \frac{S_t}{t} = \frac{\mu_{X_1}}{\mu_{T_1}} \tag{2.33}$$

With $F(x) := P(X_i < x)$ as the common distribution function of the random variables X_i, the expectation of the renewal counting process is the **renewal function** and the **renewal equation** can be proved:

$$\mu_{N_t} = F(t) + \int_0^t \mu_{N_{t-x}} dF(x) \tag{2.34}$$

2.2.6 Semi-Markov Processes

A **semi-Markov process**, also called **Markov renewal process**, is a process that may change states any time (it is a continuous time process) and the waiting times between the changes are not necessary exponentially distributed. A continuous time Markov chain is a special case of a semi-Markov process, in which the transition

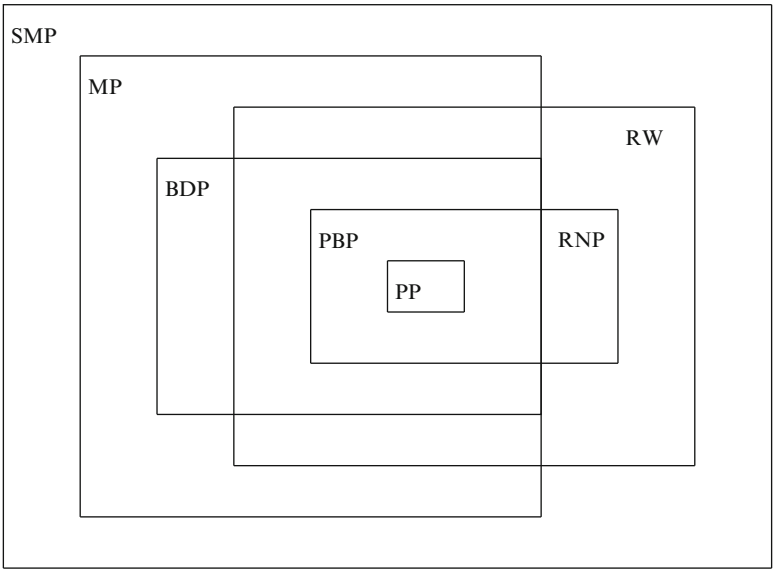


Fig. 2.1 Classification of stochastic processes

SMP: Semi-Markov Process	p_{ij} arbitrary,	F_τ arbitrary
MP: Markov Process	p_{ij} arbitrary,	F_τ memoryless
BDP: Birth-and-Death Process	$p_{ij} = 0$ if $\|i - j\| > 1$	F_τ memoryless
PBP: Pure Birth Process	$\mu_i = 0$	F_τ memoryless
RW: Random Walk	$p_{ij} = q_{j-i}$,	F_τ arbitrary
RNP: Renewal Process	$q_1 = 1$,	F_τ arbitrary
PP: Poisson Process	$\lambda_i = \lambda$	F_τ memoryless

times are exponentially distributed as shown in (2.15). Formally, a stochastic process is a semi-Markov process if the following equation holds for each $n \geq 1$, $t \geq 0$ and $i, j \in S$ with the waiting time τ_n defined as $\tau_n = t_n - t_{n-1}$:

$$\begin{aligned} &P(\tau_{n+1} \leq t, X_{n+1} = j \,|\, X_1 = i_1, X_2 = i_2, \ldots, X_n = i_n) \\ &\quad = P(\tau_{n+1} \leq t, X_{n+1} = j \,|\, X_n = i_n) \end{aligned} \tag{2.35}$$

A graphical representation of the processes defined thus far is given in Fig. 2.1.

2.3 The Standard Stochastic Model of Homogeneous Reaction Kinetics

The widely used deterministic model of chemical kinetics was described in Chap. 1. This section will introduce the stochastic equivalent of Eq. (1.1) using very similar notations and the concepts of stoichiometric (ν) and order matrices (α).

The equivalence here refers to the fact that the interactions governing the natural events are the same, the only difference is that they are considered on a molecular rather than bulk level. As already mentioned in the beginning of this chapter, the concentration vector is replaced by the particle number vector $\mathbf{N}(t) = N_\mathrm{A} V \mathbf{c}(t)$ in the standard stochastic model. Whilst the trajectory $\mathbf{c}(t)$ is unambiguously defined by the initial condition $\mathbf{c}(0)$ and the rate equation (1.1), the same is not true for sample path $\mathbf{N}(t)$, and $\{\mathbf{N}(t),\ t \in T\}$ is a stochastic process with time as the independent variable. The elements of vector $\mathbf{N}$ are the particle numbers of chemical species $\mathrm{A}_1, \mathrm{A}_2, \ldots, \mathrm{A}_n$, and are denoted $a_1, a_2, \ldots, a_n$. The number of elements in vector $\mathbf{N}$ is equal to the number of chemical species (types of particles), n.

The CDS approach to stochastic kinetics assumes that $\mathbf{N}(t)$ is a Markov process. The stochastic equivalent of the deterministic rate equation is the master equation of the process, in which the time-dependent probabilities of states $P_\mathbf{N}(t) = P(\mathbf{N}(t) = \mathbf{N})$ are the dependent variables and the independent variable is time (t). The notation with an enumeration of the particle numbers as the index of the probabilities, $P_{a_1,a_2,\ldots,a_n}(t)$, is also used instead of $P_\mathbf{N}(t)$.

2.3.1 State Space: Size and Enumeration

The state space of CDS stochastic kinetics is the values vector $\mathbf{N}$ can assume. Because all elements of this vector are natural numbers, the state space is always a subset of $\mathbb{N}^n$. As $\mathbb{N}^n$ is a countable set, so is the state space. As a fundamental principle, the state space should always be finite because, even in the absence of other reasons, the overall number of elementary particles in the Universe is known to be finite. However, using an infinite state space sometimes makes mathematical calculations easier. In these instances, the physical limitations of using the infinite model should be recognized and (preferably) clearly stated.

A state can generally be identified by giving the numbers of all particle types, such as $(a_1, a_2, \ldots, a_n)$. Some of the concepts introduced earlier about the states (e.g. accessible and absorbing states) of a stochastic process will be useful in analyzing the state space of CDS stochastic kinetics as well. States i and j are **mutually accessible** if state i is accessible from state j and state j is also accessible for state i at the same time. State i is **downstream** from state j if state i is accessible from state j, but the two states are not mutually accessible. In this case, state j is also **upstream** from state i. An absorbing state represents a final mixture in which chemical reactions are not possible any more.

The actual size of the state space is typically determined by the initial conditions. A state is **impossible** if it is not accessible from the initial state(s), whereas a **possible** state is not impossible. The initial state can be a single state or a range of states characterized by some initial distribution $P_\mathbf{N}(0)$. As the largest computational problem in CDS stochastic kinetics is posed by the typically huge number of states, it is imperative to exclude all impossible states from the state space.

Counting the number of states (denoted M) in a finite state space is usually important, but not always essential. The initial conditions usually give rise to a number of conservation equations, similarly to the time-independent linear combinations of concentrations for the deterministic model. States not satisfying these conservation equations are impossible. Therefore, counting the number of possible states is basically the combinatorial problem of determining the number of solutions for a set of simultaneous Diophantine equations. The difficulty of this task varies greatly from trivial (e.g. for $m = 1$) to immensely complex. The least common multiple of the integer multipliers in the conservation equations usually has special significance. Some guidelines are available in the mathematical literature [15]. Because of the conservation equations, unambiguous identification of a possible state seldom requires giving all the particle numbers $a_1, a_2, \ldots, a_n$. If this is the case, indexing of the probabilities is often reduced to those a_i values that are necessary as the rest of the values can be calculated from the conservation equations. For example, giving any one of the a_i values is sufficient to identify a state in a single-step reaction ($m = 1$) with a unique initial state.

Enumerating the Diophantine solutions of the conservation equations is also often necessary. For this purpose, an enumerating function, f, is needed, which gives a single and unique whole number to every possible state without repetitions or omissions. Therefore, $f(a_1, a_2, \ldots, a_n)$ should take the values $1, 2, \ldots, M$. The need for avoiding omissions in the values of the enumerating function is not theoretically necessary, but has high practical importance in order to minimize the dimensions of matrices for certain solution methods. It should also be noted that the enumerating function is far from being unique. In fact, the number of different enumerating functions for a given system and initial conditions is $M!$. Using an enumerating function, the states can be ordered and the indexing $P_{f(a_1,a_2,\ldots,a_n)}(t)$ can be used instead of $P_{\mathbf{N}}(t)$ to have positive integers as indexes rather than vectors, which is very convenient if matrix operations are used to solve the master equation.

It has already been stated that determining the number of possible states can be quite a formidable combinatorial problem even in systems of moderate complexity. The difficulties in finding enumerating functions are often much worse. An enumerating function is invertible by definition, but finding its inverse can be a third hard problem. Fortunately, computational algorithms can typically be programmed in a way that the inverse of the enumerating function is not necessary.

If all else fails, a **brute force method** can be used to determine the number of states and define an enumerating function in a relatively small state space if the maximum number of particles in possible states is known:

$$N_{\max} = \max(\max(a_1), \max(a_2), \ldots, \max(a_n)) \tag{2.36}$$

The algorithm involves generating all possible points in the set ${N_{\max}}^n$ by successively organized cycles. Adherence to the conservation equations is tested for all such points. The first point that satisfies all equations is assigned a value of 1, then the following satisfactory points are given gradually increasing values. When function f is generated in this way, it will be given by an enumeration of the entire

state space and all the values instead of a symbolic formula. This is computationally only viable for small values of M. Inverting function f is trivial in this case.

In most published papers, analyzing the state space of a problem does not receive much attention. However, this is usually worth the effort as it often reveals special features of the studied system that could aid the general solution of the master equation or accelerate the algorithms.

2.3.2 Master Equation

The master equation of CDS stochastic kinetics has the same importance as the rate equation of the deterministic approach, and has a similar structure as well. Deterministic rates should be replaced by transition rates. The same notation is used here (v_j) for the two physical quantities, but their physical dimensions are different. The dimension of rates is concentration divided by time (the most common unit is $\mathrm{mol\,dm^{-3}\,s^{-1}}$), the dimension of transition rates is inverse time (the most common unit is $\mathrm{s^{-1}}$). A sort of equivalence of the deterministic and stochastic approaches can be ensured if the functions describing the dependence of transition rates on particle numbers can be obtained by replacing concentrations $[\mathrm{A}_i]$ with particle numbers a_i, and rate constants k_j by rate parameters κ_j in the symbolic formulas. This procedure is probably the only possibility for non-power law rates or for fractional orders of reaction in power-law rates. However, this is never fully consistent with physical reality. Because of the reducibility of net chemical reactions and the molecular level of description in CDS stochastic kinetics, it is best to reduce all systems to mass action type kinetics and give the transition rate of reaction j starting from state $(a_1, a_2, \ldots, a_n)$ as follows:

$$v_j(\mathbf{n}) = v_j(a_1, a_2, \ldots, a_n) = \kappa_j \prod_{k=1}^{n} \binom{a_k}{\alpha_{j,k}} \tag{2.37}$$

The full master equation of the process is then given by summing all transition rates relevant to a given state:

$$\begin{aligned} \frac{dP_{f(a_1,a_2,\ldots,a_n)}}{dt} &= -\sum_{j=1}^{m} v_j(a_1, a_2, \ldots, a_n) P_{f(a_1,a_2,\ldots,a_n)} \\ &+ \sum_{j=1}^{m} v_j(a_1 - \nu_{j,1}, a_2 - \nu_{j,2}, \ldots, a_n - \nu_{j,n}) P_{f(a_1-\nu_{j,1},a_2-\nu_{j,2},\ldots,a_n-\nu_{j,n})} \end{aligned} \tag{2.38}$$

The master equation as stated in Eq. (2.38) can be solved directly in computer calculations by using some matrix algebra for cases when the number of states is finite (M). First, the equation is stated for the vector $\mathbf{P}(t)$, the elements of which are the individual $P_{f(a_1,a_2,\ldots,a_n)}(t)$ values:

$$\frac{d\mathbf{P}(t)}{dt} = \Omega \mathbf{P}(t) \tag{2.39}$$

The matrix Ω is composed of the transition rates and its elements are defined as follows:

$$\begin{aligned} \Omega_{f(a_1,a_2,\ldots,a_n),f(a_1,a_2,\ldots,a_n)} &= -\sum_{j=1}^{m} v_j(a_1,a_2,\ldots,a_n) \\ \Omega_{f(a_1,a_2,\ldots,a_n),f(a_1-v_{j,1},a_2-v_{j,2},\ldots,a_n-v_{j,n})} &= v_j(a_1,a_2,\ldots,a_n) \end{aligned} \tag{2.40}$$

Problems with this definition might arise if two (or more) reaction steps are described by the same set of stoichiometric coefficients. However, within mass action type kinetics where the stoichiometric coefficients determine the orders of reaction, this problem is easily handled by uniting these steps into one with a rate parameter that is equal to the sum of the original rate parameters.

In addition, Ω can also be defined using the transition probability matrix of Eq. (2.20):

$$\Omega = \lim_{s \to t} \frac{\mathbf{P}(s,t) + \mathbf{I} - 2\mathrm{diag}(\mathbf{P}(s,t))}{t-s} \tag{2.41}$$

Here, matrix $\mathrm{diag}(\mathbf{P}(s,t))$ is the diagonal square matrix that only contains the main diagonal elements of $\mathbf{P}(s,t)$. Ω is independent of time for the same reason $q_{i,j}$ values are not time-dependent.

The equivalence of the deterministic and stochastic approaches is guaranteed by a connection between deterministic rate constants and stochastic rate parameters:

$$\kappa_j = k_j (N_\mathrm{A} V)^{-\sum_{k=1}^{n} \alpha_{j,k}} \tag{2.42}$$

Although the physical dimensions of rate constants vary depending on the overall order of the reactions step, the physical unit of stochastic rate parameters is always inverse time. Chapter 3 will give numerous examples of the use of the CDS master equation. However, some significant phenomena will be illustrated here by examples that have little practical significance.

Example 2.5. Consider the irreversible second order dimerization reaction: $2\mathrm{A}_1 \longrightarrow \mathrm{A}_2$. Let N_0 and 0 be the number of A_1 and A_2 particles in the initial state. There is only one conservation equation in this system:

$$N_0 = a_1 + 2a_2 \tag{2.43}$$

In the possible states, $a_1 \leq N_0$ and a_1 has the same parity as N_0. The number of possible states is $M = \mathrm{int}(N_0/2) + 1$, where int denotes the integer part function. A suitable enumerating function is $f(a_1,a_2) = a_2 + 1$. The CDS master equation is as follows:

$$\frac{dP_{a_1,a_2}}{dt} = -\kappa_1 \frac{a_1(a_1-1)}{2} P_{a_1,a_2} + \kappa_1 \frac{(a_1+2)(a_1+1)}{2} P_{a_1+2,a_2-1} \tag{2.44}$$

A direct equivalent of the deterministic description in this case would be to have the rate term a_1^2 instead of $a_1(a_1 - 1)$. However, the a_1^2 term would imply that a single molecule of A_1 could undergo a reaction requiring two A_1 particles with a non-zero probability, which is physically impossible.

Example 2.6. Consider the irreversible half-order reaction: $A_1 \longrightarrow A_2$. Let N_0 and 0 be the number of A_1 and A_2 particles in the initial state. The only conservation equation in this system is:

$$N_0 = a_1 + a_2 \tag{2.45}$$

In the possible states, $a_1 \leq N_0$. The number of possible states is therefore $M = N_0 + 1$. A suitable enumerating function is $f(a_1, a_2) = a_2 + 1$. The CDS master equation is as follows:

$$\frac{dP_{a_1,a_2}}{dt} = -\kappa_1 \sqrt{a_1} P_{a_1,a_2} + \kappa_1 \sqrt{a_1 + 1} P_{a_1+1,a_2-1} \tag{2.46}$$

This is an example of power law kinetics, which does not adhere to mass action. Therefore, the conventions of the transition rate definition given in (2.37) cannot be used here. In general, it is better to reduce this system to a several-step scheme that follows mass action type kinetics.

Example 2.7. Consider the mass action kinetics scheme of catalyzed dimerization, which is composed of two elementary reactions: $A_1 + A_2 \longrightarrow A_3$, $A_1 + A_3 \longrightarrow A_2 + A_4$. Let N_0, C_0, 0, and 0 be the number of A_1, A_2, A_3, and , A_4 particles in the initial state. There are two conservation equations:

$$\begin{aligned} N_0 &= a_1 + a_3 + 2a_4 \\ C_0 &= a_2 + a_3 \end{aligned} \tag{2.47}$$

Giving the molecule numbers of a_1 and a_2 always unambiguously identifies a state as the conservation equations can be re-arranged to $a_3 = C_0 - a_2$ and $a_4 = 0.5(N_0 - a_1 - C_0 + a_2)$. In the possible states, $a_1 \leq N_0$, $a_2 \leq C_0$, and $N_0 - a_1 - C_0 + a_2$ is a non-negative even number. The number of possible states is:

$$\begin{aligned} M &= \text{int}\left(\frac{(N_0+2)^2}{4}\right) \quad \text{if } N_0 \leq C_0 \\ M &= \text{int}\left(\frac{C_0^2}{4}\right) + \text{int}\left(\frac{N_0 - C_0}{2} + 1\right)(C_0 + 1) \\ &\qquad \text{if } N_0 > C_0 \text{ and } N_0 \equiv C_0 (\text{mod } 2) \\ M &= \text{int}\left(\frac{(C_0+1)^2}{4}\right) + \text{int}\left(\frac{N_0 - C_0}{2} + 1\right)(C_0 + 1) \\ &\qquad \text{if } N_0 > C_0 \text{ and } N_0 \neq C_0 (\text{mod } 2) \end{aligned} \tag{2.48}$$

A suitable enumerating function is:

$$\begin{aligned} f(a_1,a_2,a_3,a_4) &= \frac{N_0 - a_1 - C_0 + a_2}{2}(C_0+1) + C_0 - a_2 + 1 \quad \text{if } a_1 \geq a_2 \\ f(a_1,a_2,a_3,a_4) &= M - \text{int}\left(\frac{(C_0 + a_1 - a_2)^2}{4}\right) - a_1 \quad \text{if } a_1 < a_2 \end{aligned} \tag{2.49}$$

The CDS master equation is as follows:

$$\begin{aligned} \frac{dP_{a_1,a_2,a_3,a_4}}{dt} &= -[\kappa_1 a_1 a_2 + \kappa_2 a_1 (C_0 - a_2)] P_{a_1,a_2,a_3,a_4} \\ &+\kappa_1 (a_1+1)(a_2+1) P_{a_1+1,a_2+1,a_3-1,a_4} \\ &+\kappa_2 (a_1+1)(C_0 - a_2 + 1) P_{a_1+1,a_2-1,a_3+1,a_4-1} \end{aligned} \tag{2.50}$$

This example shows that the formula for the number of states and the enumerating function may show some complexity even if the scheme and the master equation are simple.

2.3.3 *Connection Between Deterministic and Stochastic Kinetics: Similarities and Differences*

In the previous subsection, it has been emphasized that the CDS approach introduced is an equivalent of the usual deterministic rate equation in some sense. In fact, the applicability of the CDS approach is wider than that of deterministic kinetics as it accommodates the particulate nature of matter. The two approaches are connected formally by Kurtz's theorem [48]. The theorem basically states that the deterministic approach is the limiting case of the CDS approach for infinite volume (hence infinite molecule numbers). There are some further analogies between the structures of the two approaches, which can be listed as follows:

1. Deterministic kinetics uses autonomous differential equations, i.e. rates only depend on time through the time dependence of concentrations, but not directly. The stochastic equivalent of this property is that CDS is a Markov process, i.e. transition rates are independent of time.
2. The rate equation in the deterministic approach defines reaction rates as a function of concentrations. The stochastic equivalent is that transition rates are functions of molecule numbers.
3. In the deterministic rate equation, each differential equation contains a maximum of m additive terms because a concentration can be changed in m different reactions at most. In, the stochastic equivalent, a differential equation for the probability of any state contains all the transition rates from all preceding and

to all successor states. The maximum number of preceding states is m, the maximum number of successor states is also m. In the master equation, the equation for the time derivate of P_i contains the probabilities of m other states with positive transition rates, whereas the negative term before P_i also contains m additive terms.

There are notable differences, though. The deterministic rate equation is typically non-linear and contains differential equations for n species. The master equation is always linear and homogeneous, but the number of equations in it is the number of states M. As pointed out, M is usually huge compared to n, which gives the main difficulty of solving the master equation. No general strategy is known for obtaining useful properties from this solution. On the other hand, numerical integration of rate equations is a well-developed field and several software packages are available that calculate the concentrations of any mass action type deterministic model.

It is customary to compare the time dependence of concentrations obtained in deterministic calculations with the expectation of stochastic molecule numbers, which are defined as follows:

$$\langle a_i \rangle (t) = \sum_{\text{all states}} a_i \, P_{f(a_1,a_2,\ldots,a_n)}(t) \tag{2.51}$$

The standard deviation, or variance of the molecule number of A_i is given as:

$$\sigma_{a_i}(t) = \sqrt{\langle a_i^2 \rangle (t) - \langle a_i \rangle (t)^2} = \sqrt{\sum_{\text{all states}} a_i^2 P_{f(a_1,a_2,\ldots,a_n)}(t) - \langle a_i \rangle (t)^2} \tag{2.52}$$

The standard deviation of a molecule number has no analogue, or, with a slightly different approach, is always 0 in deterministic kinetics.

Example 2.8. Consider the first order irreversible autocatalytic reaction with the chemical equation $A_1 \longrightarrow 2A_1$. The state space is infinite in this case, but the initial presence of the $N_0 (\geq 1)$ molecules of A_1 is required. A suitable enumerating function is $f(a_1) = a_1 - N_0 + 1$ The master equation is as follows:

$$\frac{dP_{a_1-N_0+1}(t)}{dt} = k_1(a_1 - 1)P_{a_1-N_0}(t) - k_1 a_1 P_{a_1-N_0+1}(t). \tag{2.53}$$

Multiplying all equations with their specific a_1 values in Eq. (2.53) and summing them give an explicit equation for the expectation (or first moment) of the molecule numbers of A_1:

$$\frac{d \langle a_1 \rangle (t)}{dt} = k_1 \langle a_1 \rangle (t) \tag{2.54}$$

This is the same as Eq. (1.16), which describes the concentration of A_1 in the deterministic approach. The two models are said to be 'consistent in the mean' [7,8].

This is not a general property of the master equation, it is specific for the case where the overall order of reaction is 1. The same method yields the following equation for the second moment, $\left\langle a_1^2 \right\rangle$.

$$\frac{d\left\langle a_1^2 \right\rangle(t)}{dt} = 2k_1\left\langle a_1^2 \right\rangle(t) + k_1\left\langle a_1 \right\rangle(t) \tag{2.55}$$

The equation for the variance of molecule number of A_1 can be obtained by combining Eqs. (2.52), (2.54) and (2.55):

$$\frac{d\sigma_{a_1}(t)}{dt} = \sigma_{a_1}(t) + \frac{k_1\left\langle a_1 \right\rangle(t)}{2\sigma_{a_i}(t)} \tag{2.56}$$

Ordinary differential equations (2.54)–(2.56) can be solved analytically in this case.

Example 2.9. Consider the second order quadratic autocatalytic reaction $2A_1 \longrightarrow 3A_1$. The properties of the state space and the enumerating function are the same as in the previous example. The master equation of this process is:

$$\frac{dP_{a_1-N_0+1}(t)}{dt} = k_1(a_1-1)^2 P_{a_1-N_0}(t) - k_1 a_1^2 P_{a_1-N_0+1}(t). \tag{2.57}$$

The general method used in the previous example yields the following equation for the expectation:

$$\frac{d\left\langle a_1 \right\rangle(t)}{dt} = k_1\left\langle a_1^2 \right\rangle(t) \tag{2.58}$$

In this equation, a higher order moment appears. This is a general phenomenon for all systems featuring non-first order processes, which limits the applicability of the presented strategy for the calculation of expectations.

At this point, it should be also pointed out that relying on expectations can sometimes be misleading. Blomberg [9] gave an excellent example to highlight a major source of possible pitfall in understanding. This could be termed the 'fallacy of statistical expectations', which means that an expectation is not necessarily a value that actually occurs with a significant probability. To put it more vaguely, individual states close to the expectation are sometimes not expected to occur at all.

To illustrate this phenomenon, consider the following two-step scheme involving autocatalysis:

$$\begin{aligned} A_1 &\xrightarrow{k_1} A_2 \\ A_1 + A_2 &\xrightarrow{k_2} 2A_2 \end{aligned} \tag{2.59}$$

If $k_1 \ll k_2/(N_A V)$ holds with N_0 and 0 as the initial numbers of A_1 and A_2 molecules, this system is characterized by a long wait before the first molecule of A_2 forms, then suddenly all the remaining A_1 molecules are converted to the product. At intermediate times, the expectation for the molecule numbers of A_2 will be well between 0 and N_0 despite the fact that only the states with 0 and N_0 molecules of A_2 could occur with significant probability.

2.3.4 Stochastic Maps

When actual experimental data are evaluated, it must be decided whether the particular problem can be handled by the computationally less intensive deterministic approach, or the use of mathematically more demanding stochastic kinetics is inevitable. The general notion seems to be that "large" systems (not usually clearly defined) can always be described deterministically, whereas stochastic considerations are necessary for small systems. This is partially based on Kurtz's theorem [48]. However, the quoted general guideline is far from reliable under certain conditions [38].

A more satisfactory way of dealing with this problem could be called **stochastic mapping** [18, 51], which attempts to identify the part of the parameter space of a given kinetic scheme in which only the stochastic approach is viable. A convenient definition of this stochastic region is the set of parameter values for which the stochastic approach shows that the relative standard error of the target variable is larger than a pre-set critical value (often 1 % because of the usual precision of analytical methods used for concentration determination). Although there is no general proof known yet, a small standard deviation of the expectation of a variable calculated based on the stochastic approach seems to ensure that the stochastic expectation is very close to the deterministic solution. Stochastic maps are graphs that help identify the stochastic regions of kinetic models. A few examples of these maps are known in the literature.

It should be emphasized that stochastic maps critically depend on the property of interest (referred to as the target variable). Naturally, the number of reactant molecules can be given reliably by the deterministic approach in the very beginning of any chemical reaction (about identical to the initial number), whereas the same cannot be done for the number of product molecules. Therefore, it must be specified which property of the system is considered. For the comparison between the two approaches, one natural way is to compare a quantity (most often a concentration) calculated in the deterministic approach with the expectations of the same quantity derived in the stochastic method [18, 51].

2.4 Solutions of the Master Equation

The master equation can be solved symbolically, numerically, or approximate symbolic solutions can be sought. The method of choice usually depends on the size of the state space and the complexity if the reactions scheme. Several methods of solution are described in this chapter. If all else fails, Monte Carlo simulations can be carried out. This strategy is introduced briefly in Sect. 2.6.

2.4.1 What Do We Mean by Exact Analytical Solutions?

There is a question of semantics about exact analytical solutions. In usual mathematical conventions, the solution of a system of differential equations is termed "analytical" (as opposed to numerical) if the differential equation itself is solved by an analytical method leading to a formula that gives the dependent variable as a function of the independent variable without involving any differentials. In this sense, the master equations of all CDS models can be solved analytically as they are linear, first-order, ordinary differential equations. Defining the infinitesimal transition probability matrix Ω is always possible in such systems and the matrix exponential of its product with time gives a symbolic form of the analytical solution after multiplication by the initial probability vector as described in later Eq. (2.60). Nevertheless, this symbolic solution is rarely useful for practical purposes except in computations involving not more than a few thousand possible states.

The solution of a system of linear, first-order, ordinary differential equations can always be given as a linear combination of exponential functions multiplied by polynomials. This requires first the calculation of the eigenvalues of the infinitesimal transition probability matrix and then finding the multiplication factors, neither of which is necessarily possible analytically. So a narrower possible definition of an analytical solution of a CDS model could be a case when suitable analytical formulas giving the eigenvalues and multiplication factors are found. A third, even more restrictive, but not uncommon view is that a solution is only analytical when it can be given in terms of combinations of known discrete (often multinomial) probability distributions with the parameters of the distributions expressed as continuous functions of time. In this sense, analytical solutions for systems other than networks of first order reactions are only known in exceptional cases (e.g. compartmental processes).

2.4.2 Direct Matrix Operations

The solution of Eq. (2.39) can be stated in a symbolic manner by using the matrix exponential function:

$$\mathbf{P}(t) = \text{expm}(\Omega t)\mathbf{P}(0) \tag{2.60}$$

The matrix exponential function of a Ω is the matrix extension of the exponential function defined on real (or complex) numbers, and is routinely calculated by a number of widely used mathematical softwares:

$$\text{expm}(\Omega) = I + \Omega + \frac{1}{2!}\Omega\Omega + \frac{1}{3!}\Omega\Omega\Omega + \ldots. = \sum_{i=0}^{\infty} \frac{1}{i!}\Omega^i \tag{2.61}$$

Using the solution given in Eq. (2.60) is normally quite straightforward in a finite system with no more than a few thousands individual states once the transition probability matrix Ω is constructed. In matrix Ω, all of the rows should contain at least one non-zero element, otherwise the corresponding state is impossible. A column may contain zeros only if the corresponding state is absorbing.

The practicability of these direct calculations is not usually limited by computation power, but the memory needed for the storage of the elements of the transition probability matrix. In a usual chemical reaction with about 10^{20} species, the number of states easily exceeds 10^{100}. This is much larger than the estimated number of elementary particles in the visible universe (about 10^{80}), so the restrictions on the extension of stochastic kinetic calculations for higher particle numbers are of very fundamental nature and the problem is not merely a question of computing power. Nevertheless, the direct method is often useful for comparative calculations in small systems intended to demonstrate the validity of assumptions used for deriving various approximations of the exact solution. At this point, it should be noted that Ω is normally a sparse matrix, in which most of the elements are zero. In fact, the number of non-zero elements in any given row or column of Ω is not larger than the number of possible reactions in the system, which rarely exceeds a few dozen. This means that Ω is an extremely sparse matrix, with only a tiny fraction of the elements differing from zero. It is also known that Ω is a singular matrix, where all of the column sums are zero, which is again a special property of the matrix. These properties, in principle, could be used to develop numerical algorithms for calculating the matrix exponential of Ωt that are much faster and more efficient than the standard one that is developed for general matrices. In the typical cases when the initial state is certain, only one row of $\text{expm}(\Omega t)$ needs to be calculated, which could also be a factor in accelerating the numerical calculations. However, these possibilities do not seem to have been explored in any studies.

A somewhat higher number of states (often an order of magnitude) can be handled in direct numerical calculations for cases when the transition probability matrix is a lower triangular matrix, that is, when all of its elements above the main diagonal are zero. This may seem a very strict condition at first sight. However, practically irreversible reactions are quite common in chemistry and the master equation for these can always be stated with Ω arranged as a lower triangular matrix with a suitably selected enumerating function. These systems also share the characteristic that none of the states can occur twice as time progresses and they have at least one absorbing state. The master equation of a system like this in the matrix format is given as:

$$\begin{pmatrix} \frac{dP_1}{dt}(t) \\ \frac{dP_2}{dt}(t) \\ \vdots \\ \frac{dP_M}{dt}(t) \end{pmatrix} = \begin{pmatrix} -\sum_{i=2}^{M} \omega_{i,1} & 0 & \cdots 0 \\ \omega_{2,1} & -\sum_{i=3}^{M} \omega_{i,2} & \cdots 0 \\ \vdots & \vdots & \ddots \vdots \\ \omega_{M,1} & \omega_{M,2} & \cdots 0 \end{pmatrix} \begin{pmatrix} P_1(t) \\ P_2(t) \\ \vdots \\ P_M(t) \end{pmatrix} \tag{2.62}$$

The eigenvalues of the transition probability matrix in Eq. (2.62) are exactly the elements in the main diagonal. Assuming that all these eigenvalues are different (not uncommon in chemical systems), the solution can be given as follows:

$$\begin{pmatrix} P_1(t) \\ P_2(t) \\ \vdots \\ P_M(t) \end{pmatrix} = \begin{pmatrix} C_{1,1} & 0 & \cdots 0 \\ C_{2,1} & C_{2,2} & \cdots 0 \\ \vdots & \vdots & \ddots \vdots \\ C_{M,1} & C_{M,2} & \cdots 0 \end{pmatrix} \begin{pmatrix} e^{-\sum_{i=2}^{M} \omega_{i,1}} \\ e^{-\sum_{i=3}^{M} \omega_{i,2}} \\ \vdots \\ e^{0} \end{pmatrix} \tag{2.63}$$

The constants $C_{i,j}$ can be calculated from the initial conditions in a successive manner:

$$C_{1,1} = P_1(0)$$

$$C_{n,j} = \frac{\sum_{i=j}^{n-1} \omega_{n,i} C_{i,j}}{\sum_{i=n+1}^{M} \omega_{i,n} - \sum_{i=j+1}^{M} \omega_{i,j}} \quad j < n \tag{2.64}$$

$$C_{n,n} = P_n(0) - \sum_{j=1}^{n-1} C_{n,j}$$

2.4.3 Time-Independent Q-Functions

In cases when the master equation is of the form shown in Eq. (2.62), the concept of time-independent Q functions may be useful if the interesting properties themselves are independent of time [49, 50]. For systems that have no elements of reversibility, there are no mutually accessible states. Therefore, it is often informative to calculate the $Q_{f(a_1,a_2,\ldots,a_n)}$ probability of a given state ever occurring during the entire course of the reaction. It can be shown that $Q_{f(a_1,a_2,\ldots,a_n)}$ is related to the time-dependent function $P_{f(a_1,a_2,\ldots,a_n)}(t)$ as follows:

$$Q_{f(a_1,a_2,\ldots,a_n)} = \lim_{t\to\infty} P_{f(a_1,a_2,\ldots,a_n)}(t) + \int_0^{\infty} \left(\sum_{j=1}^{M} \omega_{f(a_1,a_2,\ldots,a_n),j} \right) P_{f(a_1,a_2,\ldots,a_n)}(t)dt \tag{2.65}$$

The most frequent use of Q functions is to re-arrange the master equation, or perhaps use more direct ways of thought based on the chemical characteristics of

the system, to give a later Q_i as a function of earlier Q_i quantities. This sort of recursive definition often makes it possible to use mathematical induction to prove an analytical formula for any Q_i. The typical implementation of this strategy is when there are several absorbing states in the system: the probability distribution for the final outcome of the reaction will simply be given by the Q_i values. The Q functions may also be useful in cases when time dependent characteristics are sought for a large number of molecules by a combined stochastic and deterministic approach.

2.4.4 Laplace Transformation

An already noted property of the master equation given in Eq. (2.22) or (2.39) is that it is a first-order, linear differential equation with constant coefficients. Therefore, Laplace transformation might occasionally provide a convenient way of looking for the time-dependent solution. The Laplace transform of a function $f(x)$ is defined as:

$$\mathscr{L}(f(x))(s) = \int_0^\infty e^{sx} f(x)dx \tag{2.66}$$

Two properties of the Laplace transform have high relevancy here. The first is linearity, which states that the Laplace transform of a linear combination of functions can be given as the linear combination of the individual Laplace transforms:

$$\mathscr{L}\left(\sum_i \lambda_i f_i(x)\right)(s) = \sum_i \lambda_i \mathscr{L}(f_i(x))(s) \tag{2.67}$$

The second important property is that the Laplace transform of the derivate of a function can be given in terms of the Laplace transform itself:

$$\mathscr{L}\left(\frac{df(x)}{dx}\right)(s) = s\mathscr{L}\left(f(x)\right)(s) - f(0) \tag{2.68}$$

Using these two simple properties, the Laplace transform of Eq. (2.39) is as follows:

$$\begin{pmatrix} s\mathscr{L}(P_1(t))(s) \\ s\mathscr{L}(P_2(t))(s) \\ \vdots \\ s\mathscr{L}(P_M(t))(s) \end{pmatrix} = \begin{pmatrix} P_1(0) \\ P_2(0) \\ \vdots \\ P_M(0) \end{pmatrix} + \Omega \begin{pmatrix} \mathscr{L}(P_1(t))(s) \\ \mathscr{L}(P_2(t))(s) \\ \vdots \\ \mathscr{L}(P_M(t))(s) \end{pmatrix} \tag{2.69}$$

It is seen that Laplace transformation gives rise to a system of linear equations with the independent variable p. So the Laplace transforms of the P_i functions can sometimes be obtained quite readily based on Eq. (2.69). If Ω is a lower triangular matrix as explained in Eq. (2.62), induction is sometimes possible to obtain an easily handled form of $\mathscr{L}(P_i(t))(p)$ functions. Nevertheless, the need for the inverse Laplace transformation of the results or for calculations with complex numbers may occasionally render this strategy burdensome.

2.4.5 Generating Functions

Introducing a generating function of the following form can often be a useful strategy in solving the master equation:

$$G(z_1, z_2, \ldots, z_n, t) = \sum_{\text{all states}} z^{a_1} z^{a_2} \cdots z^{a_b} P_{f(a_1,a_2,\ldots,a_n)}(t) \qquad z_i \in \mathbb{C} \quad i = 1, 2, \ldots, n \tag{2.70}$$

The rationale in constructing this function is that all the individual variables of interest in the stochastic description can be obtained from it in a relatively straightforward manner. The expectation of the number of A_i molecules can be generated using first partial derivatives:

$$\langle a_i \rangle (t) = \frac{\partial G(1, 1, \ldots, 1, t)}{\partial z_i} \tag{2.71}$$

Second order moments and correlations can be given using second and mixed partial derivatives as follows:

$$\left\langle a_i^2 \right\rangle (t) = \frac{\partial^2 G(1, 1, \ldots, 1, t)}{\partial z_i^2} + \frac{\partial G(1, 1, \ldots, 1, t)}{\partial z_i} \tag{2.72}$$

$$\left\langle a_i a_j \right\rangle (t) = \frac{\partial^2 G(1, 1, \ldots, 1, t)}{\partial z_i \partial z_j} \quad i \neq j \tag{2.73}$$

The individual state probabilities can be obtained from the generating function by successive partial derivations:

$$P_{f(a_1,a_2,\ldots,a_n)}(t) = \frac{1}{a_1! a_2! \cdots a_n!} \frac{\partial^{a_1+a_2+\cdots+a_n} G(0, 0, \ldots, 0, t)}{\partial z_1^{a_1} \partial z_2^{a_2} \cdots \partial z_n^{a_n}} \tag{2.74}$$

Usually, all the possible states are enumerated on the right hand side of Eq. (2.70), and summing all the probabilities listed there should give 1. Therefore, the following equation also holds for the generator function at any time t:

$$G(1, 1, \ldots, 1, t) = 1 \tag{2.75}$$

In the case of power law kinetics, summing the M equations in master equation (2.38) transforms it into the following partial differential equation:

$$\frac{\partial G(z_1, z_2, \ldots, z_n, t)}{\partial t} = \sum_{j=1}^{m} \kappa_j \prod_{i=1}^{n} \frac{z_i^{\alpha_{j,i}}}{\alpha_{j,i}!} \left(\prod_{i=1}^{n} z_i^{\nu_{j,i}} - 1 \right) \frac{\partial^{\alpha_{j,1}+\alpha_{j,2}+\cdots+\alpha_{j,n}} G(z_1, z_2, \ldots, z_n, t)}{\partial z_1^{\alpha_{j,1}} \partial z_2^{\alpha_{j,2}} \cdots \partial z_n^{\alpha_{j,n}}} \tag{2.76}$$

The boundary condition is stated by setting $t = 0$ in Eq. (2.70).

Although the generating function is defined very broadly in Eq. (2.70) (on the entire complex plane), its application shows that handling it in the $[0, 1]$ real interval is usually quite sufficient. In fact, the values and derivatives are usually only needed at $z_i = 0$ or $z_i = 1$. This observation means that the generating function contains more information than necessary for the full CDS description of the system. Consequently, solving partial differential equation (2.76) is usually a more difficult problem than solving the original master equation in Eq. (2.38). Yet using a generator function and Eq. (2.76) is not entirely in vain. The main advantage of this approach is that the master equation, which is composed of M ordinary differential equations, is transformed into a single partial differential equation with $(n + 1)$ variables. These variables correspond to the types of molecules present and the single variable of time. However, the user should not forget that inability to solve Eq. (2.76) by no means implies that a closed and easily handled solution does not exist for master equation (2.38).

2.4.6 *Poisson Representation*

Another possible method for solving the master equation (2.22) is called the **Poisson representation** [13, 31, 32]. The technique sets up a Fokker–Planck equation that is equivalent to master equation (2.22). To implement the Poisson representation, the state probabilities are assumed to be given as a superposition of uncorrelated Poisson distributions:

$$P_{f(a_1,a_2,\ldots,a_n)}(t) = \int \prod_{i=1}^{n} \frac{e^{-\mu_i} \mu_i^{a_i}}{a_i!} g(\boldsymbol{\mu}, t) d\boldsymbol{\mu}$$
$$\int \int \cdots \int \prod_{i=1}^{n} \frac{e^{-\mu_i} \mu_i^{a_i}}{a_i!} g(\mu_1, \mu_2, \ldots, \mu_n, t) d\mu_1 d\mu_2 \cdots d\mu_n \tag{2.77}$$

Here, the function $g(\mu_1, \mu_2, \ldots, \mu_n, t)$ is referred to as the quasiprobability distribution. Variables $\mu_1, \mu_2, \ldots, \mu_n, t$ are auxiliary variables: the limits of the integration are specified later, often real numbers between 0 and ∞. It is tempting to interpret $g(\mu_1, \mu_2, \ldots, \mu_n, t)$ as the probability density of the system showing

$\mu_1, \mu_2, \ldots, \mu_n$ as the expectation of molecule numbers of species A_1, A_2, ..., A_n at time t, but this interpretation does not account for the fact that function g may assume negative or even complex values.

At first sight, Eq. (2.77) may seem to be at odds with the properties of functions $P_{f(a_1,a_2,\ldots,a_n,t)}$. Poisson distributions are not bounded, whereas the state space is often finite. This apparent contradiction can be resolved by showing that any discrete distribution can be generated as the superposition of uncorrelated Poisson distributions. This is most conveniently proved by for the purposes of this book by giving a function $g(\mu_1, \mu_2, \ldots, \mu_n, 0)$ that is suitable for use in Eq. (2.77) for any initial conditions:

$$g(\boldsymbol{\mu}, 0) = \sum_{\text{all states}} P_{f(a_1,a_2,\ldots,a_n)}(0) \prod_{i=1}^{n} \frac{e^{a_i} a_i!}{a_i^{a_i}} \delta(\mu_i - a_i) \tag{2.78}$$

In Eq. (2.78), δ stands for the Dirac delta function, which was defined in (2.5).

Another, somewhat semantic source of misunderstanding is the use of the phrase 'uncorrelated Poisson distributions' before introducing Eq. (2.77). At the very least, the conservation equations (see Sect. 2.3.1.) in a system provide a clear source of correlation between some molecule numbers. However, function $g(\boldsymbol{\mu}, t)$ actually incorporates the conservation equations. In fact, it would be much better to state that the Poisson distributions are only correlated through the g function.

In a way, the philosophy behind the introduction of the Poisson representation in Eq. (2.77) is similar to the one used in the definition of generating functions in Eq. (2.70): instead of M distinct single variable continuous functions, the problem is transformed into one involving a single function with $n + 1$ independent (and continuous) variables. In the generating function approach, individual $P_{f(a_1,a_2,\ldots,a_n,t)}$ functions and various expectations can be obtained by differentiating $G(z_1, z_2, \ldots, z_n, t)$, whereas integration is used for the same purpose in the Poisson representation approach. In both cases, the transformed continuous function actually carries more information than the original $P_{f(a_1,a_2,\ldots,a_n,t)}$ functions, a property that is taken advantage of in order to obtain a relatively simple partial differential equation for the transformed continuous functions (G or g).

Another similarity to the generating function approach is that the expectation for a molecule number can be defined easily using function $g(\boldsymbol{\mu}, t)$:

$$\langle a_i \rangle (t) = \int \mu_i g(\boldsymbol{\mu}, t) d\boldsymbol{\mu} \tag{2.79}$$

Furthermore, higher factorial moments relevant to power law kinetics can also be stated in a very straightforward manner:

$$\left\langle \binom{a_1}{\alpha_1} \binom{a_2}{\alpha_2} \cdots \binom{a_n}{\alpha_n} \right\rangle (t) = \int \prod_{i=1}^{n} \frac{\mu_i^{\alpha_i}}{\alpha_i!} g(\boldsymbol{\mu}, t) d\boldsymbol{\mu} \tag{2.80}$$

However, there are notable differences between the generating function and Poisson representation approaches. The initial state probabilities define the $G(z_1, z_2, \ldots, z_n, 0)$ unambiguously, whereas this is clearly not the case for the $g(\boldsymbol{\mu}, 0)$ function. In fact, the generally applicable function constructed from the initial conditions in Eq. (2.78) is just one example and is by no means unique. It is not impossible that mathematical ingenuity or perhaps some special properties of a particular set of initial conditions may occasionally give rise to a $g(\boldsymbol{\mu}, 0)$ function that is much easier handle than the one shown in Eq. (2.78).

Master equation (2.38) can be transformed by substituting the definition given in Eq. (2.77):

$$\begin{aligned}\int \prod_{i=1}^{n} e^{-\mu_i} \mu_i^{a_i} \frac{\partial g(\boldsymbol{\mu}, t)}{\partial t} d\boldsymbol{\mu} = \\ -\sum_{j=1}^{m} \int \kappa_j \prod_{i=1}^{n} \frac{a_i! e^{-\mu_i} \mu_i^{a_i}}{\alpha_{j,i}!(a_i - \alpha_{j,i})!} g(\boldsymbol{\mu}, t) d\boldsymbol{\mu} \\ +\sum_{j=1}^{m} \int \kappa_j \prod_{i=1}^{n} \frac{a_i! e^{-\mu_i} \mu_i^{a_i - \nu_{j,i}}}{\alpha_{j,i}!(a_i - \nu_{j,i} - \alpha_{j,i})!} g(\boldsymbol{\mu}, t) d\boldsymbol{\mu}\end{aligned} \tag{2.81}$$

Further re-arrangements are made in this form without adding other equations. First, mass action type kinetics is assumed as given in Eq. (1.8). Then integration by parts is used successively with the aim of eliminating all molecule numbers $a_1, a_2, \ldots, a_n$ as multiplicators (but not as exponents) and stoichiometric coefficients $\nu_{j,i}$ from Eq. (2.81). Finally, the functions are set to be equal on the right and left hand site of the equation instead of the definite integrals, and the following partial differential equation can be obtained for $g(\boldsymbol{\mu}, t)$:

$$\begin{aligned}\frac{\partial g(\boldsymbol{\mu}, t)}{\partial t} = -\sum_{j=1}^{m} \kappa_j \left(\prod_{i=1}^{n} \frac{1}{\alpha_{j,i}!} \right) \left[1 - \prod_{i=1}^{n} \left(1 - \frac{\partial}{\partial \mu_i} \right)^{\alpha_{j,i}} \right] \Xi(\boldsymbol{\mu}, t) \\ \Xi(\boldsymbol{\mu}, t) = \mu_1^{\alpha_{j,1}} \mu_1^{\alpha_{j,1}} \cdots \mu_n^{\alpha_{j,n}} g(\mu_1, \mu_2, \ldots, \mu_n, t)\end{aligned} \tag{2.82}$$

Equation (2.82) is called a generalized Fokker–Planck equation. At this point, it should be noted that although Eq. (2.81) follows from Eq. (2.82), the reverse is not strictly true because of the step where the integration is dropped in during the derivation.

Interestingly, the concepts of Poisson representation and the generating function can be used simultaneously, so the generating function of the probabilities defined as superposition of uncorrelated Poisson distributions in Eq. (2.77) can be given as follows:

$$G(z_1, z_2, \ldots, z_n, t) = \int e^{\sum_{i=1}^{n} (z_i - 1)\mu_i} g(\boldsymbol{\mu}, t) d\boldsymbol{\mu} \tag{2.83}$$

An alternative derivation of Eq. (2.82) may be based on Eq. (2.83).

Again, the reader should be warned that inability to solve Eq. (2.82) does not imply that nothing useful about the solution of master equation (2.38) can be stated by other methods. In fact, it seems that the main value of the Poisson representation is that the methodology it describes can be used generally in each case to obtain the generalized Fokker–Planck equation (2.82). However, this is seldom the easiest or the most elegant route to the solution of master equation (2.38).

2.4.7 Mathematical Induction

In a few cases, mathematical induction is a useful method to prove a closed form of a solution that is conjectured based on some heuristic line of thinking. This is especially the case when time-independent Q functions are used, as they very often lead to recursive equations starting from a value known from the initial conditions. However, simply substituting the solution into the master equation often serves the same purpose as mathematical induction.

Nevertheless, mathematically informed intuition based inductive (and often heuristic) lines of thought seems to be the richest source of obtaining closed solutions of master equation (2.38) in specific cases. For these, a careful analysis of the structure of the state space is often a very useful starting point.

2.4.8 Initial Conditions Given with a Probability Distribution

Some of the strategies presented in the previous subsections assumed that the initial state is a certain single state. This assumption might seem to limit the general applicability of algorithms. However, the linear nature of the master equation is very forgiving in this respect. If it is not convenient to use the initial distribution during the solution of the master equation, it is always possible to obtain the general solution as a linear combination of the particular solutions starting from certain initial states, where the factors used in the linear combination are the initial probabilities. If the particular solution of the master equation from the certain initial state $(a_1, a_2, \ldots, a_n)$ is denoted $\mathbf{P}(t)^{(a_1,a_2,\ldots,a_n)}$, the general solution from an initial distribution $\mathbf{P}(0)$ (with elements $P_{f(a_1,a_2,\ldots,a_n)}(0)$) is simply given as:

$$\mathbf{P}(t) = \sum_{\text{all states}} P_{f(a_1,a_2,\ldots,a_n)}(0)\mathbf{P}(t)^{(a_1,a_2,\ldots,a_n)} \tag{2.84}$$

In fact, this technique based on the linear combination of single-state derived properties has already been used in this book occasionally. For example, this was the underlying logic used to construct Eq. (2.78).

2.4.9 Time Scales

As has already been stated and emphasized by the matrix formalism, master equation (2.39) is a system of simultaneous linear, first-order, ordinary differential equations. Therefore, the eigenvalues of matrix Ω, denoted here as $\lambda_1, \lambda_2, \ldots, \lambda_M$, have special significance in the solution as they each represent the exponent of one term in a multiexponential function. The reciprocals of the eigenvalues are called life times corresponding to each term. These life times give a rough idea about the characteristic time scale of changes in a system. Many systems that do not exclusively contain first order processes feature three different groups of live times, which are referred to as three different time scales [60], especially in biochemical reactions networks:

1. The time scale of individual reactions, called the molecular signaling time scale (t_{ms}) in cellular biochemistry
2. The time scale of nonlinear network dynamics (t_{nd})
3. The time scale of the transitions between domains of attraction, called cellular evolution (t_{ce})

In nonlinear deterministic kinetics (i.e. at least one reaction in the network is not first order), a long time means $t \gg t_{nd}$ but not $t > t_{ce}$. On this time scale, a system reaches an attractor determined by the initial state. At times $t \gg t_{ce}$, the system is characterized by a probability distribution between the two domains of attraction.

In addition, it is to be noted that there is a great separation of time scales between (t_{nd}) and (t_{ce}) when the populations are large. For times $t \gg t_{nd}$ but not $t > t_{ce}$, the behavior of the system follows a bifurcation diagram. On the time scale $t > t_{ce}$, this bifurcation diagram has to be modified by the Maxwell construction.

2.5 Stationary and Transient Distributions

Thus far, considerations in this chapter were focused on the time dependence of individual state probabilities. However, it is customary and also often useful to think about the state probabilities collectively as forming a discrete distribution. Maxima, minima and zero values occurring on this distribution are normally used as descriptors to characterize the solution of a problem as a whole.

2.5.1 Stationary Distributions

A stationary distribution, already defined in Eq. (2.11), is the direct stochastic analog of deterministic stationary state concentrations defined in Eq. (1.25). A stationary

distribution is a vector, usually denoted as $\boldsymbol{\pi}$, and satisfies the following equation, which is derived from Eq. (2.39) by setting the left side 0:

$$0 = \Omega \boldsymbol{\pi} \tag{2.85}$$

In the deterministic approach, Eq. (1.25) does not usually have a unique solution. However, if the conservation equations arising from the initial conditions are considered, the number of stationary states is typically finite and often one. In the stochastic approach, the conservation equations are already satisfied by limiting the state space to possible states. Yet, Eq. (2.85), which is always linear, never has a unique solution (if it had, it would be the trivial solution where all state probabilities are 0). As already explained, matrix Ω is always a singular matrix. In fact, it is a negative semidefinite matrix meaning that the only negative elements occur in the main diagonal, but these elements are dominant in each column.

In addition to satisfying Eq. (2.85), a stationary distribution $\boldsymbol{\pi}$ must fulfil two additional conditions. The first is that all the probabilities should be non-negative. The second is that the sum of all probabilities should be one (normalization):

$$\sum_{\text{all states}} \pi_{f(a_1,a_2,\ldots,a_n)} = 1 \tag{2.86}$$

Even with the non-negativity and Eq. (2.86), a stationary distribution may or may not be uniquely determined. If it is, then the stationary distribution is independent of the initial conditions. The opposite case is more common. In fact, if at least one of the states is final, its stationary probability cannot be determined from Eq. (2.85) as all of its coefficients are zero. Section 2.2.4 gives some more information on the uniqueness of stationary distributions.

In typical cases, the stationary distribution $\boldsymbol{\pi}$ can also be thought of as the limit of the vector of probability functions with time approaching infinity:

$$\pi_{f(a_1,a_2,\ldots,a_n)} = \lim_{t\to\infty} P_{f(a_1,a_2,\ldots,a_n)}(t) \tag{2.87}$$

This means that if a stationary distribution exists (which is not guaranteed), it is uniquely determined by the initial conditions. A criterion for the existence and uniqueness of the stationary distribution follows from the Karlin–McGregor condition for birth-and-death processes [45].

Although Eq. (2.85), even combined with Eq. (2.87), is a system a simultaneous linear equations, the solution is often more problematic than it would be expected. The primary source of the problems is the typical large number of states, which is translated into huge dimensions of matrix Ω. Numerical solution methods based on elimination are often less useful than iterative approaches. Sometimes, special techniques may be developed that suit a special system also making use of the fact that Ω is a sparse matrix.

Under certain conditions, the stationary distribution can be obtained without dealing with Eq. (2.85) based on the concepts of statistical mechanics:

$$\pi_{f(a_1,a_2,\ldots,a_n)} = \frac{M_{f(a_1,a_2,\ldots,a_n)} e^{E_{f(a_1,a_2,\ldots,a_n)}}}{\sum_{\text{all states}} M_{f(a_1,a_2,\ldots,a_n)} e^{E_{f(a_1,a_2,\ldots,a_n)}}} \tag{2.88}$$

In Eq. (2.88), $M_{f(a_1,a_2,\ldots,a_n)}$ is the multiplicity and $E_{f(a_1,a_2,\ldots,a_n)}$ is the energy of state $(a_1, a_2, \ldots, a_n)$. The function in the enumerator on the right is usually called **partition function**. If the model shows detailed balance, the formula written in Eq. (2.88) can always be used. In other cases, it might still be useful, sometimes for a closed subgroup of states [17, 18].

2.5.2 *Transient Distributions*

Transient distributions can be defined at any time instant by collecting all state probabilities. Unlike stationary distributions, transient distributions can only be obtained from the full solution of master equation (2.39). It is often useful to think of the full solution as the time dependence of the transient distribution. In many cases, it is sufficient to characterize the transient distribution in a semiquantitive manner: listing the number or positioning of minima or maxima is often thought to be important.

2.5.3 *Properties of Stationary and Transient Distributions: Unimodality Versus Multimodality*

The view that the stationary distribution is necessarily Poissonian was commonly held at some point [53, 59]. However later results clearly proved that such a case is exceptional [68] rather than regular [69, 74]. Still later, an interesting connection was demonstrated between the deterministic and the stochastic kinetic approaches. A sufficient condition of the existence of the product form stationary distribution is complex balancing, a property also implying the regular behavior in deterministic kinetics [2].

At this point, it should be recalled that state spaces and, consequently, stationary distributions are typically finite in CDS models. Therefore, some more explanation is needed before going into the details of comparison with infinite Poisson distributions. Another problem is that a state cannot generally be identified with a single molecule number. Earlier considerations in this book used enumerating functions to solve the similar problem of forming a single vector of the possible states, but it was also pointed out that these enumerating functions are not unique. A possible way to overcome this problem would be to use multivariate Poisson distributions. Yet, much like in the monovariate case, Poisson distributions can only be taken literally when mass conservation equations are absent. Most of the following discussion will be limited to monovariate cases.

When the size of the state space is sufficiently large, the infinite nature of a Poisson distribution has little relevance as only a miniscule part of the probability values would fall into the region that is ab ovo impossible because of the finite nature of the stationary distribution. Still, a mathematically more acceptable way to make this statements would use truncated (and finite) versions of the Poisson distribution [57]. In terms of Eq. (2.88), the Poisson distribution actually specifies the possible form of the multiplicities $M_{f(a_1,a_2,\ldots,a_n)}$. Even in the strongly restrictive case based on the assumption that the grand-canonical distribution of the particle number of an ideal mixture is Poissonian, the stationary distributions obtained were not Poissonian in general [72]. Furthermore, a particular chemical reaction exhibiting a non-Poisson distribution was also simulated [75]. A sufficient condition to obtain a Poissonian stationary distribution was found for a certain class of birth-and-death processes [73], but the assumptions were slightly different from those typical of chemical reactions. A necessary and sufficient condition for a simple birth and death type process to feature a Poissonian stationary distribution (if it actually has a unique stationary distribution) is that the process should be linear [23]. Further conclusions were drawn for polynomial simple Markov population processes [46]. When the detailed balance conditions hold, special relations between the coefficients are necessary for the stationary distribution to be Poissonian [68–70].

The property of unimodality and multimodality is connected to the number of extrema on the stationary or transient distribution. Again, multivariate definitions of uni- and multimodality must be used for general cases [16, 67], but the typical and very often only tacit assumption is that only monovariate cases are handled. In a case like this, this single variable is suitable as a trivial enumerating function and a distribution π_i is said to be unimodal if the series $\pi_{i+1}-\pi_i$ has precisely one change of sign. A number of useful theorems and conditions for uni- and multimodality are found in the work of Medgyessy [54, 55].

Earlier, the appearance of multiple stationary states in the deterministic model was assumed to be strongly connected to multimodality in the stochastic stationary distribution. In particular, the number and location of stable stationary points in the deterministic approach were ascribed to maxima in the stochastic distribution, whereas unstable stationary points were similarly correlated with minima. These assumptions were first shown to be wrong using non-kinetic examples in stochastic catastrophe theory [14, 20]. The Schlögl reaction was used particularly frequently to study the exotic behavior of deterministic and stochastic models [21, 26]. It seems likely that multistationarity and multimodality correspond to each other asymptotically with increasing volume [6, 10, 28, 43].

Useful forms for transient distributions are typically difficult to obtain symbolically. It would be natural to assume that processes leading to a unimodal stationary distribution involve unimodal transient distributions. However, this is not true without exceptions: the phenomenon of **transient bimodality** is well known from the stochastic analysis of explosive reactions [5, 29, 30].

2.6 Simulation Methods

Stochastic simulation of chemical reactions emerged as an alternative method to solve differential equations numerically for identifying reaction mechanisms [63]. Lindblad and Degn [52] introduced a discrete time discrete space stochastic (DDS) model (i.e. a Markov chain). The probability that a certain elementary reaction will be selected in a time instant is proportional to the rate constants and to the particle numbers of the reactants. Having selected a reaction, the value of the state vector is modified according to the stoichiometry. The iteration of the procedure generate a realization of a process. The algorithm was recommended to use Monte Carlo simulation in the education of chemists [56, 61].[1]

The celebrated Gillespie algorithm [34, 35, 37, 39] is based on answering two questions: "...when will the next reaction occur, and what kind of reaction will it be?..." by defining the **reaction probability density function** $P(\tau, \mu)$. Accordingly, $P(\tau, \mu)d\tau$ means the probability at time t that the *next* reaction will occur in the differential time interval $(t + \tau, t + \tau + d\tau)$, *and* it will be an R_μ reaction. It is calculated as the probability of that **no** reaction occur in the given time interval by multiplying with the probability of the occurrence of the reaction R_μ. The probability of the *no reaction* is an exponential function of the time interval τ. (The *no reaction* interval can be identified with the *waiting time* used in the theory of stochastic process. No reference was given to the Feller–Doob theorem by Gillespie.)

Simulations are usually carried out when the master equation of the process cannot be handled in any other meaningful way. Alternatively, they may be used for an initial study of a particular system and serve as a source of conjectures that can be proved properly based on the master equation.

In simulations, a number of realizations (sample paths) are generated using random numbers primarily on DDS principles [33, 34, 62]. In a way, this can be thought of as a directed random walk in an n-dimensional state space. An almost exclusive approach is based on generating evenly distributed random numbers between 0 and 1 (rnd), two such numbers are needed in each DDS jump. For simplicity, the description here relies on the stochastic transition rates defined in Eq. (2.37). Two series are generated, the first is a time series t_k, the second is a state series $\mathbf{n}_k$. The first element of the time series is the initial time, usually $t_1 = 0$, whereas the first element in the state series is the initial state i, $S_1 = i$. The successive members of these two series are defined in a recursive manner:

[1] Everybody has a story how (s)he missed to be a world champion. The senior author of this book and his friends used this algorithm [22, 25, 64, 65]. They have probabaly heard about Doob's theorem first from his book [19] from János Tóth around 1973 (an early version is due to Feller [27]. Their (our :-)) student/colleague Vera Hárs implemented an algorithm known now as the *Gillespie algorithm* in her master thesis [42]. Why we did not publish it? Well, those were different times. Our consolation is that everybody has a predecessor. Patrick Hanusse [41] published a paper in French, by using a rather similar algorithm.

$$t_{k+1} = t_k - \frac{\log(rnd)}{\sum_{j=1}^{n} v_j(\mathbf{n})} \tag{2.89}$$

$$\mathbf{n}_{k+1} = \mathbf{n}_k + (\nu_{i,1}, \nu_{i,2}, \ldots, \nu_{i,n}) \quad \text{if} \quad \frac{\sum_{j=1}^{i-1} v_j(\mathbf{n})}{\sum_{j=1}^{n} v_j(\mathbf{n})} \leqq rnd < \frac{\sum_{j=1}^{i} v_j(\mathbf{n})}{\sum_{j=1}^{n} v_j(\mathbf{n})} \tag{2.90}$$

It should be emphasized that the random numbers (rnd) used in Eqs. (2.89) and (2.90) must be independent, otherwise some correlation will be introduced between the time and state series. The values of $\nu_{i,j}$ are the stoichiometric coefficients defined in Eq. (1.2), whereas the quantity $v_j(\mathbf{n})$ is the transition rate of reaction j in state $\mathbf{n}$ as given by Eq. (2.37). The term **propensity function** is also often used for $v_j(\mathbf{n})$. The presented method itself is often referred to as **exact stochastic simulation algorithm** (or exact SSA).

Generating a single sample path by simulation usually does not serve any practical purposes. Repetitions, that is, multiple independent simulated sample paths from the same initial conditions are necessary to learn certain properties of the investigated process. For example, the absolute state probabilities can in principle be approximated: if the overall number of simulated trajectories is N, and $N_{\mathbf{n}}(t)$ of these show the system in state $\mathbf{n}$, or $(a_1, a_2, \ldots, a_n)$, at time t, $P_{\mathbf{n}}(t)$ is estimated as:

$$P_{f(a_1,a_2,\ldots,a_n)}(t) = \frac{N_{\mathbf{n}}(t)}{N} \tag{2.91}$$

Using Eq. (2.91) is seldom a particularly practicable way of characterizing simulation results. Such efforts are normally much more successful if the estimated property is some sort of cumulative probability, the particle number of a species present, or the time necessary to reach certain pre-defined target conditions.

The number of recursive simulation steps calculated in a sample path is usually rather arbitrarily decided, the only factor being the final purpose of the simulation studies. In irreversible reaction systems subject to mass conservation such as the one described by master equation (2.62), all sample paths are necessarily finite. In other systems, infinite sample paths are possible so the simulation algorithm must include testing for some sort of stop conditions. Generally, if there is an absorbing state, there are possible finite sample paths. An element of reversibility in the reactions, on the other hand, guarantees the existence of infinite sample paths.

The main advantage of simulations lies in the simplicity of calculations. Despite the formulation of Eq. (2.90), there is no need for actually enumerating all the possible states, it is enough to deal with those that are accessible from $\mathbf{n}_k$. This fact often greatly reduces the computation time requirements. For some purposes, it is not even necessary to store the elements of the two series t_k and $\mathbf{n}_k$, the recursive calculations only need the current members of the two series because of the memorylessness of Markov chains. Thus, simulations are usually not restricted by the available memory, and increasing the computational power usually results in significant improvements in terms of the size of the system that can be handled by simulations.

These advantages come at a cost: simulations usually provide information on a minute fraction of the entire state space. On the positive side, this information is normally gained exactly about the states that occur with the highest probability, which are the states worth knowing about. Less likely states can be characterized by increasing the number of repetitions carried out, but the number of possible states usually imposes a fundamental limitation on this strategy.

Elements of reversibility in the reactions studied often slow down simulations because the sample path may easily move back and forth between two states or two distinct groups of states. There are a number of published tricks for accelerating simulations in such cases.

The most common accelerating method is referred to as **tau leaping** [36]. The essence of the method is choosing a suitable time interval τ, during which the transition rates do not change appreciably. Instead of taking one step at a time for a time period $t_{k+1} - t_k$, the random number generator is used to take several reactions steps during the fixed time τ. For each reaction, Poisson random numbers are generated to simulate the number of reaction events in the time period τ. This can be achieved using the commonly available generator of uniformly distributed random numbers between 0 and 1 (rnd):

$$b_j = i \quad \text{if} \quad \sum_{k=0}^{i} \frac{(v_j \tau)^k}{k!} e^{-v_j \tau} \leqq rnd < \sum_{k=0}^{i+1} \frac{(v_j \tau)^k}{k!} e^{-v_j \tau} \tag{2.92}$$

The τ-leaping step taken then is:

$$\mathbf{n}(t + \tau) = \mathbf{n}(t) + \sum_{j=1}^{m} b_j (v_{j,1}, v_{j,2}, \ldots, v_{j,n}) \tag{2.93}$$

Obviously, the cornerstone of the tau leaping method is choosing the suitable time interval τ. Too large a time step will fail to meet the leap condition (small changes in transition rates). Too small a time step will result in a large number of leaps without any chemical reaction, which is considerably less efficient than the original simulation method described by Eqs. (2.89) and (2.90). In essence, this is a well-known problem that arises during the numerical solution of the ordinary differential equations such as deterministic rate equation (1.1): too small time steps make the algorithm impossibly slow, too large time steps lead to unreasonable results. To add some confusion to selecting suitable τ values for leaping, the paper that introduced the method itself recommended a flawed strategy [36], which led to substantial inaccuracies and occasionally caused the population(s) of one or more reactant species to go negative [12]. The solution to the problem can also be borrowed from the standard strategies of the numerical solution of ordinary differential equations. Testing a computed step usually reveals if the selected value of τ is too large. The step can be rejected and τ lowered before re-taking the step.

This is called postleap checking [1], and a favorable way to implement this is to select a target ϵ first, then calculate an efficient τ value with the following equation:

$$\tau = \min_{i=1,2,\ldots,n} \left(\frac{\max(\epsilon a_i(t)/g_i, 1)}{|\sum_{j=1}^{m} \nu_{j,1} v_j(\mathbf{n}(t))|}, \frac{[\max(\epsilon a_i(t)/g_i, 1)]^2}{\sum_{j=1}^{m} \nu_{j,1}^2 v_j(\mathbf{n}(t))} \right) \tag{2.94}$$

In Eq. (2.94), the quantity g_i is defined for each A_i species as a pre-determined function that gives the overall order of the highest-order reaction in which A_i is a reactant. After calculating a possible τ-step with this interval, the step is only accepted if it satisfies the following **leap condition** for all species ($i = 1, 2, \ldots, n$):

$$|a_i(t + \tau) - a_i(t)| \leq \max(\epsilon a_i(t)/g_i, 1) \tag{2.95}$$

If the leap condition is not met, the leaping step is retried with a suitably reduced value of τ. Occasionally, the step may simply be re-tried with a new set of Poisson random numbers. Sometimes it is also useful to test if any reaction steps have actually occurred during the interval τ to avoid excessively long computation times.

The analogy of the problem of selecting τ-values with the time step problem of ordinary differential equations is by no means accidental. In fact, the Euler method of the numerical solution of deterministic rate equation (1.1) is obtained if the Poisson random numbers in (2.92) are replaced by their expectation, $v_j \tau$.

Another accelerating trick is called **R leaping** [3]. In effect, a leap is defined here with the number of reaction steps taken rather than the time these steps take. The deterministic equivalent of this is numerical integration of the ordinary differential equations with one of the concentrations as the independent variable and considering time as a dependent variable. In programming this algorithm, Eq. (2.93) is used in a different way: a set number of reaction steps (l) is selected first so that $l = \sum_{j=1}^{m} b_j$, and then the time interval τ is assigned by generating a gamma distributed random variable with shape l and rate $\sum_{j=1}^{n} v_j(\mathbf{n})$. For $l = 1$, the gamma distribution is the same as the exponential distribution, and the time τ can be given as in Eq. (2.89) with $\tau = t_{k+1} - t_k$. The individual integer b_j values are then randomly generated based on a multinomial distribution with the individual probabilities $v_j(\mathbf{n}) / \sum_{j=1}^{n} v_j(\mathbf{n})$.

A still different sort of acceleration of the simulation method can be obtained by uniting the variables that are connected in fast equilibrium reactions [17]. For example, if species A_1 and A_2 are connected through a fast 'isomerization' equilibrium (so that the reversible reaction connecting them is mass action type), then the simulation can be more efficiently run using only the sum of the two molecule numbers: $\overline{a_{12}} = a_1 + a_2$. In this formalism, A_1 and A_2 become the two states of a generic $\overline{A_{12}}$ species. Knowledge of the equilibrium constant of the isomerization (or the rate constants of the forward and reverse reactions) enables the calculation of the time-independent probability p of finding a certain molecule in the A_1 state as opposed to the A_2 state. Therefore, whenever the molecule number for A_1 is needed, it can be given as $p\overline{a_{12}}$. Conversely, $(1 - p)\overline{a_{12}}$ can be used to

replace a_2 in calculations. In effect, this accelerating method reduces the state space of the problem. Deterministic kinetics features a direct analog of this method, which is called **steady-state assumption**.

2.7 Deterministic Continuation

As pointed out in the previous sections, stochastic kinetic calculations are seldom viable up to high particle numbers normally encountered in chemistry. It has also been stated that the stochastic approach to chemical kinetics is often only important for relatively small systems. Nevertheless, certain mechanisms may conserve or even amplify initial fluctuations. In such cases, a method called deterministic continuation [17, 50] may prove very useful.

In this method, stochastic kinetic calculations are carried out either with the master equation or with simulations until a pre-set number of molecular events occur or product molecules are formed. From this point on, the calculations are continued by the usual deterministic method described in Sect. 1.1. Basically, this means that the differential equation (1.1), which is directly equivalent to the stochastic model used, is solved with the initial concentration vector given from the results of stochastic calculations. As stochastic calculations are usually done based on particle numbers, it is often useful to formulate Eq. (1.1) in a way that particle numbers appear in it as dependent variables rather than concentrations. This can always be done by multiplying or dividing with the volume of the system and Avogadro's constant.

Sometimes, only some static property of the system (such as the final state reached) is important, and the actual time dependence of the particle numbers is not of much interest because there are no experimental data as a function of time. In a case like this, it could always be a good strategy to re-formulate Eq. (1.1) in a way that time is dropped from the list of variables, and a purposefully selected particle number is used instead. This technique is often useful in overcoming problems related to the Stiff numerical property of the differential equations. In this technique, it is important to select a particle number that changes monotonously in time as a substitute independent variable. As an example, the simultaneous differential equations of the Michaelis–Menten mechanism given in Eq. (1.21) in Sect. 1.1 will be transformed into a system not using time as a variable any more. The reaction sequence $A_1 + A_2 \longrightarrow A_3$, $A_3 \longrightarrow A_1 + A_2$, and $A_3 \longrightarrow A_1 + A_4$ shows that A_4 only occurs as a product, therefore its concentration must change monotonously in time, and is suitable for using as an alternative independent variable. Incidentally, a more insightful analysis would show that A_2 also usually changes monotonously in this system, but not A_1 and A_3. The system of transformed time-independent equations is essentially obtained by dividing the individual differential equations with that describing the concentration change of A_4:

$$\begin{aligned}
\frac{d[\mathrm{A}_1]}{d[\mathrm{A}_4]} &= \frac{-k_1[\mathrm{A}_1][\mathrm{A}_2]}{k_3[\mathrm{A}_3]} + \frac{k_2}{k_3} + 1 \\
\frac{d[\mathrm{A}_2]}{d[\mathrm{A}_4]} &= \frac{-k_1[\mathrm{A}_1][\mathrm{A}_2]}{k_3[\mathrm{A}_3]} + \frac{k_2}{k_3} \\
\frac{d[\mathrm{A}_3]}{d[\mathrm{A}_4]} &= \frac{k_1[\mathrm{A}_1][\mathrm{A}_2]}{k_3[\mathrm{A}_3]} - \frac{k_2}{k_3} - 1
\end{aligned} \tag{2.96}$$

Although this set of equations seems more complicated than the original one at first sight, in fact, it usually has much better properties for numerical integration. In effect, this approach calculates the trajectory defined by the deterministic rate law in the concentration space without specifying the time values. In this special case, the time-independent set of differential equations is autonomous. This property is not general, the concentration used as a substitute independent variable may appear in more or more of the equations on the right-hand side as well.

2.8 Continuous State Approximations

Continuous state stochastic approaches are also frequently used to describe chemical processes. A direct analog of the CDS approach would be to assume that concentration (or molecule numbers) can take any value from a range instead of being discreet. In this approach, the number of possible states is necessarily infinite, but their continuous nature opens the possibility of using new and sometimes advantageous mathematical techniques. As state probabilities should still add up to 1, only a finite number of states may have nonzero probability. At this point, it must be recalled that zero probability is not the same as an impossible event. In typical continuous state approaches, all individual state probabilities are zero and are consequently of little practical use. The description should rely on the multivariate cumulative distribution function, $F(x_1, x_2, \ldots, x_n)$ for n independent variables, which represents the probability that the random variable X_i takes on a value less than or equal to x_i for all values of index i. This function is monotonically non-decreasing and right-continuous for each of its variables. Again, the time dependence of this cumulative distribution function is of interest, the notation $F(x_1, x_2, \ldots, x_n, t)$ is often used.

The probability density function $f(x_1, x_2, \ldots, x_n, t)$ is often more useful for calculations and is defined as:

$$f(x_1, x_2, \ldots, x_n, t) = \frac{\partial^n F(x_1, x_2, \ldots, x_n, t)}{\partial x_1 \partial x_2 \cdots \partial x_n} \tag{2.97}$$

The function $f(x_1, x_2, \ldots, x_n, t)$ is the closest equivalent of the state probabilities $P_{f(a_1,a_2,\ldots,a_n)}(t)$ in the discreet approach. The definition of the cumulative distribution function $F(x_1, x_2, \ldots, x_n)$ can be used directly in the discreet state

space as well, the values will be equal to the sum of the probabilities of all states for which $a_1 \leq x_1, a_2 \leq x_2, \ldots, a_n \leq x_n$.

It should also be mentioned that spatially non-homogeneous systems are very common in the continuous state approaches, i.e. some of the variables of the cumulative distribution function or the probability density function may be spatial coordinates.

The **Kramers–Moyal–Stratonovich equation** [47, 58, 66] equation can be derived for a probability density function using the jth velocity function of conditional moments, denoted as D_j. In the physical literature, it is also called Kramers–Moyal extension and takes the following form for the single variable case:

$$\frac{\partial f(x,t)}{\partial t} = \sum_{j=1}^{\infty} \frac{1}{j!} \left(-\frac{\partial}{\partial x}\right)^j \left[D^j(x,t) f(x,t)\right] \tag{2.98}$$

An analogue of the master equation (2.22) is the **Fokker–Planck equation**, which describes the time evolution of the probability density function of the position of a particle undergoing Brownian motion in a fluid. This equation is derived from Eq. (2.98) by truncating the expansion after the second term. In one spatial dimension, $D^1(x,t)$ is called drift, $D^2(x,t)$ is diffusion, and the Fokker–Planck equation is stated as:

$$\frac{\partial}{\partial t} f(x,t) = -\frac{\partial}{\partial x}\left[D^1(x,t) f(x,t)\right] + \frac{\partial^2}{\partial x^2}\left[D^2(x,t) f(x,t)\right] \tag{2.99}$$

More generally, the time-dependent probability distribution may depend on a vector $\mathbf{x}$ of N macrovariables x_i. The general form of the Fokker–Planck equation is then:

$$\frac{\partial f(\mathbf{x},t)}{\partial t} = -\sum_{i=1}^{N} \frac{\partial}{\partial x_i}\left[D_i^1(\mathbf{x}) f(\mathbf{x},t)\right] + \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial^2}{\partial x_i\, \partial x_j}\left[D_{ij}^2(\mathbf{x}) f(\mathbf{x},t)\right] \tag{2.100}$$

where D^1 is the drift vector and D^2 the diffusion tensor; the latter results from the presence of a random force.

In statistical physics, a **Langevin equation** is a stochastic differential equation describing Brownian motion of charged particles in a potential (see also (1.28) and (1.31)).

The first Langevin equations studied were those in which the potential is constant, so that the acceleration $\mathbf{a}$ of a Brownian particle of mass $\mathbf{m}$ is expressed as the sum of a viscous force which is proportional to the particle's velocity $\mathbf{v}$ (by Stokes' law), a noise term $\boldsymbol{\eta}(t)$ (the name given in physical contexts to terms in stochastic differential equations which are stochastic processes) representing the effect of a continuous series of collisions with the atoms of the underlying fluid, and $\mathbf{F}(\mathbf{x})$ which is the systematic interaction force due to the intramolecular and intermolecular interactions:

$$m\mathbf{a}(t) = m\frac{d\,\mathbf{v}(t)}{dt} = \mathbf{F}(\mathbf{x}(t)) - \beta\mathbf{v}(t) + \boldsymbol{\eta}(t) \tag{2.101}$$

Example 2.10. A CCS process W_t is called **Wiener process** if it satisfies the following three conditions:

1. $W_0 = 0$
2. W_t is almost surely continuous,
3. W_t has independent increments with normal (Gaussian) distribution, i.e. $W_t - W_s \sim N(0, t-s)$.

Here $N(\mu, \sigma^2)$ denotes the normal distribution with expectation μ and variance σ^2. The Wiener process plays a key role in describing Brownian motion.

The basic properties of the Wiener process:

1. The expectation is zero: $\langle W_t \rangle = 0$.
2. The variance is t: $\left\langle W_t^2 \right\rangle - \langle W_t \rangle^2 = t$.
3. Its covariance and correlation are: $\operatorname{cov}(W_s, W_t) = \min(s,t) \quad \operatorname{corr}(W_s, W_t) = \frac{\min(s,t)}{\sqrt{st}} = \sqrt{\frac{\min(s,t)}{\max(s,t)}}$.
4. The unconditional probability density function at a fixed time t is: $f_{W_t}(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}$.

Example 2.11. The **Ornstein–Uhlenbeck process** [71] also known as the mean-reverting process, is a stochastic process $\{X(t)\}$ obeying the following stochastic differential equation:

$$dX(t) = \theta(\mu - X(t))\,dt + \sigma\,dW(t), \tag{2.102}$$

where $\theta > 0, \mu$ and $\sigma > 0$ are parameters and $W(t)$ denotes the Wiener process. The Ornstein–Uhlenbeck process is a special case of a Gaussian process that has a bounded variance and admits a stationary probability distribution, in contrast to the Wiener process; the difference between the two is in their drift term. For the Wiener process, the drift term is constant zero, whereas for the Ornstein–Uhlenbeck process, it is dependent on the current value of the process: if the current value of the process is less than the (long-term) mean, the drift will be positive; if the current value of the process is greater than the (long-term) mean, the drift will be negative. In other words, the mean acts as an equilibrium level for the process. This gives the process its informative name, "mean-reverting." The stationary (long-term) variance is given by $\operatorname{var}(X(t)) = \frac{\sigma^2}{2\theta}$. The Ornstein–Uhlenbeck process is the continuous-time analogue of the discrete-time first order autoregressive process.

It is possible (and often convenient) to represent $X(t)$ (unconditionally) as a scaled time-transformed Wiener process:

$$X(t) = \mu + \frac{\sigma}{\sqrt{2\theta}} W(e^{2\theta t}) e^{-\theta t} \tag{2.103}$$

or conditionally (given $X(0)$) as

$$X(t) = X(0)e^{-\theta t} + \mu(1 - e^{-\theta t}) + \frac{\sigma}{\sqrt{2\theta}} W(e^{2\theta t} - 1)e^{-\theta t} \tag{2.104}$$

The Ornstein–Uhlenbeck process can be interpreted as a scaling limit of a discrete process, in the same way that Brownian motion is a scaling limit of random walks. Consider an urn containing n red and smaragdite balls. At each step a ball is chosen at random and replaced by a ball of the opposite color. Let X_n be the number of red balls in the urn after n steps. Then $\frac{X_{[n]} - n/2}{\sqrt{n}}$ converges to an Ornstein–Uhlenbeck process, as n tends to infinity.

In the simplest case, the solution of the Langevin equation is an Ornstein–Uhlenbeck process [44, 71].

The Fokker–Planck equation describing the distribution $f(x, t)$ of the Ornstein–Uhlenbeck process is given by

$$\frac{\partial f(x,t)}{\partial t} = \theta \frac{\partial}{\partial x}((x - \mu) f(x, t)) + \frac{\sigma^2}{2} \frac{\partial^2 f(x,t)}{\partial x^2} \tag{2.105}$$

This equation has the following stationary solution:

$$f_s(x) = \sqrt{\frac{\theta}{\pi\sigma^2}}\, e^{-\theta(x-\mu)^2/\sigma^2} \tag{2.106}$$

2.9 Non-Markovian Approaches

Non-Markovian approaches to stochastic chemical kinetics are sometimes found in the literature [11, 40]. However, there is an intellectual catch here, which will be illustrated by a simple example. The outflow reaction $A_1 \longrightarrow \emptyset$ with rate constant λ, and 2 as an initial number of A_1 molecules has the following master equation:

$$\begin{aligned} \frac{dP_2(t)}{dt} &= -\lambda P_2(t) \\ \frac{dP_1(t)}{dt} &= \lambda P_2(t) - \lambda P_1(t) \\ \frac{dP_0(t)}{dt} &= \lambda P_1(t) \end{aligned} \tag{2.107}$$

Equation (2.107) clearly describes a Markov process. When this system of differential equation is solved with the deterministic initial condition $P_2(0) = 1, P_1(0) = P_0(0) = 0$, it is easy to notice that $P_2(t) = e^{-\lambda t}$ and $P_0(t) = 1 - P_2(t) - P_1(t)$. Substituting this information into Eq. (2.107) leaves to following single-variable ordinary differential equation for $P_1(t)$:

$$\frac{dP_1(t)}{dt} = \lambda e^{-\lambda t} - \lambda P_1(t) \tag{2.108}$$

As the variable time explicitly appears in Eq. (2.108), it could be argued that the approach suddenly became non-Markovian, which is in conflict with the original Markovian nature of the process. However, all that happened is that the zero point in time was defined when the solution was obtained for variable $P_2(t)$.

As pointed out in Sect. 2.2.5, the property of memorylessness is a characteristic feature of Markov chains. But this property is also very generally true for many physical or chemical processes: all the information necessary to predict future changes in a system (deterministically or using probabilities) is thought to be contained in the present state. Because of these somewhat philosophical problems, it is probably better to limit the use of the term 'non-Markovian' to the particular mathematical equations used in describing a process rather than the physical phenomenon itself.

In a work describing non-Markovian polymer reaction kinetics [40], it was pointed out that the typical reactive conformation of the polymer is more extended than the equilibrium conformation, which leads to reaction times significantly shorter than predicted by Markovian theories. Here, introducing the non-Markovian mathematics is caused by the fact the different conformers of the same polymer are thought to be the same species. Clearly, the conformers are different with respect to reactivity. Therefore, regarding different conformers different species would certainly give rise to a Markovian description. It should also be pointed out that a very similar problem has long been known in the field of photochemistry, where electronically excited states have very different reactivities than ground states, yet they only differ in the configuration of electrons. The standard solution there is to consider the excited state a separate species with its own properties (number of molecules or concentration).

Another application of non-Markovian approach (actually to gene expression) will be mentioned in Sect. 3.7.

References

1. Anderson DF (2008) Incorporating postleap checks in tau-leaping. J Chem Phys 128:054103
2. Anderson DF, Craciun G, Kurtz TG (2010) Product-form stationary distributions for deficiency zero chemical reaction networks. Bull Math Biol 72(8):1947–1970
3. Auger A, Chatelain P, Koumoutsakos P (2006) R-leaping: accelerating the stochastic simulation algorithm by reaction leaps. J Chem Phys 125:084103
4. Barabás B, Tóth J, Pályi G (2010) Stochastic aspects of asymmetric autocatalysis and absolute asymmetric synthesis. J Math Chem 48(457–489)
5. Baras F, Frankowicz M (1984) Transient bimodality: a new mechanism of stochasticity in explosive chemical reactions. In: Vidal C, Pacault A (eds) Non-equilibrium dynamics in chemical systems. Springer series in synergetics, vol 27. Springer, Berlin, p. 250-
6. Baras F, Malek-Mansour M, Pearson JE (1996) Microscopic simulation of chemical bistability in homogeneous systems. J Chem Phys 105:8257–8261
7. Bartholomay AF (1958) Stochastic models for chemical reactions: I. Theory of the unimolecular reaction process. Math Biophys 20:175–190

8. Bartholomay AF (1959) Stochastic models for chemical reactions: II. The unimolecular rate constant. Math Biophys 21:363–373
9. Blomberg C (2006) Fluctuations for good and bad: the role of noise in living systems. Phys Life Rev 3:133–161
10. Borgis D, Moreau M (1984) On the stochastic theory of a bistable chemical reaction. Physica A 123:109–123
11. Bratsun DA, Volfson DN, Hastya J, Tsimring LS (2005) Non-Markovian processes in gene regulation. In: Kish LB, Lindenberg K, Gingl Z (eds) Noise in complex systems and stochastic dynamics III. Proceedings of the SPIE, vol 5845. SPIE, Bellingham, pp. 210–219
12. Cao Y, Gillespie DT, Petzold LR (2006) Efficient step size selection for the tau-leaping simulation method. J Chem Phys 124:044109
13. Chaturvedi S, Gardiner CW (1978) The poisson representation. II. Two-time correlation functions. J Stat Phys 18:501–522
14. Cobb L (1978) Stochastic catastrophe models and multimodal distributions. Behav Sci 23:360–374
15. Dahmen W, Micchelli CA (1988) The number of solutions to linear Diphantine equations and multivariate splines. Trans Am Math Soc 308:509–532
16. Dharmadhikari SW, Jogdeo K (1976) Multivariate unimodality. Ann Stat 4:607–613
17. Dóka É, Lente G (2011) Mechanism-based chemical understanding of chiral symmetry breaking in the Soai reaction. A combined probabilistic and deterministic description of chemical reactions. J Am Chem Soc 133:17878–17881
18. Dóka É, Lente G (2012) Stochastic mapping of the Michaelis-Menten mechanism. J Chem Phys 136:054111
19. Doob JL (1953) Stochastic processes. Wiley, New York
20. Ebeling W (1978) Nonequilibrium transitions and stationary probability distributions of stochastic processes. Phys Lett A 68:430–432
21. Ebeling W, Schimansky-Geier L (1979) Stochastic dynamics of a bistable reaction system. Physica A 98:587–600
22. Érdi P, Sipos T (1972) Simulation of complex chemical reactions by Monte Carlo method. In: Proceedings of the conference of hungarian chemists. Hungarian Chemical Society, Budapest
23. Érdi P, Tóth J (1979) Some comments on Prigogine's theories. React Kinet Catal Lett 11:371–375
24. Érdi P, Tóth J (1989) Mathematical models of chemical reactions. Theory and applications of deterministic and stochastic models. Manchester University Press/Princeton University Press, Manchester/Princeton
25. Érdi P, Sipos T, Tóth J (1973) Stochastic simulation of complex chemical reaction by digital computer. Magy Kem Folyirat 79:97–108. (in Hungarian)
26. Érdi P, Tóth J, Hárs V (1981) Some kinds of exotic phenomena in chemical systems. In: Farkas M (ed) Qualitative theory of differential equations. Colloquia mathematica societatis János Bolyai, Szeged, vol 30. North-Holland/János Bolyai Mathematical Society, Budapest, pp. 205–229
27. Feller W (1940) On the integro-differential equations of purely discontinuous Markoff process. Trans Am Math Soc 48:488–515
28. Frankowicz M, Kawczynski AL (1997) Modeling of bistable chemical system by cellular automata technique. Pol J Chem 71:467–473
29. Frankowicz M, Nicolis G (1983) Transient evoltuion towards a unique stable steady state: stochastic analysis of explosive behaviour in a chemical system. J Stat Phys 33:595–609
30. Frankowicz M, Malek-Mansour M, Nicolis G (1984) Stochastic analysis of explosive behaviour: a qualitative approach. Physica 125:237–245
31. Gardiner CW (2009) Chapter 12: The poisson representation. In: Stochastic methods: a handbook for the natural and social sciences, 4th edn. Springer, Berlin, pp. 301–335
32. Gardiner CW, Chaturvedi S (1977) The poisson representation. I. A new technique for chemical master equations. J Stat Phys 17:429–468

33. Gibson MA, Bruck J (2000) Efficient exact stochastic simulation of chemical systems with many species and many channels. J Phys Chem A 104:1876–1889
34. Gillespie DT (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J Comput Phys 22:403–434
35. Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81:2340–2361
36. Gillespie DT (2001) Approximate accelerated stochastic simulation of chemically reacting systems. J Chem Phys 115:1716–1733
37. Gillespie DT (2007) Stochastic simulation of chemical kinetics. Annu Rev Phys Chem 58:35–55
38. Gillespie DT (2009) Deterministic limit of stochastic chemical kinetics. J Phys Chem B 113:1640–1644
39. Gillespie DT, Hellander A, Petzold LR (2013) Perspective: stochastic algorithms for chemical kinetics. J Chem Phys 138:170901
40. Guérin T, Bénichou O, Voituriez R (2012) Non-Markovian polymer reaction kinetics. Nat Chem 4:568–573
41. Hanusse P (1973) Simulation of chemical systems using Monte Carlo method. C R Acad Sci Paris 277C:93 (in French)
42. Hárs V (1976) On some questions of stochastic reaction kinetics. Thesis, Eötvös University, Budapest (in Hungarian)
43. Horváth Á (1985) Multimodality of discrete distributions with applications. MSc thesis, ELTE TTK, Budapest (in Hungarian)
44. Iosifescu M, Tautu P (1973) Stochastic processes and applications in biology and medicine, vol 1: theory, vol 2: models. Springer/Editura Academiei, Berlin/Heidelberg/ New York/Bucure sti
45. Karlin S, McGregor J (1957) The classification of birth and death processes. Trans Am Math Soc 86:366–400
46. Kingman JFC (1969) Markov population processes. J Appl Probab 6:1–18
47. Kramers HA (1940) Brownian motion in a filed of force and the diffusion model of chemical reactions. Physica 7:284–304
48. Kurtz TG (1972) The relationship between stochastic and deterministic models for chemical reactions. J Chem Phys 57:2976–2978
49. Lente G (2004) Homogeneous chiral autocatalysis: a simple, purely stochastic kinetic model. J Phys Chem A 108:9475–9478
50. Lente G (2005) Stochastic kinetic models of chiral autocatalysis: a general tool for the quantitative interpretation of total asymmetric synthesis. J Phys Chem A 109:11058–11063
51. Lente G (2012) Stochastic mapping of first order reaction networks: a systematic comparison of the stochastic and deterministic kinetic approaches. J Chem Phys 137:164101
52. Lindblad P, Degn H (1967) A compiler for digital computation in chemical kinetics and its application to oscillatory reaction schemes. Acta Chem Scand 21:791–800
53. Malek-Mansour M, Nicolis G (1975) A master equation description of local fluctuations. J Stat Phys 13:197–217
54. Medgyessy P (1961) Decomposition of superpositions of distributions functions. Publishing House of the Hungarian Academy of Sciences, Budapest
55. Medgyessy P (1977) Decomposition of superpositions of density functions and discrete distributions. Series disquisitiones mathematicae hungaricae. Publishing House of the Hungarian Academy of Sciences, Budapest
56. Moebs WD, Haglund EA (1976) A simple Monte Carlo method for teaching chemical kinetics. J Chem Educ 53:506–507
57. Moore PG (1952) The estimation of the Poisson parameter from a truncated distribution. Biometrika 39:247–251
58. Moyal JE (1949) Stochastic processes and statistical physics. J R Stat Soc B 11:150–210
59. Prigogine I (1978) Time, structure and fluctuations. Science 201:777–785. doi:10.1126/ science.201.4358.777

60. Qian H (2011) Nonlinear stochastic dynamics of mesoscopic homogeneous biochemical reaction systems – an analytical theory. Nonlinearity 24:R19–R49
61. Rahinovitch BJ (1969) The Monte Carlo method. Plotting the curve of complex reactions. J Chem Educ 46:262–268
62. Sanft KZF, Wu S, Roh M, Fu J, Lim RK, Petzold LR (2011) StochKit2: software for discrete stochastic simulation of biochemical systems with events. Bioinformatics 27:2457–2458
63. Schaad LJ (1963) The Monte Carlo integration of rate equations. J Am Chem Soc 85:3588–3592
64. Sipos T, Tóth J, Érdi P (1974) Stochastic simulation of complex chemical reactions by digital computer. I. The model. React Kinet Catal Lett 1:113–117
65. Sipos T, Tóth J, Érdi P (1974) Stochastic simulation of complex chemical reactions by digital computer. II. Applications. React Kinet Catal Lett 1:209–213
66. Stratonovich RL (1963) Topics in the theory of random noise: general theory of random processes, vol 1. Gordon and Breach, New York/London
67. Subasi E, Subasi M, Prékopa A (2004) On strong unimodality of multivariate discrete distributions. Rutcor Research Report RRR 47-2004
68. Tóth J (1981) Poissonian stationary distributions in a class of detailed balance reactions. React Kinet Catal Lett 18:169–173
69. Tóth J, Török TL (1980) Poissonian stationary distributions: a degenerate case of stochastic kinetics. React Kinet Catal Lett 13:167–171
70. Tóth J, Török TL, Érdi P (1983) Significance of the Poisson distribution in the stochastic model of complex chemical reactions. Alk MAt Lapok 9:175–195 (in Hungarian)
71. Uhlenbeck LS, Ornstein GE (1930) On the theory of the Brownian motion. Phys Rev 36:823–841
72. van Kampen NG (1976) The equilibrium distribution of a chemical mixture. Phys Lett 59A:333–334
73. Whittle P (1968) Equilibrium distribution for an open migration process. J Appl Probab 5:569–571
74. Zheng Q, Ross J (1991) Comparison of deterministic and stochastic kinetics for nonlinear systems. J Chem Phys 94:3644–3648
75. Zurek WH, Schieve WC (1980) Deviations from the Poisson behaviour of equilibrium fluctuatuions in a closed system. J Stat Phys 22:289–296

Chapter 3
Applications

3.1 Introductory Remarks

We certainly cannot give an extensive review of the possible applications of stochastic chemical kinetics. Obviously, the investigation of the behavior of "small systems", i.e. the case when the number of particles is "small", and fluctuation phenomena is very important. More precisely, they cannot be interpreted as the superposition of certain deterministic behaviors, as they are the very phenomena under observations.

Fluctuation phenomena are particularly significant in connection with *critical phenomena* related to instabilities, and we start the analysis of applications here, related to simple and bistable reactions, respectively. A number of specific systems will be studied in this chapter. Compartmental systems are from a formal point of view analogues of isomerization reactions, and often used in the classical "biomathematics" literature. Autocatalysis implements positive feedback in the chemical system, and they are important ingredients e.g. of chemical oscillators. Enzyme-catalyzed reactions are extensively studied now by stochastic kinetic methods due to the availability of data on single enzyme activity. In the next section, chemical reaction networks as signal generating devices are studied, among others, signal processing in olfactory systems is reviewed. Then the kinetic basis of biological chirality is reviewed. In the last 15 years, there was a huge development in studying the mechanisms of genetic regulatory networks as a consequence of the combination of new experimental technique and stochastic kinetic analysis. Noise is also a source of information, and the fluctuation-dissipation theorem of chemical kinetics offers a method to estimate rate constants from equilibrium fluctuations. The field of stochastic resonance became popular, since noise plays a beneficial role in signal amplification. Finally, we review the remarkable applicability of stochastic chemical kinetics in the theory of computation.

P. Érdi and G. Lente, *Stochastic Chemical Kinetics: Theory and (Mostly) Systems Biological Applications*, Springer Series in Synergetics, DOI 10.1007/978-1-4939-0387-0_3,

3.2 Fluctuations Near Instabilities

3.2.1 Stochastic Chemical Reaction: A Simple Example

Here is a simple example to show the importance of fluctuations. Let us consider the reaction ([56], Sect. 5.6.1):

$$A + X \xrightarrow{\lambda'} 2X \tag{3.1}$$

$$X \xrightarrow{\mu} 0. \tag{3.2}$$

where A is the external and X is the internal component, and 0 denotes a so-called zero complex. This reaction can be associated with a simple birth-and-death process. The deterministic model is the following:

$$dx(t)/dt = (\lambda - \mu)x(t); \quad x(0) = x_0, \tag{3.3}$$

(here $\lambda = \lambda'[A]$.) The solution is

$$x(t) = x_0 exp((\lambda - \mu)t). \tag{3.4}$$

If $\lambda > \mu$, i.e. the birth rate constant is greater than the death rate constant, x is exponentially increasing function of time. For the case of $\lambda = \mu$

$$x(t) = x_0. \tag{3.5}$$

The stochastic model of the reaction is

$$dP_k(t)/dt = -k(\lambda + \mu)P_k(t) + \lambda(k-1)P_{k-1}(t) + \mu(k+1)P_{k+1}(t) \tag{3.6}$$

$$P_k(0) = \delta_{kx_0}; k = 1, 2, \ldots N. \tag{3.7}$$

There are two consequences of the model:

1. The expectation coincides with the process coming from the deterministic theory, i.e.

$$E[\xi(t)] = x_0 exp(\lambda - \mu)t, \tag{3.8}$$

which in the case of $\lambda = \mu$ reduces to the form

$$E[\xi(t)] = x_0.$$

2. The variance of the process is

$$D^2[\xi(t)] = (\lambda + \mu)t. \tag{3.9}$$

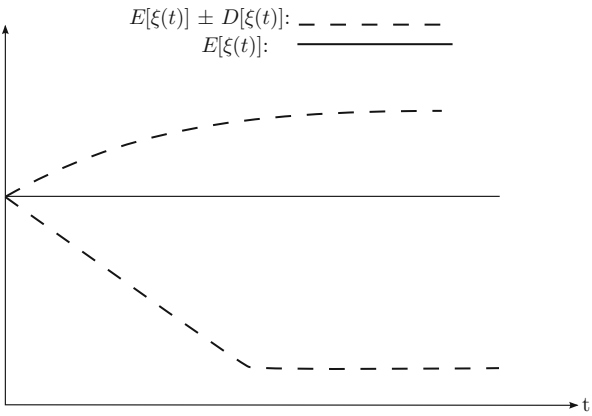


Fig. 3.1 Amplifications of fluctuations might imply instability. While the expectation is constant, the variance increases in time

For the case of $\lambda = \mu$

$$D^2[\xi(t)] = 2D\lambda t,$$

i.e. progressing with time, larger and larger fluctuations around the expectation occur (Fig. 3.1).

It is quite obvious that in this situation, it is very important to take the fluctuations into consideration. Such kinds of formal reactions are used to describe the chain reactions in nuclear reactors. In this context, it is clear that the fluctuations have to be limited since their increase could imply undesirable instability phenomena.

Fluctuations near instabilities, as an analogy to phase transitions have been studied by [137]. Near instability points (a) the amplitude of fluctuations grows; (b) the lifetime of these fluctuations becomes longer; and (c) the spatial correlation length increases.

3.2.1.1 Keizer's Paradox

Keizer [88] studied an autocatalytic system, where the deterministic and stochastic models seemingly lead to paradoxical results. Let's have the reaction

$$\mathrm{A} + \mathrm{X} \underset{\mathrm{k}_{-1}}{\overset{\mathrm{k}_1}{\rightleftharpoons}} 2\mathrm{X} \tag{3.10}$$

$$X \xrightarrow{k_2} 0. \tag{3.11}$$

The deterministic model is the following:

$$dx(t)/dt = k_1 a x(t) - k_{-1} x^2(t) - k_2 x(t); \tag{3.12}$$

There are two equilibrium points, $x_1^* = 0$ and $x_2^* = \frac{k_1 a - k_2}{k_1}$, the first is unstable and the second is stable.

In the stochastic model, the birth and death rates are defined, as

$$\mu_n = k_1 a n \text{ and } \lambda_n = \frac{k_1 n(n-1) + k - 1n}{V} + k_2 n$$

Since $\mu_0 = 0$, $n = 0$ is an absorbing state, and its stationary distribution has probability 1 for $n = 0$, i.e. extinction. It looks a paradox that $x = 0$ is an unstable fixed point of (3.12). As it was shown [88, 194] the stochastic model predicts the deterministic behavior on a reasonable time scale, which can be consistently obtained from both models. For $t_{nd} \ll t \ll t_{ce}$, the system has a quasistationary distribution centered around the non-trivial fixed point. One can obtain this distribution as the eigenfunction associated with the largest non-zero eigenvalue derived from the master equation (2.24). The time scale reaching the extinction state is proportional to $exp(-cV)$, where $c > 0$, see also [149].

3.2.2 Stochastic Theory of Bistable Reactions

3.2.2.1 Schlögl Reaction of the First-Order Phase Transition

The Schlögl reaction of the first-order phase transition has been the workhorse for studying the relationship between multistatonarity (seen in the deterministic models) and multimodality (of the stationary distribution in stochastic models). The results of [50, 57] suggest that multistationarity and multimodality correspond to each other asymptotically with increasing volume.

$$\mathrm{A} + 2\mathrm{X} \underset{\mathrm{k_2}}{\overset{\mathrm{k_1}}{\rightleftharpoons}} 3\mathrm{X} \tag{3.13}$$

$$\mathrm{B} \underset{\mathrm{k_4}}{\overset{\mathrm{k_3}}{\rightleftharpoons}} \mathrm{X},. \tag{3.14}$$

A and B are external components, X is the only internal component. The reaction model was re-investigated by [195] within the framework of non-equilibrium thermodynamics.

The deterministic model is

$$dx(t)/dt = k_1 a x^2 - k_2 x^3 - k_4 x + k_3 b; \quad x(0) = x_0. \tag{3.15}$$

The two "phases" of the system is identified by having one or three stationary points. In the latter, there are two stable stationary points separated by an unstable one, so the system is called **bistable**. The master equation is given as

$$\frac{dP_n(t)}{dt} = \lambda_{n-1} P_{n-1} + \mu_{n+1} P_{n+1} - (\lambda_n + \mu_n) P_n, \tag{3.16}$$

for $n = 1 \ldots \infty$, and

$$\frac{dP_0}{dt} = \mu_1 P_1 - \lambda_0 P_0. \tag{3.17}$$

Here $\lambda_n = \hat{k}_3 n_B + \hat{k}_1 n_A n(n-1)$ and $\mu_n = \hat{k}_4 n + \hat{k}_2 n(n-1)(n-2)$, λ_n and μ_n are the birth and death rates, respectively.

Of course, we can express the birth and death rates with the deterministic rate constants, and in that case, the volume V is explicitly taken into account by taking into account $\hat{k}_i = \frac{k_i}{V^{m-1}}$.

$$\lambda_n = \frac{a k_1 n(n-1)}{V N_A} + b k_3 V N_A, \tag{3.18}$$

and

$$\mu_n = n k_4 + \frac{k_2 n(n-1)(n-2)}{(N_A V)^2} \tag{3.19}$$

The stationary distribution is calculated by using the detailed balance assumption $\lambda_{n-1} P_{n-1}^{ss} = \mu_n P_n^{ss}$[1] as

$$P_n^{ss} = P_0^{ss} \prod_{i=0}^{n-1} \frac{\lambda_i}{\mu_{i+1}}, \; P_0^{ss} = 1 - \sum_{j=1}^{\infty} P_j^{ss}. \tag{3.20}$$

There is a remarkable difference in the behavior of the deterministic and stochastic models. The set of the initial conditions can be classified into two classes, since starting from any element of one class, the system tends to one of the two possible point attractors. As the stochastic description is concerned, two remarks should be made. First, the modality of the distribution is volume-dependent, and there might be a change in the qualitative behavior of the system, as the bifurcation diagram shows (volume is the bifurcation parameter in Fig. 3.2). Second, the stochastic model also describes that the system quickly tends towards one of the stable points, but it allows with certain probability to drive the system. This phenomenon is related to the concept of **relative stability** [93, 94, 148].

3.2.2.2 Time Spent in Each Steady State, and Time Scale of Transitions

Diffusion in Bistable Potential

Fokker–Planck equations in a double-well potential are widely used to describe phase transitions, and exact solutions for the diffusion in bistable potentials in

[1]The detailed balance condition is stronger than that required merely for a stationary distribution; that is, there are Markov processes with stationary distributions that do not have detailed balance e.g. [42].

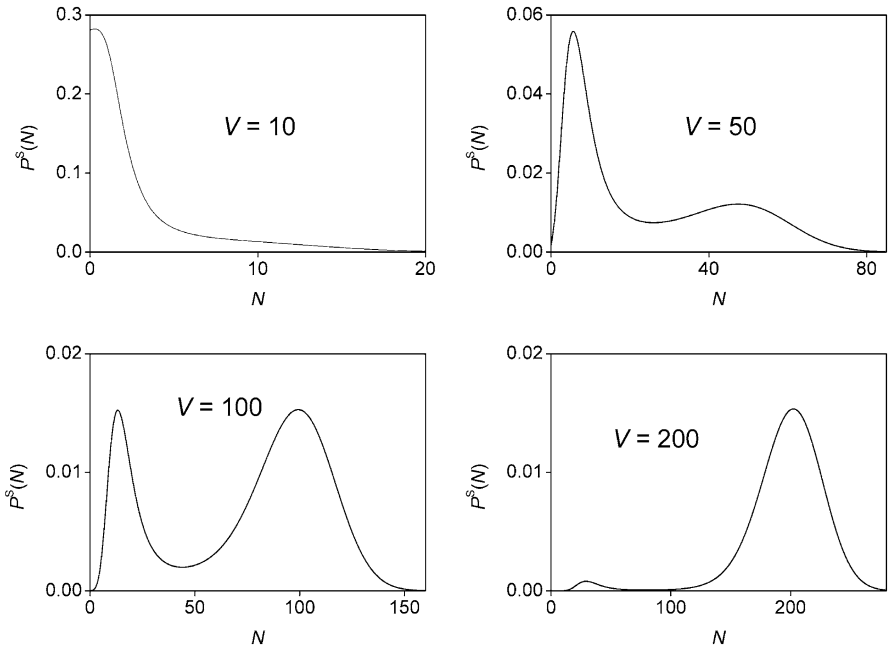


Fig. 3.2 The volume-dependence of the modality (Adapted from [50])

the case of certain potential functions were given [70, 75]. The estimation of the relaxation times from metastable to stable states was also given [117]. The Fokker–Planck equation can be written as

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{x}[A(x)P(x,t)] + (1/2V)\frac{\partial^2}{\partial x^2}[B(x)P(x,t)]. \tag{3.21}$$

The stationary distribution is

$$P^{st}(x) = (K/B(x))exp[-VU(x)], \tag{3.22}$$

where K is a normalization constant and

$$U(x) = -2\int_0^x [A(x')/B(x')]dx'. \tag{3.23}$$

U can be considered analogous to "free energy". The stationary states of the system occur at the extrema of $U(x)$. (Meta)stable states can be identified with local minima. The relaxation time of the process leading from a metastable to a stable state is considered as

$$\tau_m = 2V \int_{x_m}^{x_u} exp[V(U(x)]dx \int_{-\infty}^{x} 1/B(y)\{exp[-VU(y)]\}dy, \tag{3.24}$$

where x_u and x_m are neighboring unstable and metastable states. It can be seen that the relaxation time exponentially depends on the volume. By evaluating the integrals, the relaxation time is estimated as

$$\tau_m \sim exp[V(U(x_u) - U(x_m))]. \tag{3.25}$$

The formula shows not only the exponential dependences of the relaxation time on the system size, but also the effect of the height of the potential barrier $(U(x_u) - U(x_m))$.

Returning now to the master equation (3.16), it can be re-stated in the matrix formalism of Eq. (2.39). The composition of matrix Ω is as follows:

$$\Omega = \begin{pmatrix} -\lambda_0 & \mu_1 & 0 & \cdots \\ \lambda_0 & -\lambda_1 - \mu_1 & \mu_2 & \cdots \\ 0 & \lambda_1 & -\lambda_2 - \mu_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{3.26}$$

Infinite matrix (3.26) is singular (all the column sums are 0) and negative semidefinite. All the eigenvalues are non-positive real numbers, the highest one is zero. The largest non-zero eigenvalue differs from the others: it decays exponentially as the volume size increases and dictates the slow time scale of the system. The other eigenvalues are relatively stationary with the volume change, and are much larger in magnitude.

Considerable numerical efforts have been made [194] to estimate the mentioned largest non-zero eigenvalue. The method used was to truncate matrix (3.26) at a reasonable value to be finite, and the eigenvalues of the resulting (non-singular!) tridiagonal matrix were calculated numerically.

Stochastic bistable systems can be characterized by three different **time scales** as it was mentioned in Sect. 2.4.9, (see e.g. [149]): (i) the time scale t_{ms} of the individual reactions (under the somewhat strong assumption that the k_i constants have a common scale); (ii) the time scale t_{nd} of the whole reaction system; and (iii) the time scale t_{ce} (can be called, as evolutionary time scale) of the transition between the (meta)stable states. The "long-term" behavior of the deterministic models happens when $t \gg t_{nd}$ and $t \ll t_{ce}$, and the system "selects" an attractor depending on the initial value. (Such kinds of systems can be interpreted as classifiers of the set of the potential initial values.) On the time scale $t \gg t_{ce}$ the system is better characterized by a bimodal distribution.

As it was emphasized by Qian [149], nonlinear stochastic biochemical dynamics seems to be a new mathematical framework to combine deterministic and statistical aspects of modeling, which is capable of describing the importance of *rare events*. Such kinds of events occur with infinitesimal probability on a regular time scale associated with deterministic phenomena, but occur with probability 1 on an evolutionary time scale.

3.2.2.3 The *lac* Operon Genetic Network

Jacob and Monod outlined a network theory of genetic control in prokaryotes (prokaryotes are simple cells, which do not contain nucleus, while eukaryotes do) in 1961 [81,127]. The Operon model is the classical model for the cellular metabolism, growth and differentiation (for the legacy and historical analysis of this seminal work see [129]). Detailed deterministic models [201] (which by taking into account the network structure of the lactose operon regulatory system) are able to reflect the fundamental bistable property of the system, admittedly for large system only. However, a single cell can switch between different phenotypes by some stochastic mechanism (e.g. [4]).[2] Nowick and Weiner [138] famously recognized the "all-or-none", bimodal nature of *lac* operon, and the existence of switching mechanism between different (in their case *"induced* and *"non-induced"*) phenotypes, and hypothesized that the random expression of a single molecule of permease should be enough to trigger induction. Choi et al. [38] identified a mechanism of switching between different phenotypes triggered by a single-molecule event[3] Measurements of genetic noise will briefly be reviewed in Sect. 3.7.2.

3.3 Compartmental Systems

The stochastic kinetic description of compartmental systems has been the subject of a large number of research articles, which often appear to be independent of each other [11,12,31,43,65,66,77,78,82,90,102,108,121,123,128,171,172,177–180]. In most chemical considerations, such systems of reactions are called first order processes, which may be both open and closed. Compartmental processes are characterized by the fact that the possible changes of all single molecular (or other) entities are independent of the presence of others. In other words, no interaction between any two such entities occurs.

3.3.1 Model Frameworks

A general network of first order reactions in a closed system involves n different chemical species ($A_1, A_2, \ldots, A_n$), every one of which can convert to any other, i.e. chemical reactions are possible for all pairs of species present:

[2] As they write "...Conventional deterministic kinetics cannot be used to predict statistics of regulatory systems that produce probabilistic outcomes. Rather, a stochastic kinetic analysis must be used to predict statistics of regulatory outcomes for such stochastically regulated systems."

[3] It is known that bistable systems may show *history-dependence*, i.e. *hysteresis* and they are related to memory phenomena. For a very good reading about chance, bistability and memory, see [143].

$$\mathrm{A}_i \xrightarrow{k_{i,j}} \mathrm{A}_j \tag{3.27}$$

This sort of reaction is sometimes termed 'conversion' to distinguish it from other types of possible processes in compartmental systems.

From a chemist's point of view, a reactor can be made open with inflow and outflow by including processes that produce or consume A_i molecules without consuming or producing any other molecules. In the reactions representing the inflow, molecules are produced. These processes are sometimes termed 'production from a source'.

$$\emptyset \xrightarrow{\alpha_i} \mathrm{A}_i \tag{3.28}$$

The outflow process, sometimes termed 'degradation', it is the exact opposite of the inflow process:

$$\mathrm{A}_i \xrightarrow{\omega_i} \emptyset \tag{3.29}$$

Finally, there is a family of steps, which is termed 'catalytic production from a source', in which species A_j is produced in with a rate that is proportional to number of A_i molecules:

$$\emptyset \xrightarrow{\beta_{i,j},\mathrm{A}_i} \mathrm{A}_j \tag{3.30}$$

The notable case of $i = j$ here could be considered self-reproduction or autocatalytic formation.

It should be noted that a 'catalytic degradation', which would be an equivalent of catalytic production is not possible because a molecule cannot be degraded in a manner that is independent of its own presence.

Although the outflow process is often included in schemes as written in (3.29), this is not a conceptual necessity. In fact, in any system involving outflow defining a $(n + 1)$th 'reservoir' species, and including it as a common product of all outflow processes, will lead to a mathematically equivalent scheme. In the absence of production steps, such an equivalent system will be closed. It is also notable that open chemical reactors are sometimes conveniently formulated without the inflow or outflow processes described here in order to keep the number of molecules in the reactor finite.

Another question that arises is about stoichiometry. All equations here are written with a set 1:1 stoichiometry for each process (meaning that for each molecule of reactant produced, there is one molecule of product formed, or the inflow or outflow occurs individually for molecules) in the compartmental network. This is not necessarily the case in all examples and including stoichiometric coefficients may be needed. For 1:1 stoichiometry in a closed system without inflow and outflow, conservation of matter ensures that the sum of particle numbers is always the same.

With different stoichiometries but still in closed systems, a similar conservation relationship can be defined by using a suitable linear combination of particle numbers.

Individual states in a compartmental system can be identified by giving the numbers of different molecules. In a given state, let a_i mean the number of entities for species A_i.

For truly open systems, no mass conservation equation holds and the number of possible states is infinite. Calculations in such a system need to be analytical (e.g. using the Poisson representation) or the numerical calculations should be limited to a suitably chosen finite subgroup of states.

3.3.2 Master Equation and State Space

The master equation of the general compartmental system is given as follows:

$$\begin{aligned}
&\frac{dP_{a_1,a_2,\ldots,a_n}(t)}{dt} = \\
&-\left(\sum_{i=1}^{n}(\alpha_i + \omega_i a_i) + \sum_{i=1}^{n}\sum_{j=1}^{n}(a_i k_{i,j} + a_i \beta_{i,j})\right) P_{a_1,a_2,\ldots,a_n}(t) \\
&+\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i + 1)k_{i,j}\, P_{a_1,a_2,\ldots,a_i+1,\ldots,a_j-1,\ldots,a_n}(t) \\
&+\sum_{i=1}^{n}\sum_{j=1}^{n} a_i \beta_{i,j}\, P_{a_1,a_2,\ldots,a_i,\ldots,a_j-1,\ldots,a_n}(t) \\
&+\sum_{i=1}^{n} \alpha_i\, P_{a_1,a_2,\ldots,a_i-1,\ldots,a_n}(t) \\
&+\sum_{i=1}^{n} \omega_i\, P_{a_1,a_2,\ldots,a_i+1,\ldots,a_n}(t)
\end{aligned} \tag{3.31}$$

For cases when all α_i, $\beta_{i,j}$ and ω_i, values are zero, the conservation of mass ensures that the sum of all a_i values will be equal to the overall number of particles, N_0:

$$N_0 = \sum_{i=1}^{n} a_i \tag{3.32}$$

The number of states can be given by a combinatorial line of thought: it is identical to the number of different non-negative integer solutions of Diophantine

equation (3.32). The mathematics of this class of problems was developed in detail [41] and the number of states can be given by a binomial coefficient:

$$M = \begin{pmatrix} N_0 + n - 1 \\ n - 1 \end{pmatrix} = \frac{(N_0 + n - 1)!}{(n-1)!N_0!} \tag{3.33}$$

A suitable enumerating function for this special but very common case is given as follows:

$$f(a_1, a_2, \ldots, a_n) = 1 + \sum_{i=2}^{n} \sum_{j=1}^{a_i} \begin{pmatrix} N_0 - j + i - 1 - \sum_{h=i+1}^{n} a_h \\ i - 2 \end{pmatrix} \tag{3.34}$$

For more general cases with inflow and outflow, the size of the state space depends on how many of the parameters α_i, $\beta_{i,j}$ and ω_i are zero. If all α_i and $\beta_{i,j}$ values are zero, the state space is finite. In all other cases, the number of states is infinitely large.

3.3.3 Solutions

Master equation (3.31) can be handled by using the generating function introduced in Eq. (2.70). The partial differential equation equivalent to master equation (3.31) is as follows:

$$\begin{aligned} \frac{\partial G(z_1, z_2, \ldots, z_n, t)}{\partial t} &= \sum_{i=1}^{n} (z_i - 1)\alpha_i G(z_1, z_2, \ldots, z_n, t) \\ &- \sum_{i=1}^{n} (z_i - 1)\left(\omega_i + \sum_{j=1}^{n} k_{i,j}\right) \frac{\partial G(z_1, z_2, \ldots, z_n, t)}{\partial z_i} \\ &+ \sum_{i=1}^{n} (z_i - 1) \sum_{j=1}^{n} \frac{\partial G(z_1, z_2, \ldots, z_n, t)}{\partial z_j} (k_{i,j} + \beta_{i,j} z_j) \end{aligned} \tag{3.35}$$

The time dependences for the molecule numbers of species A_i is can be obtained from the set of ordinary differential equations fully analogous to the deterministic equations describing the evolution of concentrations in the compartmental system:

$$\frac{d \langle a_i \rangle (t)}{dt} = \alpha_i + \sum_{j=1}^{n} \langle a_j \rangle (t)(k_{j,i} + \beta_{j,i}) - \langle a_i \rangle (t)\left(\omega_i + \sum_{j=1}^{n} k_{i,j}\right) \tag{3.36}$$

An essential and often noted consequence of this equation is that the expectations of molecule numbers obtained by the stochastic approach in a compartmental

system are identical to the concentrations obtained by the deterministic approach. The equation can be re-formulated into a more concise matrix format:

$$\frac{d\,\mathbf{n}(t)}{dt} = \underline{\underline{k}}\,\mathbf{n}(t) + \boldsymbol{\alpha} \tag{3.37}$$

In this equation, the expectations are arranged in the form of vector $\underline{a}(t)$:

$$\mathbf{n}(t) = \begin{pmatrix} \langle a_1 \rangle\,(t) \\ \langle a_2 \rangle\,(t) \\ \vdots \\ \langle a_n \rangle\,(t) \end{pmatrix} \tag{3.38}$$

A similar vector of $\underline{\alpha}$ values is also defined:

$$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \tag{3.39}$$

Finally, the remaining rate constants are given in the matrix $\underline{\underline{k}}$ as follows:

$$\underline{\underline{k}} = \begin{pmatrix} \beta_{1,1} - \omega_1 - \sum_{i=1}^{n} k_{1,i} & k_{2,1} + \beta_{2,1} & \cdots & k_{n,1} + \beta_{n,1} \\ k_{1,2} + \beta_{1,2} & \beta_{2,2} - \omega_2 - \sum_{i=1}^{n} k_{2,i} & \cdots & k_{n,2} + \beta_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ k_{1,n} + \beta_{1,n} & k_{2,n} + \beta_{2,n} & \cdots & \beta_{n,n} - \omega_n - \sum_{i=1}^{n} k_{n,i} \end{pmatrix} \tag{3.40}$$

It is also possible to give a similar matrix equation for the second order moments. These moments are most easily calculated as elements of the following matrix $\underline{\underline{\Sigma}}$:

$$\underline{\underline{\Sigma}}(t) = \begin{pmatrix} s_{1,1}(t) - \langle a_1 \rangle\,(t) & s_{1,2}(t) & \cdots & s_{1,n}(t) \\ s_{2,1}(t) & s_{2,2}(t) - \langle a_2 \rangle\,(t) & \cdots & s_{2,n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ s_{n,1}(t) & s_{n,2}(t) & \cdots & s_{n,n}(t) - \langle a_n \rangle\,(t) \end{pmatrix} \tag{3.41}$$

Here, the quantity $s_{i,j}(t)$ is defined as:

$$s_{i,j}(t) = \sum_{\text{all states}} a_i a_j\, P_{a_1,a_2,\ldots,a_n}(t) \tag{3.42}$$

The equation itself is given as follows:

$$\frac{d\,\underline{\underline{\Sigma}}(t)}{dt} = \underline{\underline{k}}\,\underline{\underline{\Sigma}}(t) + (\underline{\underline{k}}\,\underline{\underline{\Sigma}}(t))^T + \underline{\underline{\Gamma}}(t) + (\underline{\underline{\Gamma}}(t))^T \tag{3.43}$$

The matrix $\underline{\underline{\Gamma}}$ in this equation is defined as:

$$\underline{\underline{\Gamma}}(t) = \begin{pmatrix} (\beta_{1,1} + \alpha_1)\langle a_1 \rangle & (\beta_{1,2} + \alpha_1)\langle a_2 \rangle & \cdots & (\beta_{1,n} + \alpha_1)\langle a_n \rangle \\ (\beta_{2,1} + \alpha_2)\langle a_1 \rangle & (\beta_{2,2} + \alpha_2)\langle a_2 \rangle & \cdots & (\beta_{2,n} + \alpha_2)\langle a_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ (\beta_{n,1} + \alpha_n)\langle a_1 \rangle & (\beta_{n,2} + \alpha_n)\langle a_2 \rangle & \cdots & (\beta_{n,n} + \alpha_n)\langle a_n \rangle \end{pmatrix} \tag{3.44}$$

The moment generating function is also useful for obtaining the distribution of A_i molecules at the steady state:

$$P_{A_i}(k,t) = \frac{1}{k!} \frac{\partial^k G(z_1, z_2, \ldots, z_n, t)}{\partial z_i^k} (z_1 = 1, z_2 = 1, \ldots, z_i = 0, \ldots, z_n, t) \tag{3.45}$$

In the case of first-order reaction networks, there is a particularly useful line of thought, which can simplify some considerations. This could be termed the method of independent molecules, which can be employed because – as a consequence of first order processes only – there are no interactions between molecules and the state of each individual molecule can be described without knowing about the states of the rest of the molecules in the system. This line of thought is particularly easily used under conditions when the initial state of the system contains only one type of molecules, which is not uncommon in practice. If this holds, even $P_{a_1,a_2,\ldots,a_n}$ functions can be given in a simple way using Π_i functions:

$$P_{a_1,a_2,\ldots,a_n}(t) = \frac{N_0!}{\prod_{i=1}^n a_i!} \prod_{i=1}^n \left(\frac{\langle a_i \rangle}{N_0} \right)^{a_i} \tag{3.46}$$

Among compartmental systems, the one describing irreversible first order decay, i.e. the case when $n = 2$, $k_{1,2} > 0$, and $k_{2,1} = 0$, is the classical example invariably present in all introductory chemical kinetics textbooks. In many cases, chemical reactions can be simplified to this mathematical description by the method flooding (i.e. using all the reactants in large excess except the limiting reagent). The molecule numbers for both the reactant and product (species A_1 and A_2) are characterized by a binomial distribution. In this case, a usual experimental problem is to estimate the quantity of $k_{1,2}t$ from measured values of molecule numbers, which could serve as a way for either determining the age of a sample using a known rate constant, or finding the rate constant of a process for which the time of experiment is known. If N particles remain in an experiment out of an initial number of N_0, the expectation for $k_{1,2}t$ can be given as follows:

$$\langle k_{1,2}t \rangle = \ln \frac{N_0 + 1}{N + 1} \tag{3.47}$$

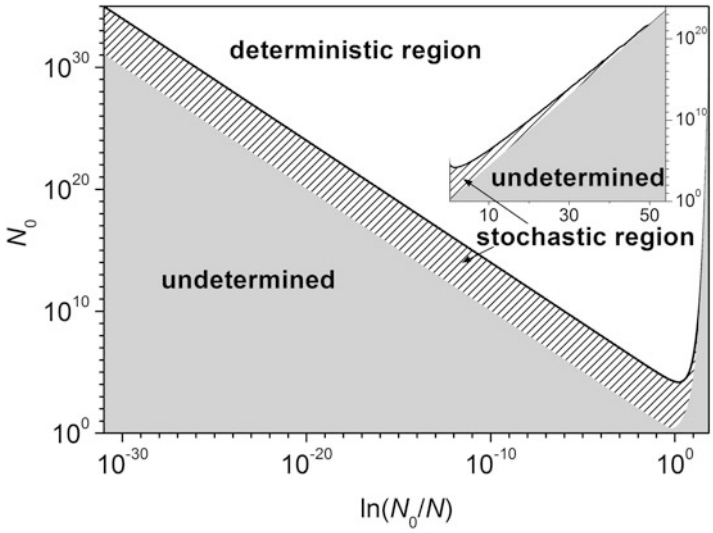


Fig. 3.3 Stochastic map of the irreversible first order reaction with $k_{1,2}t$ as the target variable (Adapted from [103])

The standard deviation can be calculated by the following equation:

$$\sigma_{k_{1,2}t} = \ln\left(1 + \sqrt{\frac{N_0 - N}{(N+1)(N_0+2)}}\right) \tag{3.48}$$

A stochastic map based on this target variable is shown in Fig. 3.3.

3.4 Autocatalysis

Autocatalysis is the phenomenon when the appearance of a product in a chemical reactions system opens a new pathway for the reaction, thereby accelerating the rate of reaction. In macroscopic observations, autocatalysis is often recognized from the presence of an induction period, which is basically the time necessary for the slow formation of the initial amount of product necessary to jump-start the process. However, an induction period is neither a necessary nor a sufficient condition for proving the autocatalytic nature of a process.

As early as 1940, Delbrück pointed out that autocatalysis may lead to macroscopically observable fluctuations under certain conditions [44]. Experimental verification of this prediction was reported in the 1980s by Nagypál and Epstein, who have shown that autocatalytic phenomena lead to observable fluctuations in the chlorite ion – thiosulfate and chlorite ion – iodide ion systems [134, 135]. The latter

reaction involves visible color change because of the formation of iodine and is sometimes referred to as the **crazy clock reaction** as opposed to the classical **clock reaction** [106] first described by Landolt [95], which also involves iodine formation without any measurable fluctuations in the sulfite ion – iodate ion system.

Conceptually, a full description of an autocatalytic system should involve at least two different types of chemical reactions. The first is an initiating process that provides a (usually slow) way of forming the first molecules of the product. The rate of this step should be independent of the presence of the product. The second is the actual autocatalytic step, whose rate is proportional to some power of the concentration (or particle number) of the product in addition to any dependence on reactant concentrations. The terms positive feedback or amplification are also sometimes used to characterize the essence of the second step. The necessity of the first type of process is not always recognized in the deterministic or stochastic considerations on autocatalysis, but can be deduced logically quite easily: in a scheme lacking this step, no reaction would occur at all if the product is not already present at the initial time.

Surprisingly, stochastic theoretical works on autocatalysis often only deal with one of these two types of chemical reactions. When the first process is absent, it is easy to resort to the argument that the product is present in some minimal quantity at the initial time for some unspecified external reason or even its deliberate addition can be assumed in order to start the process. In this case, the magnitude of this initial product amount will greatly influence any later considerations. When the autocatalytic reaction is absent, then the usual assumption is that the fluctuations generated by the initiating process will spread to later observations.

This second avenue of thought was used in Delbrück's early article, the title of which may even be considered somewhat misleading as there is no mathematical treatment given for any autocatalytic reaction. The work uses the assumption that the autocatalytic process is initiated by a simple first-order reaction and derives expectations and standard deviations for particle numbers and waiting times. These are the same as those used in the description of first-order processes (see Sect. 3.3). The arguments are then verbally extended for autocatalytic processes by noting that the macroscopically observable stochastic distributions in them should be very similar to those in the initiating reaction at very low particle numbers.

3.4.1 Autocatalytic Extinction

The lack of an initiating process was shown to give rise to a phenomenon called extinction in autocatalytic systems [49], which occurs when the molecule number of the autocatalytic species falls to zero in a system that involves a pathway for the decay of the autocatalyst. This phenomenon is unknown in deterministic kinetics, as an initially nonzero concentration can at no time be exactly zero there. A detailed study was published for a kinetic scheme that does not conserve mass. In this

work, a reversible autocatalytic step and an irreversible decay step were considered simultaneously:

$$\begin{aligned} &A_1 \rightarrow 2A_1 \quad v_1 = k_1[A_1] \\ &2A_1 \rightarrow A_1 \quad v_2 = k_1[A_1]^2 \\ &A_1 \rightarrow \emptyset \quad v_3 = k_3[A_1] \end{aligned} \tag{3.49}$$

Because of the non-mass-conserving nature of the process, there is an infinitely large number of possible states in this system. However, identifying a state is very easy by giving the number of A_1 molecules, as this is the only particle present. The master equation can be written as follows:

$$\begin{aligned} \frac{dP_i(t)}{dt} = &-\left[\kappa_1 i + \kappa_2 i(i-1) + k_3 i\right] P_i(t) + \kappa_1(i-1)P_{i-1}(t) \\ &+ \left[\kappa_2 i(i+1) + \kappa_3(i+1)\right] P_{i+1}(t) \end{aligned} \tag{3.50}$$

In the mathematical handling of this master equation, the Poisson representation was used (Sect. 2.4.6). The probability distribution over discrete chemical population configurations was expanded in terms of an overcomplete basis of analytically continued Poissonian probability distributions.

$$P_i(t) = \int d^2x f(x,t) e^{-x} \frac{x^i}{i!} \tag{3.51}$$

Here, the integral is taken over the entire complex plane. It was proven that there always exists a real, positive definite distribution $f(x,t)$ that can be used in this equation. Because of the overcompleteness of the basis, there are an arbitrarily large number of such distributions; each of which carries the same information contained as the original probability distribution P(n,t).

The master equation can be transformed into the following equation of motion for the transformed distribution:

$$\frac{\partial f(x,t)}{\partial t} = (k_3 - k_1) f(x,t). \tag{3.52}$$

This new equation of motion has exactly the same form as a Fokker–Planck equation, which describes the evolution of a probability distribution associated with a continuous-variable diffusion process. An advantage of this calculation method is that there is an equivalent representation by a set of stochastic differential equations, which can be solved to give x as a function of time.

$$dx = \left[(k_1 - k_3)x - k_2 x^2\right] dt + \sqrt{2x(k_1 - k_2 x)}\, dW(t). \tag{3.53}$$

Here, $dW(t)$ is the Wiener increment for which $\langle dW(t)\rangle = 0$ and which satisfies the autocorrelation relation $\langle dW(t)dW(t+\tau)\rangle = \delta(\tau)dt$.

The individual $P_i(t)$ functions can be obtained when the $x(t)$ function is known:

$$P_i(t) = \left\langle e^{-x(t)} \frac{x(t)^i}{i!} \right\rangle \tag{3.54}$$

Based on the previous considerations, a simple prescription can be established for calculating extinction times (T) in a numerical simulation as a function of the initial number of species (x_0):

$$T(x_0) = \int_0^\infty \langle 1 - e^{-x} \rangle \, dt \tag{3.55}$$

At this point, it should be again noted that extinction is a phenomenon that is caused by the lack of a non-autocatalytic pathway for the formation of A_1, and not by the lack of conservation of mass in the scheme. There is a published experimental example of extinction in the literature of autocatalytic enzyme reactions [7]. The non-mass-conserving nature of the scheme is responsible for another exclusively stochastic phenomenon in this system, which is termed autocatalytic runaway: there is always a finite probability that the autocatalytic population will be too large compared to reasonable physical constraints.

3.4.2 Time Dependence of the Crazy Clock Reaction

The models presented so far are too simple to give any meaningful interpretation of the experimentally observed stochastic fluctuations in autocatalytic systems. This was attempted, as an additional benefit from modeling work in chiral autocatalysis, based on a scheme that included both an initiating and an autocatalytic step [98]. The scheme was composed of a first order direct reaction and a step which was first order with respect to both the reactant and the product:

$$\begin{aligned} A_1 &\rightarrow A_2 \quad v = k_1[A_1] \\ A_1 &\rightarrow A_2 \quad v = k_2[A_1][A_2] \end{aligned} \tag{3.56}$$

A state is unambiguously identified by giving the number A_1 molecules as conservation of mass ensures $a_2 = N_0 - a_1$, where N_0 as the initial number of A_1 molecules. The master equation describing this scheme with is as follows:

$$\begin{aligned} \frac{dP_i(t)}{dt} = &-[\kappa_1 i + \kappa_2 i(N_0 - i)]\, P_i(t) \\ &+ [\kappa_1(i+1) + \kappa_2(i+1)(N_0 - i - 1)]\, P_{i+1}(t) \end{aligned} \tag{3.57}$$

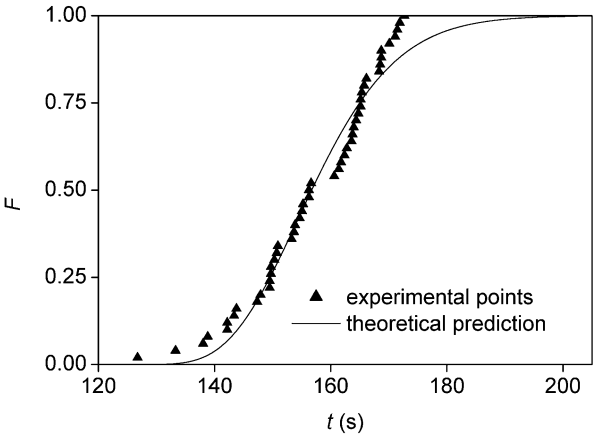


Fig. 3.4 Experimental and fitted cumulative distribution functions for the reaction time in the crazy-clock (chlorite ion-iodide ion) reaction (Adapted from [98])

It was shown that this model gives a reasonably good fit to the experimentally observed reaction times in the chlorite ion-iodide ion reaction. The goodness of the fit is illustrated by Fig. 3.4.

3.4.3 Autocatalytic Cycle Process

A more complex autocatalytic model was studied by Cianci et al. [39], the scheme of which involves the sequential autocatalytic interconverting of n species A_i and an out-of-cycle nonreactive species A_{n+1}:

$$\begin{aligned}
&A_i + A_{i+1} \to 2A_{i+1} \quad v_i = k_i[A_i][A_{i+1}] \quad i = 1, \ldots, n-1 \\
&A_n + A_1 \to 2A_1 \quad v_n = k_n[A_n][A_1] \\
&A_{n+1} \to A_i \quad v_{n+i} = k_{n+i}[A_{n+1}] \quad i = 1, \ldots, n \\
&A_i \to A_{n+1} \quad v_{2n+i} = k_{2n+i}[A_i] \quad i = 1, \ldots, n
\end{aligned} \tag{3.58}$$

Parameters k_i are the autocatalytic process rate constants, while k_{n+i} and k_{2n+i} are the rate constants at which reactive A_i molecules appear and disappear from the system. The overall number of particles in the system is denoted by N_0. The number of states is the same as for a compartmental system with $(n + 1)$ different types of molecules and a fully analogous enumerating function can be used, see Eqs. (3.33) and (3.34). This scheme does not show extinction phenomena, as all n autocatalytic species can be formed in a non-catalytic pathway. In this scheme, a state can be identified by giving the number of each A_i species as a_i, while the number of A_{i+1} species is obtained by conservation of mass as $a_{n+1} = N_0 - \sum_{i=1}^{n} a_i$. The master equation is as follows:

$$
\begin{aligned}
&\frac{dP_{a_1,a_2,\ldots,a_n}(t)}{dt} = \\
&-\left(\kappa_n a_n a_1 + \sum_{i=1}^{n-1}(\kappa_i a_i a_{i+1} + \kappa_{n+i} a_{n+1} + \kappa_{2n+i} a_i)\right) P_{a_1,a_2,\ldots,a_n}(t) \\
&+\sum_{i=1}^{n-1} \kappa_i (a_i + 1)(a_{i+1} - 1) P_{a_1,\ldots,a_i+1,a_{i+1}-1,\ldots,a_n}(t) \\
&+\kappa_n (a_n + 1)(a_1 - 1) P_{a_1+1,\ldots,a_n-1}(t) \\
&+\sum_{i=1}^{n} \kappa_{n+i} a_{n+1} P_{a_1,\ldots,a_i-1,\ldots,a_n}(t) \\
&+\sum_{i=1}^{n} \kappa_{2n+i} (a_i + 1) P_{a_1,\ldots,a_i+1,\ldots,a_n}(t)
\end{aligned}
\tag{3.59}
$$

This system of equations is solved using the van Kampen expansion method, [192] which yields the following partial differential equation:

$$
\begin{aligned}
&-\sum_{i=1}^{n} \frac{\partial \Pi}{\partial \xi_i} \frac{d[\mathrm{A}]_i}{dt} = \sum_{i=2}^{n-1} (k_i [\mathrm{A}_i][\mathrm{A}_{i+1}] - k_{i-1}[\mathrm{A}_{i-1}][\mathrm{A}_i]) \frac{\partial \Pi}{\partial \xi_i} \\
&+(k_1[\mathrm{A}_1][\mathrm{A}_2] - k_n[\mathrm{A}_n][\mathrm{A}_1]) \frac{\partial \Pi}{\partial \xi_1} + (k_n[\mathrm{A}_n][\mathrm{A}_1] - k_{n-1}[\mathrm{A}_{n-1}][\mathrm{A}_n]) \frac{\partial \Pi}{\partial \xi_n} \\
&+\sum_{i=1}^{n} (k_{2n+i}[\mathrm{A}_i] - k_i[\mathrm{A}_{n+1}]) \frac{\partial \Pi}{\partial \xi_i}
\end{aligned}
\tag{3.60}
$$

In this equation, functions $[\mathrm{A}_i](t)$ represent the deterministic solution of the same reaction scheme, for which the following ordinary differential equations hold:

$$
\begin{aligned}
\frac{d[\mathrm{A}_1]}{dt} &= k_n[\mathrm{A}_n] k_1[\mathrm{A}_i] - [\mathrm{A}_1][\mathrm{A}_2] + k_{n+1}[\mathrm{A}_{n+1}] - k_{2n+1}[\mathrm{A}_1] \\
\frac{d[\mathrm{A}_i]}{dt} &= k_{i-1}[\mathrm{A}_{i-1}][\mathrm{A}_i] - k_i[\mathrm{A}_i][\mathrm{A}_{i+1}] + k_{n+i}[\mathrm{A}_{n+1}] - k_{2n+i}[\mathrm{A}_i] \\
&\qquad i = 2, \ldots, n-1 \\
\frac{d[\mathrm{A}_n]}{dt} &= k_{n-1}[\mathrm{A}_{n-1}][\mathrm{A}_n] - k_n[\mathrm{A}_n][\mathrm{A}_1] + k_{2n}[\mathrm{A}_{n+1}] - k_{3n}[\mathrm{A}_n]
\end{aligned}
\tag{3.61}
$$

Distribution $\Pi(\xi_1, \ldots, \xi_n, t)$ is defined as:

$$
\Pi(\xi_1, \ldots \xi_k, t) = P\left(\phi_1 + \frac{\xi_1}{\sqrt{N_0}}, \ldots, \phi_n + \frac{\xi_n}{\sqrt{N_0}}, t\right) \tag{3.62}
$$

The second and third moments of ξ_i functions were characterized and an analytical solution based an approximation was compared to simulation results.

Discreteness-induced transitions were also studied in a variation of the model given in Eq. (3.58), which only contained the autocatalytic cycle with species up to A_n (i.e. k_{n+i} and k_{2n+i} rate constants were zero) and all autocatalytic process rate constants (k_i) were equal to a common value k [188]. However, the model was studied in a flow reactor, so the deterministic rate equation was as follows:

$$\begin{aligned}
\frac{d[A_1]}{dt} &= k[A_n][A_i] - k[A_1][A_2] + D([A_1]_{\text{feed}} - [A_1]) \\
\frac{d[A_i]}{dt} &= k[A_{i-1}][A_i] - k[A_i][A_{i+1}] + D([A_i]_{\text{feed}} - [A_i]) \\
& \qquad i = 2, \ldots, n-1 \\
\frac{d[A_n]}{dt} &= k[A_{n-1}][A_n] - k[A_n][A_1] + D([A_n]_{\text{feed}} - [A_n])
\end{aligned} \tag{3.63}$$

In Eq. (3.63), D is a parameter characteristic of the flow rate, and $[A_1]_{\text{feed}}$, $[A_2]_{\text{feed}}$, …, $[A_n]_{\text{feed}}$ are the concentrations of species $A_1, A_2, \ldots, A_n$ in the feed. With a further simplifying assumption $[A_1]_{\text{feed}} = [A_2]_{\text{feed}} = \ldots = [A_n]_{\text{feed}}$, the deterministic version of this model has a single stationary point, in which all concentrations are equal.

Simulations using the Gillespie algorithm showed that the stochastic behavior of this system at low molecule numbers and $n \geq 4$ is quite different from what is expected and observed at the high volume limit. Instead of each molecule number fluctuating around the common stationary value, the system underwent continuous changes between two states for $n = 4$. The first was rich in A_1 and A_3 and had 0 molecules of A_2 and A_4. In the second, A_2 and A_4 were abundant, whereas A_1 and A_3 were not present. These states resemble extinct states, but are in fact not final states as the flow can always introduce new molecules of any of the reagents. As expected, switching between these two distinct states occurred randomly and in a short time compared to the time intervals spent in the two characteristic states.

It should be noted that the key to this phenomenon is that the flow is relatively slow compared to the autocatalytic processes. Setting $D = 0$ in Eq. (3.63) reveals that even the deterministic system without in- and outflow in fact has additional stationary points: $[A_1] = [A_3] = 0$ (with $[A_2]$ and $[A_4]$ left undetermined), and $[A_2] = [A_4] = 0$ (where $[A_1]$ and $[A_3]$ are undetermined). So the stochastic simulation results actually showed the dominance of these additional stationary states, which may not be very surprising as the deterministic stationary point $[A_1]_{\text{feed}} = [A_2]_{\text{feed}} = \ldots = [A_n]_{\text{feed}}$ is an unstable one.

3.5 Enzyme Kinetics

Enzyme-catalyzed reactions, although by no means fundamentally different from non-biological forms of catalytic processes, are usually discussed separately in the kinetic literature. The main reasons for this are the very specific features of enzymes such as high selectivity, extreme efficiency and activity, as well as their typically very low concentrations in biochemical systems. A single cell can be considered as a spatially distinct reactor whose size is clearly small compared to other usual chemical systems. This small size together with the large molecular dimensions and high variety of enzymes present in living cells ensures that the amounts of at least some of the essential enzymes in a cell do not exceed a few individual molecules. As a consequence, the stochastic approach to chemical kinetics is extremely important in biochemistry. With the advance of experimental detection methods, especially fluorescence spectroscopy, studies on the activity of a single enzyme have been possible for the last decade [51, 54, 60, 62, 96, 113, 166, 184, 193, 196]. This development was also the driving force of further theoretical studies on stochastic kinetic methods specific to enzyme catalysis.[4]

3.5.1 Michaelis–Menten: Scheme and State Space

Without doubt, the single most important mechanism in enzyme kinetics is the Michaelis–Menten equation [124] and the corresponding chemical scheme developed by Briggs and Haldane [30]. The simplest and most commonly used form of this mechanism comprises two consecutive processes: the reversible reaction of the enzyme (E) and a substrate (S) to form an adduct (ES), and the formation of the product (P) with simultaneous re-generation of the reactive form of the enzyme:

$$\mathrm{E} + \mathrm{S} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{ES} \xrightarrow{k_2} \mathrm{E} + \mathrm{P} \tag{3.64}$$

The deterministic differential equations for this scheme have already been given as Example 1.8 in Chap. 1 with the notation $\mathrm{E} = \mathrm{A}_1$, $\mathrm{S} = \mathrm{A}_2$, $\mathrm{ES} = \mathrm{A}_3$, and $\mathrm{P} = \mathrm{A}_4$. This scheme has been the subject of numerous theoretical works using the stochastic approach to chemical kinetics. For the identification of a given state, the number of free enzyme molecules (e) and the number of uncomplexed substrate molecules (s) can be conveniently used. Using the conservation of mass and the initial molecule numbers e_0 and s_0, the number of enzyme-substrate adducts (es) is calculated as $es = e_0 - e$, and the number of product molecules formed is given by $p = s_0 - s - e_0 + e$. The master equation can be formulated as:

[4]The classical papers on stochastic models of enzyme kinetics were written in the first half of the 1960s [13, 14, 80].

$$\begin{aligned}\frac{dP_{e,s}(t)}{dt} = &-[\kappa_1 es + (\kappa_{-1}+\kappa_2)(e_0-e)]\,P_{e,s}(t)\\ &+\kappa_1(e+1)(s+1)P_{e+1,s+1}(t)+\\ &+\kappa_{-1}(e_0-e+1)P_{e-1,s-1}(t)\\ &+\kappa_2(e_0-e+1)P_{e-1,s}(t)\end{aligned} \tag{3.65}$$

The connection between stochastic and deterministic rate constants are easily given by considering the orders of reactions for each rate constant ($\kappa_1 = k_1 N_A/V$, $\kappa_{-1} = k_{-1}$, $\kappa_2 = k_2$). The overall number of different possible states (M) is:

$$M = (s_0 - \frac{e_0}{2} + 1) \times (e_0 + 1) \tag{3.66}$$

This formula is given for the usual case of $s_0 \geq e_0$. For $s_0 < e_0$, a fully analogous formula obtained by exchanging s_0 and e_0 can be used. A suitable enumerating function is defined as follows:

$$f(e,s) = \begin{cases} (s_0 - s - e_0 + e + 1)(e_0+1) - e & if\ e \leq s \\ (s_0 - s - e_0 + e + 1)(e_0+1) - \frac{(e-s-1)(e-s)}{2} - e & if\ e > s \end{cases} \tag{3.67}$$

A characteristic feature of the state space is that it can be divided into subsets of mutually accessible states, which have identical $s - e$ values. The subset with a lower $s - e$ value is always downstream from the subset with a higher $s - e$ value. Reactions with rate parameters κ_1 and κ_{-1} only occur within a subset, whereas the reaction with rate parameter κ_2 moves the system between subsets.

3.5.2 Michaelis–Menten: Solutions

The master equation for a single enzyme molecule ($e_0 = 1$) was solved by Arányi and Tóth [2][5] using marginal generating functions.

$$G_e(z,t) = \sum_{s=0}^{s_0-1+e} z^s P_{e,s}(t) \tag{3.68}$$

The master equation can be transformed into system of partial differential equations:

$$\begin{aligned}\frac{\partial G_e(z,t)}{\partial t} = &\kappa_1(e+1)\frac{\partial G_{e+1}(z,t)}{\partial z} - \kappa_1 ez\frac{\partial G_e(z,t)}{\partial z} - (\kappa_{-1}+\kappa_2)(1-e)G_e(z,t)\\ &+(\kappa_{-1}+\kappa_2)z(2-e)G_{e-1}(z,t)\end{aligned} \tag{3.69}$$

[5] We might be biased, but probably even objectively a somewhat overlooked very important pioneer paper.

In effect, instead of the $2s_0 + 1$ simultaneous ordinary differential equations displayed in Eq. (3.65), the problem is now transformed into two simultaneous partial differential equations.

The solution of this system of equations is:

$$\begin{aligned} G_0(z,t) &= \Gamma e^{-\kappa_{-1}(z-1)/\kappa_1} e^{-\kappa_2 t} + \overline{\overline{\Gamma}} \frac{\kappa_{-1}+\kappa_2}{\kappa_{-1}z+\kappa_2} e^{(\kappa_1+\kappa_2)t} \\ &\quad + \sum_{i=1}^{2}\sum_{n=0}^{\infty} \Gamma_i^{(n)} \left[\frac{\kappa_2 - (\kappa_2+\lambda_i^{(n)})z}{-\lambda_i^{(n)}}\right]^{q_n} e^{\lambda_i^{(n)} t} \\ G_1(z,t) &= \Gamma^{(-1)} - \overline{\Gamma} e^{-\kappa_{-1}(z-1)/\kappa_1} e^{-\kappa_2 t} - \overline{\overline{\Gamma}} e^{-(\kappa_1+\kappa_2)t} \\ &\quad - \sum_{i=1}^{2}\sum_{n=0}^{\infty} \Gamma_i^{(n)} \left[\frac{\kappa_2 - (\kappa_2+\lambda_i^{(n)})z}{-\lambda_i^{(n)}}\right]^{q_n+1} e^{\lambda_i^{(n)} t} \end{aligned} \tag{3.70}$$

where

$$q_n = \frac{\lambda_i^{(n)}\lambda_i^{(n)} + (\kappa_1+\kappa_{-1}+\kappa_2)\lambda_i^{(n)} + \kappa_1\kappa_2}{\kappa_1(\kappa_2+\lambda_i^{(n)})} \tag{3.71}$$

Values of λ can be obtained as the roots of a quadratic equation:

$$\begin{aligned} \lambda_1^{(n)} &= \frac{-\kappa_1(n+1)-\kappa_{-1}-\kappa_2}{2} \\ &\quad - \frac{\sqrt{[\kappa_1(n+1)+\kappa_{-1}+\kappa_2]^2 - 4\kappa_1\kappa_2(n+1)}}{2} \\ \lambda_1^{(n)} &= \frac{-\kappa_1(n+1)-\kappa_{-1}-\kappa_2}{2} \\ &\quad + \frac{\sqrt{[\kappa_1(n+1)+\kappa_{-1}+\kappa_2]^2 - 4\kappa_1\kappa_2(n+1)}}{2} \end{aligned} \tag{3.72}$$

Incidentally, the previous equations actually give the eigenvalues of the transition probability matrix defined by the master equation of the process. The size of this transition probability matrix is $(2s_0 + 1) \times (2s_0 + 1)$, and its structure is such that the eigenvalues can be calculated as a trivial eigenvalue of 0 and the eigenvalues of s_0 independent 2×2 matrices, each corresponding to a subsystem composed of a pair of states with identical number of product molecules. The special structure of the transition probability matrix is also observed for $e_0 > 1$. In this case, the subsystems are characterized by (s_0-e_0+1) different matrices of $(e_0+1)\times(e_0+1)$ dimensions, with additional submatrices of sizes $e_0 \times e_0$, $(e_0 - 1) \times (e_0 - 1)$, ..., 2×2 (one of each). The reason behind this special state structure is the existence of subsets of mutually accessible states.

Returning to Eq. (3.70) now, the constants Γ can be determined from the initial conditions:

$$\begin{aligned} G_0(1,t) + G_1(1,t) &= 1 \\ G_0(z,0) &\equiv 0 \\ G_1(z,0) &\equiv z^{s_0} \end{aligned} \tag{3.73}$$

The $P_{e,s}(t)$ function then can be determined from the generating function:

$$P_{e,s}(t) = \frac{1}{s!}\frac{\partial^s G_e(z,t)}{\partial z^s} \tag{3.74}$$

A further possible line of thought for single enzyme kinetics ($e_0 = 1$) in this system is to give the functions P_{1,s_0} and P_{0,s_0-1} for the most natural initial condition where only separated enzyme and substrate molecules are present, which is expressed by $P_{1,s_0}(0) = 1$ and $P_{e,s}(0) = 0$ for every other state. The solution is:

$$\begin{aligned} P_{1,s_0}(t) &= -\frac{\lambda_2^{(s_0-1)} + \kappa_1 s_0}{\lambda_1^{(s_0-1)} - \lambda_2^{(s_0-1)}} e^{\lambda_1^{(s_0-1)} t} + \frac{\lambda_1^{(s_0-1)} + \kappa_1 s_0}{\lambda_1^{(s_0-1)} - \lambda_2^{(s_0-1)}} e^{\lambda_2^{(s_0-1)} t} \\ P_{0,s_0-1}(t) &= -\frac{(\lambda_2^{(s_0-1)} + \kappa_1 s_0)(\lambda_1^{(s_0-1)} + \kappa_1 s_0)}{(\lambda_1^{(s_0-1)} - \lambda_2^{(s_0-1)})\kappa_{-1}} e^{\lambda_1^{(s_0-1)} t} \\ &\quad + \frac{(\lambda_2^{(s_0-1)} + \kappa_1 s_0)(\lambda_1^{(s_0-1)} + \kappa_1 s_0)}{(\lambda_1^{s_0-1} - \lambda_2^{(s_0-1)})\kappa_{-1}} e^{\lambda_2^{(s_0-1)} t} \end{aligned} \tag{3.75}$$

The waiting time (τ) before the formation of the first product molecule (P) assumed some special importance in this system [89]. Its expectation is calculated as:

$$\begin{aligned} \langle \tau \rangle &= \int_0^\infty t \left(-\frac{dP_{1,s_0}(t)}{dt} - \frac{dP_{0,s_0-1}(t)}{dt} \right) dt = \int_0^\infty t\kappa_2 P_{0,s_0-1}(t) dt \\ &= -\kappa_2 \frac{(\lambda_2^{(s_0-1)} + \kappa_1 s_0)(\lambda_1^{(s_0-1)} + \kappa_1 s_0)}{(\lambda_1^{(s_0-1)} - \lambda_2^{(s_0-1)})\kappa_{-1}} \int_0^\infty t e^{\lambda_1^{(s_0-1)} t} dt \\ &\quad + \kappa_2 \frac{(\lambda_2^{(s_0-1)} + \kappa_1 s_0)(\lambda_1^{(s_0-1)} + \kappa_1 s_0)}{(\lambda_1^{s_0-1} - \lambda_2^{(s_0-1)})\kappa_{-1}} \int_0^\infty t e^{\lambda_2^{(s_0-1)} t} dt \\ &= -\kappa_2 \frac{(\lambda_2^{(s_0-1)} + \kappa_1 s_0)(\lambda_1^{(s_0-1)} + \kappa_1 s_0)}{(\lambda_1^{(s_0-1)} - \lambda_2^{(s_0-1)})\kappa_{-1}} \left(\frac{1}{\lambda_1^{(s_0-1)} \lambda_1^{(s_0-1)}} - \frac{1}{\lambda_2^{(s_0-1)} \lambda_2^{(s_0-1)}} \right) \\ &= \frac{\kappa_1 s_0 + \kappa_{-1} + \kappa_2}{\kappa_1 \kappa_2 s_0} \end{aligned} \tag{3.76}$$

The nature of enzyme catalyzed reactions is such that the substrate (S) is mostly used in large excess over the catalyst enzyme and the experimental methods used to monitor the process typically have neither the time resolution nor the chemical sensitivity to follow the formation of the enzyme-substrate adduct (ES) directly. Under such conditions, which are most often described by $k_1[\mathrm{S}]_0 + k_{-1} >> k_2$, mathematical simplifications called the pre-equilibrium and steady-state approximations are often used in deterministic kinetics. In the classical approach to Michaelis–Menten kinetics, the steady-state assumption is used for the intermediate species ES.[6] The derivation gives rise to a combination parameter $K_\mathrm{M} = (k_2 + k_{-1})/k_1$ called **Michaelis constant** and the classical Michaelis–Menten equation:

$$\frac{d[\mathrm{P}]}{dt} = \frac{k_2[\mathrm{E}]_0[\mathrm{S}]}{K_\mathrm{M} + [\mathrm{S}]} \tag{3.77}$$

This example nicely illustrates that the experimental limitations behind introducing the simplifying treatment result in a loss of information: it is impossible to resolve all three parameters (k_1, k_2, k_{-1}) of the original scheme, only two of them (k_2 and combination parameter K_M) are accessible. Further and often unavoidable experimental limitations will typically result in additional loss of parametric information. For example, solubility problems with substrate S could limit experimental work to conditions $[\mathrm{S}]_0 < K_\mathrm{M}$, where the number of reliably resolvable parameters decreases to 1 (i.e. k_2/K_M).

Remarkably, the equation for the expectation of the waiting time can be transformed into a form that is seemingly fully analogous with the deterministic Michaelis–Menten equation:

$$\frac{1}{\langle\tau\rangle} = \frac{k_2[\mathrm{S}]}{K_\mathrm{M} + [\mathrm{S}]} \tag{3.78}$$

This equation is commonly referred to as the **single-molecule Michaelis–Menten equation** in the literature [89]. However, it should not be left unnoticed that the analogy is overwhelmingly accidental in this case. The deterministic Michaelis–Menten equation can only be used in a limited range of parameters because of the steady-state approximation used to derive it, whereas the single-molecule Michaelis–Menten equation is deduced from the exact mathematical solutions and is therefore free of such limitations. Furthermore, it is also tempting to extend the analogy to systems containing several enzyme molecules ($e_0 > 1$) by somehow including e_0 in Eq. (3.78), but this was shown to be incorrect: the formal analogy is limited to single-enzyme kinetics [48].

[6] Two classical papers for the mathematical analysis of the pseudo-steady state hypothesis are [71, 168].

Further theoretical work [17, 51, 54, 101, 113, 151, 167, 182, 184, 191, 200] on the Michaelis–Menten scheme included the development of a stochastic equivalent of the deterministic pre-equilibrium and steady-state approximations, which replace one of the differential equations with an algebraic equation and give the concentration of the intermediate as an explicit function of other concentrations so that the time dependence remains only implicit. A stochastic equivalent can be introduced by assuming that the function $P_{e,s}(t)$ can be obtained as a product of a time dependent R function and an S value, which is characteristic of the state but does not depend on time [48]:

$$P_{e,s}(t) = R_{s_0-s+e-e_0}(t) S_{e_0-e,s_0-s+e-e_0} \tag{3.79}$$

In essence, this assumption states that the probability of the formation of a given number of ES adducts can be obtained simply from the number initial number of enzyme molecules and untransformed product molecules without explicit inclusion of time. Thus, the deterministic pre-equilibrium and steady-state approximations decrease the number of concentrations whose time dependence needs to be calculated, whereas the stochastic equivalent reduces the number of states whose probability needs to be calculated as a function of time. With the new notation introduced by Eq. (3.79), Eq. (3.65) is transformed into the following, more compact form:

$$\frac{dR_p(t)}{dt} = k_2 a_{p-1} R_{p-1}(t) - k_2 a_p R_p(t) \tag{3.80}$$

The new quantity a_p is the steady-state expectation for the number of ES molecules when there are $p(= s_0 - s - -e_0 + e)$ molecules of product formed. Function R_p is actually the sum of $P_{e,s}(t)$ probabilities for states with identical p values:

$$R_p(t) = \sum_{i=0}^{\min(e_0,s_0-p)} P_{i,s_0+e-p-e_0}(t) \tag{3.81}$$

The value of a_p can be calculated using the S function:

$$a_p = \sum_{i=0}^{\min(e_0,s_0-p)} i S_{i,p} \tag{3.82}$$

Different works used different S and a_p functions in the approximation. In a simple possibility, the S function does not even need to be specified as $R_p(t)$ can be given based on a_p solely. An a_p function used in a few studies can be given by a (not fully justified) transfer of the deterministic Michaelis–Menten equation to the stochastic approach [116, 153]:

$$a_p = \frac{e_0(s_0 - p)}{s_0 - p + \frac{\kappa_{-1}+\kappa_2}{\kappa_1}} \tag{3.83}$$

In a slightly more advanced sequence of thought, S values can be given as individual conditional equilibrium probabilities using partition functions as described in statistical thermodynamics [48]:

$$S_{i,p} = \frac{\binom{e_0}{i} \frac{(s_0-p)!}{(s_0-p-i)!} \left(\frac{\kappa_{-1}+\kappa_2}{\kappa_1}\right)^{-i}}{\sum_{i=0}^{\min(e_0,s_0-p)} \binom{e_0}{i} \frac{(s_0-p)!}{(s_0-p-i)!} \left(\frac{\kappa_{-1}+\kappa_2}{\kappa_1}\right)^{-i}} \tag{3.84}$$

This leads to the following a_p function:

$$\begin{aligned} a_p &= \frac{\sum_{i=1}^{\min(e_0,s_0-p)} i \binom{e_0}{i} \frac{(s_0-p)!}{(s_0-p-i)!} \left(\frac{\kappa_{-1}+\kappa_2}{\kappa_1}\right)^{-i}}{\sum_{i=0}^{\min(e_0,s_0-p)} \binom{e_0}{i} \frac{(s_0-p)!}{(s_0-p-i)!} \left(\frac{\kappa_{-1}+\kappa_2}{\kappa_1}\right)^{-i}} \\ &= \min(e_0, s_0-p) \frac{{}_1\mathrm{F}_1\left(-\min(e_0,s_0-p)+1, |e_0-s_0+p|+1, -\frac{\kappa_{-1}+\kappa_2}{\kappa_1}\right)}{{}_1\mathrm{F}_1\left(-\min(e_0,s_0-p), |e_0-s_0+p|+1, -\frac{\kappa_{-1}+\kappa_2}{\kappa_1}\right)} \end{aligned} \tag{3.85}$$

The notation ${}_1\mathrm{F}_1$ here means the confluent hypergeometric function. This formula is also independently known from the stochastic description of reversible second order reactions [122]. The standard deviation corresponding to a_p is given as:

$$\sigma_{a,p} = \sqrt{\frac{a_p K_\mathrm{M}}{V N_\mathrm{A}} - (e_0 - a_p)(s_0 - p - a_p)} \tag{3.86}$$

The expectation and standard deviation for the number of enzyme-substrate adducts can be calculated as follows for any given time t:

$$\langle es \rangle (t) = \sum_{i=0}^{s_0} a_i R_i(t) \tag{3.87}$$

$$\sigma_{es}(t) = \sqrt{\sum_{i=0}^{s_0} \left[(\sigma_{a,i}^2 + [a_i]^2) R_i(t)\right] - [\langle es \rangle (t)]^2} \tag{3.88}$$

Similarly, the expectation and standard deviation for the number of product molecules are given as:

$$\langle p \rangle (t) = \sum_{i=0}^{s_0} i R_i(t) \tag{3.89}$$

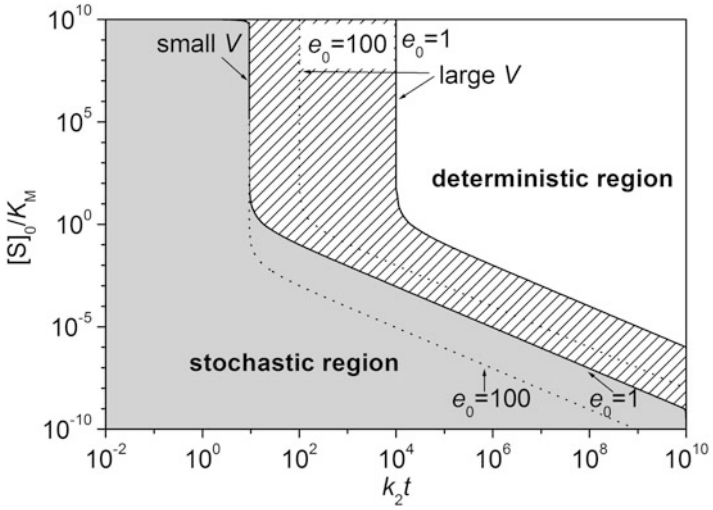


Fig. 3.5 Stochastic map of the Michaelis–Menten mechanism with the number of product molecules formed as the target variable

$$\sigma_p(t) = \sqrt{\sum_{i=0}^{s_0} i^2 R_i(t) - \left(\sum_{i=0}^{s_0} i R_i(t)\right)^2} \tag{3.90}$$

It was also pointed out that a limitation of the stochastic steady-state approximation is its inability to predict the number of ES adducts in an initial time period of the reaction. This dead time can be estimated by the following formula:

$$t_d = \frac{1}{\kappa_\Psi} \ln \frac{20\kappa_\Psi + \kappa_1 a_0}{\kappa_\Psi + \kappa_1 a_0} \qquad \kappa_\Psi = \kappa_1(e_0 + s_0 - 2a_0) + \kappa_{-1} \tag{3.91}$$

The waiting time for the first product molecule to form (τ) can also be estimated based on the approximations:

$$\langle \tau \rangle = \frac{1}{\kappa_2 a_0} \tag{3.92}$$

A stochastic map using the number of product molecules (P) formed the Michaelis–Menten scheme is shown in Fig. 3.5. The graph uses deterministic rate constants for convenience and shows composite parameters on both axes: $k_2 t$ (time in units of $1/k_2$) on the x axis and $[\mathrm{S}]_0/K_\mathrm{M}$ (initial substrate concentration in K_M units) on the y axis. In addition, the map also depends on e_0 (1 and 100 are used in Fig. 3.5) and the overall volume of the system (small V and large V limits). In fact, the extreme case of $s_0 = 1$ sets a lowest meaningful volume, $V_{min} = 1/([\mathrm{S}]_0 N_\mathrm{A})$, for the map.

An approximate solution for master equation (3.65) was given using the binomial approach for both stages of the process (reactant association and product formation) [104]:

$$P_{e,s}^{approx}(t) = \binom{w}{e_0 - e} (1 - qq(t))^{w+e-e_0} qq(t)^{e_0-e} \binom{s_0}{s + e_0 - e} p(t)^{s+e_0-e} (1 - p(t))^{s_0-s-e_0+e} \tag{3.93}$$

Two time functions, $p(t)$ and $qq(t)$ appear in this formula. The definition of the former is as follows:

$$p(t) = \frac{\kappa_{-1} + \kappa_2}{s_0 \kappa_1} W\left(\frac{s_0 \kappa_1}{\kappa_{-1} + \kappa_2} \exp\left(\frac{\kappa_1 (s_0 - e_0 \kappa_2 t)}{\kappa_{-1} + \kappa_2}\right)\right) \tag{3.94}$$

Here, W denotes the Lambert W function, which is the inverse of the xe^x function. The function $qq(t)$ can only be conveniently defined through a series of equations, which also give the meaning of quantity w in Eq. (3.93).

$$w = \min(e_0, e_0 + s - e) \tag{3.95}$$

$$qq(t) = \frac{\lambda_1(t)}{e_0} \frac{1 - e^{(\lambda_2(t) - \lambda_1(t))\kappa_1 t}}{\lambda_2(t)/\lambda_1(t) - e^{(\lambda_2(t) - \lambda_1(t))\kappa_1 t}} \tag{3.96}$$

$$\lambda_1(t) = \frac{e_0 + p(t)s_0}{2} + \frac{\kappa_{-1} + \kappa_2}{2\kappa_1} + \frac{\sqrt{(e_0 + p(t)s_0 + \kappa_{-1} + \kappa_2)^2 - 4e_0 p(t) s_0}}{2} \tag{3.97}$$

$$\lambda_2(t) = \frac{e_0 + p(t)s_0}{2} + \frac{\kappa_{-1} + \kappa_2}{2\kappa_1} - \frac{\sqrt{(e_0 + p(t)s_0 + \kappa_{-1} + \kappa_2)^2 - 4e_0 p(t) s_0}}{2} \tag{3.98}$$

The usefulness of the approximate probability values calculated by (3.93) was systematically tested in small systems ($e_0 = 6 - 10$ and $s_0 = 60 - 100$) using the exact solution of Eq. (3.65) obtained by a direct method.

3.5.3 *Other Enzyme Systems*

Stochastic sequences of thought have been applied to a number of enzymes following more complicated kinetic patterns than the Michaelis–Menten equation. In an experimental study of the oxidation of molecular hydrogen by HynSL hydrogenase from *Thiocapsa roseopersicina*, evidence of the stochastic autocatalytic phenomenon of extinction was obtained [8]: the catalytic reaction came to an end

without consuming the reactants or giving a noticeably degraded enzyme. The reaction was interpreted by a three step catalytic cycle, in which the first process was autocatalytic to an enzyme form:

$$\begin{aligned} \mathrm{E}_2 + \mathrm{E}_3 &\xrightarrow{\kappa_\mathrm{b}} 2\mathrm{E}_3 \\ \mathrm{E}_3 &\xrightarrow{\kappa_\mathrm{c}} \mathrm{E}_4 \\ \mathrm{E}_4 + \mathrm{H}_2 + 2\mathrm{M}_\mathrm{o} &\xrightarrow{\kappa_\mathrm{d}} \mathrm{E}_2 + 2\mathrm{M}_r \end{aligned} \tag{3.99}$$

E_2, E_3 and E_4 are different enzyme forms in the catalytic cycle, H_2 is hydrogen, whereas M_o and M_r are the oxidized and reduced form of the electron acceptor compound benzyl viologen. A state here is identified by giving the number of E_2, E_3 and M_r molecules as e_2, e_3 and p. The master equation can then be stated as using m and h for the number of M_r and H_2 species, and introducing the constant n as the total number of all enzyme forms ($n = e_2 + e_3 + e_4$):

$$\begin{aligned} \frac{dP_{e_2,e_3,p}(t)}{dt} = &-\left[\kappa_\mathrm{b} e_2 e_3 + \kappa_\mathrm{c} e_3 + \kappa_\mathrm{d}(n - e_2 - e_3)m(m-1)h\right] P_{e_2,e_3,p}(t) \\ &+\kappa_\mathrm{b}(e_2+1)(e_3-1)P_{e_2+1,e_3-1,p}(t) + \kappa_\mathrm{c}(e_3+1)P_{e_2+1,e_3+1,p}(t) \\ &+\kappa_\mathrm{d}(n-e_2-e_3+1)(m+2)(m-1)(h+1)P_{e_2-1,e_3,p-2(t)} \end{aligned} \tag{3.100}$$

A method for calculating individual state probabilities was devised under conditions when there is no substantial loss of either M_o nor H_2 in the system. Under such conditions, the master equation can be transformed after the introduction of two new functions:

$$S_{e_2,e_3}(t) = \sum_{i=0}^{\infty} P_{e_2,e_3,i}(t) \quad \text{and} \quad R_{e_2,e_3}(t) = \sum_{i=0}^{\infty} i P_{e_2,e_3,i}(t) \tag{3.101}$$

With the shorthand notation of $\kappa = \kappa_d(m+2)(m+1)h$, the transformed master equation can be deduced from an appropriate linear combination of Eqs. (3.100) and is written as:

$$\begin{aligned} \frac{dS_{e_2,e_3}(t)}{dt} = &-\left[\kappa_\mathrm{b} e_2 e_3 + \kappa_\mathrm{c} e_3 + \kappa_\mathrm{d}(n - e_2 - e_3)\right] S_{e_2,e_3}(t) \\ &+\kappa_\mathrm{b}(e_2+1)(e_3-1)S_{e_2+1,e_3-1}(t) + \kappa_\mathrm{c}(e_3+1)P_{e_2+1,e_3+1}(t) \\ &+\kappa_\mathrm{d}(n-e_2-e_3+1)S_{e_2-1,e_3}(t) \end{aligned} \tag{3.102}$$

$$\frac{dR_{e_2,e_3}(t)}{dt} = -\left[\kappa_b e_2 e_3 + \kappa_c e_3 + \kappa_d(n - e_2 - e_3)\right] R_{e_2,e_3}(t)$$
$$+\kappa_b(e_2+1)(e_3-1)R_{e_2+1,e_3-1}(t) + \kappa_c(e_3+1)R_{e_2+1,e_3+1}(t)$$
$$+\kappa_d(n - e_2 - e_3 + 1)R_{e_2-1,e_3}(t) + 2\kappa_d(n - e_2 - e_3 + 1)S_{e_2-1,e_3}(t) \tag{3.103}$$

The transformed master equation may seem more complicated than the original one, but in fact it is a highly reduced form because the number of states and therefore the number of differential equations depends only on the overall number of enzyme molecules (n). With this technique, the values of the rate constants that lead to highly probable extinction in the process could be found. The method itself is in principle suitable for handling any system containing a catalytic cycle and simultaneous product formation.

3.6 Signal Processing

3.6.1 Signaling with Chemical Networks: General Remarks

Chemical systems can be interpreted as signal processing devices. These systems, i.e. reaction networks, convert (generally speaking) multiple, time-dependent and noisy inputs to responses. The general goal is to determine the relationship between the input and outputs. There are direct and inverse problems. If the network of the chemical system is known, than a **white box** model approach is used. If the network is not known, the procedure called "system identification" is adopted. The situation is more complicated. To understand the mechanism of information processing of the chemical network to be studied should be the subject of decomposition, and functional modules are identified. Some illustrative examples will be given.

Input signals can influence one or more chemical components, may be constant in time or time-dependent. Periodic inputs are often used. At many times, random environment generates noisy input. (Chemical signal processing can be analyzed by deterministic models, but our concern here is stochastic modeling. Randomness has at least two sources: in addition to noisy input, internal noise due to the small number of molecules also can occur.)

An important class of (bio)chemical signal processing is related to the concept of **frequency filtering**. It is a process of selecting, or suppressing, certain frequency components of a signal. The basic types of frequency filters known from the studies of electrical circuits can be implemented by chemical reactions. Arkin [3] and Samoilov et al. [164] demonstrated how low-pass, bandpass and high pass filters, and even more complex chemical filters (such as notch filter and bump filters) can be implemented by simple chemical mechanisms. The functional form of filtering (some characteristic quantity of the output and input signals) as a function of

the input frequency determined the type of filtering. For low-pass and high-pass filtering, the function is monotonically decreasing and increasing, respectively. For a band-pass filter, there is a maximum, if it is not very sharp, there is a reasonable "band" around it. A notch filter passes all frequencies except those in a stop band centered on a center frequency. A bump filter shows two (or more) peaks, so it behaves optimally for two frequencies. Filtering procedures are particularly interesting in stochastic systems due to the intricate relationship between signal and noise.

Generally, the fundamental property of signal processing is its efficiency, and there are some statistical measures, such as mutual information and Fisher information, to characterize it. Examples will be briefly shown.

3.6.2 Signal Processing in Biochemical Networks

3.6.2.1 Evaluation of Signal Transfer by Mutual Information

In the general case, (stochastic) biochemical networks map time-dependent inputs to time-dependent outputs. The efficiency of information transmission can be well characterized by the mutual information between the input signal I and output signal O by the mutual information

$$M(I,O) = H(O) - H(O|I), \tag{3.104}$$

where

$$H(O) \equiv -\int p(O) log p(O) dO$$

is the information-theoretical entropy of the output O having $P(O)$ probability distribution, and $H(O|I) \equiv -\int p(I) dI \int p(O|I) log p(O|I) dO$ is the average (over inputs I) information-theoretical entropy of O given I, with $p(O|I)$ the conditional probability distribution of O given I, as it was studied by Tostevin and ten Wolde [189, 190].

They applied the general formalism to some simple biochemical systems with an input species S and output species X. Denoting the time dependence of the particle numbers with S(t) and X(t) (they are called "trajectories" which should be understood as realizations of stochastic processes), the mutual information between the two trajectories is written as

$$M(S,X) = \int DS(t) \int DX(t) p(S(t), X(t)) log \frac{p(S(t), X(t))}{p(S(t)) p(X(t))}. \tag{3.105}$$

Analytical results can be and were obtained by assuming small (and) Gaussian fluctuations around the stationary value $\langle S \rangle$ and $\langle X \rangle$, respectively: $s := S - < S >$ and $x := X - < x >$.

The mutual information rate defined as $R(\mathbf{s},\mathbf{x}) = \lim_{T\to\infty} M(\mathbf{s},\mathbf{x})/T$ can be calculated from the power spectra of the fluctuations (derived as the Fourier transform of the elements of the covariance matrix) as

$$R(\mathbf{s},\mathbf{x}) = -\frac{1}{4\pi}\int_{-\infty}^{\infty} d\omega \ln\left\{1 - \frac{|S_{sx}(\omega)|^2}{S_{ss}(\omega)S_{xx}(\omega)}\right\} \tag{3.106}$$

$R(\mathbf{s},\mathbf{x})$ takes into account temporal correlations between the input and output signals. Equation (3.106) is exact for linear systems with Gaussian noise. While this is not a trivial restriction for chemical systems, the theory comes from the studies of optical fiber communications [126]. However, in the case of chemical reactions, the detection of input signals may generate correlations between the signal and the intrinsic noise of the reactions. If there is no correlation, the **spectral addition rule** expressed by Eq. (3.107)

$$S_{xx}(\omega) = N(\omega) + g^2(\omega)S_{ss}(\omega) \tag{3.107}$$

is valid. Here $N(\omega)$ is the internal fluctuation and $S_{ss}(\omega)$ is the power spectrum of the input signal, furthermore $g^2(\omega) = \frac{|S_{sx}(\omega)|^2}{S_{ss}(\omega)^2}$ is the frequency-dependent gain.

The spectrum of transmitted signal is expressed as $P\omega = g^2(\omega)S_{ss}(\omega)$, so Eq. (3.106) can be rewritten as

$$R(\mathbf{s},\mathbf{x}) = -\frac{1}{4\pi}\int_{-\infty}^{\infty} d\omega \ln\left\{1 + \frac{P(\omega)}{N(\omega)}\right\} \tag{3.108}$$

Specific examples to process time continuous signals were given [189]. For all the three cases, the input signal is taken as $0 \xrightarrow{k} S$ and $S \xrightarrow{\lambda} 0$. This reaction leads to Poissonian stationary distribution to be well approximated by Gaussian, at least for a large number of particles. The reliability of the transmission is evaluated by the gain-to-noise ratio defined as $\frac{g^2(\omega)}{N(\omega)}$.

Reversible Binding

$$\mathrm{S} + \mathrm{W} \underset{\mu}{\overset{\nu=k_f\mathrm{W}}{\rightleftharpoons}} \mathrm{X} \tag{3.109}$$

This scheme describes reversible binding, it might be considered as ligand-receptor interaction, or a reaction between enzyme and its substrate. The input signal is taken to be the total number of both bound and unbound molecules, i.e. $S_T(t) = S(t) + X(t)$. This scheme (called "motif" in the literature of systems

biology and related areas) works as a low-pass filter, since the gain-to-noise ratio is constant at low frequencies, but inversely proportional ω^2 for high frequencies. The result suggests that it the processing of high-frequency input is unreliable in biochemical signal processing, since this scheme is certainly a frequently used biochemical signal detector.

Signaling Molecule Is Deactivated upon Detection

A remarkable example of the deactivation of the signaling molecule is when the activation of a receptor is followed by endocytosis:

$$S \overset{\nu}{\to} X \overset{\mu}{\to} 0. \tag{3.110}$$

In this case, the mutual information between instantaneous values of S and X is zero. However, the analysis of the mutual information between the input and output trajectories shows that the gain-to-noise ratio diverges (at least in the limit of infinitely fast reaction), so the scheme is able to transmit reliably input signals varying with high-frequency.

Coarse-Grained Model for Enzymatic Reactions

The scheme

$$S \overset{\nu}{\to} S + X \tag{3.111}$$

$$X \overset{\mu}{\to} 0. \tag{3.112}$$

is a somewhat oversimplified model for enzymatic reactions or gene activation. For this scheme, $\frac{g^2(\omega)}{N(\omega}$ does not depend on ω, so the fidelity of the transmission is not frequency-dependent.

3.6.2.2 Impact of Network Structure on the Transmission

Obviously, the network structure influences signal transmission. In particular, it was studied by processing constant input signals [203]. Specifically, they studied the maximum mutual information between the input (chemical) signal and the output (genetic) response for small networks. Actually possible networks of three chemical species were considered each under the control of one regulator. Instead of mass action kinetics, simple rational functions were adopted, and linear noise approximation was used. The fidelity of transmission was characterized basically by

(average) mutual information. While generally all small networks proved to be quite good signal transducers, networks with negative feedback work somewhat better.

A comparative study of certain networks with time-dependent inputs was given by [46]. The simplest network is a cascade reaction, and more complicated networks can be generated by supplementing it with autoregulation, feedforward and feedback connections. It is plausible that in a cascade with n nodes, the information about the input encoded in the signal at node $i+1$ cannot be greater than the information at node i. The mechanism is irreversible, so the lost information about the input cannot be recovered later in the cascade.

The analysis is based on the calculation of the gain-to-noise ratio. The details of the calculation are not repeated here, we restrict ourselves to tell the qualitative consequences.

Two-Component Cascade with Autoregulation

Autoregulation modifies the diagonal elements of the Jacobian matrix. The deterministic model of autoregulation of a simple cascade is given by Eq. (3.113a):

$$\frac{dx}{dt} = f(x)s - \mu_x(x) \tag{3.113a}$$

$$f(x) = \frac{v\beta}{K+x} \quad \begin{cases} \beta = K, & \text{negative regulation} \\ \beta = x, & \text{positive regulation} \end{cases} \tag{3.113b}$$

$$J_{xx} = -\mu_x + \langle s \rangle \left[\frac{\partial f(x)}{\partial x} \right]_{s.s.} \tag{3.113c}$$

For negative regulation, $J_x x > \mu_x$, while $J_x x < \mu_x$ holds for positive regulation. Turning to the stochastic model of the system, calculations show that, interestingly, information transmission is not influenced by the autoregulation at the output of this network scheme.

Three-Component Cascade with Autoregulation

$$\frac{dv}{dt} = f(v)s - \mu_v(v) \tag{3.114a}$$

$$\frac{dx}{dt} = \beta v - \mu_x x f(v) = \frac{v\beta}{K+x} \quad \begin{cases} \beta = K, & \text{negative regulation} \\ \beta = v, & \text{positive regulation} \end{cases} \tag{3.114b}$$

$$J_{vv} = -\mu_x + \langle s \rangle \left[\frac{\partial f(v)}{\partial v} \right]_{s.s.} \tag{3.114c}$$

The analysis of the stochastic version of the scheme shows that the gain-to-noise ratio decreases compared with the simple cascade. Negative autoregulation therefore tends to suppress slowly varying signals relative to the simple cascade, while positive autoregulation amplifies them.

Feedback from the Output Signal to an Upstream Component

$$\frac{dv}{dt} = f(x)s - \mu_v v \tag{3.115a}$$

$$\frac{dw}{dt} = \beta v - \mu_w w \tag{3.115b}$$

$$f(x) = \frac{vC^m}{K^n + x^n} \quad \begin{cases} \mathrm{C} = \mathrm{x}, & \text{positive feedback} \\ \mathrm{C} = \mathrm{K}, & \text{negative feedback} \end{cases} \tag{3.115c}$$

Two cases of feedback were studied. First, the output x sends back information to an upstream component. The feedback x to v does not increase the information, and is actually a source of noise to the $s \to v$ pathway. Since the noise is highest at low frequencies, this network scheme works as a high-pass filter.

A Four-Component Cascade with Feedback from an Intermediate Component

Second, there is a feedback from an intermediate to component precedes it in the cascade. A four-component (i.e. a three step) cascade was studied:

$$\frac{dv}{dt} = f(w)s - \mu_v \tag{3.116a}$$

$$\frac{dw}{dt} = \beta v - \mu_w w \tag{3.116b}$$

$$\frac{dx}{dt} = \gamma w - \mu_x x \tag{3.116c}$$

$$f(w) = \frac{vC^n}{K^n + w^n} \begin{cases} C = w, & \text{positive feedback} \\ C = K, & \text{negative feedback} \end{cases} \tag{3.116d}$$

At low frequencies, positive feedback amplifies the signal and the noise introduced at the levels of v and w, but not noise introduced at x. At low frequencies, the gain-to-noise ratio increases relative to the simple cascade. However, at high frequencies, the positive feedback reduces the gain and the noise upstream of x, but not the internal noise; therefore the gain-to-noise ratio is reduced compared to the simple cascade. The implication is that a network with negative feedback reduces the gain at low frequencies, reducing the gain-to-noise ratio. At high frequencies, the feedback amplifies the signal but not the internal noise, leading to an increase in the gain-to-noise ratio.

Generally speaking, the most important design principle is the autoregulation and feedback can improve information transmission, but only if they occur upstream to the noise source. The results are admittedly approximative, since among others, the linear noise approximation was used.

3.6.2.3 Further Studies

Stochastic Signaling and Noise-Induced Bistability

Samoilov et al. [165] showed by using both analytical and numerical investigation that at least in one ubiquitous class of (bio)chemical-reaction mechanisms, namely in enzymatic futile cycles, the external noise may induce a bistable oscillatory (dynamic switching) behavior that is both quantitatively and qualitatively different from what is predicted or possible deterministically. When two metabolic pathways work simultaneously in opposite directions and have no overall (mass) effect, the only effect is energy dissipation (i.e. entropy production).

The effect of external noise (on E_+) is summarized by the addition of a correction (diffusion) term to the deterministic result:

$$R_{N(\sigma)}(X_{ss}, E_+; E_-) = E_+ - \frac{k_- E - (X_0 - X_{ss})(K_+ + X_{ss})}{K_+ X_{ss}(K_{-+} X_0 - X_{ss})} + \frac{\sigma^2 k_+ K_+}{(K_{++} X_{ss})^2} = 0 \tag{3.117}$$

Here $\{E_+, E_-\}$ denote the forward and reverse (e.g., activating and deactivating) enzymes, and X, X^* stand for the concentrations of the forward substrate and product, respectively. k_s are the catalytic constants of the enzyme (the complex to product reaction rates), and K_+ and K_- are the Michaelis constants for the substrate reaction. $R_N[X^*_{ss}, E_+; E_-] = 0$ is the stationary response curve (nullcline) relationship, X_0 is the total amount of X and X^* (s^* is the signaling molecule, σ is the forward enzyme noise strength.

That is, an enzymatic futile cycle can act not only as a signal transducer but also as a stochastic amplifier. Under certain circumstances, the system is not expected to have any significant amplification in a deterministic system, but shows substantial signal gain stochastically Fig. 3.6.

As the external signal increases and begins to approach the sigmoidal region, the response level becomes bistable and begins to transiently switch between two states with a characteristic amplitude/frequency distribution. This dynamic switching may be viewed as an extra information channel through which more (accurate) signals can be passed to the downstream processes. The extra channel is obtained simply due to the stochastic nature of the chemical reactions (and it is not related to the network structure of the reaction) (Fig. 3.7).

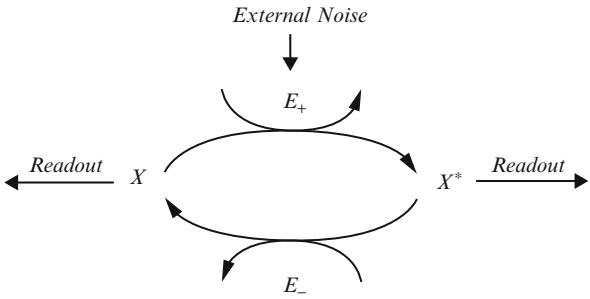


Fig. 3.6 The enzymatic futile cycle reaction mechanism (Adapted from [165])

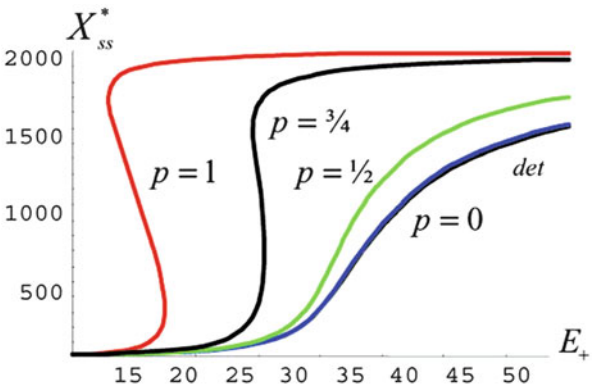


Fig. 3.7 Transition between uni- and multistationarity. The bifurcation parameter p is the exponent of E_+ in a relationship connecting the variance E_+ and E_+ (Adapted from [165])

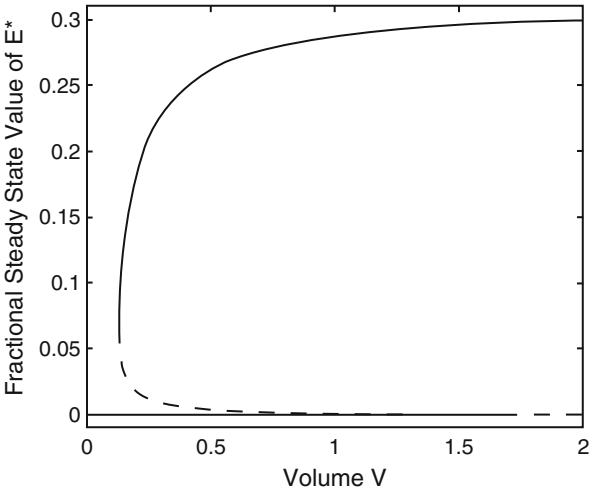


Fig. 3.8 Stochastic bifurcation plot of the fractional steady-state values of activated kinase, as a function of the volume V. The *solid curves* represent the maxima of the steady-state distribution, and the *dashed curve* represents the minimum of the distribution. $V = 1.67$ is the critical value, where the bistability disappears (From [22])

In a somewhat similar analysis, [22] showed the volume of the system V and the available free energy as bifurcation parameters in a signaling pathway (namely in driven phosphorylation-dephosphorylation cycle kinetics with autocatalytic kinase). Figure 3.8 demonstrates the existence of bistability for low volume.

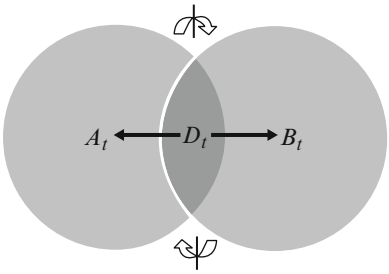


Fig. 3.9 Illustrative example of a network with two modules. Each *circle* represents a module, i.e. a group of biochemical species. The network species are grouped into three non-overlapping sets A, B and D as shown, in such a way that $A_t \leftrightarrow_{D_t} Bt$ – i.e. the trajectories (up to any time t) of A and B are independent given the trajectory of D. Hence A_t and B_t contain no mutual information given D_t, all information transfer between the two modules being conveyed via D_t (From [24])

Kinetic Independence: A Framework for Analyzing Signal Processing

Bowsher and his colleagues gave a framework to analyze information encoding and propagation by biochemical reaction networks [24–26]. They formulated a kind of **inverse problem** starting from measured time courses of different biochemical species in the reaction network and the general goal is to express the interaction of components by some probabilistic relationships. The basic concept is the **conditional independences** between species trajectories. Reaction networks are decomposed into modules by using dynamic conditional independence properties and some stoichiometric information. An algorithm, called MIDIA, was developed. It is a graphical algorithm that makes use of a new representation of the kinetics of reaction networks called the Kinetic Independence Graph (KIG). The conceptual and mathematical framework was elaborated in [24].

Network species can be decomposed into non-overlapping groups $A; D; B$ such that $A(t)$ and $B(t)$ contain no mutual information given D(t). The key concept is the **conditional independence**: $A_t \nleftrightarrow_{D_t} B_t$, which gives two modules (circles) as Fig. 3.9 shows.

Species in the overlap region, D, are informational intermediaries. The MIDIA algorithm computes exact network decompositions based on dynamic independence properties of the modules and also able to identify important biochemical intermediaries that result in the overlap of modules.

In [26], the authors provided a new framework for understanding sources of stochasticity. Fluctuations can be decomposed into multiple components, and some advice might be given to the experimentalists which component quantities should be measured.

3.6.3 Signal Processing in Olfactory Systems

3.6.3.1 Fisher Information and Optimal Signal Transmission

Signal processing in the early phase of olfactory processing is obviously chemosensory information processing by chemical reactions. In the theory of neural transmission and neural coding, the Fisher information measure is often used to characterize signal transmission. Brunel and Nadal [32] discusses the relationship between the statistical concept of Fisher information and mutual information (coming of course from information theory).

Let's assume the probability density function of random variable X with values x depends on a scalar parameter θ, so it is written as $f(x;\theta)$.

Fisher information about parameter θ in random variable X is given by Eq. (3.118):

$$J^X(\theta) = E\left(\left(\frac{\partial \ln f(X;\theta)}{\partial\theta}\right)^2\right) = \int_M \left(\frac{\partial \ln f(x;\theta)}{\partial\theta}\right)^2 f(x;\theta)d\mu(x). \quad (3.118)$$

Here μ is an additive probability measure, and M is the support (i.e. the closure of the set of possible values of a random variable). It is not a measure of information in information-theoretical sense. Fisher information tells the relationship between changes in the parameter and changes in the distribution. To put it in another way, it tells how well the parameter change can be estimated by knowing the changed distribution.

The **Cramér-Rao** inequality relates the variance of estimator $\hat{\theta}$ and Fisher information

$$\frac{1}{J^X(\theta)} \leq Var(\hat{\theta}),$$

more precisely, there is a lower bound for variance of any parameter estimator.

In [145] there is a practical introduction to use Fisher information for finding optimal signal in optimally detecting odorant concentration in olfactory system. Before discussing how to use Fisher information to optimal signal detection, the stochastic kinetic models of some odor intensity detection schemes are presented. As [146] emphasizes, the binding-activation cascade can be classified into two categories, concentration detectors and flux detectors.

3.6.3.2 Stochastic Kinetic Models of Odor Intensity Detection

Both deterministic and stochastic models exist to describe odor intensity detection, in particular the response of the system to chemical stimulus. Figure 3.10 shows the stimulus response curve in both cases.

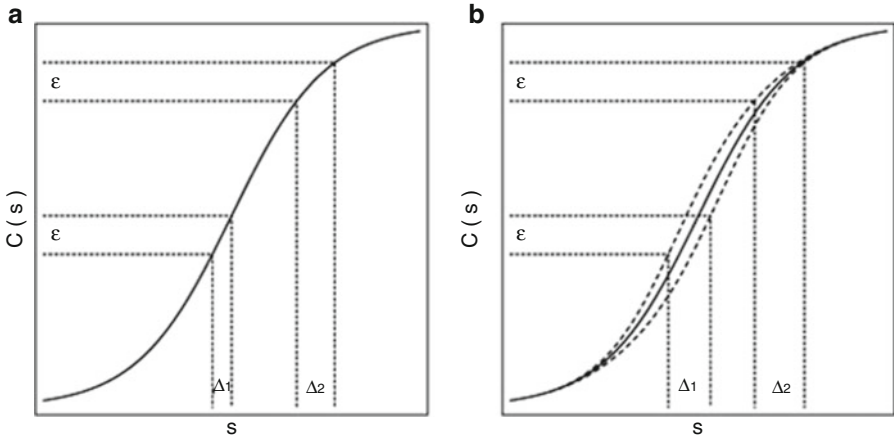


Fig. 3.10 Deterministic and the stochastic concepts. (**a**) A unique response, $C(S)$ is plotted against the odorant log-concentration, s. Two equal changes ϵ in response, the number of activated receptors $C(s)$, are caused by different changes in corresponding odorant concentrations, $\Delta_1 < \Delta_2$, because of varying slope of the input–output function. Therefore, the changes in the odorant concentration in the region around Δ_1 can be determined from the knowledge of the response $C(s)$ more precisely than in region around Δ_2. (**b**) Even the fluctuations of the response are taken into account and plotted versus the odorant log-concentration, s. Hence the changes Δ_1 and Δ_2 in odorant log-concentration are different from the situation as in (**a**). Due to the larger variability of the responses in the central part of the transfer function: Δ_1 Δ_2 (Based on [146])

Basic Model

The simplest model is based on the assumption that a receptor occupied is instantaneously activated: $\mathrm{A} + \mathrm{R} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{C}$, where A represents an unbound molecule of odorant, R unoccupied receptor and C stands for bound activated receptor (complex of the odorant molecule and the receptor), k_1 and k_{-1} are the reaction rates coefficients of association and dissociation of the odorant molecules.

Model of Simple Activation

The receptors really appear in three states: unbound, R, occupied but not activated, C^*, and occupied activated, C.

$$\mathrm{C}^+ \underset{k_{1\mathrm{N}}}{\overset{k_{-1}}{\rightleftharpoons}} \mathrm{A} + \mathrm{R} \underset{k_{-1}}{\overset{k_{1\mathrm{A}}}{\rightleftharpoons}} \mathrm{C}$$

Double-Step Model

This model also contains receptors in three states, but here it is assumed that the occupied receptor can become activated only with a delay after the binding:

$$\mathrm{A} + \mathrm{R} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{C}^{+} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \mathrm{C}$$

Flux Detector

As opposed to concentration detectors, "flux detectors" accumulate the stimulus molecules in a perireceptor compartment. If the arrival of stimulus molecules is balanced with deactivation, constant effective stimulus concentration at constant adsorptive flux of stimulus molecules is generated [83].

$$\mathrm{A_E} \overset{\mathrm{k_I}}{\rightleftharpoons} \mathrm{A}$$

$$\mathrm{A} + \mathrm{R} \underset{\mathrm{k_{-1}}}{\overset{\mathrm{k_1}}{\rightleftharpoons}} \mathrm{C} \xrightarrow{\mathrm{k_0}} \mathrm{R} + \overline{\mathrm{A}}$$

In this simple model of a flux detector in which receptor molecules themselves catalyze the deactivation, the dose-response relationship is linear (and not sigmoid).

Results

The mean and the standard deviation of the number of activated receptors as a function of odorant log-concentration can calculated either analytically or at least numerically, as Fig. 3.11 shows.

3.6.3.3 Estimation of Optimal Olfactory Signals

In the deterministic model, the optimality measure is the first derivative of the input–output function with respect to the concentration of odorant is $J_{det}(s) = \frac{\partial f}{\partial(s)}$, which measures the slope of the function. In their analysis, Pokora and Lansky [146] assumed that the particle number is sufficiently large to adopt continuous approach, and for the actual models the deterministic model coincides with the equation for the expectation, i.e. $E(C(s)) = f(s)$. Three criteria were derived. The first criterion is based on the approximative assumption for the expectation, therefore the criterion is $J_1(s) = \frac{\partial E(C(s)}{\partial(s)}$, and the optimal concentration s_1 maximizes it:

$$J_1(s) = \max_s J_1(s). \tag{3.119}$$

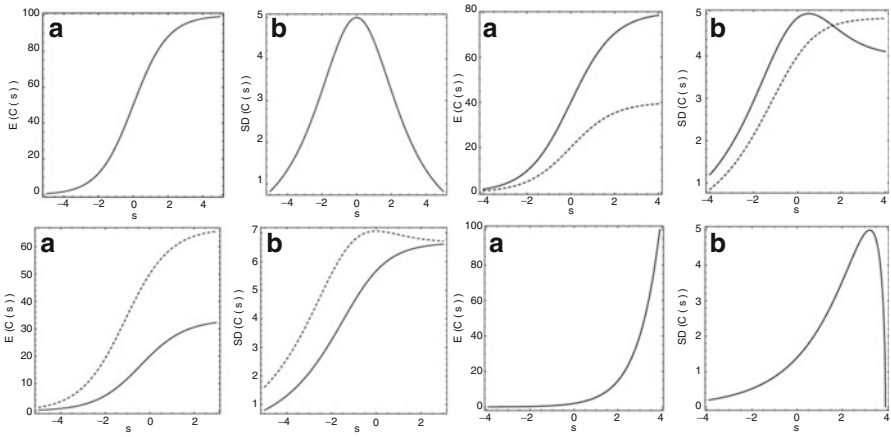


Fig. 3.11 Mean and standard deviation of the number of activated receptors in the basic model, model of simple activation, double-step model and flux detector as functions of the odorant log-concentration (Based on [146])

The second criterion $J(s)$ (keeping the notation of the original paper) is the Fisher information adopted. The assumption is that to the continuous random variable $C(s)$, there is a family of probability density functions g(x;s) where the odorant concentration s is the parameter.

The optimal concentration s is estimated as $\hat{s}$ from sampling responses $\{x_1, x_2, \ldots, x_n\}$ of $C(s)$ by using the Fisher information

$$J(s) = E\Big(\frac{\partial \ln g(s)}{\partial s}\Big)^2 = \int \frac{1}{g(x;s)}\Big(\frac{\partial g(x;s)}{\partial s}\Big)^2 dx. \tag{3.120}$$

The independence of measurements is far are from being trivial. The Rao–Cramer equality is

$$Var\hat{s} \geq J(s)^{-1}. \tag{3.121}$$

Since the higher $J(s)$ is, the better the estimation of s is, the optimal concentration s_0 maximizes $J(s)$. However, the analytical expression of $J(s)$ is generally difficult, so the third criterion J_2 is defined based on another approximation, namely assuming the knowledge of the first two moments $E(C(s))$ and $Var(C(s))$ of the distribution only:

$$J_2(s) = \frac{1}{Var(C(s)}\Big(\frac{\partial E(C(s))}{\partial s}\Big)^2 = \frac{J_1(s)^2}{Var(C(s).} \tag{3.122}$$

$J_2(s)$ is a lower bound to $J(s)$, $J_2(s) \leq J(s)$, so an optimal concentration s_2 can be defined.

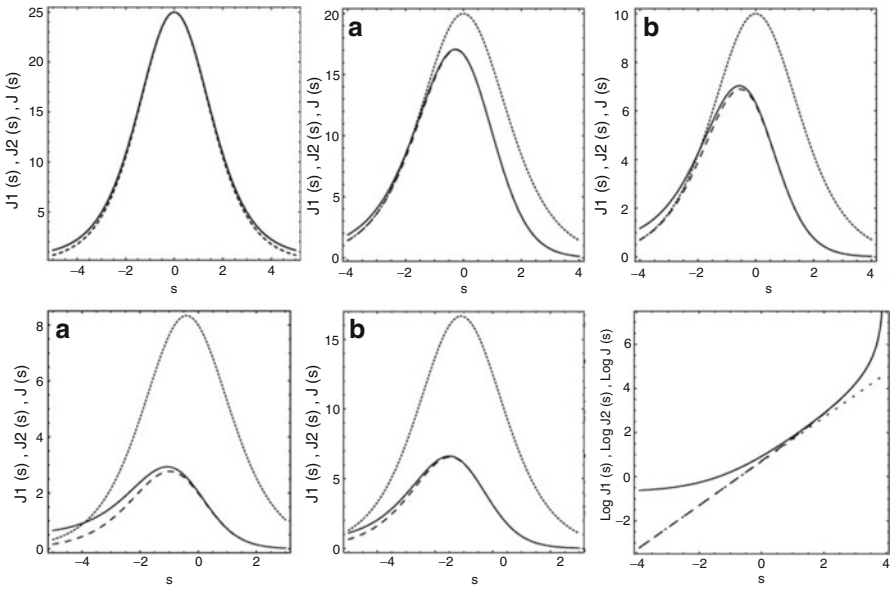


Fig. 3.12 Optimality criteria in the basic model, model of simple activation, double-step model and flux detector as functions of the odorant log-concentration (Based on [146])

Results

The optimality criteria $J(s)$, $J(s)$ and J_2 were computed, as Fig. 3.12 shows.

For the basic model, the criteria are unimodal and the deterministic and a stochastic models lead to the same result. Similarly, for the flux detector from the deterministic and stochastic models the same optimal concentration can be calculated. However, from the model of simple activation the deterministic and statistical approaches can give different results, the optimum from statistical point of view is located at lower concentrations of odorant than that obtained with the approach based on the slope of the input–output function. For the double step model, the deterministic and statistical approaches give different results, as the shift of the locations of the maxima illustrates.

3.6.4 Calcium Signaling

Many cells use changes in cytosolic Ca^{2+} concentration for signaling.

Theoretically, it is a paradigmatic example how local stochastic events (opening and closing of channels) lead to global periodic behavior. A channel type present in the endoplasmic reticulum membrane of many cells is the inositol 1, 4, 5-trisphosphate ($IP3$) receptor channel.

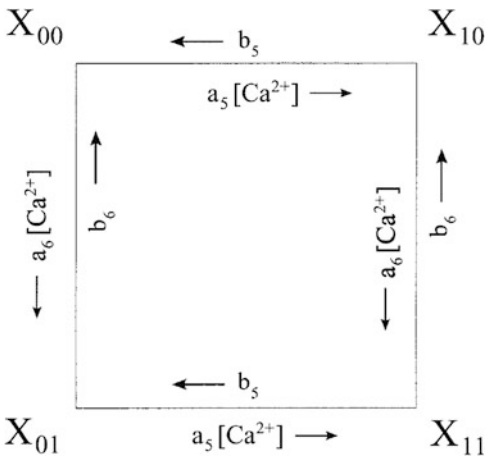


Fig. 3.13 A lumped kinetic scheme of the channel kinetics. X_{00}: state with no Ca^{2+} bound; X_{10}: activated state; X_{11} and X_{01}: inhibited states. An index is 1 if an ion is bound and 0 if not. Transition rates are shown at the edges of the *rectangle* (Based on [58])

There is a hierarchy of spatiotemporal events (e.g.):

- Single channel opening (blip)
- The opening of several closely packed channels (puff)
- Cooperation of puffs may set off a wave traveling through the cell
- Waves may occur periodically, so they can be seen as global oscillations.

It is interesting for us that for parameters taken from the nonoscillatory deterministic regime, oscillations with different frequencies emerge induced by stochastic channel dynamics.

Figure 3.13 shows a lumped kinetic scheme of the channel kinetics:

Due to the probabilistic nature of the binding and the dissociation process in this system, stochastic kinetic models are appropriate, and they are the basis of the short period oscillations. On a higher level of the hierarchy, there is a cooperation among puffs which leads to the nucleation of waves. Simulations [58] showed that such probabilistic nucleation is responsible for the generation of long period oscillations.

The studies of the stochastic aspects of intracellular Ca^{2+} dynamics demonstrated the fundamental role of fluctuations arising from the control of the release channel by Ca^{2+} and $IP3$.

In a more general way, internal noise has a special role for calcium signaling in a coupled cell system [109,202]. Addition of noise can expand the region of parameter space in which cycles occur. The effect can be maximized for particular values of the system size or noise strength. (The phenomenon is called "internal stochastic resonance", cf. with Sect. 3.10, where stochastic resonance is reviewed.)

Multi-scale modeling supports the view that global calcium signals are driven by single channel fluctuations [175], where stochastic models of channel gating is integrated into deterministic diffusion models.

"Self-organized criticality" (SOC) is known as a general phenomenon occurring in nature and society related to the emergence of macroscopic complexity from

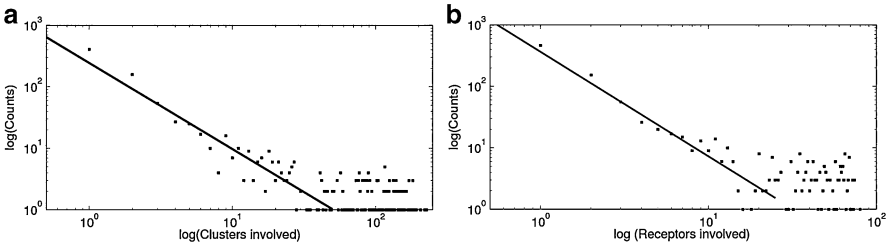


Fig. 3.14 Distributions of clusters and of receptors involved obtained with simulations of intercluster dynamics with calcium accumulation. Power law distribution can be rather well fitted (Based on [112])

spontaneous local interactions [7]. Patterns of activity characterized by different length scales can occur with a probability density that follows a power law with pattern size.

Stochastic localized activity may elicit global Ca^{2+} signals [112]. The amplitude distribution of local signals deviates from Gaussian for a significant fraction of large size events. It can be log-normal or power law. For some reasonable region the analysis of the model shows that a global elevation of the Ca^{2+} concentration plays a major role in determining whether the puff size distribution is long-tailed or not. This suggests that Ca^{2+} is a key to determine whether $IP3$-mediated Ca^{2+} signals can display a SOC-like behavior or not. For the illustration of SOC-like behavior, see Fig. 3.14.

3.7 Gene Expression

3.7.1 A Very, Very Short Review of Biochemical Background

Gene expression is the complicated process which converts genetic information from a DNA sequence into protein. Prokaryotes do not have nucleus, so DNA can be found in any part of the cell, while in eukaryotes, it is located in the nucleus. In prokaryotes there are two main processes: **transcription** and **translation**. In eukaryotes there is one more process: **splicing**.

Transcription is a complicated series of events that use DNA to synthesize messenger RNA (mRNA) by using enzyme RNA polymerase as a catalyst. The series of events contain (in prokaryotes) *binding*, *initiation*, *RNA synthesis*, *elongation*, *termination*. Specifically, a **promoter** is a region of DNA that induce binding. Eukaryote transcription is much more complicated, but some processes have the same mechanism.

Splicing is a modification of the nascent transcript in which certain nucleotide sequences (*introns*) are removed and other sequences (*exons*) are retained or joined.

Translation is a process when the mRNA is processed by a ribosome complex using the **genetic code**, which relates the DNA sequence to the amino acid sequence in proteins. It contains more elementary steps, such as *initiation*, *elongation*, *translocation* and *termination*.

Degradation While DNA is stable, RNA and protein molecules can be subject to degradation, and it is also an important step in the regulation of gene expression.[7]

3.7.2 *Measurement of Noise in Genetic and Other Biochemical Networks*

New experimental techniques give the possibility to measure the dynamics of gene expression in single cells and reveal bursts of both mRNA and protein synthesis in different types of organisms [152, 169]. Elowitz et al. [53] introduced the concepts of extrinsic and intrinsic noise in gene expression (for its mathematical analysis see [185]). Ozbudak et al. [139] showed how noise in gene expression depends on the modification of the parameters of the underlying biochemical processes. Generally speaking, while **intrinsic noise** results from the discrete nature and inherent randomness of biochemical reactions such as promoter remodeling, transcription, translation, and degradation of mRNA and protein species, **extrinsic noise** arises from intercellular differences in the amounts of cellular components (e.g., RNA polymerases and ribosomes). Single-molecule experiments (which were mentioned earlier related to the detection of single-molecule events induced bistability (Sect. 3.2.2.3) showed that proteins were synthesized in a rapid, burst-like fashion [37].

The significance of stochasticity in endogenous biochemical networks have been summarized by [169]:

- High-throughput studies have been carried out in yeast
- Three-color experiments have been used to quantify different contributions to extrinsic fluctuations
- Stochasticity has been measured in mammalian cells, both in gene expression and in the p53 network,
- In slime mold
- In HIV transactivation,

[7]There are some recent developments that suggest that gene expression might be circular [69]. Eukaryotic gene expression can be viewed as a circular process, whereby "first" (transcription) and the "last" (mRNA degradation) are interconnected.

- In bacterial chemotaxis
- In the timing of mitosis, meiosis (19), and lysis by phage lambda

Fluctuations are detected mostly in protein concentrations, but also in mRNA's.

3.7.3 Stochastic Kinetic Models of Gene Expression

3.7.3.1 General Remarks

The perspective that models of gene expression should have stochastic elements goes back to the pioneering works of Rigney and Schieve [159], D. Rigney [157, 158], and O. Berg [18], but these works came too early for main stream molecular biologists. Stochastic chemical kinetics became the lingua franca of modeling gene regulatory networks and related fields 20 years later, due to some highly cited papers [4, 119].

3.7.3.2 A Three-Stage Model of Gene Expression

Genes and proteins form transcriptional regulatory networks, and they are often present in small numbers in the cell, so CDS models proved to be the appropriate tool to study the behavior of such kinds of systems. The three-stage model written as

$$\text{inactive gene} \underset{\lambda_1^-}{\overset{\lambda_1^+}{\rightleftharpoons}} \text{active gene} \xrightarrow{\lambda_2} \text{mRNA} \xrightarrow{\lambda_3} \text{protein} \tag{3.123}$$

supplemented with two degradation steps (the second also called as proteolysis):

$$\text{mRNA} \xrightarrow{\gamma_m} 0 \quad \text{protein} \xrightarrow{\gamma_p} 0 \tag{3.124}$$

seems to be generally accepted (say [23, 142, 154, 169]). From a strict chemical point of view, all steps can be considered as "net reactions" containing several (or many) elementary reactions. The state of the system can be characterized by $\mathbf{n} = (n_1, n_2, n_3)$, where n_1, n_2 and n_3 are the number of active genes, mRNAs and proteins per cell, respectively. The standard model consists of six first order reaction steps (so the whole system is a special compartmental system). The time scale of the gene activity is $\tau_1 := \frac{1}{\lambda_1^+} + \lambda_1^-$; $\tau_2 := \frac{1}{\gamma_m}$ and $\tau_3 := \frac{1}{\gamma_p}$ are the time scale of the degradation steps.

Adopting a Markovian assumption[8]

$$\begin{aligned}\frac{dP(n_1,n_2,n_3)(t)}{dt} &= \lambda_1^+(n_1^{max}-n_1+1)P(n_1-1,n_2,n_3)-\lambda_1^+(n_1^{max}-n_1)P(n_1,n_2,n_3)\\ &+\lambda_1^-(n_1+1)P(n_1+1,n_2,n_3)-\lambda_1^- n_1 P(n_1,n_2,n_3)\\ &+\lambda_2 n_1 P(n_1,n_2-1,n_3)-\lambda_2 n_1 P(n_1,n_2,n_3)\\ &+\frac{(n_2+1)}{\tau_2}P(n_1,n_2+1,n_3)-\frac{n_2}{\tau_2}P(n_1,n_2,n_3)\\ &+\lambda_3 n_2 P(n_1,n_2,n_3-1)-\lambda_3 n_2 P(n_1,n_2,n_3)\\ &+\frac{(n_3+1)}{\tau_3}P(n_1,n_2,n_3+1)-\frac{n_3}{\tau_3}P(n_1,n_2,n_3),\end{aligned} \tag{3.125}$$

where n_1^{max} denotes a constant number of switching genes. By using the generating function methods, the time-dependent solution can be obtained. Since often stationary protein fluctuations (protein noise) can be measured, it is worth to calculate some characteristic quantities. There are two often used measures of noise in terms of the first two moments of a probability distribution, the normalized stationary variance in the number of protein molecules per cells and the Fano factor (variance divided by average).[9] The terms contributing to the normalized stationary variance:

$$\frac{\sigma_3^2}{<n_3>^2} = \boxed{\text{from individual birth and death}} + \boxed{\text{from spontaneous RNA noise}}$$

$$+ \boxed{\text{from forced mRNA noise, originating in the gene activation-inactivation.}} \tag{3.126}$$

The first term is at least approximately Poissonian (or some other unimodal) distribution.

For rapidly fluctuating genes and mRNAs, protein fluctuations can be calculated as

$$\frac{\sigma_3^2}{<n_3>^2} = \frac{1}{<n_3>} + \frac{\sigma_E^2}{<n_E>^2}\frac{\tau_E}{\tau_E+\tau_3}, \tag{3.127}$$

where E represents the external environment of either mRNAs or active genes. The two terms of the right-hand side are related to the concept of disorder

[8] A non-Markovian process may arise due to *time delay* in transcription, translation and/or in degradation [28, 29, 156]. It is known that even in deterministic models, there are time-delay induced bifurcations [91]; however, delay-induced stochastic oscillations in gene regulation were also found.

[9] For the details see e.g. Sect. 2.2 of [142].

[204] associated with fluctuation in rate constants due to slow conformational fluctuation.[10] Both transcriptional and translational bursting contribute to protein fluctuations, and the identification and separation of the these contributions both experimentally and theoretically seem to be subject of present and near-future studies (e.g. [163]). Noise can be characterized by the Fano factor taken for the protein fluctuation due to translational burst:

$$F := \frac{\sigma_3^2}{< n_3 >} = 1 + < b >, \tag{3.128}$$

where b is the number of translations per transcript.[11] For Poissonian stationary fluctuations $F = 1$, and $F > 1$ has a larger spread (larger noise).

Assuming for the gene g both transcriptional and translational bursts, noise strength $F(g)$ can be decomposed into transcriptional and translational components [163]. If $B(g)$ denotes the transcription burst size of gene and $C(g)$ is the number of proteins translated from one mRNA molecule then, ignoring any other noise contributors, the noise strength can be approximated as

$$F(G) := 1 + C(g)B(g) \tag{3.129}$$

Analytical Methods

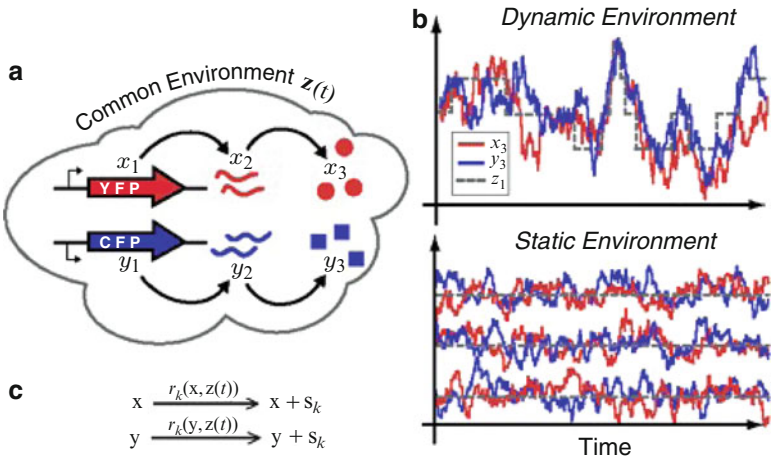
While simulation methods are most often used to study the time-dependent properties of gene regulatory systems, in the previous examples stationary variances were calculated [198]. Walczak et al. [197] reviewed the scope and limits of the more general approach of analytical methods by using master equations, and continuous approximations based on Fokker–Planck and Langevin equations. Some results were obtained both for transcriptional and translational bursts. It was shown that both types of bursts generates more noise than a simple birth-and-death mechanism leading to Poisson distribution. Stronger deviations from the Possionian distribution were also demonstrated (for transcriptional burst) [79], where bimodal and power law distributions occurred as the rate constants are varied over biologically significant time scales.

3.7.3.3 Separating Intrinsic from Extrinsic Fluctuations

As it was mentioned earlier, fluctuations in reactions leading to the production of a protein have two sources: (i) **intrinsic** noise is related to variations in protein levels

[10]The term "dynamic disorder", which comes from statistical mechanics, is nothing to do with the concept of "dynamical diseases", a sudden change in the qualitative dynamical behavior of a system due to some impairment of the physiological control system, say [114].

[11]The assumption behind the validity of the approximative formula is that the mRNA degradation is much faster than the proteolysis.



$$\frac{\mathrm{d}P(\mathrm{x},\mathrm{y}|\mathrm{z}(t))}{\mathrm{dt}} = -\sum_k r_k(\mathrm{x},\mathrm{z}(t))P(\mathrm{x},\mathrm{y}|\mathrm{z}(t)) + \sum_k r_k(\mathrm{x}-\mathrm{s}_k,\mathrm{z}(t))P(\mathrm{x}-\mathrm{s}_k,\mathrm{y}|\mathrm{z}(t))$$
$$-\sum_k r_k(\mathrm{y},\mathrm{z}(t))P(\mathrm{x},\mathrm{y}|\mathrm{z}(t)) + \sum_k r_k(\mathrm{y}-\mathrm{s}_k,\mathrm{z}(t))P(\mathrm{x},\mathrm{y}-\mathrm{s}_k|\mathrm{z}(t))$$

Fig. 3.15 The structure, dynamics and model of the dual reporter setup (From [72])

even in a population of cells with identical genotype and concentrations and states of cellular components, (ii) **extrinsic** noise due to fluctuations in the amount or activity of molecules involved in the expression of a gene, like RNA polymerase or ribosomes. A very successful interaction between the methods of new experimental techniques and of stochastic kinetics is related to the **dual reporter** method [53]. The idea was to incorporate a second independent system in the same environment, and observe both subsystems. Correlations between the subsystems reflect the influence of the common environment. Two identical and independent reporters embedded in a shared fluctuating environment can be used to identify intrinsic and extrinsic noise terms. The noise contributions identified by dual reporter methods correspond to the noise contributions predicted by correct stochastic models of either intrinsic or extrinsic mechanisms. It was found that "...the extrinsic noise from the dual-reporter method can be rigorously analyzed using models that ignore intrinsic stochasticity. In contrast, the intrinsic noise can be rigorously analyzed using models that ignore extrinsic stochasticity only under very special conditions that rarely hold in biology..." (Fig. 3.15).

It was claimed that the normalized covariance between the two reporter proteins can be used as a measure of extrinsic noise, and the remaining noise can be identified as intrinsic. Denoting by x and y the levels of the two reporter proteins, we have

$$\eta_{tot}^2 := \frac{\sigma_x^2}{< x >^2} = \eta_i nt^2 + \eta_{int}^2 + \eta_{ext}^2 \tag{3.130}$$

$$\eta_{ext}^2 = \frac{Cov(x, y)}{< x >< y >}, \tag{3.131}$$

where angled brackets denote means over the cell population. The scope and limits of the early suggestion were analyzed by [72, 73] permitting also time-dependent environment, and the whole field seems to be a very lively topic [161, 199].

3.8 Chiral Symmetry

Molecular chirality is associated with the lack of certain symmetry elements in the three dimensional structures of molecules. In effect, an object is chiral if it is not identical to its mirror image. This molecular asymmetry has extremely high biological relevance as the basic building blocks of biomolecules, amino acids and simple sugars are chiral and, therefore, biomacromolecules such as polysaccharides, peptides and nucleic acids are also chiral. In nature, the mirror image counterparts have very different roles: typically, only one of them is abundant and it cannot be exchanged with the other one. This phenomenon is called homochirality or biological chirality, the origins of which have been the subject of extensive theoretical speculations for about two centuries.

3.8.1 Racemic Mixtures

Statistical thermodynamics shows that in true chemical equilibrium, the distribution of the two mirror-image pairs of a chiral molecule (called enantiomers) is described by a binomial distribution [34–36, 99, 125, 171]. The enantiomers are labeled R and S by convention (an earlier notation is D and L). Allowing for a difference in the stability of the two enantiomers, this distribution can be given as follows:

$$P(r, s) = \binom{r+s}{r} (0.5 + \varepsilon)^r (0.5 - \varepsilon)^s \tag{3.132}$$

Here, $P(r,s)$ is the probability that r molecules of the R enantiomer and s molecules of the S enantiomer occur in an ensemble of $(r+s)$ molecules. Parameter ϵ is characteristic of the degree of inherent difference between the two enantiomers ($\epsilon \leq 0.5$), and is connected to the energy difference (ΔE) between the R and S molecules as follows:

$$\varepsilon = \frac{e^{\Delta E/RT} - 1}{2(e^{\Delta E/RT} + 1)} \tag{3.133}$$

The expectation and standard deviation for the number of R enantiomers from this distribution is given by straightforward formulae:

$$\langle r \rangle = (0.5 + \varepsilon)(r + s) \tag{3.134}$$

$$\sigma_r = \sqrt{(r + s)(0.5 + \varepsilon)(0.5 - \varepsilon)} \tag{3.135}$$

In practice, intensive parameters are preferred in the characterization of a mixture of enantiomers. One possibility is to use the molar fraction of the R enantiomer (x_R), another and even more popular option is using the enantiomeric excess (ee). Their definitions are given as follows:

$$x_{\mathrm{R}} = \frac{r}{r + s} \tag{3.136}$$

$$ee = \frac{|r - s|}{r + s} \tag{3.137}$$

The expectations and standard deviations of these parameters can be estimated as follows:

$$\langle x_{\mathrm{R}} \rangle = 0.5 + \varepsilon \tag{3.138}$$

$$\sigma_{x_{\mathrm{R}}} = \sqrt{\frac{(0.5 + \varepsilon)(0.5 - \varepsilon)}{(r + s)}} \tag{3.139}$$

$$\langle ee \rangle \geq \langle ee \rangle_{\varepsilon=0} = \frac{1}{r + s} \prod_{i=1}^{[(r+s-1)/2]} \frac{2i + 1}{2i} \tag{3.140}$$

$$\sigma_{ee} = \sqrt{\frac{1}{r + s} + 4\frac{r + s - 1}{r + s}\varepsilon^2 - \langle ee \rangle^2} \tag{3.141}$$

In most relevant works $\varepsilon = \Delta E = 0$ is assumed, which represents a sort of common sense and means the total symmetry of the two enantiomers. Nevertheless, there are numerous quantum chemical calculations based on the principle of parity violation and a consequent inherent asymmetry in atomic nuclei that predict a ΔE around 10^{-13} J mol^{-1} and an ε of 10^{-17} at room temperature. For 1 mol of chiral molecules ($r + s = 6 \times 10^{23}$) these values lead to an expected excess of the more stable enantiomer (6×10^6) that is much smaller than the standard deviation

(1.9×10^{11}), which would characterize average natural fluctuations. This small value of ΔE is still several orders of magnitude lower than the detection limits of the most advanced experimental methods, so its existence is at best an unconfirmed theoretical prediction. In addition, it has also been conclusively shown that such a small value of ϵ is without consequence when the mass of chiral material is smaller than the mass of planet Earth, and thus could not have any role in the emergence of biological chirality.

In chemical systems, the number of particles typically exceeds 10^{10}, often greatly so. Under such conditions, the binomial distribution for $\varepsilon = 0$ basically means that the two enantiomers are formed in equal amounts, and the fluctuations from this are undetectably small. This is called a racemic mixture in chemistry, and it is assumed that any synthesis method without the intervention of external chiral influences (such as chiral reagents, catalysts, mineral surfaces or circularly polarized electromagnetic radiation) will lead to a racemic product mixture.

Limited, but rather important experimental examples were published that do not adhere to this theoretical expectation [9, 10, 68, 101]. In these examples, macroscopically detectable fluctuations were observed in the enantiomeric excess of the product mixture in the absence of any external chiral influence. The most extensively studied such process is the Soai reaction [87, 173, 174, 176], which involves carbon-carbon bond formation in the reaction between an aldehyde with organometallic zinc reagents. The bromide ion induced dissociation of a trinuclear cobalt complex [6], and some chiral Mannich and aldol reactions [118] provide further example of this phenomenon, which is called absolute asymmetric synthesis or chiral symmetry breaking. The name absolute asymmetric synthesis originates from the fact that individual experiments in these processes may result in highly enriched enantiomeric excesses in products without any obvious reason. Symmetry breaking is a bit more confusing term, as it only means that non-racemic mixtures are formed in individual experiments. However, the statistical *ee* distributions are usually symmetric or close to symmetric. In this sense, these processes do not usually violate any symmetry laws. In any case, these observations are especially important because, although the compounds involved are by no means close to the abundant biomolecules (e.g. the conditions of the Soai reaction do not tolerate neither air nor water), they provide the only experimental information relevant for the formation of biological homochirality and provide direct observations in a field otherwise limited to theoretical speculations.

There are several requirements that must be fulfilled by a kinetic scheme to predict a non-binomial final distribution of the enantiomers formed. First, the direct transformation of enantiomers to each other must be vanishingly slow. This racemization reaction would lead to the binomial distribution no matter how the chiral material had been formed in previous processes. Second, the formation probability of enantiomers must be influenced by factors other than initially present (because they are completely non-chiral). The most obvious possibility is that the rates of reactions are influenced by the chiral molecules formed in the reaction. In chemical processes, such an effect is usually called autocatalysis (for an accelerating

effect) or autoinhibition (for a decelerating effect) as it is the product that influences the rate. In addition, this effect should be enantioselective, i.e. the R product formed must have different influence of the formation of R and S enantiomers. Such an interaction is not ruled out by symmetry laws. Both the phenomenon of autocatalysis and enantioselective catalysis are well known experimentally in chemistry, and a combination of these two, although may not be very common, is not impossible, either.

The possible role of enantioselective autocatalysis was probably first proposed by Frank on the basis of deterministic calculations which showed that miniscule deviations from chiral symmetry can be amplified by this mechanism [63]. Frank also coined the term mutual antagonism, which he primarily used for the interaction of R and S enantiomers that reduces the autocatalytic effect. The concepts of enantioselective autocatalysis and mutual antagonism have been invariably present in a high number of deterministic model calculations aimed at describing chiral amplification.

Stochastic theoretical considerations in this field were centered primarily on interpreting the experimental data available and this fact clearly determined the emphasis in the studies. Little effort has been devoted to describing the time dependence from any chemical models as no kinetic data have been published in the experimental examples. The studies were instead focused on predicting the final enantiomeric distributions, which is often conveniently done by using Q functions. Another characteristics of these theoretical attempts is the insistence of reaching predictions to chemically reasonable high numbers of particles (10^{19} in the case of the Soai reaction), sometimes even at the expense of mathematical precision.

3.8.2 Simple Enantioselective Autocatalysis

A particularly well studied example is a reaction in which the chiral product is formed from a nonchiral reagent A simultaneously in a direct and an autocatalytic pathway [97, 98, 100]:

$$\begin{aligned} &\mathrm{A} \rightarrow \mathrm{R} \\ &v = (0.5 + \epsilon)\kappa_u a + \kappa_c a r^{\xi} \\ &\mathrm{A} \rightarrow \mathrm{S} \\ &v = (0.5 - \epsilon)\kappa_u a + \kappa_c a s^{\xi} \end{aligned} \tag{3.142}$$

The most important case is when only A is present initially (with an initial molecules number of N_0). The overall number of possible states (M) can be given as:

$$M = \frac{(N_0 + 1)(N_0 + 2)}{2} \tag{3.143}$$

States can be identified solely by giving the numbers of R and S enantiomers (r and s, $r + s \leq N_0$) as the number of remaining A molecules ($a_0 - r - s$) can be calculated from conversion of mass. A suitable enumerating function can be given as follows:

$$f(r,s) = \frac{(r+s)(r+s+1)}{2} + r + 1 \tag{3.144}$$

The master equation of the process is the following:

$$\begin{aligned} \frac{dP_{r,s}(t)}{dt} &= -P_{r,s}(t)a\left[\kappa_u + \kappa_c r^{\xi} + \kappa_c s^{\xi}\right] \\ &+ P_{r-1,s}(t)a\left[(0.5+\varepsilon)\kappa_u + \kappa_c (r-1)^{\xi}\right] \\ &+ P_{r,s-1}(t)a\left[(0.5+\varepsilon)\kappa_u + \kappa_c (s-1)^{\xi}\right] \end{aligned} \tag{3.145}$$

For chemically first order autocatalysis ($\xi = 1$), a number of analytical formulae for the final state reached in this scheme could be derived. The use of the Q functions was rather advantageous for this problem. As the variables of interest are independent of time, only the ratio of the two rate constants ($\alpha = \kappa_c/\kappa_u$) appears as an important parameter:

$$Q(r,s) = \binom{r+s}{r} \frac{\prod_{j=0}^{r-1}(0.5+\varepsilon+\alpha j)\prod_{j=0}^{s-1}(0.5-\varepsilon+\alpha j)}{\prod_{j=0}^{r+s-1}(1+\alpha j)} \tag{3.146}$$

The expectation for x_{R} and its standard deviation can also be given:

$$\langle x_{\mathrm{R}} \rangle = 0.5 + \varepsilon \tag{3.147}$$

$$\sigma_{x_{\mathrm{R}}} = \sqrt{(0.5+\varepsilon)(0.5-\varepsilon)\frac{(r+s)^{-1}+\alpha}{1+\alpha}} \tag{3.148}$$

A formula for the expectation of ee can also be given:

$$\begin{aligned} \langle ee \rangle &= \sum_{i=0}^{[(N_0-1)/2]} \frac{N_0 - 2i}{N_0} \left(Q(N_0 - i, i) + Q(i, N_0 - i)\right) \\ &= 1 - \sum_{i=0}^{[(N_0-1)/2]} \frac{2i}{N_0} \left(Q(N_0 - i, i) + Q(i, N_0 - i)\right) \end{aligned} \tag{3.149}$$

Finally, it can be shown that the discrete distribution converges to a continuous beta distribution with parameters $(0.5 + \varepsilon)\alpha^{-1}$ and $(0.5 - \varepsilon)\alpha^{-1}$ as N_0 approaches infinity. The limiting continuous distribution function is:

$$P(x_R) = \frac{\Gamma\,(1/\alpha)}{\Gamma\,((0.5+\varepsilon)/\alpha)\,\Gamma\,((0.5-\varepsilon)/\alpha)} x_R^{(0.5+\varepsilon)/\alpha-1}(1-x_R)^{(0.5-\varepsilon)/\alpha-1} \quad (3.150)$$

For $\varepsilon = 0$, the continuous distribution with no initial chiral material present reduces to a symmetric beta distribution. In addition, the description has been extended to initial conditions where some R and/or S molecules are present in the initial mixture [170]:

$$Q(r,s) = \binom{r+s-r_0-s_0}{r-r_0} \frac{\prod_{j=r_0}^{r-1}(0.5+\alpha j)\prod_{j=s_0}^{s-1}(0.5+\alpha j)}{\prod_{j=r_0+s_0}^{r+s-1}(1+\alpha j)} \quad (3.151)$$

The expectation for the enantiomeric excess and its standard deviation have been reported as:

$$\langle ee \rangle = \frac{(N_0 + 1/\alpha)\,|r_0 - s_0|}{N_0(r_0 + s_0 + 1/\alpha)} \quad (3.152)$$

$$\sigma_{ee} = \frac{2}{N_0}\sqrt{\frac{(N_0+1/\alpha)(N_0-r_0-s_0)(r_0+0.5/\alpha)(s_0+0.5/\alpha)}{(r_0+s_0+1/\alpha+1)(r_0+s_0+1/\alpha)^2}} \quad (3.153)$$

For higher-order autocatalysis ($\xi > 1$), the probability of getting one enantiomer only, $Q(r, 0)$, can be given, which confirms that one of the enantiomers will be in overwhelming excess over the other for $\xi > 1$ and sufficiently large values of N_0. For $\varepsilon > 0$, the following expression can also be proved

$$\frac{Q(b,0)}{Q(0,b)} < e^{4\varepsilon/(0.5-\varepsilon)+2(\xi-1)\varepsilon/\alpha} \quad (3.154)$$

However, higher-order autocatalysis under reasonable conditions does not necessarily lead to the formation enantiomeric excesses close to 100% because the convergence to the unique final distribution, unlike in the case of first-order autocatalysis, may not be very fast. Numerical calculations can be carried out using the method of deterministic continuation. First, the appropriate values of $Q(r, s)$ are recursively calculated with the CDS approach for a relatively small value of $(r + s)$, e.g. 1,000 or 10,000 to obtain a discrete distribution function can be obtained in this way. The distribution of the molar fractions in the final state can then be calculated by the numerical integration of the following, deterministic differential equation:

$$-\frac{dx_R}{da} = \frac{0.5+\varepsilon+\alpha r^\xi}{(1+\alpha r^\xi+\alpha r^\xi)(N_0-a+1)} - \frac{x_R}{N_0-a+1} \quad (3.155)$$

This approach gives a discrete approximation of the final distribution the number of points in which is identical to the $(r + s)$ value up to which the CDS approach was used. This method was used to prove that second order autocatalysis can give a prediction for the distribution of enantiomeric excesses that fits the experimental observations better than the results following from first order autocatalysis.

A very similar system with a somewhat more complicated rate equation and elements of reversibility (termed recycling there) was also analyzed in some detail [162, 183]. The scheme is given as follows:

$$\begin{aligned} &\mathrm{A} \rightarrow \mathrm{R} \\ &v_1 = 0.5\kappa_0 a + \kappa_1 a r + \kappa_2 a r^2 - \lambda r \\ &\mathrm{A} \rightarrow \mathrm{S} \\ &v_2 = 0.5\kappa_0 a + \kappa_1 a s + \kappa_2 a s^2 - \lambda s \end{aligned} \tag{3.156}$$

The master equation corresponding to this scheme is as follows:

$$\begin{aligned} \frac{dP_{r,s}(t)}{dt} = &-P_{r,s}(t)(\kappa_0 a + \kappa_1 a r + \kappa_1 a s + \kappa_2 a r^2 + \kappa_2 a s^2 + \lambda r + \lambda s) \\ &+P_{r-1,s}(t)\left[0.5\kappa_0(a+1) + \kappa_1(a+1)(r-1) + \kappa_2(a+1)(r-1)^2\right] \\ &+P_{r,s-1}(t)\left[0.5\kappa_0(a+1) + \kappa_1(a+1)(s-1) + \kappa_2(a+1)(s-1)^2\right] \\ &+P_{r+1,s}(t)\lambda(r+1) + P_{r,s+1}(t)\lambda(s+1) \end{aligned} \tag{3.157}$$

As the processes are reversible in this system, none of the states is final and Q functions carry no meaning. The $\Pi_{r,s}$ stationary probabilities should be defined as the final, time-independent, stationary probability values for each state:

$$\Pi_{r,s} = \lim_{t \to \infty} P_{r,s}(t) \tag{3.158}$$

The stationary absolute state probability values can be given as:

$$\begin{aligned} &\Pi_{r,s} = \Xi \frac{a_0!}{r!s!(N_0-r-s)!} \lambda^{N_0-r-s} \\ &\prod_{j=0}^{r-1}(0.5\kappa_0 + \kappa_1 j + \kappa_2 j^2) \prod_{j=0}^{s-1}(0.5\kappa_0 + \kappa_1 j + \kappa_2 j^2) \end{aligned} \tag{3.159}$$

Here, Ξ is a normalizing constant and the appropriate products shown should simply be replaced by a number 1 if $r = 0$ or $s = 0$. Continuous approximations of the final distributions have also been given.

Finally, a random walk CDS approach was introduced on a square lattice (r, s) in a triangular region $0 \le r$, $0 \le r$, $r + s \le N_0$ for studying the system without recycling ($\lambda = 0$). The random walker can only make directed walks to the R and S directions: the possible steps are $(r, s) \to (r + 1, s)$ and $(r, s) \to (r, s + 1)$. Based on the transition rates in the master equation, the walker on site (r, s) stays on the site for the following waiting time:

$$\tau(r, s) = \frac{1}{[\kappa_0 + \kappa_1(r + s) + \kappa_2(r^2 + s^2)](N_0 - r - s)} \tag{3.160}$$

After the waiting time of τ, the jump is in either R or S direction with the respective probabilities of P_{R} and P_{S}:

$$\begin{aligned} P_{\mathrm{R}}(r, s) &= \frac{0.5\kappa_0 + \kappa_1 r + \kappa_2 r^2}{\kappa_0 + \kappa_1(r + s) + \kappa_2(r^2 + s^2} \\ P_{\mathrm{S}}(r, s) &= \frac{0.5\kappa_0 + \kappa_1 s + \kappa_2 s^2}{\kappa_0 + \kappa_1(r + s) + \kappa_2(r^2 + s^2)} \end{aligned} \tag{3.161}$$

In fact, this is identical to the Monte Carlo simulation approach to stochastic kinetics, which is also the basis of the Gillespie algorithm.

3.8.3 The Frank Model

The scheme initially originally proposed by Frank [63] has also been analyzed by the CDS approach [105]:

$$\begin{aligned} &\mathrm{A} \to \mathrm{R} \quad v_1 = 0.5\kappa_u a + \kappa_c a r^{\xi} \\ &\mathrm{A} \to \mathrm{S} \quad v_2 = 0.5\kappa_u a + \kappa_c a s^{\xi} \\ &\mathrm{R} + \mathrm{S} \to 2\mathrm{C} \quad v_3 = \kappa_d r s \end{aligned} \tag{3.162}$$

In addition to enantioselective autocatalysis, mutual antagonism in the third step is also present in this mechanism.

The overall number of possible states M is a third-order polynomial of N_0 in this system:

$$M = \binom{N_0 + 3}{3} = \frac{(N_0 + 3)(N_0 + 2)(N_0 + 1)}{6} \tag{3.163}$$

The identification of individual states requires three variables here, which are conveniently selected as a, r, and s. A suitable enumerating function is given by the following formula:

$$f(a,r,s) = \frac{a^3}{6} + \frac{aN_0(a+N_0)}{2} - a^2 + 2aN_0 + \frac{11a}{6} + \frac{r^2}{2} + r(N_0 - a) + \frac{3r}{2} + s + 1 \tag{3.164}$$

The analysis was done in a way that also allows some inflow of reactant A and an outflow of the reaction mixture. The master equation can be written as:

$$\begin{aligned}\frac{dP_{a,r,s}(t)}{dt} = &-\{a\kappa_u + ar\kappa_c + as\kappa_c + rs\kappa_d + (N_0 - a)\kappa_f\}P_{a,r,s}(t)\\ &+\{0.5(a+1)\kappa_u + (a+1)(r-1)\kappa_c\}P_{a+1,r-1,s}(t)\\ &+\{0.5(a+1)\kappa_u + (a+1)(s-1)\kappa_c\}P_{a+1,r,s-1}(t)\\ &+\{(r+1)(s+1)\kappa_d\}P_{a,r+1,s+1}(t) + (r+1)\kappa_f P_{a-1,r+1,s}(t)\\ &+(s+1)\kappa_f P_{a-1,r,s+1}(t) + (N_0 - a - r - s + 1)\kappa_f P_{a-1,r,s}(t)\end{aligned} \tag{3.165}$$

Parameter κ_f (dimension: inverse time) here corresponds to the flow rate, the equation itself also describes a closed system without flow if $\kappa_f = 0$ is set. The detailed analysis questioned the positive role that was assumed to be played by mutual antagonism in creating high enantiomeric excesses. In fact, its effect in decreasing the overall amount of chiral material in a reactor seems to be more important than the increase in enantiomeric excess values.

3.8.4 The Soai Reaction

The models described in the previous sections all share the drawback that they are chemically unreasonable in their simplicity. The mechanism of the Soai reaction is known to be much more complex. Schemes like these are not easily processed within the CDS approach primarily because of the high number of states involved. A particularly noted series of chemical reactions was proposed by Buhse to interpret experimental findings [33]:

$$
\begin{aligned}
&\mathrm{CHO} + \mathrm{Zn} \xrightarrow{k_1} \text{(R)-COZn}\\
&\mathrm{CHO} + \mathrm{Zn} \xrightarrow{k_1} \text{(S)-COZn}\\
&\text{(R)-COZn} + \text{(R)-COZn} \xrightarrow{k_2} \text{(R)(R)-(COZn)}_2\\
&\text{(R)(R)-(COZn)}_2 \xrightarrow{k_{-2}} \text{(R)-COZn} + \text{(R)-COZn}\\
&\text{(S)-COZn} + \text{(S)-COZn} \xrightarrow{k_2} \text{(S)(S)-(COZn)}_2\\
&\text{(S)(S)-(COZn)}_2 \xrightarrow{k_{-2}} \text{(S)-COZn} + \text{(S)-COZn}\\
&\text{(R)-COZn} + \text{(S)-COZn} \xrightarrow{\alpha k_2} \text{(R)(S)-(COZn)}_2\\
&\text{(R)(S)-(COZn)}_2 \xrightarrow{k_{-2}} \text{(R)-COZn} + \text{(S)-COZn}\\
&\text{(R)(R)-(COZn)}_2\text{+ CHO} \xrightarrow{k_3} \text{(R)(R)-(COZn)}_2\text{-CHO}\\
&\text{(R)(R)-(COZn)}_2\text{-CHO} \xrightarrow{k_{-3}} \text{(R)(R)-(COZn)}_2\text{+ CHO}\\
&\text{(S)(S)-(COZn)}_2\text{+ CHO} \xrightarrow{k_3} \text{(S)(S)-(COZn)}_2\text{-CHO}\\
&\text{(S)(S)-(COZn)}_2\text{-CHO} \xrightarrow{k_{-3}} \text{(S)(S)-(COZn)}_2\text{+ CHO}\\
&\text{(R)(S)-(COZn)}_2\text{+ CHO} \xrightarrow{k_3} \text{(R)(S)-(COZn)}_2\text{-CHO}\\
&\text{(R)(S)-(COZn)}_2\text{-CHO} \xrightarrow{k_{-3}} \text{(R)(S)-(COZn)}_2\text{+ CHO}\\
&\text{(R)(R)-(COZn)}_2\text{-CHO} + \mathrm{Zn} \xrightarrow{k_4} \text{(R)(R)-(COZn)}_2\text{+ (R)-COZn}\\
&\text{(S)(S)-(COZn)}_2\text{-CHO} + \mathrm{Zn} \xrightarrow{k_4} \text{(S)(S)-(COZn)}_2\text{+ (S)-COZn}\\
&\text{(R)(S)-(COZn)}_2\text{-CHO} + \mathrm{Zn} \xrightarrow{k_4} \text{(R)(S)-(COZn)}_2\text{+ (S)-COZn}\\
&\text{(R)(S)-(COZn)}_2\text{-CHO} + \mathrm{Zn} \xrightarrow{k_4} \text{(R)(S)-(COZn)}_2\text{+ (R)-COZn}
\end{aligned}
\tag{3.166}
$$

This sequence of reaction comprises four different sorts of reactions, but because of some reversibility and the existence of enantiomers, the full model contains 18 individual steps. Because of symmetry reasons, the actual number of parameters (rate constants) is only seven ($k_1, k_2, k_{-2}, k_3, k_{-3}, k_4$, and α). Final enantiomeric distributions predicted by the model were calculated by combining several computation-accelerating techniques: Monte Carlo simulations in the beginning using the stochastic analog of the steady state approximation, then deterministic continuation. A technique called symmetrization was also introduced. Symmetry ensures that the same enantiomeric excess must be formed with the same probability for both R and S enantiomers. The Monte Carlo simulation converges toward this

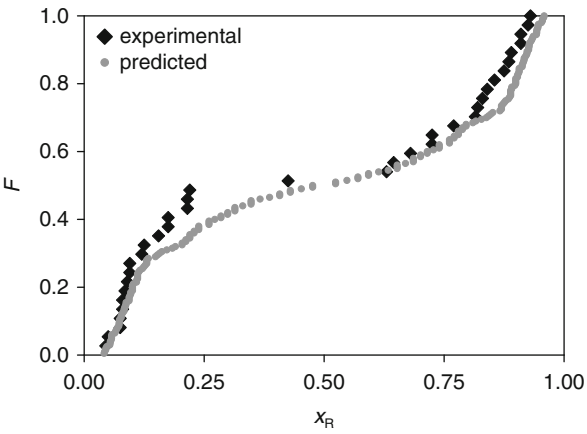


Fig. 3.16 Experimentally observed and theoretically predicted probability distributions of enantiomeric excesses in the Soai reaction. Experimental data from Soai et al. [176] (Adapted from [47])

symmetry in a probabilistic fashion after a high number of repetitions (convergence is roughly proportional to the square root of the number of repetitions done). Symmetrization is a method to force this symmetry into the calculation results: whenever a particular repetition gave a certain final enantiomeric excess value, this was interpreted as two different repetitions giving the same enantiomeric excess values, one favor of the R, the other for the S enantiomer. This technique made sure that the predicted final distribution shows the required symmetry independently of the number of repetitions carried out. With the combination of these techniques, successful prediction for the distribution of enantiomeric excesses was made based on the presented 18-step model [47]. With a suitable selection of parameter values, the experimental observations of the absolute asymmetric Soai reaction [176] could be fully interpreted as shown in Fig. 3.16.

3.9 Parameter Estimation in Stochastic Kinetic Models

3.9.1 Estimation of Rate Constants from Equilibrium Fluctuations

As we saw in Sect. 1.3.1.2, fluctuations around equilibrium are interconnected to the dissipative relaxation process to equilibrium. Fluctuations may be interpreted as spontaneous perturbations and relaxations, so in principle, they can be used to obtain kinetic information without perturbing the system externally, so they seemed to be advantageous comparing to the celebrated relaxation method [52]. Equilibrium concentration (better saying, number of particles) fluctuations served as a source of information by using methods of light scattering [19, 20], and conductance measurements [59]. For the association-dissociation reaction of the

beryllium sulfate described by the reaction $X \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} Y$, the spectrum of the electric fluctuation (which reflects the concentration fluctuation) was found to be

$$S\nu = \frac{const}{1 + (2\pi\nu(k_1 + k_2))^2}. \tag{3.167}$$

Since from deterministic equilibrium the ratio $\frac{k_1}{k_{-1}}$ is given; therefore, the individual rate constants can be calculated.

Electrical noise in biological membrane systems emerges in consequence of opening and closing of ion channels and **membrane noise analysis** was used to estimate the stoichiometric and kinetic details and discriminate among concurrent transport mechanisms [45, 55, 64, 74, 86]. Neher and Sakmann [136], by developing the patch clamp technique, were able to record small ionic currents that flow through a *single ion channel* in neuronal membranes, so they could obtain more direct information on the kinetics of channel gating.

Fluorescence spectroscopy proved to be the main experimental technique to study fluctuations, as a source of information. Single point measurements can investigate small volumes containing only few molecules. **Fluorescence correlation spectroscopy** is able to measure the fluctuation of the concentration of fluorescent particles (molecules). Temporal changes in the fluorescence emission intensity caused by single fluorophores are recorded. The autocorrelation function $C(t) := E[\xi(t)\xi(t-\tau)]$ of the signal $\xi(t)$ is calculated, and from their time-dependent decay of the fluorescence intensity, the rate parameters can be calculated. Higher order correlations $C_{mn}(t) := E[\xi(t)^m \xi(t-\tau)^n]$ [140, 150] were used to study the details of molecular aggregation. To extract more information from the available data beyond average and variance, at least two efficient methods were suggested. **Fluorescence-intensity distribution analysis** [85] is able to calculate the expected distribution of the number of photon counts, and the **photon counting histogram** [130] gives an account of the spatial brightness function. Forty years after, fluorescence fluctuation spectroscopy still is a developing method [186], for a very readable short review, see [144].

3.9.2 Parameter Estimation for Stochastic Kinetic Models: Beyond the Fluctuation-Dissipation Theorem

While the fluctuation-dissipation theorem, as a relationship between fluctuations around the equilibrium state and the dissipative relaxation process leading to it, offers a method to obtain information on the rate constants, a more general method, i.e. the **parameter estimation of stochastic processes** exploits information from time-dependent data. Interestingly, during the late 1960s and early 1970s, when

the first reviews on stochastic kinetics were written, only a very few and rather unnoticed paper dealt with the parameter estimation problems of stochastic models of chemical reactions.[12]

In the rather unnoticed papers of Mulloolly [131–133], Bartlett's techniques were applied to study the maximum likelihood estimator for the rate constant of the first order, and rth order stochastic irreversible reactions. Among others, an approximate expression for the maximum likelihood estimator of the rate constant for the stochastic rth order reaction was derived. For the reaction $rX \xrightarrow{k} Y$ starting from a molecules, and assuming error-free measurements, the asymptotic form for $a \to \infty$

$$\hat{k} = (1/t)\frac{r!}{r^2 a^{r-1}} ln \left[\frac{a - \frac{1}{2}r(r-1)}{a - \frac{1}{2}r((2n+1)r-1)} \right] \tag{3.168}$$

was derived.

Since the kinetic parameters of biochemical reactions, such as gene regulatory, signal transduction and metabolic network, generally cannot measured directly, parameter estimation techniques are developed and applied much more extensively [27, 147, 155, 187]. A number of methods, based on maximum likelihood, density function distance, cumulative density function and global optimization algorithms were adopted. Parameter estimation techniques minimize the distance measures between model predictions and the experimental data, while global optimization algorithms search for the global optima to the minimization problems.

The majority of applications are based on some version of the maximum likelihood method, and estimation criterion is the likelihood function given by

$$L(\mathbf{k}) = \prod_{j=1}^{m} \prod_{i=1}^{n} f(\mathbf{o}_i^j, \mathbf{t}_i; \mathbf{k}), \tag{3.169}$$

where the jth experimental replicates $\mathbf{o}_1^j, \mathbf{o}_2^j, \ldots \mathbf{o}_n^j$ are taken at time points t_1, t_2, $\ldots t_n$ for j = 1, 2, …, m (i.e. the experiments are done in m replicates). $f(\mathbf{o}_i^j, \mathbf{t}_i; \mathbf{k}$ is the likelihood function determined by the density function histogram constructed from the realizations of the stochastic process specified by the master equation.

The maximization of the likelihood function (actually for numerical reasons, the minimization of the negative log-likelihood function) gives the best estimated

[12]The techniques of statistical parameter estimation based on the method of **maximum likelihood** initiated by Ronald Fisher [61] was extended to stochastic (mostly for Gaussian and Markovian) process among others by Ulf Grenander [67], Maurice Bartlett [15] and Patrick Billingsley [21]. The two-volume monograph by Liptser and Shiryayev (English translation: 1977, 1978 [110,111]) describes developments in sequential estimation problems for stochastic processes. For more history see [16, 92].

parameters (i.e. it gives the greatest possible probability to the given data set): training data.

$$k^* = \underset{k}{\text{argmin}} - logL(\mathbf{k}) = \underset{k}{\text{argmin}} \sum_{j=1}^{m} \sum_{i-1}^{n} -\log P(o_i^j, t_i), \tag{3.170}$$

where $P(o_i^j, t_i)$ is the conditional probability density function reconstructed from the simulated realizations.

Most recently, there is a extensive progress in the application of parameter estimation techniques for stochastic models of biochemical reactions, and a wide range of optimization methods (such as derivative based methods, global optimization and Bayesian techniques) were applied, and we cannot attempt to review here the newer development. Probably it would be timely to write a monograph, and give a systematic analysis of the methods and applications.

3.10 Stochastic Resonance in Chemical Systems

3.10.1 General Remarks

Broadly speaking, stochastic resonance is a mechanism by which a system in a noisy environment shows an enhanced sensitivity towards small external time-dependent forcing, when the noise intensity reaches some threshold. To put in another way, the performance of a system is improved by adding some noise to it. According to the original, narrower definition, SR considered only systems where the input was a combination of a periodic single frequency signal and broadband noise. In such systems, a natural measure of system performance is the output signal-to-noise ratio (SNR) [120]. In linear systems, noise never has a "beneficial" effect, for certain combinations of nonlinearity and noise, the response of the system undergoes resonance-like behavior as a function of the noise level; hence the name stochastic resonance, see Fig. 3.17.

The curve has a similar form for frequency-dependent systems. However, frequency-dependent systems the signal/noise ratios shows maximum for a resonant frequency, here the resonance is "noise-induced".

3.10.2 Stochastic Resonance in One- and Multi-parameter System

Stochastic resonance seems to occur everywhere in nature from lasers and semiconductors, via ion channels and sensory systems via climate. Here, of

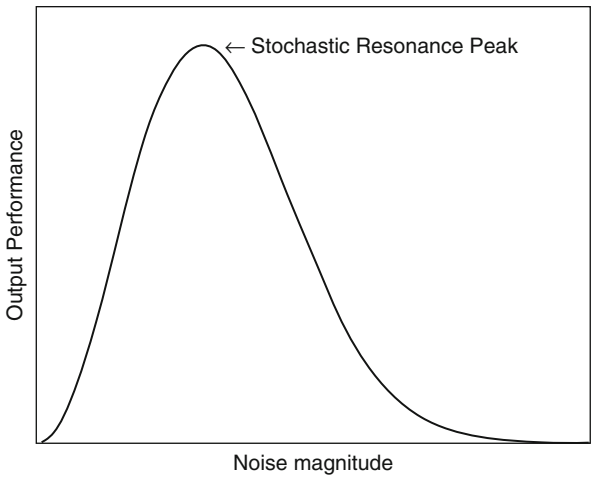


Fig. 3.17 A typical curve of stochastic resonance shows a single maximum of the output performance as a function of the intensity of the input noise

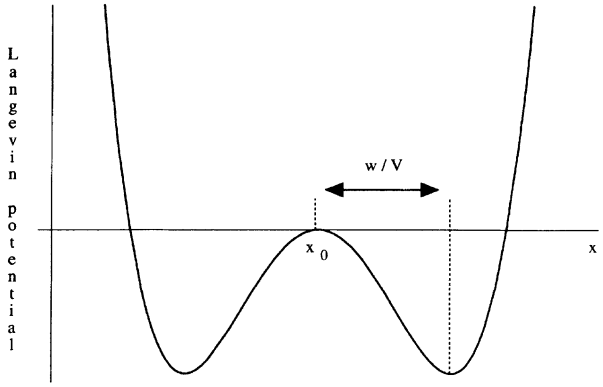


Fig. 3.18 The Langevin potential in its symmetric configuration. $x = b/V$ is the intensive variable (From [107])

course, stochastic resonance in chemical systems is reviewed. Of course, as we know, chemical noise can be generated **internally**, and it may interact with nonlinearity. Specifically, the cubic Schlögl reaction was supplemented by the periodic perturbation of a the external component C [107]. A **Langevin potential** is defined, and as Fig. 3.18 shows, it is a quartic, double-well potential. For the symmetric case ω/V measures the distance from the central barrier at x_0; here ω is the potential width and V is the volume. The periodic forcing γ induces the asymmetric rise and fall of the two minima.

The numerical solution of the assigned master equation results probability distributions with different shape by changing w. For intermediate w, there is a periodic transition between two extreme states. Figure 3.19 shows a detailed picture of the SNR, as the characteristic feature of stochastic resonance.

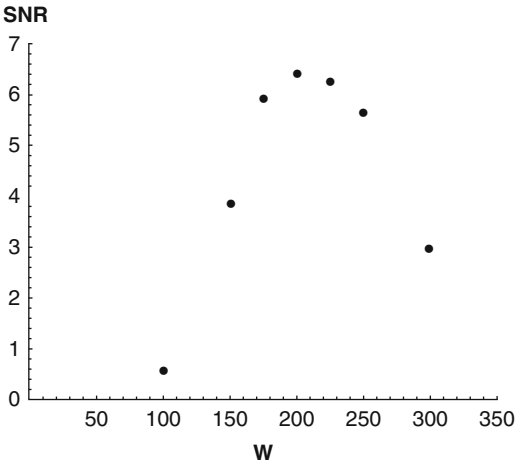


Fig. 3.19 Signal-to-noise ratio as a function of the width of the double well. Since width can be identified as the noise parameter, the peak is the feature of stochastic resonance (From [107])

Motivated by the experimental behavior of the photosensitive BZ reaction, [1] studied two-parameter stochastic resonance. It obviously shows characteristic features not observed in one-parameter systems, as it was demonstrated, see Fig. 3.20.

3.10.3 Stochastic Resonance of Aperiodic Signals

Not only periodic, but aperiodic signals might be the subject of amplification by noise, as it was demonstrated by [141] both in experimental (actually electrochemical) and model studies. Information transfer is quantified by the C_0 cross correlation function

$$C_0 = \langle (x_1 - \langle x_1 \rangle_t)(x_2 - \langle x_2 \rangle_t) \rangle_t, \tag{3.171}$$

where x_1 and x_2 represent the time series of the aperiodic input signal, and the noise induced response of the electrochemical system, respectively. $\langle\rangle$ denotes the respective time averages.

Figure 3.21 illustrates the existence of optimal noise level for information transfer.

3.11 Computation with Small Stochastic Kinetic Systems

While the main field of application of stochastic chemical kinetics is mostly related to systems biology, it emerged also as a possible model of computation in which information is stored and processed in the integer counts of molecules in a well-mixed solution [40, 181].

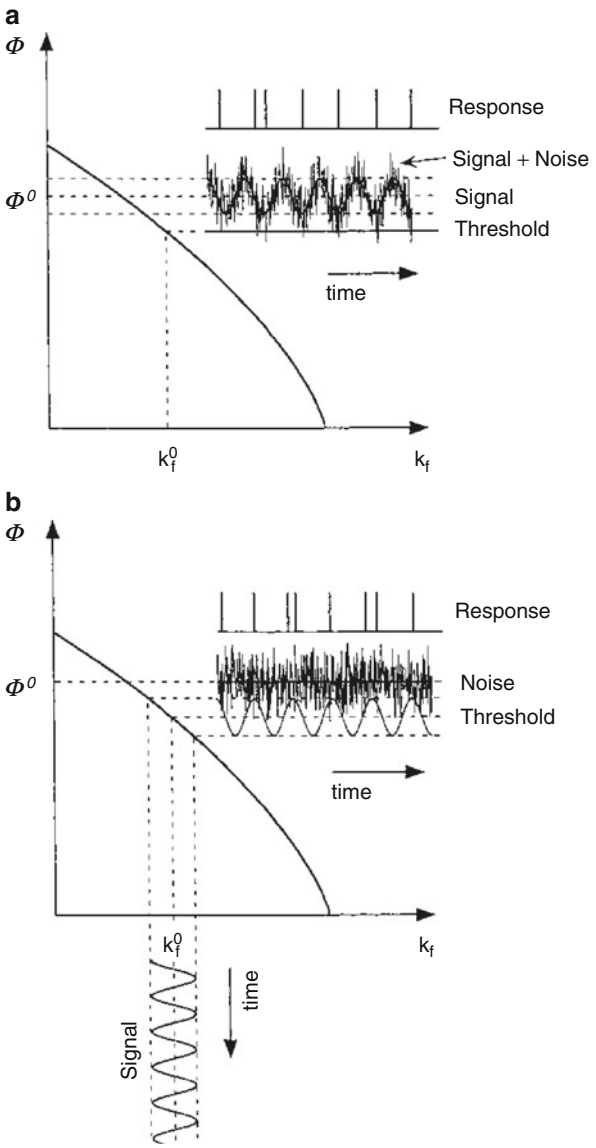


Fig. 3.20 (**a**) One-parameter system: a threshold is constant in time and the system generates a response if the sum of the signal and additive noise exceeds the threshold. (**b**) Two-parameter system: a periodic modulation in one parameter changes a threshold of the other parameter, and the system generates response when noise in the latter parameter exceeds the modulated threshold (From [1])

The applicability of chemical systems to serve as computing elements have been suggested in several different contexts. The realizability of logical functions by chemical bistable systems was suggested by Otto Rössler [160]. The combination of the fundamental logical elements can lead to complex logical networks. Chemical reactions may implement efficient computing devices [115]. However, computation with logic circuits would require more and more species – need a finite machine that

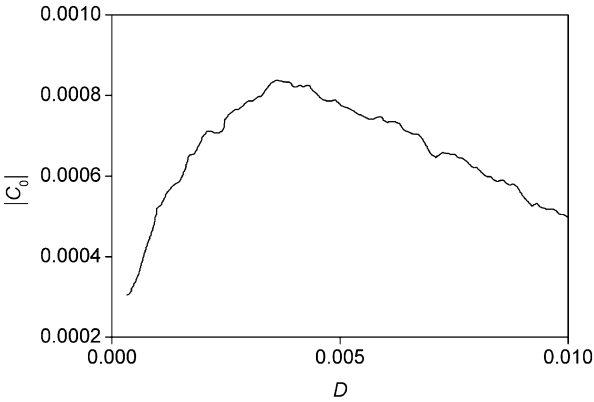


Fig. 3.21 Cross-correlation as a function of noise level (From [141])

can do infinite computation. It is a model of the so-called *non-uniform* computation: a different "program" is necessary for each input size.

Particularly, there were two questions to be answered. First, it turned out that the **reachability** question – i.e. whether state B can be reached from state A is **decidable**. The problem was formulated in a specific field of computer science called Vector Addition Systems (VASs) [84]. Second, the **probability** of the transition from A to B is undecidable.

A computation is uniform if the same program runs for all (the infinitely many) inputs. It was shown that a canonic model of computation, i.e. a *Register Machine* is a simple uniform model of computation. An algorithm was given [181] to show how a stochastic chemical reaction network implement (or simulate) Register Machine (a very simple universal Turing Machine) efficiently within a known error bound.[13]

References

1. Ameyima T, Ohmori T, Nakaiwa M, Yamaguchi T (1998) Two-parameter stochastic resonance in a model of the photosensitive belousov-zhabotinsky reaction in a flow system. J Phys Chem A 102:4537–4542
2. Arányi P, Tóth J (1977) A full stochastic description of the Michaelis–Menten reaction for small systems. Acta Biochimica et Biophysica Academiae Scientificarum Hungariae 12:375–388
3. Arkin A (2000) Signal processing by biochemical reaction networks. In: Walleczek J (ed) Self-organized biological dynamics and nonlinear control. Toward understanding complexity, chaos and emergent function in living systems, pp 112–144

[13]The problem of reachability in stochastic chemical reaction networks is related to *Primitive Recursive Functions* studied by the legendary Hungarian mathematician Rózsa Péter ("Aunt Rózsa"). The relationship between primitive recursive functions and stochastic chemical networks was also shown [40].

4. Arkin A, Ross J, McAdams HH (1998) Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected Escherichia coli cells. Genetics 149(4):1633–48
5. Arnold L (1980) On the consistency of the mathematical models of chemical reactions. In: Haken H (ed.) Dynamics of synergetic systems. Springer series in synergetics, vol 6. Springer, Berlin, pp 107–118
6. Asakura K, Ikumo A, Kurihara K, Osanai S, Kondepudi DK (2000) Random chiral asymmetry generation by chiral autocatalysis in a far-from-equilibrium reaction system. J Phys Chem A 104:2689–2694
7. Bak P, Tang C, Wiesenfeld K (1988) Self-organized criticality. Phys Rev A 38(1):364–374
8. Bankó S, Kucsma Z, Lente G, Bagyinka C (2013) The autocatalytic step is an integral part of the hydrogenase enzyme-cycle. Biochim Biophys Acta 1834:658–664
9. Barabás B, Caglioti L, Zucchi C, Maioli M, Gál E, Micskei K, Pályi G (2007) Violation of distribution symmetry in statistical evaluation of absolute enantioselective synthesis. J Phys Chem B 111:11506–11510
10. Barabás B, Caglioti L, Micskei K, Pályi G (2009) Data-based stochastic approach to absolute asymmetric synthesis by autocatalysis. Bull Chem Soc Jpn 82:1372–1376
11. Bartholomay AF (1958) Stochastic models for chemical reactions: I. Theory of the unimolecular reaction process. Math Biophys 20:175–190
12. Bartholomay AF (1959) Stochastic models for chemical reactions: II. The unimolecular rate constant. Math Biophys 21:363–373
13. Bartholomay AF (1962) A stochastic approach to statistical kinetics with application to enzyme kinetics. Biochem 1:223–230
14. Bartholomay AF (1962) Enzymatic reaction-rate theory: a stochastic approach. Ann N Y Acad Sci 96:897–912
15. Bartlett MS (1955) An introduction to stochastic processes. Cambridge University Press, Cambridge
16. Basawa IV, Prakasa Rao LS (1980) Asymptotic inference for stochastic processes. Stoch Proc Appl 9:291–305
17. Basu M, Mohanty PK (2009) Stochastic modeling of single molecule Michaelis Menten kinetics. arXiv:0901.2844
18. Berg OG (1978) A model for the statistical fluctuations of protein numbers in a microbial population. J Theor Biol 71:587–603
19. Berne BJ, Pecora R (1976) Dynamic light scattering. Dover, New York
20. Berne BJ, Deutch JM, Hynes JT, Frisch HL (1968) Light scattering from chemically reacting mixtures. J Chem Phys 49:2864–2865
21. Billingsley P (1961) Statistical inference for Markov processes. University of Chicago Press, Chicago
22. Bishop LM, Qian H (2010) Stochastic bistability and bifurcation in a mesoscopic signaling system with autocatalytic kinase. Biophys J 98:1–11
23. Blake WJ, Kaern M, Cantor CR, Collins JJ (2003) Noise in eukaryotic gene expression. Nature 422:633–637
24. Bowsher CG (2010) Stochastic kinetic models: dynamic independence, modularity and graphs. Ann Stat 38:2242–2281
25. Bowsher CG (2010) Information processing by biochemical networks: a dynamic approach. J R Soc Interface 8:186–200
26. Bowsher CG, Swain P (2012) Identifying sources of variation and the flow of information in biochemical networks. Proc Natl Acad Sci USA 109(20):E1320–E1328
27. Boys R, Wilkinson D, Kirkwood T (2008) Bayesian inference for a discretely observed stochastic kinetic model. Stat Comput 18:125–135
28. Bratsun D, Volfson D, Hasty J, Tsimring LS (2005) APS meeting. http://inls.ucsd.edu/~volfson/pub/nonmark_b.pdf
29. Bratsun D, Volfson D, Tsimring LS, Hasty J (2005) Delay-induced stochastic oscillations in gene regulation. Proc Natl Acad Sci USA 102(41):14593–14598

30. Briggs GE, Haldane JBS (1925) A note on the kinematics of enzyme action. Biochem J 19:338–339
31. Brown FLH (2003) Single-molecule kinetics with time-dependent rates: a generating function approach. Phys Rev Lett 90:028302
32. Brunel N, Nadal JP (1998) Mutual information, Fisher information and population coding. Neural Comput 10:1731–1757
33. Buhse T (2003) A tentative kinetic model for chiral amplification in autocatalytic alkylzinc additions. Tetrahedron Asymmetry 14:1055–1061
34. Caglioti L, Pályi G (2008) Chiral chemistry of single molecules. Chimica Oggi 26:41–42
35. Caglioti L, Zucchi C, Pályi G (2005) Single-molecule chirality. Chimica Oggi 23:38–43
36. Caglioti L, Hajdu C, Holczknecht O, Zékány L, Zucchi C, Micskei K, Pályi G (2006) The concept of racemates and the soai-reaction. Viva Orig 34:62–80
37. Cai L, Friedman N, Xie XS (2006) Stochastic protein expression in individual cells at the single molecule level. Nature 440:358–362
38. Choi PJ, Cai L, Frieda K, Xie XS (2008) A stochastic single-molecule event triggers phenotype switching of a bacterial cell. Science 322(5900):442–446
39. Cianci C, Di Patti F, Fanelli D, Barletti L (2012) Analytical study of non Gaussian fluctuations in a stochastic scheme of autocatalytic reactions. Eur Phys J Spec Top 212:5–22
40. Cook M, Soloveichik D, Winfree E, Bruck Y (2009) Programmability of chemical reaction networks. In: Condon A, Harel D, Kok JN, Salomaa A, Winfree E (eds) Algorithmic bioprocesses. Springer, Berlin/Heidelberg, pp 543–584
41. Dahmen W, Micchelli CA (1988) The number of solutions to linear Diophantine equations and multivariate splines. Trans Am Math Soc 308:509–532
42. Dambrine S, Moreau M (1981) On the stationary distribution of a stochastic chemical process without detailed balance. J Stat Phys 269:137–148
43. Darvey IG, Staff PJ (1966) Stochastic approach to first-order chemical reaction kinetics. J Chem Phys 44:990–997
44. Delbrück M (1940) Statistical fluctuation in autocatalytic reactions. J Chem Phys 8:120–124
45. der Chen Y (1978) Noise analysis of kinetic systems and its applications to membrane channels. Adv Chem Phys 37:67–97
46. de Ronde WH, Tostevin F, ten Wolde PR (2010) Effect of feedback on the fidelity of information transmission of time-varying signals. Phys Rev E 82:031914
47. Dóka E, Lente G (2011) Mechanism-based chemical understanding of chiral symmetry breaking in the Soai reaction. A combined probabilistic and deterministic description of chemical reactions. J Am Chem Soc 133:17878–17881
48. Dóka E, Lente G (2012) Stochastic mapping of the Michaelis–Menten mechanism. J Chem Phys 136:054111
49. Drummond PD, Vaughan TG, Drummond AJ (2010) Extinction times in autocatalytic systems. J Phys Chem A 114:10481–10491
50. Ebeling W, Schimansky-Geier L (1979) Stochastic dynamics of a bistable reaction system. Physica 98A:587–600
51. Edman L, Foldes-Papp Z, Wennmalm S, Rigler R (1999) The fluctuating enzyme: a single molecule approach. Chem Phys 247:11–22
52. Eigen M, de Maeyer L (1963) Relaxation methods. In: Weissberger A (ed) Technique of organic chemistry, vol 8, Part II. Interscience, New York, p 895
53. Elowitz MB, Levine AJ, Siggia ED, Swain PS (2002) Stochastic gene expression in a single cell. Science 297:1183–1186
54. English BP, Min W, van Oijen AM, Lee KT, Luo G, Sun H, Cherayil B, Kou SC, Xie XS (2006) Ever-fluctuating single enzyme molecules: Michaelis–Menten equation revisited. Nat Chem Biol 2:87–94
55. Ërdi P, Ropolyi L (1979) Investigation of transmitter-receptor interactions by analyzing postsynaptic membrane noise using stochastic kinetics. Biol Cybern 32:41–45
56. Érdi P, Tóth J (1989) Mathematical models of chemical reactions: theory and applications of deterministic and stochastic models. Manchester University Press, Manchester/Princeton University Press, Princeton

57. Érdi P, Tóth J, Hárs V (1981) Some kinds of exotic phenomena in chemical systems. In: Farkas M (ed) Qualitative theory of differential equations, Szeged. Colloquia mathematica societatis János Bolyai, vol 30. North-Holland, Amsterdam/New York/János Bolyai Mathematical Society, Budapest, pp 205–229
58. Falcke M (2003) On the role of stochastic channel behavior in intracellular Ca^{2+} dynamics. Biophys J 84:42–56
59. Feher G, Weissman M (1973) Fluctuation spectroscopy: determination of chemical reaction kinetics from the frequency spectrum of fluctuations. Proc Natl Acad Sci USA 70:870–875
60. Filisetti A, Graudenzi A, Serra R, Villani M, Lucrezia D, Füchslin RM, Kauffman SA, Packard N, Poli I (2011) A stochastic model of the emergence of autocatalytic cycles. J Syst Chem 2:2–10
61. Fisher RA (1925) Theory of statistical estimation. Proc Camb Philos Soc 22:700–725
62. Flomenbom O, Velonia K, Loos D, Masuo S, Cotlet M, Engelborghs Y, Hofkens J, Rowan AE, Nolte RJM, Van der Auweraer M, de Schryver FC, Klafter J (2005) Stretched exponential decay and correlations in the catalytic activity of fluctuating single lipase molecules. Proc Natl Acad Sci USA 102:2368–2372
63. Frank FC (1953) On spontaneous asymmetric synthesis. Biochim Biophys Acta 11:459–463
64. Frehland E (1982) Stochastic transport processes in discrete biological systems. Springer, Berlin/New York
65. Gadgil C, Lee CH, Othmer HG (2005) A stochastic analysis of first-order reaction networks. Bull Math Biol 67:901–946
66. Gans PJ (1960) Open first-order stochastic processes. J Chem Phys 33:691–694
67. Grenander U (1950) Stochastic processes and statistical inference. Arkiv för matematik 0004–2080; 1:17 (in Swedish). Almqvist and Wiksell, Stockholm
68. Gridnev ID, Serafimov JM, Quiney H, Brown JM (2003) Reflections on spontaneous asymmetric synthesis by amplifying autocatalysis. Org Biomol Chem 1:3811–3819
69. Haimovich G, Medina DA, Causse SZ, Garber M, Millán-Zambrano G, Barkai O, Chávez S, Pérez-Ortin JE, Darzacq X, Choder M (2013) Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. Cell 153:1000–1011
70. Hanggi P, Grabert H, Talkner P, Thomas H (1984) Bistable systems: master equation versus Fokker-Planck modeling. Phys Rev 29:371–377
71. Heineken FG, Tsuchiya HM, Aris R (1967) On the mathematical status of the pseudo-steady state hypothesis of biochemical kinetics. Math Biosci 1:95–113
72. Hilfinger A, Paulsson J (2011) Separating intrinsic from extrinsic fluctuations in dynamic biological systems. Proc Natl Acad Sci USA 108(29):12167–12172
73. Hilfinger A, Chen M, Paulsson J (2012) Using temporal correlations and full distributions to separate intrinsic and extrinsic fluctuations in biological systems. Phys Rev Lett 109:248104
74. Holden AV (1976) Models of the stochastic activity of neurones. Lecture notes in biomathematics, vol 12. Springer, Berlin
75. Hongler MO, Zheng WM (1982) Exact solution for the diffusion in bistable potentials. J Stat Phys 29(2):317–327
76. Intosalmi J, Manninen T, Ruohonen K, Linne ML (2011) Computational study of noise in a large signal transduction network. BMC Bioinform 12:252
77. Ishida K (1966) Stochastic approach to nonequilibrium thermodynamics of first-order chemical reactions. J Phys Chem 70:3806–3811
78. Ishida K (1968) Stochastic approach to nonequilibrium thermodynamics of first-order chemical reactions. II. Open systems. J Phys Chem 72:92–96
79. Iyer-Biswas S, Hayot F, Jayaprakash C (2009) Stochasticity of gene products from transcriptional pulsing. Phys Rev E Stat Nonlinear Soft Matter Phys 79(3):031911
80. Jachimowski CJ, McQuarrie DA, Russle MS (1964) A stochastic approach to enzyme-substrate reaction. Biochemistry 3:1732–1736
81. Jacob F, Monod J (1961) Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol 3:318–356
82. Jahnke T, Huisinga W (2007) Solving the chemical master equation for monomolecular reaction systems analytically. J Math Biol 54:1–26

83. Kaissling KE (1998) Flux detectors versus concentration detectors: two types of chemoreceptors. Chem Senses 23:99–111
84. Karp RM, Miller RE (1969) Parallel program schemata. J Comput Syst Sci 3:147–195
85. Kask P, Palo K, Ullmann D, Gall K (1999) Fluorescence-intensity distribution analysis and its application in biomolecular detection technology. Proc Natl Acad Sci USA 96:13756–13761
86. Katz B, Miledi R (1972) The statistical nature of the acetylcholine potential and its molecular components. J Physiol 1224:665–699
87. Kawasaki T, Suzuki K, Shimizu M, Ishikawa K, Soai K (2006) Spontaneous absolute asymmetric synthesis in the presence of achiral silica gel in conjunction with asymmetric autocatalysis. Chirality 18:479–482
88. Keizer J (1979) Nonequilibrium thermodynamics and the stability of states far from equilibrium. Acc Chem Res 12:243–249
89. Kou SC, Cherayil BJ, Min W, English BP, Xie SX (2005) Single-molecule Michaelis–Menten equations. J Phys Chem B 109:19068–19081
90. Krieger IM, Gans PJ (1960) First-order stochastic processes. J Chem Phys 32:247–250
91. Kuang Y (1993) Delay differential equations: with applications in population dynamics. Academic, Boston
92. Lai TL (2009) Martingales in sequential analysis and time series, 1945–1985. Journ@l Électronique d'Histoire des Probabilités et de la Statistique/Electron J Hist Probab Stat **5**(1):31. Article 11
93. Landauer R (1962) Fluctuations in bistable tunnel diode circuits. J Appl Phys 33:2209–2216
94. Landauer R (1978) Stability in the dissipative steady state. Phys Today 31:23–30
95. Landolt H (1886) Ueber die Zeitdauer der Reaction zwischen Jodsaeure und schwefliger Saeure. Ber Dtsch Chem Ges 19:1317–1365
96. Lee NK, Koh HR, Han KY, Lee J, Kim SK (2010) Single-molecule, real-time measurement of enzyme kinetics by alternating-laser excitation fluorescence resonance energy transfer. Chem Commun 46:4683–4685
97. Lente G (2004) Homogeneous chiral autocatalysis: a simple, purely stochastic kinetic model. J Phys Chem A 108:9475–9478
98. Lente G (2005) Stochastic kinetic models of chiral autocatalysis: a general tool for the quantitative interpretation of total asymmetric synthesis. J Phys Chem A 109:11058–11063
99. Lente G (2006) Stochastic analysis of the parity-violating energy differences between enantiomers and its implications for the origin of biological chirality. J Phys Chem A 110:12711–12713
100. Lente G (2007) The effect of parity violation on kinetic models of enantioselective autocatalysis. Phys Chem Chem Phys 9:6134–6141
101. Lente G (2010) The role of stochastic models in interpreting the origins of biological chirality. Symmetry 2:767–798
102. Lente G (2012) Stochastic modeling of the Soai reaction. In: Pályi G, Zucchi C, Caglioti L (eds) The Soai reaction and related topic. Artestampa - Accademia Nazionale di Scienze, Lettere ed Arti, Modena, pp 123–147
103. Lente G (2012) Stochastic mapping of first order reaction networks: a systematic comparison of the stochastic and deterministic kinetic approaches. J Chem Phys 137:164101
104. Lente G (2013) A binomial stochastic kinetic approach to the Michaelis-Menten mechanism. Chem Phys Lett 568–569:167–169
105. Lente G, Ditrói T (2009) Stochastic kinetic analysis of the frank model. Stochastic approach to flow-through reactors. J Phys Chem B 113:7237–7242
106. Lente G, Bazsa G, Fábián I (2007) What is and what isn't a clock reaction? New J Chem 37:1707–1707
107. Leonard DS, Reichil LE (1994) Stochastic resonance in a chemical reaction. Phys Rev E 49:1734–1737
108. Leontovich MA (1935) Basic equations of kinetic gas theory from the viewpoint of the theory 618 of random processes. J Exp Theor Phys 5:211–231

109. Li H, Bi JH (2011) The role of internal noise for optimal intracellular calcium signaling in coupled biological cell system. Fluct Noise Lett 10:31–39
110. Liptser RS, Shiryaev AN (1977) Statistics of random processes. I. General theory. Springer, Berlin (In Russian: Statistika slu'ajnych processov. Nauka, Moscow, 1974)
111. Liptser RS, Shiryaev AN (1978) Statistics of random processes. II. Applications. Springer, Berlin
112. Lopez L, Piegari E, Sigaut L, Ponce Dawson S (2012) Intracellular calcium signals display an avalanche-like behavior over multiple length-scales. Front Physiol 3:1–13
113. Lu HP, Xun L, Xie XS (1998) Single-molecule enzymatic dynamics. Science 282:1877–1882
114. Mackey MC, Glass L (1977) Oscillation and chaos in physiological control systems. Science 197:287–289
115. Magnasco MO (1997) Chemical kinetics is turing universal. Phys Rev Lett 78:1190–1193
116. Mastny EA, Haseltine EL, Rawlings JB (2007) Two classes of quasi-steady-state model reductions for stochastic kinetics. J Chem Phys 127:094106
117. Matsuo KK, Lindenberg K, Shuler KE (1978) Stochastic theory of nonlinear rate processes with multiple stationary states. II. Relaxation time from a metastable state. J Stat Phys 19:65–75
118. Mauksch M, Tsogoeva SB, Wei S, Martynova IM (2007) Demonstration of spontaneous chiral symmetry breaking in asymmetric Mannich and Aldol reactions. Chirality 19:816–825
119. McAdams HH, Arkin A (1997) Stochastic mechanisms in gene expression. Proc Natl Acad Sci USA 94:814–819
120. McDonnell MD, Abbott D (2009) What is stochastic resonance? Definitions, misconceptions, debates, and its relevance to biology. PLoS Comput Biol 5(5):e1000348
121. McQuarrie DA (1963) Kinetics of small systems. I. J Chem Phys 38:433–436
122. McQuarrie DA (1967) Stochastic approach to chemical kinetics. J Appl Probab 4:413–478
123. McQuarrie DA, Jachimowski CJ, Russell ME (1964) Kinetics of small systems. II. J Chem Phys 40:2914
124. Michaelis L, Menten ML (1913) Kinetics of invertase action. Biochem Z 49:333–369
125. Mills W (1932) Some aspects of stereochemistry. Chem Ind (Lond) 51:750–759
126. Mitra PP, Stark JB (2001) Nonlinear limits to the information capacity of optical fibre communications. Nature 411:1027–1030
127. Monod J, Jacob F (1961) Telenomic mechanisms in cellular metabolism, growth and differentiation. Cold Springer Harb Symp Quant Biol 26:389–401
128. Montroll EW, Shuler KE (1958) The application of the theory of stochastic processes to chemical kinetics. Adv Chem Phys 1:361–399
129. Morange M (2005) What history tell us. J Biosci 30:313–316
130. Müller JD, Chen Y, Gratton E (2000) Resolving heterogeneity on the single molecular level with the photon-counting histogram. Biophys J 78:474–486
131. Mulloolly P (1971) Maximum likelihood estimation for stochastic first order reactions. Bull Math Biol 33:83–96
132. Mulloolly P (1972) Maximum likelihood estimation for stochastic rth order reactions. J Appl Probab 9:32–42
133. Mulloolly P (1973) Maximum likelihood estimation for stochastic rth orderreactions, II. J Appl Probab 10:441–446
134. Nagypál I, Epstein IR (1986) Fluctuations and stirring rate effects in the chlorite-thiosulfate reaction. J Phys Chem 90:6285–6292
135. Nagypál I, Epstein IR (1988) Stochastic behavior and stirring rate effects in the chlorite-iodide reaction. J Chem Phys 89:6925–6928
136. Neher E, Sakmann B (1976) Single-channel currents recorded from membrane of denervated frog muscle fibres. Nature 260:799–802
137. Nitzan A, Ortoleva P, Deutch J, Ross J (1974) Fluctuations and transitions at chemical instabilities: the analogy to phase transitions. J Chem Phys 61:1056–1074
138. Novick A, Weiner M (1957) Enzyme induction as an all-or-none phenomenon. Proc Natl Acad Sci USA 43:553–566

139. Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A (2002) Regulation of noise in the expression of a single gene. Nat Genet 31:69–73
140. Palme AG, Thompson NL (1987) Molecular aggregation characterized by high order autocorrelation in fluorescence correlation spectroscopy. Biophys J 52:257–270
141. Parmananda P, Escalera Santos GJ, Rivera M, Showalter K (2005) Stochastic resonance of electrochemical aperiodic spike trains. Phys Rev E Stat Nonlinear Soft Matter Phys 71(3 Pt 1):031110
142. Paulsson J (2005) Models of stochastic gene expression. Phys Life Rev 2:157–175
143. Perkins TJ, Weisse AY, Swain PS (2012) Chance and memory. In: Wall ME (ed) Quantitative biology: from molecular to cellular systems. CRC, Boca Raton
144. Petrásek Z, Schwille P (2009) Fluctuations as a source of information in fluorescence microscopy. J R Soc Interface 6:S15–S25
145. Pokora O (2010) Fisher information as the measure of signal optimality in olfactory neuronal models. Acta Univ M Belii ser Math 16:53–63
146. Pokora O, Lansky P (2008) Statistical approach in search for optimal signal in simple olfactory neuronal models. Math Biosci 214:100–108
147. Poovathingal S, Gunawan R (2010) Global parameter estimation methods for stochastic biochemical systems. BMC Bioinform 11(1):414
148. Procaccia I, Ross J (1977) Stability and relative stability in reactive systems far from equilibrium. II. Kinetic analysis of relative stability of multiple stationary states. J Chem Phys 67:5765–5771
149. Qian H (2011) Nonlinear stochastic dynamics of mesoscopic homogeneous biochemical reaction systems: an analytical theory. Nonlinearity 24:R19–R49
150. Qian H, Elson EL (1990) On the analysis of high order moments of fluorescence fluctuations. Biophys J 57:375–380
151. Qian H, Elson EL (2002) Single-molecule enzymology: stochastic Michaelis–Menten kinetics. Biophys Chem 101–102:565–576
152. Raj A, van Oudenaarden A (2008) Stochastic gene expression and its consequences. Cell 35:216–226
153. Rao CV, Arkin AP (2003) Stochastic chemical kinetics and the quasi-steady-state assumption: application to the Gillespie algorithm. J Chem Phys 118:4999–5010
154. Raser JM, O'shea EK (2004) Control of stochasticity in eukaryotic gene expression. Science 304:1811–1814
155. Reinker S, Altman R, Timmer J (2006) Parameter estimation in stochastic biochemical reactions. IEE Proc Syst Biol 153:168–178
156. Ribeiro AS (2010) Stochastic and delayed stochastic models of gene expression and regulation. Math Biosci 223:1–11
157. Rigney DR (1979) Stochastic model of constitutive protein levels in growing and dividing bacterial cells. J Theor Biol 76:453–480
158. Rigney DR (1979) Note on the kinetics and Stochasticity of induced protein synthesis as influenced by various models for messenger RNA degradation. J Theor Biol 79:247–257
159. Rigney DR, Schieve WC (1977) Stochastic model of linear, continuous protein-synthesis in bacterial populations. J Theor Biol 69:761–766
160. Rössler OE (1974) Chemical automata in homogeneous and reaction-diffusion kinetics. In: Conrad M, Güttinger W, Dal Cin M (eds) Physics and mathematics of the nervous system. Lecture notes in biomathematics, vol 4. Springer, Berlin, pp 399–418, 546–582
161. Ruess J, Milias-Argeitis A, Lygeros J (2013) Designing experiments to understand the variability in biochemical reaction networks. R Soc Interface 10:20130588
162. Saito Y, Sugimori T, Hyuga H (2007) Stochastic approach to enantiomeric excess amplification and chiral symmetry breaking. J Phys Soc Jpn 76:044802
163. Salari R, Wojtowicz D, Zheng J, Levens D, Pilpel Y, Przytycka TM (2012) Teasing apart translational and transcriptional components of stochastic variations in eukaryotic gene expression. PLoS Comput Biol 8:8

164. Samoilov M, Arkin A, Ross J (2002) Signal processing by simple chemical systems. J Phys Chem 106:10205–10221
165. Samoilov M, Plyasunov S, Arkin A (2005) Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. Proc Natl Acad Sci USA 102:2310–2315
166. Samoilov M, Plyasunov S, Arkin AP (2005) Stretched exponential decay and correlations in the catalytic activity of fluctuating single lipase molecules. Proc Natl Acad Sci USA 102:2368–2372
167. Sanft KR, Gillespie DT, Petzold LR (2011) Legitimacy of the stochastic Michaelis–Menten approximation. IET Syst Biol 5:58–69
168. Segel LA (1988) On the validity of the steady state assumption of enzyme kinetics. Bull Math Biol 50:579–593
169. Shahrezaei V, Swain PS (2008) The stochastic nature of biochemical networks. Curr Opin Biotechnol 19:369–374
170. Shao J, Liu L (2007) Stochastic fluctuations and chiral symmetry breaking: exact solution of Lente model. J Phys Chem A 111:9570–9572
171. Siegert AJF (1949) On the approach to statistical equilibrium. Phys Rev 76:1708–1714
172. Singer K (1953) Application of the theory of stochastic processes to the study of irreproducible chemical reactions and nucleation processes. J R Stat Soc Ser B 15:92–106
173. Singleton DA, Vo LK (2002) Enantioselective synthesis without discrete optically active additives. J Am Chem Soc 124:10010–10011
174. Singleton DA, Vo LK (2003) A few molecules can control the enantiomeric outcome. Evidence supporting absolute asymmetric synthesis using the Soai asymmetric autocatalysis. Org Lett 5:4337–4339
175. Skupin A, Kettenmann H, Falcke M (2010) Calcium signals driven by single channel noise. PLoS Comput Biol 6:pii:e100087
176. Soai K, Sato I, Shibata T, Komiya S, Hayashi M, Matsueda Y, Imamura H, Hayase T, Morioka H, Tabira H, Yamamoto J, Kowata Y (2003) Asymmetric synthesis of pyrimidyl alkanol without adding chiral substances by the addition of diisopropylzinc to pyrimidine-5-carbaldehyde in conjunction with asymmetric autocatalysis. Tetrahedron Asymmetry 14:185–188
177. Solc M (1972) Stochastic model of reversible first-order reaction: fluctuations near to equilibrium. Z Phys Chem 79:278–287
178. Solc M (1999) A stochastic model of chemical equilibrium: case study of the reversible first-order reaction. Acta Chim Hung Models Chem 136:647–661
179. Solc M (2000) A stochastic model of chemical equilibrium in a system where the two-stage consecutive reversible first-order reaction takes place. Acta Chim Hung Models Chem 137:547–560
180. Solc M (2002) Stochastic model of the n-stage reversible first-order reaction: relation between the time of first passage to the most probable microstate and the mean equilibrium fluctuations lifetime. Z Phys Chem 216:869–893
181. Soloveichik D, Cook M, Winfree E, Bruck J (2008) Computation with finite stochastic chemical reaction networks. Nat Comput 7:615–633
182. Staff PJ (1970) A stochastic development of the reversible Michaelis–Menten mechanism. J Theor Biol 27:221–232
183. Sugimori T, Hyuga H, Saito Y (2008) Fluctuation induced homochirality. J Phys Soc Jpn 77:064606
184. Svoboda K, Mitra PP, Block SM (1994) Fluctuation analysis of motor protein movement and single enzyme kinetics. Proc Natl Acad Sci USA 91:11782–11786
185. Swain PS, Elowitz MB, Siggia FD (2002) Intrinsic and extrinsic contributions to stochasticity in gene expression. Proc Natl Acad Sci USA 99:12795–12800
186. Tetin SE (ed) (2013) Methods in enzymology; v. 518: fluorescence fluctuation spectroscopy (ffs), Part 1. Elsevier/Academic, San Diego
187. Tian T, Xu S, Gao J, Burrage K (2007) Simulated maximum likelihood method for estimating kinetic rates in gene expression. Bioinform 23:84–91

188. Togashi DT, Kaneko K (2005) Discreteness-induced transitions in autocatalytic systems. arXiv:physics/0501017
189. Tostevin F, ten Wolde PR (2009) Mutual information between input and output trajectories of biochemical networks. Phys Rev Lett 102(21):218101
190. Tostevin F, ten Wolde PR (2010) Mutual information in time-varying biochemical systems. Phys Rev E 81:061917
191. Turner TE, Schnell S, Burrage K (2004) Stochastic approaches for modelling in vivo reactions. Comput Biol Chem 28:165–178
192. van Kampen NG (2007) Stochastic processes in physics and chemistry, 3rd edn. Elsevier, Amsterdam
193. van Oijen AM, Blainey PC, Crampton DJ, Richardson CC, Ellenberger T, Xie XS (2003) Single-molecule kinetics of λ exonuclease reveal base dependence and dynamic disorder. Science 301:1235–1239
194. Vellela M, Qian H (2007) A quasistationary analysis of a stochastic chemical reaction: Keizer's paradox. Bull Math Biol 69:1727–1746
195. Vellela M, Qian H (2009) Stochastic dynamics and non-equilibrium thermodynamics of a bistable chemical system: the Schlögl model revisited. J R Soc Interface 6:925–940
196. Velonia K, Flomenbom O, Loos D, Masuo S, Cotlet M, Engelborghs Y, Hofkens J, Rowan AE, Klafter J, Nolte RJM, de Schryver FC (2005) Single-enzyme kinetics of CALB-catalyzed hydrolysis. Angew Chem Int Ed 44:560–564
197. Walczak AM, Mugler A, Wiggins CH (2012) Analytic methods for modeling stochastic regulatory networks. Methods Mol Biol 880:273–322
198. Warren PB, Tanase-Nicola S, ten Wolde PR (2006) Exact results for noise power spectra in linear biochemical reaction networks. J Chem Phys 125:144904
199. Yeung E, Beck JL, Murray RM (2013) Modeling environmental disturbances with the chemical master equation. In: 2013 conference on decision and control (CDC), Florence (Submitted)
200. Yi M, Liu Q (2010) Michaelis–Menten mechanism for single-enzyme and multi-enzyme system under stochastic noise and spatial diffusion. Physica A 389:3791–3803
201. Yildirim N, Mackey M (2003) Feedback regulation in the lactose operon: a mathematical modeling study and comparison with experimental data. Biophys J 84:2841–2851
202. Zhang J, Hou Z, Xin H (2005) Effects of internal noise for calcium signaling in a coupled cell system. Phys Chem Chem Phys 7:2225–2228
203. Ziv E, Nemenman I, Wiggins CH (2007) Optimal signal processing in small stochastic biochemical networks. PLoS One 24:e1077
204. Zwanzig R (1990) Rate processes with dynamical disorder. Acc Chem Res 23:148–152

Chapter 4
The Book in Retrospect and Prospect

In Chap. 1 we introduced the topic from different perspectives.

We started the book with the general features of chemical kinetics. The mass action type kinetic is a postulate in the phenomenological, deterministic model of chemical reaction kinetics. While we restricted ourselves to deal with spatially homogeneous systems only, what we see for the future is that systems biological applications due to the heterogeneous nature of the cell will require the consideration of spatial events, too. Macroscopic theories of chemical reactions cannot take into account the spatial course of reactions, though it is obviously relevant from a microscopic point of view. (Microscopic simulation of chemical reactions proved to be very successful [12, 15].) Therefore, a chemical reaction is handled as interactions among components being present at the "same place". (The "same place" is considered as the region of space that is sufficiently large to make possible the underlying mechanism of reaction, but sufficiently small to be able to assume that it is a "point".) A chemical reaction has no spatial "cause", and it can be considered as the rearrangement of a "point having internal structure". Traditional chemical kinetics uses the concept of concentration, which is defined asymptotically for infinite systems ($N \to \infty$ and $V \to \infty$ and $\frac{n}{V} = const$). As we see now, random fluctuations occur as a consequence of the small number of reacting molecule. As it was shown (e.g. [31]), stochastic models of **in vivo** reactions should "...include the fluctuation effects caused by the structural organisation of the cytoplasm and the limited diffusion of molecules due to macromolecular crowding." The fundamental problem of chemical kinetics is establishing rate laws based on experimental observations and interpret them on a molecular level. While continuous state space deterministic kinetic models are still very useful in chemical kinetics, there are situations, predominantly but not exclusively in "small systems" [14], where stochastic modeling is a must. "...If the number of particles of the components is small, the fluctuations taken into consideration by the stochastic

P. Érdi and G. Lente, *Stochastic Chemical Kinetics: Theory and (Mostly) Systems Biological Applications*, Springer Series in Synergetics, DOI 10.1007/978-1-4939-0387-0_4,

model only, cannot be neglected even in "zeroth approximation", because they are not superimposed upon the phenomenon but they represent the phenomenon itself" [29].

The main ancestor of any stochastic models occurring in physical, even financial (!) systems is related to Brownian motion. Theoretical works led (i) via the Einstein theory to fluctuation-dissipation theorems, (ii) via Langevin equations to the theory of stochastic differential equations, (iii) via the connection between the stochastic differential equations and the evolution equation for the probability density function to theory of diffusion processes, and (iv) more generally to the establishment of the foundations of stochastic processes, (v) to the experimental determination of the Avogadro number.

The spirit and techniques of using stochastic models in chemical kinetics arrived to chemistry by physicists (Leontovich [20] in the Soviet Union, and Delbrück [7] in the USA) but remained generally unknown for chemists. They worked on specific problems, and used analytic methods. The first complete treatment of a second-order reaction was given by the Hungarian mathematician A. Rényi [26]. This paper gives evidence that the differential equation for the expectation of the stochastic model **cannot** be identified with the differential equation associated with the usual deterministic model. The remaining part of Chap. 1 is a preview of three big questions: How to define the model? How to solve models analytically and by simulations? How the emergence of systems biology implied the renaissance of stochastic kinetics?

Chapter 2 starts with a brief summary of the mathematical foundations of stochastic processes. The most important processes for us are the continuous time discrete Markov processes. The state of the system is characterized by the absolute state probabilities (also called as absolute distribution). By using also transition probabilities and the Chapman-Kolmogorov equation, the **master equation** can be derived. It is interesting to see that important classes of Markov processes, such as birth-and-death process, Poisson process, can be implemented by chemical reactions.

The standard stochastic model of homogeneous reaction kinetics is defined and studied in Sect. 2.3. Specifically, the size and enumeration of the state space was discussed. We feel that the potential applicability of the analysis of the enumeration of state space is underestimated, and it may help to find solutions of the master equation. It is an open problem for these authors, but the whole concept is related to topic discussed in Sect. 3.11. In addition to the derivation of the master equation in the general case, we showed in some simple examples how does it works in the practice.

Since historically deterministic models preceded stochastic ones, early investigators gauged the quality of a stochastic model by the proximity of its behavior to that of the corresponding deterministic one. If one considers that the CDS model takes into consideration the discrete character of the of the state space and it does not neglect fluctuations than the appropriate question asks whether in what sense and what extent can the deterministic model be considered as a good approximation of the stochastic one?

The deepest and most far-reaching results on the relationship of the CDS and CCD models are due to Kurtz [17] and were extended by L. Arnold [2] to cell models of reaction diffusion systems. Since, in general, only (loosely speaking) linear systems have the property of consistency in the mean (that the mean of the stochastic process satisfies the equations of the associated deterministic model) other relationships were proven. Roughly speaking, they gave theorems for a large class of (i.e. conservative and reversible) reactions, where the stochastic model tends to the corresponding deterministic model in thermodynamic limit. This expression means that the number of particles and the volume of the vessel tend to infinity at the same time and in such a way that the concentration of the individual components (i.e. the ratio of the particle number and the volume) tends to a constant and the two model will be close to each other. The **law of large numbers** proved to be also valid. One of us (G.L.) developed the method of **stochastic map** to delimit the parameter space where the only stochastic models lead pragmatically acceptable results, but the result obtained so far should be generalized and proven [19].

Section 2.4 is about the solution of master equations. First we discussed the scope and limits of obtaining analytical and symbolic solutions. One of us (G.L) introduced the concept of time-independent Q functions to characterize the probability of a given state ever occurring during the entire course of the reaction. The method proved to be useful for calculating the distribution of chiral molecules. Laplace transform was used in [26] followed by [18] for a reversible bimolecular reaction. Generating functions are particularly useful to study compartmental systems, since the absolute distribution function can be expressed for every time instant as a function of rate constants and the initial conditions. The Poisson representation technique using so-called "quasi-probabilities" is based on the assumption that state probabilities are given as a superposition of uncorrelated Poisson distributions. The original [9] goal was to transform master equations to Fokker-Planck equations for giving simpler characterization of equilibrium states. A very important feature of stochastic models is related to their multiple time scales. Since the master equation is a linear equation, the eigenvalues of the coefficient matrix give direct information about the parameters of the multi-exponential time course.

Section 2.5 is about stationary and transient distributions. Historically, the role of the Poissonian distribution was overemphasized, and it turned out to have restricted significance. Stationary distributions are generally finite, while the Poisson distribution is of course continuous. Loosely speaking, we see now that unimodal stationary distributions are far from being always Poissonian. As concerns the connection between multistationarity in deterministic models, and multimodality in stochastic models, the trivial assumption is valid asymptotically with increasing volume. Stochastic simulation is the most important methods for solving stochastic models, and the Gillespie method (with its variations) is the only game in the village. (The senior author already psychically processed the missed opportunity mentioned in a previous footnote :-)). As concerns approximative solutions, deterministic continuation is particularly useful in situations when initial fluctuations are the subject of amplification. Even 25 years ago, we saw that "One of the most extensively discussed topics of the theory of stochastic physics is

whether the evolution equations of the discrete state space stochastic processes, i.e. the master equations of the jump processes, can be approximated asymptotically by Fokker-Planck equations when the volume of the system increases…". The accuracy of the approximations is still subject of ongoing debates [10], as even the title of the paper tells: "How accurate are the nonlinear chemical Fokker-Planck and chemical Langevin equations?"

Non-Markovian models are occasionally used, and two examples, one in polymer kinetics and one in gene expression model were given. Generally, the stochastic counterparts of biochemical reactions with time delay are naturally lead to non-exponential inter-event time distributions, so non-markovian process.

Chapter 3 gives a condensed overview about selected applications. The selection is obviously somewhat arbitrary, we tried to illustrate the diversity of topics and methods by alloying historically important and actually main stream applications.

In Sect. 3.2, examples are given for the particularly important role of fluctuations around instability points. First we gave a constructed reaction scheme, where the expectation coincides with the deterministic value, and for some particular relationship for the rate constants, it is constant over time, while the variance linearly increases with time. Keizer's paradox is a paradigmatic example for finding dramatic differences between deterministic and stochastic descriptions. Careful analysis should be based on the description of the metastable states and the relaxation times from them.

The motivation for studying bistable reactions came from two directions. First, a model of a Brownian particle moving in a potential well was adopted by Kramers [16] to reformulate the diffusion model of chemical reactions at the microscopic level. Second, as we briefly reviewed, how diffusion in bistable potential can be described by Fokker-Planck equation. The *lac* operon genetic network is a paradigmatic example of bistable systems, where fluctuations may play a major role for transitions among phenotypes.

Compartmental systems are often used frameworks in biomathematics, as models of spatially discrete transport processes, and they time-dependent solutions can be determined, i.e. the probability distribution can be expressed as the function of the rate constants and the initial conditions.

From the perspective of formal chemical reactions, autocatalytic reactions can be considered as a chemical implementation of positive feedback (economists like to use the term "increasing returns"), a mechanism that is appropriate to amplify microscopic fluctuations. Autocatalysis is assumed to be involved in abiogenesis as well, and is the basis of clock reactions, when classically, a change in the color of the solution occurs after a certain time delay. Under certain conditions, this time delay showed fluctuations [24, 25]. From a chemical point of view, there is also another reaction ("initiating process") that provides a (usually slow) way of forming the first molecules of the product. Autocatalytic extinction and autocatalytic runaway emerges if the initiating process is neglected. Autocatalytic cycle processes (which might have an important role in primordial organisms) are related to discreteness-induced transitions.

While enzyme kinetics is conceptually nothing else than conventional catalytic reactions, the specific nature of the catalyst explains the specific role of enzyme kinetics in chemical kinetics **per se** and its low concentration justifies the importance of the stochastic approach in biochemistry. The master equation for a single enzyme molecule ($e_0 = 1$) was solved by Arányi and Tóth [1].[1]

The stochastic map of the now 101 years old Michaelis–Menten mechanism helps to identify the parameter regions when the use of CDS model is a must.

Probably the most visible development in the application of stochastic chemical kinetics is the extensive use of the concept of signal processing. Chemical reaction systems can be considered as devices which convert (in the general case time-dependent and noisy) inputs to output. Filtering is a general concept to eliminate some unwanted components of the input, and a very important class is **frequency filtering**, and the whole procedure came from the engineering literature. The efficiency of signal processing is characterized by several statistical measures, such as mutual information and Fisher information.

Section 3.6.2 is about signal processing in biochemical networks. While biochemical networks are able to perform computations, similarly to electronic circuits, the mechanisms are rather different, and the main point is to understand how chemically interacting molecules perform these calculations? Probably there is no general theory to describe the effects of the structure of signaling reactions on the reliability of information transmission (characterized by gain-to-noise ratio). We reviewed a number of specific examples (such as reversible binding, deactivation upon detection, autoregulation, certain feedback steps). In certain situations the

[1] Its significance now noticed [11], as we cite it almost verbatim. As they basically write, Arányi and Tóth were the first to systematically study the master equation for enzyme kinetics. They considered the special case in which there is only one enzyme molecule with several substrate molecules in a closed compartment and showed that the master equation can then be solved exactly. The exact solution consists of the probability distribution of the state of the system at any time point. This is remarkable when one considers that it is impossible to solve the CCD model without imposing restrictions on the reaction conditions such as pseudo-first-order kinetics, or applying an approximation. From the exact solution of the probability distribution, Arányi and Tóth derived exact expressions for the time course of the mean substrate and enzyme concentrations and compared them with those obtained by numerical integration of the CCD model. Interestingly, they found differences of 20–30 % between the average substrate concentrations calculated using the CCD and CDS models for the same set of rate constants and for the case of one enzyme reacting with one substrate molecule. If the initial number of substrate molecules is increased to five, whilst keeping the same rate constants, then one notices that the difference between the CCD and CDS results becomes negligibly small. In general, it can be shown that the discrepancy between the two approaches stems from the fact that the mean concentrations, in chemical systems involving second-order reactions, are dependent on the size of the fluctuations in a CDS description and independent in the CCD description. The discrepancies become smaller for larger numbers of substrate molecules because fluctuations roughly scale as the inverse square of the molecule numbers. This important contribution by Arányi and Tóth went largely unnoticed at the time, because experimental approaches did not have the resolution for measuring single-enzyme-catalysed experiments to test the theoretical results.

response to the signal shows bistability, and this phenomenon is noise-induced, so no deterministic counterpart exists.

The concept of **kinetic independence** was offered as a framework for analyzing information encoding and propagation in biochemical networks to remedy at least partially the lack of general theory [4,5]. Probably it is too early to see the scope and limits of the new approach, which uses conditional independence among realization of the individual chemical components, but looks very interesting.

Odor information is processed at the early stages of olfactory system by odor-receptor interactions. Papers which presented simpler and more complex stochastic kinetic models of some odor intensity detection schemes were reviewed. Several criteria for the estimation of optimal olfactory signal (i.e. odorant concentration) exists, one of them is the Fisher information. (We have not found any paper which used mutual information to characterize the efficiency of the odor information processing.) There were some reaction schemes, where the deterministic and the stochastic model led different optimal signals.

Stochastic models are extensively used and proved to be very useful in describing intracellular calcium signaling. Noisy environment and low particle number implies the necessity of using CDS models. Calcium signaling is related to a hierarchy of spatiotemporal events from localized single channel gating to global spatiotemporal pattern formation. A new, excellent paper [27] gives a summary of stochastic models of intracellular calcium signaling. After reviewing the general modeling concepts for intracellular biochemical reactions, the author describes process at different hierarchical levels starting from microscopic events (such as kinetic models of gating) via mesoscopic processes (such as reaction-diffusion models and nonequilibrium dynamic description of single $IP3$ channels) where calcium ion feedback for a single stochastic channel is a very important step – and spatial modeling of clustered channels to macroscopic scale (such as whole-cell oscillations and waves).

Gene expression is maybe the most active field of applied stochastic chemical kinetics. Protein (and somewhat mRNA) fluctuations reflect the inherent random character of the translation (and transcription) process. Noise in gene regulatory networks were analyzed [21]. Related to the **lac operon** model, which is the first workhorse of gene regulation, the differences in the switching behavior in deterministic and stochastic models were pinpointed.[2] Generally, relatively simple,

[2]It was interesting too see, how the relationship between bistability (i.e. three-stationarity) and bimodality was commented by [21]:"... It is often assumed that bistability of deterministic mass action kinetics is associated with bimodality in the steady-state solution to the master equation. However, this is often not the case-we can have bistability without obvious bimodality and bimodality without bistability. In fact, steady-state solutions to either the mass action kinetics or the master equation can be very misleading-we cannot ignore the dynamics...." Cobb illustrated in 1978 [6] with the aid of a non-kinetic model that there is **no** one-to-one correspondence between the location of the equilibrium points and of the extrema of stationary distributions. Somewhat artificial kinetic examples were constructed [8] to support the view that all the four possible cases among uni- and multistationarity and uni- and bimodality may occur.

weakly realistic stochastic kinetic models have been studied analytically, much more by simulations. We might be wrong, but probably we are not too far from the reality to state that a general study of stochastic gene expression model from the perspective of the theory of stochastic kinetic is missing. Separation of intrinsic and extrinsic ("internal" and "external" in the older kinetic literature) is an important topic [13]. There are admitted mathematical difficulties in the case of multiplicative noise and for reactions, when the deterministic model describes nonlinear kinetics. (The neglect of the latter in stochastic gene expression model is somewhat unjustified.)

Section 3.8 investigates the stochastic kinetic aspects of chiral symmetry (i.e. associated with the lack of certain symmetry elements). Due to its remarkable importance in biochemistry, the formation of **homochirality** was extensively studied. Many biological macromolecules have this property, i.e. when mirror image counterparts have very different roles: typically, only one of them is abundant and it cannot be exchanged with the other one.

In racemic mixtures, the two mirror-image pairs (enantiomers) occur in equal quantities (in deterministic approximation), or characterized by a binomial distribution by using statistical approach. In (even not too) large systems the fluctuations are below the detectability threshold. The basic question was to find kinetic schemes which generate non-binomial stationary distributions of the enantiomers. *Enantioselective autocatalysis* (where both words are emphasized) is supposed to be a chemically realistic amplifying mechanism to produce homochirality. A simple enantioselective autocatalysis model was first presented followed by the description of the celebrated (originally purely deterministic) Frank model (1953). Finally, a chemically realistic system, namely the Soai reaction was briefly reviewed. The origin of biological homochirality is still under extensive investigations (see e.g. [3, 23], and the appearance of more realistic mechanisms for the amplification of small initial deviations due to fluctuations can be expected.

Since there were relatively few measurements for chemical fluctuations, the development of techniques of parameter estimation for stochastic models was a neglected fields. The spirit of the fluctuation-dissipation theorem was used to calculate individual rate constants from equilibrium fluctuations. Since fluctuation is much better visible in $2D$ systems, it is understandable that noise analysis was much more frequent related to cell membrane gating processes than to reactions in conventional chemical solutions. With the appearance of time-dependent fluctuation data obtained by high resolution fluorescence microscopy parameter estimation techniques for stochastic models of biochemical reactions are developing rapidly. There is active research on both derivative based and derivative free methods [30]. While some methods seem computationally effective and more reliable than others :-), the computational problem is difficult. Probably there is a long way to go from toy models to real life problems to prove the pleasant features of the new methods.

Stochastic resonance is a mechanism by which a system in a noisy environment shows an enhanced sensitivity towards small external time-dependent forcing, and the optimal performance of certain nonlinear systems happens at a certain larger-than-zero noise intensity. While there are misconceptions and debate about use and

misuse of the concept of stochastic resonance [22] some evolved version of the concept [32] might be useful also in (chemical) signal detection.

The very last Sect. 3.11 briefly reviews a fascinating and very original idea, the use of stochastic kinetics in the theory of computation. It was demonstrated that it is possible to design chemical computers that to implement Turing-universal computation. We (the authors!) don't know, however, whether or not there are any studies about the eventual super-Turing character, as neural networks have [28]. It is too early to see the details of the intimate relationship between the computational power of stochastic chemical reaction networks and its role in biological computation. But this is a story of the future.

References

1. Arányi P, Tóth J (1977) A full stochastic description of the Michaelis-Menten reaction for small systems. Acta Biochimica et Biophysica Academiae Scientificarum Hungariae 12:375–388
2. Arnold L (1980) On the consistency of the mathematical models of chemical reactions. In: Dynamics of synergetic systems. Springer series in synergetics. Springer, Berlin/New York, pp 107–118
3. Blackmond DG (2010) The origin of biological homochirality. Cold Spring Harb Perspect Biol 2(5):a002147
4. Bowsher CG (2010) Stochastic kinetic models: dynamic independence, modularity and graphs. Ann Stat 38:2242–2281
5. Bowsher CG (2010) Information processing by biochemical networks: a dynamic approach. J R Soc Interface 8:186–200
6. Cobb L (1978) Stochastic catastrophe models and multimodal distributions. Behav Sci 23:360–374
7. Delbrück M (1940) Statistical fluctuation in autocatalytic reactions. J Chem Phys 8:120–124
8. Érdi P, Tóth J, Hárs V (1981) Some kinds of exotic phenomena in chemical systems. In: Farkas M (ed) Qualitative theory of differential equations (Szeged). Colloquia Mathematica Societatis János Bolyai, vol 30. North-Holland/János Bolyai Mathematical Society, Budapest, pp 205–229
9. Gardiner CW, Chaturvedi S (1977) The poisson representation. I. A new technique for chemical master equations. J Stat Phys 17:429–468
10. Grima R, Thomas P, Straube AV (2011) How accurate are the nonlinear chemical Fokker-Planck and chemical Langevin equations? J Chem Phys 135:084103
11. Grima R, Walter NG, Schnell S (2014) Single-molecule enzymology à la Michaelis-Menten. FEBS J. doi:10.1111/febs.12663
12. Hersbach DR (1987) Molecular dynamics of elementary chemical reactions. Angew Chem 26:1121–1143
13. Hilfinger A, Chen M, Paulsson J (2012) Using temporal correlations and full distributions to separate intrinsic and extrinsic fluctuations in biological systems. Phys Rev Lett 109(24):248104
14. Hill T (1964) Thermodynamics of small systems – two volumes bound as one. Dover, New York
15. Karplus M, McCammon JA (2002) Molecular dynamics simulations of biomolecules. Nat Struct Biol 9:646–52
16. Kramers HA (1940) Brownian motion in a field of force and the diffusion model of chemical 605 reactions. Physica 7:284–304

17. Kurtz TG (1972) The relationship between stochastic and deterministic models for chemical reactions. J Chem Phys 57:2976–2978
18. Laurenzi IJ (2000) An analytical solution of the stochastic master equation for reversible bimolecular reaction kinetics. J Chem Phys 113:3315
19. Lente G (2005) Stochastic kinetic models of chiral autocatalysis: a general tool for the quantitative interpretation of total asymmetric synthesis. J Phys Chem A 109:11058–11063
20. Leontovich MA (1935) Basic equations of kinetic gas theory from the viewpoint of the theory 618 of random processes. J Exp Theor Phys 5:211–231
21. Lestas I, Paulsson J, Ross NE, Vinnicombe G (2008) Noise in gene regulatory networks. IEEE Trans Autom Control 53:189–200
22. McDonnell MD, Abbott D (2009) What is stochastic resonance? Definitions, misconceptions, debates, and its relevance to biology. PLoS Comput Biol 5:e1000348
23. Morneau BM, Kubala JM, Barratt C, Schwartz PM (2014) Analysis of a chemical model system leading to chiral symmetry breaking: implications for the evolution of homochirality. J Math Chem 52:268–282
24. Nagypál I, Epstein IR (1986) Fluctuations and stirring rate effects in the chlorite-thiosulfate reaction. J Phys Chem 90:6285–6292
25. Nagypál I, Epstein IR (1988) Stochastic behavior and stirring rate effects in the chlorite-iodide Reaction. J Chem Phys 89:6925–6928
26. Rényi A (1953) Kémiai reakciók tárgyalása a sztochasztikus folyamatok elmélete segítségével (in Hungarian). (Treating chemical reactions using the theory of stochastic process.) MTA Alk Mat Int Közl 2:83–101
27. Rüdiger S (2014) Stochastic models of intracellular calcium signals. Phys Rep 534:39–87
28. Siegelmann HT (1995) Computation beyond the Turing limit. Science 268:545–548
29. Sipos T, Tóth J, Érdi P (1974) Stochastic simulation of complex chemical reactions by digital computer. I. The model. React Kinet Catal Let 1:113–117
30. Srivastava R, Rawlings JB (2014) Parameter estimation in stochastic chemical kinetic models using derivative free optimization and bootstrapping. Comput Chem Eng 63:152–158
31. Turner TE, Schnell S, Burrage K (2004) Stochastic approaches for modelling in vivo reactions. Comput Biol Chem 28:165–178
32. Zhang X, Hu N, Cheng Z, Hu L (2012) Enhanced detection of rolling element bearing fault based on stochastic resonance. Chin J Mech Eng 25:1287–1297

Index

P. Érdi and G. Lente, *Stochastic Chemical Kinetics: Theory and (Mostly) Systems Biological Applications*, Springer Series in Synergetics, DOI 10.1007/978-1-4939-0387-0,