Vorlesung: Sommersemester 2013
270151 VO Stochastische Kinetik. 2 Std. 3.0 ECTS credits
Kapitel 27.03; 44.01, F 300.
Vortragender: Peter Schuster
Ort: Institut für Theoretische Chemie, Seminarraum 3. Stock
Vorbesprechung und erste Vorlesungsstunde: **Freitag, 08.03.2013, 9.00 – 10:30 Uhr**

Weitere Vorlesungstermine: Do. 14.03. und Fr. 15.03.
Do. 21.03. und Fr. 22.03.
Fr. 17.05. und Di. 28.05.
Mo. 03.06. und Do. 06.06.
Mo. 10.06., Fr. 14.06. und Mo. 17.06.
Mo. 24.06. und Fr. 28.06., jeweils 9:00 – 10:30 Uhr.

Vorlesungsunterlagen:
`http://www.tbi.univie.ac.at/~pks/Preprints/stochasticity.pdf`

Preface

Statistics and stochastic processes are often neglected mathematical disciplines in the education of chemists and biologists, although modern experimental techniques allow for investigations on small sample sizes down to single molecules and most measured data are sufficiently accurate to allow for direct detection of fluctuations. The progress in the development of new techniques and the improvement in the resolution of conventional experiments has been enormous within the last fifty years. Indeed, molecular spectroscopy provided hitherto unimaginable insights into processes down to the hundred attosecond range and current theory in physics, chemistry, and the life sciences cannot be successful without a deeper understanding of randomness and its causes. Sampling of data and reproduction of processes are doomed to produce artifacts in interpretation unless the observer has a solid background in the mathematics of limited reproducibility. As a matter of fact stochastic processes are much closer to observations than deterministic descriptions in modern science and everyday life. Exceptions are the motions of planets and moons as encapsulated in celestial mechanics, which stood at the beginnings of science and modeling by means of differential equations. Fluctuations are so small that they cannot be detected even in highest precision measurements: Sunrise, sunset, and solar eclipses are predictable with practically no scatter. Processes in the life sciences are often entirely different. A famous and characteristic historical example are Mendel's laws of inheritance: Regularities are detectable only in sufficiently large samples of individual observations, and the influence of stochasticity is ubiquitous. Processes in chemistry are between the extremes: The deterministic approach in conventional chemical reaction kinetics has neither suffered a loss in applicability nor did the results become less reliable in the light of modern experiments. What has increased rather dramatically is the accessible resolution in detectable amounts, space, and time. Deeper insights into mechanisms provided new access to molecular information for theory and practice.

Biology is currently in a state of transition: The molecular connection to chemistry revolutionized the sources of biological data and is setting the stage for a new theoretical biology. Historically biology was based on observation and theory in biology was engaged only in interpretations of the observed regularities. The development of biochemistry at the end of the nineteenth and the first half of twentieth century introduced quantitative thinking in terms of chemical kinetics into some biological subdisciplines. Biochemistry attributed also a new dimension to experiments in biology in the form of *in vitro* studies on isolated and purified biomolecules. A second import of mathematics into biology came in the form of population genetics, which was created in the nineteen twenties as a new theoretical discipline uniting Darwin's natural selection and Mendelian genetics more than twenty years before evolutionary biologists completed the so-called *synthetic theory* performing the same goal. Beginning in the second half of the twentieth century molecular biology started to build a comprehensive bridge from chemistry to biology and enormous progress in experimental techniques created a previously unknown situation in biology insofar as new procedures were required for data handling, analysis, and interpretation since the volume of information is drastically exceeding the capacities of human mind. Biological cells and whole organisms are now accessible to complete description at the molecular level and the overwhelming amount of information thought to be required for a deeper understanding of biological objects is simply a consequence of the complexity of biology and the lack of a universal theoretical biology.

The current flood of results from molecular genetics and genomics to systems biology and synthetic biology requires apart from computer science techniques primarily suitable statistical methods and tools for verification and evaluation of data. Analysis, interpretation, and understanding of experimental results, however, is impossible without proper modeling tools. These tools were so far mainly based on differential equations but it has been realized within the last few years that an extension of the available repertoire by methods derived from stochastic processes is inevitable. Moreover, the enormous complexity of the genetic and metabolic networks in the cell calls for radically new methods of modeling that resemble the mesoscopic level of description in solid state physics. In mesoscopic models the overwhelming and for many purposes dispensable wealth of detailed molecular information is cast into a partially probabilistic description in the spirit of *dissipative particle dynamics*, and such a description cannot be successful without a solid background in stochastic methods. The field of stochastic processes has not been bypassed by the digital revolution. Numerical calculation and computer simulation play a decisive role in present day stochastic modeling in physics, chemistry and biology. Speed of computation and digital storage capacities are growing exponentially since the nineteen sixties with an approximate doubling time of eighteen month, a fact that is commonly addressed as Moore's law. It is not so well known, however, that the spectacular exponential growth in computer power has been overshadowed by the progress in numerical mathematics that led to an enormous increase in the efficiency of algorithms. We give just one example, which was reported by Martin Grötschel: "*The solution of a benchmark production planning model by linear programming would have taken 82 years CPU time in 1988, using the computers and the linear programming algorithms of the day. In 2003 -- fifteen years later -- the same model could be solved in one minute and this means an improvement by a factor of about 43 million. Out of this, a factor of roughly 1000 resulted from the increase in processor speed whereas a factor of 43000 was due to improvement in the algorithms, and many other examples of similar progress in the design of algorithms can be given."* Understanding, analyzing, and designing high-performance numerical methods, however, requires a firm background in mathematics. The availability of cheap computing power has also changed the attitude towards exact results in terms of complicated functions: It does not take so much more computer time to compute a sophisticated hypergeometric function than to calculate an ordinary trigonometric function for an arbitrary argument, and operations

on confusing expressions are enormously facilitated by symbolic computation. In this way the present day computational facilities have also large impact on the analytical work.

In the past biologists had often quite mixed feelings for mathematics and reservations against the use of theory. The recent developments in molecular biology, computation, and applied mathematics, however, seem to initiate a change in biological thinking since there is practically no chance to shape modern biology without mathematics, computer science and theory as the biologist Sydney Brenner, an early pioneer of molecular life sciences, points out ``*... it is clear that the prime intellectual task of the future lies in constructing an appropriate theoretical framework for biology. Unfortunately, theoretical biology has a bad name because of its past. ... Even though alternatives have been suggested, such as computational biology, biological systems theory and integrative biology, I have decided to forget and forgive the past and call it theoretical biology.*'' He and others are calling for a *theoretical biology new* that allows for handling the enormous complexity. Manfred Eigen stated very clearly what can be expected from theory ``*Theory cannot remove complexity but it can show what kind of `regular' behavior can be expected and what experiments have to be done to get a grasp on the irregularities.*'' Theoretical biology will have to find the appropriate way to combine randomness and deterministic behavior in modeling and it is not very risky to guess that it will need a strong anchor in mathematics in order to be successful.

In this course an attempt is made to collect the necessary mathematical background material for understanding stochastic processes. In the sense of Albert Einstein's version of Occam's razor ``*... Everything should be made as simple as possible, but not simpler. ...*'', dispensable deep dwelling in higher mathematics has been been avoided. Some sections that are not required if one is primarily interested in applications are marked for skipping by readers who are willing to accept the basic results without explanations. On the other hand the derivations of analytical solutions for the selected examples are given in full length because the reader who is interested to apply the theory of stochastic processes in practice should be brought in the position to derive solutions on his own. An attempt was made to use a largely uniform notation throughout the. We refrained from preparing a separate section with exercises, instead case studies, which may serve as good examples for calculations by the reader, will be indicated. Sources from literature were among others the text books by Kai Lai Chung, Crispin W. Gardiner, and Robert V. Hogg & Elliot A. Tanis. For a brief and concise introduction we recommend the booklet by Kurt Jacobs.

This scriptum is derived from the manuscripts of courses in stochastic chemical kinetics for graduate students of chemistry and biology held in the years 1999, 2006, and 2011. Critical comments by the students of all four courses were very helpful in the preparation of this text and are gratefully acknowledged.