Summary

Mathematical modeling and analysis of biochemical systems contribute to a better insight in the functional properties of cellular networks, facilitate drug development, and assist in experimental design. Many of the models used are relatively simple, because they have been made with as few assumptions as possible according to the idea of Occam's razor and with the objective to make them tractable by simple mathematics. In recent years however, the development of more realistic models of much higher complexity has been stimulated due to the improved understanding of the components of inherently complex biochemical systems and the development of experimental techniques to verify the models. Indeed, the great advance of biochemistry and molecular biology enabling the understanding of how individual macromolecules work, together with the advances in functional genomics enabling the identification, quantification and purification of virtually all molecules of any living organism, are now increasing the pressure to take the ultimate step to understand the molecular basis of the functioning of organisms. Such understanding should be founded on realistic models of all the components of living organisms together (so-called silicon cells/organisms). The high complexity of the already existing exemplars of parts of such realistic models however often hampers the understanding of what they model and restrains the mathematical analysis. Indeed, a complete realistic model of reality is not yet the understanding of that reality. Mathematical analyses to achieve simplified descriptions of the dynamics are amongst the tools that help such understanding. In this thesis we explore methods to make the realistic biochemical models as tractable as possible while aiming for an accurate description of the complex behavior of the biochemical systems.

The work presented in this thesis concerns two main approaches towards the simplification of biochemical models. First, we study one of the best-known reduction methods deriving from the area of control theory, consisting of so-called balancing and truncation. This is a linear method, i.e., it is applicable to linear models only and the resulting simpler models are also linear. This method has been widely used for reduction of models in other sciences but its application to biochemical models has not been explored much. The second method, which the main part of the work in this thesis is related to, is a nonlinear method (i.e., applicable to and resulting in nonlinear models) based on the so-called zero-derivative principle (ZDP). This method has been applied to only few models before, mainly for theoretical analysis and demonstration of the method. In this thesis we develop a framework for the use of this method in a biochemical and systems biological context.
Balancing and truncation is a method which is applicable to models which are formulated in terms of dynamic input and output variables in addition to the state variables, i.e., the dynamic variables that describe the state of the system. Balancing is the reformulation of a model in terms of new state variables that are linear combinations of the old ones such that the new model, expressed in these new variables, achieves certain properties: the so-called reachability and observability Grammians of the model, i.e., the matrices pertaining to the maps from input to state and state to output, respectively, become equal and diagonal, with the eigenvalues arranged in decreasing order. This implies that the new state variables become ordered according to their importance for the dynamics of the map from input to output. The truncation is the subsequent elimination of those variables that have least importance for this dynamics. In Chapter 2 we apply this method to a realistic nonlinear model of glycolysis in yeast with 13 state variables. Since the method is a linear method, the glycolysis model had to be linearized, i.e. approximated by a linear model. This approximation however resulted in relatively large errors when comparing the time courses of the linear model with those of the original, nonlinear one. We then employed information obtained implicitly from the balancing procedure to estimate the importance of the state variables for the dynamics of the model. Because the analysis was based on linearization, which had proved to yield inaccurate time courses, we performed the analysis at several points of linearization in order to account for possible effects of the nonlinearities. We found, in particular, that the state variables representing the concentrations of NADH and acetaldehyde have little influence on the dynamics of the model. Reducing the nonlinear model by approximating these state variables by their steady state values resulted in a nonlinear reduced model which accurately approximates the original model. Apparently, information obtained from analysis of ill-fitting linear models can still be of use in designing good non-linear approximations of the original models.

In Chapters 3–5 we treat simplifications obtained by implementing the ZDP. These simplifications are based on the fact that biochemical systems often exhibit a large variety of timescales at which the various processes in the system proceed: certain processes change very quickly during a short initial phase while others are almost stagnant. At the end of this phase, the initially fast variables have slowed down and in the view of the fast timescale, the state resembles a steady state although still changing, but now at the slow timescale. In the state space, this behavior amounts to attraction of trajectories by a low-dimensional manifold (subspace) on which the dynamics is slow, a so-called slow invariant manifold (SIM). As a consequence of this behavior, the system spends a large part of the time close to the SIM and, therefore, simplified models describing only the dynamics on this manifold would assist to approximate the behavior of the system. Simplifications of this type are not new to biochemistry; the same idea rests behind the quasi-steady-state approximation (QSSA) which has been widely used in biochemistry for almost a century. The QSSA has been used to derive the well-known enzyme kinetic rate laws of Michaelis–Menten type from mass action kinetics. These rate laws are highly useful for the modeling of enzymatic reactions in cases when the concentration of the enzyme is much lower than that of its substrate since the timescale separation is large under this condition. However, in vivo, this condition often does not hold and the timescale separation is
only moderate. The ZDP method is an extension of the QSSA which provides high accuracy, also in the cases when the timescale separation is moderate.

The ZDP is the principle of setting the time derivatives of any given order of the state variables that change quickly during the short initial phase, equal to zero. The resulting set of algebraic equations defines a manifold in the state space which is an approximation of a SIM. A simplified model is given by the differential equations that describe the dynamics of the slow state variables together with these algebraic constraints for the fast variables. The zeroth order ZDP, i.e., the condition obtained from setting the first-order time derivatives to zero, is equivalent with the QSSA. The use of higher order derivatives provides higher accuracy of the approximation of the SIM and of the simplified system.

In Chapter 3 we first outline the behavior of systems exhibiting timescale disparities and we describe the theory of ZDP for a biochemical readership. We then apply first-order ZDP (i.e. the case when second derivatives are set to zero) to the mass action kinetic expressions for a reversible reaction and derive an enzyme kinetic rate law. We show that under conditions of realistic yet low substrate concentrations, the Michaelis–Menten rate expression may readily be off by more than a factor of four when predicting the variation of concentrations with time while our first-order ZDP based kinetics gives a much more accurate description of the true enzyme dynamics.

We then use ZDP to explore the dynamics in a model of the phosphotransferase system (PTS) involved in glucose uptake, metabolism, and signalling in enteric bacteria. This model has nine state variables. We find that this system can be understood accurately in terms of the behavior of a single state variable with the remaining eight constrained to follow that behavior, a clear example of an extensive simplification of a complex system without loss of understanding. The one-dimensional SIM in this model was given by eight algebraic equations which were hard to solve analytically and hence we determined the SIM numerically. We also employed this manifold to integrate numerically over the reduced model and in this way reduce the simulation time by a factor between 5 and 25, depending on the initial condition, as compared to integrating with a standard stiff integrator over the full model. The algorithm that we developed for the numerical calculation of approximations of SIMs based on ZDP, which uses Newton’s method to solve for points on the manifolds over a grid, is described in Chapter 4 where also the Mathematica code is provided.

In Chapter 5 we further explore the slow dynamics of the PTS by investigating the values of the reaction rates along the manifolds. We find that the rates in this network exhibit a behavior which is much simpler than it could have been: the rates are partitioned into collectives within each of which all rates assume the same magnitude during the slow phase, while rates belonging to different collectives assume different magnitudes. Such relaxation behavior of the rates was observed also upon a large variety of perturbations of the parameter set albeit other constitutions of the collectives was obtained. These results indicate that grouping of the rates is likely to be encountered for different environmental conditions and, also, in other signal transduction pathways possessing structures similar to that of the PTS. Moreover, the difference between the constitutions of the collectives suggests that biological evolution may have involved this behavior in the regulation of cell function. The observed behavior also suggests new reduction procedures that utilize inherent, possibly functional sim-
plicity of the system and calls for further investigations in order to understand the dynamics in more depth.

In summary, the simplification approaches taken in this thesis facilitate the handling of the realistic, complex models in systems biology in several different ways. One issue that we have approached is the need of new rate laws to describe the cellular reaction mechanisms for which traditional enzyme kinetics is inappropriate. We derived a rate law based on ZDP for a reversible enzyme-catalyzed reaction which describes the reaction kinetics accurately also in cases of low substrate concentrations when the traditional enzyme kinetics fails to describe the dynamics; rate laws for other mechanisms which hold under similar conditions may be derived by the same principle. We also addressed the issue of long computation times. In particular, we showed that for the model of the PTS, we reduced the times it takes to integrate over the model, substantially. Last, and perhaps most importantly, we demonstrated that simplified models may contribute to further insight into the unintuitive dynamics of complex biochemical systems: we explored the behavior on different timescales in the PTS and we found that the rates in this pathway exhibit an interesting relaxation behavior on the slow timescale.