Introduction

While computer power is becoming ever more increasing, there are many problems in physics, chemistry, biology—like climate modeling, protein folding etc.—where the length and time scales of interest remain entirely beyond the capacity currently available, and will remain out of reach in the foreseeable future. As a consequence, there is an increasing need for simplified, reduced–order descriptions. Reduced–order models may provide insight and numerical simulations for larger length scales and longer time scales, but of course at the cost of discarding some level of detail. Instead of simply neglecting some degrees of freedom, one is rather interested in reduced models that incorporate into their dynamical behavior the effective influence of the neglected coordinates. There a two basic approaches for the reduction of complexity, one is based on *elimination* like, e.g., elimination of fast degrees of freedom [3], or stochastic approches like Brownian and Langevin models [87], and the other one is based on *remodeling* like, e.g., base pair and rod models [58] in the biomolecular context. The design of new reduced model systems is a growing field of research. In order to verify the approximation quality of reduced models or even to automatically construct reduced model systems the numerical approximation of essential features of dynamical systems becomes an important task.

Recently, efficient techniques for the numerical approximation of the essential statistical behavior of deterministic and stochastic dynamical systems have been developed [13, 14, 69]. They are based on the fact that, when modeling the overall dynamics in terms of some transfer operator, certain features of the dynamics are related to its eigenvalues on and close to the unit circle, and can be identified by exploiting the corresponding eigenfunctions. A thereon based strategy has first been proposed by Dellnitz and Junge [13] to analyze almost invariant subsets, attractors and (almost) cyclic behavior of discrete deterministic dynamical systems subject to small additive noise. It has been successfully applied to examine metastable behavior of deterministic Hamiltonian systems by Deuflhard et al. [14]. Although the numerical results of the latter approach were intriguing and seemed to catch the essential features of the molecular system, the deterministic Hamiltonian model appeared to be unsatisfactory for both theoretical discrepancies and computational complexity [72]. Guided by concepts of statistical physics and Monte Carlo techniques, Schütte et al. introduced in [68, 69] a substantially remodeled stochastic Hamiltonian model. It is based on a special discrete–time Markov process that can be understood as a Hamiltonian systems with randomized momenta. Its reliable application to biomolecular systems is demonstrated in [35, 69].

Modeling, theory and numerics presented herein are motivated by the

successful study of biomolecular systems within the stochastic Hamiltonian model and the desire to extend the concepts to a broader class of systems. It is based on a series of preceding studies, where we investigated in detail the Hamiltonian system with randomized momenta [35, 72] and subsequently extended the approach to other molecular systems [34, 71, 72]. This thesis presents a unified and extended transfer operator based approach to metastability of general Markovian systems. It addresses the question of how modeling, theory and algorithmic aspects should be generalized from the weakly perturbed deterministic and from the Hamiltonian setting to the class of Markovian systems. We contribute a profound analysis of metastability and a theoretical justification of the algorithmic strategy for the identification of metastable subsets. This is achieved by combining results on Markov processes by Meyn & Tweedie [52], Markov operators by Lasota & Mackey [46], Markov semigroups by Davies [9, 10] and Singleton [73, 74], and (weakly) compact operators by Weis [79, 80, 81, 82, 83]. As a consequence, we are able to establish new links between spectral properties of transfer operators and well–established Doeblin and ergodicity conditions on Markov processes and operators. This turns out to be particularly advantageous when aiming at a theoretical justification of the algorithmic approach for new model systems, as we consider herein. This thesis investigates for the first time the essential statistical behavior of the Langevin and the Smoluchowski equation in comparison with the Hamiltonian systems with randomized momenta. All in all, the transfer operator based approach to metastability has proven to be very powerful. Its application to the small biomolecule r(ACC), in comparison with other techniques to study biomolecular conformations, is documented in [35].

Most applications of molecular dynamics are in the context of thermodynamics, not only because most experiments measure thermodynamic quantities, but also since most biomolecular processes can only be understood within a thermodynamical context. In the macroscopic theory of equilibrium thermodynamics, the so–called canonical ensemble describes the distribution of microscopic systems under the condition of constant temperature, volume and number of particles [30]. The canonical ensemble is stationary and hence does not change in time. But at the same time each microscopic single systemevolves in time, causing internal fluctuations within the ensemble. The characterization and identification of the most relevant fluctuations is of main interest. From a biochemical point of view, these fluctuations are related to the conformational dynamics of a biomolecule. In this setting, a conformation describes a metastable global state of the molecule, in which the large scale geometric structure is understood to be conserved, whereas on smaller scales the molecule may well vibrate, oscillate or deform (see Figure 1). As a consequence, we model conformations as metastable subsets

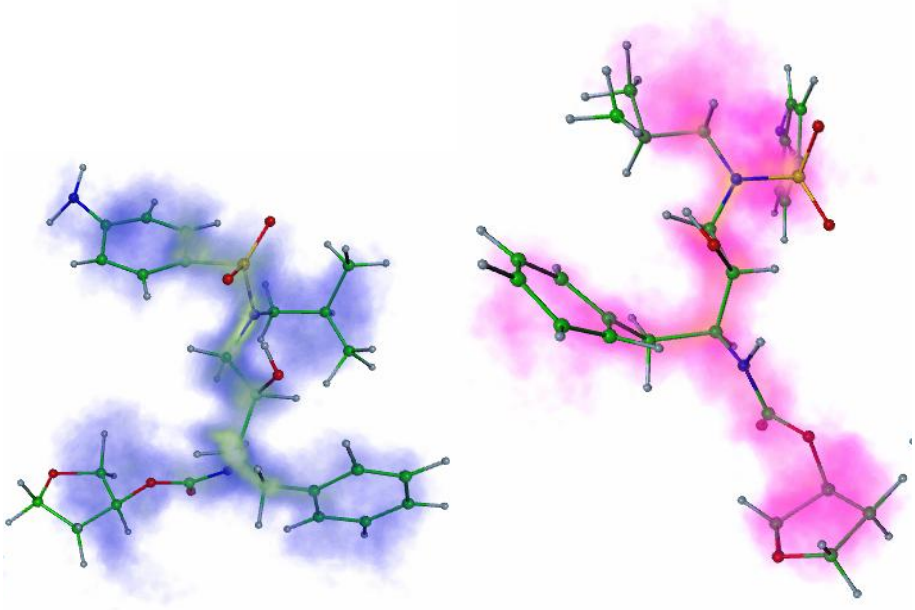

Figure 1: Visualization of two different conformations of a small biomolecule. The so–called ball and stick representations (configurations) correspond to two different conformations of the molecule. The surrounding densities indicate the flexibility within each conformation. Visualization by amira [42].

of the state space and hence include geometric as well as dynamical properties of the system. Both, conformations and their dynamics are expected to gain further insight into the nature of biomolecules and their influence in biochemical reactions.

From a theoretical point of view, the analysis of metastability permits a statistical description of the essential behavior of dynamical systems. The identification of metastable subsets is based on the following idea: Describe internal fluctuations within the invariant distribution by means of a transfer operator defined in terms of the dynamical system. Then the state space can be decomposed into metastable subsets and the essential statistical behavior can be identified by exploiting eigenfunctions of the transfer operator corresponding to eigenvalues close to 1 [13, 17, 69]. Following [72], we give a new theoretical justification of the algorithmic approach in terms of a simple and intriguing relation between the existence of metastable subsets and eigenvalues close to 1 (see Theorem 3.1). We want to emphasize that metastability, as considered herein, is defined w.r.t. some fixed invariant distribution, which in the biomolecular application context is given by the canonical ensemble. This might differ from other approaches to metastability, e.g., the approach via exit times.

The identification strategy requires two particular conditions on the

transfer operator in order to be theoretically justifiable and numerically applicable. Stated in terms of the spectrum of the transfer operator these are (i) the essential spectral radius is less than 1 and (ii) the eigenvalue 1 is simple and dominant (see Sec. 3.2 for a discussion). Within the stochastic Hamiltonian context, Schütte proved in [68] that these two conditions can be reduced to a property of the Hamiltonian flow, which he called momentum–invertibility, and some mixing condition on the dynamics. Having in mind a generalization to Markovian systems, we have to look for alternative conditions independent of particular properties of special model systems. Here the so–called stochastic transition function—a family of probability measures—will play a key role. On the one hand it uniquely characterizes the Markov process representing the microscopic dynamics of single systems. On the other hand it defines the transfer operator modeling the macroscopic evolution of ensembles. In a first step, we are going to relate the two spectral conditions—which are purely functional analytically—to properties of the stochastic transition function. Then, in the second step, these properties will be transformed into more probabilistic conditions on the transfer operator or the Markov process. This enables us to exploit the rich and powerful literature on Markov operators (e.g., Lasota & Mackey [46]) and Markov processes (e.g, Meyn & Tweedie [52]). As a result, we combine results from either of the mathematical theories like, e.g., the fundamental Theorem 4.13, which relates a bound on the essential spectral radius, uniform constrictiveness and the Doeblin–condition.

When dealing with transfer operators, we have to specify the space of functions, the operator is regarded to act on. In the stochastic Hamiltonian approach [68] Schütte considered a weighted Hilbert space of square integrable functions L^2. For the general Markovian setting, however, the natural space is a weighted Banach space of integrable functions L^1 that includes all probability densities on the state space. Thus, a particular emphasis lies on a detailed analysis of transfer operators acting on L^1. Nevertheless, we do study transfer operators on L^2 for the special class of reversible Markov processes. Reversibility describes the property that the Markov process and its time–reversed counterpart are statistically the same. It has the advantageous consequence that then the transfer operator is self–adjoint in L^2.

Within this extended transfer operator based approach to metastability, we analyze theoretically as well as numerically four Markovian systems for molecular dynamics: the deterministic Hamiltonian system, the Hamiltonian system with randomized momenta, the Langevin and the Smoluchowski equation. For the first time we investigate the essential statistical behavior of the Langevin and the Smoluchowski equation in comparison with the Hamiltonian system with randomized momenta. The numerical results give detailed insight into the model systems and prove the transfer operator based

approach to metastability as very powerful.

Acknowledgment. It is a pleasure to thank all those, who supported and encouraged me during the last three and a half years of research. First of all, Christof Schütte and Peter Deuflhard for offering a stimulating research atmosphere, continuous support and for their immense patience. It is their merit to have guided me towards the fascinating field of interdisciplinary research. Many thanks go to the staff at Konrad–Zuse–Zentrum (ZIB), where I spent the first two and a half years during my Ph.D. work. I enjoyed the many discussions, helpful comments and mathematical criticism, whose originated many times from a small cup of espresso in the kitchen. In particular, I would like to thank Frank Cordes, Tobias Galliat, Peter Nettesheim, Daniel Runge, Johannes Schmidt–Ehrenberg and Lin Zschiedrich. At the mathematical institute of the Free University Berlin, where I moved to a year ago, I appreciated the stimulating mixture of teaching and research. Special thanks to Alexander Fischer for fruitful discussions and questioning theory. Finally, I would like to thank Andrew Stuart for illuminating discussions about Markov processes and Dirk Werner for functional analytical advise. Last but for sure not least, I would like to thank Katja and Rosa for simply being there and showing how wonderful live can be.