

1. Introduction

This thesis explores different routes to model reduction in the context of classical molecular dynamics. Adopting a reductionist’s point of view, our main concern is the sound causal explanation of observable macroscopic properties, e.g., activation energies or dynamical stability of conformations in terms of the microscopic physical model. Notwithstanding computational aspects, the difficulty lies in the sheer complexity of the microscopic models with their vastly different spatial and temporal scales.

Roughly speaking, reduced modelling comes in two varieties: elimination of specific (e.g., fast) degrees of freedom (also: *modes*) from the original model, or parametrization of certain simplified models. The first approach is usually referred to as *mode reduction*, whereas the latter is often termed *remodelling*. In this thesis we mainly focus on mode reduction, since sticking to the original microscopic model means to keep as much of the problem’s physics as possible. (The microscopic models are based on profound physical and chemical knowledge, both theoretical considerations and experimental data, which constitutes their empirical adequacy and predictive power.) In doing so, we extend and refine available methods of mode reduction such as averaging or projection operator techniques and set them in context with one another. Nevertheless we also allude to aspects of parametrized models.

Model reduction, as it is understood here, relies on the knowledge of a possibly multidimensional *reaction coordinate*. Of course the problem of finding good (i.e., physically meaningful) reaction coordinates is highly sensitive to boundary conditions such as temperature or pressure and hence cannot be addressed without referring to a specific situation. However for medium-sized molecules such as polypeptides or Lennard-Jones clusters there are often few natural candidates for good reaction coordinates, e.g., torsion angles or radii of gyration. Difficulties arise, if solvent effects play a role, for then the molecules’ configuration space has to be extended to incorporate the solvent which may lead to systems with varying number of particles. However we do not address the reaction coordinate problem here.

Central paradigm in reduced modelling: free energy Free energy is arguably one of the most important and prevalent concepts in molecular dynamics (see the reviews [1, 2]). According to the general view, free energy describes the tendency of a molecular system to associate and react. By definition free energy encodes statistical information about such activated processes, provided the reaction coordinate is suitably chosen. Hence statistical equilibrium properties such as conformational weights can be expressed in terms of free energy. It is less obvious that also many dynamical properties such as transition rates are related to a particular variant of free energy, as has been pointed out on various occasions, e.g., [3, 4, 5]; for the original works on transition state theory we refer to the papers of Eyring [6] and Wigner [7]. Interestingly enough, we find that this specific free energy also appears as an effective molecular potential in most of the reduced models, which casts it a fundamental dynamical concept. For reasons that will become clear below, we term this second type of free energy *geometric free energy*, whereas the first one (which reflects the equilibrium statistics) will be referred to as *standard free energy*.

Literature and previous developments The reader may believe that giving a complete overview of the relevant free energy literature is hopeless. During the last few years progress has been made towards algorithms that sample standard free energy

profiles from their derivatives using constrained simulations, exploiting the dichotomy of *free energy as the potential of mean force*; the free energy is recovered afterwards by numerical integration. This method is known as *Thermodynamic Integration* and goes back to Kirkwood [8]. The idea of relating the derivative of the free energy to the averaged force of constraint appears in the work of Mülders *et al.* [9] for the first time; however these authors derive a wrong expression emanating from a wrong definition of the conditional expectation. Correct expressions have been established in, e.g., [10, 11, 12, 13, 14]. Most of these authors omit the problem of sampling the respective conditional expectation which is crucial for actual computations; recent work in that direction is [15, 16, 17]. Later articles by Vanden-Eijnden and co-workers [13, 5] address the problem of relations among different the free energy definitions, geometric and standard free energy. Current work that exploits this relations for the development of efficient algorithms has been done by the author [18].

Mode elimination techniques for ordinary or stochastic differential equations with multiple scales have a long lasting tradition in celestial mechanics, especially in the Russian literature, e.g., [19, 20], but also in the climate modelling community; for instance, see the proceedings [21]. In celestial mechanics typical problems boil down to finding an appropriate set of action-angle variables [22], whereas climate problems are often described by stochastic differential equations with slow and fast variables [23]. In either case the reduced models are obtained upon averaging over the random perturbations induced by the fast degrees of freedom; the relevant reference regarding the Averaging Principle is the textbook by Freidlin and Wentzell [24]. (A good and systematic overview of the current multiscale literature can be found in the review by Givon *et al.* [25]). Instances of the just mentioned averaging methods are rare in molecular dynamics, however. One such case that is studied by Bornemann & Schütte [26, 27] or Reich [28, 29] is the elimination of fast bond vibrations by introducing holonomic constraints (rigid bond approximations). Another current example that is treated in Yanao *et al.* [30] is the dynamics of gyration radii as collective variables in Hamiltonian system; yet these authors pursue a purely deterministic approach which is more in the spirit of De Leon *et al.* [31] or Uzer *et al.* [32]. The only approach known to the author that addresses stochastic dynamics is a projection operator type method in E & Vanden-Eijnden [13]. For such systems it may happen that the averaged dynamics is trivial on the typical observation time scale, whereas relevant effects appear on longer time scales only. In this case (which, however, does not appear in the examples considered by us) averaging theorems on diverging time intervals come into play which were originally stated by Khas'minskii [33]; see also the recent article [34]. The related problem of large deviations from the averaged equations (for example, if the unresolved system contains essential barriers) is considered in, e.g., [35].

A significant part of the model reduction literature deals with simplified parametrized models, typically linear differential equations. Examples involve rigid base or rigid base-pair models in DNA modelling [36], diffusion models for protein folding [37], or stochastic differential equations that are coupled to Hidden Markov Models [38, 39]. Related work on nonlinear Langevin equations which is based on a Maximum-Likelihood principle is [40]. We abstain from presenting an exhaustive list of references and instead refer to the bibliography in [41]. For the sake of completeness we also mention the equation-free approach that has been developed in [42].

Quite often model reduction is also understood in the sense of clustering or state space decomposition. Most of these methods aim at classifying an essential subspace onto which the full molecular time series is projected. By this, one obtains a dimension-

reduced time series which is easier to analyze. Methods of this kind are known, e.g., by the name of Principal Component Analysis [43, 44] or Proper Orthogonal Decomposition [45]. Yet another promising approach that allows for the identification of essential subspaces that are dynamically relevant is the transfer operator approach by Schütte *et al.* [46]; see also the book of Weber [47]. Essential subspace techniques can be easily linked with methods of mode reduction. For example, a popular approach in the optimal control community (e.g., see [48]) is to truncate the modes orthogonal to the essential subspace, a method known by the name of *Galerkin projection*. Applications to molecular problems are not known to us though.

Another important part of the literature is concerned with projection operator techniques that have been established by Mori [49] and Zwanzig [50] in the context of non-equilibrium statistical mechanics. (See also the review by Hynes [51] or the monograph by Evan & Morriss [52].) In molecular applications such methods amount to the derivation of the (non-Markovian) *generalized Langevin equation*. Examples can be found in [53, 54]; regarding a systematic study of the projection operator ansatz for molecular problems we refer to [55]. Contemporary mathematical works concern closure schemes for the generalized Langevin equation [56], Markovian approximations [57], or issues related to existence and uniqueness [58]. A good overview can be found in the new textbook by Chorin & Hald [59]. Related problems such as applications to Kac-Zwanzig heat bath models are discussed in [60]. A Markovian variant of the projection operator approach, which can be regarded as least-square approximation in some suitably defined function space, is called *optimal prediction*. Typically optimal prediction is applied to problems involving partial differential equations, for which the method works quite well (e.g., Burgers' equation [61], Korteweg-deVries-Burgers equation [62]). However the application to deterministic Hamiltonian systems yields rather poor results as has been repeatedly demonstrated, for example, by Hald & Kupferman [63] or Chorin *et al.* [64], and instances of stochastic Hamiltonian systems are not known to us. The only molecular dynamics application we are aware of is in the article of Seibold [65]; however therein the author mainly focuses on low temperature asymptotics and aspects of computational efficiency.

Issues addressed in this thesis This thesis deals with very different aspects of model reduction. The original models range from deterministic mechanical models on the one hand to stochastic differential equations such as Brownian motion or Langevin dynamics on the other hand. Each of these models comes along with its own formalism (covariant formulations in mechanics, Itô calculus for stochastic differential equations, etc.) which makes it difficult to handle all problems within a unifying framework. Moreover many problems in molecular dynamics are of genuinely thermodynamical nature which calls for an appropriate mathematical description of statistical concepts such as free energy. Here we adopt a more geometric language that is common to classical mechanics on manifolds. This may seem unusual, especially for readers that are familiar with stochastic differential equations. But in fact, many problems in molecular dynamics are problems on manifolds that have an interesting underlying geometric structure: constrained dynamics on a configuration submanifold, curvilinear reaction coordinates, and many more. Furthermore the covariant formalism of mechanics allows for straightforward generalization of statistical mechanics problems to curved spaces which is suitable, e.g., for sampling certain probability measures subject to holonomic constraints (cf. Section 4.2).

The general mathematical framework is established in Section 2. We basically

follow the relevant literature on geometric mechanics by Abraham & Marsden [66] and review ideas for Hamiltonian systems with randomized momenta that have been put forward in the work of Schütte [67]. It turns out that the covariant formalism of mechanics easily extends to stochastic differential equations which leads, among others, to a geometric version of Itô’s formula which proves that the Itô stochastic differential equation transforms like a second-order vector field (Lemma 2.11). The basic language of *stochastic differential geometry*, on which our considerations are based, is developed in the books by Stroock [68] and Hsu [69]. By expanding the ideas therein to Langevin processes, we reveal that the Langevin equation has some interesting transformation properties as compared to general hypo-elliptic diffusion processes which are due to its Hamiltonian origin (Lemma 2.10). In particular we find that the Itô-Stratonovich ambiguity vanishes, if we confine our attention to point transformations. The geometric viewpoint of molecular dynamics thus highlights that such different systems as second-order mechanical systems, first-order Brownian motion and stochastic Langevin equations exhibit common transformation properties. Additionally it gives rise to a physically intuitive and unifying picture for what is called *entropic effects* in conformation dynamics: in case of second-order mechanical systems, these stem from inertial contributions due to the kinetic energy, but they can be likewise explained by the interplay between ordinary diffusion and certain conformational degrees of freedom. In either case these effects are actuated by the underlying Riemannian structure (see Section 5 for some examples).

Section 3.1 tries to shed some light on the different free energy definitions that circulate in the literature; cf. the review [13], and see also the schematic overview in Appendix A. On a purely formal level, Federer’s co-area formula [70] links standard and geometric free energy by relating the underlying conditional probability densities. From a physical point of view, the standard free energy can be expressed as a sum of geometric free energy and an appropriately defined Fixman potential. Neither relation is actually new, but they both have useful practical implications for Thermodynamic Integration algorithms that, to the best of our knowledge, have not been taken into account so far: First of all, we explicate that the famous Blue Moon formula by Carter *et al.* [71] for the conditional expectation is an instance of the co-area formula (Section 3.1.2). The Blue Moon ensemble method is a popular and widely-used technique for the sampling of conditional expectations by means of constrained simulations with an appropriate reweighting strategy. Yet there has been (and still is) some confusion (e.g., see [16, 72, 73]) about whether the weight is affected by the presence of momenta or velocities in the system. But as we will argue below, reweighting is an issue for any type of constrained dynamics — no matter if the system involves momenta or not. Secondly, we demonstrate that geometric free energy can be viewed as the *potential of mean constraint force* (see Section 3.1.1 or Section 4.1 regarding holonomic constraints). By using Thermodynamic Integration it is hence possible to compute the derivative of the geometric free energy by simply averaging over the Lagrange multipliers (forces of constraint) that are explicitly available during the simulations without further function evaluations; cf. also the recent article [17]. Last but not least, the Fixman potential that marks the difference between geometric and standard free energy can be directly computed from constrained simulations without computing second derivatives of the reaction coordinate. This yields a remarkably simple formula for computing standard free energies that does neither require second derivatives nor reweighting à la Blue Moon (see Remark 4.14 or the recent article [18] by the author). To the best of the author’s knowledge all available algorithms do in fact require the calculation of the

reaction coordinate's Hessian; e.g., see [10, 11, 12, 16]. The various contributions to geometric and standard free energy have concise geometric and physical interpretations as is worked out in detail in Section 3.1.1. In Section 3.1.2 we also give an answer to the question in which sense free energy can be understood as a *potential of mean force*: if one takes up the position that force is understood as a differential 1-form, then only the derivative of the geometric free energy qualifies as a (mean) force (viz., the mean constraint force), while the derivative of the standard free energy exhibits additional gauge dependencies with regard to transformations of the reaction coordinate.

Section 3.2 is dedicated to the application of the Averaging Principle to diffusion models, i.e., stochastic differential equations with non-degenerate noise term. If the reaction coordinate is linear in the configuration variables, then the reduced equations describe simple diffusion in the free energy landscape; in this particular case, standard and geometric free energy coincide. If the reaction coordinate is nonlinear, the application of the Averaging Principle requires that we write the equation globally in terms of the resolved and the unresolved coordinates. Not only is this difficult (or even impossible), but it also makes the resulting equations in some respects intransparent. Therefore we take advantage of the fact that the reaction coordinate foliates configuration space and consider only local averages of the dynamics on each leaf, where each leaf is defined by a specific value of the reaction coordinate. The global picture can then be reconstructed by endowing the locally averaged equations with an appropriate Riemannian metric that is induced by the reaction coordinate and is defined for all of its possible values (Section 3.2.2). Although the just described approach is no longer covered by the Averaging Principle, it turns out that the reduced system is again an Itô equation which is covariant under transformations of the reaction coordinate and has a straightforward physical interpretation as a diffusion equation on a Riemannian manifold. In particular the effective potential energy is given by the geometric free energy. We briefly illustrate the method by suitable examples (see Examples 3.10 and 3.11) and discuss its relation to a related approach [13].

For mechanical systems the situation is more complicated, since the equations are essentially second-order, and thus the requirement that the fast dynamics exhibits a unique invariant measure for all values of the slow coordinates is difficult to handle analytically and numerically (see Example 3.12). Therefore we resort to projection operator techniques or least-square approximations such as optimal prediction [64]. Though similar to standard averaging, these methods account for the fact that the equations are second-order. As Chorin & Hald have proved in [56], optimal prediction for Hamiltonian systems leads to reduced models that are again Hamiltonian. Based on considerations therein, we derive a new and simple expression (3.63) for the effective total energy that allows for a lucid interpretation as a mechanical system on a Riemannian manifold which is spanned by the reaction coordinate. As in the Brownian dynamics case, we demonstrate that the effective potential is given by the geometric free energy. More sophisticated projection operator techniques like the Mori-Zwanzig procedure involve the derivation of a *generalized Langevin equation*, which is a suggestive way to rewrite Hamilton's equations as a Langevin-like equation that is formally equivalent [49, 50]. However we emphasize that the equivalence is only formal, for the derivation relies on the tacit (but wrong) assumption that the Hamiltonian system has a unique invariant measure. Although theoretically appealing, the ideas of Mori and Zwanzig have barely any practical relevance for studying complicated molecular processes; here we mention it only for the sake of completeness, while pointing out certain difficulties in connexion with the derivation of the generalized

Langevin equation (see Remark 3.17 and the preceding paragraph).

Finally in Section 3.4, we propose an *ad-hoc* alternative to averaging and optimal prediction that is based on the observation that the unresolved degrees of freedom often have small amplitude and can thus be approximated by harmonic motions. By averaging over these modes one ends up with semi-analytic reduced models (both diffusive and mechanical) that live only on the essential subspace, but still have few free parameters. Since the parametrization involves only the unresolved parts of a molecule, these models are easy to compute at the cost of restricted physical interpretability. Two interesting aspects emerge in connexion with the semi-analytic models: First, the reduced models are relevant in the context of stiff-bond approximations, since they explain how the dynamics is altered, when infinitely stiff bonds are replaced by rigid bonds (i.e., constraints). Work in this direction has been done by Hinch [74] for diffusive systems and by van Kampen & Lodder [75] and Reich [28] for mechanical systems. Although we do not contribute new results regarding rigid-bond approximations, knowledge about practical implications thereof are still not widespread in the molecular dynamics community (e.g., see [76]). Second, there are some interesting relations to adiabatic perturbation problems in mechanics [27]. For instance, it is well-known that averaging problems for small oscillations may suffer from resonances between the oscillators' frequencies [20]. It seems, however, that resonances do not play a role, if the system is appropriately thermalized. Since investigating resonance effects in stochastic Hamiltonian systems or Brownian motion in detail is far beyond the scope of this thesis, we provide only numerical evidence for this claim in Section 3.4.1. However more careful studies would be desirable.

To some extent reduced modelling results in the calculation of (geometric) free energy profiles. In fact there is a bunch of literature that addresses standard free energy calculation by means of Thermodynamic Integration, e.g., [71, 77]. However Thermodynamic Integration proceeds by constrained integration, and it is by no means clear how standard thermostating techniques fit constrained integration. It is striking that the question of how to sample the correct probability measure (constrained Gibbs measure) is typically ignored; e.g., see [10, 11, 78]. In particular there is a lot of confusion in the literature whether constrained Hamiltonian systems inherit fundamental thermodynamical properties from their unconstrained counterparts. For example, it is common sense in the molecular dynamics community that constrained Hamiltonian flows do not preserve phase space volume; e.g., see [73, 72, 79, 80]. Additionally there is an ongoing discussion [16] concerning the impact of so-called *hidden constraints* on the invariant distribution of constrained second-order systems and its relation to first-order systems. (Again this remark alludes to the co-area formula and the problem of Blue Moon reweighting for first-order systems.) We provide the theoretical background regarding constrained mechanical systems in Section 4.1. To this end, we basically review available results from the literature [66, 81]; in particular we adopt an argument in [82] that proves that constrained Hamiltonian systems are symplectic and therefore volume-preserving (Lemma 4.3). Taking advantage of this property, we then construct a novel hybrid Monte-Carlo (HMC) scheme that can be used together with the RATTLE symplectic integrator for constrained Hamiltonian systems. Following an idea in [83], we can prove that the corresponding discrete Markov chain is ergodic with respect to the constrained Gibbs measure on configuration space (Proposition 4.12). Related results for Brownian motion have recently become available in the work of Lelièvre *et al.* [17]. Therein, however, the authors prove ergodicity only for the time-continuous process, while

disregarding discretization issues. In Section 4.2.3 we generalize their results and construct a constrained Langevin dynamics, equation (4.27), that conserves a given holonomic constraint and that preserves the constrained Gibbs measure. (For this purpose we once more take advantage of the similarity between Hamilton and Langevin dynamics, borrowing ideas from index reduction techniques for differential-algebraic equations.) Furthermore we suggest a discretization scheme, equations (4.29)–(4.32), that can be regarded as a stochastic modification of the RATTLE algorithm. We should mention that an almost identical algorithm has been published by Vanden-Eijnden & Ciccotti [84] during the course of this thesis, where the authors could even prove that the algorithm is second-order accurate. However the article does not address issues of invariant measures and constrained probability distributions. We conclude with Section 4.3 by discussing the application of the various sampling schemes to the calculation of free energy profiles. We especially propose a novel algorithm that does neither require Blue Moon reweighting of the expectation values nor calculating second derivatives of the reaction coordinate (see Remark 4.14).

We illustrate the reduction schemes as well as the constrained hybrid Monte-Carlo sampling by means of several examples in Section 5. Both averaged Brownian motion and optimal prediction perform remarkably well in terms of dynamical observables such as transition rates or decay of correlations. (Especially for the latter approach this comes rather unexpected as optimal prediction for deterministic Hamiltonian system is known to yield fairly poor results; e.g., see Chorin *et al.* [64].) Moreover optimal prediction reveals an interesting (and yet unknown) physical mechanism that explains the backbone dynamics of a chain-like molecule: for *n*-butane, for example, we observe that the angular kinetic energy favours the *trans* conformation, which is characterized by a rather slim shape with respect to the principal axis of inertia and which should be contrasted with the bulky *cis* conformations. *Prima facie* this seems counter-intuitive, since one could expect that the mass distribution of a rotating molecule tends to spread out due to centrifugal forces. However here the situation is different, for the backbone rotation is an internal motion of the molecule. Since the kinetic energy tends to stabilize the more compact *trans* conformation by slightly increasing the total energy of the *cis* conformations, we term the induced force *internal centripetal force*. The same rotation mechanism explains the different conformational stabilities of the glycine dipeptide analogue, for which we study free energy landscapes and the optimal prediction Hamiltonian along the two central backbone angles: also here the kinetic energy stabilizes the extended C5 conformations by slightly lowering their total energy as compared to the bulky C7 conformations. Our calculations also reveal that the kinetic energy preserves the molecular potential’s symmetry under parity transformations in the Ramachandran plane, but exhibits an even higher symmetry itself: the matrix elements of the effective inverse metric are (approximately) invariant under reflections of the two backbone angles independently, where the slight perturbation of the symmetry reflects the non-uniform mass distribution along the peptide’s backbone. To the best of the author’s knowledge symmetry-breaking effects of the peptide backbones’ mass-distribution have not been studied so far, and more careful studies would be desirable.

Finally, we survey known results from the literature that deal with corrections to the Averaging Principle for non-degenerate stochastic differential equations. Problems involve moderate [85, 86] and large deviations [24, 35] or deviations on long time scales [33, 34]. We mention them for the sake of completeness and encourage their application to molecular dynamics problems in the future.

Summary of the main achievements

- The transformation properties of mechanical systems, non-degenerate diffusion equations (Brownian motion) and hypo-elliptic Langevin equations are studied. Restricting our attention to point transformations, it is a common feature of all such systems that they transform as second-order vector fields. In case of the diffusion equation, this property reflects the well-known Itô formula. For Langevin equations the Itô-Stratonovich ambiguity vanishes.
- The various contributions to standard and geometric free energy have concise geometric and physical interpretations. In particular we reveal that the geometric free energy is the potential of mean constraint force. As a by-product we provide an alternative version of the famous Blue Moon ensemble method that turns out to be an instance of Federer’s co-area formula. It therefore also applies to constrained first-order systems such as Brownian motion.
- For both averaging and optimal prediction we find that the geometric free energy appears in the reduced equations as an effective potential, which casts it a fundamental dynamical quantity. For all practical purposes the optimal prediction Hamiltonian can be approximated by a sum of kinetic and potential energy (i.e., geometric free energy), where the kinetic energy is defined with respect to an averaged Riemannian metric that is induced by the reaction coordinate.
- Hybrid Monte-Carlo (HMC) for constrained mechanical systems is a novel algorithm for sampling constrained Gibbs measures and free energy profiles. We prove a Law of Large Numbers for the time-discrete HMC Markov chain that holds for any stable step-size. Exploiting the close relationship between Hamilton and Langevin equations, we derive a constrained version of the Langevin dynamics that preserves the constrained canonical distribution and allows for calculating free energy profiles without reweighting or computing second derivatives.
- The performance of the different reduction schemes is demonstrated by means of two molecular examples: *n*-butane and the glycine dipeptide analogue. The reduced systems reproduce essential dynamical observables such as correlations or transition probabilities. Even more important, the models reveal a common rotation mechanism for the molecules’ conformational dynamics that can be explained by the interplay between geometric free energy and the (extrinsic) geometry of the reaction coordinate.

Some matter of notation We will make extended use of abstract index notations. Often we will use lowercase Greek indices α, β, γ and Latin indices i, j, k to distinguish between different types of coordinates (e.g., resolved and unresolved coordinates) with respect to an unspecified basis. Note, however, that this distinction is sometimes relaxed; then we use Latin indices h, l, m, n to label arbitrary coordinates. Moreover we use Einstein’s summation convention, that is, we sum over double upper and lower indices, where the range of the respective indices should be clear from the context. Using a particular coordinate system, for example, polar coordinates (r, φ, ϑ) we may also write g^{12} or $g^{r\varphi}$ to denote the $(1, 2)$ component of a contravariant tensor g (e.g., the inverse metric tensor). The reader should be aware of some other abusive notations that are common in the physical literature. For example, we will use the dot (time derivative) to denote tangent vectors. That is, if $q = (q^1, \dots, q^n)$ denotes coordinates on a configuration manifold Q , then we write (\dot{q}, \dot{q}) to denote the respective coordinates

on the tangent space TQ . We employ so-called mass-scaled coordinates which allows us to set the molecular mass to unity throughout this thesis. This is convenient, for it considerably simplifies the notation, and it allows us to identify tangent and cotangent space in the sense that $\dot{q} = p$. In doing so the velocity vector \dot{q} on the left is an element of the tangent space T_qQ , whereas the momentum on right hand side is from the dual space T_q^*Q . In general this identification will be procured by the metric tensor g on Q , but in the Euclidean case we will exploit this identification without further comment.

The various types of differential operators appearing in the text may seem a little confusing. Often it is important to distinguish between derivatives that are denoted by df or ∇f , and the gradient of a function f that is written as $\text{grad } f$ (the latter is a vector field, whereas the former denotes a 1-form). For vector fields X, Y we will sometimes use the symbols $\nabla_X Y$ for the covariant derivative between vector fields, or $dY(X) = \nabla Y \cdot X$ to denote the directional derivative of a vector field along a vector. Moreover the bold face symbol $\mathbf{D}f$ means the Jacobian of a vector-valued function f . We may also use the notation $\mathbf{D}_1 f(\cdot, \cdot)$ or $\mathbf{D}_2 f(\cdot, \cdot)$ to indicate derivatives with respect to the first or second slot (e.g., slow and fast coordinates) of a function.

Acknowledgments It is my pleasure to thank all those who supported and encouraged this thesis: First of all, my advisor, Christof Schütte, for directing my dissertation research, pointing out many interesting mathematical questions, giving constant encouragement and steady support. Eric Vanden-Eijnden, for providing many insightful remarks about free energy concepts that were invaluable for this work and for inviting me to the Courant Institute and UC Berkeley. Special thanks goes to Tony Lelièvre, Pete Kramer and John Maddocks for fruitful discussions on problems with constraints, and for persistently asking further questions. I also enjoyed the vibrant criticism by Giovanni Ciccotti (not to mention the excellent dinners).

I have appreciated the stimulating interdisciplinary research atmosphere at Freie Universität Berlin. In particular I very much thank my colleagues: Wilhelm Huisinga for introducing me to molecular dynamics, Jessika Walter and Burkhard Schmidt for stimulating discussions about my work (and their work, too), Tobias Jahnke, Oliver Sander and Heiko Berninger for any kind of sideline mathematics, and Illia Horenko for regularly joining me at the cafeteria. I am moreover grateful to Tim Conrad, Alexander Fischer and Eike Meerbach for helping me with the computer simulations.

Finally, this thesis would not be on hand without those persons, my family and friends, who accompanied me during the last few years. I am deeply indebted to my parents for backing me all the years, granting unrestricted freedom to follow my interests, and to Gebke for just being there.

The work was funded by the DFG Research Center MATHEON "Mathematics for Key Technologies" (FZT86) in Berlin.