Konrad-Zuse-Zentrum für Informationstechnik Berlin
Takustr. 7, D-4195 Berlin-Dahlem

P. Deuflhard M. Dellnitz[*,†] O. Junge[†] Ch. Schütte

Computation of Essential Molecular Dynamics by Subdivision Techniques I: Basic Concept

[†] Mathematisches Institut
Universität Bayreuth
95440 Bayreuth

[*] Konrad–Zuse–Fellow

Preprint SC 96–45 (December 1996)

Computation of Essential Molecular Dynamics by Subdivision Techniques I: Basic Concept

P. Deuflhard, M. Dellnitz*, O. Junge*, and Ch. Schütte

Abstract

The paper presents the concept of a new type of algorithm for the numerical computation of what the authors call the *essential dynamics* of molecular systems. Mathematically speaking, such systems are described by Hamiltonian differential equations. In the bulk of applications, individual trajectories are of no specific interest. Rather, time averages of physical observables or relaxation times of conformational changes need to be actually computed. In the language of dynamical systems, such information is contained in the natural invariant measure (infinite relaxation time) or in almost invariant sets ("large" finite relaxation times). The paper suggests the direct computation of these objects via eigenmodes of the associated Frobenius-Perron operator by means of a multilevel subdivision algorithm. The advocated approach is different to both Monte-Carlo techniques on the one hand and long term trajectory simulation on the other hand: in our setup long term trajectories are replaced by short term sub-trajectories, Monte-Carlo techniques are just structurally connected via the underlying Frobenius-Perron theory. Numerical experiments with a first version of our suggested algorithm are included to illustrate certain distinguishing properties. A more advanced version of the algorithm will be presented in a second part of this paper.

*Research partly supported by the Deutsche Forschungsgemeinschaft under Grant De 448/5-2

Contents

1 Introduction

Reliable modelling and simulation of molecular processes is one of the really grand challenges of today. Computational answers to chemical and biochemical questions are as important for industry as they are hard to get.
The classical microscopic description of molecular processes leads to a mathematical model in terms of Hamiltonian differential systems. Discretization of such systems allows, in principle, a simulation of the dynamics. However, both forward and backward analysis of numerical discretizations restrict such simulations to only short time spans and comparatively small discretization time steps. Fortunately, most questions of chemical relevance just require the computation of averages of physical observables. For the computation of such averages several algorithmic approaches are popular at present:

(i) time averages computed by “large” discretization steps — physically motivated via the physical ergodicity hypothesis,

(ii) ensemble averages via Monte Carlo methods — physically motivated by prescribed canonical ensembles which model a heat bath embedding of the molecule.

In the present paper, we advocate a new computational approach on the basis of the mathematical theory of dynamical systems. We directly discretize the eigenvalue problem of the Frobenius–Perron operator, which is associated with any dynamical system. Without any physical a-priori assumptions we thus are able to compute:

(i) the associated (natural) *invariant measure* (corresponding to the eigenvalue $\lambda = 1$), which determines the time averages of any physical observable,

(ii) almost invariant sets (for eingenvalues $\lambda \approx 1$), which correspond to the *essential dynamics* of the molecular system.

In what follows, we first examine the problem of MD simulation from the points of view of Numerical Mathematics (Sec. 2.1), of Statistical Physics (Sec. 2.2), and of Dynamical Systems (Sec. 2.3). In Section 3 we work out that conformations of molecular systems mathematically correspond to almost invariant sets — which, in turn, are related to eigenmeasures of the Frobenius–Perron operator for $\lambda \approx 1$. The discretization of the eigenvalue problem is a modification of techniques recently developed for *hyperbolic* systems [7, 9, 8] — see Section 4. This discretization requires short term trajectory simulations only. Preliminary numerical experiments are presented in Section 5. A more advanced version of the algorithm will be presented in a second part of this paper.

2 Time Averages versus Ensemble Averages

In classical MD (cf. textbook [1]) a molecule is modelled as a collection of classical mass points with masses m_k, positions $q_k \in \mathbb{R}^3$, and momenta $p_k \in \mathbb{R}^3$, $k = 1, \ldots, N$. The interaction of these mass points is characterized by a Hamiltonian function

$$H(q,p) \,=\, \frac{1}{2}\, p^T M^{-1} p \,+\, V(q),$$

where $q = (q_1^T, \ldots, q_N^T)$, $p = (p_1^T, \ldots, p_N^T)$, $M = \mathrm{diag}(m_1, \ldots, m_N)$, and a differentiable potential V. The Hamiltonian H is defined on the phase space $\Gamma \subset \mathbb{R}^{6N}$. The corresponding canonical equations of motion

$$\begin{array}{rcl} \dot{q} & = & M^{-1} p \\ \dot{p} & = & -\mathrm{grad}\, V \end{array} \tag{2.1}$$

describe the dynamics of the molecule. The formal solution of (2.1) with initial state $x_0 = (q(0), p(0))$ is given by $x(t) = (q(t), p(t)) = \Phi^t x_0$, where Φ^t denotes the flow.

2.1 Long Term Trajectory Simulation

Suppose we want to predict the motions of the molecular system by numerical integration of (2.1). This means that we replace the exact trajectories $x(t) = \Phi^t x_0$ by discrete approximations. Thus, assuming one-step discretization for simplicity, the (exact) flow Φ^t is replaced by a *discrete flow* Ψ^τ, so that the discrete solution can be written as

$$x_{k+1} \,=\, \Psi^\tau x_k \quad \Rightarrow \quad x_k \,=\, \left(\Psi^\tau\right)^k x_0,$$

with τ being the applied stepsize (assumed to be constant for the time being) . An important feature of molecular processes is that *long term predictions* are required, which means predictions over periods much longer than the applied time steps. We are therefore led to discuss long term numerical integration of (2.1) in terms of accuracy and stability of the selected discretizations.

Forward Analysis

In this type of analysis, we are interested in the propagation of initial perturbations δx_0 along the flow Φ^t of (2.1), i.e., in the growth of the perturbations $\delta x(t; x_0) = \Phi^t(x_0 + \delta x_0) - \Phi^t x_0$. The *condition number* $\kappa(t)$ is defined as the maximal error propagation factor (cf. textbook [10]), so that, in first order perturbation analysis and with a suitable norm $|\cdot|$, we have

$$|\delta x(t; x_0)| \,<\, \kappa(t) |\delta x_0| \qquad \text{for all } x_0.$$

By definition, the number $\kappa(t)$ characterizes the worst case *analytical* error amplification *independent of any discretization*. Long term accumulation of discretization errors is just a special case.

From this point of view, long term integration is only feasible as long as $\kappa(T)$ is small enough for $0 \leq t \leq T$. For integrable systems (such as the popular Kepler problem) it is known that $\kappa(T) \sim T$ [22], which allows for quite long term simulations. Unfortunately, for real life MD problems, κ is exponentially increasing. As an illustration, test simulations for the Butane molecule are presented in Fig. 1. As can be seen, global error propagation totally spoils any initial information after a time span, which is significantly shorter than time spans of physical interest.

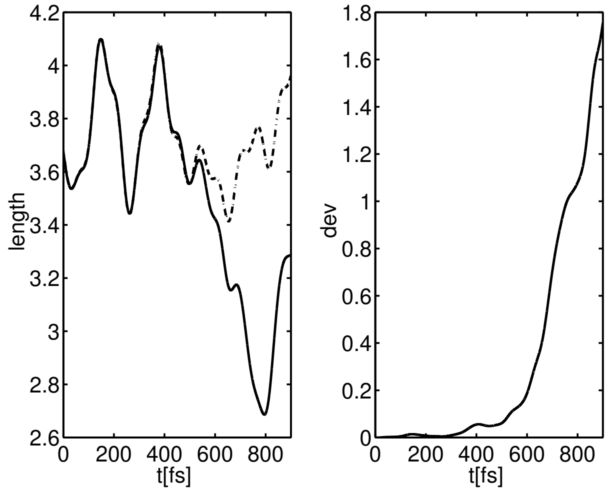

Figure 1: Comparison of two different MD simulations of the Butane molecule (Verlet discretization with stepsize $\tau = 0.005$fs) starting from two nearly identical initial states (initial spatial deviation: 10^{-4}Å) . Left hand figure: Evolution of the length (Å) of the Butane molecule for the two cases. Right hand side: Spatial deviation (Å) of the two trajectories versus time.

Backward Analysis

In this type of analysis, the *discrete* solution is regarded as an exact solution of a perturbed problem. In particular, backward analysis of *symplectic discretizations* of Hamiltonian systems has recently achieved a considerable amount of attention (see [33] and references therein). The most prominent symplectic discretization is the popular Verlet scheme.
These discretizations give rise to the following nice feature: the corresponding discrete solution of a Hamiltonian system is "exponentially close" to the exact solution of a perturbed system that is again Hamiltonian. To be more precise, let $x_k = (\Psi^\tau)^k x_0$ be the discrete solution to (2.1) computed via a symplectic discretization Ψ^τ with order of consistency p and stepsize τ. Then, the perturbed Hamiltonian is [20, 2]

$$\tilde{H} \;=\; H \;+\; \sum_{k=0}^{N} \tau^{p+k}\, H_k \tag{2.2}$$

where the components H_k are composed of derivatives of H up to order k. A nice consequence is the fact that the discrete solution nearly conserves the Hamiltonian $\tilde{H}$ and, thus, conserves H up to $\mathcal{O}(\tau^p)$. This is the reason for the superior long–term energy conservation property of symplectic integrators in MD applications.

Closer inspection reveals the following situation: If H is analytic, then the truncation index N in (2.2) is arbitrary. In general however, the series diverges as $N \to \infty$. For the behavior of the solutions the following result holds:

THEOREM 2.1 (HAIRER/LUBICH [21]) *Let H be analytic and $\tilde{x} = \tilde{x}(t)$ the (exact) solution of the perturbed Hamiltonian system corresponding to $\tilde{H}$ with $\tilde{x}(0) = x_0$. There exists some $\tau_* > 0$, so that for all $\tau < \tau_*$ the numerical solution $x_k = (\Psi^\tau)^k x_0$ and the exact solution $\tilde{x}$ of the perturbed system remain exponentially close in the sense that*

$$x_k - \tilde{x}(k\tau) = \mathcal{O}(e^{-1/\tau})$$

over a time interval $T = \mathcal{O}(|\log \tau|/\tau)$, i.e., for all $k\tau < T$.

Thus, closeness can only be guaranteed over finite time spans which decrease with increasing τ. Unfortunately, the theorem does not state how small τ_* might be for a given problem. In fact, as MD simulations show, the critical stepsize τ_* may be several orders of magnitude smaller than desirable stepsizes τ. This is illustrated in Fig. 2 via the time average of the length of a Butane molecule in an MD simulation over 200 picoseconds. We observe $\tau_* \approx 10^{-2}$fs as opposed to typical MD stepsizes of interest such as $\tau \approx 10$fs.

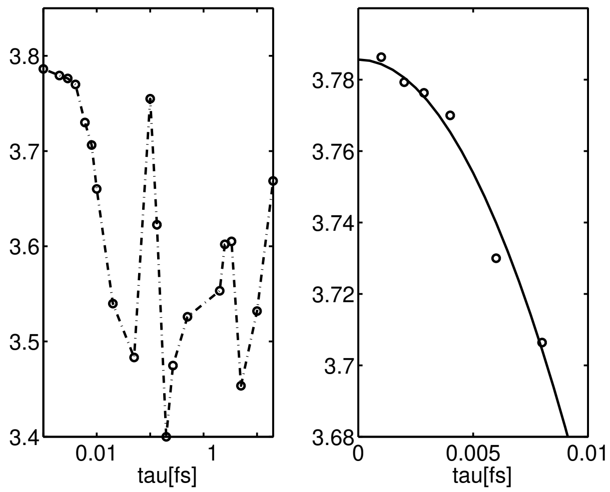

Figure 2: Left hand figure: Dependence of the time average (time scale $T = 200$ps) of the length of a Butane molecule on the stepsize τ of the discretization (symplectic Verlet–discretization with order of consistency $p = 2$). Right hand figure: Zoom of the asymptotic domain ($\tau < 10^{-2}$ fs) and quadratic fit.

Summarizing, both forward and backward analysis lead to the insight that trajectory simulation is appropriate only for short time intervals even with symplectic discretizations.

Essential Dynamics

Fortunately, in most of the applications, details of individual MD trajectories of a molecular system are of only minor importance. We begin with an illustrative example due to GRUBMÜLLER [18] documented in Figure 3. It describes the dynamics of a polymer chain of 100 CH_2 groups. The figure presents six different zoom levels, each of which scales up in time by a factor of 10. On the small time scales (upper levels) the dynamical behavior is characterized by nonlinear *oscillations* around certain "equilibrium positions". On larger and larger time scales these oscillations become more and more unimportant. On the largest time scale (lowest level) we observe an "essential" dynamical behavior as a kind of flip–flop between two "conformations".

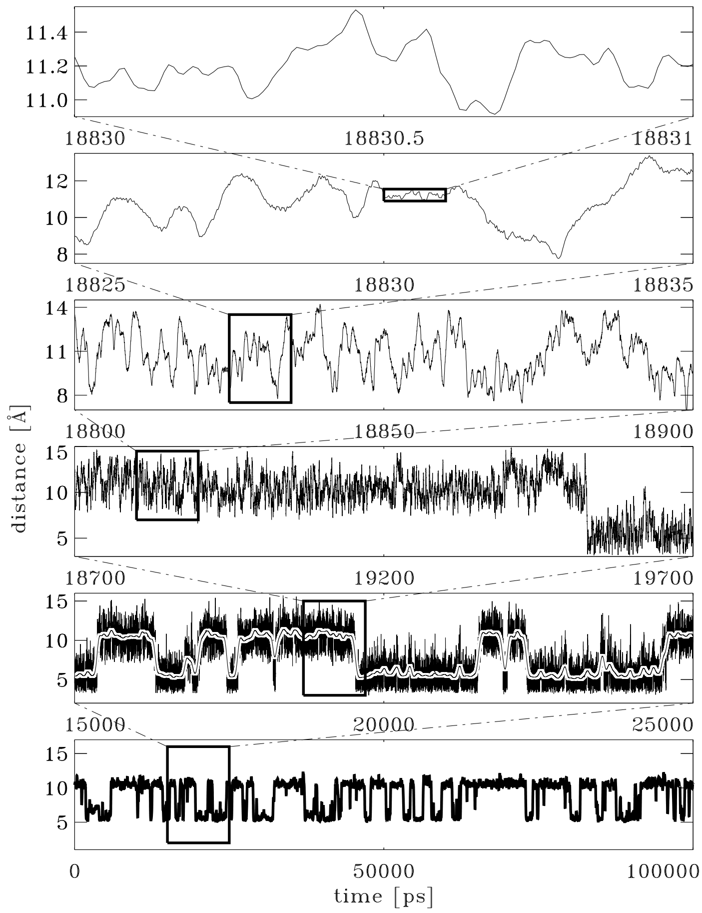

Figure 3: MD simulation of a polymer chain of 100 CH_2 groups. The picture is taken from [18]. It shows the dynamics of the distance between two CH_2–groups (# 12 and # 36). The series of plots illustrates the oscillations of the distance at time scales increasing by a zoom factor of 10 at each level.

At this point it is visible that the essential dynamics of the molecular process could as well be modelled by probabilities describing the average durations of stay within the different conformations of the system. Possible stepsizes ($\tau < 10$fs) for numerical integration are confined by the fast oscillations. Time scales of physical interest range between 10^3 and 10^5 picoseconds, which is a factor $10^5 - 10^7$ larger.
These observations explain why, in MD applications, interest mainly concentrates on computing *statistical* properties of the molecules under consideration such as in time averages of physical observables. Let $A : \Gamma \to \mathbb{R}^d$ denote a physical *observable*, then

its *time average* $\bar{A}$ is given in terms of the flow Φ^t as

$$\bar{A}(x) = \lim_{T\to\infty} \frac{1}{T} \int_0^T A(\Phi^t x)\, dt. \tag{2.3}$$

Recall from Figure 2 that the actual evaluation of $\bar{A}$ will also require extremely short time steps which lead to extraordinarily long computing times for long term averages. More detailed information about the dynamics can be obtained by introducing *running averages* over typical *finite* time scales. Techniques to eliminate the smallest time scales have recently been investigated by homogenization analysis (cf. [3, 34]) or by a statistical representation [31, 16]. In view of Figure 2 the whole MD problem can be seen to be of a *multi-scale nature*, so that mere elimination of the smallest time scales is not enough. In this situation, we herein aim at a different algorithmic approach to compute averages directly by means of well–conditioned subtrajectory simulations only.

2.2 Ensemble Averages in Statistical Physics

In many situations, the theory of "Statistical Mechanics" lights a way to avoid the problem of reliable statistical characterization of the essential dynamics: the statistical properties of the system are directly desribed in terms of the Hamiltonian H without explicit use of the flow Φ^t. Time averages $\bar{A}$ are then replaced by *ensemble averages* of A. Ensembles are defined via given probability measures μ on the phase space Γ. The associated expectation values

$$\langle A\rangle_\mu = \int_\Gamma A(x)\, d\mu(x) \tag{2.4}$$

describe the mean values of A over a statistical ensemble of identically prepared systems with Hamiltonian H. Probability measures on Γ, which are of particular physical importance, define certain statistical equilibrium densities. The corresponding ensemble averages are interpreted as thermodynamic equilibrium quantities. In what follows we discuss the *canonical* and the *microcanonical ensembles*.

Canonical Ensemble Averages

The canonical ensemble γ_T is shown to be connected to an equilibrium embedding of the molecule into a "heat bath" of temperature T, i.e., the expectation values with respect to γ_T include the statistics of a stationary interaction with a surrounding of constant temperature [26, 19]. Its density f_{γ_T} is defined as

$$f_{\gamma_T}(x) = \frac{1}{Q} \exp(-\beta H(x)) \quad \text{with} \quad Q = \int_\Gamma \exp(-\beta H(x))\, d\lambda(x),$$

and $\beta = 1/(k_B T)$, where k_B is Boltzmann's constant.
The evaluation of expectation values with respect to the canonical ensemble γ_T is mostly done via thermalization methods [29] or by means of various Monte Carlo techniques.

Monte Carlo (MC) methods (Metropolis algorithm) The recursive scheme of all MC algorithms is the construction of a sequence $(x_k)_{k\in\mathbb{N}} \subset \Gamma$ of states with two essential properties:

- Numerically, the state x_{k+1} can be computed using its predecessor x_k only and without any evaluation of the density f_{γ_T} itself.
- The subsequent mean values of the sequence converge to the expectation value desired, i.e.,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} A(x_k) \; = \; \langle A \rangle_{\gamma_T}, \tag{2.5}$$

  for the observable A under consideration.

The algorithmic realization of the step $x_k \to x_{k+1}$ consists of two parts:

(i) The *update step* $x_k \to \tilde{x}_k = Q(x_k)$: herein, the numerical operations necessary for realizing the "update operator" $Q : \Gamma \to \Gamma$ should be computationally cheap, e.g., should exclude evaluation of f_{γ_T}. There is a single theoretical restriction: Q must be *irreducible* [35].

(ii) The *acceptance step*: evaluate $\Delta E = H(\tilde{x}_k) - H(x_k)$, set

$$a = \min\{1, \exp(-\beta \Delta E)\}, \tag{2.6}$$

and compute r randomly equidistributed from $[0,1]$. The state $\tilde{x}_k$ is accepted as x_{k+1} if $r \le a$, otherwise $x_{k+1} = x_k$ is kept.

In the context of the theoretical justification of the Metropolis algorithm the sequence (x_k) is discussed as a realization of a Markov chain [5]. The transition operator of this Markov chain proves to be an irreducible *Frobenius–Perron operator*, the largest eigenvalue of which is $\lambda = 1$. The corresponding eigenspace is one–dimensional and the normalized eigenfunction proves to be the canonical density f_{γ_T}. Thus, the sequence (x_k) is the result of a fixed point iteration, which converges to the dominant eigenfunction f_{γ_T}, which, in turn, guarantees the convergence (2.5). In MC simulations, eigenvectors to smaller eigenvalues of the Frobenius–Perron operator are *not* taken into account (apart from the fact that the second largest eigenvalue may be considered to estimate the convergence rate [36]).
As is widely known, MC simulations for ensemble averages may suffer from possible "critical slowing down" [27]. This phenomenon occurs when the iteration $x_k \to x_{k+1}$ gets trapped near a local potential minimum so that a proper sampling of the phase space within reasonable computing times is prevented. In order to overcome such a trapping, large steps in the potential energy landscape (so–called global updates) would be required, which are extremely hard to construct. In [35] a multigrid approach to MC has been advocated designed to treat this multiscale phenomenon on sufficiently coarse grids. The algorithm to be designed herein will have some flavor of multigrid methods as well, but within a rather different underlying mathematical concept.

Hybrid Monte Carlo (HMC) methods The basic idea of HMC is the realization of the update step via short discrete subtrajectories [14]. The update step starting with $x = (q, p) \in \Gamma$ is realized via a discretization Ψ^τ of (2.1): first, new momenta $\tilde{p}$ are determined randomly according to the canonical momentum distribution $\exp(-\beta\, p^T M^{-1} p/2)$. Secondly, a trajectory is computed yielding the proposal

$$Q(x) = \tilde{x} = (\Psi^\tau)^m\ (q, \tilde{p}),$$

for the next state, with m and τ being free parameters. Q is irreducible, iff Ψ^τ is reversible and symplectic. The parameters m and τ have to be adjusted according to two requirements: On one hand, τ must be not too large, so that the energy variation is small enough to guarantee a sufficiently large rate of acceptance. On the other hand, m and τ have to be large enough, otherwise the update Q remains to act essentially locally with the undesirable consequences described above. It should be noted, that in HMC the discrete flow Ψ^τ need not be a good approximation of the exact flow Φ^τ. It is only used as a technique for proposing the next state and energy stability is the only requirement on τ.
HMC is reported to be an appropriate approach for several problems concerning polymerization [23, 28, 15]. However, for most MD applications, the sampling of the phase space still remains to be local.

Microcanonical Ensemble Averages

For introducing the microcanonical ensemble, consider an *energy cell* defined by

$$\Gamma_\delta(E) = \{x \in \Gamma,\, |H(x) - E| < \delta\}$$

for $E \in \mathbb{R}$ and $\delta \geq 0$. This reduces to an *energy surface* for $\delta = 0$. In the following, all energy cells are assumed to be bounded sets. Let m be the Lebesgue–measure on Γ. Then, the microcanonical ensemble μ_E is defined as the limit of measures $\mu_{E,\delta}$ with densities

$$f_{E,\delta}(x) = \begin{cases} 0, & x \notin \Gamma_\delta(E) \\ \frac{1}{m(\Gamma_\delta(E))}, & x \in \Gamma_\delta(E) \end{cases},$$

for $\delta \to 0$. Thus, μ_E corresponds to an *equidistribution* on $\Gamma_0(E)$. In statistical mechanics, the corresponding density f_E is interpreted as the equilibrium density for the states of the molecule at energy E when the system is *closed*, i.e., without any interaction to an exterior, and when it is large enough (i.e., in the "thermodynamical limit" [26]).
The microcanonical ensemble is intimately linked with the *physical ergodicity hypothesis*, which assumes that the expectation value of A with respect to this ensemble is equal to the time average $\bar{A}$ of each macro–observable A [1]:

$$\bar{A}(x_0) = \langle A \rangle_{\mu_E} = \int\limits_\Gamma A(x)\, d\mu_E, \tag{2.7}$$

where $E = H(x)$. Obviously, $\bar{A}$ is a function of E only and does not depend on the initial state x_0 of the time average. "Macro–observables" are observables like total energy or entropy, which do not only measure *local* properties of molecules.
On the basis of this hypothesis we are free to either compute $\bar{A}$ via (2.7) by simply integrating A over the energy surface under consideration or to compute $\langle A \rangle_{\mu_E}$ by an approximation of the time average $\bar{A}(x)$ via (2.3) starting with an arbitrary $x \in \Gamma_0(E)$. Typically, the latter option is taken using symplectic discretizations (compare the discussion in Section 2.1) to generate a *time series*. Its mean value is then taken as an approximation to $\bar{A}$. The stepsizes applied are as large as possible – the only restriction being the *stability* of the discrete iteration and *not* any *accuracy* requirement with respect to an exact trajectory. As a justification of this procedure, it is claimed that the time series "samples" the phase space with respect to the equidistribution of the microcanonical measure. Frankly speaking, however, the present authors are not aware of any more rigorous justification.
Finally, we want to emphasize that the "physical ergodicity hypothesis" should not be mixed up with the mathematical ergodic theory of dynamical systems (see Section 2.3). In fact, the underlying equidistribution hypothesis may even be wrong as will be exemplified in Section 5.

2.3 Invariant Measures of Dynamical Systems

We now turn to the question of how time averages $\bar{A}$ can be described in the theory of dynamical systems. For simplicity, we again restrict our attention to symplectic discretizations of (2.1). The discrete flow Ψ^τ as an approximation of the continuous flow Φ^τ can meet any accuracy requirement by a suitable choice of the time steps. In view of Section 2.1, we will only consider well–conditioned short–term subtrajectories. These restrictions lead us to discrete dynamical systems of the form

$$x_{j+1} = f(x_j), \quad j = 0, 1, 2, \ldots, \tag{2.8}$$

where $f = \Psi^\tau : \Gamma \to \Gamma$.

Mathematical Ergodicity

The long term behavior of any system (2.8) is described by so-called *invariant measures*: a probability measure μ is *invariant* iff $\mu(f^{-1}(B)) = \mu(B)$ for all measurable subsets $B \subset \Gamma$. In other words, invariant measures describe the recurrence behavior of the dynamical system:

Theorem 2.2 (Poincaré Recurrence Theorem) *Let μ be an invariant measure. Then for any measurable set B almost all points $x \in B$ (with respect to μ) return to B under some iterate.*

Obviously, invariant measures are deeply connected to *invariant sets*, i.e., subsets $B \subset \Gamma$ with $f^{-1}(B) = B$. An invariant measure is *ergodic* if $\mu(B) \in \{0, 1\}$ for every invariant set $B \subset \Gamma$. The following theorem implies that ergodic measures have particularly nice recurrence properties. Roughly speaking, it will state that *the time average is equal to the spatial average* for μ-almost all initial conditions.

THEOREM 2.3 (BIRKHOFF ERGODIC THEOREM) *Let μ be an ergodic measure. Then for μ almost all points $x \in \Gamma$ we have that*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} A(f^j(x)) = \int A \, d\mu$$

for each integrable function A.

This theorem is fundamental in mathematical Ergodic Theory. However, there is an essential drawback, which we are now going to illustrate by an example.

EXAMPLE 2.4 Suppose that the dynamical system (2.8) has a stable fixed point p, that is $f(p) = p$, and there is a neighborhood U of p such that all points inside that neighborhood converge to p in the course of the iteration. The ergodic measure related to p is the Dirac measure $\mu = \delta_p$, and the Birkhoff Ergodic Theorem implies that *for δ_p almost all $x \in \Gamma$*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} A(f^j(x)) = \int A \, d\mu = A(p) \tag{2.9}$$

for each integrable function A. In particular, this theorem just provides information on the temporal behavior of the point p itself, since this is the only point where (2.9) can be applied to. On the other hand, since p is stable, (2.9) would be satisfied for all points inside the neighborhood U of p.

The previous example illustrates that the stability property of an invariant set – in that case the stable fixed point p – is not taken into account in the notion of an ergodic measure. However, from the application point of view it would be much more satisfactory if the existence of an invariant measure were guaranteed which provides equality of the temporal and the spatial average for a "large" set of points. This observation leads naturally to the notion of an *SBR-measure* named after **S**inai, **B**owen and **R**uelle (cf. [32, 4]).

DEFINITION 2.5 An ergodic measure μ is an *SBR-measure* if there exists a subset $U \subset \Gamma$ with $m(U) > 0$ and such that for all $x \in U$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} A(f^j(x)) = \int A \, d\mu \tag{2.10}$$

for each continuous function A. Once again, m denotes Lebesgue measure.

Frobenius-Perron Operator

The most important observation in connection with the numerical computation of invariant measures is the fact that it is equivalent to the solution of an eigenvalue problem. To make this relationship more precise, we introduce the *Frobenius-Perron operator* on the set $\mathcal{M}$ of probability measures on Γ.

DEFINITION 2.6 The *Frobenius-Perron operator* $P : \mathcal{M} \to \mathcal{M}$ is defined by

$$(P\mu)(B) = \mu(f^{-1}(B)) \quad \text{for all measurable } B \subset \Gamma \text{ and arbitrary } \mu \in \mathcal{M}.$$

By the definition of the Frobenius-Perron operator,

a probability measure $\mu \in \mathcal{M}$ is invariant if and only if $P\mu = \mu$.

Hence a promising strategy will be to discretize this operator equation in such a way that the matrix approximation P_d of P has an eigenvector v_d with $P_d v_d = v_d$ which is close to an invariant measure. More precisely, suppose that – by backward analysis – the numerical discretization is exact for a stochastic perturbation of the original system. Then it is reasonable to assume that an SBR-measure is approximated by our discretization, since these measures are robust with respect to stochastic perturbations. We will explain the discretization in more detail in Section 4.1 (see also Appendix 5.2). There, we will also see that the entries of the stochastic matrix P_d can indeed be evaluated by a collection of short–term subtrajectories — as desired by considerations in Section 2.1.

Summarizing, we will aim at constructing an algorithm based on a multiscale discretization of the eigenvalue problem for the Frobenius-Perron operator. In contrast to the Statistical Physics approach, where certain measures μ are prescribed on the basis of physical model considerations, the solution of the eigenvalue problem for $\lambda = 1$ will supply approximations of the invariant measure μ *without any physical modelling assumptions.*

3 Conformational Changes

From a chemist's point of view, biomolecular systems are characterized by different "conformations". This term simultaneously describes both distinguishable geometric configurations and the associate chemical "functionality". In a conformation the large-scale geometric structure of the molecule is understood to be conserved, while the system may well rotate, oscillate or fluctuate on small spatial scales. Fluctuations of a molecule are only of interest in the transient phase from one conformation to another (compare the flip-flops between conformations in Fig. 3). Typically the duration of stay within a conformation is long enough to make the conformation an object of chemical interest or, equivalently, to make a significant contribution to any (statistical) averages. Consequently, conformational changes are rare events, which normally can only be observed in long–term simulations.

Mathematically speaking, conformations are described as special subsets of phase space. Invariant sets of MD systems, which correspond to infinite relaxation times, typically consist of a number of different subsets describing different conformations. These conformation subsets, in turn, correspond to finite relaxation times and may therefore be denoted as "almost invariant" sets. Of course, the main interest focusses on those conformations with the largest relaxation times.

3.1 Illustrative Example

Let us introduce a suitably simple example in order to illustrate the notion of almost invariant sets for Hamiltonian systems. For $p = (p_1, p_2), q = (q_1, q_2) \in \mathbb{R}^2$ consider the potential

$$V_4(q) = \left(\frac{3}{2}q_1^4 + \frac{1}{4}q_1^3 - 3q_1^2 - \frac{3}{4}q_1 + 3\right) \cdot \left(2q_2^4 - 4q_2^2 + \alpha\right) \quad \text{with } \alpha = 3. \tag{3.1}$$

As illustrated in Fig. 4, this potential comprises four local minima at the points $(\pm 1, \pm 1)$ (named A, B, C, D), which are separated by four saddlepoints. The energy barrier between A and B is significantly higher than the other three ones. The dynamical behavior of the system consists of oscillations around the local minima and, if the total energy is large enough for the system to cross the barriers, of motions from one minimum to the other. If the energy is not too large, there will be two kinds of "long term" dynamical behavior:

(a) oscillations in the neighborhood of the four different minima,

(b) back and forth oscillations between two different minima: $A \leftrightarrow D$, $B \leftrightarrow C$, and $C \leftrightarrow D$.

This is observed in simulations of the dynamics. Fig. 4 presents a solution which starts with an oscillation between A and D, followed by an oscillation around C, a long period of oscillations between A and D and so on. The similarity of the trajectories shown in Fig. 4 (right) and Fig. 3 illustrates that we are actually looking at the same kind of phenomena. Thus, for the case presented in Fig. 4, the neighborhoods of the different minima should turn out to be "almost invariant sets" as well as neighborhoods of the pairs of minima (A, D), (C, D) and (B, C) together with regions around the corresponding saddlepoints "between" them. In Section 5 we will see that this fact can indeed be justified.

3.2 An Eigenvalue Approach to Almost Invariant Sets

In Section 2.3, we have already seen that invariant measures correspond to fixed points of the Frobenius-Perron operator. Hence, eigenmeasures of the Frobenius-Perron operator corresponding to (real) eigenvalues close (but not equal) to 1 should be related to almost invariant sets as described above.

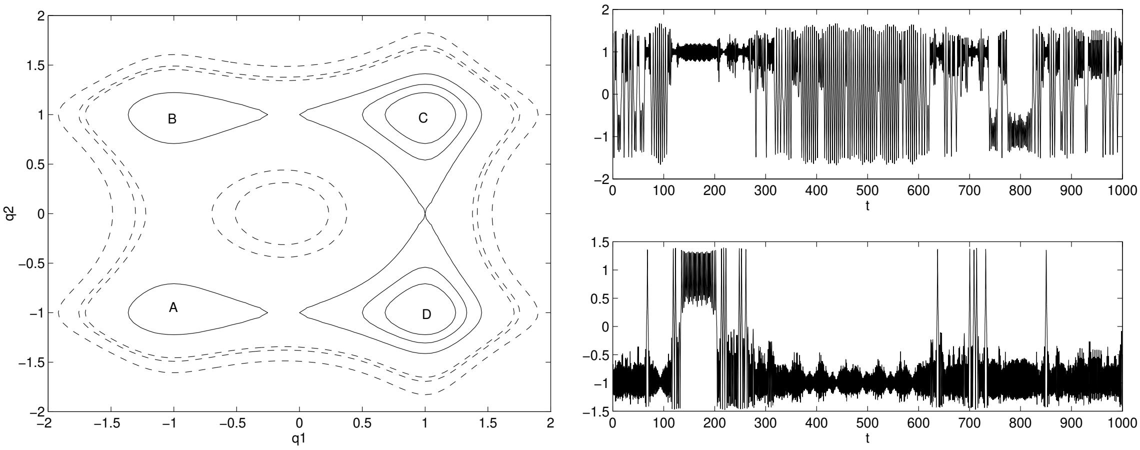

Figure 4: The left hand side figure shows a contour plot of the potential energy landscape due to V_4 with equipotential lines of the energies $E = 1.5, 2, 3$ (solid lines) and $E = 7, 8, 12$ (dashed lines). There are minima at the four points $(\pm 1, \pm 1)$ (named A to D), a local maximum at $(0, 0)$, and saddlepoints in between the minima. The right hand figure illustrates a solution of the corresponding Hamiltonian system with total energy $E = 4.5$ (positions q_1 and q_2 versus time t). See Section 3 for more details.

Let us illustrate this fact using the simple test system given by V_4 in (3.1). First suppose that our system has two disjoint invariant sets B_1 and B_2 with corresponding invariant measures μ_1 and μ_2. For total energy $E = 4.5$ this is the case, if we choose $\alpha \geq 4.5$ in (3.1). Then B_1 may be chosen as the neighborhood of the pair of minima (B, C) and B_2 as the neighborhood of (A, D) (cf. Fig. 4). The invariant measures μ_1 and μ_2 are two independent eigenmeasures of the Frobenius-Perron operator corresponding to the (at least) double eigenvalue 1. They may be chosen so that $\mu_k(B_j) = 0$ for $(k, j) = (1, 2)$ and $(k, j) = (2, 1)$. Assume that the corresponding eigenspace E_1 is two–dimensional. Then, $\mu_* = (\mu_1 + \mu_2)/2$ and $\nu_* = (\mu_2 - \mu_1)/2$ are also a basis of E_1 with the property that

$$\nu_*(B_1) = -1/2 \quad \text{and} \quad \nu_*(B_2) = 1/2. \tag{3.2}$$

This fact will be of importance in the following.
Next suppose that we vary the control parameter α in (3.1) such that the two invariant sets B_1 and B_2 merge for a certain value of the parameter ($\alpha = 4.5$) leading to a “confluent” invariant set $B \approx B_1 \cup B_2$. For $\alpha \leq 4.5$ trajectories can move from B_1 to B_2 and vice versa, but this will happen rarely and they will stay in each of these components for quite a long time. (In fact, this is precisely the flip-flop behavior between the almost invariant sets that we have observed in the previous examples — see the lowest level of Fig. 3.) In the merging process of the two invariant sets B_1 and B_2, one of the two eigenvalues 1 has to move away from 1 along the real line into the unit circle. We now show how to extract information on the “almost invariant components” from the magnitude of this eigenvalue and its corresponding eigenmeasure.

Let us begin with a mathematically precise definition of an almost invariant set. Let $\rho \in \mathcal{M}$ be a probability measure. We say that the set B is *δ-almost invariant with respect to ρ* if

$$\frac{\rho(f^{-1}(B) \cap B)}{\rho(B)} = \delta.$$

Thus, δ is the probability that points in B are mapped into B under f. In particular, if B is an invariant set, that is $f^{-1}(B) = B$, then $\delta = 1$. The main purpose is to relate the magnitude of eigenvalues of the Frobenius-Perron operator P which are close to 1 to this probability δ.

Once $\delta = \delta_B$ has been computed for a given set $B \subset \Gamma$ using the stepsize τ in the discrete dynamical system with $f = \Psi^\tau$, the system's probability of staying within B for time T can be estimated to be

$$p_B(T) = \delta_B^{T/\tau}. \tag{3.3}$$

It is clear that this information is of utmost chemical importance.

We assume that $\lambda \neq 1$ is an eigenvalue of P with corresponding real valued eigenmeasure $\nu \in \mathcal{M}_{\mathbb{C}}$, that is,

$$P\nu = \lambda\nu, \quad \text{where } \nu \text{ is scaled so that } |\nu| \in \mathcal{M}.$$

(We denote by $\mathcal{M}_{\mathbb{C}}$ the set of bounded complex valued measures.) Obviously, ν cannot be a probability measure – in fact, it is easy to see that $\nu(\Gamma) = 0$, see Appendix 5.2. However, if the eigenvalue λ is close to one it is reasonable to assume that the probability measure $|\nu|$ is close to the invariant measure μ of the system. To see this observe that $|\nu| \approx |\nu_*| = \mu_* \approx \mu$ when $\lambda \approx 1$.

We return to our example from above to illustrate these facts. In Figure 5 we present the eigenmeasure ν for $\alpha = 3$. Since B_1 and B_2 were taken to be the regions around (B, C) and (A, D), respectively, we observe $\nu(B_1) < 0$ and $\nu(B_2) > 0$ so that $\nu(\Gamma) = \nu(B_1 \cup B_2) = 0$. Obviously, we find $\nu(B_2) = -\nu(B_1) = 1/2$ since $|\nu|$ is a probability measure. Thus, the perturbation via the variation of α conserves the property (3.2) for the two emerging almost invariant sets. Consequently, it is reasonable to look for the almost invariant sets B among the sets with the property $\nu(B) = \pm 1/2$.

We now state a result concerning the relationship between probabilities, by which sets are almost invariant, and the eigenvalue λ. A proof of the result in the context of *small random perturbations* can be found in [9].

PROPOSITION 3.1 *Let $B \subset \Gamma$ be a set with $\nu(B) = \frac{1}{2}$. If B is δ_1-almost invariant and $\Gamma - B$ is δ_2-almost invariant with respect to $|\nu|$, then*

$$\delta_1 + \delta_2 = \lambda + 1. \tag{3.4}$$

If there are more than two almost invariant sets, then a set B as in Proposition 3.1 may itself be the union of several almost invariant sets. In order to identify all these components one has to consider all the eigenmeasures corresponding to eigenvalues which are close to one. We will illustrate this in our numerical examples in Section

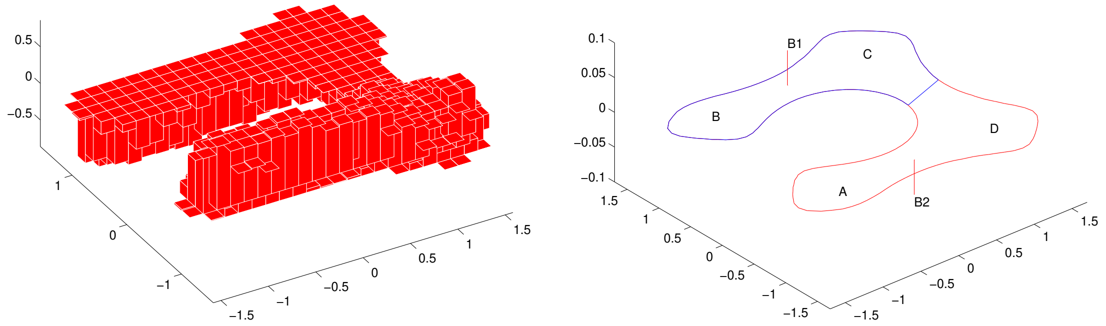

Figure 5: Eigenmeasure ν of the Frobenius–Perron operator to the eigenvalue $\lambda = 0.9963$ for the test system (3.1) with $\alpha = 3$. ν was computed via our new subdivision algorithm (cf. Section 5.2).

5. (An alternative numerical approach is presented in [8].) In that context also the following consideration will be useful. Suppose that we are given the probabilities δ_X and δ_Y for two separate almost invariant sets X and Y. Then we will be interested in computing $\delta_{X\cup Y}$. This can be done by means of the following lemma.

Lemma 3.2 *Let $\rho \in \mathcal{M}$ be a probability measure and let X and Y disjoint sets which are δ_X- resp. δ_Y-almost invariant with respect to ρ. Moreover suppose that $f^{-1}(X) \cap Y = \emptyset$ and $f^{-1}(Y) \cap X = \emptyset$. Then $X \cup Y$ is $\delta_{X\cup Y}$-almost invariant with respect to ρ where*

$$\delta_{X\cup Y} = \frac{\rho(X)\delta_X + \rho(Y)\delta_Y}{\rho(X) + \rho(Y)}. \tag{3.5}$$

Proof: We calculate

$$\begin{aligned}\delta_{X\cup Y} &= \frac{\rho(f^{-1}(X\cup Y)\cap(X\cup Y))}{\rho(X\cup Y)}\\ &= \frac{\rho((f^{-1}(X)\cap X)\cup(f^{-1}(Y)\cap Y))}{\rho(X)+\rho(Y)} = \frac{\rho(X)\delta_X + \rho(Y)\delta_Y}{\rho(X)+\rho(Y)}.\end{aligned}$$

$\square$

In (3.4), both δ_1 and δ_2 appear and in general there will be no relation between these constants. However, if the underlying system possesses an additional symmetry – as in (3.1), since the Hamiltonian system is equivariant under the transformation $(q_2, p_2) \to -(q_2, p_2)$ –, then we can express one of these numbers in terms of the other one. To illustrate this fact, let us consider the simplest case where we have a symmetry transformation κ in the problem with $\kappa^2 = id$. Then one can show (see again [9]):

Corollary 3.3 *In addition to the assumptions in Proposition 3.1 suppose that*

(i) the set B satisfies $\kappa B = \Gamma - B$, and

(ii) the measure $|\nu|$ is κ-symmetric, that is $\kappa^*|\nu| = |\nu|$.

Then $\Gamma - B$ is δ-almost invariant with respect to $|\nu|$ if and only if B is δ-almost invariant. In particular

$$\delta = \frac{\lambda + 1}{2}. \tag{3.6}$$

Example 3.4 Consider the Hamiltonian system discussed in Section 3.1 above. The symmetry κ is given by $\kappa(q_1, q_2, p_1, p_2) = (q_1, -q_2, p_1, -p_2)$. We define our phase space to be $\mathbb{R}^4$ without the fixed point space of κ, that is

$$\Gamma = \mathbb{R}^4 - \{(q, p) \in \mathbb{R}^4 : (q_2, p_2) = (0, 0)\},$$

and set

$$B = \{(q, p) \in \Gamma : q_2 \geq 0 \text{ and } p_2 > 0 \text{ if } q_2 = 0\}.$$

Obviously $\kappa B = \Gamma - B$, and therefore condition (i) in Corollary 3.3 is satisfied. Now, consider the real valued eigenmeasure ν from Fig. 5 with corresponding eigenvalue $\lambda \approx 0.9963$. $|\nu|$ has the symmetric support $B_1 \cup \kappa B_1 = B_1 \cup B_2$. Then Corollary 3.3 supplies that B_1 (the neighborhood of the pair of minima (B, C)) is δ-almost invariant with respect to $|\nu|$ if and only if $B_2 = \kappa B_1$ (the neighborhood of $(\kappa B, \kappa C) = (A, D)$) is δ-almost invariant. Moreover, δ is given by (3.6) which results in $\delta = (\lambda + 1)/2 \approx 0.9981$. For a detailed discussion see Section 5.2 below.

4 Subdivision Algorithms

Until now subdivision techniques have been used to analyze the long term dynamical behavior of *hyperbolic* dynamical systems. In particular, they turned out to be very useful in the computation of invariant measures and invariant manifolds [7, 9, 8]. In this section we describe how these techniques can be modified to apply to *Hamiltonian* dynamical systems. Both for completeness and in order to make the differences more transparent, we briefly review the techniques for hyperbolic systems in Sections 4.1.

4.1 Hyperbolic Systems

The central mathematical object to be approximated by the subdivision algorithm due to [7] is the so-called *relative global attractor*,

$$A_Q = \bigcap_{j \geq 0} f^j(Q), \tag{4.1}$$

where $Q \subset \mathbb{R}^n$ is a compact subset. Roughly speaking, the set A_Q should be viewed as the *union of unstable manifolds of invariant objects inside* Q. In particular, A_Q may contain subsets of Q which cannot be approximated by direct simulation.

Covering of the Relative Global Attractor

The numerical realization involves the discretization of the Frobenius-Perron operator. For that purpose we first need to determine a sequence of box coverings $\mathcal{B}_k$ of the relevant dynamics in phase space. The subdivision algorithm for the approximation generates a sequence $\mathcal{B}_0, \mathcal{B}_1, \mathcal{B}_2, \ldots$ of finite collections of boxes with the property that for all integers k the set $Q_k = \bigcup_{B \in \mathcal{B}_k} B$ is a covering of the relative global attractor A_Q under consideration. The sequence of coverings is constructed in such a way that the diameter of the boxes,

$$\operatorname{diam}(\mathcal{B}_k) = \max_{B \in \mathcal{B}_k} \operatorname{diam}(B)$$

converges to zero for $k \to \infty$.
Given an initial collection $\mathcal{B}_0$, one recursively obtains $\mathcal{B}_k$ from $\mathcal{B}_{k-1}$ for $k = 1, 2, \ldots$ in two steps.

(i) *Subdivision:* Construct a new collection $\hat{\mathcal{B}}_k$ such that

$$\bigcup_{B \in \hat{\mathcal{B}}_k} B = \bigcup_{B \in \mathcal{B}_{k-1}} B \qquad \text{and} \qquad \operatorname{diam}(\hat{\mathcal{B}}_k) \leq \theta \operatorname{diam}(\mathcal{B}_{k-1})$$

for some $0 < \theta < 1$.

(ii) *Selection:* Define the new collection $\mathcal{B}_k$ by

$$\mathcal{B}_k = \left\{ B \in \hat{\mathcal{B}}_k : f^{-1}(B) \cap \hat{B} \neq \emptyset \quad \text{for some} \quad \hat{B} \in \hat{\mathcal{B}}_k \right\}.$$

The following proposition establishes a general convergence property of this algorithm.

Proposition 4.1 ([7]) *Let A_Q be the global attractor relative to the compact set Q, and let $\mathcal{B}_0$ be a finite collection of closed subsets with $Q_0 = Q$. Then*

$$\lim_{k \to \infty} h\left(A_Q, Q_k\right) = 0,$$

where we denote by $h(B, C)$ the usual Hausdorff distance between two compact subsets $B, C \subset \mathbb{R}^n$.

Remarks 4.2 (a) Observe that for the convergence result in Proposition 4.1 it is not necessary to assume that the underlying system is hyperbolic. However, for estimates concerning the speed of convergence the existence of a hyperbolic structure is very useful (see [7]).

(b) A drawback of the described subdivision technique is that at each level boxes of the same size are taken into account without any use of information about the underlying dynamical behavior. In particular, also those boxes are further subdivided that are dynamically irrelevant in the sense that their (natural) invariant measure is zero. For this reason, an *adaptive* version of the subdivision algorithm has recently been developed in [8], which led to a considerable computational speed-up.

Discretization of the Frobenius-Perron Operator

Following [12, 9] we use a Galerkin method to approximate the Frobenius-Perron operator. This treatment is theoretically justified in Appendix 5.2 where, roughly speaking, it is shown that the Frobenius-Perron operator becomes a compact operator on L^2 if the system is stochastically perturbed. In this case, invariant measures have L^∞-densities, and L^2 is simply a nice choice for a corresponding function space. Having this in mind we describe the discretization of the Frobenius-Perron operator P viewed as an operator acting on densities in L^2.
Let V_d, $d \geq 1$, be a sequence of d–dimensional subspaces of L^2 and let $\{\varphi_i\}$, $i = 1, 2, \ldots, d$, be a basis of V_d such that

$$\sum_{i=1}^{d} \varphi_i(x) = 1 \quad \text{for all } x \in \Gamma. \tag{4.2}$$

For $g \in L^2$ the Galerkin projection $Q_d : L^2 \to V_d$ is defined by

$$(Q_d g, \varphi_i) = (g, \varphi_i) \quad \text{for } i = 1, \ldots, d,$$

where $(\cdot, \cdot)$ is the usual inner product in L^2. Observe that Q_d converges pointwise to the identity on L^2.
In most of the applications V_d consists of functions which are locally constant. More precisely, let B_i, $i = 1, \ldots, d$, denote the boxes contained in the covering $\mathcal{B}_k$. Then we can choose

$$\varphi_i = \chi_{B_i}, \quad i = 1, 2, \ldots, d.$$

With this choice the discretized Frobenius-Perron operator $P_d = Q_d P$ is represented by the stochastic matrix

$$v = P_d u, \quad v(B_i) = \sum_{j=1}^{d} \frac{m(f^{-1}(B_i) \cap B_j)}{m(B_j)} u(B_j), \quad i = 1, \ldots, d, \tag{4.3}$$

where m denotes Lebesgue measure. Now a fixed point u_d of P_d provides an approximation to the SBR-measure of f.

Remark 4.3 Until now it is theoretically not clear whether the fixed points obtained by the discretization described above indeed converge to an SBR-measure for $d \to \infty$. Except for expanding maps (and other specific situations) this is not even known under the additional assumption of the existence of an SBR-measure. However, convergence to SBR-measures can be proved in the context of *small random perturbations* of the underlying system. We will outline these results for systems with a hyperbolic structure in Appendix 5.2.

4.2 Hamiltonian Systems

The above subdivision algorithm has proved to be well suited for hyperbolic dynamical systems where the long term dynamical behavior is confined to a low-dimensional object in phase space. In the Hamiltonian context, however, energy surfaces are dynamically invariant and it is of interest to derive information on the dynamical behavior of the system restricted to this manifold. Of course, the long term dynamics again may be confined to a lower-dimensional subset of the energy surface under consideration. Consequently, the main algorithmic steps for the approximation of the dynamical behavior of a Hamiltonian system are as follows:

(i) Construction of an approximate covering of the energy surface;

(ii) extraction of the subset containing the long-term dynamics;

(iii) setting up the Frobenius-Perron operator with respect to this subset.

We now describe each of these steps in more detail.

Covering of the Energy Surface

Given a compact energy surface $\Gamma_0(E) \subset \Gamma$ we want to construct a collection $\mathcal{B}$ of compact subsets of Γ such that $\Gamma_0(E)$ is contained in the union Q of these subsets. We require $\mathcal{B}$ to be a good approximation in the sense that the Hausdorff-distance between Q and $\Gamma_0(E)$ is smaller than a prescribed accuracy δ.
The algorithm, by which this collection $\mathcal{B}$ is constructed, is very similar to the standard subdivision algorithm except for one modification. The *selection*-step must be replaced by the following rule: define the new collection $\mathcal{B}_k$ by

$$\mathcal{B}_k = \{B \in \hat{\mathcal{B}}_k : B \cap \Gamma_0(E) \neq \emptyset\}. \tag{4.4}$$

One easily verifies that, if $\Gamma_0(E) \subset Q_0$, then the union Q_k of the boxes of the collection $\mathcal{B}_k$ covers the energy surface $\Gamma_0(E)$ under consideration for every $k = 1, 2, \ldots$. Furthermore the Hausdorff-distance between Q_k and $\Gamma_0(E)$ can be seen to approach zero for $k \to \infty$.
Equation (4.4) is difficult to check in practice. Therefore, in the actual implementation, we employ a somewhat different procedure: instead of an energy surface we aim at covering an *energy cell* $\Gamma_{\delta_k}(E)$ in the k-th step of the algorithm. The sequence (δ_k) is chosen to decrease with increasing k. With this modification, (4.4) can be checked by calculating the energy for a heuristically determined (fixed) number of points within each box. In each subdivision step k, the parameter δ_k is adapted to the size of the boxes.

Extraction of the Location of the Long-Term Dynamics

Once we have constructed a covering $\mathcal{B}$ of the energy surface $\Gamma_0(E)$ under consideration, this collection may be used as the initial collection for the standard subdivision algorithm. Note, however, that still one modification has to be taken into account:

since we are interested in the dynamics on $\Gamma_0(E)$ we just have to map those points in the boxes of the current collection that lie on $\Gamma_0(E)$. Hence the *selection*-step becomes: define the new collection $\mathcal{B}_k$ by

$$\mathcal{B}_k = \left\{ B \in \hat{\mathcal{B}}_k : f(\hat{B} \cap \Gamma_0(E)) \cap B \neq \emptyset \quad \text{for some} \quad \hat{B} \in \hat{\mathcal{B}}_k \right\}. \tag{4.5}$$

Again, we encounter the same problem for the realization of (4.5) as for (4.4) before. But since we are approximating energy *cells* (as described above), we just need to check additionally whether a given point is contained in the energy cell or not.

Discretization of the Frobenius-Perron Operator

The previous two steps led to a collection $\mathcal{B} = \{B_1, \ldots, B_d\}$ covering the global attractor relative to a certain energy surface. We may now use this covering for the computation of a discretized Frobenius-Perron operator as described in Section 4.1. As an example consider again the case of locally constant basis functions

$$\varphi_i = \chi_{G_i}, \quad i = 1, 2, \ldots, d,$$

where we have set $G_i = B_i \cap \Gamma_0(E)$. Then, as in (4.3), the discretized Frobenius-Perron operator $v = P_d u$ can be written componentwise as

$$v(G_i) = \sum_{j=1}^{d} p_{ij} u(G_j), \quad p_{ij} = \frac{m(f^{-1}(G_i) \cap G_j)}{m(G_j)}, \quad i = 1, \ldots, d. \tag{4.6}$$

In the implementation we are faced with the calculation of the transition probabilities p_{ij} for $G_i = B_i \cap \Gamma_{\delta_k}(E)$. This is done via a Monte-Carlo approximation,

$$p_{ij} = \frac{1}{m(G_j)} \int_{G_j} \chi_{G_i}(f(x))dx \approx \frac{1}{N} \sum_{n=1}^{N} \chi_{G_i}(f(x_n)),$$

where the x_n are chosen randomly and uniformly distributed in G_j.
After the assembling of the stochastic matrix P_d we have to solve the associated *non–selfadjoint eigenvalue problem.* In order to catch the multiscale flavor of the whole problem, the new adaptive multigrid algorithm due to Friese [17] is most appropriate. This algorithm is an extension of the adaptive multigrid methods for selfadjoint eigenvalue problems as published in [11]. Up to now, apart from first tests, our preliminary numerical results presented in Section 5 have been computed with `speig` (by Radke and Sørensen) in Matlab. An important feature of all these algorithms is that they allow for a simultaneous subspace iteration to compute eigenmodes associated with eigenvalue clusters. (Here we are, of course, interested in the cluster around $\lambda = 1$.)

5 Illustrative Numerical Experiments

In this section we want to illustrate certain features of a first version of our subdivision algorithm SubMD. As derived above, the objects of interest will be invariant measures and almost invariant sets.

5.1 Invariant Measures

Recall that invariant measures are approximated by the eigenvectors of the discretized Frobenius-Perron operator according to the eigenvalue $\lambda = 1$.

Counterexample to the Physical Ergodicity Hypothesis

For $x = (q, p) \in \mathbb{R}^2$ consider the "double-well" potential

$$V(q) = \left(q^2 - 1\right)^2 .$$

All solutions are periodic (cf. Fig. 6). The periodicity allows a reliable evaluation of the invariant measure via direct simulation.

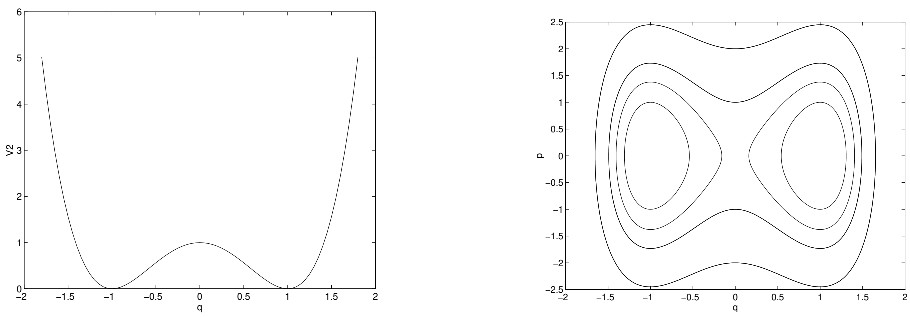

Figure 6: The double well potential V_2 and the corresponding phase portrait.

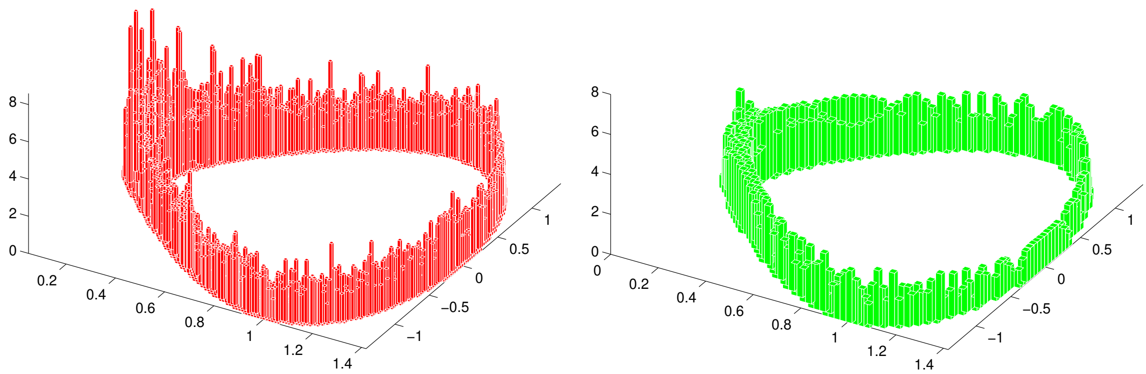

Figure 7: The density of the invariant measure of the double well potential for total energy $E = 0.95$. Left hand side: computed via the subdivision algorithm. Right hand side: computed via direct simulation.

As can be seen in Fig. 7, both our subdivision algorithm and direct simulation yield comparable approximations of the invariant measure. Note that the measure is *not equidistributed on the energy cell* in clear contradiction to the physical ergodicity hypothesis. Instead, there is a significant maximum of the probability density near the point $(0, 0)$. This is caused by the fact that the energy chosen is nearly critical and the mass particle is creeping slowly near the turning point.

Inefficiency of Direct Simulation

Recall example (3.1) from Section 3.1. We want to compute the corresponding invariant measure μ_4. A direct analytical solution does not exists. Direct long term simulation by symplectic discretization of (2.1) yields the discrete solution $(x_k)_{1,\ldots,N}$. For N large enough and a box $B \subset \Gamma$ one takes

$$p = \frac{1}{N\,|B|} \sum_{j=1}^{N} \chi_B(x_j)$$

as an approximation of the density $f^4|_B$. If the system were ergodic, the convergence of this algorithm would be guaranteed. Even in this case the convergence could be arbitrarily slow, when the iteration gets trapped within an almost invariant set of the system – compare the sequence of results obtained by direct simulation in Fig. 9.
Our global subdivision approach is not sensitive to such a situation. Over sufficiently long run times of direct simulation both methods eventually yield roughly the same results, see Fig. 8.

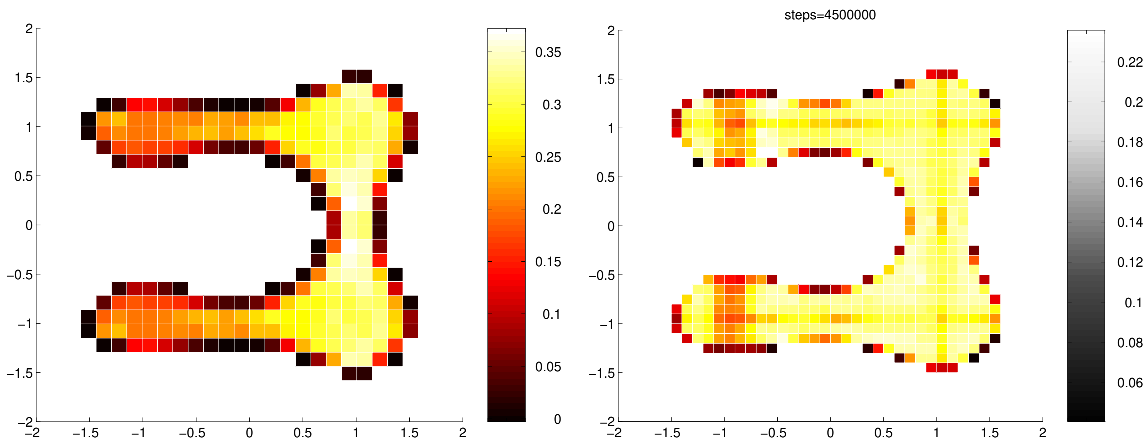

Figure 8: The density of the invariant measure of the potential V_4 for total energy $E = 4.5$. Results of the subdivision approach (left) and a direct simulation with about 4.5 million steps for stepsize $\tau = 1/30$ (right).

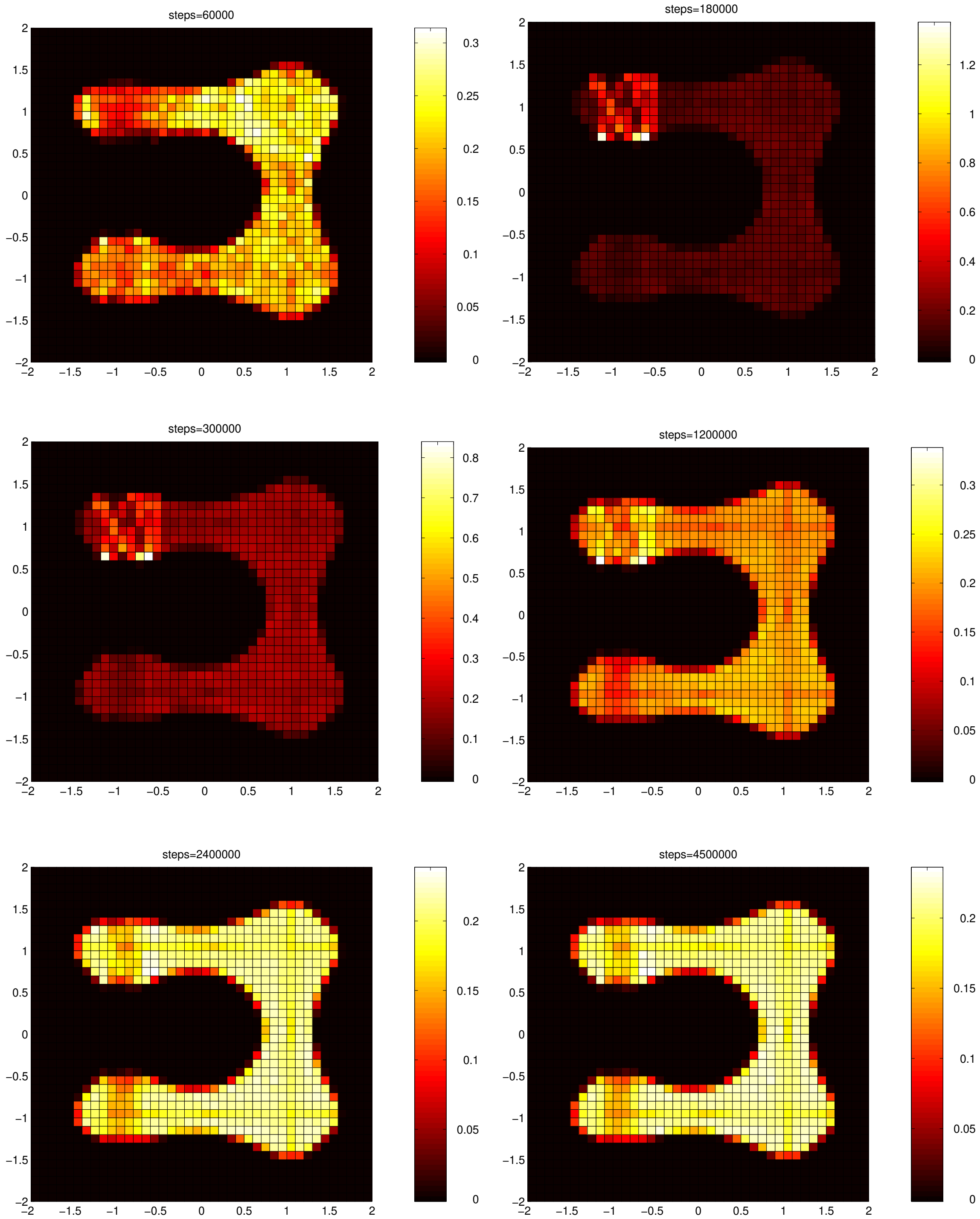

Figure 9: Discrete spatial density $h(q) = \int h^4(q,p)dp$ of the invariant measure of the potential V_4 for total energy $E = 4.5$. Results of direct simulation with stepsize $\tau = 1/30$ leading to an energy variation of about 1 percent. Between the 60000th and the 105000th step the discrete solution gets repeatedly trapped in one of the minima of the potential. The approximation has still not recovered from this event after 1 million steps.

5.2 Almost Invariant Sets

Recall that the relevant almost invariant sets correspond to eigenvalues $\lambda \approx 1$ with $|\lambda| < 1$ of the associated Frobenius–Perron operator.
Again we consider the example of Section 3. Based on observations concerning the dynamical behavior we already conjectured that there exist seven almost invariant sets – a conjecture that we now want to check numerically. We employ the subdivision algorithm for stepsize $\tau = 0.1$. The final box-collection corresponding to the total energy $E = 4.5$ after 18 subdivision steps consists of 18963 boxes.
A simultaneous computation of the four largest eigenvalues $\lambda_1, \ldots, \lambda_4$ leads to the following table:

Number	Eigenvalue
1	1.0000
2	0.9963
3	0.9891
4	0.9782

The invariant measure ν_1 corresponding to $\lambda_1 = 1$ has already been shown in Fig. 8. Next, we discuss the information provided by the eigenmeasure ν_2 corresponding to λ_2. The box coverings in the two parts of Fig. 10 approximate two sets, where the discrete density of ν_2 is positive resp. negative. In other words each of these sets is a candidate for a set B mentioned in the assumptions of Corollary 3.3. Thus, by this result, both of these sets are almost invariant with probability $\delta = (\lambda_2 + 1)/2 = (0.9963 + 1)/2 = 0.9981$. Observe that these almost invariant sets confirm the observation made in Section 3 that dynamically there exist "long term" oscillations between the minima $A \leftrightarrow D$ and $B \leftrightarrow C$.

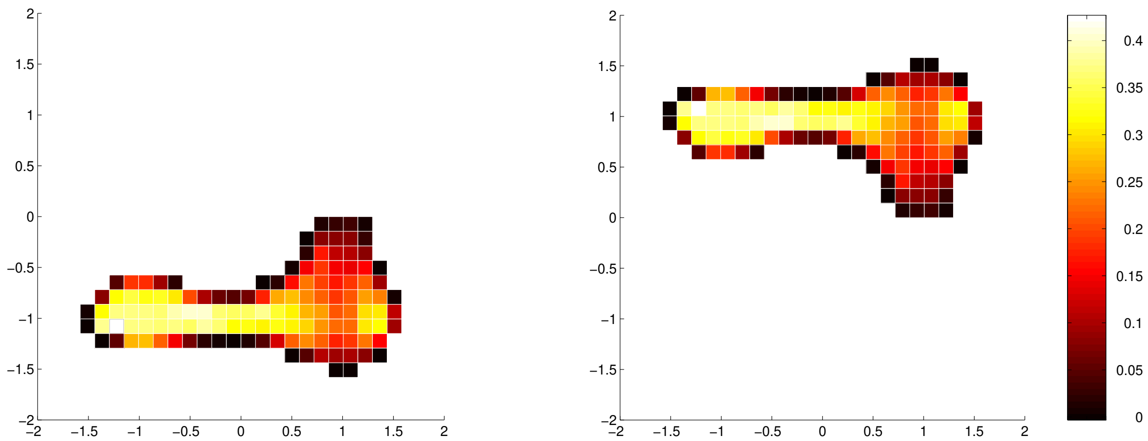

Figure 10: Illustration of two almost invariant sets with respect to the probability measure $|\nu_2|$. The coloring is done according to the magnitude of the discrete density.

The third eigenmeasure ν_3 corresponding to λ_3 provides information about three additional almost invariant sets: on the left hand side in Fig. 11 we have the set corresponding to the oscillation $C \leftrightarrow D$, whereas on the right hand side the two almost invariant sets around the equilibria A and B are identified. Again the boxes shown in the two parts of Fig. 11 approximate two sets where the discrete density of ν_3 is positive resp. negative. In this case we can use Proposition 3.1 and the fact that A and B are symmetrically related to conclude that for all these almost invariant sets $\delta \geq \lambda_3 = 0.9891$.

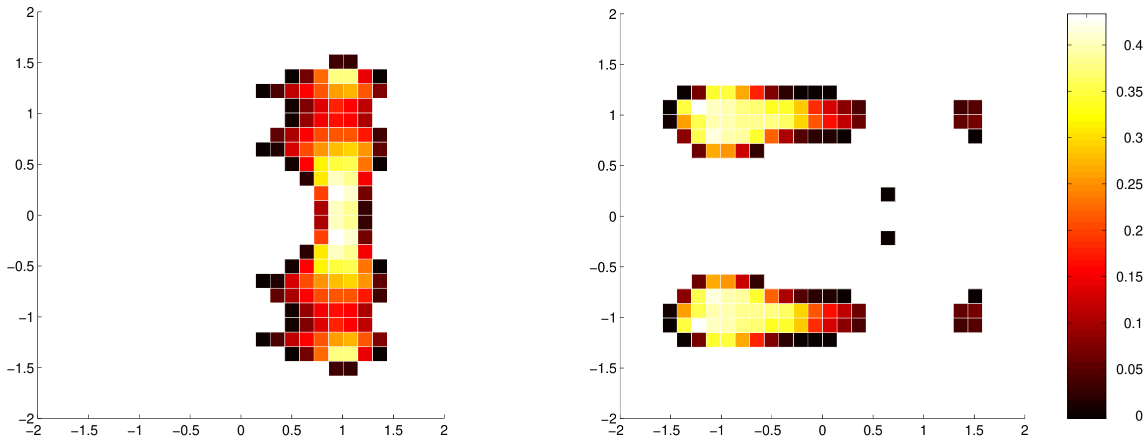

Figure 11: Illustration of three almost invariant sets with respect to the probability measure $|\nu_3|$. The coloring is done according to the magnitude of the discrete density.

Finally, the information on the remaining almost invariant sets in the neighborhood of the equilibria C and D can be extracted using the eigenvalue λ_4 with the eigenmeasure ν_4 (see Fig. 12). In the two parts of Fig. 12 we show again the boxes, which approximate two sets, where the discrete density of ν_4 is positive resp. negative. Let us denote by Y the union of the boxes around equilibrium B in the first part of the figure and by X the boxes around D. (We ignore the isolated box in the left lower corner, which we regard as a numerical artefact.) We now use Lemma 3.2 to derive a lower bound for δ_X. Numerically we obtain the values $|\nu_4(X)| = 0.3492$ and $|\nu_4(Y)| = 0.1508$. Note that $|\nu_4(X \cup Y)| = 0.5$ and $\lambda_4 + 1 = 2\delta_{X\cup Y}$ (using again the symmetry and Corollary 3.3) which leads to the estimate

$$\delta_X \;=\; \frac{0.5\delta_{X\cup Y} - \rho(Y)\delta_Y}{\rho(X)} = \frac{\lambda_4 + 1 - 4\rho(Y)\delta_Y}{4\rho(X)} \geq \frac{\lambda_4 + 1 - 4\rho(Y)}{4\rho(X)} = 0.9844$$

In all calculations done so far a fixed stepsize $\tau = 0.1$ has been used. Hence an application of formula (3.3) leads to the following table concerning flip–flop probabilities between different conformations.

conformation	probability to stay within for			
	0.1 sec.	1 sec.	10 sec.	100 sec.
$A \leftrightarrow D,\ B \leftrightarrow C$	0.9981	0.9812	0.8268	0.1493
$C \leftrightarrow D,\ A,\ B$	0.9891	0.8962	0.3342	< 0.0002
C,D	0.9844	0.8545	0.2076	$< 10^{-6}$

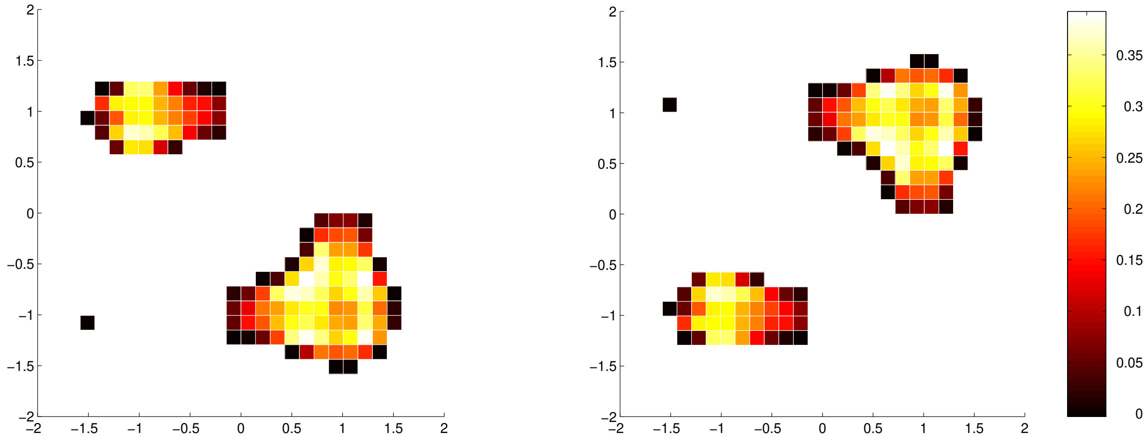

Figure 12: Illustration of four almost invariant sets with respect to the probability measure $|\nu_4|$. The coloring is done according to the magnitude of the discrete density.

These numbers indicate that it is very unlikely for the system to stay in C and D for more than 100 seconds, whereas for an oscillation $A \leftrightarrow D$ or $B \leftrightarrow C$ this may well be the case. In particular, these results are in nice agreement with Fig. 4 (right): there we observe an oscillation $A \leftrightarrow D$ for about 200 seconds, whereas the longest stay in the neighborhood of the minimum C only lasts about 60 seconds.

References

[1] M.P. Allen and D.J. Tildesley. *Computer Simulations of Liquids*. Clarendon Press, Oxford, 1990.

[2] G. Benettin and A. Giorgilli. On the Hamiltonian interpolation of near to the identity symplectic mappings with applications to symplectic integration algorithms. *J. Stat. Phys.*, 74, 1994.

[3] F. A. Bornemann and Ch. Schütte. A mathematical approach to smoothed molecular dynamics: Correcting potentials for freezing bond angles. *(to appear in Physica D, 1997)*, (SC 95–30).

[4] R. Bowen and D. Ruelle. The ergodic theory of axiom A flows. *Invent. math.*, **29**:181–202, 1975.

[5] K.L. Chung. *Markov chains with stationary transition probabilities*, volume 104 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, Göttingen, Heidelberg, 1960.

[6] M. Dellnitz and A. Hohmann. The computation of unstable manifolds using subdivision and continuation. In H.W. Broer, S.A. van Gils, I. Hoveijn, and F. Takens, editors, *Nonlinear Dynamical Systems and Chaos*, pages 449–459. Birkhäuser, *PNLDE* **19**, 1996.

[7] M. Dellnitz and A. Hohmann. A subdivision algorithm for the computation of unstable manifolds and global attractors. To appear in *Numerische Mathematik*, 1996.

[8] M. Dellnitz and O. Junge. Adaptive box refinement in subdivision techniques for the approximation of dynamical behavior. Submitted, 1996.

[9] M. Dellnitz and O. Junge. On the approximation of complicated dynamical behavior. Submitted, 1996.

[10] P. Deuflhard and F. Bornemann. *Numerische Mathematik II — Integration gewöhnlicher Differentialgleichungen*. Walter de Gruyter, Berlin, New York, 1994.

[11] P. Deuflhard, T. Friese, F. Schmidt, R. März, and H.-P. Nolting. Effiziente Eigenmodenberechnung für den Entwurf integriert–optischer Chips. In W. Jäger, Th. Lohmann, and H. Schunck, editors, *Mathematik – Schlüsseltechnologie für die Zukunft*. Springer Verlag.

[12] J. Ding, Q. Du, and T. Y. Li. High order approximation of the Frobenius-Perron operator. *Appl. Math. Comp.*, **53**:151–171, 1993.

[13] J.L. Doob. *Stochastic Processes*. Wiley, 1960.

[14] S. Duane, A.D. Kennedy, B.J. Pendleton, and D. Roweth. Hybrid Monte Carlo. *Phys. Letters B*, 195(2):216–222, 1987.

[15] A.M. Ferrenberg and R.H. Swendsen. New Monte Carlo technique for studying phase transitions. *Phys. Rev. Letters*, 61(23):2635–2638, 1988.

[16] M. Fixman. Classical statistical mechanics of constraints: a theorem and applications to polymers. *Proc. Nat. Acad. Sci.*, 71:3050–3053, 1974.

[17] T. Friese. Eine nichtlineare Mehrgitter–Methode für das Eigenwertproblem linearer nicht–selbstadjungierter Operatoren. *Work done in preparation of a dissertation*, 1997.

[18] H. Grubmueller and P. Tavan. Molecular dynamics of conformational substates for a simplified protein model. *J. Chem. Phys.*, 101, 1994.

[19] R. Haberlandt, S. Fritzsche, G. Peinel, and K. Heinzinger. *Moleküldynamik*. Vieweg, Braunschweig, Wiesbaden, 1995.

[20] E. Hairer. Backward analysis of numerical integrators and symplectic methods. *Annals of Numerical Mathematics*, 1, 1994.

[21] E. Hairer and Ch. Lubich. The life-span of backward error analysis for numerical integrators. Report, 1996.

[22] E. Hairer and D. Stoffer. Reversible long–term integration with variable step sizes. Report, 1995.

[23] D.W. Heermann and L. Yixue. A global-update simulation method for polymer systems. *Makromol. Chem., Theory Simul.*, 2:299–308, 1993.

[24] Yu. Kifer. General random perturbations of hyperbolic and expanding transformations. *J. Analyse Math.*, **47**:111–150, 1986.

[25] Yu. Kifer. Computations in dynamical systems via random perturbations. 1996. Preprint.

[26] R. Kurt. *Axiomatics of Classical Statistical Mechanics*. Pergamon Press, Oxford, New York, 1980.

[27] E. Leontidis, B.M. Forrest, A.H. Widmann, and U.W. Suter. Monte Carlo algorithms for the atomistic simulation of condensed polymer phases. *J. Chem. Soc. Faraday Trans.*, 91(16):2355–2368, 1995.

[28] B. Mehlig, D.W. Heermann, and B.M. Forrest. Hybrid Monte Carlo method for condensed-matter systems. *Phys. Review B*, 45(2):679–685, 1992.

[29] S. Nose. A molecular dynamics methods for simulations in the canonical ensemble. *Mol. Phys.*, 52, 1984.

[30] J.E. Osborn. Spectral approximation for compact operators. *Math. Comp.*, **29**(131):712–725, 1975.

[31] S. Reich. Smoothed dynamics of highly oscillatory Hamiltonian systems. *Physica D*, 89:28–42, 1995.

[32] D. Ruelle. A measure associated with axiom A attractors. *Amer. J. Math.*, **98**:619–654, 1976.

[33] J.M. Sanz-Serna and M.P. Calvo. *Numerical Hamiltonian Systems*. Chapman and Hall, London, Glasgow, New York, Tokyo, 1994.

[34] Ch. Schütte and F. A. Bornemann. Homogenization approach to smoothed molecular dynamics. In V. Lakshmikantham, editor, *World Congress of Nonlinear Analysts '96. Proceedings of the Second Congress held in Athens, Greece*. Walter de Gruyter, Berlin, New York (to appear).

[35] A.D. Sokal. Monte Carlo methods in statistical mechanics. Lecture note, Department of Physics, New York University, 1989.

[36] G. Winkler. *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*. Springer-Verlag, Berlin, Heidelberg, 1995.

Appendix A: Theoretical Background

Theoretically it is known that in the hyperbolic case the discretization of the Frobenius-Perron operator by the Galerkin method as outlined in this article indeed leads to an approximation of an underlying SBR-measure – if it exists. In this appendix we briefly summarize the related results from [9].

A.1 Stochastic Transition Functions

The theoretical results rely on the concept of *small random perturbations* of dynamical systems. Hence we begin by recalling some basic notions and results on Markov processes that will be needed later on. For a detailed introduction the reader is referred to [13].

Invariant Measures

First we describe the notion of an invariant measure in the stochastic framework. We assume that Γ is compact equipped with a σ-algebra $\mathcal{A}$.

Definition A.1 A function $p : \Gamma \times \mathcal{A} \to \mathbb{R}$ is a *stochastic transition function*, if

(i) $p(x, \cdot)$ is a probability measure for every $x \in \Gamma$,

(ii) $p(\cdot, B)$ is Lebesgue-measurable for every measurable B.

Example A.2 Let δ_y denote the Dirac measure supported on the point $y \in \Gamma$. Then $p(x, B) = \delta_{h(x)}(B)$ is a stochastic transition function for every m-measurable function h. Moreover, the specific choice $h = f$ represents the deterministic situation in this more general set up.

Definition A.3 Let p be a stochastic transition function. If $\mu \in \mathcal{M}$ satisfies

$$\mu(B) = \int p(x, B)\, d\mu(x)$$

for all measurable B, then μ is an *invariant measure of* p.

Absolutely Continuous Stochastic Transition Functions

Now we assume that for every $x \in \Gamma$ the probability measure $p(x, \cdot)$ is absolutely continuous with respect to the Lebesgue measure m. Hence we may write $p(x, \cdot)$ as

$$p(x, B) = \int_B k(x, y)\, dm(y) \quad \text{for all measurable } B,$$

with an appropriate *transition density function* $k : \Gamma \times \Gamma \to \mathbb{R}$. Obviously,

$$k(x, \cdot) \in L^1(\Gamma, m) \quad \text{and} \quad k(x, y) \geq 0.$$

In this case we also call the stochastic transition function p *absolutely continuous*. Note that

$$\int k(x,y)\,dm(y) = p(x,\Gamma) = 1 \quad \text{for all } x \in \Gamma.$$

The Frobenius–Perron Operator

DEFINITION A.4 Let p be a stochastic transition function. Then the *Frobenius–Perron operator* $P : \mathcal{M}_{\mathbb{C}} \to \mathcal{M}_{\mathbb{C}}$ is defined by

$$P\mu(B) = \int p(x,B)\,d\mu(x),$$

where $\mathcal{M}_{\mathbb{C}}$ is the space of bounded complex valued measures. If p is absolutely continuous with density function k then we may define the Frobenius–Perron operator P on L^1 by

$$Pg(y) = \int k(x,y)g(x)\,dm(x) \quad \text{for all } g \in L^1.$$

REMARK A.5 By definition a measure $\mu \in \mathcal{M}$ is invariant if and only if it is a fixed point of P. In other words, as in the deterministic case invariant measures correspond to eigenmeasures of P for the eigenvalue one.
Moreover, let $\lambda \in \mathbb{C}$ be an eigenvalue of P with corresponding eigenmeasure ν, that is, $P\nu = \lambda\nu$. Then in particular

$$\lambda\nu(\Gamma) = P\nu(\Gamma) = \int p(x,\Gamma)\,d\nu(x) = \nu(\Gamma)$$

since $p(x,\Gamma) = 1$ for all $x \in \Gamma$. It follows that $\nu(\Gamma) = 0$ if $\lambda \neq 1$.

A.2 Convergence to SBR-measures in the Hyperbolic Case

Small Random Perturbations

Recall that the purpose is to approximate the Frobenius-Perron operator of a *deterministic* dynamical system represented by a diffeomorphism f. Hence the approximating *stochastic* system that we consider should be a small perturbation of the original deterministic system.
For $\varepsilon > 0$ we set

$$k_\varepsilon(x,y) = \frac{1}{\varepsilon^n m(D)}\chi_D\left(\frac{1}{\varepsilon}\Big(y - x\Big)\right), \quad x,y \in \Gamma. \tag{A.1}$$

Here $D = D_0(1)$ denotes the open ball in $\mathbb{R}^n$ of radius one and χ_D is the characteristic function of D. Obviously $k_\varepsilon(f(x),y)$ is a transition density function and we may define a stochastic transition function p_ε by

$$p_\varepsilon(x,B) = \int_B k_\varepsilon(f(x),y)\,dm(y). \tag{A.2}$$

REMARK A.6 Note that $p_\varepsilon(x,\cdot) \to \delta_{f(x)}$ for $\varepsilon \to 0$ uniformly in x in a weak*–sense. Hence the Markov process defined by any initial probability measure μ and the transition function p_ε is a *small random perturbation* of the deterministic system f in the sense of Yu. Kifer ([24]).

Now observe that

$$\iint |k_\varepsilon(f(x),y)|^2 \, dm(x)dm(y) \le \left(\frac{m(\Gamma)}{\varepsilon^n m(D)}\right)^2 < \infty,$$

and therefore the corresponding Frobenius-Perron operator $P_\varepsilon : L^2 \to L^2$ is compact. Since the invariant densities are not just in L^1 but even in L^∞ the restriction of P_ε to L^2 is perfectly reasonable.

Approximation of SBR-Measures

The idea is to combine classical convergence results for compact operators with results from Ergodic Theory on the convergence of invariant measures of small random perturbations to SBR-measures with decreasing magnitude of the perturbation. Let us be more precise. Suppose that the diffeomorphism f possesses a hyperbolic attractor Λ with an SBR-measure μ_{SBR}, and let p_ε be a small random perturbation of f. Then, under certain hypotheses on p_ε, it is shown in [24] that the invariant measures of p_ε converge in a weak*–sense to μ_{SBR} as $\varepsilon \to 0$. On the other hand standard results on compact operators (see [30]) guarantee that the relevant eigenmeasures of P_ε are approximated by our Galerkin projection, and this leads to the desired convergence result.

THEOREM A.7 ([9]) *Suppose that the diffeomorphism f has a hyperbolic attractor Λ, and that there exists an open set $U_\Lambda \supset \Lambda$ such that*

$$k_\varepsilon(x,y) = 0 \quad \text{if } x \in \overline{f(U_\Lambda)} \text{ and } y \notin U_\Lambda.$$

Then the transition function p_ε in (A.2) has a unique invariant measure π_ε with support on Λ and the approximating measures

$$\mu_d^\varepsilon(A) = \int_A g_d^\varepsilon \, dm$$

obtained by the Galerkin method described in Section 4.1 converge in a weak–sense to the SBR–measure μ_{SBR} of f as $\varepsilon \to 0$ and $d \to \infty$,*

$$\lim_{\varepsilon\to 0} \lim_{d\to\infty} \mu_d^\varepsilon = \mu_{SBR}. \tag{A.3}$$

REMARK A.8 Recently Yu. Kifer [25] has obtained a similar convergence result by a discretized version of his methods and results on stochastic perturbations of dynamical systems.