6 Theoretical and Numerical Investigations

Based on the results of the last two sections, we want to theoretically and numerically analyze our four model systems. We will see that under suitable conditions on the potential energy function all model systems but the deterministic Hamiltonian system satisfy the two basic conditions (C1) and (C2). The application to a three–well potential show significant similarities for the model systems on a mesoscopic level.

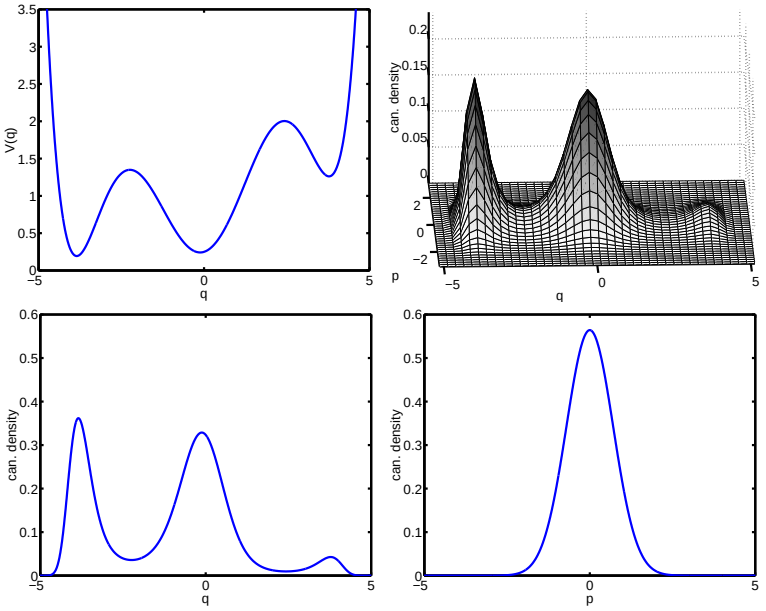

Figure 5: Three well potential (top left) and corresponding canonical density f_{can} for $\beta = 2$ (top right). The positional canonical density $\mathcal{Q}$ (bottom left) and canonical density of momenta $\mathcal{P}$ (bottom right).

In the following we consider the three well potential $V : \Omega \to \mathbf{R}$ with

$$V(q) \quad = \quad \frac{1}{400}(q^6 - 30q^4 + 238q^2 + 56q + 100) \tag{63}$$

as our test system. We choose $\Omega = [-5, 5]$ as the position space with periodic boundary conditions, and modify the potential function at the boundary ± 5 in such a way that it is smooth (cf. Remark below). Its graph, the canonical distribution f_{can} corresponding to $\beta = 2$ and its two marginal distributions $\mathcal{Q}$ and $\mathcal{P}$ are shown in Figure 5. We choose $\tau = 1$ as the observation time span. Intuitively, we would expect to exist three metastable subsets around the (local) minima of the potential function for moderate temperature.

Remark. The positional canonical density Q already indicates that it is very unlike to stay near the periodic boundary at ± 5; this can also be

observed from the realization shown in Figure 6. Hence, numerically it might be difficult to resolve the difference between an unbounded position space and one with periodic boundary conditions. Yet, in special situations this is possible, as we are going to exemplify for the Hamiltonian system with randomized momenta in application to a harmonic potential test system (see end of Section 6.2).

6.1 Deterministic Hamiltonian System

Denote by P_τ the propagator corresponding to the deterministic Hamiltonian system.

Proposition 6.1 *For both bounded systems and periodic systems, the propagator $P_\tau : L^r(\mu_{\mathrm{can}}) \to L^r(\mu_{\mathrm{can}})$ neither satisfies condition (C1) nor condition (C2) in $L^r(\mu_{\mathrm{can}})$ with $r = 1, 2$.*

Proof: For the statement on $r = 1$, we observe that the stochastic transition function $p_\tau(x, \mathrm{d}y) = \delta_{\Phi^\tau x}(\mathrm{d}y)$ is singular w.r.t. the invariant measure $\mu_{\mathrm{can}}(\mathrm{d}y) = f_{\mathrm{can}}(y)\mathrm{d}y$ implying $\Delta(P_\tau) = 1$ by Theorem 4.9. Since $p_\tau^n(x, \mathrm{d}y) = \delta_{\Phi^{n\tau} x}(\mathrm{d}y)$, we more generally have $\Delta(P_\tau^n) = 1$ for every $n \in \mathbf{Z}_+$ and hence $r_{\mathrm{ess}}(P_\tau) = 1$ by Theorem 4.6. This violates condition (C1). Now, define for an arbitrary smooth function $\mathcal{F} : \mathbf{R} \to \mathbf{R}_+$ the density $f : \mathbf{X} \to \mathbf{R}$ by $f(x) = \mathcal{F}(H(x))$ and the thereby induced measure μ_f by $\mu_f(\mathrm{d}y) = f(y)\mathrm{d}y$. Since the Hamiltonian flow is energy–preserving, μ_f is a finite invariant measure for every $f \in L^1(\mu_{\mathrm{can}})$. This way we may construct arbitrarily many invariant probability measures, which violates condition (C2). The statement for $r = 2$ follows from the fact that P_τ is unitary in $L^2(\mu_{\mathrm{can}})$, as stated in Section 2.1. □

Within the proof of Proposition 6.1 we have shown that the deterministic Hamiltonian system admits infinitely many invariant probability measures. Due to this ambiguity, in our context pure Hamiltonian dynamics seems not to be appropriate for modeling internal fluctuations within one specific stationary ensemble, in our case the canonical ensemble.

6.2 Hamiltonian System with Randomized Momenta

Denote by P_τ the propagator corresponding to the Hamiltonian system with randomized momenta. We first state under which conditions on the potential function the two requirements (C1) and (C2) on P_τ hold. Then, we numerically analyze the induced essential statistical behavior w.r.t. the positional canonical ensemble.

Proposition 6.2 *For periodic systems with position space $\Omega \subset \mathbf{R}^d$, some fixed observation time span $\tau > 0$ and smooth periodic potential function*

$V : \Omega \to \mathbf{R}$, the propagator $P_\tau : L^r(\mu_\mathcal{Q}) \to L^r(\mu_\mathcal{Q})$ satisfies the conditions (C1) and (C2) in $L^r(\mu_\mathcal{Q})$ with $r = 1, 2$.

Proof: The statement for $r = 1$ is based on results in [68]. Due to [68, Lem. 4.51 and Prop. 4.18] the Lebesgue decomposition of the stochastic transition function $p_\tau(q, \mathrm{d}y) = p_a(q, y)\mu_\mathcal{Q}(\mathrm{d}y) + p_s(q, \mathrm{d}y)$ has the following two properties:

(i) the absolutely continuous part satisfies: $\operatorname*{ess\,sup}_{(q,y)\in\Omega} p_a(q, y) < \infty$.

(ii) the singular part satisfies: $\operatorname*{ess\,sup}_{q\in\Omega} p_a(q, \Omega) < 1$.

Application of Theorem 4.9 proves that condition (C1) holds. Since P_τ is asymptotically stable according to [68, Lem. 4.51], condition (C2) is a consequence of Corollary 4.22. In order to prove the statement for $r = 2$ we note that due to [68, Lem. 4.31] the Markov process corresponding to the Hamiltonian system with randomized momenta is $\mu_\mathcal{Q}$–irreducible. Therefore, the statement is a consequence of Theorem 4.32. Using a different approach, the statement for $r = 2$ was already proved by Schütte in [68]. □

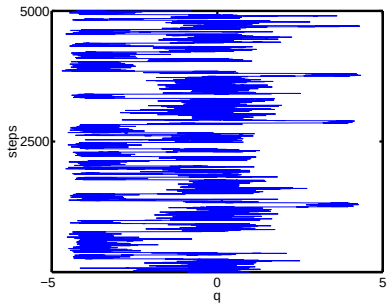

Figure 6: Typical realization of the Hamiltonian system with randomized momenta for $\beta = 2$, observation time span $\tau = 1$ and initial distribution $Q_0 \sim 1$.

By Proposition 6.2 the application of our algorithmic approach to the test systems is theoretically justified. In order to discretize the propagator, we proceed according to Example 5.1 using the Trapezoid rule with $N = 300000$, $M = 1$ and the Leapfrog discretization [65] of the Hamiltonian flow with internal step size $\Delta t = 0.02$. A typical realization is shown in Figure 6. We observe that the Markov process stays for some time close to one of the three (local) minima, then suddenly jumps close to another minimum, stays there for a while, jumps again and so on. Hence, by looking at the realization we visually identify three metastable subsets. Discretizing the state space $\Omega = [-5, 5]$ with periodic boundary conditions into 30 equal–sized intervals, we obtain a 30×30 stochastic transition matrix S. Solving the eigenvalue

problem for S yields:

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	...	λ_{30}
1.000	0.975	0.958	0.599	0.490	0.369	...	-0.435

Evaluating the indicator for $r_{\text{ess}}(P_\tau)$ we get $[\Delta(P_\tau)] = 0.557$—for a further analysis of the indicator within a hierarchical context see below. Looking at the spectrum of S, we identify a cluster of three eigenvalues $\{\lambda_1, \lambda_2, \lambda_3\}$ close to 1 that is well separated from the remaining part of the spectrum by a gap. Hence, in view of our algorithmic strategy we look for a decomposition into three metastable subsets. The eigenfunctions corresponding to the largest eigenvalues are shown in Figure 7 (left). We observe *almost constant levels* around the three minima for the first three eigenfunctions, while the fourth eigenfunction does not show this particular structure. This almost constant level structure is exploited by the identification algorithm outlined in Section 5.4. Application to our example yields a clustering $\{C_1, C_2, C_3\}$ with $C_1 = \{q \leq -2.1\}$, $C_2 = \{-2.1 < q \leq 1.8\}$ and $C_3 = \{1.8 < q\}$. The statistical weights $\mu(C_k)$ within the canonical ensemble $\mu_{\mathcal{Q}}$ and the metastabilities $p(\tau, C_k, C_k)$ are given by the following table:

metastable subset	C_1	C_2	C_3
statistical weight	0.324	0.616	0.060
metastability	0.966	0.973	0.908

The essential statistical behavior, i.e., the probability of transitions between the metastable subsets, is described by the coupling matrix $C = (c_{jk})_{j,k=1,2,3}$ with $c_{jk} = p(\tau, C_j, C_k)$. For our example, we obtain

$$C = \begin{pmatrix} 0.966 & 0.034 & 0 \\ 0.018 & 0.973 & 0.009 \\ 0 & 0.092 & 0.908 \end{pmatrix}.$$

Analyzing only the coupling matrix C and the metastability of the clusters, we would predict that a typical realization of the Markov process would stay most of the time in C_2, sometimes moving to C_1, stay there for some time, then moving back and so on. Rarely, there will be transitions to C_3, where in addition the Markov process is unlikely to stay for a while. This is what we observed for the realization shown in Figure 6. In this sense the clustering, its metastabilities and the corresponding coupling matrix allow to describe the essential statistical behavior of the model system.

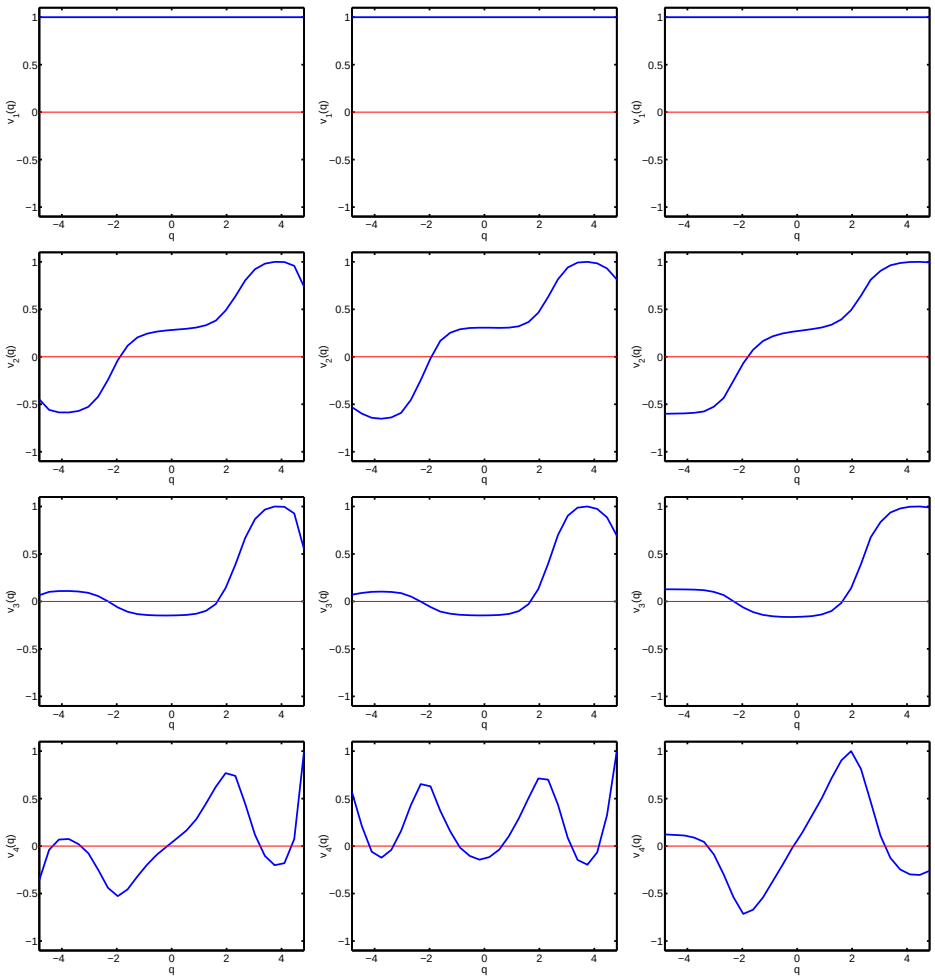

Figure 7: The four dominant eigenfunctions of the propagator P_τ for different model systems. Left: Hamiltonian system with randomized momenta corresponding to the eigenvalues $1.000, 0.975, 0.958, 0.599$ (from top to bottom). Middle: Langevin equation for $\gamma = 1.0$ corresponding to the eigenvalues $1.000, 0.969, 0.949, 0.430$; eigenfunctions *projected* on the position space. Right: Smoluchowski equation for $\gamma = 1.0$ corresponding to the eigenvalues $1.000, 0.950, 0.915, 0.387$.

Let us have a closer look at the indicator $[\Delta(P_\tau)]$. As suggested in Section 5.1, the indicator should be used within some hierarchy of decompositions. The following table shows the indicator for our test system evaluated in a hierarchy of decompositions into m equal–sized intervals:

m	30	50	100	200	500	1000
$[\Delta(P_\tau)]$	0.557	0.423	0.315	0.207	0.152	0.117

We deduce that the essential spectral radius is likely to be much less than the initially value of $[\Delta(P_\tau)] = 0.557$ indicated. Moreover, while $[\Delta(P_\tau)]$ drastically decays for finer decompositions of the state space, the dominant eigenvalues of P_τ are quite insensitive w.r.t. refinements. This was already illustrated for a small molecule in [69]. However, we do not know the exact value of $r_{\text{ess}}(P_\tau)$ in order to validate the results.

This is possible for the systems presented in Example 4.10 (ii), where we considered a harmonic potential on the position space $\Omega = [-1, 1]$ with periodic boundary conditions. Let us choose $\beta = 2$ and consider the corresponding propagator P_τ. Analytically, we have $\Delta(P_\tau) = 0.841$ according to Example 4.10. The following table shows the indicator $[\Delta(P_\tau)]$ based on different decompositions of the state space into m equal–sized intervals:

m	30	50	100	200
$[\Delta(P_\tau)]$	0.846	0.849	0.850	0.859

We observe that a finer decomposition of the state space does not significantly change the value of $[\Delta(P_\tau)]$; the results are in good agreement with the analytical value. For the unbounded state space $\Omega = \mathbf{R}$ described in Example 4.10 (i) we have $P_\tau = \text{Id}$ and hence $r_{\text{ess}}(P_\tau) = 1$. This perfectly illustrates the influence of periodic boundary conditions and also demonstrates that it is possible to resolve the difference to the unbounded state space numerically.

6.3 Langevin Equation

Let P_τ denote the propagator corresponding to the Langevin Markov process and p_τ its stochastic transition function. Up to now, the stability properties of the Langevin Markov process are only partially understood. While (μ–a.e.) geometric or V–uniform ergodicity can be proved for both bounded systems and periodic systems under reasonable conditions on the potential function [51, 76], little is known about (μ–a.e.) uniform ergodicity. It is believed that the Markov process is not (μ–a.e.) uniformly ergodic for bounded systems, while the periodic case is even less understood [33]. These topics are subject to current investigations. Since the Langevin Markov process

is not reversible, it is an ideal test system to demonstrate the effects of time–symmetrization. Hence, we analyze both the "original" as well as the time–symmetrized Markov process.

Proposition 6.3 *The following holds for the Langevin equation:*

(i) Fix some observation time space $\tau > 0$ and consider a periodic system *with position space $\Omega \subset \mathbf{R}^d$ and smooth potential function $V : \Omega \to \mathbf{R}$. Then, μ_{can} is the unique invariant probability measure of the Langevin Markov process $X_n = \{X_{n\tau}\}_{n \in \mathbf{Z}_+}$ and its stochastic transition function p_τ is (μ_{can}–a.e.) geometrically and V–uniformly ergodic.*

(ii) Fix some observation time space $\tau > 0$ and consider a bounded system *with position space $\Omega = \mathbf{R}^d$ and smooth potential function $V : \Omega \to \mathbf{R}$ satisfying $V(q) \geq 0$ for $q \in \Omega$ and growing at infinity like $\|q\|^{2l}$ for some positive integer l. Then, μ_{can} is the unique invariant probability measure of the Langevin Markov process $X_n = \{X_{n\tau}\}_{n \in \mathbf{Z}_+}$ and its stochastic transition function p_τ is (μ_{can}–a.e.) geometrically and V–uniformly ergodic.*

Proof: Statement (i) is an immediate result of Theorem 3.1 by Stuart [76], while statement (ii) immediately follows from Theorem 3.2 by Mattingly et al. [51]. □

We remark that according to Mattingly et al. [51] the condition on the growth rate of the potential function in the case of bounded systems can be further weakened (see cond. 3.1 in [51]). While we are able to prove (μ_{can}–a.e.) geometric and V–uniform ergodicity for the Langevin equation, reversibility fails to hold, since the infinitesimal generator $\mathcal{L}$ and therefore the propagator P_τ is not self–adjoint in $L^2(\mu_{\text{can}})$. In this case, Theorem 4.31 states nothing about the validity of the conditions (C1) and (C2) in $L^2(\mu_{\text{can}})$. Although the application of our algorithmic approach is not (yet) theoretically justified, the numerical results presented below are very promising.

In order to discretize the propagator in application to our test system, we proceed according to Example 5.2, hence exploit a realization of the discrete time Markov process X_n. We use the Leapfrog discretization for the deterministic part with internal step size $\Delta t = 0.02$ and random variables $N_n \sim \sigma\sqrt{\Delta t}\,\mathcal{N}(0,1)$ for a realization of the white noise (within each internal step). We discretize the state space $\mathbf{X} = \Omega \times \mathbf{R}$ with $\Omega = [-5, 5]$ and periodic boundary conditions into $m = 900$ subsets as follows: discretize the position space $\Omega = [-5, 5]$ into 30 equal–sized intervals and the momenta space $\mathbf{R}$ by partitioning $[-3, +3]$ into 28 equal–sized intervals and adding the two infinite intervals $(-\infty, -3]$ and $[3, \infty)$. Note that in view of the momenta

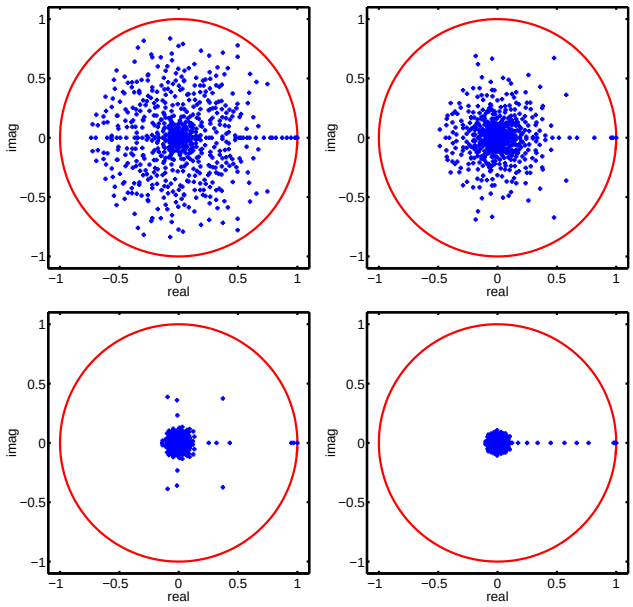

Figure 8: Spectrum of P_τ for the Langevin equation for $\gamma = 0.01$ (upper left), $\gamma = 0.16$ (upper right), $\gamma = 1.0$ (lower left) and $\gamma = 4.0$ (lower right).

distribution shown in Figure 5, the statistical weight of the infinite intervals is negligible. This way, we obtain a 900×900 non self–adjoint stochastic transition matrix S. Solving the eigenvalue problem for S yields

λ_1	λ_2	λ_3	λ_4	$\lambda_{5/6}$	...
1.000	0.969	0.949	0.430	$0.369 \pm 0.372i$	...

with $[\Delta(P_\tau)] = 0.750$ for $m = 400$ discretization subsets and $[\Delta(P_\tau)] = 0.667$ for $m = 900$. Hence, $[\Delta(P_\tau)]$ indicates that the first three dominant eigenvalues belong to the discrete spectrum and are separated by a spectral gap from the remaining part of the spectrum. This is in agreement with the fact that the dominant eigenvalues are rather insensitive w.r.t. further refinements of the decomposition. The next table show the eigenvalues of *maximal modulus* for different decompositions into m subsets:

m	λ_1	λ_2	λ_3	λ_4
400	1.000	0.968	0.947	0.412
900	1.000	0.969	0.949	0.430
1600	1.000	0.969	0.949	0.436
2500	1.000	0.969	0.950	0.440

We deduce from this results that numerically the two conditions (C1) and (C2) are satisfied and proceed with the case $m = 900$. The eigenfunctions corresponding to the largest eigenvalues are shown in Figure 10 (middle),

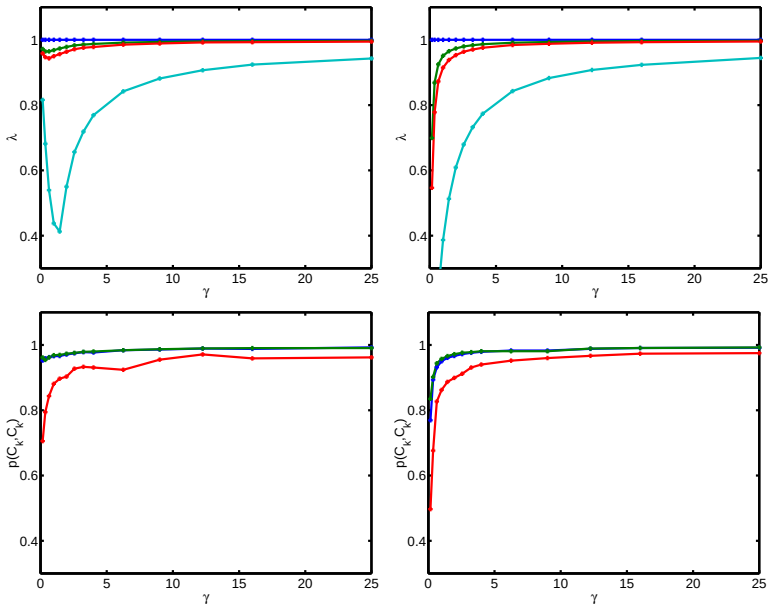

Figure 9: First four dominant eigenvalues (maximal real part) of P_τ for the Langevin equation (top left) and Smoluchowski equation (top right) for different values of γ. Metastabilities of the three subsets for the Langevin equation (bottom left) and Smoluchowski equation (bottom right) for different values of γ.

their projections onto the position space are depicted in Figure 7 (middle). We remark that, since the first three eigenvalues are real, the corresponding eigenfunctions can be chosen real–valued. Moreover, because the largest scalar product between either two different normalized eigenfunctions is of order 10^{-3}, the first three eigenfunctions are almost orthogonal, although the propagator is not reversible. The same holds for the projected eigenfunctions. Hence, the dominant part of P_τ is nearly self–adjoint and we would identify metastable subsets by application of the identification algorithm. However, we proceed in a slightly different way, since our aim is to identify conformations induced by the Langevin equation.

Recall from Section 1.1 that conformations are thought to be objects in the position space. For reduced models acting only on the position space Ω like, e.g., the Hamiltonian system with randomized momenta, the notions of conformations and metastable subsets coincide (see end of Sec. 1.1). This is different for models acting on the phase space $\Gamma = \Omega \times \mathbf{R}$ (positions and momenta) like, e.g., the Langevin equation. In this case we characterized conformations as special metastable subsets of the form $C = \hat{C} \times \mathbf{R} \subset \Gamma$ with $\hat{C} \subset \Omega$. Hence, for every position $q \in \hat{C}$, the conformation C contains all states of the form (q, p) with $p \in \mathbf{R}$. Our strategy to identify conformations induced by the Langevin equation is therefore as follows: In the position space Ω, run the identification algorithm based on the first three *projected*

eigenfunctions and extend the identified metastable subsets $\hat{C}_j \subset \Omega$ to the "cylindrical" subsets $C_j = \hat{C}_j \times \mathbf{R} \subset \mathbf{X}$. Finally check for metastability of the C_j w.r.t. the Langevin Markov process in Γ.

Applying this strategy[13], we end up with a clustering $\{C_1, C_2, C_3\}$ with $C_1 = [q \leq -2.0] \times \mathbf{R}$, $C_2 = [-2.0 < q \leq 1.7] \times \mathbf{R}$ and $C_3 = [1.7 < q] \times \mathbf{R}$. The statistical weights $\mu(C_k)$ within the canonical ensemble μ_{can} and the metastabilities $p(\tau, C_k, C_k)$ are given by the following table:

metastable subset	C_1	C_2	C_3
statistical weight	0.350	0.594	0.055
metastability	0.963	0.966	0.884

(64)

Calculating the coupling matrix yields

$$C = \begin{pmatrix} 0.963 & 0.037 & 0 \\ 0.023 & 0.966 & 0.011 \\ 0 & 0.116 & 0.884 \end{pmatrix}. \tag{65}$$

So far we have analyzed the Langevin equation for the specific choice of friction constant $\gamma = 1$ and stochastic excitation $\sigma = 1$, resulting in the inverse temperature $\beta = 2$. Recalling that $\beta = 2\gamma/\sigma^2$, we find that there is a one–parameter family of Langevin equations corresponding to the same inverse temperature: choose $\sigma = \sqrt{2\gamma/\beta}$ for arbitrary $\gamma > 0$. Figure 8 illustrates the dependence of the spectrum on γ. It shows the spectrum of the stochastic transition matrix S obtained from discretizing the propagator P_τ corresponding to $\beta = 2$ and different values of γ. We observe that for small values of γ the spectrum of S is spread all over the unit disc, while it concentrates more and more on the interval $[0, 1]$ for larger values of γ. This reflects the fact that the Langevin equation is similar to the deterministic Hamiltonian system for $\gamma \approx 0$ [23], while it is similar to the Smoluchowski equation for $\gamma \gg 1$. Hence, the Langevin equation exhibits different behavior depending on the friction constant γ. This might also explain the dependence of the dominant eigenvalues on γ, as shown in Figure 9 (top). The eigenvalues accumulate in $\lambda = 1$ for $\gamma \approx 0$ as well as for $\gamma \gg 1$. The former effect reflects the fact that the deterministic Hamiltonian system admits infinitely many invariant probability distributions (see Sec. 6.1), while the latter effect is due to the behavior of the Smoluchowski equation for large γ.

[13] We could think of different strategies like, e.g., projecting the Langevin Markov process onto the position space Ω and considering the corresponding propagator. Alternatively, we could modify the Langevin model similar to the Hamiltonian system with randomized momenta in order to obtain a Langevin system with randomized momenta in the position space only. Our strategy is motivated by the fact that it fits best our context, as we are going to see below and in Section 6.5.

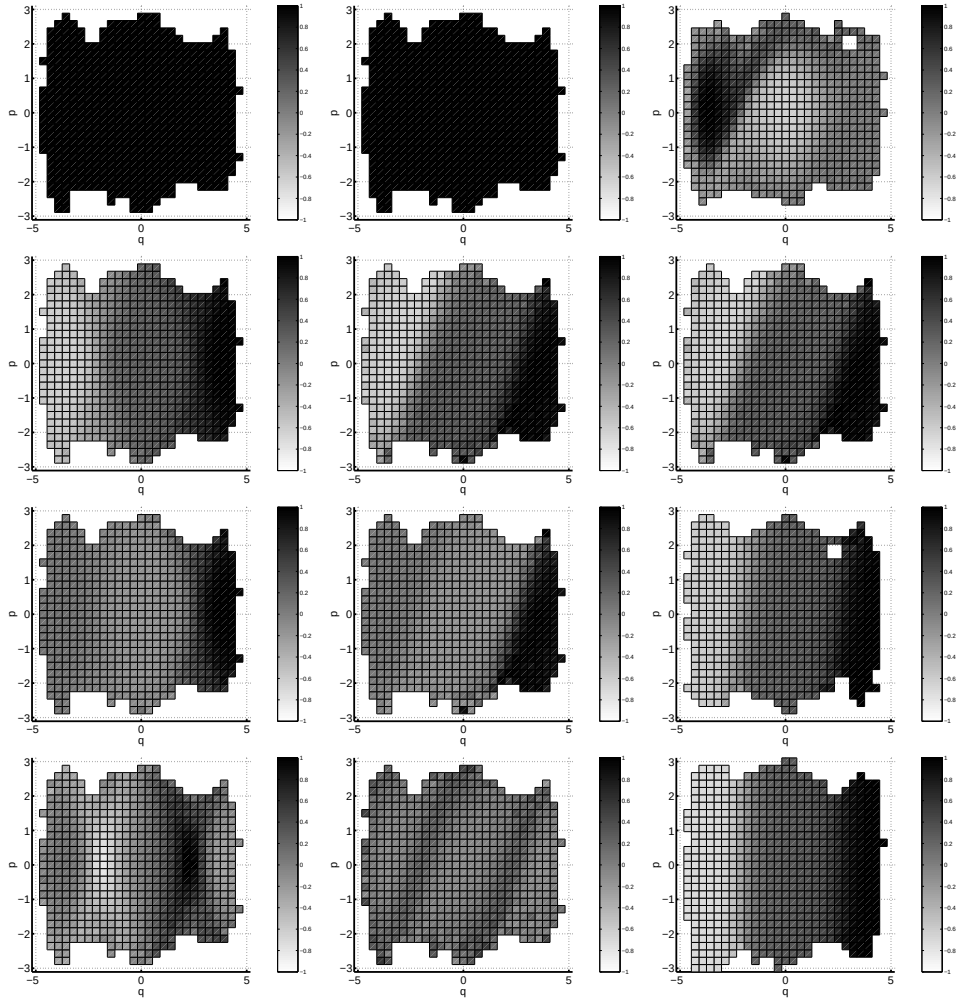

Figure 10: Left and middle: Dominant eigenfunctions for the Langevin equation for $\gamma = 1.0$. Eigenfunctions of the time–symmetrized discretization of P_τ corresponding to the eigenvalues $1.000, 0.984, 0.974, 0.701$ (left) and eigenfunctions corresponding to the discretization of P_τ corresponding to $1.000, 0.969, 0.949, 0.430$ (middle). Right: Second eigenfunction of P_τ for $\gamma = 0.16, 1.0, 4.0, 16.0$ (top to bottom).

Seemingly in contrast to the behavior of the eigenvalues is the decay of metastability of the corresponding clusterings, as shown in Figure 9 (bottom). This effect is due to our definition of conformations as cylindrical metastable subsets (see Sec. 1.1). Due to energy conservation in the deterministic Hamiltonian case, we would expect that subsets of the state space $\mathbf{X} = \Gamma$ corresponding to a certain energy range get more and more metastable for small γ. Since these subsets follow contour lines of H or, equivalently, of f_{can} (see Figure 5), every "cylindrical" subset will necessarily only partially intersect with these energy range subsets. This causes loss of metastability and explains the decay of metastability for small values of γ. Nevertheless, we would expect to find metastable subsets—subject to no restriction—for small γ, since the energy fluctuation of the Langevin equation decay for γ tending to zero. The mentioned structure due to nearly energy conservation is already visible in Figure 10 for the top left eigenfunction (see also [49]).

In the last part of this section we want to study the effects of time–symmetrization. Since the Langevin equation is not reversible, it is an ideal test model. In order to facilitate a comparison of the original and the time–symmetrized approach, we use the same realization of the Langevin Markov process X_n for the subsequent numerical discretization as above. The next proposition is a direct application of Theorem 4.31.

Proposition 6.4 *Fix some observation time span $\tau > 0$ and assume that the stochastic transition function r_τ corresponding to the time–symmetrized Langevin Markov process $Z_n = \{Z_n\}_{n \in \mathbf{Z}_+}$ is geometrically ergodic. Then the propagator $P_\tau : L^2(\mu_{\text{can}}) \to L^2(\mu_{\text{can}})$ corresponding to the time–symmetrized Markov process satisfies the conditions (C1) and (C2) in $L^2(\mu_{\text{can}})$.*

The assumption in Proposition 6.4 is in particular satisfied, if the stochastic transition function corresponding to the original Langevin Markov process can proved to be uniformly ergodic. We now discretize the propagator P_τ corresponding to the time–symmetrized Langevin Markov process according to Example 5.2. Proceeding as in the first part, we end up with a *self–adjoint* 900×900 stochastic transition matrix S. Solving the eigenvalue problem for S yields:

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	...
1.000	0.984	0.974	0.701	0.649	0.539	...

In view of our algorithmic strategy, we are looking for a decomposition into three metastable subsets. The eigenfunctions corresponding to the largest eigenvalues are shown in Figure 10 (left). Application of our identification strategy yields a clustering $\{C_1, C_2, C_3\}$ with $C_1 = [q \leq -1.9] \times \mathbf{R}$, $C_2 =$

$[-1.9 < q \leq 1.8] \times \mathbf{R}$ and $C_3 = [1.8 < q] \times \mathbf{R}$. The statistical weights $\mu(C_k)$ within the canonical ensemble μ_{can} and the metastabilities $p(\tau, C_k, C_k)$ are given by the following table:

metastable subset	C_1	C_2	C_3
statistical weight	0.358	0.585	0.057
metastability	0.975	0.977	0.921

(66)

The coupling matrix based on the realization of the *original* Langevin Markov process X_n w.r.t. the clusters based on the time–symmetrized Markov process is given by

$$C = \begin{pmatrix} 0.975 & 0.025 & 0 \\ 0.015 & 0.977 & 0.008 \\ 0 & 0.079 & 0.921 \end{pmatrix}. \tag{67}$$

As predicted by Theorem 3.2 the metastability w.r.t. the time–symmetrized Markov process Z_n (see table (66)) and the metastability w.r.t. the original Markov process X_n (see table (64)) are identical. However, comparing the two coupling matrices (65) and (67), we see that a clustering based on the time–symmetrized process might differ from results based on the original process. This a consequence of the different resulting propagotors and corresponding (projected) eigenfunctions.

6.4 Smoluchowski Equation

Denote by P_τ the propagator corresponding to the Smoluchowski Markov process. We first state under which conditions on the potential function the two requirements (C1) and (C2) on P_τ hold. Then, we numerically analyze the induced essential statistical behavior w.r.t. the positional canonical ensemble $\mu_{\mathcal{Q}}$.

Proposition 6.5 *The following holds for the Smoluchowski equation:*

(i) For periodic systems *with position space $\Omega \subset \mathbf{R}^d$, some fixed observation time space $\tau > 0$ and smooth potential function $V : \Omega \to \mathbf{R}$, the propagator $P_\tau : L^1(\mu_{\mathcal{Q}}) \to L^1(\mu_{\mathcal{Q}})$ satisfies $r_{\text{ess}}(P_\tau) = 0$ and is asymptotically stable, hence conditions (C1) and (C2) are fulfilled in $L^1(\mu_{\mathcal{Q}})$.*

(ii) For bounded systems *with position space $\Omega = \mathbf{R}^d$, some fixed observation time space $\tau > 0$ and smooth potential function $V : \Omega \to \mathbf{R}$ satisfying for some integer $\alpha > 0$ the growth conditions $V(q) \sim \|q\|^{2\alpha}$, $\nabla V(q) \sim \|q\|^{2\alpha-1}$ and $\partial^2 V(q) \sim \|q\|^{2\alpha-2}$ as $\|q\| \to \infty$, the propagator $P_\tau : L^2(\mu_{\mathcal{Q}}) \to L^2(\mu_{\mathcal{Q}})$ satisfies the conditions (C1) and (C2) in $L^2(\mu_{\mathcal{Q}})$.*

Proof: For periodic systems, application of Theorem 3 of Ichihara and Kunita [36] yields that the stochastic transition function $p(t, q, \mathrm{d}y)$ is absolutely continuous w.r.t. $\mu_{\mathcal{Q}}$ admitting a C^∞ density. Hence, the density is bounded, since Ω is compact, and therefore $r_{\mathrm{ess}}(P_\tau) = 0$ by Theorem 4.9. Furthermore, we get asymptotic stability of P_τ according to [46], hence conditions (C1) and (C2) hold in $L^1(\mu_{\mathcal{Q}})$. For bounded systems, the statement follows by Theorem 4.31, if we can prove that the stochastic transition function is V–uniformly ergodic. But this is an immediate result of Theorem 5.3 by Mattingly et al. [51]. □

Due to Mattingly et al. [51], the condition on the growth rate of the potential function in the case of bounded systems can be weakened (see cond. 5.1 in [51]). For an analysis of the Smoluchowski equation from a symmetric Markov semigroup point of view see Davies [8].

By Proposition 6.5 the application of the algorithmic approach to our test systems is theoretically justified. In order to discretize the propagator, we proceed according to Example 5.1 using the Trapzoid rule with $N = 300000$, $M = 1$ and the forward Euler or Euler–Maruyama scheme [41] with internal step size $\Delta t = 0.02$. A realization looks comparable to Figure 6. Discretizing the state space $\Omega = [-5, 5]$ with periodic boundary conditions into 30 equal–sized intervals, we obtain a 30×30 stochastic transition matrix S. Solving the eigenvalue problem for S yields:

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	...
1.000	0.950	0.915	0.387	0.227	0.125	...

Evaluating the indicator for $r_{\mathrm{ess}}(P_\tau)$ we get $[\Delta(P_\tau)] = 0.361$—for a further analysis of the indicator within a hierarchical context see below. As for the preceding model systems, we look for a decomposition into three metastable subsets. The eigenfunctions corresponding to the largest eigenvalues are shown in Figure 7 (right). Applying the identification algorithm, we end up with a clustering $\{C_1, C_2, C_3\}$ with $C_1 = \{q \leq -2.1\}$, $C_2 = \{-2.1 < q \leq 1.7\}$ and $C_3 = \{1.7 < q\}$. The statistical weights $\mu(C_k)$ within the positional canonical ensemble $\mu_{\mathcal{Q}}$ and the metastabilities $p(\tau, C_k, C_k)$ are given by the following table:

metastable subset	C_1	C_2	C_3
statistical weight	0.353	0.589	0.058
metastability	0.948	0.956	0.867

The essential statistical behavior is given by the coupling matrix

$$C = \begin{pmatrix} 0.950 & 0.050 & 0 \\ 0.030 & 0.957 & 0.013 \\ 0 & 0.138 & 0.862 \end{pmatrix}.$$

From a theoretical point of view we know that $r_{\text{ess}}(P_\tau) = \Delta(P_\tau) = 0$ holds for periodic systems due to Prop. 6.5. Does the indicator $[\Delta(P_\tau)]$ reproduce this result? The following table shows the indicator based on different decompositions of the state space into m equal–sized intervals:

m	30	50	100	200
$[\Delta(P_\tau)]$	0.361	0.224	0.116	0.070

(68)

The values are in good agreement with the theoretical value $r_{\text{ess}}(P_\tau) = 0$. The previous results are based on a discretization of the propagator according to Example 5.1. This was possible, since the state space is very low–dimensional. In higher dimensions, we will have to use a discretization procedure according to Example 5.2 that is based on a realization of the Markov process $X_n = \{X_{n\tau}\}$ for some fixed time $\tau > 0$. Exploiting a realization of X_n we see that the dominant eigenvalues, the resulting clusterings, their statistical weights and their metastabilities are almost indistinguishable from the results previously obtained, while the indicator $[\Delta(P_\tau)]$ behaves quite differently:

m	30	50	100	200
$[\Delta(P_\tau)]$	0.409	0.389	0.500	0.500

The values for $[\Delta(P_\tau)]$ decrease first and then stay constant for a decomposition into more than 100 subsets. This seems to contradict the theoretical results, but can be understood in the context mentioned at the end of Section 5.1: Since the Smoluchowski process is reversible, we used the original as well as the reversed sampling to discretize the propagator. Analyzing the stochastic transition matrix S for, e.g., $m = 100$ we see that $[\Delta(P)] = S_{2,13} = S_{2,16} = 0.5$. Since the statistical weight of the 2nd subset is $3e-6$, it implies that only one sampling point hit the second subset. Taking into account also the reversed sampling, we see that the second subset was exactly hit two times, which after normalization results in the value $S_{2,13} = S_{2,16} = 0.5$. In this case, $[\Delta(P_\tau)]$ indicates an insufficient sampling of the 2nd subset rather than an upper bound of the essential spectral radius, as was already discussed at the end of Section 5.1. A refinement of the sampling would decrease the value of $[\Delta(P_\tau)]$, as we already know from table (68).

As in the Langevin case, for a fixed inverse temperature β there is a one–parameter family of Smoluchowski equations parameterized by γ (or σ). The corresponding family of infinitesimal generators $\{\mathcal{L}_\gamma\}_{\gamma>0}$ has a very

simple form, since for fixed β and arbitrary $\gamma > 0$

$$\mathcal{L}_\gamma \;=\; \frac{1}{\gamma}\left(\underbrace{\frac{1}{\beta}\Delta_q - \nabla_q V(q)\cdot\nabla_q}_{\mathcal{L}_\beta}\right). \tag{69}$$

Hence, each generator $\mathcal{L}_\gamma$ is simply a multiple of $\mathcal{L}_\beta$ and we get the following relationship for the eigenvalues and -functions:

$$\mathcal{L}_\beta\, v \;=\; \lambda v \quad\Leftrightarrow\quad \mathcal{L}_\gamma\, v \;=\; \frac{\lambda}{\gamma}\, v.$$

We deduce that the eigenfunctions of $\mathcal{L}_\gamma$ are *independent* of γ. For the corresponding propagator $P_t v = \exp(t\mathcal{L}_\gamma)v$ we deduce

$$\mathcal{L}_\gamma\, v \;=\; \frac{\lambda}{\gamma}\, v \quad\Rightarrow\quad P_t v \;=\; \exp\left(\frac{\lambda}{\gamma} t\right)\, v. \tag{70}$$

Figure 9 (top right) shows the eigenvalues of $P_\tau = \exp(\tau\mathcal{L}_\gamma)$ for different values of γ. From (70) we would expect a logarithmic decay with $1/\gamma$, which can be observed in a semi–logarithmic plot (not shown). Since the eigenvalues decay with decreasing γ, we expect that the metastabilities of the resulting decompositions also decay with decreasing γ due to the upper bound in Thm. 3.1. This phenomenon can be observed in Figure 9 (bottom right).

6.5 Comparison of Model Systems

In broad terms the Hamiltonian system with randomized momenta, the Langevin equation and the Smoluchowski equation behave quite similar (for the chosen model parameters). The numerical investigations show comparable results for the clusterings into metastable subsets, their statistical weights and the corresponding coupling matrices reflecting the essential statistical behavior. Qualitative different behavior is displayed by the fourth eigenfunctions (bottom line of Figure 7). But since for each model dynamics the fourth eigenvalue is well separated from the three dominant ones by a spectral gap, the influence on the dynamical behavior is expected to be less important.

The Hamiltonian system with randomized momenta is uniquely determined by specifying a potential function and some inverse temperature, hence specifying the canonical ensemble. In contrast to that, for fixed β, the Langevin and the Smoluchowski equation still depend on the friction constant γ, which can be related to the viscosity of the surrounding. As a consequence, the results the dynamical behavior depends on γ. Comparing the Langevin and the Smoluchowski model, we observe increasing

agreement of eigenvalues and metastabilities for increasing values of γ (see Figure 9). This is what we should expect, since the Smoluchowski equation was derived as a high–friction approximation of the Langevin equation (Section 2.4). This can also be understood from a dominant eigenfunction point of view: Figure 10 (right) shows the second dominant eigenfunctions for different values of γ; the not depicted third eigenfunctions behaves similarly. For small friction the eigenfunctions still display features of the deterministic Hamiltonian system, while for larger values of γ the eigenfunctions convert more and more to a special product form, namely the product of a function acting on positions times a *constant* function in the momenta. Since the "loss of information" by projecting such a function to the position space is negligible, this might explain the good approximation quality of the Langevin by the Smoluchowski equation in the case of high friction. This is also reflected by the observation that the spectrum of the Langevin propagator becomes more and more "real–valued" (see Figure 8), as this is the case for the Smoluchowski equation. However, for small γ, the Langevin and the Smoluchowski equation exhibit different behavior. While in the Langevin case some eigenvalues accumulate in $\lambda = 1$ for $\gamma \ll 1$, this is not the case for the Smoluchowski equation, where all eigenvalues but $\lambda = 1$ tend to zero due to relation (70).

So far, we have presented an analysis of metastability for some fixed observation time span $\tau > 0$. How do the results depend on τ? For the Hamiltonian system with randomized momenta, first investigations are documented in [68]. Exemplified for a small molecule, the dependence of the eigenvalues on τ is analyzed[14]. It looks almost exponential. For the Langevin and Smoluchowski equation the exponential dependence of the eigenvalues on τ is theoretically deducible on the basis of the semigroup property $P_\tau = \exp(\tau\mathcal{L})$. We further conclude that in this case the corresponding eigenfunctions and hence the thereon based metastable subsets are *independent* of any observation time span.

[14] Actually, Schütte analyzed in [68] the dependence of the eigenvalues on the inverse temperature β, which in combination with Section 3.7.2 of [68], where an (inverse) temperature scaling is related to a rescaling of the observation time span τ, gives the stated result.