2 The Model Systems

We introduce four popular models for molecular dynamics: the deterministic Hamiltonian system, the Hamiltonian system with randomized momenta, the Langevin and the Smoluchowski equation. While the deterministic Hamiltonian system and the Langevin equation are classical models on the phase space Γ (positions and momenta), the Hamiltonian system with randomized momenta and the Smoluchowski equation are reduced models acting only on the position space Ω. For each model system we assume that Ω belongs to one of the two fundamentally different cases:

1. **Bounded system**: The position space Ω is unbounded, typically $\Omega = \mathbf{R}^{3N}$, and the potential energy function V is smooth, bounded from below, and satisfies $V \to \infty$ for $|q| \to \infty$. Such systems are called bounded, since the energy surfaces $\{(q,p) \in \Gamma : H(q,p) = E\}$ are bounded subsets of Γ for every energy E.

2. **Periodic systems**: The position space Ω is some $3N$–dimensional torus and the potential energy function V is continuous on Ω and thus bounded. There is an intensive discussion concerning the question of whether V can also be assumed to be smooth as we will do herein, see [68, Sec. 2] for details.

Both cases are typical for molecular dynamics applications. Periodic systems in particular include the assumption of periodic boundaries, which is by far the most popular modeling assumption for biomolecular systems.

2.1 Deterministic Hamiltonian System

The **deterministic Hamiltonian system**

$$\dot{q} \;=\; p, \qquad \dot{p} \;=\; -\nabla_q V(q). \tag{13}$$

defined on the state space $\mathbf{X} = \Gamma$ models an energetically closed system, whose total energy is given by the Hamiltonian H as defined in (1); its energy is conserved under the dynamics. The Markov process $X_t = \{X_t\}_{t \in \mathbf{R}_+}$ defined by the deterministic Hamiltonian system coincides with the flow Φ^τ associated with (13); hence $X_t = \Phi^t x_0$ for the initial distribution $X_0 \sim x_0$. This allows us to denote the stochastic transition function as

$$p(t,x,C) \;=\; \mathbf{1}_C\big(\Phi^t x\big) \;=\; \delta_{\Phi^t x}(C) \tag{14}$$

for every $t \in \mathbf{R}_+$ and $C \in \mathcal{A}$. It is well known that canonical ensemble μ_{can} is invariant w.r.t. X_t. The evolution of densities $v = v(x,t)$ w.r.t. μ_{can} is governed by the **Liouville equation**

$$\partial_t v \;=\; \Bigg(\underbrace{-p \cdot \nabla_q \;+\; \nabla_q V \cdot \nabla_p}_{i\mathcal{L}} \Bigg)\, v, \tag{15}$$

where $\mathcal{L}$ denotes the Liouville operator defined on some suitable subspace of $L^1(\mu_{\text{can}})$. Since the solution of (15) satisfies $v(x, t+s) = v(\Phi^{-t}x, s)$ for every $t, s \in \mathbf{R}_+$, the semigroup of propagators $P_t : L^1(\mu_{\text{can}}) \to L^1(\mu_{\text{can}})$ with $t \in \mathbf{R}_+$ is defined by

$$P_t v(x) \;=\; \exp(it\mathcal{L})v(x) \;=\; v\left(\Phi^{-t}x\right). \tag{16}$$

This is exactly the definition of the Frobenius–Perron operator corresponding to the Hamiltonian flow Φ^t [46]. Since $\mathcal{L}$ is self–adjoint w.r.t. the scalar product in $L^2(\mu)$ the operator P_t is unitary in $L^2(\mu)$. The semigroup of backward transfer operators $T_t : L^\infty(\mu) \to L^\infty(\mu)$ with $t \in \mathbf{R}_+$ is given by

$$T_t u(x) \;=\; u\left(\Phi^t x\right), \tag{17}$$

which is identical with the Koopman operator corresponding to Φ^t [46].

2.2 Hamiltonian System with Randomized Momenta

Aiming at a conformational analysis of biomolecular systems, Schütte introduced in [68] some kind of stochastic Hamiltonian system. It is a reduced dynamics defined solely on the position space and derived from the deterministic Hamiltonian system by "randomizing the momenta" and integrating for some fixed observation time span τ.

Let us briefly sketch the derivation. Fix some observation time span $\tau > 0$ and denote by $p_\Gamma(\tau, x, A)$ the stochastic transition function corresponding to the full deterministic Hamiltonian system on Γ (for a comment on the time τ see remark below). In view of the relation between metastable sub–ensembles and conformations, we are interested in a simplified model describing the dynamics between "cylindric" subsets $B \times \mathbf{R}^d$ and $C \times \mathbf{R}^d$. Inserting these special subsets into definition (14) yields

$$\begin{aligned}
&p_\Gamma(\tau, B \times \mathbf{R}^d, C \times \mathbf{R}^d) \\
&\quad= \frac{1}{\int_{B\times\mathbf{R}^d} \mu_{\text{can}}(\mathrm{d}x)} \int_{B\times\mathbf{R}^d} \mathbf{1}_{C\times\mathbf{R}^d}(\Phi^\tau(x))\mu_{\text{can}}(\mathrm{d}x) \\
&\quad= \frac{1}{\int_B \mu_{\mathcal{Q}}(\mathrm{d}q)} \int_B \underbrace{\int_{\mathbf{R}^d} \mathbf{1}_C(\Pi_q \Phi^\tau(q,p))\mu_{\mathcal{P}}(\mathrm{d}p)}_{p^\tau_\Omega(1,q,C)}\; \mu_{\mathcal{Q}}(\mathrm{d}q). \qquad (18) \\
&\quad= p^\tau_\Omega(1, B, C),
\end{aligned}$$

where $\Pi_q : \Gamma \to \Omega$ denotes the projection onto the position space. Equation (18) defines a *one–step* stochastic transition function, whose n–step version is determined via the Chapman–Kolmogorov equation (5). The associated discrete time Markov process $Q_n = \{Q_n\}_{n\in\mathbf{Z}_+}$, defined on the state

space $\mathbf{X} = \Omega$, solves the **Hamiltonian system with randomized momenta** [68]

$$Q_{n+1} \;=\; \Pi_q \Phi^\tau(Q_n, P_n); \qquad n \in \mathbf{Z}_+ \tag{19}$$

where P_n is chosen randomly from the canonical distribution of momenta $\mathcal{P}$, as defined in (3). As shown in [68] the positional canonical ensemble $\mu_\mathcal{Q}$ is invariant w.r.t. the Markov process Q_n. The semigroup of propagators $P_n : L^1(\mu_\mathcal{Q}) \to L^1(\mu_\mathcal{Q})$ for $n \in \mathbf{Z}_+$ is given by $P_n = (P_1)^n$ with

$$P_1 v(q) \;=\; \int_{\mathbf{R}^d} v(\Pi_q \Phi^{-\tau}(q,p)) \mathcal{P}(p) \mathrm{d}p. \tag{20}$$

Exploiting that $\mu_\mathcal{Q}$ is invariant and Φ^τ is reversible and symplectic, it is shown in [68] that P_1, and thus the semigroup, is self–adjoint in $L^2(\mu_\mathcal{Q})$.

Remark. For arbitrary, but fixed $\tau > 0$ we have defined in (19) the *one–step* transition function $p^\tau(1, q, D)$. Changing the observation time to $\sigma > 0$ results in a new one–step transition function $p^\sigma(1, q, D)$. In general we will have $p^{2\tau}(1, q, D) \neq p^\tau(2, q, D)$ and, consequently, $P_1^{2\tau} \neq P_2^\tau$, where the superscripts indicate the corresponding observation time spans (for an example see [68, Sec. 3.7.1]). In terms of the Hamiltonian system with randomized momenta, this is not surprising, since $P_1^{2\tau}$ includes only one choice of momenta according to $\mathcal{P}$, while P_2^τ does include two.

2.3 Langevin Equation

The most popular model for an open system stochastically interacting with its environment is the **Langevin equation**[3] [61]

$$\dot{q} \,=\, p, \qquad \dot{p} \,=\, -\nabla_q V(q) \,-\, \gamma\, p \,+\, \sigma\, \dot{W} \tag{21}$$

corresponding to some friction constant $\gamma > 0$ and external force $F_{\mathrm{ext}} = \sigma \dot{W}$ given by a standard $3N$-dimensional Brownian motion W. Eq. (21) defines a continuous time Markov process $X_t = \{(Q_t, P_t)\}_{t \in \mathbf{T}}$ on the state space $\mathbf{X} = \Gamma$.

In the Langevin model, the effects of solvent molecules not explicitly present in the system being simulated are approximated in terms of a frictional drag on the solute as well as random collisions associated with the thermal motion of the solvent molecules. The Hamiltonian H describes the internal energy of the systems, which is not conserved due to energy

[3] In our context, the notion $\dot{W}$ is a convenient form of the more common $\mathrm{d}W$. Hence, the Langevin equation (21) should be understood as $\mathrm{d}q = p\mathrm{d}t$ and $\mathrm{d}p = -\nabla_q V(q)\mathrm{d}t - \gamma p \mathrm{d}t + \sigma \mathrm{d}W$, which moreover is just the common abbreviation for the corresponding integral notion, see e.g., [57]. For convenience we will henceforth use the "dot" without further comments.

transfer with the surrounding, the so–called heat bath. Yet, the interplay between stochastic excitation and damping equilibrates the internal energy to $\beta = 2\gamma/\sigma^2$. As a result, the canonical ensemble $\mu_{\rm can}$ corresponding to the inverse temperature β is invariant w.r.t. the Markov process defined by (21). The evolution of densities $v = v(t,x)$ w.r.t. $\mu_{\rm can}$ is governed by the **Fokker–Planck equation**

$$\partial_t v \;=\; \left(\underbrace{\frac{\sigma^2}{2}\Delta_p \,-\, p\cdot\nabla_q \,+\, \nabla_q V\cdot\nabla_p \,-\, \gamma p\cdot\nabla_p}_{\mathcal{L}} \right) v \tag{22}$$

regarded on some suitable subspace of $L^1(\mu_{\rm can})$. Therefore, $\mathcal{L}$ is the infinitesimal generator of the semigroup of propagators $P_t : L^1(\mu_{\rm can}) \to L^1(\mu_{\rm can})$ with $t \in \mathbf{R}_+$ defined by

$$P_t v \;=\; \exp(t\mathcal{L})v \tag{23}$$

for $v \in L^1(\mu_{\rm can})$. In general, P_t is not self–adjoint in $L^2(\mu)$.

Remark. The evolution of some physical density[4] $v_{\rm phys} = v f_{\rm can} \in L^1({\rm d}x)$ with $v \in L^1(\mu_{\rm can})$ is governed by the so–called **forward Kolmogorov equation** $\partial_t v_{\rm phys} = \mathcal{A}_{\rm fw} v_{\rm phys}$ with

$$\mathcal{A}_{\rm fw} \;=\; \frac{\sigma^2}{2}\Delta_p - p\cdot\nabla_q + \nabla_q V\cdot\nabla_p + \gamma p\cdot\nabla_p + \gamma$$

acting on a suitable subspace of $L^1({\rm d}x)$ [38, Chapter 5.1]. It permits to define the semigroup of propagators $P_t^{\rm fw} : L^1({\rm d}x) \to L^1({\rm d}x)$ by

$$P_t^{\rm fw} v_{\rm phys} \;=\; \exp(t\mathcal{A}_{\rm fw}) v_{\rm phys}.$$

As a consequence of the invariance of μ, we obtain the relation

$$P_t^{\rm fw}(v f_{\rm can}) \;=\; (P_t v) f_{\rm can}; \qquad v \in L^1(\mu_{\rm can}) \tag{24}$$

between the two semigroups of propagators. To derive the evolution equation (22) for v, we insert $v_{\rm phys} = v f_{\rm can}$ into the forward Kolmogorov equation and obtain after simple manipulations

$$\partial_t(v f_{\rm can}) \;=\; \mathcal{A}_{\rm fw}(v f_{\rm can}) \;=\; (\mathcal{L}v) f_{\rm can},$$

which is the infinitesimal version of (24). Exploiting time–independence and positivity of $f_{\rm can}$, we finally end up with the Fokker–Planck equation (22).

[4] See remark about our mathematical model at the end of Section 1.3.

In some sense dual to the forward Kolmogorov equation is the **backward Kolmogorov equation** $\partial_t u = \mathcal{A}_{\mathrm{bw}} u$ with

$$\mathcal{A}_{\mathrm{bw}} \;=\; \frac{\sigma^2}{2}\Delta_p + p\cdot\nabla_q - \nabla_q V\cdot\nabla_p - \gamma p\cdot\nabla_p$$

acting on a suitable subspace of $L^1(\mathrm{d}x)$. For bounded and periodic systems, we have $\langle \mathcal{A}_{\mathrm{fw}} v, u\rangle_2 = \langle v, \mathcal{A}_{\mathrm{bw}} u\rangle_2$ on the Hilbert space $L^2(\mathrm{d}x)$ and hence $\mathcal{A}_{\mathrm{fw}}$ and $\mathcal{A}_{\mathrm{bw}}$ are adjoint to each other. Therefore neither $\mathcal{A}_{\mathrm{fw}}$ nor $\mathcal{A}_{\mathrm{bw}}$ is self–adjoint in $L^2(\mathrm{d}x)$. The generator $\mathcal{A}_{\mathrm{bw}}$ permits to define the semigroup of backward transfer operators $T_t^{\mathrm{bw}} : L^1(\mathrm{d}x) \to L^1(\mathrm{d}x)$ by

$$T_t^{\mathrm{bw}} u(x) \;=\; \mathbf{E}_x[u(X_t)] \;=\; \exp(t\mathcal{A}_{\mathrm{bw}})u(x),$$

see, e.g., [38, Chapter 5.1]. We remark that although the formal definition of the two semigroups of backward transfer operators T_t and T_t^{bw} via expectation is the same, they differ in the space of functions regarded to act on.

2.4 Smoluchowski Equation

As a second reduced model system, we introduce the Smoluchowski equation. It is derived from the Langevin equation by considering the high friction limit $\gamma \to \infty$. In contrast to the Langevin equation it defines a reversible Markov process.

Write the Langevin equation (21) in second order form

$$\ddot{q} \;=\; -\nabla_q V(q) - \gamma\dot{q} + \sigma\dot{W}. \tag{25}$$

For the high friction limit, we introduce some smallness parameter $\epsilon > 0$ and transform the friction constant to γ/ϵ; in order to conserve the inverse temperature $\beta = 2\gamma/\sigma^2$ of the surrounding heat bath and hence the canonical ensemble, we simultaneously have to scale the white noise constant to $\sigma/\sqrt{\epsilon}$. This yields

$$\ddot{q} \;=\; -\nabla_q V(q) - \frac{\gamma}{\epsilon}\dot{q} + \frac{\sigma}{\sqrt{\epsilon}}\dot{W}.$$

After rescaling the time according to $t \mapsto \epsilon t$ we finally get

$$\epsilon^2\ddot{q} \;=\; -\nabla_q V(q) - \gamma\dot{q} + \sigma\dot{W}.$$

Since the white noise process $\dot{W}$ is unbounded, we cannot simply assume that the acceleration $\epsilon^2\ddot{q}$ is small for $\epsilon \ll 1$. However, investigations by Nelson [53] show that the solution $q^\epsilon_{\mathrm{Lan}}(t; q_0, p_0)$ of the Langevin equation (26) and the solution $q_{\mathrm{Smol}}(t; q_0)$ of the **Smoluchowski equation**

$$\dot{q} = -\frac{1}{\gamma}\nabla_q V(q) + \frac{\sigma}{\gamma}\dot{W} \tag{26}$$

are close to each other for high friction γ.

Theorem 2.1 *[53, Theorem 10.1] Assume that $\nabla_q V(q)$ is global Lipschitz. Then, for every p_0, with probability one*

$$\lim_{\epsilon \to 0} |q_{\mathrm{Smol}}(t; q_0) - q^{\epsilon}_{\mathrm{Lan}}(t; q_0, p_0)| \quad = \quad 0$$

uniformly for t in compact subintervals of $[0, \infty)$.

Hence, in the high friction case, the Smoluchowski dynamics is a good approximation of the Langevin dynamics. This fact will be analyzed in the following section.

The stochastic differential equation (26) defines a continuous time Markov process $Q_t = \{Q_t\}_{t \in \mathbf{R}_+}$ on the state space $\mathbf{X} = \Omega$ with transition kernel $p = p(t, q, C)$ and invariant probability measure $\mu_{\mathcal{Q}}$ [61]. The evolution of densities $v = v(t, q)$ w.r.t. $\mu_{\mathcal{Q}}$ is governed by the **Fokker–Planck equation** (see Remark below for derivation)

$$\partial_t v \quad = \quad \left(\underbrace{\frac{\sigma^2}{2\gamma^2} \Delta_q - \frac{1}{\gamma} \nabla_q V(q) \cdot \nabla_q}_{\mathcal{L}} \right) v \tag{27}$$

regarded on some suitable subspace of $L^1(\mu_{\mathcal{Q}})$. Therefore, $\mathcal{L}$ is the infinitesimal generator of the semigroup of propagators $P_t : L^1(\mu_{\mathcal{Q}}) \to L^1(\mu_{\mathcal{Q}})$ with $t \in \mathbf{R}_+$ defined by

$$P_t v \quad = \quad \exp(t\mathcal{L}) v \tag{28}$$

for $v \in L^1(\mu)$.

Proposition 2.2 *The semigroup of propagators $P_t : L^2(\mu_{\mathcal{Q}}) \to L^2(\mu_{\mathcal{Q}})$ is self–adjoint w.r.t. the scalar product $\langle \cdot, \cdot \rangle_{\mu_{\mathcal{Q}}}$ in $L^2(\mu_{\mathcal{Q}})$.*

Proof: We prove that $\mathcal{L}$ is self–adjoint in $L^2(\mu_{\mathcal{Q}})$. The statement of the theorem then follows from [8, Thm. 4.6]. Consider $v, u \in L^2(\mu_{\mathcal{Q}})$ in the domain of $\mathcal{L}$. Then $\mathcal{Q}^{\frac{1}{2}} v, \mathcal{Q}^{\frac{1}{2}} u \in L^2(\mathrm{d}q)$ and

$$\langle \mathcal{L} v, u \rangle_{\mu_{\mathcal{Q}}} \quad = \quad \langle (\mathcal{Q}^{\frac{1}{2}} \mathcal{L} \mathcal{Q}^{-\frac{1}{2}})(\mathcal{Q}^{\frac{1}{2}} v), \mathcal{Q}^{\frac{1}{2}} u \rangle_2 .$$

A simple calculation proves that $\mathcal{Q}^{\frac{1}{2}} \mathcal{L} \mathcal{Q}^{-\frac{1}{2}} = \mathcal{L}_s$ for the generator $\mathcal{L}_s$ defined in (30) below. Since $\mathcal{L}_s$ is a so–called Schrödinger operator, it is self–adjoint in $L^2(\mathrm{d}q)$ and we obtain

$$\langle \mathcal{L} v, u \rangle_{\mu_{\mathcal{Q}}} \quad = \quad \langle \mathcal{Q}^{\frac{1}{2}} v, (\mathcal{Q}^{\frac{1}{2}} \mathcal{L} \mathcal{Q}^{-\frac{1}{2}})(\mathcal{Q}^{\frac{1}{2}} u) \rangle_{\mu_{\mathcal{Q}}} \; = \; \langle v, \mathcal{L} u \rangle_{\mu_{\mathcal{Q}}},$$

hence $\mathcal{L}$ is self–adjoint in $L^2(\mu_{\mathcal{Q}})$. □

There is a strong relation of our approach to the theory of large deviations and first exit times (see, e.g., Freidlin and Wentzell [23]); a brief outline is given in [72]; further investigations are part of a current diploma thesis [49].

Remark.[5] The evolution of some physical density $v_{\rm phys} = v\mathcal{Q} \in L^1(\mathrm{d}q)$ with $v \in L^1(\mu_\mathcal{Q})$ is governed by the so–called **forward Kolmogorov equation** $\partial_t v_{\rm phys} = \mathcal{A}_{\rm fw} v_{\rm phys}$ with

$$\mathcal{A}_{\rm fw} \;=\; \frac{\sigma^2}{2\gamma^2}\Delta_q + \frac{1}{\gamma}\nabla_q V(q)\cdot\nabla_q + \frac{1}{\gamma}\Delta_q V(q)$$

acting on a suitable subspace of $L^1(\mathrm{d}q)$ [38, Chapter 5.1]. It permits to define the semigroup of propagators $P_t^{\rm fw} : L^1(\mathrm{d}q) \to L^1(\mathrm{d}q)$ by

$$P_t^{\rm fw} v_{\rm phys} \;=\; \exp(t\mathcal{A}_{\rm fw}) v_{\rm phys}.$$

As a consequence of the invariance of μ, we obtain the relation

$$P_t^{\rm fw}(v\mathcal{Q}) \;=\; (P_t v)\mathcal{Q}; \qquad v \in L^1(\mu_\mathcal{Q}) \tag{29}$$

between the two semigroups of propagators. To derive the evolution equation (27) for v, we insert $v_{\rm phys} = v\mathcal{Q}$ into the forward Kolmogorov equation and obtain after simple manipulations

$$\partial_t(v\mathcal{Q}) \;=\; \mathcal{A}_{\rm fw}(v\mathcal{Q}) \;=\; (\mathcal{L}v)\mathcal{Q},$$

which is the infinitesimal version of (29). Exploiting time–independence and positivity of $\mathcal{Q}$, we finally end up with the Fokker–Planck equation (27). In some sense dual to the forward Kolmogorov equation is the **backward Kolmogorov equation** $\partial_t u = \mathcal{A}_{\rm bw} u$ with

$$\mathcal{A}_{\rm bw} \;=\; \frac{\sigma^2}{2\gamma^2}\Delta_q - \frac{1}{\gamma}\nabla_q V(q)\cdot\nabla_q$$

acting on a suitable subspace of $L^1(\mathrm{d}q)$. Although $\mathcal{L}$ and $\mathcal{A}_{\rm bw}$ are formally the same, they are regarded to act on different function spaces. For bounded and periodic systems, we have $\langle \mathcal{A}_{\rm fw} v, u\rangle_2 = \langle v, \mathcal{A}_{\rm bw} u\rangle_2$ on the Hilbert space $L^2(\mathrm{d}q)$ and hence $\mathcal{A}_{\rm fw}$ and $\mathcal{A}_{\rm bw}$ are adjoint to each other. However, neither $\mathcal{A}_{\rm fw}$ nor $\mathcal{A}_{\rm bw}$ are self–adjoint in $L^2(\mathrm{d}q)$, while $\mathcal{L}$ is self–adjoint in $L^2(\mu_\mathcal{Q})$. The generator $\mathcal{A}_{\rm bw}$ permits to define the semigroup of backward transfer operators $T_t^{\rm bw} : L^1(\mathrm{d}q) \to L^1(\mathrm{d}q)$ according to

$$T_t^{\rm bw} u(q) \;=\; \mathbf{E}_q[u(X_t)] \;=\; \exp(t\mathcal{A}_{\rm bw}) u(q),$$

[5] We are aware of the fact that large parts of this remark are analogous to those made for the Langevin equation and could hence be abbreviated. Yet, for sake of clarity, we prefer to state everything explicitly, since there are also important differences.

see, e.g., [38, Chapter 5.1]. We remark that although the formal definition of the two semigroups of backward transfer operators T_t and T_t^{bw} via expectation is the same, they differ in the space of functions regarded to act on.

Sometimes it is advantageous to consider yet another generator

$$\mathcal{L}_s \quad = \quad \frac{\sigma^2}{2\gamma^2}\Delta_q - \left(\underbrace{\frac{1}{2\sigma^2}(\nabla V(q))^2 - \frac{1}{2\gamma}\Delta V(q)}_{U(q)} \right) \tag{30}$$

regarded to act on a suitable subspace of $L^2(\mathrm{d}q)$. It is defined in terms of the potential function $U : \Omega \to \mathbf{R}$ and allows to apply the powerful theory of Schrödinger operators (see, e.g., Reed and Simon [59]); for a brief outline see [71]. The **Schrödinger operator** $\mathcal{L}_s$ is related to the generators $\mathcal{L}$ by the identity $\mathcal{Q}^{\frac{1}{2}}\mathcal{L}v = \mathcal{L}_s(\mathcal{Q}^{\frac{1}{2}}v)$ for every $v \in L^1(\mu_{\mathcal{Q}})$.