2. Modelling molecular motion

2.1. Lagrangian and Hamiltonian mechanics

In classical molecular dynamics, a molecule is mostly modelled as a natural mechanical system [87, 88]. The two basic formulations are those of Lagrange and Hamilton: From a systematic point of view, Lagrangian mechanics seems more fundamental as it is based on a variational principle that directly leads to a coordinate-invariant formulation of mechanics. Intriguingly the *least action principle* was discovered by Hamilton itself decades after Lagrange stated the Euler-Lagrange equations, and which is thus also termed Hamilton's principle [89, 90]. However the equivalent Hamiltonian formulation which should be distinguished from the variational principle, is better suited to establish statistical thermodynamics from mechanical principles or to point out the fundamental relationship with quantum mechanics. The related ensemble concepts are crucial ingredients of molecular models. We essentially follow the outline in the textbook [66]; see also [81] for some historical remarks.

Atomic interaction potential The key ingredient to all classical molecular models is the atomic interaction potential that contains all physically relevant interactions, where the word *physically* is to be understood within the scope of classical molecular dynamics. Here we consider only potentials that do not involve any velocity interactions and that are at least twice continuously differentiable.

Typically the molecular potential is modelled as a sum of contributions of different physical origin. For the sake of convenience we endeavour the multi-body notation with the bold face symbol $\mathbf{q}_i \in \mathbf{R}^3$, $i = 1, \ldots, N$ for the position of the i-th atom. Roughly speaking, the molecular interaction can be divided into two parts: *local* interactions as induced by the bond structure of the molecule, and *long-range* interactions like electrostatic interactions, for instance. One rather simple but widely-used example is the Ryckaert-Bellemans potential for modelling alkane chains [91]. The local interactions are

- harmonic stretching of the covalent bond between the atoms i and $i + 1$ as described by $V_{\mathrm{bd}}(\mathbf{q}_i, \mathbf{q}_{i+1}) \propto (r_i - r_{\mathrm{eq}})^2$ with $r_i = \|\mathbf{q}_{i+1} - \mathbf{q}_i\|$,
- harmonic vibrations of the bond angle formed by the covalent bonds between three successive atoms which are modelled by $V_{\mathrm{ba}}(\mathbf{q}_{i-1}, \mathbf{q}_i, \mathbf{q}_{i+1}) \propto (\psi_i - \psi_{\mathrm{eq}})^2$ with the bond angle $\psi_i = \sphericalangle(\mathbf{q}_{i-1} - \mathbf{q}_i, \mathbf{q}_{i+1} - \mathbf{q}_i)$,
- motion of the torsion angle ω_i between two planes each of which is spanned by three atoms; this potential has a multi-well structure (see Figure 1), and it depends upon positions of four successive atoms $V_{\mathrm{ts}} = V_{\mathrm{ts}}(\mathbf{q}_{i-1}, \mathbf{q}_i, \mathbf{q}_{i+1}, \mathbf{q}_{i+2})$.

Local interactions that involve more that four atoms are not captured by this model. The corresponding non-bonded interaction are

- electrostatic interaction from the charges of the atoms j and k, that are described by a Coulomb potential $V_{\mathrm{C}}(\mathbf{q}_j, \mathbf{q}_k) \propto 1/d_{jk}$ with $d_{jk} = \|\mathbf{q}_j - \mathbf{q}_k\|$,
- van der Waals interactions between polarizable atoms that are modelled by a Lennard-Jones potential $V_{\mathrm{LJ}}(\mathbf{q}_j, \mathbf{q}_k) \propto 1/d_{jk}^{12} - 1/d_{jk}^6$ that contains both short-range and long-range interactions.
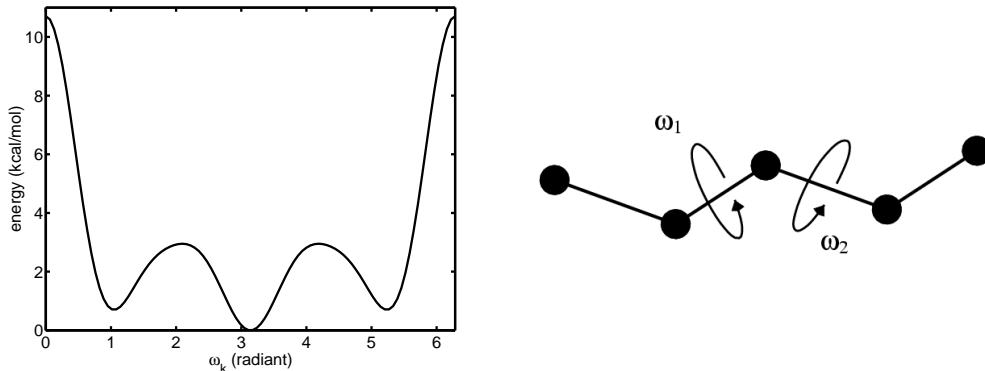
Figure 1. Ryckaert-Bellemans torsion angle potential as a function of the torsion angle ω_k; see [91] for the details. The right panel shows the only two proper torsion angles (dihedral angles) of the *united atoms* pentane molecule.

This results in the following molecular potential

$$V(\mathbf{q}) = \sum_i V_{\text{bd}}(\mathbf{q}_i, \mathbf{q}_{i+1}) + \sum_j V_{\text{ba}}(\mathbf{q}_{j-1}, \mathbf{q}_j, \mathbf{q}_{j+1})$$
$$+ \sum_k V_{\text{ts}}(\mathbf{q}_{k-1}, \mathbf{q}_k, \mathbf{q}_{k+1}, \mathbf{q}_{k+2})$$
$$+ \sum_{k,l} V_{\text{C}}(\mathbf{q}_k, \mathbf{q}_l) + \sum_{k,l} V_{\text{LJ}}(\mathbf{q}_k, \mathbf{q}_l),$$

where $\mathbf{q} = (\mathbf{q}_1, \ldots, \mathbf{q}_N)^T$ and the sums run over all the atoms that contribute to the respective interaction. Other possibly relevant interactions that are not directly contained in the potential are accounted for by adjusting the parameters in the potential in an appropriate way. In this sense the molecular potential is semi-empirical, and in fact using different potentials or *force fields* may produce vastly different results. See [92, 44] for the general force field methodology and [93] for a recent comparison of different force fields for the conformation dynamics of trialanine in water.

Lagrangian mechanics Consider a molecule with configuration space $Q \subseteq \mathbf{R}^n$ and molecular configurations $q = (q^1, \ldots, q^n)^T$. Realistic molecular systems typically involve a large number N of atoms, therefore the spatial dimension $n = 3N$ is large. We denote by $\dot{q} = dq/dt$ the corresponding velocities, and by $TQ \cong Q \times \mathbf{R}^n$ the tangential bundle over the configuration space Q.

Let further $V : Q \to \mathbf{R}$ be a smooth molecular interaction potential. Throughout this thesis we shall assume that any of the two conditions is met: Either the system is *bounded* in the sense that $V \to \infty$ as $\|q\| \to \infty$, i.e., the configuration space is unbounded, but the potential is such that a particle cannot escape to infinity. Or the system is *periodic* in the sense that $Q \cong \mathbf{T}^n$. That is, Q is isomorphic to a flat n-dimensional torus. The latter is typically assumed for periodic boundary conditions or to bound the potential energy. In any case the Lagrange function $L : TQ \to \mathbf{R}$ is of the form *kinetic minus potential energy*,

$$L(q, \dot{q}) = \frac{1}{2} \langle M\dot{q}, \dot{q} \rangle - V(q), \qquad (2.1)$$

11

where $M \in \mathbf{R}^{n \times n}$ is the diagonal, and positive-definite mass matrix, and $\langle \cdot, \cdot \rangle$ is the standard inner product between (tangent) vectors in $\mathbf{R}^n$. Consider a curve $q(t)$ in Q with $t \in [a, b]$ and fixed endpoints $q(a) = q_a$, $q(b) = q_b$. Hamilton's principle states that if a curve $q(t)$ minimizes the action integral

$$\int_a^b L(q(t), \dot{q}(t)) \, dt \,, \tag{2.2}$$

then it is a solution of the corresponding Euler-Lagrange equations

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = 0 \,, \quad i = 1, \ldots, n \,. \tag{2.3}$$

For a multi-body system in Cartesian space, the Euler-Lagrange equations reduce to the familiar Newton equations: Let again $\mathbf{q}_i \in \mathbf{R}^3$ denote the position vector of the i-th atom in Cartesian space $Q = \mathbf{R}^{3N}$, and let m_i be the corresponding atomic mass, such that $M = \mathrm{diag}(m_1, m_1, m_1, \ldots, m_N, m_N, m_N)$. Then (2.3) becomes

$$m_i \ddot{\mathbf{q}}_i = -\frac{\partial V}{\partial \mathbf{q}_i} \,, \quad i = 1, \ldots, N \,.$$

A convenient property of the Euler-Lagrange equations (2.3) is that they can be easily put into a form which is invariant under coordinate transforms (covariant); clearly Newton's equations above are not covariant. The covariant formulation is particularly important, e.g., if a molecule is rigid or partially rigid [36].

Assume that Q is a n-dimensional Riemannian or pseudo-Riemannian manifold which is endowed with a coordinate-dependent metric $g(q) : T_q Q \times T_q Q \to \mathbf{R}$. Here $T_q Q$ is the tangent plane attached to $q \in Q$ which we identify with $\mathbf{R}^n$ as usual. Alternatively we will denote the metric by $\langle \cdot, \cdot \rangle_g = \langle g(q)\cdot, \cdot \rangle$ and write $g^\flat$ (or just g) for the associated linear map: the *metric tensor* with elements g_{ij}. The elements of the respective inverse map $g^\sharp$ are labeled g^{ij}.

The proof of the basic statement below is standard and can be found in various textbooks on geometric mechanics [66]. We give it for the sake of illustration.

Lemma 2.1. *Let* $L : TQ \to \mathbf{R}$ *be a Lagrange function. Then in coordinates* $(q, \dot{q}) \in TQ$ *the Euler-Lagrange equations (2.3) have the covariant form*

$$\ddot{q}^i + \Gamma^i_{jk} \dot{q}^j \dot{q}^k + g^{ij} \frac{\partial V}{\partial q^j} = 0 \,, \tag{2.4}$$

where

$$\Gamma^i_{jk} = \frac{1}{2} g^{il} \left(\frac{\partial g_{jl}}{\partial q^k} + \frac{\partial g_{kl}}{\partial q^j} - \frac{\partial g_{jk}}{\partial q^l} \right) \tag{2.5}$$

are the symmetric Christoffel symbols on TQ, and we have set in force the summation convention which means that we sum over double upper and lower indices.

Proof. Consider the Lagrange function on TQ

$$L(q, \dot{q}) = \frac{1}{2} g_{ij} \dot{q}^i \dot{q}^j - V(q) \,,$$

which gives rise to the standard Euler-Lagrange equations (2.3), that now become

$$\frac{d}{dt} \left(g_{ij} \dot{q}^j \right) - \frac{1}{2} \frac{\partial g_{jk}}{\partial q^i} \dot{q}^j \dot{q}^k + \frac{\partial V}{\partial q^i} = 0 \,.$$

12

Hence

$$g_{ij}\ddot{q}^j + \left(\frac{\partial g_{ij}}{\partial q^k} - \frac{1}{2}\frac{\partial g_{jk}}{\partial q^i}\right)\dot{q}^j\dot{q}^k + \frac{\partial V}{\partial q^i} = 0\,,$$

which upon multiplying by the inverse metric tensor $g^\sharp$ and rearranging indices gives

$$\ddot{q}^i + \hat{\Gamma}^i_{jk}\dot{q}^j\dot{q}^k + g^{il}\frac{\partial V}{\partial q^l} = 0$$

with the non-symmetric Christoffel symbols

$$\hat{\Gamma}^i_{jk} = g^{il}\left(\frac{\partial g_{jl}}{\partial q^k} - \frac{1}{2}\frac{\partial g_{jk}}{\partial q^l}\right)\,.$$

Taking into account the symmetry of the summation indices j and k in the Euler-Lagrange equations above, the quadratic expression $\hat{\Gamma}^i_{jk}\dot{q}^j\dot{q}^k$ can be rewritten as

$$g^{il}\left(\frac{\partial g_{jl}}{\partial q^k} - \frac{1}{2}\frac{\partial g_{jk}}{\partial q^l}\right)\dot{q}^j\dot{q}^k = \frac{1}{2}g^{il}\left(\frac{\partial g_{jl}}{\partial q^k} + \frac{\partial g_{kl}}{\partial q^j} - \frac{\partial g_{jk}}{\partial q^l}\right)\dot{q}^j\dot{q}^k$$

where the right hand side of the equation contains the symmetric Christoffel symbols Γ^i_{jk} just defined. From this the assertion follows. $\quad\square$

Remark 2.2. *Sometimes it is required to write the Euler-Lagrange equations in* coordinate-free *form. This is easily done by means of the lemma above using the language of covariant derivatives which is common in differential geometry [94, 95]. If $c(t)$ is a curve, and X is a vector field, the covariant derivative $\nabla_{\dot{c}(t)}X$ is declared as the directional derivative of X along $c(t)$ which reads*

$$\left(\nabla_{\dot{c}(t)}X\right)^i = \frac{d}{dt}X^i(c(t)) + \Gamma^i_{jk}(c(t))X^j(c(t))\dot{c}^k(t)\,.$$

See [66] for the details. If $c(t)$ is an integral curve of X, and X is the Lagrangian vector field, as defined by the Euler-Lagrange equations (2.4), we obtain the following coordinate-free representation

$$\nabla_{\dot{c}(t)}\dot{c}(t) + \mathrm{grad}\,V(c(t)) = 0\,,$$

of the Euler-Lagrange equations, where the gradient of V is defined in coordinates as $(\mathrm{grad}\,V)^i = g^{ij}\partial V/\partial q^j$. In fact, if we multiply the last equation by the metric tensor g (generalized mass matrix) we obtain Newton's second axiom, $F = ma$. For a modern survey of Riemannian geometry and related applications see [96].

Hamiltonian mechanics Given a molecular Langrangian (2.1) we can proceed to the Hamiltonian formalism by introducing the conjugate momentum variable

$$p_i = \frac{\partial L}{\partial \dot{q}^i}\,, \quad i = 1,\ldots,n\,. \tag{2.6}$$

We call the coordinate pair (q,p) a set of *conjugate variables*. We assume that the transformation $(q,\dot{q}) \mapsto (q,p)$ is invertible which requires the partial Hessian matrix $(\partial^2 L/\partial \dot{q}^i\partial \dot{q}^j)$ to be non-singular; so far this is the mass matrix M or the metric tensor g, respectively. We introduce the Hamiltonian as the Legendre transform of L

$$H(q,p) = \dot{q}^i p_i - L(q,\dot{q})\,,$$

13

where we have to substitute the velocities by the respective momentum variables. Keep in mind that the sum is taken over the double upper and lower index i. Once the Hamiltonian is defined we directly obtain Hamilton's equations of motion

$$\dot{q}^i = \frac{\partial H}{\partial p_i}$$

$$\dot{p}_i = -\frac{\partial H}{\partial \dot{q}^i}\,, \quad i = 1, \dots, n\,. \tag{2.7}$$

It is a straight consequence of the above definition that

$$\frac{\partial H}{\partial q^i} = -\frac{\partial L}{\partial q^i} \tag{2.8}$$

which together with (2.6) shows the equivalence of Hamilton's equations and the Euler-Lagrange equations. Furthermore (2.6) assigns the phase space T^*Q with coordinates (q, p) to the configuration space Q with coordinates q; the bundle $T^*Q \cong Q \times \mathbf{R}^n$ is called the cotangent bundle over Q. The mapping from the tangent space TQ to the cotangent space T^*Q is provided by the linear map $g^\flat$,

$$g^\flat : T_q Q \to T_q^* Q\,, \quad p_i = g_{ij}\dot{q}^j\,, \quad i = 1, \dots, n$$

which is a slightly more sophisticated way to expresses the well-known relationship $p = M\dot{q}$ between Cartesian velocities and momenta. By the regularity assumption on the Lagrangian the matrix of $g^\flat$ is invertible, so $g^\sharp$ exists and takes T^*Q back to TQ. On the cotangent space the Hamiltonian has the generic form

$$H(q, p) = \frac{1}{2}g^{ij}p_i p_j + V(q)$$

and generates Hamilton's equations (2.7). A coordinate-free representation of Hamilton's equations can be obtained using Poisson brackets; see [81] for details.

Hamiltonian flows have some important conservation and symmetry properties: Let $\Psi_t : T^*Q \to T^*Q$ denote the one-parameter group of diffeomorphisms that is generated by the Hamiltonian. That is, $(q(t), p(t)) = \Psi_t(q, p)$ is the solution of the initial value problem (2.7) with $(q(0), p(0)) = (q, p)$.

First of all, the flow conserves the total energy, $H = H \circ \Psi_t$. Secondly, it is reversible in the sense that $\Psi_t(q, p) = \Psi_{-t}(q, -p)$. Last but not least, it conserves phase space volume, i.e., $\det \mathbf{D}\Psi_t = 1$. The last feature follows directly from the fact that the vector field (2.7) is divergence-free; an even stronger attribute is that Hamiltonian flows are *symplectic* which will be explained in the following.

Definition 2.3. *Let M be an even-dimensional smooth manifold. A symplectic structure on M is a closed non-degenerate differential 2-form $\Omega : T_m M \times T_m M \to \mathbf{R}$, i.e., the following conditions hold true*

$$d\Omega = 0 \quad and \quad \forall \xi \neq 0\, \exists \eta : \, \Omega(\xi, \eta) \neq 0\,.$$

The pair (M, Ω) is called a symplectic manifold.

As a consequence of Darboux's Theorem [97] there is a natural symplectic form on the cotangent space T^*Q that is given by

$$\Omega = dq^i \wedge dp_i\,. \tag{2.9}$$

where the skew-symmetric wedge product of two differentials is defined as

$$\Omega(\xi, \eta) = dq^i(\xi)dp_i(\eta) - dq^i(\eta)dp_i(\xi)\,.$$

14

with $dq^i(\zeta) = \zeta_1^i$, $dp_i(\zeta) = \zeta_2^i$ for $\zeta \in T_z T^*Q$ and $z = (q, p)$. If Q is one-dimensional, Ω is a volume form. For arbitrary dimension n, the *oriented Liouville volume* is obtained by taking n-fold exterior products of the symplectic form,

$$\Lambda = \frac{(-1)^{n(n-1)/2}}{n!} \underbrace{\Omega \wedge \ldots \wedge \Omega}_{n \text{ times}} \tag{2.10}$$

which yields in coordinates

$$\Lambda = dq^1 \wedge \ldots \wedge dq^n \wedge dp_1 \wedge \ldots \wedge dp_n \,.$$

We will also need [66]

Definition 2.4. *Let (M, Ω) and (N, Ξ) be symplectic manifolds. A smooth map $\phi : M \to N$ is called symplectic, if Ω is the pull-back of Ξ by ϕ. That is,*

$$\phi^* : \Omega_z(\xi, \eta) = \Xi_{\phi(z)}(T_z\phi \cdot \xi, T_z\phi \cdot \eta) \,, \quad for \quad \xi, \eta \in T_z M \,,$$

where the tangent map

$$T_z\phi : T_z M \to T_{\phi(z)}N \,, \quad T_z\phi \cdot \zeta = \left. \frac{d}{ds}\phi(\gamma(s)) \right|_{s=0}$$

explains the derivative of ϕ at $z \in M$; here $\gamma(s)$ denotes a curve in M that satisfies $\gamma(0) = z$ and $\dot{\gamma}(0) = \zeta$, and the vector $T_z\phi \cdot \zeta$ only depends on ζ but not on the curve.

We consider the particular case of the Hamiltonian flow map $\Psi_t : T^*Q \to T^*Q$ with standard symplectic form $\Omega = dq^i \wedge dp_i$ for which the next statement holds [98].

Proposition 2.5 (Arnold 1989)**.** *Hamiltonian flows are symplectic.*

Although it is common knowledge, we shall prove the *Liouville Theorem* as following from the symplecticness property for the readers convenience:

Corollary 2.6. *Hamiltonian flows preserve the Liouville volume $\Psi_t^* \Lambda = \Lambda$.*

Proof. The wedge product commutes with the pull-back, i.e., the pull-back of the wedge product is the wedge product of the pull-back. Hence (2.10) gives

$$\Psi_t^* \Lambda = \frac{(-1)^{n(n-1)/2}}{n!}(\Psi_t^* \Omega) \wedge \ldots \wedge (\Psi_t^* \Omega) \,,$$

but as $\Psi_t^* \Omega = \Omega$, according to Proposition 2.5, we conclude that $\Psi_t^* \Lambda = \Lambda$. $\qquad\square$

Assumption 2.7. *For the sake of convenience we shall introduce mass-scaled coordinates $q \mapsto M^{-1/2}q$ and $p \mapsto M^{-1/2}p$, where M is the symmetric positive-definite mass matrix. This is certainly a symplectic transform, i.e., it preserves the wedge product, and it allows us to set the molecular masses to one in what follows.*

2.1.1. Statistical mechanics: ensemble concepts So far we have considered a single molecule only; on physical grounds however this is often not the appropriate perspective: First of all, real experiments mostly deal with ensembles of molecules rather than with a single molecule. Secondly, it is doubtful that an experimentalist can precisely prepare or measure the initial state of a single molecule at room temperature.

Hence we study ensembles of molecules. The ensemble will be represented by a probability distribution of initial states, where the choice of the respective distribution will depend on the problem under consideration; situations where chemical reactions occur and the number of particles may change, require other ensemble concepts than

ensembles with a constant number of particles. We shall mainly treat the case of an ensemble at constant temperature with a constant number of particles and constant volume, the so-called *canonical ensemble*. The time evolution of the initial ensemble is then governed by the dynamics of the single molecules [67]. This can be understood in the following way: Let us abbreviate $z = (q, p)$, and let $f_0(z)$ be the initial preparation of the statistical ensemble. Moreover we denote by Ψ_t the Hamiltonian flow in phase space $E = T^*Q$. Then the density in time $f(z, t)$ will be the initial density that is transported along the orbits of the flow Ψ_t, that is,

$$f(z, t) = (f_0 \circ \Psi_{-t})(z), \quad \text{with} \quad f_0 = f(\cdot, 0). \tag{2.11}$$

Thus, the density f at time t is the *push-forward* of the initial density f_0 by the flow map Ψ_t. This is a consequence of Liouville's Theorem which implies conservation of probability. The last equation can be rephrased in terms of an evolution equation for f. If we denote by X_H the Hamiltonian vector field that is defined by (2.7), then f is governed by the Liouville equation

$$\partial_t f(z, t) = -\mathcal{L}f(z, t), \quad f(\cdot, 0) = f_0, \tag{2.12}$$

where the Liouville operator $\mathcal{L}$ is defined as

$$\mathcal{L} = X_H(z) \cdot \nabla, \tag{2.13}$$

where ∇ is the derivative with respect to $z^1, \ldots, z^{2n}$. The Liouvillian is skew-adjoint, if it is considered to act on an appropriate subspace of the Hilbert space $L^2(dz)$. For further reading the interested reader is referred to the original work of Koopman [99]. See also [100] and the references therein.

It is easy to check that the transported density (2.11) satisfies the Liouville equation, and in fact the Liouvillian is the generator of the corresponding *semigroup*,

$$f_0 \circ \Psi_{-t} = \exp(-t\mathcal{L})f_0. \tag{2.14}$$

Accordingly a stationary density is a function ρ for which $\rho(z) = (\rho \circ \Psi_{-t})(z)$ for all $z \in E$ at all instances of time t. Of course, *any* function of the Hamiltonian, $\rho(z) = f(H(z))$, is a stationary solution of the Liouville equation. One prominent representative of this class is the canonical density or Boltzmann density

$$\rho_{\text{can}}(z) = \frac{1}{Z} \exp(-\beta H(z)), \tag{2.15}$$

where

$$Z = \int_E \exp(-\beta H(z)) \, dz, \tag{2.16}$$

is the normalization constant (partition function), $\beta = 1/T$ is the inverse temperature, and dz denotes the Lebesgue measure associated with the Liouville form (2.10).

The semigroup notation above emphasizes the mathematical equivalence of the particle dynamics and the ensemble dynamics [101]. Indeed the relation (2.14) allows for likewise studying the statistical properties of the Hamiltonian system (2.7) in terms of the flow map Ψ_t: Consider the set $\mathcal{M}$ of probability measures on E, that are absolutely continuous with respect to the Lebesgue measure dz (Liouville measure). According to the Liouville Theorem, dz is preserved by the flow and we can introduce the *transfer operator* or Frobenius-Perron operator P_t that acts on measures $\mu \in \mathcal{M}$,

$$(P_t\mu)(B) = (\mu \circ \Psi_{-t})(B) \quad \text{for all measurable sets } B \subset E$$

by means of the action on the associated smooth density,

$$(P_t f_0) = f_0 \circ \Psi_{-t}.$$

This amounts to interpreting the transfer operator in a probabilistic sense for a single particle, and we can rephrase the stationarity of the Gibbs density, stating that

$$\mu_{\text{can}}(dz) = \rho_{\text{can}}(z)\,dz$$

is an invariant measure of the Hamiltonian dynamics, that is, it satisfies $P_t\mu_{\text{can}} = \mu_{\text{can}}$. Consult [67, 102] for a detailed discussion in the context of metastability.

Stochastic Hamiltonian systems In general, statistical mechanics is about computing average quantities with respect to some probability measure. So far we have omitted the problem of how to compute such averages. In molecular dynamics the expectation value of an observable $A(z)$ is typically computed via time averages, exploiting the notion of ergodicity: a typical trajectory will eventually visit all possible states with $\mu > 0$. In a more formal language this means

$$\mathbf{E}A(z) = \int_E A(z)\mu(dz) = \lim_{T\to\infty} \frac{1}{T}\int_0^T A(\Psi_t z)\,dt\,.$$

The main difficulty here is that the deterministic flow Ψ_t has no unique invariant measure, since any function of the Hamiltonian yields a stationary solution of the Liouville equation. Even worse, the only candidates for ergodic invariant measures, like the equidistribution on the surface of constant energy, the microcanonical measure, are singular with respect to the Liouville measure. Conversely this means that we cannot compute averages with respect to the canonical measure μ_{can} from single particle trajectories [67]. In fact very few Hamiltonian systems are known to be ergodic: certain billiards and geodesic flows on surfaces of constant negative curvature [103, 104].

However it is possible to define a *stochastic* Hamiltonian system by averaging out the momenta from the deterministic flow, by which we obtain a dynamical system on configuration space only. For simplicity, we shall make two arrangements: Firstly, we consider only observables $A(q)$ that do not depend on the momenta $p \in T_q^*Q$. Secondly, we fix an observation time span $\tau > 0$ and define the *discrete* flow map

$$z_{k+1} = \Psi_\tau z_k\,, \quad z_k = (q(t_k), p(t_k))\,.$$

Notice that here the subscript denotes a time index that should not be confused with the respective coordinate label; anyway the distinction should be clear from the context. Further note that both restrictions are not too severe, as the restriction on the observables can mostly be circumvented by analytically doing the momentum averaging, and the discrete flow can be thought of stemming from a symplectic discretization of the Hamiltonian equations of motion [105, 82].

Let π denote the natural bundle projection $\pi : T^*Q \to Q$ from the cotangent bundle onto its base, and we let the parametric function

$$\varrho_q(p) = \frac{1}{Z(q)} \exp\left(-\frac{\beta}{2}g^{ij}(q)p_i p_j\right)$$

denote the probability density of the momentum variable p according to the Gibbs distribution ρ_{can}. The function $Z(q)$ normalizes the momentum density to one. Clearly the momentum distribution depends on the configuration q by means of the inverse metric tensor and the normalization constant; nevertheless it is locally Gaussian. We now introduce a *stochastic* Hamiltonian system as iterates of the map

$$q_{k+1} = (\pi \circ \Psi_\tau)(q_k, p_k) \tag{2.17}$$

17

with p_k randomly chosen according to the momentum distribution $\varrho_{q_k}(\cdot)$. The respective discrete *spatial* transfer operator S_τ that takes probability densities on Q forward in time, can then be defined as

$$S_\tau f(q) = \int (f \circ \pi \circ \Psi_\tau)(q, p)\, \varrho_q(p)\, dp\,.$$

Note the following important difference between the spatial and the full transfer operator P_t: The full propagator satisfies the semigroup property $P_{t+s} = P_s P_t$. However, the definition of the spatial propagator involves a momentum average, and in general averaging over the initial momenta and integrating up to time, say, 2τ gives a different result than averaging at the initial time and the intermediate time τ. In particular this means that S_τ has no generator.

In the following we will consider S_τ on the Hilbert space $L^2(\nu)$, that is weighted by the reduced probability measure $\nu(dq) = \varrho_0(q)dq$, where ϱ_0 is the marginal

$$\varrho_0(q) = \int_{T_q^* Q} \rho_{\mathrm{can}}(q, p)\, dp\,.$$

The weighted Hilbert space is defined as

$$L^2(\nu) = \left\{ v : Q \to \mathbf{R} \,\Big|\, \int_Q (v(q))^2\, \nu(dq) < \infty \right\} \tag{2.18}$$

with the scalar product

$$\langle u, v \rangle_\nu = \int_Q u(q) v(q)\, \nu(dq)\,. \tag{2.19}$$

The invariant probability measure that seems to be naturally associated with the stochastic Hamiltonian dynamical system (2.17) is ν. This means for the associated propagator that it must satisfy $S_\tau \mathbf{1} = \mathbf{1}$, as $\mathbf{1}$ denotes the invariant density with respect to ν. And indeed, it follows immediately from the fact that the canonical density is normalized that $S_\tau \mathbf{1} = \mathbf{1}$. Moreover the following is true [67]:

Proposition 2.8 (Schütte 1998). *Let $Q = \mathbf{T}^n$ be periodic. Then the transfer operator $S_\tau : L^2(\nu) \to L^2(\nu)$ has a simple eigenvalue $\lambda = 1$ that is bounded away from the remaining spectrum. Hence $\mathbf{1}$ is the unique invariant density in $L^2(\nu)$.*

2.2. Stochastic Langevin dynamics

Using deterministic dynamics has several advantages. First of all deterministic Hamiltonian systems are widely used in the molecular dynamics community and there is a variety of efficient numerical algorithms available [92, 106]. Moreover there is a well-established covariant formalism which is flexible enough so as to allow for the treatment of infinite dimensional problems, systems with symmetries or systems on manifolds [81]. However one severe drawback for molecular applications is the fact that there is no unique invariant phase space measure.

The *sampling problem* is mostly addressed by means of certain thermostatting techniques like Nosé-Hoover, Berendsen or stochastic Andersen thermostats [88, 107]. Mostly, these algorithms modify the equations of motion in such a way that the dynamics samples the canonical density, provided the Hamiltonian flow is ergodic with respect to the microcanonical measure. This is a very strong assumption, and it is well-known that the ordinary Nosé-Hoover thermostat suffers from ergodicity problems for certain classes of Hamiltonians [108, 109]. This pathology can be

removed by employing extensions to the single-oscillator chain or by imposing constant temperature constraints [110, 111, 112]. But even then, the sampling works well *only if* the dynamics is ergodic, and conditions to guarantee ergodicity are still lacking. Additionally all these more sophisticated methods have in common that due to their complexity they are relatively hard to implement, and they require a careful adjustment of the parameters involved. See the recent survey article [113].

A promising alternative is stochastic Langevin dynamics or Brownian (Smoluchowski) dynamics [13]. These systems are proven to be ergodic under sufficiently weak assumptions which are similar to those from the last section, i.e., periodic or bounded configuration space [114, 115]. A Langevin system can be regarded as a mechanical system with additional noise and dissipation (friction). The noise can be thought of modelling the influence of a heat bath surrounding the molecule and the dissipation is chosen such as to counterbalance the energy fluctuations due to the noise. In its traditional form the equations of motion read [116]

$$\dot{q}(t) = M^{-1}p(t)$$
$$\dot{p}(t) = -\nabla V(q(t)) - \gamma M^{-1}p(t) + \sigma \dot{W}(t),$$

where $\gamma, \sigma \in \mathbf{R}^{n \times n}$ are positive-definite matrices, and $W(t)$ is the standard n-dimensional Wiener process [117]. Here we use the expedient notation $\dot{W}(t)$ rather than $dW(t)$ for the increment of the Wiener process at the price of being less exact with regard to its correct interpretation [118]; so far it is sufficient to state that we understand the last equation in the sense of Itô, and therefore we will continue using this notation without further comments and omitting the time argument in most cases.

The fluctuation-dissipation theorem of Kubo [119] states that, if friction and noise satisfy the fluctuation-dissipation relation $2\gamma = \beta\sigma\sigma^T$, then the Langevin equation defines a Markov process that is ergodic with respect to the canonical measure μ_{can}. For our purposes it is convenient to rewrite the Langevin equation in a slightly different form: Let $H : T^*\mathbf{R}^n \to \mathbf{R}$ be the molecular Hamiltonian in mass-scaled coordinates

$$H(q,p) = \frac{1}{2}\langle p, p \rangle + V(q)$$

with conjugate coordinates $z = (q,p)$. Then the Langevin equation can be written as

$$\dot{z}(t) = (\mathbb{J} - \Gamma)\nabla H(z(t)) + \Sigma \cdot \dot{W}(t), \tag{2.20}$$

where ∇ denotes the derivative in $\mathbf{R}^{2n}$, and we have introduced the $2n \times 2n$ matrices

$$\mathbb{J} = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix}, \quad \Gamma = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \gamma \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma \end{pmatrix}.$$

Remark 2.9. *In principle the mass scaling $q \mapsto M^{-1/2}q$, $p \mapsto M^{-1/2}p$ affects both friction and noise terms in the Langevin equation in the way that the coefficients γ, σ scale according to $\gamma \mapsto M^{-1/2}\gamma M^{-1/2}$ and $\sigma \mapsto M^{-1/2}\sigma$. Hence the friction and noise coefficients become mass-dependent as well [39]. However it is easy to see that the mass scaling preserves the fluctuation-dissipation relation, and we refrain from making the scaling explicit. The reader should apologize this abuse of notation.*

Covariant form of the Langevin equation: Hamiltonian formulation
Unfortunately there is no canonical way that would lead to a covariant formulation of the Langevin equation in the spirit of the Euler-Lagrange equations or Hamilton's equations. As we shall see below the Langevin equation can be understood either as an Itô or as a Stratonovich equation, as long as we restrict our attention to

point transformations (i.e., symplectic lifts of configuration space maps to phase space); the distinction becomes irrelevant. The restriction to point transformations is not too severe, for the vast majority of physical problems that involve changes of variables are of that type. We will see below that equation (2.20) transforms like a dissipative Hamiltonian vector field, no matter whether it is interpreted as an Itô or a Stratonovich equation.[1] That is, the Langevin equation has some very specific properties as compared to general hypo-elliptic diffusion equations which are due to its Hamiltonian origin.

To the best of the author's knowledge the available approaches for covariant representations use projection operator techniques, where the system is coupled to a heat bath [121, 122]. These approaches involve very specific assumptions concerning timescale separation between system and bath variables or are based on Markov approximations of the friction kernel [123]. In principle one could simply apply Itô's formula to the Langevin equation, and recognize that it boils down to the familiar change-of-variables formula. Here we adopt an alternative approach by considering the associated evolution equation: The time evolution of a function $u \in L^1(dz)$ is governed by the Kolmogorov backward equation

$$\partial_t u(z,t) = \mathcal{A}_{\mathrm{bw}} u(z,t), \quad u(\cdot,0) = u_0, \tag{2.21}$$

where the backward generator is defined as

$$\mathcal{A}_{\mathrm{bw}} = \frac{1}{2}\Delta_{\Sigma,\mathbf{R}^{2n}} + (\mathbb{J} - \Gamma)\nabla H(z) \cdot \nabla. \tag{2.22}$$

where $\Delta_{\Sigma,\mathbf{R}^{2n}}$ is the degenerate Laplacian with respect to the covariance matrix, i.e.,

$$\Delta_{\Sigma,\mathbf{R}^{2n}} = \mathrm{tr}(\Sigma\Sigma^T\nabla^2).$$

Now we can easily study the transformation properties of the Langevin equation by means of the corresponding backward equation: Let $M = T^*Q$ be a symplectic manifold of dimension $2n$ that is endowed with a Riemannian metric h_0, and understand ∇ as the covariant derivative in TM. Recall that the Laplace-Beltrami can be computed as the trace of the Hessian [68]. Further recall that the Hessian $\nabla^2 f : T_z M \to T_z M$ is defined as the matrix associated with the bilinear form $d^2 f(X,Y)$, where $X, Y \in TM$ are two vector fields that are evaluated point-wise at $z \in M$. It follows from Leibniz' rule that

$$\nabla_X df(Y) = d^2 f(X,Y) + df(\nabla_X Y),$$

and we obtain in coordinates:

$$d^2 f(X_i, X_j) = \frac{\partial^2 f}{\partial z^i \partial z^j} - \Gamma_{ij}^k \frac{\partial f}{\partial z^k}, \quad X_l = \frac{\partial}{\partial z^l}.$$

Here Γ_{ij}^k are the symmetric Christoffel symbols associated with the Riemannian metric h_0 on TM that contains the inverse covariance matrix. Now we define the following sub-Riemannian metric

$$h(z) = F(z)\left(\Sigma\Sigma^T\right)^{-1} F^T(z),$$

where F is obtained by a Cholesky decomposition of the phase space metric $h_0 = F^T F$. (The metric h is sub-Riemannian because only the momentum block is non-zero; see [124] for a more general discussion on sub-Riemannian geometry.) The corresponding symmetric Hessian matrix then reads in coordinates

$$(\nabla^2 f)_j^i = h^{il}\left(\frac{\partial^2 f}{\partial z^l \partial z^j} - \Gamma_{lj}^k \frac{\partial f}{\partial z^k}\right). \tag{2.23}$$

---

[1]For the geometry of dissipative Hamiltonian vector fields without noise, see [120].

Double contraction (trace) of the Hessian yields the degenerate Laplace-Beltrami operator $\Delta_{\Sigma,M}$ on the phase space $M = T^*Q$, in other words: the diffusion generator of the Langevin equation on the phase space M. We state:

Lemma 2.10. *Let Q be a Riemannian manifold that is endowed with a metric g which is of the form $g = B^T B$, and let $H : T^*Q \to \mathbf{R}$ be the Hamiltonian*

$$H(q,p) = \frac{1}{2} \left\langle g(q)^{-1}p, p \right\rangle + V(q) \,.$$

Moreover we denote by $\tilde{\gamma}$ and $\tilde{\sigma}$ the usual friction and noise coefficients that satisfy the fluctuation-dissipation relation $2\tilde{\gamma} = \beta\tilde{\sigma}\tilde{\sigma}^T$. Abbreviating the canonical coordinates by $z = (q,p)$, the coordinate-invariant form of the Langevin equation is

$$\dot{z}(t) = (\mathbb{J} - \Gamma) \nabla H(z(t)) + \Sigma \cdot \dot{W}(t) \,, \tag{2.24}$$

where the $2n \times 2n$ matrices are defined as

$$\mathbb{J} = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix}, \quad \Gamma = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & B^T\tilde{\gamma}B \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & B^T\tilde{\sigma} \end{pmatrix}.$$

Unwrapping the matrix-vector notation, (2.24) is equivalent to

$$\begin{aligned} \dot{q}^i &= \frac{\partial H}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial H}{\partial q^i} - \gamma_{ij}\frac{\partial H}{\partial p_j} + \sigma_{ij}\dot{W}^j \,, \quad i = 1, \ldots, n \end{aligned} \tag{2.25}$$

where $\gamma = B^T\tilde{\gamma}B$, and $\sigma = B^T\tilde{\sigma}$ are q-dependent friction and noise matrices that satisfy the fluctuation-dissipation relation $2\gamma = \beta\sigma\sigma^T$.

Proof. The assertion concerning fluctuation-dissipation is trivial. For the first part it is sufficient to consider the evolution equation associated with the Langevin equation (2.20). To this end consider the molecular Hamiltonian

$$\tilde{H}(\eta,\xi) = \frac{1}{2} \left\langle \xi, \xi \right\rangle + \tilde{V}(\eta) \,, \quad (\eta,\xi) \in T^*\mathbf{R}^n \cong \mathbf{R}^n \times \mathbf{R}^n \,,$$

which gives rise to the Langevin equation (2.20). We shall study the transformation properties of (2.20) by means of the corresponding backward equation

$$\partial_t u = \left(\frac{1}{2}\Delta_{\Sigma,\mathbf{R}^{2n}} + \left(\mathbb{J} - \tilde{\Gamma} \right) \tilde{\nabla}\tilde{H} \cdot \tilde{\nabla} \right) u \,.$$

Here $\tilde{\nabla} = (\partial/\partial\eta^1, \ldots, \partial/\partial\xi_n)$ is the derivative in $\mathbf{R}^{2n}$, and $\tilde{\Gamma}, \tilde{\Sigma} \in \mathbf{R}^{2n \times 2n}$ are defined in an obvious way. Consider a diffeomorphism $\phi : Q \to \mathbf{R}^n$ with $\eta = \phi(q)$ which has a symplectic lift $T^*\phi : T^*Q \to T^*\mathbf{R}^n$ to phase space that is given by $(\eta,\xi) = (\phi(q), \mathbf{D}\phi(q)^{-T}p)$. In coordinates (q,p) the Hamiltonian reads

$$H(q,p) = \frac{1}{2} \left\langle g(q)^{-1}p, p \right\rangle + V(q) \quad (V = \tilde{V} \circ \phi) \,,$$

with the induced metric tensor $g = B^T B$, $B = \mathbf{D}\phi$. Writing the backward equation in the new coordinates is straightforward: using (2.23) we obtain for the diffusion part

$$\frac{1}{2}\Delta_{\Sigma,T^*Q} = \frac{1}{2} \operatorname{tr} \left(B^T\tilde{\sigma}\tilde{\sigma}^T B \, \mathbf{D}_2^2 \right) =: \frac{1}{2} \operatorname{tr} \left(\Sigma\Sigma^T \, \nabla^2 \right) \,,$$

where $\mathbf{D}_2$ denotes the derivatives with respect to the second (momentum) slot, and ∇ denotes the derivatives with respect to the new coordinates. Also the drift is

easily accounted for: since (q, p) are conjugate, the Hamiltonian vector field keeps its canonical form; moreover the dissipation part transforms like an ordinary vector field on $T^*Q \cong \mathbf{R}^{2n}$. The thus transformed drift part reads

$$(\mathbb{J} - \Gamma)\nabla H = \mathbb{J}\nabla H - \begin{pmatrix} \mathbf{0} \\ B^T \tilde{\gamma} B^{-T} p \end{pmatrix}$$

from which the assertion follows upon noting that $B^T \tilde{\gamma} B^{-T} = B^T \tilde{\gamma} B g^{-1} = \gamma g^{-1}$. $\quad\square$

The proof reveals the origin of the Itô-Stratonovich equivalence in case of the Langevin equation without using the Itô formula: regarding the backward operator only the momentum derivatives contributes to the Laplace-Beltrami, and the momentum transforms linearly under point transformations. In the language of Itô calculus this means that the second order momentum derivatives in the Itô formula, which make up the additional terms, all vanish.

The time evolution of physical densities $f \in L^1(dz)$ is governed by the Kolmogorov forward equation, that is in some sense dual to the backward equation [102]

$$\partial_t f(z, t) = \mathcal{A}_{\mathrm{fw}} f(z, t), \quad f(\cdot, 0) = f_0, \tag{2.26}$$

where the forward generator,

$$\mathcal{A}_{\mathrm{fw}} = \frac{1}{2}\Delta_{\Sigma, M} - \nabla \cdot (\mathbb{J} - \Gamma)\nabla H. \tag{2.27}$$

is the formal adjoint of the backward generator on the Hilbert space $L^2(dz)$. (Note that the diffusion part of the operator is self-adjoint in $L^2(dz)$ since $\Delta_{\Sigma, M}$ involves only momentum derivatives, but the metric merely depends on the configuration variables.) The forward operator can then be rewritten according to

$$\mathcal{A}_{\mathrm{fw}} = \frac{1}{2}\Delta_{\Sigma, M} - (\mathbb{J} - \Gamma)\nabla H \cdot \nabla + \nabla \cdot (\Gamma\nabla H),$$

where we have used that the Hamiltonian vector field is divergence-free. Alternatively the rightmost divergence can be written as $\mathrm{tr}(\gamma g^{-1})$. Provided that the fluctuation-dissipation relation is met, the canonical (Boltzmann) density is a stationary solution of the forward equation, independently of the chosen coordinate system:

$$\mathcal{A}_{\mathrm{fw}}\rho_{\mathrm{can}} = 0, \quad \rho_{\mathrm{can}} = \frac{1}{Z}\exp(-\beta H).$$

2.3. High-friction dynamics: Brownian motion

The Smoluchowski equation be understood as the high-friction limit of the Langevin equation. Unfortunately there is some ambiguity in the denotations throughout the literature: we shall use the term *Brownian motion* synonymously with *Smoluchowski equation* which describes a non-degenerate (elliptic) diffusion process in contrast to the degenerate (hypo-elliptic) Langevin equation. For the sake of illustration we shall briefly sketch the transition from Langevin to Smoluchowski for a one-dimensional system. To this end we write the Langevin equation in second-order form

$$M\ddot{q} = -\nabla V(q) - \gamma\dot{q} + \sigma\dot{W}.$$

We introduce a small parameter $\epsilon > 0$ by scaling the friction constant according to $\gamma \mapsto \gamma/\epsilon$; keeping the correct invariant density requires the scaling $\sigma \mapsto \sigma/\sqrt{\epsilon}$ for the

noise coefficient. In other words, the high-friction limit amounts to letting ϵ go to zero, so that we arrive at the singularly perturbed problem

$$M\ddot{q}_\epsilon = -\nabla V(q_\epsilon) - \frac{\gamma}{\epsilon}\dot{q}_\epsilon + \frac{\sigma}{\sqrt{\epsilon}}\dot{W},$$

So far the limit $\epsilon \to 0$ would yield the boring result $\dot{q} = 0$, as can be seen upon multiplying the whole equation by ϵ. However this changes if we consider the dynamics on the *diffusive timescale* $t \mapsto t/\epsilon$, on which the equations of motion take the form

$$\epsilon^2 M\ddot{q}_\epsilon = -\nabla V(q_\epsilon) - \gamma\dot{q}_\epsilon + \sigma\dot{W}.$$

Taking the limit $\epsilon \to 0$ formally leads to the Smoluchowski equation

$$\gamma\dot{q}_0 = -\nabla V(q_0) + \sigma\dot{W}. \tag{2.28}$$

The little word *formally* should be taken literally, because the white noise in the equation is unbounded, and therefore we cannot be sure that $\epsilon^2\ddot{q}_\epsilon$ goes to zero as ϵ goes to zero. Nevertheless it can be shown that the sample paths $q_\epsilon(t)$ converge to $q_0(t)$ with probability one [125, 126].

Covariant form of the Smoluchowski equation: geometric Itô formula
Loosely speaking the Smoluchowski equation can be understood as some sort of diffusive limit of the Langevin equation. However we have to take care that the Smoluchowski equation (SE) transforms in a way that is consistent with transformation properties of the Langevin equation (LE). Given a change of variables $\phi: \mathbf{R}^n \to Q$, this property is expressed in the following commutative diagram

$$
\begin{array}{ccc}
(\text{LE}, T^*\mathbf{R}^n) & \xrightarrow{\;T^*\phi: T^*Q \to T^*\mathbf{R}^n\;} & (\text{LE}, T^*Q) \\
{\scriptstyle \gamma\to\infty}\big\downarrow & & \big\downarrow{\scriptstyle \gamma\to\infty} \\
(\text{SE}, \mathbf{R}^n) & \xrightarrow{\;\phi: Q \to \mathbf{R}^n\;} & (\text{SE}, Q)
\end{array}
$$

Unlike the Langevin equations we understand the Smoluchowski equation strictly in the sense of Itô, because here the distinction really matters as we will see below. Without loss of generality we set $\gamma = \mathbf{1}$. Then the fluctuation-dissipation relation implies that $\sigma\sigma^T \propto \mathbf{1}$. In other words, the noise matrix σ is proportional to an orthogonal matrix. We may even choose $\sigma = \sqrt{2\beta^{-1}}$ scalar, since the white noise is invariant under orthogonal transformations [127]. Now the multidimensional Smoluchowski equation takes the form[2]

$$\dot{q} = -\operatorname{grad} V(q) + \sqrt{2\beta^{-1}}\dot{W}, \quad q \in \mathbf{R}^n. \tag{2.29}$$

Note that we have tacitly replaced the derivative ∇ by the gradient; the first one is a one-form, whereas the latter defines a vector field on $\mathbf{R}^n$. This might seem somehow arbitrary but the replacement is owed to the fact that the friction matrix in the Langevin equation transforms according to $\gamma \mapsto B^T\gamma B$ under a change of coordinates. Hence $\gamma\dot{q}$ is an element of the cotangent space, and suppressing γ amounts to replacing ∇V by the gradient. For integrable functions $v \in L^1(dq)$ let us introduce the Kolmogorov backward equation that is associated with (2.29),

$$\partial_t v(q,t) = \mathcal{A}_{\mathrm{bw}}\, v(q,t), \quad v(\cdot, 0) = v_0$$

---

[2]Setting $\gamma = \mathbf{1}$ is along the lines of the mass scaling, for we can always make γ vanish in the Smoluchowski equation by scaling the coordinates according to $q \mapsto \gamma^{-1/2}q$.

with the elliptic backward generator

$$\mathcal{A}_{\mathrm{bw}} = \beta^{-1}\Delta - \operatorname{grad} V(q) \cdot \nabla \,.$$

We shall prove the following statement which is due to [69]:

Lemma 2.11. *Let Q be a configuration manifold of dimension n that is equipped with a Riemannian metric $g = \langle \cdot, \cdot \rangle_g$ and coordinates $q = (q^1, \ldots, q^n)$. Then the covariant Smoluchowski equation in Itô form reads*

$$\dot{q} = -\operatorname{grad} V(q) + b(q) + a(q) \cdot \dot{W} \tag{2.30}$$

with the coefficients

$$b^i(q) = -\beta^{-1} g^{jk} \Gamma^i_{jk} \quad and \quad a(q) = \sqrt{2\beta^{-1} g(q)^{-1}} \tag{2.31}$$

where the Γ^i_{jk} are the matrices of the symmetric Christoffel symbols on TQ,

$$\Gamma^i_{jk} = \frac{1}{2} g^{il} \left(\frac{\partial g_{kl}}{\partial q^j} + \frac{\partial g_{lj}}{\partial q^k} - \frac{\partial g_{jk}}{\partial q^l} \right) \,.$$

The Stratonovich equivalent is obtained by omitting the additional drift vector $b(q)$.

Proof. Recall the definition of the diffusion backward generator

$$\mathcal{A}_{\mathrm{bw}} = \beta^{-1}\Delta - \operatorname{grad} V \cdot \nabla \,.$$

Now recall that the gradient of a function $V : Q \to \mathbf{R}$ is defined by the equation $dV(w) = \langle \operatorname{grad} V, w \rangle_g$ for all vectors $w \in T_qQ$. In coordinates this means

$$(\operatorname{grad} V)^i = g^{ij} \frac{\partial V}{\partial q^j} \,.$$

Accordingly, we know that the diffusion on Q is generated by the Laplace-Beltrami operator on Q which is obtained as the double contraction (trace) of the Hessian matrix (2.23) or, equivalently, $\Delta = \operatorname{div} \circ \operatorname{grad}$. In coordinates the Laplacian reads [128]

$$\Delta v = g^{jk} \left(\frac{\partial^2 v}{\partial q^j \partial q^k} - \Gamma^i_{jk} \frac{\partial v}{\partial q^i} \right) \,,$$

where Γ^i_{jk} are the symmetric Christoffel symbols associated with the metric g. We observe that the first term in the Laplace-Beltrami appears as an ordinary diffusion, whereas the other one has the formal structure of a drift term. Thus the Laplacian can be split into two parts, and so the backward generator admits the representation

$$\mathcal{A}_{\mathrm{bw}} = \mathcal{A}_{\mathrm{diff}} + \mathcal{A}_{\mathrm{drift}}$$

with the single operators

$$\mathcal{A}_{\mathrm{diff}} = \beta^{-1} g^{ij} \frac{\partial^2}{\partial q^i \partial q^j}$$

$$\mathcal{A}_{\mathrm{drift}} = - \left(\beta^{-1} g^{ij} \Gamma^k_{ij} + g^{kl} \frac{\partial V}{\partial q^l} \right) \frac{\partial}{\partial q^k}$$

Writing down the associated Itô stochastic differential equation is straightforward. The Stratonovich form is obtained by simply considering the Smoluchowski equation as an ordinary first-order vector field which amounts to erasing all terms that involve Christoffel symbols. This proves the assertion. □

24

Remark 2.12. *We point out an important distinction: on the Euclidean configuration space $U \subseteq \mathbf{R}^n$ consider the Brownian motion of a free particle*

$$\dot{u}(t) = \dot{W}(t).$$

Now let $\phi : Q \rightarrow U$ be a change of coordinates, such that $u = \phi(q)$. The map ϕ induces a Riemannian metric $g = J_\phi^T J_\phi$ on Q, where $J_\phi(q) = \mathbf{D}\phi(q)$ is the Jacobian of the transformation. According to Lemma 2.11 the respective Itô equation on Q becomes

$$\dot{q}^i(t) = -\frac{1}{2}g^{jk}\Gamma^i_{jk} + a^{ij}\dot{W}_j(t),$$

where a is the uniquely defined positive-definite matrix square root of g^{-1}. However note that a is not the only admissible choice for the diffusion matrix, that stems from reinterpreting the diffusion part $\mathcal{A}_{\text{diff}} = g^{ij}\partial^2/\partial q^i \partial q^j$ in the associated backward generator in terms of white noise in the differential equation. In particular if we demand that the two processes are related pathwise by the coordinate map, $u(t) = \phi(q(t))$, then the noise term is altered due to:

$$\dot{q}^i(t) = -\frac{1}{2}g^{jk}\Gamma^i_{jk} + J_\phi^{ij}\dot{W}_j(t),$$

where J_ϕ^{ij} are the entries of the inverse Jacobian J_ϕ^{-1}. (In many cases, J_ϕ is uniquely related to the metric $g = J_\phi^T J_\phi$ by a Cholesky decomposition. If ϕ is conformal, it may even happen that J_ϕ is nontrivial, although $g = \mathbf{1}$ and thus $\sigma = a$.)

As before, it is the Kolmogorov forward equation which governs the time evolution of densities. For functions $f \in L^1(dq)$ the forward equation reads [129]

$$\partial_t f(q,t) = \mathcal{A}_{\text{fw}} f(q,t), \quad f(\cdot, 0) = f_0.$$

The forward operator is defined as the formal adjoint of the backward operator in the Hilbert space $L^2(d\sigma)$, where $d\sigma = \sqrt{\det g(q)}\,dq$ denotes the volume element on Q. It is easy to see, using integration by parts, that the forward operator has the form[3]

$$\mathcal{A}_{\text{fw}} = \beta^{-1}\Delta + \operatorname{grad} V(q) \cdot \nabla + \Delta V(q).$$

It is then straightforward to show that the Gibbs density

$$\rho(q) = \frac{1}{Z}\exp(-\beta V(q))$$

is a stationary solution of the forward equation, i.e., $\mathcal{A}_{\text{fw}}\rho = 0$. In turn it follows that

$$\nu(dq) = \frac{1}{Z}\exp(-\beta V(q))\sqrt{\det g(q)}\,dq$$

is the invariant measure (Gibbs measure) of the Smoluchowski equation (2.30). (Notice that the Gibbs density ρ is the density of the Gibbs measure ν with respect to the volume on Q.) There is yet another operator which we will often consider, and that is formally equivalent to the backward operator but acts on a different function space. Following [102], we consider the Fokker-Planck equation that evolves densities in time with respect to the invariant probability measure $\nu(dq)$

$$\partial_t w(q,t) = \mathcal{L}\,w(q,t), \quad w(\cdot, 0) = w_0$$

with the (backward) Fokker-Planck operator

$$\mathcal{L} = \beta^{-1}\Delta - \operatorname{grad} V(q) \cdot \nabla,$$

which is regarded on a suitable subspace of $L^1(\nu)$. Clearly $\mathcal{L}\mathbf{1} = 0$, since the constant function $\mathbf{1}$ is the invariant Gibbs density with respect to ν. Moreover it can be shown that the Fokker-Planck generator is self-adjoint:

---

[3]It follows directly from two iterations of Stokes' Theorem (i.e., Green's formula) that the Laplace-Beltrami operator is essentially self-adjoint in the Hilbert space $L^2(d\sigma)$, but not in $L^2(dq)$.

Proposition 2.13. *Let the weighted Hilbert space be defined according to (2.18) with the respective weighted scalar product $\langle \cdot, \cdot \rangle_\nu$. Then $\langle \mathcal{L}u, v \rangle_\nu = \langle u, \mathcal{L}v \rangle_\nu$ for functions $u, v \in L^2(\nu)$, i.e., $\mathcal{L}$ is self-adjoint in $L^2(\nu)$.*

Proof. The proof for the Euclidean case is standard can be found, e.g., in [102]. Exploiting that $\mathcal{L}$ can be written as a Schrödinger operator in $L^2(d\sigma)$, the proof carries over to the manifold case considered here at almost no further expense. $\square$

2.4. Essential degrees of freedom

Most systems of nonlinear differential equations, either deterministic or stochastic, that model *real-world* problems are characterized by vastly different timescales, on which certain dynamical effects happen [130]. In many molecular systems such timescales are well-separated, since different physical interactions in the molecular potential induce different characteristic timescales. Roughly speaking, we can assign a natural timescale τ to a physical process that is inversely proportional to its average energy. Accordingly, the vibrations of the molecular bonds (about 1 femtosecond) are typically the fastest modes in the system. We introduce the ratio

$$\epsilon = \frac{\tau_{\text{fast}}}{\tau_{\text{slow}}} \ll 1$$

which serves as a measure for the separation of timescales in the system. Often things become more complicated as the degrees of freedom are coupled among each other. Then the fast variables may induce slow motions elsewhere in the system, such that the assignment of timescales to certain modes gets difficult. Dividing a molecular system into slow and fast degrees of freedom is motivated by the observation that the conformational dynamics of a molecule is typically slow, as compared to the remaining degrees of freedom, since transitions between conformations are rare events. However this statement is in some respects circular, since the distinguishability of conformations pretty much depends on the choice of the right essential (slow) coordinate.

We understand conformation dynamics as taking place in configuration space; in particular in the Hamiltonian context, the restriction to the configuration variables was owed to generating the Gibbs measure as the unique invariant measure of the dynamics. But even then one might imagine that the distinction between conformational degrees of freedom and those which are of minor interest is not along temporal scales but along spatial scales, in case of which we could be interested in those degrees of freedom that have the largest amplitude [43]. In fact conformational behaviour is a spatial property of a molecule, in the sense that it is related to the its geometric shape. Nevertheless the dynamics between conformations is often driven by either crossing of energy barriers or other temperature-dependent entropic mechanisms, both of which are inherent temporal effects [131, 132]. Clearly one might think of other possible subdivisions: the principal decision to be made is whether the velocities or momenta should be taken into account; this is possible in principle; standard references for such purely deterministic approaches that treat the full phase space are [31, 32]; see also [133, 134] for some recent developments.

2.4.1. Spatial decomposition methods It is not the aim of this thesis to discuss the problem of finding good reaction coordinates in great detail or even to resolve it; for our purposes it is sufficient to assume that we are given a set of slow coordinates with significant spatial amplitude, for instance, by sufficient physical insight a priori or by
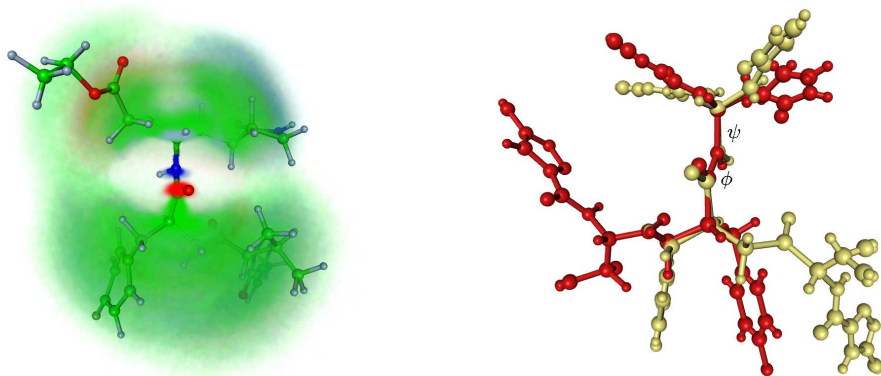
Figure 2. The left panel of the figure shows a volume rendering plot of the SARS-protease inhibitor at — allowedly unrealistic — temperature T=1500K. The clouds indicate the flexibility (fast oscillations) of the two main conformations which are depicted on the right, and which are well described by the two central torsion angles ϕ, ψ. See [136] for the simulation details.

statistical tools that operate on the produced time series *a posteriori*. Accordingly this subsection mainly surveys the available approaches for the identification of essential degrees in molecular systems, i.e., spatially extended motions that are slow in some appropriate sense. At the very beginning of this chapter we have seen that molecular interaction forces are defined in terms of internal degrees of freedom of a molecule. Hence it seems natural to suspect the essential coordinates among the internal coordinates. Indeed in many cases the conformation dynamics can be well described by specific dihedral (torsion) angles which link certain approximately rigid subunits of the molecule; see Figure 2 below for illustration. Other possible candidates for essential degrees of freedom are radii of gyration, certain intramolecular distances between selected endgroups of the molecule, or distances between parts of a molecule and some specific solvent molecules [135]. In any event the reader should keep in mind that essential degrees of freedom are not intrinsically defined by the molecule under consideration, but rather depend on the specific problem.

In case the essential degrees of freedom are not known a priori they can be possibly detected by statistical analysis of simulation data. There are plenty of such data-based methods and algorithms that all have in common that they sort out a linear subspace of the data space which is in some sense optimal, although the methods might be rather different in detail. This subspace is then assumed to be spanned by the essential variables [137]. Mostly these techniques work on Euclidean data only, and it would be nice to apply them in a more general framework to data that lies on a manifold; we shall discuss this issue below.

The most famous approach is certainly the method of Principal Component Analysis (PCA) which is also known as Karhunen-Loève expansion or Proper Orthogonal Decomposition (POD) to name just a few [138, 45].

The idea is to find a best-approximating k-dimensional subspace to a given set of data. We may characterize this subspace $S \subset \mathbf{R}^n$ by a projection operator $\mathscr{P}$,

mapping $\mathbf{R}^n$ onto S. We give a brief derivation of the method to provide some geometric intuition for the following discussion. Assume the data consist of sampled measurements $\{x(t_1), \ldots, x(t_N)\}$ of a trajectory $x(t) \in \mathbf{R}^n$. Finding the optimal subspace S amounts to solving the least-square problem[4]

$$\min_{\mathscr{P}} \sum_{l=1}^{N} \|x(t_l) - \mathscr{P}x(t_l)\|_2^2$$

where the rank of $\mathscr{P}$ is k. This problem can easily be reformulated as a related variational problem: Let S be a one dimensional subspace in $\mathbf{R}^n$ spanned by the vector $w \in \mathbf{R}^n$, such that the projection $\mathscr{P}$ becomes the linear map $\mathscr{P} = \langle \cdot, w \rangle w$. Since $\|x\|^2 = \|x - \mathscr{P}x\| + \|\mathscr{P}x\|^2$ holds true for any projection, the least-square problem is equivalent to maximizing the *energy* of the projection,

$$E[w] = \sum_{l=1}^{N} \langle x(t_l), w \rangle^2 - \lambda(\|w\|_2^2 - 1),$$

where the Lagrange multiplier imposes normalization upon the vector w. Taking the variation of $w = w_\epsilon$ keeping x fixed it follows by chain rule $\delta E_\epsilon = \langle \mathbf{d}E, \delta w_\epsilon \rangle$ that the critical points of E have to satisfy the equation

$$2\sum_{l=1}^{N} \langle x(t_l), w \rangle \langle x(t_l), \delta w_0 \rangle - 2\lambda \langle w, \delta w_0 \rangle = 0.$$

Since the last equation clearly must hold for arbitrary variations δw_ϵ we conclude that

$$\left(\sum_{l=1}^{N} \langle x(t_l), \cdot \rangle x(t_l) \right) w = \lambda w,$$

where (modulo normalization) the sum on the left hand side is an estimator of the symmetric covariance matrix of the data with accuracy $\mathcal{O}(N^{-1/2})$. Assuming that the covariance matrix has maximum rank n, the eigenvectors form a complete orthogonal set of vectors $w_1, w_2, \ldots, w_n$. Consequently, we may repeat this whole procedure on the orthogonal complement to $w = w_1$, and find that then the best approximation is given by the second eigenvector w_2, and so on. Finally the rank-k approximant S is spanned by the first k eigenvectors which is optimal in the sense that

$$\min_{\mathscr{P}} \sum_{l=1}^{N} \|x(t_l) - \mathscr{P}x(t_l)\|_2^2 = \sum_{j=k+1}^{n} \lambda_j$$

is the minimum achieved by any k-plane in $\mathbf{R}^n$, where the λ_j are the ordered eigenvalues of the covariance matrix; to put this differently, the chosen subspace contains the maximum variance of all k-dimensional linear subspaces [45]. By this procedure we find the optimal subspace; if we want to find the optimal *affine* subspace, we can exploit that it must pass through the mean of the data, and thus simply centre the data by subtracting the mean before computing the covariance.

Although POD is optimal at approximating a given data set it is not necessarily so for describing the dynamics that generates the particular data, for features of low

---

[4]There is some freedom in choosing a vector norm with respect to which the minimization is carried out. Frequent choices, for instance, are energy-based norms that are induced by an inner product $\langle u, v \rangle_A = \langle Au, v \rangle$, where $A = M$ may be the molecular mass matrix or the matrix of a quadratic Lyapunov function (notice that the former is equivalent to using mass-scaled coordinates). The specific choice of an inner product should be suited to the problem under consideration.

variance (energy) may induce important effects. In the molecular dynamics community a large amount of variance in the subspace is often taken as equivalent to the statement that the conformational dynamics takes place in this subspace. This is often wrong, as has been pointed out on various occasions; see [139, 140], for instance. Although in the derivation above we have nowhere assumed that the data follows an unimodal or even Gaussian distribution, it is self-evident that the method is reliable only for unimodal data. Hence a one-dimensional subspace containing, say, 99% of the variance is not necessarily the essential subspace with regard to the conformation dynamics.

Apparently, choosing the essential degrees of freedom merely according to their variance is not a good idea in general. Nevertheless we can consider the full set of eigenvectors $w_1, w_2, \ldots, w_n$ as a new basis for $\mathbf{R}^n$ that gives an ordering according to variance. This may serve as a hint what the essential subspace could be, and we know that the slow modes are good candidates for the conformational degrees of freedom. Hence we introduce characteristic timescales for the thus rotated modes,

$$ z(t) = R \cdot x(t) \, , \quad R = (w_1, \ldots, w_n) \in O(n) \, , $$

by the respective *decorrelation time* τ which is defined as the integral

$$ \tau = \int_0^\infty |\rho_i(s)| \, ds \, , \quad \rho_i(s) = \mathrm{cor}(z_i(0), z_i(s)) \, . $$

Loosely speaking, the decorrelation time indicates when $z(t)$ and $z(t+s)$ are effectively independent. We stress that the thus defined characteristic timescale is a technical notion rather than a sound and uniquely defined mathematical term; there are alternative definitions: in particular in the climate modelling community the integral is sometimes computed using the square of the autocorrelation function or the function itself [141]. However we claim that a slowly decaying autocorrelation function with $\tau \gg 1$ together with a large variance indicates possibly relevant dynamical behaviour. Of course this approach is not completely tight, but it is based on the fundamental observation that conformational changes in a molecule are related to spatially large and relatively slow rearrangements of the configuration variables.

Diagonalizing the data covariance matrix is not the only way to proceed; in particular the assignment of timescales can be linked with any other method as well. Another technique takes advantage of the insight that *correlated motions* in molecules, in particular proteins, are ubiquitous and often essential for biomolecular function. In principle, the covariance analysis can be extended to using the correlation matrix instead which means using an appropriately *normalized* covariance matrix. This approach, however, heavily relies on a quasi-harmonic treatment of the configurational ensemble, as it detects only linearly correlated motions [142, 143].

A quite promising approach is the Full Correlation Analysis which is based on the information theoretical concept of mutual information. It allows to detect and quantify any correlated motion from the Cartesian molecular dynamics trajectories. As in the previous cases the methods singles out a linear subspace of $\mathbf{R}^n$ which is spanned by maximally uncorrelated basis vectors which are the solution of a generalized nonlinear eigenvalue problem [54]. This method has proven superior to the classical correlation analysis on some occasions [144]. Related approaches are Independent Component Analysis [145] and Non-Gaussian Component Analysis [146].

Yet another dynamics-based technique is the method of Principal Interaction Patterns (PIP), as introduced by Hasselmann [147]. It is predominantly used in the climate modelling community, and it takes into account the dynamical system upon

which a reduced dynamical model is built: the aim is to minimize the difference

$$J[\Theta] = \int_0^T \|x(t) - x_\Theta(t)\| \, dt \,,$$

where $x_\Theta(t)$ is the solution of a reduced dynamical system that is described by a set of functions or parameters Θ; for instance, one may think of Θ as the linear projection $\mathscr{P}$ from above, such that the reduced system is simply the Galerkin projection of the original equations. Now PIP involves two steps, the first of which is an ensemble average over the initial values $x(0), x_\Theta(0)$, followed by minimization with respect to Θ. Eventually, finding the essential degrees of freedom x_Θ boils down to solving a nonlinear optimization problem in $\mathscr{P}$. Alternatively one might imagine that Θ parametrizes a reduced model [148]. Although theoretically appealing we believe that this method does not lead to computationally tractable problems for most molecular systems; we quote it for the sake of completeness; see also [149].

We finally mention that the problem of maximizing the decorrelation time can be addressed directly, which is known by the name of Optimal Persistence Patterns. This technique aims at identifying the linear subspace that has the slowest decay of correlation. There are several possible ways to measures persistence (e.g., by means of the decorrelation time τ), and it is shown in [150] that finding the optimally persistent subspace results in solving a generalized eigenvalue problem. However since it is our strong believe that conformation dynamics manifests itself as macroscopic behaviour in both space and time, we prefer the combined approach of least-square approximations like POD together with the analysis of decorrelation times.

Transfer operator approach The transfer operator approach to metastability rests upon the observation that conformation dynamics can be understood as flipping dynamics between distinct subsets of configuration space that are *almost invariant* under the dynamics [67, 151]. Hence the problem of identifying conformations amounts to the identification of almost invariant sets in configuration space, where *almost* is understood in a way that transitions between those sets are rare. The method exploits the close relationship between the flow of a dynamical system, either Langevin, Smoluchowski or stochastic Hamiltonian, with its associated Frobenius-Perron operator S_t on configuration space.

The key idea is that the natural *invariant* sets and measures are given by the eigenvectors or eigenfunctions to the eigenvalue $\lambda_0 = 1$, whereas the eigenfunctions to $\lambda_k < 1$ that are close to $\lambda_0 = 1$ correspond to the *almost invariant* sets [152]. Suppose that the probability of the system to be in configuration $q \in Q$ at time $t = 0$ is given by the unique invariant density $f_0 = \rho_{\text{can}}$. Then the transition probability $p(\tau, A, B)$ from $A \subset Q$ to $B \subset Q$ within the observation time τ is given by the fraction of the ensemble that has started in A at $t = 0$, and which has ended up in B at time $t = \tau$. Hence $p(\tau, A, B)$ is the conditional probability

$$\mathbf{P}_\rho \left[q(\tau) \in B \mid q(0) \in A \right] = \frac{\mathbf{P}_\rho \left[q(\tau) \in B \cap q(0) \in A \right]}{\mathbf{P}_\rho \left[q(0) \in A \right]} \,,$$

where subscript indicates that the initial preparation is given by ρ_{can}. Finally, a metastable set A is characterized by the requirement that $p(\tau, A, A) \approx 1$. This is to say that during a fixed observation time τ the system is likely to stay within the set. The algorithmic strategy that is put forward in [67] is to identify metastable subsets from the eigenfunctions of the Frobenius-Perron operator S_τ that correspond

to eigenvalues $\lambda < 1$ close to the Perron root $\lambda_0 = 1$. The number of metastable sets is then equal to the number of eigenvalues close to one, including $\lambda_0 = 1$ and counting multiplicity. The justification of this strategy is given in [102, 153].

2.4.2. Reaction coordinates The term *reaction coordinate* is typically used in a very loose sense and synonymous with *essential degrees of freedom*. More precise, a reaction coordinate can be understood as defining a family of isocommittor surfaces, such that all trajectory launched from the same surface have equal probability to first reach one metastable set before the other [154, 155]. The isocommittor surfaces are the codimension-one level sets of the reaction coordinates which, within the framework of Brownian motion, can be systematically computed from the Kolmogorov backward equation. For simplicity assume that there are only two conformations, or metastable sets $A \subset Q$ and $B \subset Q$. If $\mathcal{A}_{\mathrm{bw}}$ is the backward generator associated with the molecular diffusion process, then the solution of the boundary value problem

$$\mathcal{A}_{\mathrm{bw}}\Phi = 0\,, \quad \Phi|_{\partial A} = 0\,, \Phi|_{\partial B} = 1$$

yields the reaction coordinate in the sense that the probability that a process starting from $q \in Q$ reaches B before A is $\Phi(q)$; so to speak, the level sets $\Phi(q) = \xi$ with ξ between 0 and 1 are the isocommittor surfaces of the process [13]. This clearly presumes that the metastable sets have been identified in advance, e.g., by the transfer operator algorithm as presented in the last paragraph.

In principle the procedure works in the same way for processes that are generated by Langevin equations, with the only difference that the reaction coordinate becomes a function on phase space [156]. But notice that *reaction coordinate* with its probabilistic meaning lacks a physical interpretation in terms of the intramolecular motion, i.e., it does not *explain* the transition mechanism. This issue is addressed by algorithms like the String Method [157] or the Nudged Elastic Band method [158]. Further note that although the direct identification of reaction coordinates is theoretically appealing, it requires solving a high-dimensional PDE which is part of a current PhD thesis [159].

2.5. Problems related to symmetry and further generalizations

As we have illustrated any of the introduced methods can provide a set of essential variables, and it depends on the specific problem which method to take. However we have concealed that all of the methods require some sort of preprocessing, before they can be applied to the raw Cartesian data. The reason is that in the absence of external forces, the intramolecular force field is translationally and rotationally invariant; therefore also the equations of motion are equivariant with respect to the Euclidean group $SE(3) \cong \mathbf{R}^3 \times SO(3)$ which consists of translations and proper rotations in $\mathbf{R}^3$ (rigid body symmetry). But since the rigid body symmetry carries no interesting information for the conformation analysis, the symmetry group is factored out before the data analysis starts; otherwise the motion that is associated with the symmetry deteriorates the statistical analysis, in particular the covariance analysis.

Typically the symmetry reduction is done by means of a molecular alignment of the simulation data, also sometimes called *fitting*, which removes the overall translations and rotations from the Cartesian data. This has several drawbacks: First of all, molecular alignment algorithms operate with respect to an arbitrarily chosen reference configuration of the molecule, such that all rotations and translations are carried out such as to minimize some distance to the reference state [160]. Often only a

specific group of atoms is chosen as reference, with respect to which the data is aligned, and, as the reader can imagine, the subspace identification depends on the choice of the reference configuration. Secondly and even worse, the preprocessing destroys the correspondence between data and equations of motion, for the symmetry reduction is performed on the data indeed, but the equations of motion are left untouched. This inconsistency has the effect that, e.g., the dichotomy *free energy as the potential of mean force* for a reaction coordinate that is not invariant under translations and rotation is annihilated. As a consequence the free energy as computed from the distribution of the symmetry-reduced data is different from the free energy that is computed by means of Thermodynamic Integration.

Therefore we want to apply the techniques to general non-Euclidean data which could, e.g., be obtained from nonlinear transformations of the original data. On the one hand these transformations can be chosen such as to respect the symmetry, like the transformation to internal coordinates. On the other hand this requires that we know how to compute the covariance matrix in a meaningful way, say, for data lying on a torus. We shall explain this idea in more detail: Consider a set of curvilinear coordinates that are represented by the following map

$$\varphi : \mathbf{R}^n \to M \subset \mathbf{R}^m \quad \text{with} \quad \varphi(t) := \varphi(x(t)) \,,$$

where in most cases $m \leq n$, but also the reverse case can be dealt with [161]. Given a series of observation data $x(t)$, we define a centered data set by simply subtracting the mean $\varphi(t) \mapsto \varphi(t) - \bar{\varphi}$, where the term *mean* has to be specified in a way that fits the problem; see [162] and the references therein. It seems appealing to decompose M directly. However neither do we believe that there is a straightforward generalization of the above mentioned concepts to manifolds nor do we assume that such a procedure would lead to computationally tractable problems. Instead we favour the following approach going back to considerations from nonlinear elasticity [163]: the idea is to embed M into a linear space V of higher dimension, and apply a linear subspace decomposition to V; once a subspace $U \subset V$ is chosen, we can construct the approximant $S \subset M$ as the intersection $M \cap U$. The last step is to be understood as follows: Let v be a local coordinate map on V, and let $\mathscr{P}v \in U$, where $\mathscr{P}$ is the usual projection from V to $U \subset V$. Moreover let the map $\tau : M \to V$ denote the embedding $M \subset V$. Then $\tau^{-1} \circ \mathscr{P} \circ \tau$ maps $\varphi \in M$ to points on the approximant.

The particular embedding τ is open to choice, and should be motivated by the physics of the problem. For an example using a polar decomposition of torsion space, see [164]; this particular embedding is quite problematic, for all points lie on a hypersphere S^{2m-1}, where m is the number of torsion angles. There is yet another difficulty: Notice that the approximant is a regular submanifold of codimension s in M; but then, considered as a submanifold of $\mathbf{R}^n$ it has codimension s, too. Hence the approximant has dimension $n - s$ which is still high-dimensional, since mostly $m \ll n$.

It is a good idea to change the view slightly: consider the non-approximating subspaces which are defined as the fibres over the essential variables or reaction coordinates, and which are given by a family of configurational submanifolds

$$\Sigma_\xi = \{x \in \mathbf{R}^n \,|\, \Phi : \mathbf{R}^n \to \mathbf{R}^{m-s}, \, \Phi(x) = \xi\} \,, \tag{2.32}$$

with the reaction coordinate

$$\Phi : \mathbf{R}^n \to \mathbf{R}^{m-s}, \quad \Phi = \chi \circ \tau^{-1} \circ \mathscr{P} \circ \tau \circ \varphi \,,$$

where $\chi : \mathbf{R}^m \to \mathbf{R}^{m-s}$ is a local coordinate map on S. By construction, $\mathbf{D}\Phi$ has rank $m - s$ which is equivalent to stating that the fibres $\Phi^{-1}(\xi)$ are smooth submanifolds

of $\mathbf{R}^n$, and so $\Phi(x) = \xi$ defines a foliation of the Euclidean configuration space. The reaction coordinates can then be thought of as *spanning* the essential configuration space [42]. Nevertheless it is important to bear in mind that the object of interest is the approximating subspace rather than the respective reaction coordinates. Moreover it is no longer the case that Σ_ξ is the orthogonal complement of S in $\mathbf{R}^n$ as was true in the linear scenario. However proceeding this way has the advantage that a clever choice of both φ and τ can lead to relatively low-dimensional essential variables, e.g., linear combinations of a few torsion angles or radii of gyration (cf. Figure 2).