LANGEVIN DYNAMICS WITH CONSTRAINTS AND COMPUTATION OF FREE ENERGY DIFFERENCES

TONY LELIÈVRE, MATHIAS ROUSSET, AND GABRIEL STOLTZ

Abstract. In this paper, we consider Langevin processes with mechanical constraints. The latter are a fundamental tool in molecular dynamics simulation for sampling purposes and for the computation of free energy differences. The results of this paper can be divided into three parts. (i) We propose a simple discretization of the constrained Langevin process based on a standard splitting strategy. We show how to correct the scheme so that it samples exactly the canonical measure restricted on a submanifold, using a Metropolis rule in the spirit of the Generalized Hybrid Monte Carlo (GHMC) algorithm. Moreover, we obtain, in some limiting regime, a consistent discretization of the overdamped Langevin (Brownian) dynamics on a submanifold, also sampling exactly the correct canonical measure with constraints. (ii) For free energy computation using thermodynamic integration, we rigorously prove that the longtime average of the Lagrange multipliers of the constrained Langevin dynamics yields the gradient of a rigid version of the free energy associated with the constraints. A second order time discretization using the Lagrange multipliers is proposed. (iii) The Jarzynski-Crooks fluctuation relation is proved for Langevin processes with mechanical constraints evolving in time. An original numerical discretization without time discretization error is proposed. Numerical illustrations are provided for (ii) and (iii).

1. Introduction and main results

In this section, we present the main results and give the outline of the paper. We define only briefly the concepts we need in this general introduction, and refer the reader to the following sections (in particular Section 2) for more precisions on the mathematical objects at hand.

We consider mechanical systems with constraints. The configuration of a classical $N$-body system is denoted by $(q, p) \in \mathbb{R}^{6N}$. The results of the paper can be generalized mutatis mutandis to periodic boundary conditions ($q \in \mathbb{T}^{3N}$), or to systems with positions confined in a domain $q \in \mathcal{D} \subset \mathbb{R}^{3N}$. The mass matrix of the system is a constant strictly positive symmetric matrix $M$, usually of the form $M = \text{Diag}(m_1 \text{Id}_3, \ldots, m_N \text{Id}_3) \in \mathbb{R}^{3N \times 3N}$. The interaction potential is a smooth function $V : \mathbb{R}^{3N} \to \mathbb{R}$. The Hamiltonian of the system is assumed to be separable:

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

Positions are subject to a $m$-dimensional mechanical constraint denoted by

$$\xi(q) = (\xi_1(q), \ldots, \xi_m(q))^T = z \in \mathbb{R}^m.$$

Constrained systems appear in computational statistical physics in two kinds of contexts (see e.g. Chapter 10 in [31], [9] and [28] for applications to the computation of free energy differences, and [3] [24] for mathematical textbooks dealing with constrained Hamiltonian dynamics):

(i) for free energy computations, where $\xi$ is a given reaction coordinate parameterizing a transition between "states" of interest;

(ii) when the system is subject to molecular constraints such as rigid covalent bonds, or rigid bond angles in molecular systems.

Date: June 23, 2010.

2000 Mathematics Subject Classification. Primary 82B80, 65C30; Secondary 82B35.

Key words and phrases. Constrained stochastic differential equations, free energy computations, nonequilibrium dynamics.

This work is supported by the Agence Nationale de la Recherche, under the grant ANR-09-BLAN-0216-01 (MEGAS), and by the ARC Hybrid from INRIA.
In any case, the position of the system is constrained onto the submanifold of co-dimension \( m \):
\[
\Sigma(z) = \left\{ q \in \mathbb{R}^{3N} \mid \xi(q) = z \right\},
\]
and the associated phase space is the cotangent bundle denoted by
\[
T^*\Sigma(z) = \left\{ (q, p) \in \mathbb{R}^{6N} \mid q \in \Sigma(z), \nabla \xi(q)^T M^{-1} p = 0 \right\}.
\]
For a given \( q \in \Sigma(z) \), the set of cotangent momenta is denoted by
\[
T^*_q \Sigma(z) = \left\{ p \in \mathbb{R}^{3N} \mid \nabla \xi(q)^T M^{-1} p = 0 \right\}.
\]
The orthogonal projection on \( T^*_q \Sigma(z) \) with respect to the scalar product induced by \( M^{-1} \) is denoted
\[
P_M(q) = \text{Id} - \nabla \xi(q) G_M^{-1}(q) \nabla \xi(q)^T M^{-1},
\]
where \( G_M(q) \) is the Gram matrix associated with the constraints
\[
G_M(q) = \nabla \xi(q)^T M^{-1} \nabla \xi(q).
\]
Throughout the paper, we assume that \( G_M \) is invertible everywhere on \( \Sigma(z) \) (for all \( z \)). It is easily checked that \( P_M \) satisfies the projector property \( P_M(q)^2 = P_M(q) \), and the orthogonality property
\[
M^{-1} P_M(q) = P_M(q)^T M^{-1}.
\]
In the present paper, the focus is on the canonical ensemble, which is the equilibrium probability distribution of microscopic states of a system at fixed temperature (fixed average energy). For systems without constraints, this ensemble is characterized by the probability distribution
\[
\mu(dq dp) = Z^{-1} e^{-\beta H(q,p)} dq dp,
\]
where \( Z \) is the normalizing constant ensuring that \( \mu \) is indeed a probability distribution. For constrained systems, the associated canonical distribution is defined by
\[
\mu_{T^* \Sigma(z)}(dq dp) = Z^{-1}_{z,0} e^{-\beta H(q,p)} \sigma_{T^* \Sigma(z)}(dq dp),
\]
where \( \sigma_{T^* \Sigma(z)}(dq dp) \) is the phase space Liouville measure of \( T^* \Sigma(z) \), and \( Z_{z,0} \) the normalizing constant (\( z \) refers to the position constraint, and 0 to the velocity or momentum constraint, see \((2.15)\) below). See Section \((2.3)\) for precise definitions. A dynamics admitting this measure as an invariant equilibrium measure is the following Langevin process (for a given initial condition \((q_0, p_0) \in T^* \Sigma(z)\)):
\[
\begin{aligned}
dq_t &= M^{-1} p_t dt, \\
dp_t &= -\nabla V(q_t) dt - \gamma(q_t) M^{-1} p_t dt + \sigma(q_t) dW_t + \nabla \xi(qt) d\lambda_t, \\
\xi(q_t) &= z,
\end{aligned}
\]
where the \( \mathbb{R}^m \)-valued adapted process \( t \mapsto \lambda_t \) is the Lagrange multiplier associated with the (vectorial) constraint \( \langle C_q \rangle \), \( W_t \) a standard \( 3N \)-dimensional Brownian motion, and \( \gamma(q), \sigma(q) \) are \( 3N \times 3N \) real matrices. Note that \((q_t, p_t) \in T^* \Sigma(z) \) for all \( t \geq 0 \). The standard fluctuation-dissipation identity
\[
\sigma(q) \sigma^T(q) = \frac{2}{\beta} \gamma(q),
\]
should be imposed in order for the canonical distribution \((1.7)\) to be invariant under the dynamics \((1.8)\). Then, averages of an observable \( A : \mathbb{R}^{6N} \rightarrow \mathbb{R} \) with respect to the distribution \((1.7)\) can be obtained as longtime averages along any trajectory of the dynamics \((1.8)\) (when \( P_M \gamma P_M^T \) is symmetric positive on \( \Sigma(z) \)):
\[
\lim_{T \to +\infty} \frac{1}{T} \int_0^T A(q_t, p_t) dt = \int_{T^* \Sigma(z)} A \, d\mu_{T^* \Sigma(z)} \quad \text{a.s.}
\]
1\text{i.e.} a random variable depending only on the past values of the Brownian motion.
This is detailed in Section 3. Several recent studies (e.g. \cite{18, 19, 35, 6}) have analyzed dynamics similar to (1.8) and some appropriate discretization of the process in order to approximate the left-hand side of (1.10). In this paper, a classical discretization of the dynamics (1.8) using a splitting strategy between the Hamiltonian and the thermostat part is considered (see also \cite{5} in the unconstrained case). The originality lies in the two following remarks: (i) for some choice of the parameters, an Euler discretization of the overdamped Langevin (Brownian) dynamics with a projection step associated with the constraints is obtained (see Equation (3.23) and Proposition 3.6); (ii) it can be completed by a Metropolis rule to obtain a Generalized Hybrid Monte Carlo (GHMC) method sampling exactly (i.e. without any bias) the canonical distribution on the submanifold (1.7) (Algorithm 3.3). The so-obtained numerical scheme is close to the ones proposed in \cite{19, 18, 17}. See also \cite{11, 29} for historic references on Hybrid Monte Carlo methods, and \cite{21} for GHMC.

The free energy $F : \mathbb{R}^m \to \mathbb{R}$ associated with the reaction coordinates $\xi$ can be identified with the opposite of the log-density of the marginal probability distribution of the reaction coordinates $\xi$ under the canonical distribution (1.4). Explicitly, it is defined through the following relation for any test function $\phi : \mathbb{R}^m \to \mathbb{R}$:

$$
\int_{\mathbb{R}^m} \phi(z) e^{-\beta F(z)} dz = \int_{\mathbb{R}^N} \phi(\xi(q)) \mu(dq dp).
$$

In simulations, computing the free energy profile $z \mapsto F(z)$ (up to an additive constant independent of $z$), or free energy differences between two states $F(z_2) - F(z_1)$, is very important, and yields the relative probability of different "states" parameterized by $\xi$. A state should be understood here as the collection of all possible microscopic configurations $(q, p)$, distributed according to the canonical measure (1.6), and satisfying the macroscopic constraint $\xi(q) = z$. Since we only focus on computing free energy differences, $F$ is defined up to an additive constant (independent of $z$, denoted by $C$ below, and whose value may vary from line to line) and can be rewritten as:

$$
F(z) = -\frac{1}{\beta} \ln \int_{\Sigma(z) \times \mathbb{R}^N} e^{-\beta H(q,p)} \delta_{\xi(q) \rightarrow z}(dq dp)
$$

$$
= -\frac{1}{\beta} \ln \int_{\Sigma(z)} e^{-\beta V(q)} \delta_{\xi(q) \rightarrow z}(dq) + C,
$$

where $\delta_{\xi(q) \rightarrow z}$ denotes the conditional measure on $\Sigma(z)$ verifying the following identity of measures in $\mathbb{R}^{3N}$: $dq = \delta_{\xi(q) \rightarrow z}(dq) dz$ (see Section 2.3 for more precisions on this relation).

However, when using constrained simulations in phase space, the momentum variable of the dynamical system is also constrained, and a modified free energy (called "rigid free energy" in the sequel) is more naturally computed. This is detailed in Section 3. The latter is defined as

$$
F^{rgd}_M(z) = -\frac{1}{\beta} \ln \int_{T^*\Sigma(z)} e^{-\beta H(q,p)} \sigma_{T^*\Sigma(z)}(dq dp).
$$

The superscript $M$ indicates that this free energy depends on the considered mass matrix, even though this is not clear at this stage (see (1.3) below). The above two definitions of free energy are related through the identity:

$$
F(z) - F^{rgd}_M(z) = -\frac{1}{\beta} \ln \int_{T^*\Sigma(z)} (\det G_M)^{-1/2} d\mu_{T^*\Sigma(z)} + C,
$$

where $\mu_{T^*\Sigma(z)}$ is the equilibrium distribution with constraints (1.7). The relation (1.14), already proposed in \cite{9}, is proved in the beginning of Section 1. For any value of the reaction coordinate, the difference $F(z) - F^{rgd}_M(z)$ can then be easily computed with any method sampling the probability distribution $\mu_{T^*\Sigma(z)}$, such as (1.8). Several methods have been suggested in the literature to compute either $F$ or $F^{rgd}_M$ from the Lagrange multipliers of a constrained process similar to (1.8). We refer for instance to \cite{9} (and references therein) for the Hamiltonian case, and to \cite{6} (and references therein) for the overdamped case. It is a second purpose of the present paper to prove that the longtime average of the Lagrange multipliers in (1.8) converges to the gradient of the
rigid free energy (1.13) (the so-called mean force), in the sense that:

(1.15) \[ \lim_{T \to +\infty} \frac{1}{T} \int_0^T d\lambda_t = \nabla z F_{\text{rigid}}^M(z) \quad \text{a.s.} \]

As compared to [3], where a formal proof for the Hamiltonian case is proposed, we use an explicit calculation that does not require the use of the Lagrangian structure of the problem, or a change of coordinates. Using (1.15) with (1.14), thermodynamic integration can then be performed without computing second order derivatives of \( \xi \) which appear in the analytical expression of the mean force, see for instance (4.11)-(4.12). We propose a discretization of the Langevin processes (1.8) and of the formula (1.13) which allows to compute the free energy gradients with second order accuracy (i.e. up to \( O(\Delta t^2) \) error terms, where \( \Delta t \) is the time-step).

Finally, we study nonequilibrium methods for free energy computations using a Hamiltonian or Langevin dynamics with constraints subject to time evolution. Such methods rely on a nonequilibrium fluctuation equality, the so-called Jarzynski-Crooks relation. See [22] for a pioneering work, as well as [7] [8] for an extension. They are termed “nonequilibrium” since the transition from one value of the reaction coordinate \( \xi \) to another one is imposed a priori, in a finite time \( T \), and with a given smooth deterministic schedule \( t \in [0, T] \mapsto z(t) \in \mathbb{R}^m \). In particular, it may be arbitrarily fast. Therefore, even if the system starts at equilibrium, it does not remain at equilibrium. The out-of-equilibrium Langevin process is given by the following equations of motion:

(1.16) \[
\begin{align*}
dq_t &= -\nabla V(q_t) dt - \gamma_P(q_t) M^{-1} p_t dt + \sigma_P(q_t) dW_t + \nabla \xi(q_t) d\lambda_t, \\
\xi(q_t) &= z(t), \\
\sigma(q_t, p_t) &= \nabla \xi(q_t)^T M^{-1} p_t = \dot{z}(t).
\end{align*}
\]

where \( t \mapsto \lambda_t \in \mathbb{R}^m \) is an adapted process enforcing the constraints \( \{C_q(t)\} \) (the Lagrange multipliers). Initial conditions are sampled from the canonical distribution consistent with the constraints \( \xi(q) = z(0) \) and \( \sigma(q, p) = \nabla \xi(q)^T M^{-1} p = \dot{z}(0) \). We assume here that the fluctuation-dissipation matrices are of the form

(1.17) \[ (\sigma_P, \gamma_P) := (P_M \sigma, P_M \gamma P_M^T), \]

where \( \gamma, \sigma \in \mathbb{R}^{3N \times 3N} \) satisfy the fluctuation-dissipation identity (1.9), so that \( \gamma_P, \sigma_P \) also verify (1.9). The Hamiltonian case is included in this general framework upon choosing \( \gamma = 0 \). The dynamics (1.16) is a natural extension of the equilibrium Langevin dynamics. It is different from the dynamics proposed in [24], which is a Langevin dynamics associated with a modified Hamiltonian with projected momenta, driven by a forcing term along \( \nabla \xi \) which acts directly on the position variable. The specific choice (1.17) (rather than the general case \( \gamma, \sigma \in \mathbb{R}^{3N \times 3N} \)) leads to a simpler analysis and more natural numerical schemes (based again on a splitting procedure).

As explained in Section 5.3, it is possible to define the work exerted on the system between time 0 and \( T \) as the displacement multiplied by the constraining force:

(1.18) \[ W_{0,T} \left\{ \{q_s, p_s\}_{0 \leq s \leq T} \right\} := \int_0^T \dot{z}^T(s) d\lambda_s. \]

It is the third and last purpose of the present paper to derive the general Crooks-Jarzynski relation (see (1.20) below, and Theorem 5.3 for the general statement) associated with the nonequilibrium dynamics (1.16) and the work defined in (1.18). The main corollary is given by the following result. Consider the corrector

(1.19) \[ G(t, q) = \frac{1}{2\beta} \ln \left( \det G_M(q) \right) - \frac{1}{2} \dot{z}(t)^T G_M^{-1}(q) \dot{z}(t), \]

where \( \frac{1}{2\beta} \ln \det G_M(q) \) is the Fixman entropic term due to the geometry of the position constraints (see (1.14) and Remark 5.3), and \( \frac{1}{2} \dot{z}(t)^T G_M^{-1}(q) \dot{z}(t) \) is the kinetic energy term due to the velocity of the switching. Then, the free energy profile can be computed through the following fluctuation...
identity (see (5.23)):

\[ F(z(T)) - F(z(0)) = \frac{1}{\beta} \ln \left( \frac{\mathbb{E} \left( e^{-\beta W_{0, T} \left( \{q_t, p_t\}_{0 \leq t \leq T} \right) + C(T, q_T) \right)} }{\mathbb{E} \left( e^{-\beta C(0, q_0)} \right)} \right) , \]

where the expectation is with respect to canonical (equilibrium) initial conditions and for all realizations of the dynamics (1.16). An original numerical scheme is proposed, which allows to compute free energy differences according to (1.16)-(1.20) without time discretization error (see Theorem 5.5). Moreover, for some choice of the parameters, the latter scheme yields a Jarzynski-Crooks relation for an Euler discretization of the overdamped Langevin (Brownian) dynamics with a projection step associated with the evolving constraints without time discretization error. This can be seen as an extension of the scheme formerly proposed in [27] (see Equation (5.53) and Proposition 5.6). We also check the consistency of the various free energy estimators we introduce.

This paper is organized as follows. We start with an introduction to the mathematical concepts required for mechanically constrained systems in Section 2. Section 3 presents the case of mechanically constrained Langevin processes defined by (1.8), and the problem of sampling the canonical distribution (1.7). Thermodynamic integration with constrained Langevin processes is detailed in Section 4. Section 5 discusses nonequilibrium constrained Langevin processes (1.16) and the associated Jarzynski-Crooks fluctuation identity (1.20). Finally, some technical lemmas are gathered in Section 6.

2. Preliminaries

After making precise our notation for matrices and matrix valued functions in Section 2.1 we introduce some additional concepts required to describe constrained systems in Section 2.2 and define the phase space measures with constraints in Section 2.3.

2.1. Notation. Throughout the paper, the following notation is used:

- Vectors and vector fields are by convention of column type. When vectors are written as a line, they should be understood as the corresponding column version. For instance, \((q, p) \in \mathbb{R}^{6N}\) should be understood as \((q^T, p^T)^T\), where \(q, p \in \mathbb{R}^{3N}\) are both column vectors.
- Gradients in \(\mathbb{R}^{3N}\) (or \(\mathbb{R}^{6N}\)) of \(m\)-dimensional vector fields are by convention \(3N \times m\)-matrices, for instance:

\[ \nabla \xi(q) = \begin{pmatrix} \nabla \xi_1(q) & \cdots & \nabla \xi_m(q) \end{pmatrix} \in \mathbb{R}^{3N \times m}, \]

where \(\nabla \xi_i(q) \in \mathbb{R}^{3N}\) is a column vector for any \(i = 1, \ldots, m\). Gradients in the space of constraints parameters \(z \in \mathbb{R}^m\) or \(\zeta \in \mathbb{R}^{2m}\) are denoted with the associated subscripts, namely \(\nabla_z\) and \(\nabla_\zeta\).
- Second order derivatives in \(\mathbb{R}^{3N}\) of \(m\)-dimensional vector fields are characterized through the Hessian bilinear form:

\[ \text{Hess}_q(\xi)(v_1, v_2) = \begin{pmatrix} v_1^T \nabla^2 \xi_1(q)v_2 \\ \vdots \\ v_1^T \nabla^2 \xi_m(q)v_2 \end{pmatrix} \in \mathbb{R}^m, \]

where \(v_1, v_2 \in \mathbb{R}^{3N}\) are test vectors.
- The canonical symplectic matrix is denoted by:

\[ J := \begin{pmatrix} 0 & \Id_{3N} \\ -\Id_{3N} & 0 \end{pmatrix} \in \mathbb{R}^{6N \times 6N}. \]

For any smooth test functions \(\varphi_1 : \mathbb{R}^{6N} \to \mathbb{R}^{n_1}\) and \(\varphi_2 : \mathbb{R}^{6N} \to \mathbb{R}^{n_2}\), the Poisson bracket is the \(n_1 \times n_2\) matrix

\[ \{\varphi_1, \varphi_2\} = (\nabla \varphi_1)^T J \nabla \varphi_2 \in \mathbb{R}^{n_1 \times n_2}. \]
- For two matrices \(A, B \in \mathbb{R}^{n \times n}\), \(A : B = \text{Tr}(A^T B)\).
2.2. Constraints. Two useful concepts to study constrained Hamiltonian systems (in particular, to use the co-area formula in phase space, as well as the Poisson bracket formulation of the Liouville equation) are the effective velocity $v_\xi$ and the effective momentum $p_\xi$ associated with the constrained degrees of freedom $\xi$:

$$v_\xi(q,p) = \nabla \xi(q)^T M^{-1} p \in \mathbb{R}^m,$$

and

$$p_\xi(q,p) = G_M^{-1}(q) v_\xi(q,p) = G_M^{-1}(q) \nabla \xi(q)^T M^{-1} p \in \mathbb{R}^m.$$

The expression of the effective velocity is obtained by deriving the constraint $\xi$ along an unconstrained trajectory of the Hamiltonian dynamics

$$
\begin{align*}
\frac{d\hat{q}_t}{dt} &= M^{-1} \hat{p}_t, \\
\frac{d\hat{p}_t}{dt} &= -\nabla V(\hat{q}_t),
\end{align*}
$$

since $\frac{d\xi(\hat{q}_t)}{dt} = v_\xi(\hat{q}_t, \hat{p}_t)$. The term $G_M^{-1}(q)$ in the expression of the effective momentum may be interpreted as the effective mass of $\xi$. This can be motivated by a decomposition of the kinetic energy of the system into a tangential and an orthogonal part. The latter reads for $(q,p) \in \mathbb{R}^{6N}$:

$$E_{\text{kin}}^T(q,p) := \frac{1}{2} p^T \left( \text{Id} - P_M(q) \right)^T M^{-1} \left( \text{Id} - P_M(q) \right) p = \frac{1}{2} v_\xi(q,p)^T G_M^{-1}(q) v_\xi(q,p) = \frac{1}{2} p_\xi(q,p)^T G_M(q) p_\xi(q,p).$$

The constraints on a mechanical system can also be rewritten in the more general form

$$\Xi(q,p) = \zeta \in \mathbb{R}^{2m},$$

where either (i) the effective momentum is constrained, in which case $\Xi = (\xi, p_\xi)$ and $\zeta = (z, p_z)$; or (ii) the effective velocity is constrained, in which case $\Xi = (\xi, v_\xi)$ and $\zeta = (z, v_z)$. The phase space associated with such constraints is denoted by

$$\Sigma_\Xi(\zeta) = \left\{ (q,p) \in \mathbb{R}^{6N} \mid \Xi(q,p) = \zeta \right\}.$$

A position $q \in \Sigma(z)$ being given, the affine space of constrained momenta verifying (2.6) is then denoted by

$$\Sigma_{v_\xi(q,\cdot)}(v_z) = \left\{ p \in \mathbb{R}^{3N} \mid v_\xi(q,p) = v_z \right\}$$

in the effective velocity case, and by $\Sigma_{p_\xi(q,\cdot)}(p_z)$ in the effective momentum case. This notation is very important for nonequilibrium methods where the constraints evolve in time according to a predefined schedule, see Section 5. Note that the phase space of mechanical constraints, defined by (2.2), is simply $T^* \Sigma(z) = \Sigma_{\xi,v_\xi}(z,0) = \Sigma_{\xi,p_\xi}(z,0)$.

We can now define the skew-symmetric Gram tensor of dimension $2m \times 2m$ associated with the constraints:

$$\Gamma(q,p) = (\Xi, \Xi)(q,p) = \nabla \Xi^T(q,p) J \nabla \Xi(q,p) \in \mathbb{R}^{2m \times 2m}.$$

The Gram matrix $\Gamma$ associated with the generalized constraints (2.6) can be explicitly computed by block. Indeed, for $\Xi = (\xi, p_\xi)^T$.

$$\Gamma = \begin{pmatrix} 0 & \text{Id} \\ -\text{Id} & \nabla p_\xi^T J \nabla p_\xi \end{pmatrix}.$$

Therefore, $\det(\Gamma) = 1$ in this case. In the case $\Xi = (\xi, v_\xi)^T$, the Gram matrix reads

$$\Gamma = \begin{pmatrix} 0 & G_M^T \\ -G_M & \nabla v_\xi^T J \nabla v_\xi \end{pmatrix},$$
and \( \det(\Gamma) = \det(M_G)^2 \). Note that in both cases \( \det(\Gamma) > 0 \). The constrained symplectic (skew-
symmetric) matrix is now defined by

\[
J_z(q, p) = J - J \nabla \Xi(q, p) \Gamma^{-1}(q, p) \nabla \Xi^T(q, p) J,
\]

and the Poisson bracket associated with generalized constraints (2.6) by:

\[
\{ \varphi_1, \varphi_2 \}_\Xi = \nabla \varphi_1^T J_z \nabla \varphi_2.
\]

It is easily checked that \( \{ \cdot, \cdot \}_\Xi \) verifies the characteristic properties of Poisson brackets, namely the skew-
symmetry, Jacobi’s identity, and Leibniz’ rule. Therefore, the flow associated with the evolution equation

\[
\frac{d}{dt} \begin{pmatrix} q_t \\ p_t \end{pmatrix} = J_z \nabla H(q_t, p_t),
\]

which is (1.8) in the case \((\gamma, \sigma) = (0, 0)\), defines a symplectic map. Recall that (see [16] Section VII.1.2) a map \( \phi : \Sigma(\zeta) \to \Sigma(\zeta) \) is symplectic if for any \((q, p) \in \Sigma(\zeta) \) and \( u, v \in T(q, p) \Sigma(\zeta) \),

\[
u^T \nabla \phi(q, p)^T J \nabla \phi(q, p)v = u^T J v.
\]

A consequence of the symplectic structure is the divergence formula (2.25) relating the phase space

\[
\text{evolution equation}
\]

conditioning formula: for any test function \( \phi : \mathbb{R}^{6N} \to \mathbb{R} \),

\[
\int_{\mathbb{R}^{6N}} \phi(q, p) \, dq \, dp = \int_{\mathbb{R}^{2m}} \int_{\Sigma(\zeta)} \phi(q, p) \delta_{\Xi(q, p) - \zeta} \, dq \, dp \, d\zeta.
\]

2.3.1. Definitions. The phase space measure (also termed Liouville measure) on the phase space

\[
T^*\Sigma(z) \quad \text{(or more generally on } \Sigma(\zeta) \text{)} \quad \text{of constrained mechanical systems is denoted by } \sigma_{T^*\Sigma(z)} \quad \text{(or more generally } \sigma_{\Sigma(\zeta)} \text{).}
\]

The latter is induced by the symplectic, or skew-symmetric 2-form on \( \mathbb{R}^{6N} \) defined by the canonical skew-symmetric matrix \( J \) in \( \mathbb{R}^{6N} \). More precisely, it can be defined through the volume form \(|\det \mathcal{G}(u(q, p))|^{1/2}\), where

\[
\mathcal{G}_{a, b}(u) = (u_a)^T J u_b, \quad a, b = 1, \ldots, 6N - 2m,
\]

and \((u_1(q, p), \ldots, u_{6N-2m}(q, p))\) is a basis of tangential vectors of the submanifold \( T^*\Sigma(z) \) (or \( \Sigma(\zeta) \)) at a given point \((q, p)\).

Surface measures induced by an Euclidean scalar product will also be of interest. We denote by \( \sigma_{\Sigma(z)}^M(dq) \) the surface measure on \( \Sigma(z) \) induced by the scalar product \( \langle q, \tilde{q} \rangle_M = q^T M \tilde{q} \) on \( \mathbb{R}^{4N} \), and, for a given \( q \in \Sigma(z) \), by \( \sigma_{\Sigma_{\xi, \nu}(q), (p)}^{M^{-1}}(dp) \) and \( \sigma_{\Sigma_{\nu}(q)}^{M^{-1}}(dp) \) the surface measures on the affine spaces \( \Sigma_{\xi, \nu}(q) \) and \( \Sigma_{\nu}(q) \) respectively, induced by the scalar product \( \langle p, \tilde{p} \rangle_{M^{-1}} = p^T M^{-1} \tilde{p} \) on \( \mathbb{R}^{4N} \). For a precise definition of these measures, we refer to [23] Sections 3.2.1 and 3.3.2] and the references therein.

It is now possible to define a generalized canonical distribution (1.7) as follows:

\[
\begin{cases}
\mu_{\Sigma_{\zeta}, \nu}(z, v_z)(dq, dp) := \frac{e^{-\beta \mathcal{L}(q, p)}}{Z_{z, v_z}} \sigma_{\Sigma_{\zeta}, \nu}(z, v_z)(dq, dp), \\
Z_{z, v_z} := \int_{\Sigma_{\zeta, \nu}(z, v_z)} e^{-\beta \mathcal{L}} \, d\sigma_{\Sigma_{\zeta}, \nu}(z, v_z).
\end{cases}
\]

The distribution (2.15) is associated with the generalized constraints \( \Xi = (\xi, \nu)^T \), and is used in Section 5 for nonequilibrium methods. Note that \( \mu_{\Sigma_{\zeta, \nu}(z, 0)} = \mu_{T^*\Sigma(z)} \) defined in (1.7).

2.3.2. Co-area decompositions. The co-area formula (see [2][14]) relates the phase space or surface

measures, and the conditional measures. Conditional measures are defined in \( \mathbb{R}^{6N} \) by the following conditioning formula: for any test function \( \phi : \mathbb{R}^{6N} \to \mathbb{R} \),

\[
\int_{\mathbb{R}^{6N}} \phi(q, p) \, dq \, dp = 
\int_{\mathbb{R}^{2m}} \int_{\Sigma(\zeta)} \phi(q, p) \delta_{\Xi(q, p) - \zeta} \, dq \, dp \, d\zeta.
\]
In the same way in $\mathbb{R}^{3N}$, conditional measures are defined, for any test function $\phi : \mathbb{R}^{3N} \to \mathbb{R}$, by

$$\int_{\mathbb{R}^{3N}} \phi(q) \, dq = \int_{\mathbb{R}^n} \int_{\Sigma(z)} \phi(q) \, \delta_{\xi(q) - z}(dq) \, dz.$$ 

A more concise notation for the above equalities is $dq \, dp = \delta_{\Xi(q,p) - z}(dq \, dp) \, d\zeta$ and $dq = \delta_{\xi(q) - z}(dq) \, dz$.

**Proposition 2.1** (Co-area). Let $\Sigma(z)$ be the submanifold defined by the constraints $\xi(q) = z$, and assume that $G_M$ defined in (2.13) is non-degenerate in a neighborhood of $\Sigma(z)$. Then, in the sense of measures on $\mathbb{R}^{3N}$:

$$\delta_{\xi(q) - z}(dq) = (\det M)^{-1/2} | \det G_M(q) |^{-1/2} \sigma_M^{\Sigma(z)}(dq).$$

Let $\Sigma(\zeta)$ be the phase space defined by generalized constraints (2.9). Assume that $\Gamma$ defined in (2.9) is non-degenerate in a neighborhood of $\Sigma(\zeta)$. Then, in the sense of measures on $\mathbb{R}^{6N}$:

$$\delta_{\Xi(q,p) - \zeta}(dq \, dp) = | \det (\Gamma(q,p)) |^{-1/2} \sigma_{\Sigma(\zeta)}(dq \, dp).$$

We refer for example to Chapter 3 in [29] for an elementary proof. An equivalent of (2.18)-(2.17) for momenta reads, for constrained effective momenta:

$$dp = \delta_{\nu(q,p) -\nu_z}(dp) \, dp_z = \det(M) \, \frac{1}{2} | \det G_M(q) | \sigma_{\Sigma_{\nu(q,p),\nu_z}}^{-1}(dp) \, dp_z,$$

and for constrained effective velocities:

$$dp = \delta_{\nu(q,p) -\nu_z}(dp) \, dv_z = \det(M) \, \frac{1}{2} | \det G_M(q) | \sigma_{\Sigma_{\nu(q,p),\nu_z}}^{-1}(dp) \, dv_z.$$

Using the co-area formulas (2.18)-(2.19), and the expressions of symplectic Gram matrices (2.10)-(2.11), we obtain the following expressions of the phase space measures:

(i) The phase space measure on $\Sigma_{\zeta,\nu_z}(z,\nu_z)$ can be identified with the conditional measure defined in (2.10):

$$\sigma_{\Sigma_{\zeta,\nu_z}(z,\nu_z)}(dq \, dp) = \delta_{\xi(q) - z,\nu_z}(dq \, dp).$$

while the phase space measure on $\Sigma_{\zeta,v_z}(z,\nu_z)$ is related to the corresponding conditional measure as

$$\sigma_{\Sigma_{\zeta,v_z}(z,\nu_z)}(dq \, dp) = \det(M) \, \delta_{\xi(q) - z,\nu_z}(dq \, dp).$$

(ii) The phase space measures are given by the product of surface measures:

$$\sigma_{\Sigma_{\zeta,\nu_z}(z,\nu_z)}(dq \, dp) = \sigma_{\Sigma_{\nu_z}(q,p)}^{-1}(dp) \sigma_{\Sigma(z)}(dq),$$

and

$$\sigma_{\Sigma_{\zeta,v_z}(z,\nu_z)}(dq \, dp) = \sigma_{\Sigma_{\nu_z}(q,p)}^{-1}(dp) \sigma_{\Sigma(z)}(dq).$$

Equations (2.22)-(2.24) are a consequence of the fact that

$$\delta_{\xi(q) - z,\nu_z}(dq \, dp) = \delta_{\nu_z}(dp) \sigma_{\Sigma(z)}(dq).$$

and a similar relation for $v_{\xi}$.}

2.3.3. **Divergence formulas.** We end this section with an important formula, which is used to show the invariance of the canonical measure in the proof of Proposition 3.2.

**Proposition 2.2** (Divergence theorem in phase space). Consider the Poisson bracket $\{\cdot, \cdot\}_{\Xi}$ defined by (2.13), and an open neighborhood $\mathcal{O}$ of $\Sigma(\zeta) \subset \mathbb{R}^{6N}$ where $\Gamma$ is invertible. Then for any smooth test functions $\varphi_1, \varphi_2 : \mathbb{R}^{6N} \to \mathbb{R}$ with compact support in $\mathcal{O}$,

$$\int_{\Sigma(\zeta)} \{\varphi_1, \varphi_2\}_{\Xi} \, d\sigma_{\Sigma(\zeta)} = 0.$$
The divergence formula (2.26) can be proved using Darboux’s theorem and internal coordinates, or directly using the co-area formula (see Section 3.3 in [28]).

We will also need the classical divergence formula on affine spaces (see for instance Section 3.3 in [28]): for a fixed \( q \in \mathbb{R}^{3N} \), for any smooth vector field \( \phi(q,p) \in \mathbb{R}^{3N} \),

\[
\int_{\Sigma_\epsilon(q_\epsilon)} \text{div}_p \left( P_M(q) \phi(q,p) \right) \sigma_{M-1,\epsilon}^M \, dp = 0.
\]

3. Constrained Langevin processes and sampling

We first give some precisions on the constrained Langevin equation (1.8) in Section 3.1, then propose some numerical schemes to discretize it in Section 3.2, and finally consider the overdamped limit in Section 3.3.

3.1. Properties of the dynamics. We consider the dynamics (1.8). By differentiating with respect to time the constraint \( \xi(q_\epsilon) = z \), the Lagrange multipliers can be computed explicitly (see for instance Section 3.3 in [28]):

\[
d\lambda_t = -G_M^{-1}(q_t) \left[ \text{Hess}_q(\xi)(M^{-1}p_t, M^{-1}p_t) \, dt + \nabla \xi(q_t)^T M^{-1} \left( -\nabla V(q_t) \, dt - \gamma(q_t) M^{-1} p_t \, dt + \sigma(q_t) \, dW_t \right) \right]
\]

\[
= f_{\text{rgd}}^M(q_t, p_t) \, dt + G_M^{-1}(q_t) \nabla \xi(q_t)^T M^{-1} \left( \gamma(q_t) M^{-1} p_t \, dt - \sigma(q_t) \, dW_t \right),
\]

where the constraining force \( f_{\text{rgd}}^M \in \mathbb{R}^m \) is defined as:

\[
f_{\text{rgd}}^M(q, p) = G_M^{-1}(q) \nabla \xi(q)^T M^{-1} \nabla V(q) - G_M^{-1}(q) \text{Hess}_q(\xi)(M^{-1}p, M^{-1}p).
\]

Thus, using the fact that \( P_M(q)^T M^{-1} p = M^{-1} p \) when \( p \in T_q^\ast \Sigma(z) \), the dynamics (1.8) can be recast in a more explicit form as

\[
\begin{cases}
dq_t &= M^{-1} p_t \, dt, \\
dp_t &= -\nabla V(q_t) \, dt + \nabla \xi(q_t) f_{\text{rgd}}^M(q_t, p_t) \, dt - \gamma_p(q_t) M^{-1} p_t \, dt + \sigma_p(q_t) \, dW_t,
\end{cases}
\]

where we introduced the notation \( (\sigma_p, \gamma_p) := (P_M \sigma, P_M \gamma P_M^T) \). The constraint therefore has two effects: (i) the matrices \( \gamma, \sigma \) in the dissipation and fluctuation terms are replaced by their projected counterparts \( \gamma_p, \sigma_p \), and (ii) an orthogonal constraining force \( \nabla \xi f_{\text{rgd}}^M \) is introduced.

The generator of this stochastic Langevin dynamics is the operator \( \mathcal{L}_\Xi \) which appears in the Kolmogorov evolution equation: for \( (q_t, p_t) \) satisfying (1.8) or (3.3), and for any smooth test function \( \varphi \),

\[
\frac{d}{dt} \mathbb{E}(\varphi(q_t, p_t)) = \mathbb{E}(\mathcal{L}_\Xi(\varphi)(q_t, p_t)).
\]

The expression of \( \mathcal{L}_\Xi \) can be obtained using Itô calculus, as detailed in the following proposition.

**Proposition 3.1.** Consider either the effective momentum (2.3) or the effective velocity (2.4), denoted with the general constraints \( \Xi = (\xi, v) \) or \( \Xi = (\xi, p) \) (see (2.10)). The solution of the constrained dynamics (1.8) (or equivalently (2.11) belongs to \( \Sigma_\Xi(z, 0) \), and the generator of this Markov process reads (whatever the value of \( z \))

\[
\mathcal{L}_\Xi = \{\cdot, H\}_\Xi + \mathcal{L}_{\Xi}^{\text{thm}},
\]

where the fluctuation-dissipation part is

\[
\mathcal{L}_{\Xi}^{\text{thm}} = \frac{1}{2} \text{div}_p \left( \sigma_p \sigma_p^T \nabla_p \cdot \right) - p^T M^{-1} \gamma_p \nabla_p,
\]

with \( (\sigma_p, \gamma_p) \) defined in (1.17). Using the fluctuation-dissipation relation (1.9), the generator \( \mathcal{L}_\Xi^{\text{thm}} \) can be rewritten more compactly as

\[
\mathcal{L}_\Xi^{\text{thm}} = \frac{1}{\beta} e^{\beta H} \text{div}_p \left( e^{-\beta H} \gamma_p \nabla_p \cdot \right).
\]
Proof. We perform the computation in two steps: (i) We compute the generator of the Hamiltonian part of the constrained Langevin dynamics, which is \(1.8\) in the case \((\sigma, \gamma) = (0, 0)\); (ii) we compute the generator of the "thermostat" part of \((1.8)\), which is an Ornstein-Uhlenbeck process on momentum variable (corresponding to the second equation in \((1.8)\) with \(V = 0\)). We actually use the reformulation \((3.3)\) for this second part.

Let us first consider (i), with \(\Xi = (\xi, v_\xi)\) (the case \(\Xi = (\xi, p_\xi)\) being similar). Note that

\[
\{\Xi, H\}(q, p) = \left(\text{Hess}_q(\xi)(M^{-1}p, M^{-1}p) - \nabla\xi(q)^T M^{-1} \nabla V(q)\right),
\]

where the Hessian operator Hess is defined in \((2.1)\). Now, \((2.11)\) implies that

\[
(3.7) \quad \Gamma^{-1} = \begin{pmatrix}
G_M^{-1} \nabla v_\xi \nabla v_\xi G_M^{-1} - G_M^{-1} & \\
0 & 
\end{pmatrix}.
\]

Besides, \(v_\xi(q, p_t) = 0\) along a trajectory, since \((q_t, p_t)\) \(\in T^*\Sigma(z)\). Therefore,

\[
(3.8) \quad \forall (q, p) \in T^*\Sigma(z), \quad \Gamma^{-1} \{\Xi, H\}(q, p) = \left(f_{\text{rgd}}^M(q, p)\right),
\]

where the notation \(f_{\text{rgd}}^M\) is introduced in \((3.2)\). Consider a test function \(\varphi : \mathbb{R}^6 \to \mathbb{R}\), and remark that

\[
(3.9) \quad \{\varphi, \Xi\}(a) = -a^T \nabla\xi^T \nabla p \varphi,
\]

so that, for any \(a \in \mathbb{R}^m\),

\[
\{\varphi, \Xi\} \Gamma^{-1} \{\Xi, H\} = -(f_{\text{rgd}}^M)^T \nabla\xi^T \nabla p \varphi.
\]

Finally, for all \((q, p) \in T^*\Sigma(z)\),

\[
(3.10) \quad \{\varphi, H\}_\Xi(q, p) = -\nabla V(q)^T \nabla p \varphi(q, p) + f_{\text{rgd}}^M(q, p)^T \nabla\xi(q)^T \nabla p \varphi(q, p) + p^T M^{-1} \nabla q \varphi(q, p).
\]

The operator \((3.10)\) is the generator of the Hamiltonian part in \((3.3)\).

We turn to (ii). The diffusive part arises from the fluctuation term \(\sigma_p(q_t) dW_t\) in \((3.3)\), and its expression

\[
\frac{1}{2} \text{div}_p \left(P_M \sigma \sigma^T P_M^T \nabla_p \cdot \right)
\]

is obtained directly from the standard Itô calculus. Similarly, the dissipation operator is

\[
-\left(\gamma_p M^{-1} p\right)^T \nabla_p = -p^T M^{-1} P_M \gamma P_M^T \nabla p.
\]

The addition of these two contributions gives the expression of \(L_{\Xi}^{\text{thm}}\).

With the expression \((3.3)\) of the generator at hand, it is easily checked that the process \((1.8)\) satisfies the following equilibrium properties:

**Proposition 3.2.** When the fluctuation-dissipation relation \((1.9)\) holds, the constrained Langevin dynamics \((1.8)\) on \(T^*\Sigma(z)\) admits the Boltzmann-Gibbs distribution \((1.7)\) as a stationary measure, and is reversible up to momentum reversal with respect to \((1.7)\): If \(\text{Law}(q_0, p_0) = \mu_{T^*\Sigma(z)}\), then, for any \(T > 0\),

\[
\text{Law}(q_t, p_t; 0 \leq t \leq T) = \text{Law}(q_{T-t}, -p_{T-t}; 0 \leq t \leq T).
\]

Moreover, if \(P_M(q)\gamma P_M(q)^T\) is everywhere strictly positive in the sense of symmetric matrices on \(T_q^*\Sigma(z)\), then the process \((1.8)\) is ergodic: for any smooth test function \(\varphi\),

\[
\lim_{T \to +\infty} \frac{1}{T} \int_0^T \varphi(q_t, p_t) dt = \int_{T^*\Sigma(z)} \varphi d\mu_{T^*\Sigma(z)} \quad \text{a.s.}
\]
Proof. The stationarity and reversibility properties follow from the following detailed balance condition up to momentum reversal (see for instance Section 2.2 in [28]): for any test functions $\varphi_1, \varphi_2$,

$$
(3.11) \quad \int_{T^*\Sigma(z)} \varphi_1 L_{\Xi}(\varphi_2) \, d\mu_{T^*\Sigma(z)} = \int_{T^*\Sigma(z)} \left( \varphi_2 \circ S \right) L_{\Xi}(\varphi_1 \circ S) \, d\mu_{T^*\Sigma(z)},
$$

where $S : (q, p) \mapsto (q, -p)$ is the momentum flip. In view of the expression of the generator, this condition should be verified for the operators $\{\cdot, H\}_\Xi$ and $L_{\Xi}^{\text{harm}}$.

For the Hamiltonian part $\{\cdot, H\}_\Xi$, the expression (3.11) yields

$$
\{\varphi \circ S, H\}_\Xi(q, p) = - \{\varphi, H\}_\Xi(q, -p) = - \{\varphi, H\}_\Xi(S(q, p)),
$$

which states the time symmetry under momentum reversal of the Hamiltonian part of the equations of motion (1.8). On the other hand,

$$
e^{-\beta H} \{\cdot, H\}_\Xi = - \frac{1}{\beta} \{\cdot, e^{-\beta H}\}_\Xi,$$

so that

$$
e^{-\beta H} (\varphi_2 \circ S) \{\varphi_1 \circ S, H\}_\Xi = - \left( e^{-\beta H} \varphi_2 \{\varphi_1, H\}_\Xi \right) \circ S = \left( e^{-\beta H} \varphi_1 \{\varphi_2, H\}_\Xi + \left\{ \varphi_2 \varphi_1, \frac{e^{-\beta H}}{\beta} \right\} \right) \circ S,$$

and the divergence formula (2.25) yields the balance condition (3.11) for the Hamiltonian part, in view of the invariance of the distribution $\sigma_{T^*\Sigma(z)}$ under the momentum flip $S$.

For the thermostat part, it is easily checked, that

$$L_{\Xi}^{\text{harm}}(\varphi \circ S) = L_{\Xi}^{\text{harm}}(\varphi) \circ S$$

for any smooth test function $\varphi$, so that the detailed balance condition up to momentum reversal (3.11) follows from the following more general detailed balance condition, in the case $v_z = 0$ ($\mu_{\zeta_v,v_z}(z,v_z)$ being defined in (2.13)):

$$
(3.12) \quad \int_{\Sigma_{\zeta_v,v_z}(z,v_z)} \varphi_1 L_{\Xi}^{\text{harm}}(\varphi_2) \, d\mu_{\Sigma_{\zeta_v,v_z}(z,v_z)} = \int_{\Sigma_{\zeta_v,v_z}(z,v_z)} \varphi_2 L_{\Xi}^{\text{harm}}(\varphi_1) \, d\mu_{\Sigma_{\zeta_v,v_z}(z,v_z)}.
$$

It is interesting to prove (3.12) for a general $v_z \in \mathbb{R}$ since it will be used in the proof of Theorem 5.3 below. Consider the divergence formula (2.25) in the affine space for the variable $p$ (the position $q$ being fixed), with

$$\phi = \gamma P_M \nabla_p (\varphi_2) e^{-\beta H} \varphi_1.$$

After integration in $q$, using the formula (2.24) for $L_{\Xi}^{\text{harm}}$ and (2.25), an expression symmetric in $(\varphi_1, \varphi_2)$ is obtained:

$$
\int_{\Sigma_{\zeta_v,v_z}(z,v_z)} \varphi_1 L_{\Xi}^{\text{harm}}(\varphi_2) \, d\mu_{\Sigma_{\zeta_v,v_z}(z,v_z)} = - \int_{\Sigma_{\zeta_v,v_z}(z,v_z)} \nabla_p \varphi_1 P_M \cdot \nabla_p \varphi_2 \, d\mu_{\Sigma_{\zeta_v,v_z}(z,v_z)},
$$

hence the detailed balance condition (3.12).

Ergodicity comes from the hypo-ellipticity of the operator $L_{\Xi}$ on $T^*\Sigma(z)$ (Hörmander’s criterion is satisfied, see [28]), which is itself a consequence of the fact that $P_M(q)\gamma P_M(q)^T$ is strictly positive on each $T_q^*\Sigma(z)$. The proof can be carried out using local coordinates and the results from [29]. □

Remark 3.3 (Infinite stiffness limit). Constrained dynamics are not, except in the very special case of affine constraints, the limit of a highly oscillatory system with slow manifold $\Sigma(z)$, for instance with potential energies of the form

$$V_z(q) = V(q) + \frac{1}{\varepsilon^2} |\xi(q) - z|^2.$$

The infinite stiffness limit ($\varepsilon \to 0$) of highly oscillatory dynamics with random perturbations has been studied in [32], where it is shown that adiabatic effective potentials (derived from the conservation of the ratio energy over frequency of fast modes) are required to describe the limiting dynamics. However, a formal argument based on “over-damping” the fast modes leads to some Markovian effective dynamics. This effective dynamics, called “softly constrained”, is obtained
by changing the potential $V$ of the “rigidly constrained” Langevin dynamics to an effective potential $V + V_{\text{fix}}$. The additional term

$$V_{\text{fix}}(q) = \frac{1}{2\beta} \ln \left( \det G_M(q) \right),$$

is sometimes called the Fixman corrector (see [15]).

If $(q_t, p_t)$ satisfies the original potential $V$ (rigidly constrained dynamics), then $q_t$ samples (in the longtime limit) the probability measure proportional to $e^{-\beta V(q)} \sigma_{\Sigma(z)}(dq)$. On the other hand, if $(q_t, p_t)$ satisfies the modified potential $V + V_{\text{fix}}$ (softly constrained dynamics), then $q_t$ samples (in the longtime limit) the probability measure proportional to $e^{-\beta V(q)} \delta_{\xi(q) - z}(dq)$ (see (2.18)). The difference between soft and rigid constraints can therefore be seen as the counterpart on dynamics of the difference between: (i) the canonical distribution (1.6) with positions conditioned by $\xi(q) = z$ (whose marginal on positions is proportional to $e^{-\beta V(q)} \delta_{\xi(q) - z}(dq)$); and (ii) the canonical distribution (1.7) with constraints on both positions and momenta (whose marginal on positions is from (2.23) proportional to $e^{-\beta V(q)} \sigma_{\Sigma(z)}(dq)$).

### 3.2. Numerical implementation

We consider in this section a numerical scheme based on a splitting of the Langevin dynamics into a Hamiltonian part (Section 3.2.1) and a fluctuation-dissipation part acting only on the momentum (Section 3.2.2). For simplicity, we restrict ourselves to constant matrices $\gamma$ and $\sigma$. Generalizations to position dependent matrices are straightforward.

The Hamiltonian part of the Langevin dynamics (1.8) (namely (1.8) with $(\sigma, \gamma) = (0, 0)$) is discretized using a velocity-Verlet scheme with constraints, which yields (3.17) below. The fluctuation-dissipation part on momentum variable in (1.8) is the following Ornstein-Uhlenbeck process (for a fixed given $q \in \Sigma(z)$):

$$dp_t = -\gamma M^{-1} p_t \, dt + \sigma \, dW_t + \nabla \xi(q) \, d\lambda_t^{\text{OU}}, \quad \nabla \xi(q) M^{-1} p_t = 0, \quad (C_p)$$

which can be rewritten as (see (3.3))

$$dp_t = -\gamma_P(q) M^{-1} p_t \, dt + \sigma_P(q) \, dW_t.$$

This equation can be explicitly integrated on $[0, t]$ to obtain:

$$p_t = e^{-t \gamma_P(q) M^{-1}} p_0 + \int_0^t e^{-(t-s) \gamma_P(q) M^{-1}} \sigma_P(q) \, dW_s.$$

However, the matrix exponential $e^{-t \gamma_P(q) M^{-1}}$ may be difficult to compute in practice (except for certain choices of $\gamma$ and $M$, see the discussion at the end of Section 3.2.2). Instead of performing an exact integration, (3.14) can be discretized using a midpoint Euler scheme, which yields (3.16) and (3.18) below.
The numerical scheme we investigate, termed midpoint Euler-Verlet-midpoint Euler splitting, is therefore the following:

\[
\begin{align*}
\begin{cases}
p^{n+1/4} = p^n - \frac{\Delta t}{4} M^{-1}(p^n + p^{n+1/4}) + \sqrt{\frac{\Delta t}{2}} \sigma \mathcal{G}^n \\
\quad + \nabla \xi(q^n) \lambda^{n+1/4}, \\
\nabla \xi(q^n)^T M^{-1} p^{n+1/4} = 0, \\
(3.16) \\
p^{n+1/2} = p^{n+1/4} - \frac{\Delta t}{2} \nabla V(q^n) + \nabla \xi(q^n) \lambda^{n+1/2}, \\
q^{n+1} = q^n + \Delta t M^{-1} p^{n+1/2}, \\
\end{cases}
\end{align*}
\]

\[
\begin{align*}
\begin{cases}
\xi(q^{n+1}) = z, \\
p^{n+3/4} = p^{n+1/2} - \frac{\Delta t}{2} \nabla V(q^{n+1}) + \nabla \xi(q^{n+1}) \lambda^{n+3/4}, \\
\nabla \xi(q^{n+1})^T M^{-1} p^{n+3/4} = 0, \\
(3.17) \\
p^{n+1} = p^{n+3/4} - \frac{\Delta t}{4} M^{-1}(p^{n+3/4} + p^{n+1}) + \sqrt{\frac{\Delta t}{2}} \sigma \mathcal{G}^{n+1/2} \\
\quad + \nabla \xi(q^{n+1}) \lambda^{n+1}, \\
\nabla \xi(q^{n+1})^T M^{-1} p^{n+1} = 0, \\
(3.18) \\
\end{cases}
\end{align*}
\]

where \((\mathcal{G}^n)_{n \geq 0}\) and \((\mathcal{G}^{n+1/2})_{n \geq 0}\) are sequences of independently and identically distributed (i.i.d.) Gaussian random variables of mean 0 and covariance matrix \(\text{Id}_{3N}\).

Note that when \(\gamma = 0\) and \(\sigma = 0\), the scheme (3.16)-(3.17)-(3.18) becomes deterministic, and reduces to (3.17), which is a scheme for the deterministic Hamiltonian equations of motion with position constraints \(\xi(q) = z\). The latter scheme is referred to as the "Hamiltonian scheme (3.17)" below.

3.2.1. Comments on the Hamiltonian scheme (3.17). The Hamiltonian part (3.17) of the scheme, often called 'RATTLE' in the literature, is an explicit integrator, and is the velocity version of the classical 'SHAKE' algorithm (see Chapter VII.1 in [16], or Chapter 7 in [25] for more precisions and historical references). In (3.17), \(\lambda^{n+1/2} \in \mathbb{R}^m\) are the Lagrange multipliers associated with the position constraints \((C_p)\), and \(\lambda^{n+3/4} \in \mathbb{R}^m\) are the Lagrange multipliers associated with the velocity constraints \((C_p)\). The nonlinear constraints \((C_q)\) are typically enforced using Newton’s algorithm. In (3.17), the (linear) momentum projection \((C_p)\) is always well defined by assuming the invertibility of the Gram matrix \(G_M(q)\), whereas the nonlinear projection used to enforce the position constraints \(\xi(q^{n+1}) = z\) is in general well defined only on a subset of phase space.

**Definition 3.4** (Domain \(D_{\Delta t}\)). The domain \(D_{\Delta t} \subset T^*\Sigma(z)\) is defined as the set of configurations \((q^n, p^{n+1/4}) \in T^*\Sigma(z)\) such that there is a unique solution \((q^{n+1}, p^{n+3/4})\) verifying (3.17).

Solving the position constraints \((C_q)\) consists in projecting onto \(\Sigma(z)\) a point in a \(\Delta t\)-neighborhood of \(q^n\). Thus, by the implicit function theorem, the domain \(D_{\Delta t}\) verifies:

\[
\lim_{\Delta t \to 0} D_{\Delta t} = T^*\Sigma(z).
\]

It may happen that there is no solution if the time-step is too large, and, even for small time-steps, that several projections exist, see for instance Example 2 in Chapter 7 of [25]. In practice, \(D_{\Delta t}\) can be chosen to be the set of \((q^n, p^{n+1/4})\) such that the Newton algorithm enforcing the constraints \((C_q)\) has converged within a given precision threshold and a limited number of iterations.

As for the Verlet scheme in the unconstrained case, the associated numerical flow shares two important qualitative properties with the exact flow: It is time reversible and symplectic (see [20]). This implies quasi-conservation of energy, in the sense that energy is conserved within a given precision threshold over exponentially long times, see [10, 25].
3.2.2. Comments on the fluctuation-dissipation part (3.10) and (3.18). The new momentum \( p_{n+1/4} \in T^*\Sigma(z) \) in (3.10) (or \( p^{n+1} \) in (3.18)) may be obtained by first integrating the unconstrained dynamics with a midpoint scheme, and then computing the Lagrange multiplier \( \lambda_{n+1/4} \) (or \( \lambda^{n+1} \)) by solving the following linear system implied by the constraints (3.11):

\[
\nabla \xi(q^n)T^{-1} \left( \text{Id} + \frac{\Delta t}{4} \gamma M^{-1} \right)^{-1} \left( \left( \text{Id} - \frac{\Delta t}{4} \gamma M^{-1} \right) p^n + \sqrt{\frac{\Delta t}{2}} \sigma \mathcal{G}^n + \nabla \xi(q^n) \lambda^{n+1/4} \right) = 0.
\]

A sufficient criteria for stability is

\[
\frac{\Delta t}{4} \gamma \leq M.
\]

It can be checked (see Sections 2.3.2 and 3.3.5 in [28]) that the Markov chain induced by the fluctuation-dissipation part of the scheme (3.11) (or (3.13)) verifies a detailed balance equation (both in the plain sense and up to momentum reversal) with respect to the stationary measure \( \kappa_{T^*\Sigma(z)}^{-1}(dp) \). The latter is defined as the kinetic probability distribution

\[
\kappa_{T^*\Sigma(z)}^{-1}(dp) = \left( \frac{\beta}{2\pi} \right)^{(3N-m)/2} \exp \left( -\frac{\beta}{2} p^T M^{-1} p \right) \sigma_{T^*\Sigma(z)}^{-1}(dp),
\]

and is the marginal in the momentum variable of the canonical distribution \( \mu_{T^*\Sigma(z)}(dq dp) \) conditioned by a given \( q \in \Sigma(z) \). Moreover, if \( \gamma_p := P_{\Sigma} \gamma P_{\Sigma}^T \) is strictly positive in the sense of symmetric linear transformations of \( T_q\Sigma(z) \), then the Markov chain on momentum variable induced by (3.10) (or (3.18)) alone is ergodic with respect to \( \kappa_{T^*\Sigma(z)}^{-1}(dp) \).

Finally, an important simplification occurs in the integration of (3.14) in the special case when \( \gamma \) and \( M \) are equal up to a multiplicative constant (so that \( \gamma M^{-1} \) is proportional to identity). Indeed in this case the equality \( (\gamma p(q)M^{-1})^n = P_{\Sigma}(q) (\gamma M^{-1})^n \) holds for any \( n \geq 0 \), and (3.14) simplifies to

\[
p_t = P_{\Sigma}(q) \left( e^{-t \gamma M^{-1}} p_0 + \int_0^t e^{-(t-s)\gamma M^{-1}} \sigma dW_s \right).
\]

The numerical integration of (3.14) can thus be carried out in two steps: (i) exactly integrating (3.14) without constraint, and then (ii) projecting the result onto \( T_q\Sigma(z) \).

3.2.3. Metropolis-Hastings correction. Usually, the invariant probability distribution sampled by the solution of a numerical scheme is biased by the time discretization. Relying on (i) the time symmetry (up to momentum reversal) and (ii) the preservation of the phase space measure \( \sigma_{T^*\Sigma(z)}(dq dp) \) by the solution of the RATTLE scheme (3.17), it is possible to eliminate the time discretization error in the splitting scheme (3.10)-(3.17)-(3.18) by resorting to a Generalized Hybrid Monte Carlo algorithm.

**Algorithm 3.5** (GHMC with constraints). Consider an initial configuration \((q^0, p^0) \in T^*\Sigma(z)\), and a sequence \((\mathcal{G}^n, \mathcal{G}^{n+1/2})_{n\geq0}\) of independently and identically distributed standard Gaussian vectors. Iterate on \( n \geq 0 \):

1. Evolve the momentum according to the midpoint Euler scheme (3.10), and compute the energy \( E^n = H(q^n, p^{n+1/4}) \) of the new configuration;
2. Integrate the Hamiltonian part according to the RATTLE scheme (3.17), denote \((\tilde{q}^{n+1}, \tilde{p}^{n+3/4})\) the resulting state, and set \( E^{n+1} = H(\tilde{q}^{n+1}, \tilde{p}^{n+3/4}) \).
3. Accept the proposal \((q^{n+1}, p^{n+3/4}) := (\tilde{q}^{n+1}, \tilde{p}^{n+3/4})\) with probability

\[
\min \left( e^{-\beta(E^{n+1}-E^n)}, 1 \right).
\]

Otherwise, reject and flip the momentum: \((q^{n+1}, p^{n+3/4}) = (q^n, -p^{n+1/4})\).
4. Evolve the momentum according to the midpoint Euler scheme (3.18).
By construction, the GHMC algorithm with constraints leaves invariant the equilibrium distribution $\mu_{T\cdot\Sigma(z)}(dq\,dp)$ (see Section 3.3.5 in [28]).

In the above, we implicitly assume that the RATTLE scheme (3.17) is everywhere well defined. In practice, however, it is necessary to modify Algorithm 3.5 by restricting the sampled configurations to $D_{\Delta t}$. This can be achieved by introducing additional tests in steps (1), (2) and (4), and rejecting the states that have gone outside the set $D_{\Delta t} \subset T^*\Sigma(z)$ where the position constraint $(C_q)$ is well defined. By doing so, the global algorithm has an invariant equilibrium distribution given by $\mu_{T\cdot\Sigma(z)}(dq\,dp)$ conditioned on the set of states $D_{\Delta t}$. This invariant distribution can be written explicitly as follows:

$$\frac{1}{Z_{z,0,\Delta t}} e^{-\beta H(q,p)} 1_{(q,p) \in D_{\Delta t} \cap \Sigma(T^*\Sigma(z))}(dq\,dp).$$

Alternatively, the rejection tests in steps (1), (2) and (4) of Algorithm 3.5 can be performed with a cut-off parameter $R_{\Delta t} > 0$ on the momentum variable, chosen so that the position constraint $(C_q)$ in (3.17) is everywhere well defined when $\frac{1}{2}p^T M^{-1} p \leq R_{\Delta t}$. This can be achieved when there exists $R_{\Delta t} > 0$ small enough so that $\Sigma(z) \times \{ \frac{1}{2}p^T M^{-1} p \leq R_{\Delta t} \} \subset D_{\Delta t} \subset T^*\Sigma(z)$. Since this is useful for later purposes (see the discussion at the end of Section 4.3), we provide a rough estimate of $R_{\Delta t}$ in terms of $\Delta t$, assuming for simplicity that $\Sigma(z)$ is compact. First, by the implicit function theorem, there exists $\alpha > 0$ such that, for all $q \in \Sigma(z)$ and $\delta q$ with norm $\|\delta q\| < \alpha$, there is a unique $\lambda \in \mathbb{R}^m$ satisfying

$$\xi(q + M^{-1}(\delta q + \nabla \xi(q)\lambda)) = z.$$

Therefore, there exists $\alpha > 0$ small enough such that, when $\|p^{n+1/4}\| \leq \alpha / \Delta t$, the RATTLE scheme in (3.17) is well defined, namely there exists a unique $q^{n+1}$ satisfying the constraint $(C_q)$. This shows that

$$R_{\Delta t} \geq A \Delta t^{-2}$$

for some $A > 0$.

The invariant probability distribution of the Markov chain generated by GHMC with the additional rejection steps ensuring $\frac{1}{2}p^T M^{-1} p \leq R_{\Delta t}$, is given by (3.21), and actually reads

$$\frac{1}{Z_{z,0,\Delta t}} e^{-\beta H(q,p)} 1_{\frac{1}{2}p^T M^{-1} p \leq R_{\Delta t} \cap \Sigma(T^*\Sigma(z))}(dq\,dp),$$

and the marginal distribution in the position variable is then exactly given by:

$$\frac{1}{Z_z} e^{-\beta V(q)} \sigma^M_{\Sigma(z)}(dq).$$

This is also the marginal distribution in the position variable of $\mu_{T\cdot\Sigma(z)}$. Note however, that if $R_{\Delta t}$ is too small, only small momenta will be sampled in step (1) of Algorithm 3.5 and the correlation time of the sampling will be large. In practice, the threshold $R_{\Delta t}$ should be tuned in preliminary computations so that: (i) $R_{\Delta t}$ is small enough so that the maximal number $N_{\max}$ of iterations for the Newton algorithm used to enforce $(C_q)$ in (3.17) is never reached; (ii) $R_{\Delta t}$ is large enough so that the correlation time of the sampling is as small as possible.

Let us end this section with a warning: It is now known that the correction of the bias in discretizations of the Langevin dynamics by a Metropolization of the scheme may reduce the efficiency of the sampling, see for instance [11].

3.3. Exact sampling on a submanifold with overdamped dynamics. Constrained overdamped Langevin processes (or Brownian dynamics) are solutions of the stochastic differential equation (see also [28, 6])

$$\begin{cases} dq_t = -\nabla V(q_t) \, dt + \sqrt{\frac{2}{\beta}} \, dW_t + \nabla \xi(q_t) \, d\lambda_t, \\ \xi(q_t) = z, \end{cases}$$

(3.23)
where $\lambda_t$ is an adapted stochastic process. Equivalently, (3.23) can be rewritten in the Stratonovitch form as
\[
dq_t = -P(q_t)\nabla V(q_t)\, dt + \sqrt{\frac{2}{\beta}} P(q_t) \circ dW_t,
\]
where $\circ$ denotes the Stratonovitch integration, and $P$ is the projector defined by (1.3) with the choice $M = \text{Id}$. It can be shown that (3.23) verifies the detailed balance condition for (and is ergodic with respect to) the invariant distribution
\[
(3.24)
Z^{-1}_z e^{-\beta V(q)\sigma_{\Sigma(z)}^\text{Id}}(dq),
\]
which is the marginal in the $q$-variable of the canonical distribution with constraints (1.7) for the choice $M = \text{Id}$. The constrained overdamped Langevin process (3.23) may be obtained from a scaling limit of the constrained Langevin dynamics (1.8) (in the limit when either the mass goes to zero, or the damping $\gamma$ goes to infinity), using a formal argument as in the proof of Proposition 2.5 in [28].

Likewise, at the discrete level, an Euler-Maruyama discretization of the overdamped process (3.23) can be obtained as a particular case of the numerical discretization (3.16)-(3.17)-(3.18) for the Langevin equation (1.8), yielding a Markov chain $(q^n)_{n \geq 0}$ on positions. This is detailed in the following proposition.

**Proposition 3.6.** Suppose that the following relation is satisfied:
\[
(3.25)
\frac{\Delta t}{4} \gamma = M = \frac{\Delta t}{2} \text{Id}.
\]
With a slight abuse of notation, the mass matrix and the friction matrix are rewritten as $M \text{Id}$ and $\gamma \text{Id}$ with $M, \gamma \in \mathbb{R}$. Then the splitting scheme (3.16)-(3.17)-(3.18) yields the following Euler scheme for the overdamped Langevin constrained dynamics (3.23):
\[
(3.26)
\begin{align*}
q^{n+1} &= q^n - \Delta t \nabla V(q^n) + \sqrt{\frac{2 \Delta t}{\beta}} G^n + \nabla \xi(q^n) \lambda^{n+1}_{\text{od}}, \\
\xi(q^{n+1}) &= z,
\end{align*}
\]
where $(G^n)_{n \geq 0}$ are independent and identically distributed centered and normalized Gaussian variables, and $(\lambda^{n+1}_{\text{od}})_{n \geq 1}$ are the Lagrange multipliers associated with the constraints $(\xi(q^n) = z)_{n \geq 1}$. Moreover, the Lagrange multipliers in (3.17) verify:
\[
(3.27)
2\lambda^{n+1/2} = G^{-1}(q^n) \left( \nabla \xi(q^n)^T (q^{n+1} - q^n) + \Delta t \nabla \xi(q^n)^T \nabla V(q^n) \right)
\]
\[
(3.28)
\lambda^{n+1}_{\text{od}} = \frac{2 \Delta t}{\beta} G^{-1}(q^n) \nabla \xi(q^n)^T G^n,
\]
as well as
\[
(3.29)
2\lambda^{n+3/4} = G^{-1}(q^{n+1}) \left( \nabla \xi(q^{n+1})^T (q^n - q^{n+1}) + \Delta t \nabla \xi(q^{n+1})^T \nabla V(q^{n+1}) \right).
\]
In the above, $G$ denotes the Gram matrix (1.5) with the choice $M = \text{Id}$. 

**Proof.** Irrespective of $p^n$, the choice (3.25) in the scheme (3.16)-(3.17)-(3.18) leads to
\[
p^{n+1/4} = \sqrt{\frac{\Delta t}{8}} \sigma G^n + \frac{1}{2} \nabla \xi(q^n) \lambda^{n+1/4},
\]
where $\lambda^{n+1/4}$ is associated with the constraints $\nabla \xi(q^n)^T p^{n+1/4} = 0$. This gives
\[
p^{n+1/2} = -\frac{\Delta t}{2} \nabla V(q^n) + \sqrt{\frac{\Delta t}{8}} \sigma G^n + \nabla \xi(q^n) \left( \frac{1}{2} \lambda^{n+1/4} + \lambda^{n+1/2} \right),
\]
where $\lambda^{n+1/2}$ is such that $\xi(q^{n+1}) = z$. The fluctuation-dissipation relation (1.9) can be reformulated in this context as
\[
\sigma \sigma^T = \frac{2}{\beta} \gamma = \frac{4}{\beta} \text{Id},
\]
and the scheme (3.26) is recovered by taking the associated Lagrange multiplier equal to $\lambda_{n+1} = \lambda_{n+1/4} + 2\lambda_{n+1/2}$. Finally, remarking that $G_M = \frac{1}{2\beta} G$ and computing explicitly $\lambda_{n+1/2}$ and $\lambda_{n+1/4}$ in (3.17) yields (3.27)-(3.28)-(3.29).

\[ \square \]

**Remark 3.7.** Proposition 3.6 can be generalized to the less restrictive relation

\[ \Delta t \gamma = M \propto \text{Id}, \]

upon considering (3.26) with $\Delta t$ replaced by the effective time-step (3.30)

\[ \Delta s = \Delta t^2/2M = 2\Delta t/\gamma. \]

\[ \square \]

This point of view allows to construct a Metropolis correction to the Euler scheme (3.26), using the Generalized Hybrid Monte Carlo scheme (Algorithm 3.5) with the time-step chosen according to (3.25). In this way, assuming that the position constraint ($C_q$) in (3.17) is everywhere well defined, we obtain a Markov chain $(q^n)_{n \geq 0}$ discretizing the overdamped dynamics (3.26) which exactly samples the invariant distribution (3.24). Deriving such a Metropolis-Hastings correction to the Euler scheme (3.26) without resorting to phase-space dynamics does not seem to be natural.

4. THERMODYNAMIC INTEGRATION WITH CONSTRAINED LANGEVIN DYNAMICS

In this section, we focus on the computation of the gradient of the rigid free energy (1.13), using a numerical discretization of the constrained Langevin process (1.8). The numerical approximation uses the Lagrange multipliers needed to integrate (1.8).

As explained in the introduction, we may concentrate on the computation of the rigid free energy (1.13), since the standard free energy (1.12) can be computed from the latter using (1.14). The relation (1.14) can be proved with the co-area formula (2.18). Indeed, the free energy defined in (1.12) can be rewritten as (where $C$ denotes a constant which may vary from line to line):

\[ F(z) = -\frac{1}{\beta} \ln \int \Sigma(z) \times \mathbb{R}^N e^{-\beta H(q,p)} \delta_{\xi(q)} q(z)(dq) dp \]

\[ = -\frac{1}{\beta} \ln \int \Sigma(z) e^{-\beta V(q)} (\det G_M(q))^{-1/2} \sigma_M(q)(dq) + C \]

\[ = -\frac{1}{\beta} \ln \int T^* \Sigma(z) e^{-\beta H(q,p)} (\det G_M(q))^{-1/2} \sigma_{T^* \Sigma(z)}(dp dp) + C \]

\[ = F_{rgd}^M(z) - \frac{1}{\beta} \ln \int T^* \Sigma(z) (\det G_M(q))^{-1/2} d\mu_{T^* \Sigma(z)} + C, \]

where surface measures are defined in Section 2.3. Note that the rigid free energy $F_{rgd}^M$, defined in (1.13), indeed depends explicitly on the mass matrix since

\[ F_{rgd}^M(z) = -\frac{1}{\beta} \ln \int \Sigma(z) e^{-\beta V(q)} \sigma_{\Sigma(z)}(dq) + C. \]

This section is organized as follows. First, we show how systems with molecular constraints and systems with constrained values of the reaction coordinate can be treated in a unified framework (Section 4.1). We then relate the Lagrange multipliers arising in the constrained Langevin dynamics, and the gradient of the rigid free energy (the so-called mean force) in Section 4.2. We consider the numerical computation of the mean force in Section 4.3 where we prove consistency results for the corresponding approximation formulas. Finally, some numerical results on a model system illustrate the approach in Section 4.4.
4.1. Molecular constraints. We discuss here how to generalize all the computations to systems with molecular constraints. This section can be considered as independent of the remainder of the paper and may therefore be omitted in a first reading.

In practice, many systems are subject to molecular constraints, such as fixed lengths for covalent bonds, or fixed angles between covalent bonds. The reader is referred to [31] for practical aspects related to the simulation of molecular constraints. In the context of free energy computations, two types of constraints are therefore considered: first, the molecular constraints, and second, the reaction coordinates denoted in this section by \( \xi_{rc} : \mathbb{R}^{3N} \rightarrow \mathbb{R}^m \), with \( m + m < 3N \). The submanifold of molecular constraints is denoted by

\[
\Sigma_{mc} = \{ q \in \mathbb{R}^{3N} \mid \xi_{mc}(q) = 0 \},
\]

and the submanifold associated with the reaction coordinates by

\[
\Sigma_{rc}(z_{rc}) = \{ q \in \mathbb{R}^{3N} \mid \xi_{rc}(q) = z_{rc} \}.
\]

It is assumed that the full Gram matrix:

\[
G_{M}^{mc,rc} := \nabla (\xi_{mc}, \xi_{rc})^T M^{-1} \nabla (\xi_{mc}, \xi_{rc}) \in \mathbb{R}^{(m+m) \times (m+m)}
\]

is everywhere invertible on \( \Sigma_{mc} \cap \Sigma_{rc}(z_{rc}) \). Likewise, we denote

\[
G_{M}^{rc} := \nabla \xi_{rc}^T M^{-1} \nabla \xi_{rc} \in \mathbb{R}^{m \times m},
\]

and

\[
G_{M}^{mc} := \nabla \xi_{mc}^T M^{-1} \nabla \xi_{mc} \in \mathbb{R}^{(m+m)}.
\]

Assuming rigid mechanical constraints on the molecular constraints \( \xi_{mc} \), we are led to considering the canonical distribution

\[
\mu_{T \cdot \Sigma_{mc}}(dq \, dp) = \frac{1}{Z_{mc}} e^{-\beta H(q,p)} \delta_{\xi_{mc}(q), p_{\xi_{mc}}(q)}(dq \, dp) = \frac{1}{Z_{mc}} e^{-\beta H(q,p)} \sigma_{T \cdot \Sigma_{mc}}(dq \, dp),
\]

(4.4)

to describe systems with molecular constraints at a fixed temperature. The measure \( \sigma_{T \cdot \Sigma_{mc}} \) denotes the phase space measure on \( T^\ast \Sigma_{mc} \), equal by (2.21) to the conditional measure \( \delta_{\xi_{mc}(q), p_{\xi_{mc}}(q)}(dq \, dp) \) associated with the constraints \( \{ \xi_{mc}(q) = 0, p_{\xi_{mc}}(q, p) = 0 \} \), where \( p_{\xi_{mc}} \) is the effective momentum (2.15) associated with \( \xi_{mc} \).

Remark 4.1 (On the choice of the distribution (4.4)). The distribution \( \mu_{T \cdot \Sigma_{mc}} \) in (4.4) is obtained by constraining rigidly \( \xi_{mc}(q) \) to 0, and not "softly" (in which case \( \delta_{\xi_{mc}(q), p_{\xi_{mc}}(q)}(dq \, dp) \) would be replaced by \( \delta_{\xi_{mc}(q)}(dq \, dp) \), see also Remark 3.3). As explained in Section 3, the distribution (4.4) is the equilibrium distribution of a Langevin process (thermostated Hamiltonian dynamics) with rigid position constraints \( \xi_{mc}(q) = 0 \). Two remarks are in order. First, it is possible to rewrite the remainder of this section by considering the softly constrained potential rather than the rigidly constrained potential, up to an appropriate modification of (4.6) below, similar to the introduction of the Fixman corrective potential (3.38). Second, we prefer to stick to the rigidly constrained potential since, for reasons coming from quantum mechanics, systems with stiff molecular constraints typically “do not oscillate” in the vicinity of the submanifolds defined by \( \xi_{mc}(q) = 0 \). Indeed, the characteristic energy associated with a quantum oscillator is \( \hbar \omega \) where \( \omega \) is the frequency of the oscillator. For molecular bonds, \( \hbar \omega \gg k_B T \) at usual temperatures; for instance \( \hbar \omega / k_B = 7600 \, K \) for H-C bonds, and \( \hbar \omega / k_B = 2300 \, K \) for C-C bonds (see [33], Table 8.1). As a consequence, the thermal energy is not sufficient to change the quantum mechanical energy of the oscillator, which thus remains in its fundamental state. Since the classical limit of the fundamental state is loosely speaking the minimal energy position, it is reasonable to assume that such constraints strictly remain at their fixed, equilibrium values. \( \square \)
By associativity of the conditioning of measures, the distribution $\mu_{T^*\Sigma_{mc}}$ conditioned by a value of the reaction coordinates $\xi_{rc}(q) = z_{rc}$ is given, up to a normalizing factor, by:
\[
e^{-\beta H(q,p)}\delta_{\xi_{mc}(q),p_{\xi_{mc}}(q,p),\xi_{rc}(q) - z_{rc}}(dq \, dp).
\]
Therefore, considering the marginal probability distribution of the reaction coordinates $\xi_{rc}(q)$ leads to the following definition of the free energy associated with $\xi_{rc}$:
\[
F^{mc}(z_{rc}) = -\frac{1}{\beta} \ln \int_{T^*\Sigma_{mc}^r(\Sigma_{rc}(z_{rc}) \times \mathbb{R}^{3N})} e^{-\beta H(q,p)}\delta_{\xi_{mc}(q),p_{\xi_{mc}}(q,p),\xi_{rc}(q) - z_{rc}}(dq \, dp).
\]
The conditional distribution can be decomposed as follows, using the co-area formulas [2.13]–[2.20] and the definition of effective momentum [2.5]:
\[
\delta_{\xi_{mc}(q),p_{\xi_{mc}}(q,p),\xi_{rc}(q) - z_{rc}}(dq \, dp)
= \delta_{p_{\xi_{mc}}(q,p)}(dp)\delta_{\xi_{mc}(q),\xi_{rc}(q) - z_{rc}}(dq)
= (\det G_{T_q}^M(q))^{1/2} \sigma_{T_q}^{M^{-1}}(dq) \left(\det G_{T_q}^{mc,rc}(q)\right)^{-1/2} \sigma_{\Sigma_{rc}(z_{rc})\cap\Sigma_{mc}}(dq).
\]
Integrating out the momentum in the linear space $T_q^*\Sigma_{mc}$ with scalar product $\langle p_1, p_2 \rangle_{M^{-1}} = p_1^T M^{-1} p_2$, the free energy can be rewritten as:
\[
F^{mc}(z_{rc}) = -\frac{1}{\beta} \ln \int_{\Sigma_{rc}(z_{rc})\cap\Sigma_{mc}} e^{-\beta V(q)} \left(\frac{\det G_{T_q}^M(q)}{\det G_{T_q}^{mc,rc}(q)}\right)^{1/2} \sigma_{\Sigma_{rc}(z_{rc})\cap\Sigma_{mc}}(dq) + C.
\]
As a consequence, the free energy $F^{mc}$ can be computed from the generalized rigid free energy:
\[
F^{mc, M}(z_{rc}, 0) = -\frac{1}{\beta} \ln \int_{T^*\Sigma_{rc}(z_{rc})\cap\Sigma_{mc}} e^{-\beta H(q,p)} \sigma_{T^*\Sigma_{rc}(z_{rc})\cap\Sigma_{mc}}(dp \, dq),
\]
using the following formula, similar to (1.13):
\[
F^{mc}(z_{rc}) = F^{mc, M}(z_{rc}, 0) - \frac{1}{\beta} \ln \int_{T^*\Sigma_{rc}(z_{rc})\cap\Sigma_{mc}} \frac{(\det G_{T_q}^{mc,rc})^{1/2}}{(\det G_{T_q}^M)^{1/2}} \mu_{T^*\Sigma_{rc}(z_{rc})\cap\Sigma_{mc}}(dq) + C.
\]
In the above, $\mu_{T^*\Sigma_{rc}(z_{rc})\cap\Sigma_{mc}}$ is defined similarly to (1.7). The case of molecular constraints can therefore be treated within the general framework considered in this paper, the sampling of the canonical measure $\mu_{T^*\Sigma_{rc}(z_{rc})\cap\Sigma_{mc}}$ and the computation of the rigid free energy [4.5] being the problems at hand.

4.2. The mean force and the Lagrange multipliers. In this section, the average of the constraining force [3.2] is related to the gradient of the rigid free energy [1.13] (or mean force). We also give a similar result for the following generalized rigid free energy:
\[
F_{\Sigma}^{\Xi}(\zeta) = -\frac{1}{\beta} \ln \int_{\Sigma_{\Xi}(\zeta)} e^{-\beta H(q,p)} \sigma_{\Sigma_{\Xi}(\zeta)}(dq \, dp).
\]

**Proposition 4.2.** The constraining force $f_{\Sigma}^{M, M}: T^*\Sigma(z) \to \mathbb{R}^m$ defined in [3.2] yields on average the rigid free energy derivative:
\[
\nabla_z f_{\Sigma}^{M, M}(z) = \int_{T^*\Sigma(z)} f_{\Sigma}^{M, M}(q,p) \mu_{T^*\Sigma(z)}(dq \, dp).
\]
Moreover, for general constraints [2.6] and the associated generalized free energy [4.7], the formula can be extended as follows: The generalized constraining force is
\[
\left(\frac{f_{\Sigma}^{\Xi}}{g_{\Sigma}^{\Xi}}\right) := \Gamma^{-1} \{\Xi, H\},
\]
where $\Gamma$ is defined in [2.9], and the rigid mean force is
\[
\nabla_\zeta F_{\Sigma}^{\Xi}(\zeta) = \frac{1}{Z_\zeta} \int_{\Sigma_{\Xi}(\zeta)} \left(\frac{f_{\Sigma}^{\Xi}}{g_{\Sigma}^{\Xi}}\right) e^{-\beta H} d\sigma_{\Sigma_{\Xi}(\zeta)},
\]

where \( Z_\xi = \int_{\Sigma(\xi)} e^{-H} d\sigma_{\Sigma(\xi)} \), and \( F_{rgd}^\Sigma(\xi) \) is defined in (4.7). When \((q,p)\) verifies \( p_\xi(q,p) = v_\xi(q,p) = 0 \), then \( g^\Sigma(q,p) = 0 \) and \( f^\Sigma(q,p) = f_{rgd}^M(q,p) \).

**Proof.** Formulas (4.9) and (4.10) are obtained directly by replacing \( \varphi \) by \( e^{-\beta H} \) in Lemma 6.2 (see the Appendix). The fact that \((f^\Sigma(q,p),g^\Sigma(q,p)) = (f_{rgd}^M(q,p),0)\) in the tangential case (namely when \( p_\xi(q,p) = v_\xi(q,p) = 0 \)) is a consequence of (3.8).

The following lemma gives a momentum-averaged version of the constraining force (a similar formula exists in the overdamped case, see Equations (4.8)-(4.9) in [6], for example).

**Lemma 4.3.** The rigid mean force (4.8) can be rewritten as:

\[
(4.11) \quad \nabla_z F_{rgd}(z) = \int_{T^*\Sigma(z)} \overline{T}_{rgd}^M(q) \mu_{T^*\Sigma(z)}(dq dp),
\]

where

\[
(4.12) \quad \overline{T}_{rgd}^M(q) = G^{-1}_M(q) \nabla_\xi(q)^T M^{-1} \nabla V(q) - \beta^{-1} G^{-1}_M(q) \text{Hess}_q(\xi) : (M^{-1} P_M(q)).
\]

**Proof.** Consider the Gaussian distribution \( \kappa_{T^*_q\Sigma(z)}^{M^{-1}}(dp) \) defined in (4.19), which is the marginal distribution in the momentum variable of the canonical distribution \( \mu_{T^*\Sigma(z)}(dq dp) \), conditioned by a given \( q \in \Sigma(z) \). Proving Lemma 4.3 amounts to showing that the average of the constraining force \( f_{rgd}^M \) with respect to \( \kappa_{T^*_q\Sigma(z)}^{M^{-1}}(dp) \) yields \( \overline{T}_{rgd}^M \):

\[
\overline{T}_{rgd}^M(q) = \int_{T^*_q\Sigma(z)} f_{rgd}^M(q,p) \kappa_{T^*_q\Sigma(z)}^{M^{-1}}(dp).
\]

First, we compute the covariance matrix

\[
C := \text{cov}(\kappa_{T^*_q\Sigma(z)}^{M^{-1}})
\]

of the Gaussian distribution \( \kappa_{T^*_q\Sigma(z)}^{M^{-1}}(dp) \). Since \( \kappa_{T^*_q\Sigma(z)}^{M^{-1}}(dp) \) is a centered Gaussian distribution, \( C \) satisfies, for all \( p_1, p_2 \in \mathbb{R}^{3N} \),

\[
p_1^T M^{-1} C M^{-1} p_2 := \int_{T^*_q\Sigma(z)} (p^T M^{-1} p_1) (p^T M^{-1} p_2) \kappa_{T^*_q\Sigma(z)}^{M^{-1}}(dp)
\]

\[
= \int_{T^*_q\Sigma(z)} (p^T M^{-1} P_M(q) p_1) (p^T M^{-1} P_M(q) p_2) \kappa_{T^*_q\Sigma(z)}^{M^{-1}}(dp).
\]

Denoting \( \langle p_1, p_2 \rangle_{M^{-1}} = p_1^T M^{-1} p_2 \), this yields

\[
p_1^T M^{-1} C M^{-1} p_2 = \int_{T^*_q\Sigma(z)} \langle p, P_M(q) p_1 \rangle_{M^{-1}} \langle p, P_M(q) p_2 \rangle_{M^{-1}} \frac{e^{-\frac{1}{2}(p,p)_{M^{-1}}}}{(2\pi)^{3N/2} \text{det} M^{-1}} \kappa_{T^*_q\Sigma(z)}^{M^{-1}}(dp)
\]

\[
= \beta^{-1} \langle P_M(q) p_1, P_M(q) p_2 \rangle_{M^{-1}},
\]

so that

\[
(4.13) \quad C = \beta^{-1} P_M(q) M.
\]

This gives

\[
\int_{T^*_q\Sigma(z)} \text{Hess}_q(\xi)(M^{-1} p, M^{-1} p) \kappa_{T^*_q\Sigma(z)}^{M^{-1}}(dp) = \beta^{-1} \text{Hess}_q(\xi) : (M^{-1} P_M(q)).
\]

Averaging (3.2) over momenta thus leads to the desired result.

Free energy derivatives can also be obtained from the Lagrange multipliers of the Langevin constrained process (1.3):

\[
\text{(Free energy derivatives)}
\]
Theorem 4.4. Consider the rigidly constrained Langevin process solution of (1.8), with associated Lagrange multipliers $\lambda_t$. Assume that $\nabla \xi$, $G_M^{-1}$ and $\sigma$ are bounded functions on $\Sigma(z)$, and $\gamma_M$ is strictly positive (in the sense of symmetric matrices). Then, the almost sure convergence claimed in the introduction holds:

$$\lim_{T \to +\infty} \frac{1}{T} \int_0^T d\lambda_t = \nabla \xi F^M_{\text{rgd}}(z) \quad \text{a.s.}$$

A similar result holds for the ‘Hamiltonian part’ of the Lagrange multipliers, defined by:

$$d\lambda^\text{ham}_t = d\lambda_t + G_M^{-1} \nabla \xi(q_t) T M^{-1} \left(-\gamma(q_t) M^{-1} p_t dt + \sigma(q_t) dW_t\right) = f_{\text{rgd}}^M(q_t, p_t) dt.$$

Indeed, the following almost sure convergence holds:

$$\lim_{T \to +\infty} \frac{1}{T} \int_0^T d\lambda^\text{ham}_t = \nabla \xi F^M_{\text{rgd}}(z) \quad \text{a.s.}$$

The estimator based on (4.14) has a smaller variance than the estimator based on (1.15), since only the bounded variation part is retained, and the martingale part due to the Brownian increments and the dissipation term are subtracted out. Similar results on variance reduction where obtained in the overdamped case in [6].

Proof. Recall the expression of the Lagrange multipliers, which can be decomposed as the sum of the constraining force, a dissipation term and a martingale (fluctuation) term:

$$d\lambda_t = f_{\text{rgd}}^M(q_t, p_t) dt + G_M^{-1} \nabla \xi(q_t) T M^{-1} \left(-\gamma(q_t) M^{-1} p_t dt - \sigma(q_t) dW_t\right).$$

The result follows from three facts. First, the process is ergodic with respect to the equilibrium distribution $\mu_{T, \Sigma(z)}(dq dp)$ and averaging $f_{\text{rgd}}^M$ yields the rigid free energy derivative in view of Proposition 4.2. This already shows (4.16).

Second, the Gaussian distribution of $\mu_{T, \Sigma(z)}(dq dp)$ with respect to momentum variables is centered:

$$\int_{T \cdot \Sigma(z)} G_M^{-1}(q) \nabla \xi(q) T M^{-1}\gamma(q) M^{-1} p \mu_{T, \Sigma(z)}(dq dp) = 0.$$

Third, the variance estimate can be uniformly bounded as

$$\mathbb{E} \left| \frac{1}{\sqrt{T}} \int_0^T G_M^{-1}(q_t) \nabla \xi(q_t) T M^{-1}\sigma(q_t) dW_t \right|^2 \leq \| \text{Tr}(G_M^{-1} \nabla \xi T M^{-1}\sigma T M^{-1} \nabla \xi G_M^{-1}) \|_{\infty}.$$

This implies the almost sure convergence

$$\lim_{T \to +\infty} \frac{1}{T} \int_0^T G_M^{-1}(q_t) \nabla \xi(q_t) T M^{-1}\sigma(q_t) dW_t = 0,$$

see for example Theorem 1.3.15 in [12].

The fact that averaging the Lagrange multiplier in (4.14) indeed yields the mean force may seem miraculous. In the following remark, we would like to mention another context where the Lagrange multipliers are related to the variation of some quantity with respect to a constraint, in the framework of optimization.

Remark 4.5 (Marginal cost interpretation of the Lagrange multiplier). For simplicity, we assume $m = 1$. Consider the following problem:

$$A(z) = \inf_{q \in \mathbb{R}} \left\{ V(q), \quad \xi(q) = z \right\}.$$

Under suitable assumptions (see for example [4]), it is known that, for a given $z$, the solution $q(z)$ to this constrained optimization problem satisfies the Euler-Lagrange equations: There exists $\lambda(z) \in \mathbb{R}$ such that

$$\begin{cases} 
\nabla V(q(z)) = \lambda(z) \nabla \xi(q(z)), \\
\xi(q(z)) = z,
\end{cases}$$

which yields the rigid free energy derivative in view of (4.14).
and \( A(z) = V(q(z)) \). It can then be checked that
\[
A'(z) = \lambda(z).
\]
Indeed, using the Euler-Lagrange equations (4.18):
\[
A'(z) = \nabla V(q(z)) \cdot q'(z) = \lambda(z)\nabla \xi(q(z)) \cdot q'(z) = \lambda(z) \frac{d}{dz} \left( \xi(q(z)) \right) = \lambda(z).
\]
This basic fact is somewhat related to what we showed in this section, since formally, in the limit of zero temperature, the dynamics (1.8) indeed samples the minimizers of \( V \) under the constraint \( \xi(q) = z \).

4.3. Numerical discretization of the mean force. Estimates of the mean force based on either (4.18), (4.11) or (4.16) can be obtained.

4.3.1. Averaging local rigid mean forces. Free energy derivatives can be computed by averaging \( T_{\text{rgd}}^M(q) \) or \( f_{\text{rgd}}^M(q, p) \) with respect to the distribution \( \mu_T \cdot \Sigma(q, p) \), for instance using the estimators:
\[
\lim_{K \to +\infty} \frac{1}{K} \sum_{k=0}^{K-1} T_{\text{rgd}}^M(q^k)
\]
or
\[
\lim_{K \to +\infty} \frac{1}{K} \sum_{k=0}^{K-1} f_{\text{rgd}}^M(q^k, p^k).
\]
The functions \( T_{\text{rgd}}^M(q) \) and \( f_{\text{rgd}}^M(q, p) \) may thus be called "rigid local mean forces". Note that using the averaged version \( T_{\text{rgd}}^M \) instead of the original \( f_{\text{rgd}}^M \) is likely to reduce the variance since the fluctuation of the momentum variable has been averaged out analytically. See Table 1 below for some numerical results illustrating this claim.

Assuming the convergence of the constrained splitting scheme (3.16) to the limiting Langevin process (1.8), the convergence of these estimators to \( \nabla z F_{\text{rgd}}^M(z) \) is ensured, when taking first the limit \( \Delta t \to 0 \) with \( K = N \Delta t \) such that \( N \Delta t \Delta t \to T \), and then \( T \to \infty \).

4.3.2. Averaging the Lagrange multipliers. Free energy derivatives can also be computed using the Lagrange multipliers of a Langevin constrained process according to (1.15) or (4.10). This technique avoids the possibly cumbersome computation of second order derivatives \( \text{Hess}_q(\xi) \) of the reaction coordinate, which appear in the expressions of \( f_{\text{rgd}}^M \) or \( T_{\text{rgd}}^M \). Again, the computation can be performed with a longtime simulation of the splitting scheme (3.16)–(3.17)–(3.18) discretizing the Langevin process with constraints. The following approximation formula can for instance be used:
\[
(4.19) \quad \nabla z F_{\text{rgd}}^M(z) \simeq \frac{1}{K \Delta t} \sum_{k=0}^{K-1} \left( \lambda^{k+1/2} + \lambda^{k+3/4} \right)
\]
where \( (\lambda^{k+1/2}, \lambda^{k+3/4}) \) are the Lagrange multipliers in the Hamiltonian part (3.17). The consistency of this estimator is given by the following proposition.

**Proposition 4.6 (Consistency).** The approximation formula (4.19) is consistent. More precisely, the Lagrange multipliers \( (\lambda^{n+1/2}, \lambda^{n+3/4}) \) in (3.16)–(3.17)–(3.18) are both equivalent when \( \Delta t \to 0 \) to the constraining force defined in (3.2):
\[
\begin{align*}
\lambda^{n+1/2} &= f_{\text{rgd}}^M(q^n, p^{n+1/2}) \frac{\Delta t}{2} + O(\Delta t^2), \\
\lambda^{n+3/4} &= f_{\text{rgd}}^M(q^{n+1}, p^{n+1/2}) \frac{\Delta t}{2} + O(\Delta t^2).
\end{align*}
\]

\(^2\)This convergence is also called weak convergence in probability theory. The proof of convergence in the present case may be carried out using classical results, see e.g. [13].
Moreover, the following second order consistency holds for the sum of the Lagrange multipliers:

\[ \lambda^{n+1/2} + \lambda^{n+3/4} = \frac{\Delta t}{2} \left( f_{\text{rgd}}^M(q^n, p^{n+1/2}) + f_{\text{rgd}}^M(q^{n+1}, p^{n+3/4}) \right) + O(\Delta t^3), \]

(4.20)

together with the variant:

\[ \lambda^{n+1/2} + \lambda^{n+3/4} = \frac{\Delta t}{2} \left( f_{\text{rgd}}^M(q^n, p^{n+1/2}) + f_{\text{rgd}}^M(q^{n+1}, p^{n+3/4}) \right) + O(\Delta t^3). \]

(4.21)

The variant \[(4.21),\] which involves positions and momenta at the beginning and at the end of the Hamiltonian steps only, is used in \[(4.23)\] below to estimate the time discretization error in the thermodynamic integration method based on the estimator \[(4.16)\].

**Proof.** For sufficiently small time-steps \( \Delta t \), the implicit function theorem ensures that the two projection steps associated with the nonlinear constraints in \[(3.16)-(3.17)-(3.18)\] have a unique smooth solution. A Taylor expansion with respect to \( \Delta t \) of the position constraints gives

\[
z = \xi(q^{n+1}) = \xi(q^n + \Delta t M^{-1} p^{n+1/2})
\]

\[
= \xi(q^n) + \Delta t \nabla \xi(q^n)^T M^{-1} p^{n+1/2} + \frac{\Delta t^2}{2} \text{Hess}_{q^n}(\xi)(M^{-1} p^{n+1/2}, M^{-1} p^{n+1/2})
\]

\[
+ \frac{\Delta t^3}{6} D^3_{\xi}(\xi)(M^{-1} p^{n+1/2}, M^{-1} p^{n+1/2}, M^{-1} p^{n+1/2}) + O(\Delta t^4),
\]

where \( D^3_{\xi}(\xi)(x, y, z) \in \mathbb{R}^m \) denotes the order 3 differential of \( \xi \) computed at \( q \) and evaluated with the vectors \( x, y, z \in \mathbb{R}^{3N} \). We denote

\[
\alpha^{n+1/2}(q) := G_M^{-1}(q) D^3_{\xi}(\xi)(M^{-1} p^{n+1/2}, M^{-1} p^{n+1/2}, M^{-1} p^{n+1/2}).
\]

Then, the fact that \( z = \xi(q^{n+1}) = \xi(q^n) \) and the identity

\[
\nabla \xi(q^n)^T M^{-1} p^{n+1/2} = -\frac{\Delta t}{2} \nabla \xi(q^n)^T M^{-1} \nabla V(q^n) + G_M(q^n) \lambda^{n+1/2}
\]

yield the following expansion of \( \lambda^{n+1/2} \) in terms of \((q^n, p^{n+1/2})\):

\[
\lambda^{n+1/2} = f_{\text{rgd}}^M(q^n, p^{n+1/2}) + \frac{\Delta t^2}{6} \alpha^{n+1/2}(q^n) + O(\Delta t^3).
\]

By time symmetry, the same computation holds for \( \lambda^{n+3/4} \), starting from \((q^{n+1}, p^{n+3/4})\) and by formally replacing \( \Delta t \) by \(-\Delta t\). This can be double checked by Taylor expanding with respect to \( \Delta t \) the position constraints, as done above for \( \lambda^{n+1/2} \). It thus holds:

\[
\lambda^{n+3/4} = f_{\text{rgd}}^M(q^{n+1}, p^{n+3/4}) + \frac{\Delta t^2}{6} \alpha^{n+1/2}(q^{n+1}) + O(\Delta t^3).
\]

The sum of the multipliers therefore reads

\[
\lambda^{n+1/2} + \lambda^{n+3/4} - f_{\text{rgd}}^M(q^n, p^{n+1/2}) - f_{\text{rgd}}^M(q^{n+1}, p^{n+3/4}) = \frac{\Delta t^2}{6} \left( \alpha^{n+1/2}(q^{n+1}) - \alpha^{n+1/2}(q^n) \right) + O(\Delta t^3) = O(\Delta t^3),
\]

which gives \[(4.20)\]. Now, using the previous calculations, we remark that:

\[
\begin{cases}
\quad p^{n+1/2} = p^{n+1/4} - \frac{\Delta t}{2} \nabla V(q^n) + \frac{\Delta t}{2} \nabla \xi(q^n) f_{\text{rgd}}^M(q^n, p^{n+1/2}) + O(\Delta t^2), \\
\quad p^{n+3/4} = p^{n+3/4} + \frac{\Delta t}{2} \nabla V(q^{n+1}) - \frac{\Delta t}{2} \nabla \xi(q^{n+1}) f_{\text{rgd}}^M(q^{n+1}, p^{n+1/2}) + O(\Delta t^2).
\end{cases}
\]
Thus, it holds
\[
\begin{align*}
f^{M}_{\text{rgd}}(q^n, p^{n+1/2}) &+ f^{M}_{\text{rgd}}(q^{n+1/2}, p^{n+1/2}) = f^{M}_{\text{rgd}}(q^n, p^{n+1/4}) + f^{M}_{\text{rgd}}(q^{n+1/4}, p^{n+3/4}) \\
+ \nabla_p f^{M}_{\text{rgd}}(q^n, p^{n+1/4}) &\left(- \frac{\Delta t}{2} \nabla V(q^n) + \frac{\Delta t}{2} \nabla \xi(q^n) f^{M}_{\text{rgd}}(q^n, p^{n+1/2}) \right) \\
- \nabla_p f^{M}_{\text{rgd}}(q^{n+1}, p^{n+3/4}) &\left(- \frac{\Delta t}{2} \nabla V(q^{n+1}) + \frac{\Delta t}{2} \nabla \xi(q^{n+1}) f^{M}_{\text{rgd}}(q^{n+1}, p^{n+1/2}) \right) + O(\Delta t^2) \\
= f^{M}_{\text{rgd}}(q^n, p^{n+1/4}) + f^{M}_{\text{rgd}}(q^{n+1}, p^{n+3/4}) + O(\Delta t^2).
\end{align*}
\]

This gives the claimed second order consistency of the sum of the Lagrange multipliers \(4.21\).

Let us discuss the convergence of the approximation \(4.19\). Assuming again that the constrained splitting scheme \(3.16\)-\(3.17\)-\(3.18\) converges in the probability distribution sense to the strained splitting scheme \(3.16\)-\(3.17\)-\(3.18\) converges in the probability distribution sense to the limiting Langevin process \(1.8\), the following convergence in probability distribution occurs when \(\Delta t \to 0\) and \(N_{\Delta t} \Delta t \to T\):
\[
\lim_{\Delta t \to 0} \text{Law} \left( \frac{1}{N_{\Delta t} \Delta t} \sum_{n=0}^{N_{\Delta t}-1} (\lambda^{n+1/2} + \lambda^{n+3/4}) \right) = \text{Law} \left( \frac{1}{T} \int_0^T d\lambda^\text{ham}_t \right).
\]

This shows the convergence of the estimate \(4.19\) of the mean force when taking first the limit \(\Delta t \to 0\) and then \(T \to \infty\).

4.3.3. Estimates relying on the Metropolized scheme. When the scheme \(3.16\)-\(3.17\)-\(3.18\) is complemented with a Metropolis step (see Algorithm 3.5), it is possible to prove a result on the long-time limit of trajectoryal averages (i.e. letting first the number of iterations go to infinity, and then taking the limit \(\Delta t \to 0\)), upon assuming the irreducibility of the numerical scheme.

Indeed, let us consider the Markov chain \((q^k, p^k)\) generated by the GHMC scheme in Algorithm 3.5, and assume (i) the irreducibility of the Markov chain, and (ii) that appropriate rejections outside the set \(\overline{D}_{\Delta t} = \Sigma(z) \times \{ \frac{1}{2} p^T M^{-1} p \leq R_{\Delta t} \}\) are made in the steps (1)-(2)-(4) of the algorithm. In particular, the projection steps associated with the nonlinear constraints in Step (2) of Algorithm 3.5 are well defined.

Then, by ergodicity, an average of the analytic expression of the local rigid mean force \(\overline{f}^{M}_{\text{rgd}}\) given in \(4.12\) yields an estimate of the free energy without time discretization error:
\[
\lim_{K \to +\infty} \frac{1}{K} \sum_{k=0}^{K-1} \overline{f}^{M}_{\text{rgd}}(q^k) = \nabla_z F^{M}_{\text{rgd}}(z) \quad \text{a.s.}
\]

If \(f^{M}_{\text{rgd}}\) is used instead of \(\overline{f}^{M}_{\text{rgd}}\), then the mean force is computed with some exponentially small error: almost surely,
\[
\lim_{K \to +\infty} \frac{1}{K} \sum_{k=0}^{K-1} f^{M}_{\text{rgd}}(q^k, p^k) = \frac{\int_{\overline{D}_{\Delta t}} f^{M}_{\text{rgd}}(q, p) \mu_T \cdot \Sigma(z) \, dq \, dp}{\int_{\overline{D}_{\Delta t}} \mu_T \cdot \Sigma(z) \, dq \, dp} + O(e^{-\alpha \Delta t^2})
\]
\[
= \frac{\int_{T^* \Sigma(z)} f^{M}_{\text{rgd}}(q, p) \mu_T \cdot \Sigma(z) \, dq \, dp}{\int_{T^* \Sigma(z)} \mu_T \cdot \Sigma(z) \, dq \, dp} + O(e^{-\alpha \Delta t^2})
\]
\[
= \nabla_z F^{M}_{\text{rgd}}(z) + O(e^{-\alpha \Delta t^2})
\]

for some \(\alpha > 0\). The error arising from replacing \(\overline{D}_{\Delta t}\) with \(T^* \Sigma(z)\) is indeed exponentially small in view of \(5.22\) and using the fact that the marginal distribution in the \(p\)-variable is Gaussian.
Likewise, for estimates based on Lagrange multipliers, the following longtime averaging holds: almost surely,

\[
\lim_{K \to +\infty} \frac{1}{K \Delta t} \sum_{k=0}^{K-1} (\lambda^{k+1/2} + \lambda^{k+3/4}) = \int_{\Omega} f^M_{\text{rgd}}(q, p) \mu_{\Sigma(z)}(dq dp) + O(\Delta t^2),
\]

where we have used the estimate (4.21) on the Lagrange multipliers. The limit \(\Delta t \to 0\) is obtained by a dominated convergence argument:

\[
\lim_{\Delta t \to 0} \lim_{K \to +\infty} \frac{1}{K \Delta t} \sum_{k=0}^{K-1} (\lambda^{k+1/2} + \lambda^{k+3/4}) = \nabla_z F^M_{\text{rgd}}(z) \quad \text{a.s.}
\]

Note that, due to the Metropolis correction in Algorithm 3.5, the time discretization error in the sampling of the invariant measure is removed. The only remaining time discretization errors come from (i) the approximation of the Lagrange multipliers (this is a second order error), and (ii) the integration domain being \(\overline{\Omega}_{\Delta t}\) instead of \(T^* \Sigma(z)\) (as discussed above, this is an exponentially small error in \(\Delta t\)). In conclusion, the left-hand side of (4.23) is an approximation of \(\nabla_z F^M_{\text{rgd}}(z)\) up to a \(O(\Delta t^2)\) error term.

4.3.4. Overdamped limit. Finally, let us emphasize that free energy derivatives can be computed with the estimator (4.19) within the overdamped Langevin framework, using the scheme (3.26) and the expressions (5.27), (5.28), (5.29) of Proposition 3.6. Let us recall that the latter are equivalent to the scheme (4.10), (4.11), (4.13) with fluctuation-dissipation matrices satisfying \(\Delta t \gamma = M = 2 \sqrt{t} \text{Id}\). This leads to the original free energy estimator

\[
\nabla_z F^M_{\text{rgd}}(z) \simeq \frac{1}{K \Delta t} \sum_{k=0}^{K-1} (\lambda^{k+1/2} + \lambda^{k+3/4}),
\]

which can be seen as a variant of the variance reduced estimator proposed directly for the overdamped scheme (3.26) in [6]:

\[
\nabla_z F^M_{\text{rgd}}(z) \simeq \frac{1}{K \Delta t} \sum_{k=0}^{K-1} \left( \lambda^k_{od} + \sqrt{\frac{2 \Delta t}{\beta}} G^{-1}(q^k) \nabla \xi(q^k)^T G^k \right) = \frac{1}{K \Delta t} \sum_{k=0}^{K-1} 2 \lambda^{k+1/2}.
\]

The rigorous justification of the consistency of (4.24) in the limit \(\Delta t \to 0\) follows from the results of [6]. See also Section 5.5.2 below for similar results.

4.4. Numerical illustration. We consider a system composed of \(N\) particles in a 2-dimensional periodic box of side length \(L\), interacting through the purely repulsive WCA pair potential, which is a truncated Lennard-Jones potential:

\[
V_{\text{WCA}}(r) = \begin{cases} 
4 \varepsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right] + \varepsilon & \text{if } r \leq r_0, \\
0 & \text{if } r > r_0,
\end{cases}
\]

where \(r\) denotes the distance between two particles, \(\varepsilon\) and \(\sigma\) are two positive parameters and \(r_0 = 2^{1/6} \sigma\). Among these particles, two (numbered 1 and 2 in the following) are designated to form a dimer while the others are solvent particles. Instead of the above WCA potential, the interaction potential between the two particles of the dimer is a double-well potential

\[
V_S(r) = h \left[ 1 - \frac{(r - r_0 - w)^2}{w^2} \right]^2,
\]

where \(h\) and \(w\) are two positive parameters. The total energy of the system is therefore, for \(q \in (LT)^{dN}\) with \(d = 2\),

\[
V(q) = V_S(|q_1 - q_2|) + \sum_{3 \leq i < j \leq N} V_{\text{WCA}}(|q_i - q_j|) + \sum_{i=1, 2} \sum_{3 \leq j \leq N} V_{\text{WCA}}(|q_i - q_j|).
\]

See [10, 34] for instance for other computational studies using this model.
The potential $V_S$ exhibits two energy minima, one corresponding to the compact state where the length of the dimer is $r = r_0$, and one corresponding to the stretched state where this length is $r = r_0 + 2\varepsilon$. The energy barrier separating both states is $h$. The reaction coordinate used to describe the transition from the compact to the stretched state is the normalized bond length of the dimer molecule:

\begin{equation}
\xi(q) = \frac{|q_1 - q_2| - r_0}{2\varepsilon},
\end{equation}

where $q_1$ and $q_2$ are the positions of the two particles forming the dimer. The compact state (resp. the stretched state) corresponds to the value $z = 0$ (resp. $z = 1$) of the reaction coordinate.

The inverse temperature is set to $\beta = 1$, with $N = 100$ particles ($N - 2$ solvent particles and the dimer) with solvent density $\rho = (1 - 2/N)a^{-2} = 0.436$, since there are $N - 2$ solvent particles in a square box of side length $L = a\sqrt{N}$ with $a = 1.5$. The parameters describing the WCA interactions are set to $\sigma = 1$ and $\varepsilon = 1$, and the additional parameters for the dimer are $w = 2$ and $h = 2$.

For this system, $M = \text{Id}$ and $|\nabla \xi|$ is constant, and the rigid free energy $F_{\text{rgd}}^M(z)$ is therefore equal to the free energy $F(z)$.

The mean force is estimated at the values $z_i = z_{\min} + i\Delta z$, with $z_{\min} = -0.2$, $z_{\max} = 1.2$ and $\Delta z = 0.014$, by ergodic averages obtained with the projected dynamics with Metropolis correction (Algorithm 3.5, where in the simple case considered here, the fluctuation-dissipation part can be integrated exactly). For each value of $z$, we integrate the dynamics on a time $T = 2 \times 10^4$ with a step size $\Delta t = 0.02$, using a scalar friction coefficient $\gamma = 1$.

The resulting mean force profile is presented in Figure 1, together with the associated free energy profile. Figure 2 compares the analytical constraining force $f_{\text{rgd}}^M(q^n, p^n)$ and the Lagrange multipliers, see Proposition 4.6. In Figure 2, the $x$-axis represents the blocks of $10^5$ simulation steps, concatenated for the 101 different values of $z_i$.

It can be checked numerically that the differences $|\lambda^{n+1/2} - f_{\text{rgd}}^M(q^n, p^{n+1/2})\Delta t/2|$ and $|\lambda^{n+3/4} - f_{\text{rgd}}^M(q^{n+1}, p^{n+1/2})\Delta t/2|$ are indeed of order $\Delta t^2$, and that the difference

\[|\lambda^{n+1/2} + \lambda^{n+3/4} - f_{\text{rgd}}^M(q^n, p^{n+1/2})\Delta t/2 - f_{\text{rgd}}^M(q^{n+1}, p^{n+1/2})\Delta t/2|\]

is indeed of order $\Delta t^3$ (by computing the average of these elementary differences for various step sizes). The Lagrange multipliers are in any case very good approximations to the constraining force $f_{\text{rgd}}^M$.

Let us finally discuss the efficiency of the different estimators of the mean force, in terms of their variances. They can be written as the empirical average of the following random sequences:

\[\left(f_{\text{rgd}}^M(q^n, p^n), f_{\text{rgd}}^M(q^n, \frac{\lambda^{n+1/2} + \lambda^{n+3/4}}{\Delta t})\right),\]
where \( q^n, p^n, \lambda^{n+1/2}, \lambda^{n+3/4} \) are given by the numerical scheme \([3.16]-[3.17]-[3.18]\). The correlations in time (between the iterates) are very similar for the three methods, and we therefore simply compute the variance over all the samples. Table 1 compares the so-obtained standard errors over \(10^5\) time-steps with \(\Delta t = 0.02\) (simulation time \(T = 2000\) for each value of the reaction coordinate). The results show that the different estimators are more or less equivalent. This is related to the fact that the essential source of variance comes from the sampling of the positions, and not the sampling of the velocities. Note however that, for the smallest value of the reaction coordinate, the estimator based on the averaged local mean force \(\overline{f_{\text{rgd}}^M(q^n)}\) appears to be better in terms of variance.

| \(z\) | \(f_{\text{rgd}}^M(q^n, p^n)\) | \(\frac{\lambda^{n+1/2} + \lambda^{n+3/4}}{\Delta t}\) | \(\overline{f_{\text{rgd}}^M(q^n)}\) |
|---|---|---|---|
| -0.2 | 22.1 | 21.9 | 14.7 |
| 0.0 | 16.0 | 15.5 | 15.4 |
| 0.2 | 23.1 | 22.5 | 22.9 |
| 0.4 | 21.1 | 20.4 | 21.0 |
| 0.6 | 21.4 | 20.7 | 21.3 |
| 0.8 | 21.6 | 20.9 | 21.5 |
| 1.0 | 21.4 | 20.6 | 21.4 |
| 1.2 | 21.0 | 20.3 | 20.9 |

Table 1. Standard error (square-root of the variance) of three mean force estimators, with correlations in time neglected, for different values \(z\) of the reaction coordinate.

5. Hamiltonian and Langevin nonequilibrium dynamics

This section presents nonequilibrium Hamiltonian and Langevin dynamics with time-evolving constraints. We thus consider \((q_t, p_t)\) solution to the dynamics \([1.16]\), which we recall for convenience:

\[
\begin{align*}
    dq_t &= M^{-1} p_t \, dt, \\
    dp_t &= -\nabla V(q_t) \, dt - \gamma p_t M^{-1} p_t \, dt + \sigma_p(q_t) \, dW_t + \nabla \xi(q_t) \, d\lambda_t, \\
    \xi(q_t) &= z(t).
\end{align*}
\]
We prove in particular the fluctuation identity \([1.20]\). Recall (see \([1.17]\)) that, for simplicity, we assume in this section that the fluctuation-dissipation matrices are assumed to be of the form 
\[
(\sigma_P, \gamma_P) = (P_M \sigma, P_M \gamma P_M^T)
\]
with \(\gamma, \sigma \in \mathbb{R}^{3N \times 3N}\). At variance with the previous sections, we do not assume that \(\gamma_P\) is strictly positive. Actually, \(\gamma_P = 0\) corresponds to an interesting case: the Hamiltonian dynamics.

This section is organized as follows. We first define the generalized free energy which is naturally computed with \([1.19]\), and relate it to the standard free energy \([1.12]\) in Section 5.1. Then, we give some precisions on the nonequilibrium dynamics \([1.16]\) in Section 5.2. Next, we prove an appropriate version of the Jarzynski-Crooks fluctuation equality in Section 5.3. A numerical discretization of the nonequilibrium dynamics is proposed in Section 5.4, together with various approximations of the work. In particular, we propose a numerical strategy to obtain a Jarzynski-Crooks identity without time discretization error (see Section 5.4.3). We then consider the overdamped limit when the mass \(M\) goes to 0 (see Section 5.5). Finally, in Section 5.6, we present some numerical results for the model system already considered in Section 4.4.

5.1. Generalized free energy. For \((q_t, p_t)\) solution to the Langevin dynamics \([1.19]\), the reaction coordinate evolution \(\xi(q_t) = z(t)\) implies that \(v_z(q_t, p_t) = \dot{z}(t)\), so that, at each time \(t \geq 0\) the system \((q_t, p_t)\) belongs to the state space \(\Sigma_{\xi,v_z}(z(t), \dot{z}(t))\). As a consequence, the free energy difference computed in this section by the Jarzynski relation without correction (see \([5.21]\) below), is in fact the generalized rigid free energy \(F_{\Xi, \text{rgd}}\) defined in \([4.7]\), in the special case \(\Xi = (\xi, v_z)^T\).

The latter free energy is associated to the normalization \(Z_{z(t), \dot{z}(t)}\) of the distribution \(\mu_{\Sigma_{\xi,v_z}(z(t), \dot{z}(t))}\) defined by \([2.15]\). The generalized rigid free energy \([4.7]\) can be explicitly related to the usual free energy as follows. First, remark that, for a fixed \(q\),
\[
\int_{\Sigma_{\xi,v_z}(v_z)} \exp\left(-\frac{\beta}{2} p^T M^{-1} p\right) \sigma_{\Sigma_{\xi,v_z}(v_z)}^{-1} (dp) = \exp\left(-\frac{\beta}{2} v_z^T G_M^{-1}(q)v_z\right) \int_{T_{z,S}(z)} \exp\left(-\frac{\beta}{2} p^T M^{-1} p\right) \sigma_{T_{z,S}(z)}^{-1} (dp);
\]
\[
= \left(\frac{2\pi\beta}{N}\right)^{-\frac{3N}{2}} \exp\left(-\frac{\beta}{2} v_z^T G_M^{-1}(q)v_z\right).
\]

In the above, the change of variable \(p \rightarrow p - \nabla \xi(q) G_M^{-1}(q)v_z\) has been used, in the space
\[
\Sigma_{\xi,v_z}(v_z) = \left\{ p \in \mathbb{R}^{3N} \mid \nabla \xi(q)^T M^{-1} \left(p - \nabla \xi(q) G_M^{-1}(q)v_z\right) = 0 \right\}.
\]

Note that \(\frac{1}{2} v_z^T G_M^{-1}(q)v_z\) can be interpreted as the kinetic energy of the reaction coordinate \(\xi\). Using the decomposition of measures \([2.24]\) and the above calculations, an alternative expression of the generalized free energy is:
\[
F^{\xi,v_z}_{\Xi, \text{rgd}}(z,v_z) = -\frac{1}{\beta} \ln \int_{T_{z,S}(z)} \exp\left(-\beta V(q) - \frac{\beta}{2} v_z^T G_M^{-1}(q)v_z\right) \sigma_{T_{z,S}(z)}(dq) + C,
\]
where, as usual, \(C\) denotes a generic constant (independent of \(z\)) whose value may vary from line to line. As a consequence, the standard free energy \([1.12]\) is easily recovered from the generalized free energy, using relations similar to \([1.13]\). Indeed, using \([5.1]\), and with computations similar to the ones leading to \([1.22]\), the difference of the two free energies writes:
\[
F(z) - F^{\xi,v_z}_{\Xi, \text{rgd}}(z,v_z) = -\frac{1}{\beta} \ln \int_{\Sigma_{\xi,v_z}(z,v_z)} (\det G_M(q))^{-1/2} \exp\left(\frac{\beta}{2} v_z^T G_M^{-1}(q)v_z\right) \mu_{\Sigma_{\xi,v_z}(z,v_z)}(dq, dp) + C.
\]

In practical nonequilibrium computations, the profile \(t \mapsto F(z(t))\) can then be computed by adding a corrector to the work value in the Jarzynski estimator computing \(F^{\xi,v_z}_{\Xi, \text{rgd}}(z(t), \dot{z}(t))\). This yields the identity \([1.20]\) mentioned in the introduction and proved below (see the discussion after Theorem \([5.3]\)).
5.2. Dynamics and generators. The explicit computation of Lagrange multipliers in (1.16) is obtained similarly to the computation of (5.4) for the case without switching, by differentiating twice the constraints over time:

\[ \frac{d^2}{dt^2} \dot{\xi}(q_t) = \ddot{z}(t). \]

In view of the special structure of \((\sigma_P, \gamma_P)\), this leads to

\begin{align*}
&d\lambda_t = f_{rgd}(q_t, p_t) \, dt + G_M^{-1}(q_t) \dot{\xi}(t) \, dt \\
&\quad + G_M^{-1} \nabla \xi(q_t)^T M^{-1} \left( \gamma_P(q_t) M^{-1} p_t \, dt - \sigma_P(q_t) \, dW_t \right) \\
&= f_{rgd}(q_t, p_t) \, dt + G_M^{-1}(q_t) \dot{\xi}(t) \, dt.
\end{align*}

The expression (5.5) does not depend on the fluctuation-dissipation tensors \((\sigma_P, \gamma_P)\). This will lead to simplified computations and motivates the special form of the latter matrices. The momentum evolution (1.16) thus simplifies as

\[ dp_t = -\nabla V(q_t) \, dt + \nabla \xi(q_t) f_{rgd}(q_t, p_t) \, dt + \nabla \xi(q_t) G_M^{-1}(q_t) \dot{\xi}(t) \, dt - \gamma_P(q_t) M^{-1} p_t \, dt + \sigma_P(q_t) \, dW_t. \]

Let us denote by \( \mathcal{L}_t^b \) the generator of the forward dynamics \( t \mapsto (q_t, p_t) \) defined in (1.16). The latter has a backward switching version,

\[ t' \mapsto (q_{t'}, p_{t'}), \]

obtained by using a time reversed switching \( t' \mapsto z(T - t') \), and by reversing the momentum first in the initial condition, and then reversing them back after the time evolution. More precisely, the backward dynamics can be defined through its generator obtained similarly to the computation of (3.1) for the case without switching, by differentiating twice the constraints over time:

\begin{align*}
&\left\{ dq_{t'}^b = -M^{-1} p_{t'}^b \, dt', \\
&dp_{t'}^b = \nabla V(q_{t'}) \, dt' - \gamma_P(q_{t'}) M^{-1} p_{t'}^b \, dt' + \sigma_P(q_{t'}) \, dW_{t'} + \nabla \xi(q_{t'}) \, d\lambda_{t'}, \\
&\xi(q_{t'}) = z(T - t').
\right.
\end{align*}

In the following proposition, the expressions of \( \mathcal{L}_t^f \) and \( \mathcal{L}_t^b \) are written explicitly.

**Proposition 5.1.** Consider \( \zeta(t) = (z(t), \dot{z}(t)) \). Then, the generator of the forward process (1.16) at time \( t \in [0, T] \) reads:

\[ \mathcal{L}_t^f = \{ \cdot, H \}_\Xi + \mathcal{L}_{\Xi}^{thm} + \{ \cdot, \Xi \} \Gamma^{-1} \dot{\zeta}(t), \]

and the generator of the backward process (5.6) at time \( t' \in [0, T] \) reads:

\[ \mathcal{L}_t^b = -\{ \cdot, H \}_\Xi + \mathcal{L}_{\Xi}^{thm} - \{ \cdot, \Xi \} \Gamma^{-1} \dot{\zeta}(T - t'), \]

where \( \mathcal{L}_{\Xi}^{thm} \) is the fluctuation-dissipation operator defined in (5.4).

**Proof.** First, let us consider the terms in (1.16) arising from the Hamiltonian evolution and from the switching (i.e. without fluctuation-dissipation, which amounts to setting \( \gamma_P = 0 \) and \( \sigma_P = 0 \) in (1.16)). Since during this dynamics \( \nu_\zeta(q_t, p_t) = \dot{z}(t) \), (5.6) yields:

\[ \{ \Xi, H \}(q_t, p_t) = \left( \text{Hess}_{q_t}(\dot{z}(t)) M^{-1} p_t, M^{-1} p_t \right) - \left( \nabla \xi^T M^{-1} \nabla V \right)(q_t), \]

so that, using (5.7),

\[ \Gamma^{-1}(q_t, p_t) \left( \{ \Xi, H \}(q_t, p_t) - \dot{\zeta}(t) \right) = \left( G_M^{-1}(q_t) \ddot{z}(t) + f_{rgd}(q_t, p_t) \right). \]
With (3.9), we then obtain:
\begin{equation}
\{\phi, \Xi\} \Gamma^{-1} \left(\zeta(t) - \{\Xi, H\}\right) (q_t, p_t) = (G_M^{-1}(q_t) \tilde{z}(t) + f_{rigd}(q_t, p_t))^T \nabla q_t T \nabla \phi (q_t, p_t).
\end{equation}

Now, the Hamiltonian part of the switched dynamics (1.10) (see also (5.4)) can be recognized in \(5.10\), so that the generator \(L^f_t\) when \((\gamma_P, \sigma_P) = (0, 0)\) reads: for any smooth test function \(\phi\),
\begin{equation}
\begin{aligned}
L^f_t(\phi) &= (\nabla \xi f_{rigd}^M + \nabla \xi G_M^{-1} \tilde{z}(t))^T \nabla \phi \nabla \phi' - (\nabla V)^T \nabla \phi' + p^T M^{-1} \nabla \phi' \\
&= \{\phi, \Xi\} \Gamma^{-1} \zeta(t) - \{\Xi, H\} + \{\phi, H\} \\
&= \{\phi, H\}_{\Xi} + \{\phi, \Xi\} \Gamma^{-1} \zeta(t).
\end{aligned}
\end{equation}

The full expression of the generator \(L^f_t\) is then obtained by adding the terms arising from the fluctuation-dissipation. These terms are directly obtained from the terms involving \(\gamma_P\) and \(\sigma_P\) in \(5.4\), as in the proof of Proposition 3.1.

The generator of the backward switching process given by (5.6) can be obtained from similar computations. First, the thermostat parts in (5.6) and in (1.16) are the same. Consider now the Hamiltonian part. For example, taking \((\gamma_P, \sigma_P) = (0, 0)\) in the dynamics (5.6). By definition of the backward dynamics, and the expression (5.11) of the forward dynamics, the Hamiltonian part reads
\begin{equation}
\begin{aligned}
L^b_t(\phi)(q, p) &= \mathcal{R} L^f_{T-t'}(\mathcal{R}(\phi))(q, p) \\
&= (\nabla \xi(q) f_{rigd}^M(q, p) + \nabla \xi(q) G_M^{-1}(q, p) \tilde{z}(T - t'))^T (-\nabla p \phi) - \nabla V(q)^T (-\nabla p \phi) - p^T M^{-1} \nabla \phi,
\end{aligned}
\end{equation}
so that
\begin{equation}
\begin{aligned}
L^b_t \phi = -L^f_{T-t'} \phi = -\{\phi, H\}_{\Xi} - \{\phi, \Xi\} \Gamma^{-1} \zeta(T - t') + \{\phi, H\}_{\Xi} + \{\phi, \Xi\} \Gamma^{-1} \zeta(T - t').
\end{aligned}
\end{equation}
This gives \(5.8\).

5.3. Jarzynski-Crooks identity. We define the work \((W_t)_{t \geq 0}\) associated with the constraining force \(\nabla \xi(q_t) \, d\lambda_t\) in \(1.10\) as the physical displacement multiplied by the force:
\begin{equation}
\begin{aligned}
dW_t &:= \left(\frac{dq_t}{dt}\right)^T \nabla \xi(q_t) d\lambda_t = \left(\frac{d\hat{q}_t}{dt}\right)^T \nabla \xi(q_t) \circ d\lambda_t = \hat{z}(t) \circ d\lambda_t \\
&= \hat{z}(T) d\lambda_t.
\end{aligned}
\end{equation}

By convention, \(W_0 = 0\). In the above computations, we used successively the fact that \(t \mapsto \xi(q_t)\), and then \(t \mapsto z(t)\) are differentiable processes, so that Stratonovitch and Itô integrations are equivalent. Let us introduce the deterministic version of the nonequilibrium process (1.10) (i.e. \((\gamma_P, \sigma_P) = (0, 0)\)):
\begin{equation}
\begin{aligned}
d\hat{q}_t &= M^{-1} \hat{p}_t \, dt, \\
d\hat{p}_t &= -\nabla V(q_t) \, dt + \nabla \xi(q_t) \, d\lambda_t, \\
\xi(q_t) &= z(t), \quad (C_q(t))
\end{aligned}
\end{equation}
and denote by \(\Phi_{t, t+h} : \Sigma_{\xi, v_c}(z(t), \tilde{z}(t)) \rightarrow \Sigma_{\xi, v_c}(z(t+h), \tilde{z}(t+h))\) the associated flow between time \(t \in [0, T]\) and \(t+h \in [0, T]\). The work can now be written out more explicitly using the following lemma:

**Lemma 5.2.** The infinitesimal variations of the work (5.12) reads:
\begin{equation}
\begin{aligned}
dW_t = w(t, q_t, p_t) \, dt,
\end{aligned}
\end{equation}
where for all \(t \in [0, T]\) and all \((q, p) \in \Sigma_{\xi, v_c}(z(t), \tilde{z}(t))\),
\begin{equation}
\begin{aligned}
w(t, q, p) &= \hat{z}(t)^T \Gamma^{-1} \{\Xi, H\}(q, p) \\
&= \hat{z}(t)^T (G_M^{-1}(q) \tilde{z}(t) + f_{rigd}^M(q, p)) \\
&= \left. \left(\frac{d}{dh} H \circ \Phi_{t, t+h}\right)\right|_{h=0} (q, p).
\end{aligned}
\end{equation}
The total exchanged work is then a time integral associated with the path \( t \mapsto (q_t, p_t) \), and is denoted by:

\[
W_{0,T} \left( \{q_t, p_t\}_{0 \leq t \leq T} \right) = W_T - W_0 = \int_0^T w(t, q_t, p_t) \, dt.
\]

Note that the expression (5.11) can be interpreted as the energy variation of the system during the switching when the stochastic thermostat is turned off.

**Proof.** The expression of the Lagrange multipliers in (5.3) yields (5.16):

\[
\dot{\lambda}(t) = \lambda(t) \left( G^{-1}_M(q_t) \dot{z}(t) + f^{M}_{rgd}(q_t, p_t) \right) dt.
\]

Moreover, (5.9) gives:

\[
\dot{\lambda}_t = \lambda(t) \left( G^{-1}_M(q_t) \dot{z}(t) + f^{M}_{rgd}(q_t, p_t) \right) dt = \lambda(t) \left( \{ \Xi, H \} (q_t, p_t) - \dot{\lambda}(t) \right)
\]

where in the last line we have used \( \dot{\lambda}(t) \dot{z}(t) = 0 \). This gives (5.14). To prove (5.16), we compute the variations of the energy \( H(\tilde{q}_t, \tilde{p}_t) \) for \( (\tilde{q}_t, \tilde{p}_t) \) solution of (5.13) with initial condition \( (q, p) \):

\[
dH(\tilde{q}_t, \tilde{p}_t) = \tilde{p}_t^T M^{-1} \tilde{d}_p + \tilde{p}_t^T M^{-1} \nabla V(\tilde{q}_t) dt
\]

\[
= \dot{\lambda}(t) \dot{z}(t) dt
\]

\[
= \dot{\lambda}_t \dot{z}(t) dt = \dot{\lambda}_t \dot{z}(t) \left( G^{-1}_M(q_t) \dot{z}(t) + f^{M}_{rgd}(q_t, p_t) \right) dt.
\]

The last equality is obtained using the computation of the Lagrange multipliers in (5.3). This yields (5.16). \( \Box \)

We are now in position to state the main result of this section.

**Theorem 5.3** (Jarzynski-Crooks fluctuation identity). Consider the normalization \( Z(\tilde{z}(t), \dot{z}(t)) \) for the canonical distribution \( \mu_{\Sigma, \dot{z}(t), \dot{z}(t)} \) defined in (2.14). Denote by \( \{q_t, p_t\}_{0 \leq t \leq T} \) the solution of the forward Langevin dynamics (1.10) with initial conditions distributed according to (5.17)

\[
(q_0, p_0) \sim \mu_{\Sigma, \dot{z}(0), \dot{z}(0)}(dq, dp),
\]

and by \( \{q^b_t, p^b_t\}_{0 \leq t \leq T} \) the solution of the backward Langevin process (5.6) with initial conditions distributed according to

\[
(q_0^b, p_0^b) \sim \mu_{\Sigma, \dot{z}(0), \dot{z}(0)}(dq, dp).
\]

Then, the following Jarzynski-Crooks identity holds on \([0, T]\): for any bounded path functional \( \varphi_{[0, T]} \),

\[
\frac{Z_{z(T), \dot{z}(T)}}{Z_{z(0), \dot{z}(0)}} = \frac{\mathbb{E} \left( \varphi_{[0, T]} \left( \{q_t, p_t\}_{0 \leq t \leq T} \right) e^{-\beta W_{0,T}} \left( \{q_t, p_t\}_{0 \leq t \leq T} \right) \right)}{\mathbb{E} \left( \varphi_{[0, T]} \left( \{q^b_t, p^b_t\}_{0 \leq t \leq T} \right) \right)},
\]

where \((\cdot)^r\) denotes the composition with the operation of time reversal of paths:

\[
\varphi^r_{[0, T]} \left( \{q^b_t, p^b_t\}_{0 \leq t \leq T} \right) = \varphi_{[0, T]} \left( \{q^b_{T-t}, p^b_{T-t}\}_{0 \leq t \leq T} \right).
\]

Note that the theorem still holds in the Hamiltonian case, i.e. when \((\gamma, \sigma, \sigma_p) = (0, 0)\). Besides, upon choosing a path functional \( \exp(\vartheta \beta W_{0,T}) \), it is possible to obtain a family of free energy estimators, parameterized by \( \theta \) and where both forward and backward paths are weighted by the exponential of some work. Moreover, the standard Crooks equality on ratios of probability density functions of work values is also a consequence of (5.19), see Section 4.2.2 in [28].

The choice \( \varphi_{[0, T]} = 1 \) in (5.19) leads to the following work fluctuation identity:

\[
F^{E, \nu}_\text{rgd} (z(T), \dot{z}(T)) - F^{E, \nu}_\text{rgd} (z(0), \dot{z}(0)) = -\frac{1}{\beta} \ln \left[ \mathbb{E} \left( e^{-\beta W_{0,T}} \left( \{q, p\}_{0 \leq t \leq T} \right) \right) \right].
\]
Besides, the choice \( \varphi_{[0,T]}(q,p) = \phi(q_T, p_T) \) leads to the following representation of the canonical distribution \( \mu_{\xi,v}(z(T),\dot{z}(T)) \):

\[
(5.22) \quad \frac{\mathbb{E} \left( \phi(q_T, p_T) e^{-\beta W_{o,T} \{ (q_T, p_T) \}_{t \in [0,T]} } \right)}{\mathbb{E} \left( e^{-\beta W_{o,T} \{ (q_T, p_T) \}_{t \in [0,T]} } \right)} = \int_{\Sigma_{\xi,v}(z(T),\dot{z}(T))} \phi(q, p) \mu_{\xi,v}(z(T),\dot{z}(T)) (dq dp).
\]

The usual free energy profile \( z \mapsto F(z) \) can therefore be computed using the relations (1.19), (1.20) presented in the introduction. Indeed, Equation (1.20) can be proved by combining (5.2) and (5.21)–(5.22) as follows:

\[
F(z(T)) - F(z(0)) = \left( F(z(T)) - F_{\text{rgd}}^\xi,\nu(z(T),\dot{z}(T)) \right) - \left( F(z(0)) - F_{\text{rgd}}^\xi,\nu(z(0),\dot{z}(0)) \right)
- \frac{1}{\beta} \ln \mathbb{E} \left( e^{-\beta W_{o,T} \{ (q_T, p_T) \}_{t \in [0,T]} } \right)
- \frac{1}{\beta} \ln \mathbb{E} \left( \left( \det G_M(q_T) \right)^{-1/2} e^{\frac{1}{2} \dot{z}(T)^T G_M(z(T)) \dot{z}(T)} e^{-\beta W_{o,T} \{ (q_T, p_T) \}_{t \in [0,T]} } \right)
= \frac{1}{\beta} \ln \left( \frac{\mathbb{E} \left( e^{-\beta \left[ W_{o,T} \{ (q_T, p_T) \}_{t \in [0,T]} + C(T,q_T) \} \right] } }{\mathbb{E} \left( e^{-\beta \left[ C(0,q_0) \right] } \right) } \right),
\]

where the corrector \( C(t,q) \) is defined in (1.19). Estimators of the free energy based on (1.20) can then be constructed, see Chapter 4 in [28] for a review.

We now give the general lemma which enables to deduce the Jarzynski-Crooks fluctuation identity from a nonequilibrium detailed balance condition.

**Lemma 5.4.** Let \( (q_t,p_t)_{0 \leq t \leq T} \) (resp. \( (q^{b}_t,p^{b}_t)_{0 \leq t \leq T} \)) be a Markov process with infinitesimal generator \( L_t^i \) (resp. \( L_t^b \)) and initial conditions distributed according to (5.17) (resp. (5.18)). Let us assume that the following nonequilibrium detailed balance condition is satisfied: for any two smooth test functions \( \varphi_1, \varphi_2 \),

\[
(5.24) \quad \int_{\Sigma_{\xi,v}(z(t),\dot{z}(t))} \left( \varphi_1 L_t^i(\varphi_2) - \varphi_2 L_t^{i-1}(\varphi_1) \right) e^{-\beta H} d\sigma_{\xi,v}(z(t),\dot{z}(t))
= \int_{\Sigma_{\xi,v}(z(t),\dot{z}(t))} \beta w(t,\cdot) \varphi_1 \varphi_2 e^{-\beta H} d\sigma_{\xi,v}(z(t),\dot{z}(t))
+ \frac{d}{dt} \int_{\Sigma_{\xi,v}(z(t),\dot{z}(t))} \varphi_1 \varphi_2 e^{-\beta H} d\sigma_{\xi,v}(z(t),\dot{z}(t)).
\]

Then the Jarzynski-Crooks fluctuation identity (5.19) holds.

**Proof.** We use in this proof the short-hand notation \( Z_t, \pi_t \) and \( S_t \) for the partition function \( Z(z(t),\dot{z}(t)) \), the (unnormalized) distribution \( Z(z(t),\dot{z}(t)) \mu_{\xi,v}(z(t),\dot{z}(t)) = e^{-\beta H} \sigma_{\xi,v}(z(t),\dot{z}(t)) \) and the submanifold \( \Sigma_{\xi,v}(z(t),\dot{z}(t)) \) respectively.

Let us introduce the following weighted transition operators: for any bounded test function \( \varphi \),

\[
(5.25) \quad P_{t,T}^i(\varphi)(q,p) = \mathbb{E} \left( \varphi(q_T, p_T) e^{-\beta W_{i,T} \{ (q_T, p_T) \}_{s \in [t,T]} } \right) \mid (q_t,p_t) = (q,p),
\]

\[
(5.26) \quad P_{t,T}^b(\varphi)(q,p) = \mathbb{E} \left( \varphi(q^b_T, p^b_T) \mid (q^b_t,p^b_t) = (q,p) \right),
\]

where \( (q_t,p_t)_{0 \leq t \leq T} \) (resp. \( (q^{b}_t,p^{b}_t)_{0 \leq t \leq T} \)) is a Markov process with infinitesimal generator \( L_t^i \) (resp. \( L_t^b \)), and \( W_{o,T} = W_{o,T} - W_{o,t} \).

We assume that these operators are well defined and smooth with respect to time for sufficiently smooth test functions defined in an open neighborhood of \( \Sigma_{\xi,v}(z(t),\dot{z}(t)) \) and \( \Sigma_{\xi,v}(z(t'),\dot{z}(t')) \) respectively (for any \( t,t' \in [0,T] \)).
The transition operators satisfy the following backward Kolmogorov evolution equations:

\[
\begin{align*}
\partial_t P^f_{t,T} &= -\mathcal{L}^f_{t,T} P^f_{t,T} + \beta w(t,\cdot) P^f_{t,T}, \\
\partial_v P^b_{v,T} &= -\mathcal{L}^b_{v,T} P^b_{v,T}, \\
P^{f}_{T,T} &= \text{Id}, \\
P^{b}_{T,T} &= \text{Id}.
\end{align*}
\]

Consider now two test functions \( \varphi_0 \) and \( \varphi_T \). The balance condition (5.24) implies

\[
\frac{d}{dt} \left( \int_{S_t} P^f_{t,T}(\varphi_T) P^b_{T-t,T}(\varphi_0) d\pi_t \right) = 0.
\]

Integrating this equality on \([0,T]\) yields

\[
(5.27) \quad \int_{S_0} P^f_{0,T}(\varphi_T) \varphi_0 \, d\pi_0 = \int_{S_T} \varphi_T P^b_{0,T}(\varphi_0) \, d\pi_T,
\]

which is the Crooks identity (5.19) for path functionals of the form

\[
\varphi_{[0,T]}(q,p) = \varphi_0(q_0,p_0) \varphi_T(q_T,p_T).
\]

Indeed,

\[
\int_{S_0} P^f_{0,T}(\varphi_T) \varphi_0 \, d\pi_0 = Z_0 \mathbb{E}\left[ \varphi_T(q_T,p_T) \varphi_0(q_0,p_0) e^{-\beta \mathcal{W}_{0,T}((q,p),e_{[0,T]})} \right],
\]

while

\[
\int_{S_T} \varphi_T P^b_{0,T}(\varphi_0) \, d\pi_T = Z_T \mathbb{E}\left[ \varphi_T(q_0^b,p_0^b) \varphi_0(q_T^b,p_T^b) \right].
\]

Then, using the Markov property of the forward and backward processes, Crooks identity (5.19) can be extended to finite-dimensional path functionals of the form:

\[
(5.28) \quad \varphi_{[0,T]}(q,p) = \varphi_0(q_0,p_0) \cdots \varphi_k(q_{t_k},p_{t_k}) \cdots \varphi_K(q_T,p_T)
\]

with \( 0 = t_0 < t_1 \cdots < t_K = T \) by repeatedly using a variant of (5.24) on time subintervals \([t_k,t_{k+1}]\). This allows to conclude since finite dimensional time marginal laws characterize the distribution on continuous paths, see for instance [13].

We are now in position to write the

Proof of Theorem 5.3 We prove Theorem 5.3 by proving the nonequilibrium detailed balance (5.24) for the Markov processes \((q_t,p_t)_{0 \leq t \leq T}\) and \((q_T^b,p_T^b)_{0 \leq t \leq T}\) solution to (1.16) and (5.10), with generators \(\mathcal{L}^f_t\) and \(\mathcal{L}^b_t\) defined by (5.7) and (5.8).

First, using Lemma 5.2 we compute the variation of the unnormalized canonical equilibrium distribution with constraints with respect to the switching:

\[
\frac{d}{dt} \left( \int_{\Sigma_x,v \xi(z(t),\dot{z}(t))} \varphi_1 \varphi_2 \, e^{-\beta H} d\sigma_{\Sigma_x,v \xi(z(t),\dot{z}(t))} \right)
\]

(5.29)

\[
= \int_{\Sigma_x,v \xi(z(t),\dot{z}(t))} \dot{\xi}(t)^T \Gamma^{-1} \{ \xi, \varphi_1 \varphi_2 \, e^{-\beta H} \} \, d\sigma_{\Sigma_x,v \xi(z(t),\dot{z}(t))}.
\]

On the other hand, (5.14) and Proposition 5.1 give

\[
\varphi_1 \mathcal{L}^f_t(\varphi_2) - \varphi_2 \mathcal{L}^f_{t-1}(\varphi_1) - \beta w(t,\cdot) \varphi_1 \varphi_2
\]

(5.30)

\[
= \{ \varphi_1 \varphi_2, H \}_\xi + e^{\beta H} \{ \varphi_1 \varphi_2, e^{-\beta H} \}_\xi \Gamma^{-1} \dot{\xi}(t)
\]

\[
+ \varphi_1 \frac{1}{\beta} e^{\beta H} \text{div}_p \left( e^{-\beta H} \gamma_p \nabla_p \varphi_2 \right) - \varphi_2 \frac{1}{\beta} e^{\beta H} \text{div}_p \left( e^{-\beta H} \gamma_p \nabla_p \varphi_1 \right).
\]

Now, (5.24) can be verified in two steps. First, the last two terms in (5.30) (the "thermostat" terms) cancel out after integration with respect to \(e^{-\beta H} d\sigma_{\Sigma_x,v \xi(z(t),\dot{z}(t))}\) thanks to the detailed
balance condition (3.12). Then, an integration of (5.30) with respect to \( e^{-\beta H} \, d\sigma_{\xi,v}(z(t),\dot{z}(t)) \) gives, in view of (5.29) and (2.25),

\[
\int_{\Sigma_{\xi,v}(z(t),\dot{z}(t))} \left( \varphi_1{\mathcal L}^a_\gamma(\varphi_2) - \varphi_2{\mathcal L}^b_{\text{T}^{-1}}(\varphi_1) - \beta w(t,\cdot)\varphi_1\varphi_2 \right) e^{-\beta H} \, d\sigma_{\xi,v}(z(t),\dot{z}(t))
\]

\[
= \int_{\Sigma_{\xi,v}(z(t),\dot{z}(t))} \left\{ \varphi_1\varphi_2 e^{-\beta H}, \Xi \right\} \Gamma^{-1} \dot{\zeta}(t) \, d\sigma_{\xi,v}(z(t),\dot{z}(t))
\]

\[
= \frac{d}{dt} \left( \int_{\Sigma_{\xi,v}(z(t),\dot{z}(t))} \varphi_1\varphi_2 e^{-\beta H} \, d\sigma_{\xi,v}(z(t),\dot{z}(t)) \right),
\]

which is indeed (5.24). Note that the time-regularity on the evolution semi-groups (5.25)-(5.26) required to make these computations rigorous is proved in the overdamped case in the proof of Theorem A.5 in [27]. A similar proof can be carried out for constrained Langevin equations. 

5.4. Numerical schemes. In this section, a numerical scheme for the nonequilibrium dynamics (1.16) and the associated free energy estimator are presented. As for Langevin processes with constraints (Section 3.2), a splitting between the Hamiltonian part and the thermostat part of the dynamics (1.16) leads to a simple and natural scheme (see (5.31)- (5.32) below). Note that a consistent numerical scheme in the case of Hamiltonian dynamics can be obtained by considering only (5.32) (this corresponds to \( \gamma = \sigma = 0 \)). Besides, we propose a discrete Jarzynski-Crooks identity without time discretization error, see Section 5.4.3.

The reaction coordinate path is first discretized as \( \{z(0), \ldots, z(t_N)\} \) where \( t_N \) is the number of time-steps. For simplicity, equal time increments are used, so that \( \Delta t = \frac{T}{t_N} \) and \( t_n = n\Delta t \). The deterministic Hamiltonian part in the equations of motion (1.10) with switched position constraints \( \xi(q) = z(t) \) can be integrated by a velocity-Verlet algorithm with constraints similar to (3.17). The fluctuation-dissipation term in (1.10) can be integrated similarly to the constrained case without switching (3.16)-(3.18), using an Ornstein-Uhlenbeck process on the momentum variable approximated by a midpoint Euler scheme. Finally, the splitting scheme for the Langevin dynamics with time-evolving constraints reads as follows: Take initial conditions \((q^0, p^0)\) distributed according to \( \mu_{\Sigma_{\xi,v}}(z(0),\dot{z}(0)) \) and iterate on \( 0 \leq n \leq t_N - 1 \):

\[
\begin{align*}
\{p_n^{n+1/4} &= p_n - \frac{\Delta t}{4} \gamma_P(q^n) M^{-1} (p_n^{n+1/4} + p^n) + \frac{\sqrt{\Delta t}}{2} \sigma_P(q^n) \mathcal{G}_n, \\
p_n^{n+1/2} &= p_n^{n+1/4} - \frac{\Delta t}{2} \nabla V(q^n) + \nabla \xi(q^n) \lambda_n^{n+1/2}, \\
q_n^{n+1} &= q^n + \Delta t \, M^{-1} p_n^{n+1/2}, \\
\xi(q_n^{n+1}) &= z(t_{n+1}), \\
p_n^{n+3/4} &= p_n^{n+1/2} - \frac{\Delta t}{2} \nabla V(q^{n+1}) + \nabla \xi(q^{n+1}) \lambda_n^{n+3/4}, \\
\nabla \xi(q_n^{n+1})^T M^{-1} p_n^{n+3/4} = & \frac{z(t_{n+2}) - z(t_{n+1})}{\Delta t}, \\
p_n^{n+1} &= p_n^{n+3/4} - \frac{\Delta t}{4} \gamma_P(q^{n+1}) M^{-1} (p_n^{n+3/4} + p^{n+1}) + \frac{\sqrt{\Delta t}}{2} \sigma_P(q^{n+1}) \mathcal{G}_n^{n+1/2},
\end{align*}
\]

where \( (\mathcal{G}_n) \) and \( (\mathcal{G}_n^{n+1/2}) \) are sequences of i.i.d. Gaussian random variables of mean 0 and covariance matrix \( \text{Id}_{3N} \). Note that the momenta obtained from (5.31)-(5.32) satisfy

\[
\nabla \xi(q_n^{n+1})^T M^{-1} p_n^{n+1/4} = \nabla \xi(q^n)^T M^{-1} p^n = \nabla \xi(q^n)^T M^{-1} p^{n-1/4} = \frac{z(t_{n+1}) - z(t_n)}{\Delta t},
\]

so that constraints on momenta are automatically enforced, and no Lagrange multiplier is needed in (5.31) and (5.33).
We comment in the subsequent sections on the different parts of the scheme.

5.4.1. **Comments on the Hamiltonian scheme** \([5.32]\). The Lagrange multipliers \(\lambda^{n+1/2}\) are associated with the position constraints \((C_q)\), and the Lagrange multipliers \(\lambda^{n+3/4}\) are associated with the velocity constraints \((C_p)\). In \((C_p)\), the velocity of the switching at time \(t_{n+1}\) is discretized as:

\[
\dot{z}(t_{n+1}) \simeq \frac{z(t_{n+2}) - z(t_{n+1})}{\Delta t}.
\]

The latter choice is motivated by the following observation: The position after one step of an unconstrained motion, given by

\[
\dot{q}^{n+1} = q^n + \Delta t \, M^{-1} p^{n+1/4} - \frac{\Delta t^2}{2} M^{-1} \nabla V(q^n),
\]

already satisfies \((C_q)\) up to, error terms of order two with respect to \(\Delta t\). Indeed, using \((5.34)\):

\[
\xi(\dot{q}^{n+1}) = \xi(q^n) + \Delta t \nabla \xi(q^n)^T M^{-1} p^{n+1/4} + O(\Delta t^2) = z(t_{n+1}) + O(\Delta t^2).
\]

This property is useful to ensure a fast convergence of the numerical algorithm solving the nonlinear constraints \((C_q)\).

The numerical flow associated with \([5.32]\) is denoted in the sequel as

\[
\Phi^n : \begin{cases}
\Sigma_{\xi,q}(z(t_n), \frac{z(t_{n+1}) - z(t_n)}{\Delta t}) & \rightarrow & \Sigma_{\xi,q}(\frac{z(t_{n+2}) - z(t_{n+1})}{\Delta t}) \\
(q^n, p^{n+1/4}) & \rightarrow & (q^{n+1}, p^{n+3/4})
\end{cases}
\]

It can be proven that \(\Phi^n\) is a symplectic map. The proof is indeed exactly the same as for the symplecticity of the classical RATTLE scheme, see \([16\text{ Sections VII.1.3}]\) for an explicit computation for symplectic Euler and \([16\text{ Sections VII.1.4}]\) for an extension to RATTLE. As a consequence, \(\Phi^n\) transports the phase space measure \(\sigma_{\xi,q}(\frac{z(t_{n+2}) - z(t_{n+1})}{\Delta t})\) to the phase space measure \(\sigma_{\xi,q}(\frac{z(t_{n+1}) - z(t_n)}{\Delta t})\).

5.4.2. **Comments on the fluctuation-dissipation part** \([5.31]-[5.33]\). In practice, \((5.31)\) may be rewritten in a form more suited to numerical computations. Of course, similar considerations hold for \((5.33)\). Since \(\gamma_P(q) = P_M(q)\gamma_P(q)^T\) and \(\sigma_P(q) = P_M(q)\sigma\), \((5.31)\) is indeed equivalent to:

\[
\begin{cases}
p^{n+1/4} = p^n - \Delta t p^{n+1/4} - 2 \nabla \xi G_M^{-1}(q^n) \frac{z(t_{n+1}) - z(t_n)}{\Delta t} \\
+ \sqrt{\frac{\Delta t}{2}} \gamma \sigma^n + \nabla \xi(q^n)^T \lambda^{n+1/4},
\end{cases}
\]

\[
\nabla \xi(q^n)^T M^{-1} p^{n+1/4} = \frac{z(t_{n+1}) - z(t_n)}{\Delta t},
\]

where the Lagrange multiplier \(\lambda^{n+1/4}\) is associated with the constraint \((C_p)\). The equivalence between \((5.31)\) and \((5.33)\) can be checked by multiplying \((5.33)\) by \(P_M(q^n)\) and using \((5.34)\).

The Lagrange multiplier \(\lambda^{n+1/4}\) in \((5.33)\) is obtained by multiplying the above equation by \(\nabla \xi(q^n)^T M^{-1} (\text{Id} + \frac{\Delta t}{4} \gamma M^{-1})^{-1}\), and solving the following linear system:

\[
\frac{z(t_{n+1}) - z(t_n)}{\Delta t} = \nabla \xi(q^n)^T M^{-1} \left( \text{Id} + \frac{\Delta t}{4} \gamma M^{-1} \right)^{-1} \left( \text{Id} - \frac{\Delta t}{4} \gamma M^{-1} \right) p^n \\
+ \nabla \xi(q^n)^T M^{-1} \left( \text{Id} + \frac{\Delta t}{4} \gamma M^{-1} \right)^{-1} \left( \gamma M^{-1} \nabla \xi(q^n) G_M^{-1}(q^n) \frac{z(t_{n+1}) - z(t_n)}{2} + \sqrt{\frac{\Delta t}{2}} \gamma \sigma^n \right) \\
+ \nabla \xi(q^n)^T M^{-1} \left( \text{Id} + \frac{\Delta t}{4} \gamma M^{-1} \right)^{-1} \nabla \xi(q^n) \lambda^{n+1/4}.
\]

This system is well posed. Indeed, the matrix \(\nabla \xi(q^n)^T M^{-1} (\text{Id} + \frac{\Delta t}{4} \gamma M^{-1})^{-1} \nabla \xi(q^n)\) can be rewritten as \(\nabla \xi(q^n)^T S \nabla \xi(q)\) with \(S = M^{-1} (\text{Id} + \frac{\Delta t}{4} \gamma M^{-1})^{-1}\). Both \(M\) and \(\gamma\) are symmetric
and non-negative, so that $S$ is symmetric, positive and invertible. Finally, the invertibility of $\nabla \xi(q)^T S \nabla \xi(q)$ follows from the invertibility of $G_M(q)$.

In the special case when $\gamma$ and $M$ are equal up to a multiplicative constant, the numerical integration can be simplified using the explicit formula (5.20) and the method described below (5.20), which still holds for the tangential part of the momentum. See Section 5.6 below for further precisions.

5.4.3. Discretization of backward process (5.36). The splitting scheme for the backward Langevin dynamics with time-evolving constraints (5.6) reads as follows: Denote $n' = N_T - n$, take initial conditions $(q^{b,0}, p^{b,0})$ distributed according to $\mu_{\xi, v} \left( z(t_{N_T}), \frac{z(t_{N_T+1}) - z(t_{N_T})}{\Delta t} \right)$ and iterate on $0 \leq n' \leq N_T - 1$,

\begin{equation}
\begin{aligned}
\left\{ \begin{array}{l}
p_{b,n'+1/4} = p_{b,n'} - \frac{\Delta t}{4} \gamma_p(q^{b,n'}) M^{-1} (p_{b,n'+1/4} + p_{b,n'}) \\
+ \sqrt{\frac{\Delta t}{2}} \sigma_p(q^{b,n'}) G_b,n',
\end{array} \right.
\end{aligned}
\end{equation}

\begin{equation}
\begin{aligned}
p_{b,n'+1/2} = p_{b,n'+1/2} = \frac{\Delta t}{2} \nabla V(q^{b,n'}) + \nabla \xi(q^{b,n'}) \lambda_b n'+1/2, \\
q^{b,n'+1} = q^{b,n'} - \Delta t M^{-1} p_{b,n'+1/2}, \\
\xi(q^{b,n'+1}) = z(t_{N_T-n'-1}),
\end{aligned}
\end{equation}

\begin{equation}
\begin{aligned}
\nabla \xi(q^{b,n'+1}) T M^{-1} p_{b,n'+3/4} = \frac{z(t_{N_T-n'}) - z(t_{N_T-n'-1})}{\Delta t},
\end{aligned}
\end{equation}

\begin{equation}
\begin{aligned}
\left\{ \begin{array}{l}
p_{b,n'+1} = p_{b,n'+3/4} - \frac{\Delta t}{4} \gamma_p(q^{b,n'+1}) M^{-1} (p_{b,n'+3/4} + p_{b,n'+1}) \\
+ \sqrt{\frac{\Delta t}{2}} \sigma_p(q^{b,n'+1}) G_b,n'+1/2,
\end{array} \right.
\end{aligned}
\end{equation}

where $(G^{b,n'})$ and $(G^{b,n'+1/2})$ are sequences of i.i.d. Gaussian random variables of mean 0 and covariance matrix $\text{Id}_{3N}$. The numerical flow associated with (5.38) is denoted

\begin{equation}
\Phi^{b,n'} : \Sigma_{\xi, v} \left( z(t_n), \frac{z(t_{n+1}) - z(t_n)}{\Delta t} \right) \rightarrow \Sigma_{\xi, v} \left( z(t_{n-1}), \frac{z(t_n) - z(t_{n-1})}{\Delta t} \right),
\end{equation}

where we recall $n' = N_T - n$. Assuming that the flow $\Phi^n$ given by (5.38) and $\Phi^{b,n'}$ are both well-defined, the following reversibility property is easily checked (extending the symmetry property of the standard RATTLE scheme, see for instance [16, Section VII.1.4]):

\begin{equation}
\Phi^{b,N_T-n} \circ \Phi^{n-1} = \text{Id}.
\end{equation}

5.4.4. Work discretization and free energy computations. The work (5.12) can be approximated using the Lagrange multipliers in (5.30):

\begin{equation}
\begin{aligned}
\mathcal{W}^0 = 0, \\
\mathcal{W}^{n+1} = \mathcal{W}^n + \left( \frac{z(t_{n+1}) - z(t_n)}{\Delta t} \right)^T \left( \lambda^{n+1/2} + \lambda^{n+3/4} \right),
\end{aligned}
\end{equation}

for $n = 0 \ldots N_T - 1$. The (formal) consistency of the work discretization (5.43) in the time continuous limit is a direct consequence of the work expression (5.12).

An estimator of the free energy profile is then obtained by using $K$ independent realizations of the switching process, computing the work $\mathcal{W}_{N_T,k}^N$ for each realization $k \in \{1, \ldots, K\}$, and
approximating \([5.23]\), rewritten up to an unimportant additive constant (independent of \(T\)), as
\[
F(z(T)) = -\frac{1}{\beta} \ln \mathbb{E} \left( e^{-\beta [\mathcal{W}^{N_T} + C^{N_T}(q^{N_T})]} \right),
\]
with empirical averages such as
\[
-\frac{1}{\beta} \ln \left( \frac{1}{K} \sum_{k=1}^{K} \exp \left[ -\beta \left( \mathcal{W}^{N_T,k} + C^{N_T}(q^{N_T,k}) \right) \right] \right).
\]
In the above, the discretization \(C^n(q)\) of the corrector \([1.19]\) is
\[
C^n(q) = \frac{1}{2\beta} \ln \left( \det G_M(q) - \frac{1}{2} \left( \frac{2}{\Delta t} \int_{t_n}^{t_{n+1}} z(t_{n+1}) - z(t_n) \right) G_M^{-1}(q) \left( \frac{z(t_{n+1}) - z(t_n)}{\Delta t} \right) \right).
\]
We refer to Chapter 4 in \([28]\) for more background on free energy estimators for nonequilibrium dynamics. In particular, it is possible to compute a work associated with the backward switching using approximations such as \([5.41]\) in the Jarzynski-Crooks identity introduces a time discretization error. We show in the next section how to eliminate this error.

### 5.4.5. Discrete Jarzynski-Crooks identity

It turns out that a discrete version of the Jarzynski-Crooks identity \([5.19]\) can be obtained. This enables estimation of free energy differences using nonequilibrium simulation \textit{without time discretization error}. The discrete equality \([5.47]\) below may be seen as an extension of the corresponding equality obtained for alchemical transitions performed with Metropolis-Hastings dynamics (see \([17]\) and Remark 4.5 in \([28]\)).

For this purpose, we consider a discretization of the work \(\mathcal{W}_{0,T}\) using the interpretation \([5.16]\) of the work as the energy variation of the Hamiltonian part of the Langevin dynamics. This leads to the following definition of the work at the discrete level:

\[
\mathcal{W}^n = 0, \quad \mathcal{W}^{n+1} = \mathcal{W}^n + H(q^{n+1}, p^{n+3/4}) - H(q^n, p^{n+1/4}),
\]
for \(n = 0 \ldots N_T - 1\). This work discretization leads to a Jarzynski-Crooks identity without time discretization error.

**Theorem 5.5** (Discrete Jarzynski-Crooks fluctuation identity). Consider the distribution \(\mu_{\Sigma_{t,v}\{z(t),\dot{z}(t)\}}\) and its normalization \(Z_{z(t),\dot{z}(t)}\) defined in \([2.13]\). Denote by \(\{q^n, p^n\}_{0 \leq n \leq N_T}\) the solution of the forward discretized Langevin dynamics \([5.31]-[5.33]\) with initial conditions distributed according to

\[
(q^0, p^0) \sim \mu_{\Sigma_{t,v}\{z(t_0), \frac{z(t_1) - z(t_0)}{\Delta t}\}}(dq\,dp),
\]
and by \(\{q^b, p^b\}_{0 \leq n' \leq N_T}\) the solution of the discretized backward Langevin dynamics \([5.37]-[5.39]\) distributed according to

\[
(q^{b,0}, p^{b,0}) \sim \mu_{\Sigma_{t,v}\{z(t_{N_T}), \frac{z(t_{N_T+1}) - z(t_{N_T})}{\Delta t}\}}(dq\,dp).
\]

Then, the following Jarzynski-Crooks identity holds on \([0, N_T]\): for any bounded discrete path functional \(\varphi_{[0,N_T]}\),

\[
\frac{Z_{z(t_{N_T}), \frac{z(t_{N_T+1}) - z(t_{N_T})}{\Delta t}}}{Z_{z(t_0), \frac{z(t_1) - z(t_0)}{\Delta t}}} = \mathbb{E} \left( \varphi_{[0,N_T]}(\{q^n, p^n\}_{0 \leq n \leq N_T}) e^{-\beta \mathcal{W}^n} \right) / \mathbb{E} \left( \varphi^{r}_{[0,N_T]}(\{q^{b,n'}, p^{b,n'}\}_{0 \leq n' \leq N_T}) \right),
\]
where \(\mathcal{W}^n\) is computed according to \([5.43]\), and \((\cdot)^r\) denotes the composition with the operation of time reversal of paths:

\[
\varphi^{r}_{[0,N_T]}(\{q^{b,n'}, p^{b,n'}\}_{0 \leq n' \leq N_T}) = \varphi_{[0,N_T]}(\{q^{b,N_T-n}, p^{b,N_T-n}\}_{0 \leq n \leq N_T}).
\]
The (formal) consistency of the work discretization (5.41) in the time continuous limit is a direct consequence of the work expression (5.10). Free energy estimators based on the identity (5.40) are obtained as described in Section 5.4.4. Let us emphasize once again that there is no error related to the finiteness of the time-step Δτ in this estimator, and that the only source of approximation is due to the statistical error.

**Proof.** With a slight abuse of notation, we denote in the same way the random variables \((q^n, p^n)\), \((q^{b,n}, p^{b,n})\), etc. in (5.31)-(5.32)-(5.33) or (5.37)-(5.38)-(5.39), and the integration variables in the definition of probability distributions. We divide the proof into three steps.

**Step 1:** The phase space conservation of \(Φ^n\) and \(Φ^{b,n'}\) and the reversibility property (5.40) imply

\[
\delta_{Φ^n(q^n,p^{n+1/4})}(dq^{n+1} dp^{n+3/4}) \sigma_{\Sigma_{\xi,v}(\epsilon(t_n),\frac{z(t_{n+1})-z(t_n)}{\Delta t})} (dq^n dp^{n+1/4})
\]

\[
= \delta_{Φ^{b,n'}(q^{n+1/4},p^{n+1/4})}(dq^{n+1} dp^{n+3/4}) \sigma_{\Sigma_{\xi,v}(\epsilon(t_{n+1}),\frac{z(t_{n+2})-z(t_{n+1})}{\Delta t})} (dq^{n+1} dp^{n+3/4}),
\]

where \(p_\parallel = P_M(q^n)p\) and \(p_\perp = (\text{Id} - P_M(q^n))p\). The Markov chain induced by the parallel part of the momentum is the same as the one induced by the scheme (3.10) (or (3.18)) defined in Section 3.2. The latter verifies a detailed balance equation (both in the plain sense and up to momentum reversal) with respect to the stationary measure \(K_{T,Σ(z)\Sigma(z)}^{-1}(dp)\) defined by (3.19) (see Sections 2.3.2 and 3.3.5 in [28]). We recall that this measure is defined as the kinetic probability distribution in the momentum variable of the canonical distribution \(μ_{T,Σ(z)}(dq dp)\) on the tangential space, conditioned by a given \(q \in Σ(z)\). Adding the (invariant) orthogonal part of momentum, the following detailed balance condition is satisfied:

\[
\exp \left( -\frac{β}{2} (p^n)^T M^{-1} p^n \right) K_{OU}(q^n, p^n, dp^{n+1/4}) \sigma_{M^{-1} \Sigma_{v}(q^n)}(\frac{z(t_{n+1})-z(t_n)}{\Delta t}) (dp^n)
\]

\[
= \exp \left( -\frac{β}{2} (p^{n+1/4})^T M^{-1} p^{n+1/4} \right) K_{OU}(q^{n+1/4}, p^{n+1/4}, dp^{n+1/4}) \sigma_{M^{-1} \Sigma_{v}(q^n)}(\frac{z(t_{n+2})-z(t_{n+1})}{\Delta t}) (dp^{n+1/4}).
\]

**Step 3:** Denote by \(K^f(q^n, p^n; dq^{n+1}, dp^{n+1/4}, dp^{n+3/4}, dp^{n+1})\) the probability distribution of the variables \((q^{n+1}, p^{n+1/4}, p^{n+3/4}, p^{n+1})\) given the variables \((q^n, p^n)\) in the scheme (5.31)-(5.32)-(5.33); and by \(K^b(q^n, p^n; dq^{b,n'+1}, dp^{b,n'+1/4}, dp^{b,n'+3/4}, dp^{b,n'+1})\) the probability distribution of the variables \((q^{b,n'+1}, p^{b,n'+1/4}, p^{b,n'+3/4}, p^{b,n'+1})\) given the variables \((q^{b,n'}, p^{b,n'})\) in the scheme (5.37)-(5.38)-(6.39). The splitting structure yields:

\[
K^f(q^n, p^n; dq^{n+1}, dp^{n+1/4}, dp^{n+3/4}, dp^{n+1}) = K_{OU}(q^{n+1}, p^{n+3/4}, dp^{n+1}) \delta_{Φ^n(q^n,p^{n+1/4})}(dq^{n+1} dp^{n+3/4}) K_{OU}(q^n, p^n, dp^{n+1/4}),
\]

as well as

\[
K^b(q^{b,n'}, p^{b,n'}; dq^{b,n'+1}, dp^{b,n'+1/4}, dp^{b,n'+3/4}, dp^{b,n'+1}) = K_{OU}(q^{b,n'+1}, p^{b,n'+3/4}, dp^{b,n'+1}) \delta_{Φ^{b,n'}(q^{b,n'},p^{b,n'+1/4})}(dq^{b,n'+1} dp^{b,n'+3/4}) \times K_{OU}(q^{b,n'}, p^{b,n'}, dp^{b,n'+1/4}).
\]
Combining the detailed balance conditions (5.48) and (5.50) of Steps 1 and 2, and using the decomposition (2.24) of phase space measures, it follows
\[
e^{-\beta(H(q^{n+1},p^{n+1/4})-H(q^n,p^{n+1/4}))}K_{t,n}(q^n,p^n; dq^{n+1} dp^{n+3/4} dp^{n+1})
\times e^{-\beta H(q^n,p^n) \sigma_{\xi,v_t} \left(z(t_n) \frac{z(t_{n+1})-z(t_n)}{\Delta t}\right)}(dq^n dp^n)
= K^{b,N_T-n-1}(q^{n+1},p^{n+1}; dq^n dp^{n+3/4} dp^{n+1/4} dp^n)
\times e^{-\beta H(q^{n+1},p^{n+1}) \sigma_{\xi,v_t} \left(z(t_{n+1}) \frac{z(t_{n+2})-z(t_{n+1})}{\Delta t}\right)}(dq^{n+1} dp^{n+1}),
\]
which can be seen as the Jarzynski-Crooks identity over one time-step. Iterating the argument, it is easy to obtain:
\[
e^{-\beta W^{N_T}} K^{t,0}(q^0,p^0; dq^1 dp^{1/4} dp^{3/4} dp^1) \cdots K^{t,N_T-1}(q^{N_T-1},p^{N_T-1}; dq^{N_T} dp^{N_T-3/4} dp^{N_T-1/4} dp^{N_T})
\times e^{-\beta H(q^0,p^0) \sigma_{\xi,v_t} \left(z(t_0) \frac{z(t_1)-z(t_0)}{\Delta t}\right)}(dq^0 dp^0)
= K^{b,N_T-1}(q^1,p^1; dq^0 dp^{3/4} dp^{1/4} dp^0) \cdots K^{b,0}(q^{N_T},p^{N_T}; dq^{N_T-1} dp^{N_T-1/4} dp^{N_T-3/4} dp^{N_T-1})
\times e^{-\beta H(q^{N_T},p^{N_T}) \sigma_{\xi,v_t} \left(z(t_{N_T}) \frac{z(t_{N_T+1})-z(t_{N_T})}{\Delta t}\right)}(dq^{N_T} dp^{N_T}),
\]
which yields (5.44).

5.5. The overdamped limit.

5.5.1. An exact free energy estimator for the overdamped Langevin dynamics. The splitting scheme (5.31)-(5.33) can be performed with an overdamped scheme, using the method of Proposition 5.6, i.e. by choosing
\[
\frac{\Delta t}{4} \gamma = M = \frac{\Delta t}{2} \text{Id},
\]
which implies \( \gamma_P = 2P_T^T P_M \) and \( \sigma_P = \frac{\sigma}{\sqrt{\beta}} P_M \). For this choice of parameters, the continuous limit of the numerical scheme is the following variant of the stochastic differential equation (3.23):
\[
\begin{cases}
dq_t = -\nabla V(q_t) dt + \sqrt{\frac{2}{\beta}} dW_t + \nabla \xi(q_t) d\lambda_{t}^{od}, \\
\xi(q_t) = z(t),
\end{cases}
\]
where \( \lambda_{t}^{od} \) is an adapted stochastic process such that \( \xi(q_t) = z(t) \). We then obtain the following Jarzynski-Crooks relation for discretized overdamped dynamics, without time discretization error.

**Proposition 5.6.** Suppose that the relation (5.51) is satisfied. With a slight abuse of notation, the mass matrix and the friction matrix are rewritten as \( M \text{Id} \) and \( \gamma \text{Id} \) with \( M, \gamma \in \mathbb{R} \). Then the splitting scheme (5.31)-(5.32)-(5.33) yields the following Euler discretization of the overdamped Langevin constrained dynamics (5.52):
\[
\begin{cases}
q^{n+1} = q^n - \Delta t \nabla V(q^n) + \sqrt{\frac{2 \Delta t}{\beta}} G^n + \nabla \xi(q^n) \lambda_{n+1}^{od}, \\
\xi(q^{n+1}) = z(t_{n+1}),
\end{cases}
\]
where \( (G^n)_{n \geq 0} \) are independent and identically distributed centered and normalized Gaussian variables, and \( (\lambda_{n}^{od})_{n \geq 1} \) are the Lagrange multipliers associated with the constraints \( \xi(q^n) = z(t_n) \) for \( 0 \leq n \leq N_T \). In the same way, the backward process (5.37)-(5.38)-(5.39) yields the following Euler scheme:
\[
\begin{cases}
q^{b,n+1} = q^{b,n} - \Delta t \nabla V(q^{b,n}) + \sqrt{\frac{2 \Delta t}{\beta}} G^{b,n} + \nabla \xi(q^{b,n}) \lambda_{n+1}^{od}, \\
\xi(q^{b,n+1}) = z(t_{N_T-n-1}).
\end{cases}
\]
Consider the work update

\begin{equation}
W^{n+1} = W^n + V(q^{n+1}) - V(q^n) + \frac{1}{\Delta t} \left( |p^{n+3/4}|^2 - |p^{n+1/4}|^2 \right),
\end{equation}

for \( n = 0 \ldots N_T - 1 \), where

\begin{align*}
2 p^{n+1/4} &= \sqrt{\frac{2 \Delta t}{\beta}} P(q^n) G^n + \nabla \xi(q^n) G^{-1}(q^n) (z(t_{n+1}) - z(t_n)), \\
2 \lambda^{n+1/2} &= \lambda_{od}^{n+1} - G^{-1}(q^n) (z(t_{n+1}) - z(t_n)) + \sqrt{\frac{2 \Delta t}{\beta}} G^{-1}(q^n) \nabla \xi(q^n)^T G^n,
\end{align*}

with \( G = \nabla \xi^T \nabla \xi \), and the scheme (5.53) is rewritten as:

\begin{align*}
p^{n+1/2} &= p^{n+1/2} - \frac{\Delta t}{2} \nabla V(q^n) + \nabla \xi(q^n) \lambda^{n+1/2}, \\
q^{n+1} &= q^n + 2 p^{n+1/2}, \\
\xi(q^{n+1}) &= z(t_{n+1}), \quad (C_q) \\
p^{n+3/4} &= p^{n+3/4} - \frac{\Delta t}{2} \nabla V(q^{n+1}) + \nabla \xi(q^{n+1}) \lambda^{n+3/4}, \\
\nabla \xi(q^{n+1}) p^{n+3/4} &= \frac{z(t_{n+2}) - z(t_{n+1})}{2}, \quad (C_p)
\end{align*}

Then the Jarzynski-Crooks relation (5.40) holds under the assumptions (5.44) and (5.45) on the initial conditions of the schemes (5.53) and (5.54) respectively.

The proof is a direct consequence of the reformulation of (5.53) into (5.56), and a direct application of Theorem 5.3 with the parameters (5.59).

Note that the free energy estimator

\begin{equation}
F(z(T)) = -\frac{1}{\beta} \ln \mathbb{E} \left( e^{-\beta [W^{N_T} + C^{N_T}(q^{N_T})]} \right),
\end{equation}

based on the work (5.55) and the corrector

\begin{equation}
C^n(q) = \frac{1}{2\beta} \ln \left( \det G(q) \right) - \frac{\Delta t}{4} \left( \frac{z(t_{n+1}) - z(t_n)}{\Delta t} \right)^T G^{-1}(q) \left( \frac{z(t_{n+1}) - z(t_n)}{\Delta t} \right)
\end{equation}

is exact (there is no time discretization error). This free energy estimator can be seen as a variant of the estimator proposed in [27], which was derived directly for the scheme (5.53), and reads (up to an unimportant additive constant):

\begin{equation}
F(z(T)) \simeq -\frac{1}{\beta} \ln \mathbb{E} \left( e^{-\beta [\tilde{W}^{N_T} + \tilde{C}(q^{N_T})]} \right),
\end{equation}

where the work is defined as

\begin{equation}
\begin{align*}
\tilde{W}^0 &= 0, \\
\tilde{W}^{n+1} &= \tilde{W}^n + \left( \frac{z(t_{n+1}) - z(t_n)}{\Delta t} \right)^T \tilde{\lambda}^{n+1}_{od},
\end{align*}
\end{equation}

with

\begin{align*}
\tilde{\lambda}^{n+1}_{od} &= 2 \lambda^{n+1/2} = \lambda^{n+1}_{od} - G^{-1}(q^n) (z(t_{n+1}) - z(t_n)) + \sqrt{\frac{2 \Delta t}{\beta}} G^{-1}(q^n) \nabla \xi(q^n)^T G^n,
\end{align*}

and the modified corrector is defined without the kinetic energy term:

\begin{equation}
\tilde{C}(q) = \frac{1}{2\beta} \ln \left( \det G(q) \right).
\end{equation}

There is a bias due to the time discretization error in the estimator (5.59).
5.5.2. Consistency analysis of three free energy estimators. In this section, we would like to discuss the consistency of three free energy estimators introduced above: (5.56)–(5.61) (based on the direct discretization of the overdamped dynamics proposed in [27]), (5.57)–(5.61) (which uses the Lagrange multipliers to approximate the work) and (5.59)–(5.60) (based on the discrete Jarzynski equality).

The limiting continuous-in-time version of the Jarzynski relation is:

\[ F(z(T)) = -\frac{1}{\beta} \ln \mathbb{E} \left( e^{-\beta W_{T}^{\text{od}}(\{q_{t}\}_{0 \leq t \leq T})} \right), \]

where the work for the overdamped dynamics (5.52) reads (see [27]):

\[ W_{0,T}^{\text{od}}(\{q_{t}\}_{0 \leq t \leq T}) = \int_{0}^{T} \dot{z}(t)^{T} d\lambda_{t}^{\text{od}}, \]

with

\[ d\lambda_{t}^{\text{od}} = d\lambda_{t}^{\text{od}} - G^{-1}(q_{t})z'(t) dt + \sqrt{\frac{2}{\beta}} G^{-1}(q_{t})\nabla \xi(q_{t})^{T} dW_{t}. \]

The consistency of (5.59)–(5.60) with (5.62)–(5.63) was already proven in [27].

Concerning the consistency of \( C^{n} \) with \( \bar{C} \) (see (5.57) and (5.59)), note that in the overdamped scaling \( (M = \frac{\Delta t}{2} \text{Id}) \), the difference

\[ \bar{C}(q) - C^{n}(q) = \frac{\Delta t}{4} \left( \frac{z(t_{n+1}) - z(t_{n})}{\Delta t} \right)^{T} G^{-1}(q) \left( \frac{z(t_{n+1}) - z(t_{n})}{\Delta t} \right) = O(\Delta t) \]

vanishes when \( \Delta t \to 0 \). This difference can therefore be neglected when analyzing the consistency of the scheme in the continuous-in-time limit. We henceforth concentrate on the consistency of the works (5.41) and (5.53) with (5.63).

In the sequel, we denote the anticipating stochastic integration of the integrand \( Y_{t} \) with respect to \( dX_{t} \) by \( Y_{t} \circ dX_{t} = Y_{t} \circ dX_{t} - Y_{t} dX_{t} \), where \( \circ \) is the Stratonovitch symmetric integration, and \( d \) the Itô integration. The symbol \( \rightsquigarrow \) denotes the formal time continuous limit.

Consistency of (5.41). Let us justify the consistency of the work expression (5.41) with (5.63). Remark that the Lagrange multipliers in (5.59) verify:

\[ 2\lambda^{n+1/2} = G^{-1}(q^{n}) \left[ \nabla \xi(q^{n})^{T} (q^{n+1} - q^{n}) - (z(t_{n+1}) - z(t_{n})) + \Delta t \nabla \xi(q^{n})^{T} \nabla V(q^{n}) \right] \]

and

\[ 2\lambda^{n+1/2} = \lambda^{n+1}_{\text{od}} - G^{-1}(q^{n})(z(t_{n+1}) - z(t_{n})) + \sqrt{\frac{2\Delta t}{\beta}} G^{-1}(q^{n})\nabla \xi(q^{n})^{T} G^{n}, \]

as well as

\[ 2\lambda^{n+3/4} = G^{-1}(q^{n+1}) \left[ \nabla \xi(q^{n+1})^{T} (q^{n} - q^{n+1}) + (z(t_{n+2}) - z(t_{n+1})) + \Delta t \nabla \xi(q^{n+1})^{T} \nabla V(q^{n+1}) \right]. \]

The expressions (5.64) and (5.66) yield

\[ 2\lambda^{n+1/2} \rightsquigarrow G^{-1}(q_{t}) \left( \nabla \xi(q_{t})^{T} .dq_{t} - z'(t) dt + \nabla \xi(q_{t})^{T} \nabla V(q_{t}) dt \right), \]

as well as

\[ 2\lambda^{n+3/4} \rightsquigarrow G^{-1}(q_{t}) \left( -\nabla \xi(q_{t})^{T} .dq_{t} + z'(t) dt + \nabla \xi(q_{t})^{T} \nabla V(q_{t}) dt \right). \]

Moreover the constraints imply that

\[ d\xi(q_{t}) = z'(t) dt = \nabla \xi^{T}(q_{t}) \circ dq_{t} = \frac{1}{2} (\nabla \xi^{T}(q_{t}) .dq_{t} + \nabla \xi^{T}(q_{t}) dq_{t}) \]

so that \( \lambda^{n+1/2} \) and \( \lambda^{n+3/4} \) yield the same time continuous limit, that is to say

\[ 2\lambda^{n+3/4} \rightsquigarrow G^{-1}(q_{t}) \left( \nabla \xi(q_{t})^{T} .dq_{t} - z'(t) dt + \nabla \xi(q_{t})^{T} \nabla V(q_{t}) dt \right). \]

Eventually, (5.65) implies

\[ 2\lambda^{n+1/2} \rightsquigarrow d\lambda_{t}^{\text{od}}. \]
the same holding true for $2\lambda^{n+3/4}$. The work expression (5.41) is thus formally consistent with (5.63). This concludes the proof of the consistency of (5.57), (5.41) with (5.62) (5.63).

Consistency of (5.56). We now prove the consistency of the work expression (5.56) with (5.63). Let us denote $f^n = -\frac{\Delta t}{2} \nabla V(q^n) + \nabla \xi(q^n)\lambda^{n+1/2}$ and $f^{n+1} = -\frac{\Delta t}{2} \nabla V(q^{n+1}) + \nabla \xi(q^{n+1})\lambda^{n+3/4}$.

The expression (5.56) yields using (5.56):

$$W^{n+1} - W^n = V(q^{n+1}) - V(q^n) + \frac{1}{\Delta t} \left( |p^{n+1/2} + f^{n+1}|^2 - |p^{n+1/2} - f^n|^2 \right)$$

$$= V(q^{n+1}) - V(q^n) + \frac{1}{\Delta t} (f^n + f^{n+1}) \cdot (q^{n+1} - q^n - f^n + f^{n+1})$$

$$= V(q^{n+1}) - V(q^n) - \frac{1}{2}(\nabla V(q^n) + \nabla V(q^{n+1})) \cdot (q^{n+1} - q^n)$$

$$+ I^n + \frac{1}{\Delta t} (f^n + f^{n+1}) \cdot (f^{n+1} - f^n),$$

where

$$I^n = \frac{1}{\Delta t} \left( \nabla \xi(q^n)\lambda^{n+1/2} + \nabla \xi(q^{n+1})\lambda^{n+3/4} \right) \cdot (q^{n+1} - q^n).$$

First, since $V(q^{n+1}) - V(q^n) \sim \nabla V(q^n) \circ dq_t$ and $\frac{1}{2}(\nabla V(q^n) + \nabla V(q^{n+1})) \cdot (q^{n+1} - q^n) \sim -\nabla V(q^n) \circ dq_t$, the limit of the terms in (5.71) is zero. Second, using the expressions (5.63) and (5.66), and similarly to (5.67) and (5.69), it holds (5.73)

$$f^n - f^{n+1} = (\nabla \xi G^{-1})(q^n) \left[ \nabla \xi(q^n)^T (q^{n+1} - q^n) - (z(t_{n+1}) - z(t_n)) + \Delta t \nabla \xi(q^{n+1})^T \nabla V(q^n) \right]$$

$$- (\nabla \xi G^{-1})(q^{n+1}) \left[ \nabla \xi(q^{n+1})^T (q^n - q^{n+1}) + (z(t_{n+2}) - z(t_{n+1})) + \Delta t \nabla \xi(q^{n+1})^T \nabla V(q^{n+1}) \right]$$

$$+ \frac{\Delta t}{2}(\nabla V(q^{n+1}) - \nabla V(q^n)) = o(\Delta t)$$

since $(\nabla \xi G^{-1} \nabla \xi(q^n)) + (\nabla \xi G^{-1} \nabla \xi(q^{n+1}))^T (q^{n+1} - q^n) \sim 2\nabla \xi G^{-1}(q(t))z'(t) dt$ by (5.68). Expanding in higher order powers of $\Delta t$, it can be checked that there exists two functions $a$ and $b$ such that

$$\frac{f^n - f^{n+1}}{\Delta t} \sim a(t, q_t) dt + b(t, q_t) dq_t.$$}

Therefore, since (in the limit $\Delta t \to 0$) the martingale part of $f^n + f^{n+1}$ arises only from the term $\nabla \xi G^{-1} \nabla \xi^T(q_t).dq_t$, one obtains

$$(f^n + f^{n+1}) \cdot \frac{f^n - f^{n+1}}{\Delta t} \sim d \left\{ \int_0^t \nabla \xi G^{-1} \nabla \xi^T(q_t).dq_t, \int_0^t b(t, q_t).dq_t \right\}_t$$

$$= \frac{2}{\beta} d \left\{ \int_0^t \nabla \xi G^{-1} \nabla \xi^T(q_t)P(q_t).dW_t, \int_0^t b(t, q_t)P(q_t).dW_t \right\}_t = 0$$

since $\nabla \xi^T(q_t)P(q_t) = 0$. In conclusion, the second term in (5.72) has a zero contribution to the continuous-in-time limit.

As a consequence, the formal time continuous limit of $W^{n+1} - W^n$ is the same as the one of $I^n$. Computations similar to the one performed above yield

$$J^n = \frac{1}{2\Delta t} (q^{n+1} - q^n)^T (\nabla \xi(q^{n+1}) - \nabla \xi(q^n)) (\lambda^{n+1/2} - \lambda^{n+3/4}) = o(\Delta t).$$

Indeed, $\lambda^{n+1/2} - \lambda^{n+3/4} = o(\Delta t)$ as in (5.73), while $(q^{n+1} - q^n)^T (\nabla \xi(q^{n+1}) - \nabla \xi(q^n)) = O(\Delta t)$. The formal time continuous limit of $I^n$ is therefore the same as the limit of

$$I^n + J^n = \frac{1}{2\Delta t} (q^{n+1} - q^n)^T (\nabla \xi(q^n) + \nabla \xi(q^{n+1})) \left( \lambda^{n+1/2} + \lambda^{n+3/4} \right).$$

Since (5.68) implies

$$\left( \nabla \xi(q^n) + \nabla \xi(q^{n+1}) \right)^T (q^{n+1} - q^n) \sim z'(t) dt,$$
we get in the end that the formal time continuous limit of $I^n$ and $W^{n+1} - W^n$ is the same as:

$$z'(t_n) T \left( \lambda^{n+1/2} + \lambda^{n+3/4} \right) \sim z'(t) T d\tilde{\lambda}_t^d,$$

where we have used (5.70). This concludes the proof of the consistency of (5.57)-(5.55) with (5.62)-(5.63).

5.6. Numerical illustration. We present some free energy profiles obtained with nonequilibrium switching dynamics for the model system and the parameters described in Section 4.4. The switching schedule reads

$$z(t) = z_{\text{min}} + (z_{\text{max}} - z_{\text{min}}) \frac{t}{T},$$

with $z_{\text{min}} = -0.1$ and $z_{\text{max}} = 1.1$. The time-step is $\Delta t = 0.01$. The initial conditions are obtained by first subsampling a constrained dynamics with $\xi(q) = z_{\text{min}}$ and $v_\xi(q, p) = 0$, with a time spacing $T_{\text{sample}} = 1$; and then adding the required component $\nabla \xi(q) G^{-1}_M \dot{z}(0)$ to the momentum variable (with $\dot{z}(0) = (z_{\text{max}} - z_{\text{min}})/T$).

In the specific case at hand, the corrector term (1.19) is constant, and free energies differences are equal to differences of rigid free energies. The dynamics used to integrate the nonequilibrium dynamics is based on a splitting strategy, analogous to (5.31)-(5.32)-(5.33), except that the midpoint integration of the Ornstein-Uhlenbeck part is replaced by an exact integration for the unconstrained dynamics, followed by a projection. This can be done here since we choose a friction matrix of the form $\gamma \text{Id}$ (recall also that $M = \text{Id}$). More precisely, the corresponding scheme is obtained by replacing (5.31) (and likewise for (5.33)) with

$$\tilde{p}^{n+1/4} = \alpha p^n + \sqrt{\frac{1 - \alpha^2}{\beta}} G^n,$$

where $\alpha = e^{-\gamma \Delta t}$, and $p^{n+1/4} = \tilde{p}^{n+1/4} + \lambda^{n+1/4} \nabla \xi(q^n)$ with $\lambda^{n+1/4}$ chosen such that

$$\nabla \xi(q^n)^T M^{-1} p^{n+1/4} = \frac{z(t_{n+1}) - z(t_n)}{\Delta t}.$$

Figure 3 presents estimates obtained with $M$ independent realizations of the switching dynamics for different switching times $T$, using the estimator presented in Section 5.4.4 with the work discretization (5.41). In all cases, the product $MT$ is kept constant. The free energy profile becomes closer to the reference curve as $T$ is decreased, and the profile obtained for $T = 100$ is in excellent agreement with the result obtained with thermodynamic integration. When the switching time is small, more realizations should be considered to reduce the statistical errors (bias and variance) and obtain estimates in better agreement with the reference profile. Of course,
more reliable profiles could be obtained with backward switchings and Crooks relation, see \[30\] and Section 4.2 in \[28\].

6. Appendix: Some Technical Results

We give in this appendix two technical lemmas, used in the proof of Proposition \[4.2\]. The first lemma can also be used to prove the divergence formula \[2.25\].

Lemma 6.1. For any \( a \in \{1, \ldots, 2m\} \):

\[
\sum_{b=1}^{2m} \left\{ |\det \Gamma|^{1/2} (\Gamma^{-1})_{a,b}, \Xi_b \right\} = 0,
\]

where \( \Gamma \) is defined in \[2.20\].

Proof. The proof relies on the following computation rules for any family of invertible square matrices \( \theta \mapsto A_\theta \):

\[
\frac{d}{d\theta} \left( \ln |\det A_\theta| \right) = \text{tr} \left( A_\theta^{-1} \frac{d}{d\theta} A_\theta \right),
\]

and

\[
A_\theta \frac{d}{d\theta} (A_\theta^{-1}) = -\left( \frac{d}{d\theta} A_\theta \right) A_\theta^{-1}.
\]

Fix \( a \in \{1, \ldots, 2m\} \). First, using \(6.3\) with \( A_\theta \) replaced by \( \Gamma \) and \( \frac{d}{d\theta} \) replaced by \( \{\cdot, \Xi_c\} \), we obtain

\[
\sum_{b,c=1}^{2m} \Gamma_{a,b} \left\{ (\Gamma^{-1})_{b,c}, \Xi_c \right\} = -\sum_{b,c=1}^{2m} \left\{ \Gamma_{a,b}, \Xi_c \right\} (\Gamma^{-1})_{b,c},
\]

so that by the skew-symmetry of \( \Gamma^{-1} \) and \( \Gamma \),

\[
\sum_{b,c=1}^{2m} \Gamma_{a,b} \left\{ (\Gamma^{-1})_{b,c}, \Xi_c \right\} = \sum_{b,c=1}^{2m} -\frac{1}{2} \left( \{ \Gamma_{a,b}, \Xi_c \} + \{ \Gamma_{c,a}, \Xi_b \} \right) (\Gamma^{-1})_{b,c}.
\]

Jacobi’s identity for Poisson brackets and \(6.2\) then yield

\[
\sum_{b,c=1}^{2m} \Gamma_{a,b} \left\{ (\Gamma^{-1})_{b,c}, \Xi_c \right\} = \frac{1}{2} \sum_{b,c=1}^{2m} \left\{ \{ \Xi_b, \Xi_c \}, \Xi_a \right\} (\Gamma^{-1})_{b,c} = -\frac{1}{2} \sum_{b,c=1}^{2m} \left\{ \Gamma_{c,b}, \Xi_a \right\} (\Gamma^{-1})_{b,c}
\]

\[
= -\frac{1}{2} \left\{ \ln |\det \Gamma|, \Xi_a \right\} = -|\det \Gamma|^{-1/2} \left\{ |\det \Gamma|^{1/2}, \Xi_a \right\}
\]

\[
= -\sum_{b,c=1}^{2m} |\det \Gamma|^{-1/2} \Gamma_{a,b} (\Gamma^{-1})_{b,c} \left\{ |\det \Gamma|^{1/2}, \Xi_c \right\}
\]

since \( \Gamma_{a,b} (\Gamma^{-1})_{b,c} = \delta_{a,c} \), where \( \delta_{i,j} \) is the Kronecker symbol. Finally, the left hand and right hand sides of the last equality can be factorized as

\[
\sum_{b,c=1}^{2m} |\det \Gamma|^{-1/2} \Gamma_{a,b} \left\{ |\det \Gamma|^{1/2} (\Gamma^{-1})_{b,c}, \Xi_c \right\} = 0.
\]

Since \( |\det \Gamma| > 0 \) and \( \Gamma \) is invertible, it follows

\[
\sum_{c=1}^{2m} \left\{ |\det \Gamma|^{1/2} (\Gamma^{-1})_{b,c}, \Xi_c \right\} = 0
\]

for all \( b = 1, \ldots, 2m \), which is \(6.1\). □
Lemma 6.2. For any compactly supported smooth test function \( \varphi \) on \( \mathbb{R}^{6N} \):

\[
\nabla_\zeta \left( \int_{\Sigma_\Xi(\zeta)} \varphi \, d\sigma_{\Sigma_\Xi(\zeta)} \right) = \int_{\Sigma_\Xi(\zeta)} \Gamma^{-1} \{ \Xi, \varphi \} \, d\sigma_{\Sigma_\Xi(\zeta)},
\]

where the phase space \( \Sigma_\Xi(\zeta) \) is defined in (2.19), and the Gram matrix \( \Gamma \) in (2.19).

Proof. Consider a test function \( \phi : \mathbb{R}^{2m} \to \mathbb{R} \). An integration by parts and the co-area formula (2.10) give:

\[
I := \int_{\mathbb{R}^{2m}} \phi(\zeta) \nabla_\zeta \left( \int_{\Sigma_\Xi(\zeta)} \varphi(q, p) \sigma_{\Sigma_\Xi(\zeta)}(dq, dp) \right) \, d\zeta
\]

\[
= - \int_{\mathbb{R}^{2m}} \nabla_\zeta \phi(\zeta) \left( \int_{\Sigma_\Xi(\zeta)} \varphi(q, p) \sigma_{\Sigma_\Xi(\zeta)}(dq, dp) \right) \, d\zeta
\]

\[
= - \int_{\mathbb{R}^{6N}} \Gamma^{-1} \{ \Xi, \phi \circ \Xi \} \varphi \det(\Gamma)^{1/2} \, dq, dp,
\]

where in the last line the following chain rule has been used:

\[
\{ \Xi, \phi \circ \Xi \}(q, p) = \{ \Xi, \Xi \}(q, p) \nabla_\zeta \phi(\Xi(q, p)) = \Gamma(q, p) \nabla_\zeta \phi(\Xi(q, p)).
\]

Now an integration by parts with respect to \( dq, dp \), together with (6.1), leads to

\[
I = \sum_{b=1}^{2m} \int_{\mathbb{R}^{6N}} \phi \circ \Xi \left( \Xi_b, \det(\Gamma)^{1/2} \Gamma^{-1} \nabla_\zeta \varphi \right) \, dq, dp
\]

\[
= \int_{\mathbb{R}^{6N}} \phi \circ \Xi \Gamma^{-1} \{ \Xi, \varphi \} \det(\Gamma)^{1/2} \, dq, dp
\]

\[
= \int_{\mathbb{R}^{2m}} \phi(\zeta) \left( \int_{\Sigma_\Xi(\zeta)} \Gamma^{-1} \{ \Xi, \varphi \} \, d\sigma_{\Sigma_\Xi(\zeta)} \right) \, d\zeta,
\]

which gives the result. \( \square \)

References

[1] E. Akhmatskaya and S. Reich, GSHMC: an efficient method for molecular simulation, J. Comput. Phys. 227 (2008), no. 10, 4934–4954.
[2] L. Ambrosio, N. Fusco, and D. Pallara, Functions of bounded variation and free discontinuity problems, Oxford Science Publications, 2000.
[3] V. I. Arnol’d, Mathematical methods of classical mechanics, Graduate Texts in Mathematics, vol. 60, Springer-Verlag, 1989.
[4] J. F. Bonnans, J. C. Gilbert, C. Lemaréchal, and C. A. Sagastizábal, Numerical optimization, Springer, 2002.
[5] N. Bou-Rabee and H. Owhadi, Long-run behavior of variational integrators in the stochastic context, SIAM J. Numer. Anal. 48 (2010), 278–297.
[6] G. Ciccotti, T. Lelivre, and E. Vanden-Eijnden, Projection of diffusion on submanifolds: Application to mean force computation, Commun. Pure Appl. Math. 61 (2008), no. 3, 371–408.
[7] G. E. Crooks, Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems, J. Stat. Phys. 90 (1998), no. 5, 1481–1487.
[8] , Entropy production fluctuation theorem and the nonequilibrium work relation for free-energy differences, Phys. Rev. E 60 (1999), no. 3, 2721–2726.
[9] E. Darve, Thermodynamic integration using constrained and unconstrained dynamics, Free Energy Calculations (C. Chipot and A. Pohorille, eds.), Springer, 2007, pp. 119–170.
[10] C. Dellago, P. G. Bolhuis, and D. Chandler, On the calculation of reaction rate constants in the transition path ensemble, J. Chem. Phys. 110 (1999), no. 14, 6617–6625.
[11] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Hybrid Monte-Carlo, Phys. Lett. B 195 (1987), no. 2, 216–222.
[12] M. Duflo, Random iterative models, Springer, 1997.
[13] S. N. Ethier and T. G. Kurtz, Markov processes: Characterization and convergence, Wiley Series in Probability and Statistics, John Wiley & Sons, 1986.
[14] L. C. Evans and R. F. Gariepy, Measure theory and fine properties of functions, Studies in Advanced Mathematics, CRC Press, 1992.
[15] M. Fixman, Simulation of polymer dynamics. I. General theory, J. Chem. Phys. 69 (1978), 1527–1537.
[16] E. Hairer, C. Lubich, and G. Wanner, Geometric numerical integration: Structure-preserving algorithms for ordinary differential equations, Springer Series in Computational Mathematics, vol. 31, Springer-Verlag, 2006.

[17] C. Hartmann, An ergodic sampling scheme for constrained Hamiltonian systems with applications to molecular dynamics, J. Stat. Phys. 130 (2008), no. 4, 687–711.

[18] C. Hartmann and C. Schütte, A constrained Hybrid Monte Carlo algorithm and the problem of calculating the free energy in several variables, Z. Angew. Math. Mech. 85 (2005), no. 10, 709–719.

[19] , A geometric approach to constrained molecular dynamics and free energy, Commun. Math. Sci. 3 (2005), no. 1, 1–20.

[20] L. Hörmander, Hypoelliptic second order differential equations, Acta Math. 119 (1967), 147–171.

[21] A. M. Horowitz, A generalized guided Monte Carlo method, Phys. Lett. B 268 (1991), 247–252.

[22] C. Jarzynski, Nonequilibrium equality for free energy differences, Phys. Rev. Lett. 78 (1997), no. 14, 2690–2693.

[23] W. Kliemann, Recurrence and invariant measures for degenerate diffusions, Ann. Probab. 15 (1987), no. 2, 690–707.

[24] J. Latorre, C. Hartmann, and Ch. Schütte, Free energy computation by controlled langevin processes, 2010, to appear in Procedia Computer Science.

[25] B. J. Leimkuhler and S. Reich, Simulating hamiltonian dynamics, Cambridge Monographs on Applied and Computational Mathematics, vol. 14, Cambridge University Press, 2005.

[26] B. J. Leimkuhler and R. D. Skeel, Symplectic numerical integrators in constrained Hamiltonian systems, J. Comput. Phys. 112 (1994), no. 1, 117–125.

[27] T. Lelièvre, M. Rousset, and G. Stoltz, Computation of free energy differences through nonequilibrium stochastic dynamics: The reaction coordinate case, J. Comput. Phys. 222 (2007), no. 2, 624–643.

[28] , Free energy computations. A mathematical perspective, Imperial College Press, 2010.

[29] D. D. L. Minh and A. B. Adib, Optimized free energies from bidirectional single-molecule force spectroscopy, Phys. Rev. Lett. 100 (2008), 180602.

[30] D. C. Rapaport, The art of molecular dynamics simulations, Cambridge University Press, 1995.

[31] S. Reich, Smoothened Langevin dynamics of highly oscillatory systems, Physica D 138 (2000), 210–224.

[32] T. Schlick, Molecular modeling and simulation, Springer, 2002.

[33] J. E. Straub, M. Borkovec, and B. J. Berne, Molecular-dynamics study of an isomerizing diatomic in a Lennard-Jones fluid, J. Chem. Phys. 89 (1988), no. 8, 4833–4847.

[34] E. Vanden-Eijnden and G. Ciccotti, Second-order integrators for Langevin equations with holonomic constraints, Chem. Phys. Lett. 429 (2006), no. 1-3, 310–316.

Université Paris Est, CERMICS and INRIA, MICMAC project-team Ecole des Ponts ParisTech, 6 & 8 Av. Pascal, 77455 Marne-la-Vallée, France
E-mail address: lelievre@cermics.enpc.fr

INRIA Lille - Nord Europe, Parc Scientifique de la Haute Borne, 40 avenue Halley, Bt.A Park Plaza, 59650 Villeneuve d’Ascq, France
E-mail address: mathias.rousset@inria.fr

Université Paris Est, CERMICS and INRIA, MICMAC project-team Ecole des Ponts ParisTech, 6 & 8 Av. Pascal, 77455 Marne-la-Vallée, France
E-mail address: stoltz@cermics.enpc.fr