Nosé-Hoover dynamics in a shaker

Erwan Faou
INRIA Rennes, Campus Beaulieu
F-35042 Rennes Cedex, France
Erwan.Faou@irisa.fr

We introduce a new class of systems based on the Nosé-Hoover equations. We show that we can add time-dependent terms without destroying the measure and energy conservation properties of the initial system. These "shakers" are typically pseudoperiodic in time, i.e. depend on a collection of harmonic oscillators. We show by numerical examples that it strengthens the sampling properties of the initial system with respect to the Gibbs measure, and helps the computation of averages in the canonical ensemble.

I. INTRODUCTION

Molecular simulations at constant temperature can be performed by using the Nosé extended-Lagrangian method[12,13], and its real-time formulation due to Hoover[5]. These systems are continuous, preserve the canonical Boltzmann-Gibbs measure, conserve an energy, but suffer from a lack of ergodicity for small or stiff systems (see[4,8,10,13]). To remedy this deficiency, Martyna, Klein and Tuckerman introduced the Nosé-Hoover chains (NHC) method where thermostats are added to the system[10]. If this method effectively improves the sampling properties of the system in some cases (see[4,10]), it still exhibits metastability or non ergodicity phenomenons for bad choices of the parameters (see the numerical examples below). Recently, a lot of progress has been made to generalize this method: see in particular the Gaussian Moment Thermostatting Method of Liu and Tuckermann, see[9], the Generalized Dynamical Thermostat Technique of Laird and Leimkuhler, see[7] and the work of Branka et al.[2].

The aim of this paper is to show that we can introduce time-dependent terms in the Nosé-Hoover equation without destroying the measure and energy conservation. The role of these terms is to break the possible hidden invariants of the system, and to reinforce its chaotic behavior. We call them "shakers" as they are independent of the dynamics itself but preserve the advantage of the original Nosé-Hoover approach. We show that it is possible to adapt the existing algorithms to integrate these systems. By numerical examples, we show that the introduction of shakers improves the sampling properties of the systems, and make them produce correct distributions. Moreover, we note that this method is rather general and can be applied to other situations (in particular for microcanonical sampling).

In this paper, we consider a system of N particles in a space of dimension d, of masses

m_i, positions $q_i \in \mathbb{R}^d$ and impulses $p_i \in \mathbb{R}^d$, $i = 1, \ldots, N$ interacting in a potential $V(\boldsymbol{q})$ at constant temperature T, in the framework defined by Nosé and Hoover. We denote by $\boldsymbol{q}$ and $\boldsymbol{p}$ the vectors $(q_i)_{i=1}^N$ and $(p_i)_{i=1}^N$ respectively. The classical Nosé-Hoover equations are written

$$\begin{aligned} \dot{\boldsymbol{q}} &= M^{-1}\boldsymbol{p} \\ \dot{\boldsymbol{p}} &= -\partial_{\boldsymbol{q}} V(\boldsymbol{q}) - \lambda \boldsymbol{p} \\ \dot{\lambda} &= \frac{1}{Q}\left(\boldsymbol{p}^T M^{-1}\boldsymbol{p} - N_f kT\right), \end{aligned} \tag{I.1}$$

where $N_f = d \times N$ is the number of degrees of freedom of the system. Here, $\partial_{\boldsymbol{q}} V(\boldsymbol{q})$ is the N_f-dimensional vector with components $\partial_{q_i} V(\boldsymbol{q})$. The matrix M is the diagonal matrix with coefficients m_i. The constant k is the Boltzmann constant, and T denotes the temperature. The number Q is a free parameter of the problem. It is well known that this system preserves the (extended) Boltzmann-Gibbs measure, see[4,5,12],

$$\exp\left(-\frac{1}{kT}\left(\frac{1}{2}\boldsymbol{p}^T M^{-1}\boldsymbol{p} + V(\boldsymbol{q}) + \frac{Q\lambda^2}{2}\right)\right) \mathrm{d}\boldsymbol{q}\, \mathrm{d}\boldsymbol{p}\, \mathrm{d}\lambda. \tag{I.2}$$

If we add the equation

$$\dot{\xi} = \lambda \tag{I.3}$$

to the equations (I.1), we get the following system, written in matrix form:

$$\begin{pmatrix} \dot{\boldsymbol{q}} \\ \dot{\xi} \\ \dot{\boldsymbol{p}} \\ \dot{\lambda} \end{pmatrix} = \begin{pmatrix} 0 & 0 & \mathrm{Id} & 0 \\ 0 & 0 & 0 & 1/Q \\ -\mathrm{Id} & 0 & 0 & -\boldsymbol{p}/Q \\ 0 & -1/Q & \boldsymbol{p}^T/Q & 0 \end{pmatrix} \begin{pmatrix} \partial_{\boldsymbol{q}} V(\boldsymbol{q}) \\ N_f kT \\ M^{-1}\boldsymbol{p} \\ Q\lambda \end{pmatrix}. \tag{I.4}$$

As the matrix in the right-hand side of (I.4) is skew-symmetric, we easily see that the system (I.1)-(II.1) conserves the energy

$$H(\boldsymbol{q}, \xi, \boldsymbol{p}, \lambda) = \frac{1}{2}\boldsymbol{p}^T M^{-1}\boldsymbol{p} + V(\boldsymbol{q}) + \frac{Q\lambda^2}{2} + N_f kT\xi. \tag{I.5}$$

For small or stiff systems, numerical examples indicate that the system (I.1) is in general not ergodic for the measure (I.2): see in particular[4,8,10] and the numerical examples below.

II. DESCRIPTION OF THE METHOD

Let $A(t) = (A_{ij}(t))_{i,j=1}^{N_f}$ be a time dependent $N_f \times N_f$ matrix and $\alpha(t) = (\alpha_i(t))_{i=1}^{N_f}$ be a time dependent vector of size N_f. We consider the following equations:

$$\begin{aligned} \dot{\boldsymbol{q}} &= A(t)M^{-1}\boldsymbol{p} + Q\alpha(t)\lambda \\ \dot{\boldsymbol{p}} &= -A(t)^T\partial_{\boldsymbol{q}}V(\boldsymbol{q}) - \lambda\boldsymbol{p} \\ \dot{\lambda} &= \tfrac{1}{Q}\left(\boldsymbol{p}^TM^{-1}\boldsymbol{p} - N_fkT\right) - \alpha(t)^T\partial_{\boldsymbol{q}}V(\boldsymbol{q}). \end{aligned} \tag{II.1}$$

The case where $A(t) = \mathrm{Id}$ and $\alpha(t) = 0$ corresponds to the standard Nosé-Hoover equations. In the following, we assume that the applications $t \mapsto A(t)$ and $t \mapsto \alpha(t)$ are smooth (and in particular, "deterministic"), to ensure the local existence and uniqueness of a smooth solution to (II.1). The matrix $A(t)$ and the vector $\alpha(t)$ are called *shakers*, as they can change strongly the dynamics without breaking the original measure and energy conservation of the system as we now show: For a given $t_0 \in \mathbb{R}$ and $\boldsymbol{y}_0 = (\boldsymbol{q}_0, \boldsymbol{p}_0, \lambda_0) \in \mathbb{R}^{N_f} \times \mathbb{R}^{N_f} \times \mathbb{R}$, we write $\Gamma(t, t_0, \boldsymbol{y}_0) = (\boldsymbol{q}(t), \boldsymbol{p}(t), \lambda(t))$ the solution of (II.1) satisfying $\boldsymbol{q}(t_0) = \boldsymbol{q}_0$, $\boldsymbol{p}(t_0) = \boldsymbol{p}_0$ and $\lambda(t_0) = \lambda_0$. The main result is the following:

Theorem II.1 *The measure* (I.2) *is invariant by the flow* $\Gamma(t, t_0, \cdot)$. *Moreover, if we add the equation* (I.3) *to the system* (II.1) *the energy* (I.5) *is conserved along the solution of* (I.3)-(II.1).

Proof. If we write the system (II.1) as $\dot{\boldsymbol{y}} = F(t, \boldsymbol{y})$, we compute easily that

$$\forall\, t \in \mathbb{R}, \quad \forall\, \boldsymbol{y} = (\boldsymbol{q}, \boldsymbol{p}, \lambda) \in \mathbb{R}^{N_f} \times \mathbb{R}^{N_f} \times \mathbb{R}, \quad \mathrm{div}_{\boldsymbol{y}} F(t, \boldsymbol{y}) = -N_f\lambda.$$

Let $G(\boldsymbol{q}, \boldsymbol{p}, \lambda) = \frac{1}{2}\boldsymbol{p}^TM^{-1}\boldsymbol{p} + V(\boldsymbol{q}) + \frac{1}{2}Q\lambda^2$. We have that

$\forall\, t, \quad \forall\, \boldsymbol{y} = (\boldsymbol{q}, \boldsymbol{p}, \lambda)$,

$$\mathrm{div}_{\boldsymbol{y}}(e^{-G(\boldsymbol{y})/kT}F(t, \boldsymbol{y})) = \frac{1}{kT}e^{-G(\boldsymbol{y})/kT}\left(-\nabla G(\boldsymbol{y})^TF(t, \boldsymbol{y}) + kT\mathrm{div}_{\boldsymbol{y}}F(t, \boldsymbol{y})\right). \tag{II.2}$$

Besides

$$\nabla G(\boldsymbol{y}) = \begin{pmatrix} \partial_{\boldsymbol{q}}V(\boldsymbol{q}) \\ M^{-1}\boldsymbol{p} \\ Q\lambda \end{pmatrix}$$

and hence

$$\nabla G(\boldsymbol{y})^T F(t,\boldsymbol{y}) = \partial_{\boldsymbol{q}} V(\boldsymbol{q})^T A(t) M^{-1}\boldsymbol{p} + Q\lambda \partial_{\boldsymbol{q}} V(\boldsymbol{q})^T \alpha(t)$$
$$- \boldsymbol{p}^T M^{-T} A(t)^T \partial_{\boldsymbol{q}} V(\boldsymbol{q}) - \lambda \boldsymbol{p}^T M^{-T}\boldsymbol{p} + \lambda \boldsymbol{p}^T M^{-1}\boldsymbol{p} - \lambda N_f kT - Q\lambda \alpha(t)^T \partial_{\boldsymbol{q}} V(\boldsymbol{q}). \quad \text{(II.3)}$$

This yields (as $M^T = M$),

$$\nabla G(\boldsymbol{y})^T F(t,\boldsymbol{y}) = -\lambda N_f kT$$

and thus we have the property:

$$\mathrm{div}_{\boldsymbol{y}} \left(e^{-G(\boldsymbol{y})/kT} F(t,\boldsymbol{y}) \right) = 0.$$

For a given t_0, this implies that for all $t > t_0$, the application $\boldsymbol{y} \mapsto \Gamma(t, t_0, \boldsymbol{y})$ preserves the measure $e^{-G(\boldsymbol{y})/kT} \mathrm{d}\boldsymbol{y}$ that is (I.2).
If we add the equation (I.3) to the equations (II.1), we get the system

$$\begin{pmatrix} \dot{\boldsymbol{q}} \\ \dot{\xi} \\ \dot{\boldsymbol{p}} \\ \dot{\lambda} \end{pmatrix} = \begin{pmatrix} 0 & 0 & A(t) & \alpha(t) \\ 0 & 0 & 0 & 1/Q \\ -A(t)^T & 0 & 0 & -\boldsymbol{p}/Q \\ -\alpha(t)^T & -1/Q & \boldsymbol{p}^T/Q & 0 \end{pmatrix} \begin{pmatrix} \partial_{\boldsymbol{q}} V(\boldsymbol{q}) \\ N_f kT \\ M^{-1}\boldsymbol{p} \\ Q\lambda \end{pmatrix}. \quad \text{(II.4)}$$

As the matrix in the right-hand side is skew-symmetric for all time, the system (I.3)-(II.1) conserves the energy (I.5). ■

Note that this proof is equivalent to the application of the Non-Hamiltonian methodology developped in[15].

The choice of the matrix $A(t)$ and the vector $\alpha(t)$ can be arbitrary. A typical choice can be

$$A(t) = \mathrm{Id} + \sum_{k=1}^{K} A_k \cos(\omega_k t) \quad \text{and} \quad \alpha(t) = \sum_{k=1}^{K} \alpha_k \cos(\beta_k t) \quad \text{(II.5)}$$

where A_k, α_k, K, ω_k and β_k are to be chosen (A_k are N_f-dimensional matrices). Equivalently, we could also assume $A(t)$ and $\alpha(t)$ to depend on a collection of harmonic oscillators. We give some examples in the numerical experiments below. In order to expect good sampling properties, it is advisable to choose the collection $\Omega := (\omega_1, \ldots, \omega_K, \beta_1, \ldots, \beta_M)$ rationnaly independent to avoid resonnances and KAM behavior (see[3] for a similar case of study). The coefficients can also be randomly chosen at the beginning of the simulation.

Another practical possibility is to take a rotation matrix $Q(t) = \exp(t\Omega)$ where Ω is a skew-symmetric matrix, and to set

$$A(t) = \mathrm{Id} + Q(t), \quad \text{and} \quad \alpha(t) = Q(t)\alpha_0$$

where α_0 is a given vector (the vector α_0 and the matrix Ω can be determined randomly at the beginning of the simulations).

Note that (II.5) implies that

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T A(t)\mathrm{d}t = \mathrm{Id} \quad \text{and} \quad \lim_{T\to\infty} \frac{1}{T}\int_0^T \alpha(t)\mathrm{d}t = 0,$$

which means that the time-average of the vector field defining (II.1) is equal to the Nosé-Hoover vector field. In the same spirit, we could consider the more general system:

$$\begin{aligned} \dot{\boldsymbol{q}} &= A(t)M^{-1}\boldsymbol{p} + Q\alpha(t)\lambda \\ \dot{\boldsymbol{p}} &= -A(t)^T\partial_{\boldsymbol{q}}V(\boldsymbol{q}) - \gamma(t)\lambda\boldsymbol{p} \\ \dot{\lambda} &= \tfrac{\gamma(t)}{Q}\left(\boldsymbol{p}^T M^{-1}\boldsymbol{p} - N_f kT\right) - \alpha(t)^T\partial_{\boldsymbol{q}}V(\boldsymbol{q}). \end{aligned} \tag{II.6}$$

where $\gamma(t)$ is a function of the time. The same proof as in Theorem II.1 shows that (II.6) preserves the measure (I.2) and the energy (I.5). However, after a change of time $\mathrm{d}s = \gamma(t)\mathrm{d}t$, we are back to (II.1) provided that $\gamma(t) \neq 0$. This assumption will be automatically satisfied if we assume $\gamma(t)$ to be of the form $\gamma(t) = 1 + \varepsilon\sum_{k=1}^{K}\gamma_k\cos(\sigma_k t)$ for sufficiently small ε and fixed numbers γ_k and σ_k.

III. NUMERICAL APPROXIMATION

A standard method used to integrate the Nosé-Hoover equation is given by the following reversible scheme, which is an adaptation of the Verlet scheme (see[4]):

$$\begin{aligned}
\boldsymbol{p}^{n+1/2} &= \boldsymbol{p}^n - \frac{\Delta t}{2}\left(\partial_{\boldsymbol{q}} V(\boldsymbol{q}^n) + \lambda^n \boldsymbol{p}^n\right) \\
\lambda^{n+1/2} &= \lambda^n + \frac{\Delta t}{2Q}\left((\boldsymbol{p}^n)^T M^{-1}\boldsymbol{p}^n - N_f kT\right) \\
\boldsymbol{q}^{n+1} &= \boldsymbol{q}_i^n + \Delta t M^{-1}\boldsymbol{p}^{n+1/2} \\
\xi^{n+1} &= \xi^n + \Delta t \lambda^{n+1/2} \\
\boldsymbol{p}^{n+1} &= \boldsymbol{p}^{n+1/2} - \frac{\Delta t}{2}\left(\partial_{\boldsymbol{q}} V(\boldsymbol{q}^{n+1}) + \lambda^{n+1}\boldsymbol{p}^{n+1}\right) \\
\lambda^{n+1} &= \lambda^{n+1/2} + \frac{\Delta t}{2Q}\left((\boldsymbol{p}^{n+1})^T M^{-1}\boldsymbol{p}^{n+1} - N_f kT\right).
\end{aligned} \tag{III.1}$$

This scheme admits the straightforward reversible extension to approximate the equations (II.1)-(I.3): Let $(\boldsymbol{q}^n, \xi^n, \boldsymbol{p}^n, \lambda^n)$ be given as approximation of the solution at the time $t_n = t_0 + n\Delta t$, we define successively

$$A^n = A((t_n + t_{n+1})/2) \quad \text{and} \quad \alpha^n = \alpha((t_n + t_{n+1})/2). \tag{III.2}$$

and

$$\begin{aligned}
\boldsymbol{p}^{n+1/2} &= \boldsymbol{p}^n - \frac{\Delta t}{2}\left((A^n)^T \partial_{\boldsymbol{q}} V(\boldsymbol{q}^n) + \lambda^n \boldsymbol{p}^n\right) \\
\lambda^{n+1/2} &= \lambda^n + \frac{\Delta t}{2}\left(\frac{1}{Q}\left((\boldsymbol{p}^n)^T M^{-1}\boldsymbol{p}^n - N_f kT\right) - (\alpha^n)^T \partial_{\boldsymbol{q}} V(\boldsymbol{q}^n)\right) \\
\boldsymbol{q}^{n+1} &= \boldsymbol{q}^n + \Delta t\left(A^n M^{-1}\boldsymbol{p}^{n+1/2} + Q\alpha^n \lambda^{n+1/2}\right) \\
\xi^{n+1} &= \xi^n + \Delta t \lambda^{n+1/2} \\
\boldsymbol{p}^{n+1} &= \boldsymbol{p}^{n+1/2} - \frac{\Delta t}{2}\left((A^n)^T \partial_{\boldsymbol{q}} V(\boldsymbol{q}^{n+1}) + \lambda^{n+1}\boldsymbol{p}^{n+1}\right) \\
\lambda^{n+1} &= \lambda^{n+1/2} + \frac{\Delta t}{2}\left(\frac{1}{Q}\left((\boldsymbol{p}^{n+1})^T M^{-1}\boldsymbol{p}^{n+1} - N_f kT\right) - (\alpha^n)^T \partial_{\boldsymbol{q}} V(\boldsymbol{q}^{n+1})\right).
\end{aligned} \tag{III.3}$$

When $A(t) = \mathrm{Id}$ and $\alpha(t) = 0$, we recover the previous algorithm. This scheme is reversible and implicit.

In the case where $\alpha(t) = 0$ in (II.1), we can adapt all the existing reversible and explicit algorithms existing for the Nosé-Hoover equations (see e.g.[1,6,11,14]). For instance, the following

scheme is an explicit and reversible adaptation of the leapfrog scheme presented in[6]:

$$A^n = A((t_n + t_{n+1})/2) \tag{III.4}$$

and

$$\begin{aligned}
\boldsymbol{p}^{n+1/2} &= \boldsymbol{p}^n - \frac{\Delta t}{2}\left((A^n)^T \partial_{\boldsymbol{q}} V(\boldsymbol{q}^n) + \lambda^n \boldsymbol{p}^{n+1/2}\right) \\
\boldsymbol{q}^{n+1} &= \boldsymbol{q}^n + (\Delta t) A^n M^{-1} \boldsymbol{p}^{n+1/2} \\
\lambda^{n+1} &= \lambda^n + \frac{\Delta t}{Q}\left((\boldsymbol{p}^{n+1/2})^T M^{-1} \boldsymbol{p}^{n+1/2} - N_f kT\right) \\
\xi^{n+1} &= \xi^n + \frac{\Delta t}{2}(\lambda^n + \lambda^{n+1}) \\
\boldsymbol{p}^{n+1} &= \boldsymbol{p}^{n+1/2} - \frac{\Delta t}{2}\left((A^n)^T \partial_{\boldsymbol{q}} V(\boldsymbol{q}^{n+1}) + \lambda^{n+1} \boldsymbol{p}^{n+1/2}\right).
\end{aligned} \tag{III.5}$$

A. One-dimensional harmonic oscillator

It is a striking fact that one of the most difficult system to sample using the Nosé-Hoover method is the one-dimensional harmonic oscillator, see[4,8,10,12]. In our numerical experiments, we thus first consider one harmonic oscillator with frequency 1, i.e. $V(q) = \frac{q^2}{2}$, and mass $m = 1$. The values of the parameters are $d = 1$, $N = 1$, $kT = 1$, $Q = 1$. For all the computations, we take the initial values $q(0) = 1$, $p(0) = 1$, $\lambda(0) = 1$ and $\xi(0) = 1$, and we use the numerical scheme (III.3). The coefficients $A(t)$ and $\alpha(t)$ are equal to

$$A(t) = 1 + \varepsilon \cos(t) \quad \text{and} \quad \alpha(t) = \varepsilon \cos(\pi t). \tag{III.6}$$

In Figure 1, we computed the position distributions after 2.10^6 iterations with a stepsize equal to $\Delta t = 0.01$. We made the computations with $\varepsilon = 0$ and $\varepsilon = 1$. In the first case (see figure 1, left) the system yields the position distribution for the solution of the Nosé-Hoover system. The distribution does not fit to the expected Gaussian distribution $(\sqrt{2\pi})^{-1} e^{-q^2/2}$. In the second case (see figure 1, right), we see a good agreement of the numerical solution with the exact distribution.

In Figure 2, we show the (p, q)-distribution after 10^5 steps for a stepsize $\Delta t = 0.1$ for $\varepsilon = 0$ (Nosé-Hoover system) and $\varepsilon = 1$.

In Figure 3 we compute the time average $\langle q^2 \rangle_t := \frac{1}{t} \int_0^t q(s)^2 ds$ and $\langle p^6 \rangle_t$ in the case where $\varepsilon = 1$ and $\Delta t = 0.01$. The exact limits should be $\langle q^2 \rangle = 1$ and $\langle p^6 \rangle = 15$.

B. Double-well potential

We consider now two particles of masses $m_1 = m_2 = 1$ interacting through the two-dimensional potential

$$V(q_1, q_2) = \nu\big[(q_1^2 - 1)^2 + q_2^2\big]. \tag{III.7}$$

The number ν measures the depth of the wells (and hence the height of the barrier between the two wells). We integrate the problem for $d = 1$, $N = 2$, $Q = 1$, $kT = 1$, $\nu = 5$,

$$A(t) = \mathrm{Id}_2 + \varepsilon \begin{pmatrix} \cos(t) & \cos(\pi t) \\ \cos(\sqrt{2}t) & \cos(\pi^2 t) \end{pmatrix} \quad \text{and} \quad \alpha(t) = 0, \tag{III.8}$$

and we take $\varepsilon = 0.6$. The choice of the frequencies in (III.8) is done to avoid possible resonnances in the matrix $A(t)$ (see the corresponding strong non-resonance assumption in[3]). We use the explicit integration scheme (III.5) with $\Delta t = 0.01$ and 2.10^6 iterations. As initial values, we take $p_1 = 1$, $p_2 = 0$, $q_1 = 1$, $q_2 = 0$, $\xi = 0$ and $\lambda = 0$. In Figure 4 we show the numerical results for the computation of the average of the variable p_1^6 and the corresponding p_1 distribution. As expected, the time-average of $p_1^6(t)$ tends to the theoretical limit $\langle p_1^6 \rangle = 15$, while the distribution of p_1 fits the exact distribution $(\sqrt{2\pi})^{-1} e^{-p_1^2/2}$.

For comparison purposes, we compute the same results for the associated Nosé-Hoover chains systems with M chains using the explicit and reversible method in[6]. The system is written (see[10]):

$$\begin{aligned}
\dot{\boldsymbol{q}} &= M^{-1}\boldsymbol{p} \\
\dot{\boldsymbol{p}} &= -\partial_{\boldsymbol{q}} V(\boldsymbol{q}) - \lambda_1 \boldsymbol{p} \\
\dot{\lambda}_1 &= \frac{1}{Q_1}\left(\boldsymbol{p}^T M^{-1}\boldsymbol{p} - N_f kT\right) - \lambda_2\lambda_3 \\
\dot{\lambda}_j &= \frac{1}{Q_j}\left(Q_{j-1}\lambda_{j-1}^2 - kT\right) - \lambda_{j+1}\lambda_j, \quad j = 2, \ldots, M-1, \\
\dot{\lambda}_M &= \frac{1}{Q_M}\left(Q_{M-1}\lambda_{M-1}^2 - kT\right).
\end{aligned} \tag{III.9}$$

In a first simulation, we take the parameters $Q_1 = kT/(4\nu)$ and $Q_j = kT/(8\nu)$ for $j = 2, \ldots, M$, $\Delta t = 0.01$, $kT = 1$, $d = 1$, $N = 2$ and the potential (III.7). We take as initial values $p_1 = 1$, $p_2 = 1$, $q_1 = 1$, $q_2 = 0$, and $\lambda_j = 0$, $j = 1, \ldots, M$. In Figure 5 we show the results after 2.10^6 iterations for $\nu = 5$, with $M = 4$, 8 and 14 thermostats. The systems produce the correct distributions and yield the right average values. The mass parameter Q_j are close to the optimal

values (see for instance[11]). Note that the stepsize is chosen so that the energy conservation is ensured during the simulations.

In a second simulation, we take all the parameters Q_j, $j = 1, \ldots, M$ equal to 1 and the same parameters as above.

For $\nu = 1$, the NHC systems produce the correct distributions and yield the right average values. However, for $\nu = 5$, the system becomes stiff and the averages are wrong. In Figure 6, we show the same results as in Figure 5 for these new mass parameters. We see that the NHC systems do not yield the correct average of p_1^6. Moreover, the p_1 distributions plotted on the right are wrong in these cases.

In Figure 7, we show the q_1 trajectories of the NHC system with $M = 8$ thermostats applied to the double well potential with $\nu = 6$ (left) and the same parameters as above. We also show the corresponding trajectory for the Nosé-Hoover system with shakers (the stepsize and initial data are the same as previously). We see that after 2.10^6, the solution of the NHC system with these bad mass parameters jumps only one time from the well centered around $q_1 = 1$ while the system with shakers changes more frequently.

C. Collection of Harmonic oscillators

For bigger systems, metastability problem may occur: this is typically the case when the trajectory of the Nosé-Hoover chain system is trapped in a deep potential well. The dynamics is then close to the dynamics of a collection of harmonic oscillators. In Figure 8, we consider the case of $N = 10$ harmonic oscillators with frequencies varying from $\sqrt{6}$ to $2\sqrt{6}$ and of masses 1. The potential is written

$$V(\boldsymbol{q}) = \sum_{i=1}^{N} 3\omega_i^2 q_i^2$$

where $\omega_i = 1 + (i-1)/N$ for $i = 1, \ldots, N$. We compute the p_1 distribution after 2.10^6 iteration with stepsize $\Delta t = 0.01$. In the left figure, the distribution of the Nosé-Hoover system with shakers is plotted. We take $Q = 1$, $kT = 1$, $d = 1$, $N = 10$, $\alpha(t) = 0$ and $A(t)$ of the form

$$A(t) = \mathrm{Id} + \sum_{i,j=1}^{N} \delta_{ij} \cos(\beta_i t)$$

where δ_{ij} is the matrix with entries 0 except for the coefficient (i, j) which is equal to 1. The coefficient β_i are random numbers determined at the beginning of the simulation (we take $\beta_i \sim \mathcal{N}(0, 1)$). We use the integration scheme (III.5). The distribution agrees with the

theoretical result. In the right figure, we plot the same result for the simulation with the NHC method with $M = 4$ (solid line), $M = 8$ (dashed-dot line) and $M = 14$ (dashed line) thermostats. According to rule for optimal coupling[16], (see for instance[11]) the parameters Q_j are $Q_1 = N_f kT/12$ and $Q_j = kT/12$, $j = 2, \ldots, M$, while $d = 1$, $N = 10$ and $kT = 1$. The exact distribution is plotted as a thin solid line. In both simulations, the initial values are $p_i = 1$ and $q_i = 0$ for $i = 1, \ldots, N$ and all the thermostats $\lambda_i = 0, i = 1, \ldots, N$, and the number of steps is 2.10^6.

IV. EXTENSIONS

A. Microcanonical shakers

The introduction of shakers can be made in many different systems. Consider in particular a Hamiltonian system

$$\dot{\boldsymbol{y}}(t) = J \nabla H\left(\boldsymbol{y}(t)\right) \tag{IV.1}$$

where J is the skew-symmetric matrix:

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

This system preserves the Euclidean volume, and the energy surface $\{H(\boldsymbol{y}) = H(\boldsymbol{y}(0))\}$. Hence, it preserves the Liouville measure $d\sigma(\boldsymbol{y})/\|\nabla H(\boldsymbol{y})\|$ on the isosurfaces. If $J(t)$ is a time-dependent matrix such that for all time t, $J(t)$ is skew symmetric, we can show that the solution of the system

$$\dot{\boldsymbol{y}}(t) = J(t) \nabla H\left(\boldsymbol{y}(t)\right) \tag{IV.2}$$

has the same preservation properties as the Hamiltonian system (IV.1): Along the solution of (IV.2) the energy $H(\boldsymbol{y}_0)$ is constant and the Euclidean volume is conserved. This implies that the measure $d\sigma(\boldsymbol{y})/\|\nabla H(\boldsymbol{y})\|$ is invariant for the exact solution of (IV.2). The proof is the same as in Theorem II.1. Again, we expect that the introduction of shakers inside the symplectic matrix destroys the possible hidden invariants of the system. Notice that in this situation, there exist ergodicity results in the case of a quadratic Hamiltonian (i.e. a collection of harmonic oscillators), see[3].

This idea can be applied in particular to the case of the initial Nosé extended Hamiltonian

(see[12]) and to the Nosé-Poincaré systems (see[1]). However, as the constant symplectic structure is destroyed, the use of symplectic scheme cannot guarantee the energy conservation for the numerical solutions.

B. NHC with shakers

The introduction of shakers can also be made in the NHC system (III.9). Indeed, for a given time-dependent matrix $A(t)$, we can consider systems of the form

$$\begin{aligned}
\dot{\boldsymbol{q}} &= A(t)M^{-1}\boldsymbol{p} \\
\dot{\boldsymbol{p}} &= -A(t)^T\partial_{\boldsymbol{q}}V(\boldsymbol{q}) - \lambda_1\boldsymbol{p} \\
\dot{\lambda}_1 &= \frac{1}{Q_1}\left(\boldsymbol{p}^TM^{-1}\boldsymbol{p} - N_fkT\right) - \lambda_2\lambda_3 \\
\dot{\lambda}_j &= \frac{1}{Q_j}\left(Q_{j-1}\lambda_{j-1}^2 - kT\right) - \lambda_{j+1}\lambda_j, \quad j = 2,\ldots,M-1, \\
\dot{\lambda}_M &= \frac{1}{Q_M}\left(Q_{M-1}\lambda_{M-1}^2 - kT\right).
\end{aligned} \tag{IV.3}$$

The numerical algorithms employed to integrate (III.9) can then easily be adapted to this case using the same method as before. Again, we expect better sampling properties for these new systems than the original ones. However, for the simple examples treated above, there is no differences with the case of $M = 1$ thermostat.

---

[1] S. Bond, B. Laird, B. Leimkuhler, *The Nosé-Poincaré Method for Constant Temperature Molecular Dynamics*, Journal of Computational Physics , 151 (1999) 114-134.

[2] A.C. Branka, M. Kowalik, K. W. Wojciechowski, *Generalization of the NosHoover approach*, J. Chem. Phys., 119 (2003), 1929–1936.

[3] L. H. Eliasson, *Ergodic skew-systems on* $\mathbb{T}^d \times SO(3,\mathbb{R})$ Ergod. Th. & Dynam. Sys. 22 (2002), 1429–1449

[4] D.Frenkel, B. Smit, *Understanding Molecular Simulation. From Algorithms to Applications*, Academic Press, Boston (1996 1st Ed.) (2002, 2nd Ed)

[5] W. G. Hoover, *Canonical dynamics: Equilibrium phase-space distributions*, Phys. Rev. A, 31 (1985), 1695–1697.

[6] S. Jang, G.A. Voth, *Simple reversible molecular dynamics algorithms for Nosé-Hoover chain dynamics*, J. Chem. Phys., 107 (1997), 9514–9526.

[7] B.B. Laird, B.J. Leimkuhler, *Generalized dynamical thermostatting* Phys. Rev. E 68, 016704 (2003).

[8] F. Legoll, M. Luskin, R. Moeckel, *Non-ergodicity of the Nosé-Hoover thermostatted harmonic oscillator*, Preprint.

[9] Y. Liu, M. E. Tuckermann, *Generalized Gaussian moment thermostatting: A new continuous dynamical approach to the canonical ensemble*, J. Chem. Phys. 112 (2000), 1685–1700.

[10] G. J. Martyna, M. L. Klein, M. E. Tuckermann, *Nosé-Hoover chains: The canonical ensemble via continuous dynamics*, J. Chem. Phys. 97 (1992), 2635–2645.

[11] G. J. Martyna, M. E. Tuckermann, D. Tobias, M. L. Klein, Mol. Phys. 87 (1996), 1117.

[12] S. Nosé, *A molecular dynamics method for simulations in the canonical ensemble*, Mol. Phys. 52 (1984), 255–258.

[13] S. Nosé, *A unified formulation of the constant temperature molecular dynamics methods*, J. Chem. Phys. 81 (1984), 511–519.

[14] S. Toxvaerd, Mol. Phys. 72 (1991).

[15] M. E. Tuckerman, Y. Liu, G. Ciccotti, G. J. Martyna, *Non-Hamiltonian molecular dynamics: Generalizing Hamiltonian phase space principles to non-Hamiltonian systems* J. Chem. Phys. 115 (2001), pp. 1678-1702.

[16] In this particular case, the use of "massive thermostatting" techniques could be more appropriate.

Acknowledgments. The author is glad to thank Philippe Chartier, Christophe Chipot, Frédéric Legoll and Tony Lelièvre for helpful comments and discussions.

V. FIGURE CAPTIONS

Figure 1: Position distribution of the Nosé-Hoover system for the harmonic oscillator ($\Delta t = 0.01$): without shakers (left, $\varepsilon = 0$) and with shakers (right, $\varepsilon = 1$).

Figure 2: (p, q) distribution of the Nosé-Hoover system for the harmonic oscillator after 10^5 steps ($\Delta t = 0.1$): without shakers (left, $\varepsilon = 0$) and with shakers (right, $\varepsilon = 1$).

Figure 3: Time average of q^2 and p^6 after 2.10^6 steps ($\Delta t = 0.01$) for the harmonic oscillator with shakers ($\varepsilon = 1$).

Figure 4: Left: time average of $p_1^6(t)$ for the double well potential with shakers. Right: corresponding p_1 numerical distribution (dashed line) and theoretical (solid line).

Figure 5: NHC computations with optimal mass parameters for the double well potential ($\nu = 5$) with M = 4 (dashed-dot), M= 8 (dashed) and M = 14 (solid line) thermostats.

Figure 6: NHC computations with bad mass parameters for the double well potential ($\nu = 5$) with M = 4 (dashed-dot), M= 8 (dashed) and M = 14 (solid line) thermostats.

Figure 7: q_1 trajectories for the double well potential ($\nu = 6$): NHC system with M = 8 thermostats and bad parameters (left) and Nosé-Hoover system with shakers (right).

Figure 8: p_1 distribution for 10 harmonic oscillators: with shakers (left) and using the NHC method with $M = 4$ (solid line), $M = 8$ (dashed-dot) and $M = 14$ (dashed) thermostats.

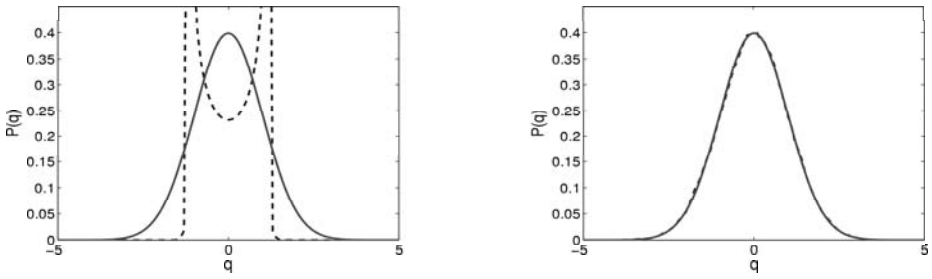

FIG. 1:

Figure 1

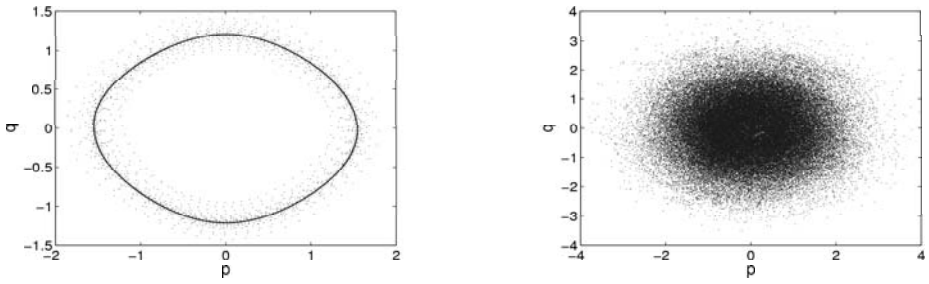

FIG. 2:

Figure 2

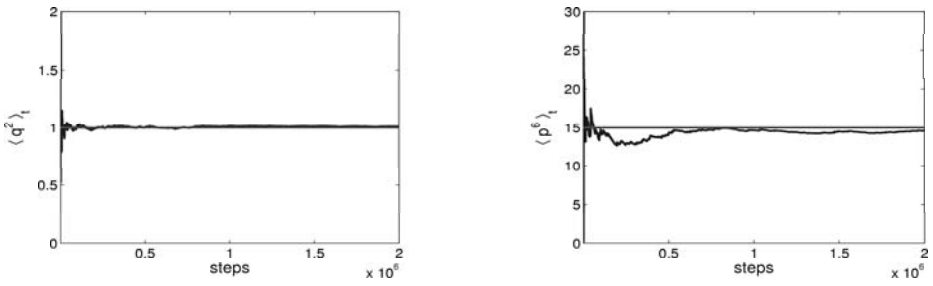

FIG. 3:

Figure 3

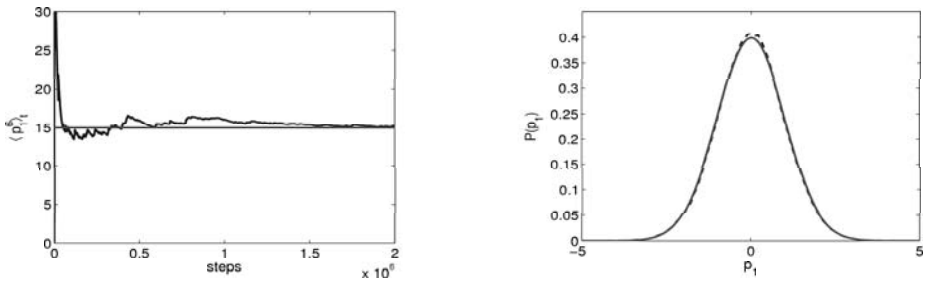

FIG. 4:

Figure 4

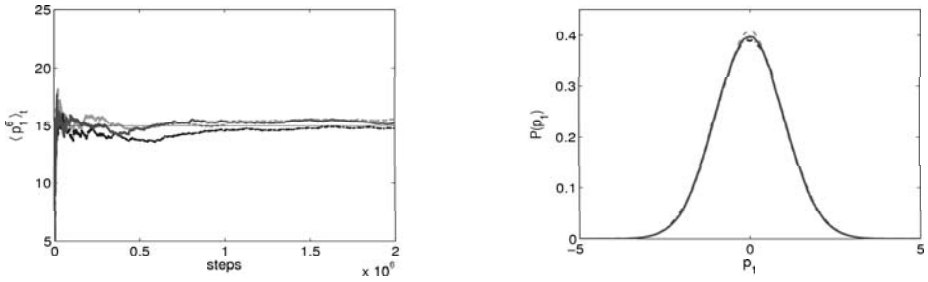

FIG. 5:

Figure 5

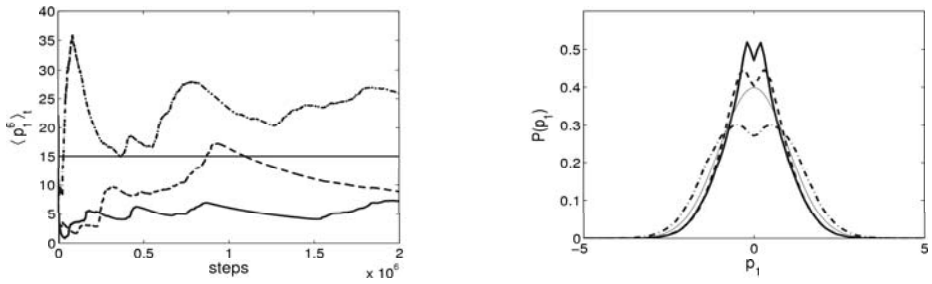

FIG. 6:

Figure 6

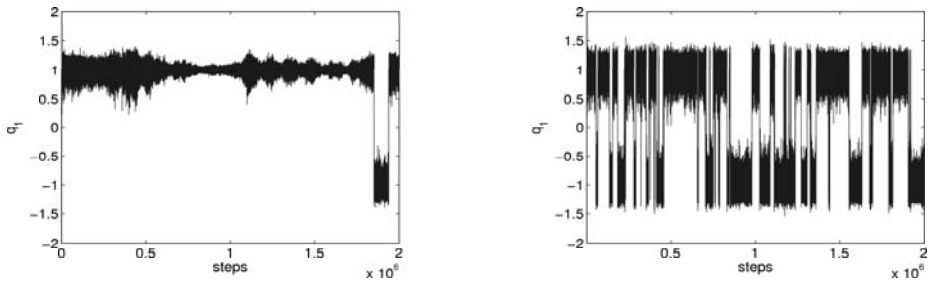

FIG. 7:

Figure 7

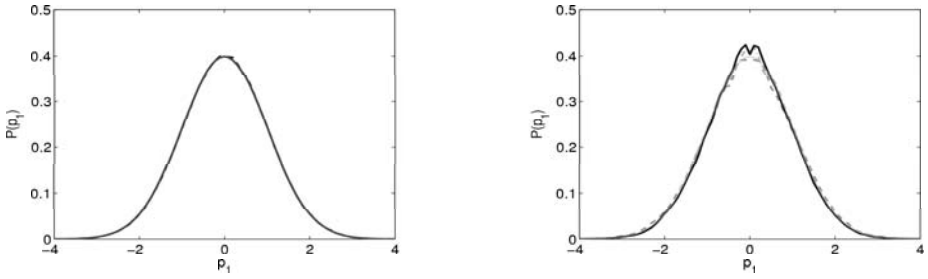

FIG. 8:

Figure 8