Non-equilibrium steady state and subgeometric ergodicity for a chain of three coupled rotors

N Cuneo¹, J-P Eckmann¹,² and C Poquet³

¹ Département de physique théorique, Université de Genève, Geneva, Switzerland
² Section de mathématiques, Université de Genève, Geneva, Switzerland
³ Dipartimento di Matematica, Università di Roma Tor Vergata, Roma, Italy

Received 7 November 2014, revised 30 March 2015
Accepted for publication 9 April 2015
Published 19 June 2015

Abstract
We consider a chain of three rotors (rotators) whose ends are coupled to stochastic heat baths. The temperatures of the two baths can be different, and we allow some constant torque to be applied at each end of the chain. Under some non-degeneracy condition on the interaction potentials, we show that the process admits a unique invariant probability measure, and that it is ergodic with a stretched exponential rate. The interesting issue is to estimate the rate at which the energy of the middle rotor decreases. As it is not directly connected to the heat baths, its energy can only be dissipated through the two outer rotors. But when the middle rotor spins very rapidly, it fails to interact effectively with its neighbours due to the rapid oscillations of the forces. By averaging techniques, we obtain an effective dynamics for the middle rotor, which then enables us to find a Lyapunov function. This and an irreducibility argument give the desired result. We finally illustrate numerically some properties of the non-equilibrium steady state.

Keywords: non-equilibrium, stochastic process, convergence to steady state, chain of rotors
Mathematics Subject Classification: 82C05, 60H10, 34C15, 34C29, 60B10

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
1. Introduction

Hamiltonian chains of mechanical oscillators have been studied for a long time. Several models describe a linear chain of masses, with polynomial interaction potentials between adjacent masses, and pinning potentials which tie the masses down in the laboratory frame. Under the assumption that the interaction is stronger than the pinning, it was shown in [5] that the model has an invariant probability measure when the chain is attached at each extremity to two heat baths at different temperatures. That paper, and later developments, see e.g. [3], relied on analytic arguments, showing in particular that the infinitesimal generator has compact resolvent in a suitable function space.

Two elements were added later in the paper [12]: first, the authors used a more probabilistic approach, based on Harris recurrence as developed by Meyn and Tweedie [10]. Second, a detailed analysis allowed them to understand the transfer of energy from the central oscillators to the (dissipative) baths. In that case the convergence to the stationary state is of exponential rate. In [1], this reasoning was extended to more general contexts.

The dynamics of the chain is very different when the pinning potential is stronger than the interaction potential. In that case the chain may have breathers, i.e. oscillators concentrating a lot of energy, which is transferred only very slowly to their neighbours. This may lead to subexponential ergodicity, as shown by Hairer and Mattingly [7] in the case of a chain of 3 oscillators with strong pinning.

In this paper, we discuss a model with three rotors (see figure 1), each given by an angle \( q_i \in \mathbb{T} = \mathbb{R}/2\pi\mathbb{Z} \) and a momentum \( p_i \in \mathbb{R}, i = 1, 2, 3 \). The phase space is therefore \( \Omega = \mathbb{T}^3 \times \mathbb{R}^3 \), and we will consider the measure space \((\Omega, B)\), where \( B \) is the Borel \( \sigma \)-field over \( \Omega \). We will denote the points of \( \Omega \) by \( x = (q, p) \) with \( q = (q_1, q_2, q_3) \) and \( p = (p_1, p_2, p_3) \).

We introduce the Hamiltonian

\[
H(q, p) = \sum_{i=1}^{3} \left( \frac{1}{2} p_i^2 + U_i(q_i) \right) + \sum_{b=1,3} W_b(q_2 - q_b),
\]

with some smooth interaction potentials \( W_b : \mathbb{T} \to \mathbb{R}, b = 1, 3 \), and some smooth pinning potentials \( U_i : \mathbb{T} \to \mathbb{R}, i = 1, 2, 3 \). We now let the two outer rotors (i.e. the rotors 1 and 3) interact with Langevin-type heat baths at temperatures \( T_1, T_3 > 0 \), and with coupling constants \( \gamma_1, \gamma_3 > 0 \). Moreover, we apply some constant (possibly zero) external forces \( \tau_1 \) and \( \tau_3 \) to the two outer rotors. Introducing \( w_b = W'_b \) and \( u_i = U'_i \), we obtain the system of SDE:

\[
\begin{align*}
\text{d}q_i(t) &= p_i(t) \text{d}t, \quad i = 1, 2, 3, \\
\text{d}p_2(t) &= - \sum_{b=1,3} w_b(q_2(t) - q_b(t)) \text{d}t - u_2(q_2(t)) \text{d}t,
\end{align*}
\]

(1.1)

Figure 1. A chain of three rotors with two external torques \( \tau_1 \) and \( \tau_3 \) and two heat baths at temperatures \( T_1 \) and \( T_3 \).
For all sufficiently small density proportional to \( e^{\int \sum b \psi_b(t) \, dt} \), we write

\[
\frac{\partial p_b(t)}{\partial t} = \left( w_b(q_2(t) - q_b(t)) + \tau_b - u_b(q_b(t)) - \gamma_b p_b(t) \right) \, dt \\
+ \sqrt{2\gamma_b T_b} \, dB_b^b, \quad b = 1, 3,
\]

where \( B_1 \) and \( B_3 \) are standard independent Brownian motions.

**Notation.** In the sequel, the index \( b \) always refers to the rotors 1 and 3 at the boundaries of the chain, and we write \( \sum_b \) instead of \( \sum_{b=1,3} \).

**Remark 1.1.** Our model can be viewed as an extreme case of that studied in [7]. A key factor in that paper is to realise how the frequency of one isolated pinned oscillator depends on its energy. Indeed, for an isolated oscillator with Hamiltonian \( p^2/2 + q^2k/(2k) \), the frequency grows like the energy to the power \( \frac{1}{2} \). When \( k \to \infty \), the exponent converges to \( \frac{1}{2} \). In this limit, the pinning potential formally becomes an infinite potential well, so that the variable \( q \) is constrained to a compact interval. In our model, the position (angle) of a rotor lives in a compact space, and its frequency scales like its momentum, i.e. like the square root of its energy. Therefore, we can view our rotor model as some kind of ‘infinite pinning’ limit.

We make the following non-degeneracy assumption (clearly satisfied for e.g. \( w_1 = w_3 = \sin \)):

**Assumption 1.2.** There is at least one \( b \in \{1, 3\} \) such that for each \( s \in \mathbb{T} \), at least one of the derivatives \( w^{(k)}_b(s) \), \( k \geq 1 \) is non-zero.

For all initial conditions \( x \in \Omega \) and all times \( t \geq 0 \), we denote by \( P^t(x, \cdot) \) the transition probability of the Markov process associated to (1.1). Since the coefficients of the SDE (1.1) are globally Lipschitz, the solutions are almost surely defined for all times and all initial conditions, so that \( P^t(x, \cdot) \) is well-defined as a probability measure on \( (\Omega, \mathcal{B}) \).

We now introduce the main theorem, in which we write

\[ \|v\|_{f} = \sup_{|k| \leq f} \int_{\Omega} h |v| \, dv \]

for any continuous function \( f > 0 \) on \( \Omega \) and any signed measure \( v \) on \( (\Omega, \mathcal{B}) \).

**Theorem 1.3.** Under assumption 1.2, the following holds for the Markov process defined by (1.1):

(i) The transition kernel \( P^t \) has a density \( p_t(x, y) \) in \( C^\infty((0, \infty) \times \Omega \times \Omega) \).

(ii) The process admits a unique invariant measure \( \pi \), which has a smooth density.

(iii) For all sufficiently small \( \beta > 0 \) and all \( \beta' \in [0, \beta) \), there are constants \( C, \lambda > 0 \) such that for all \( t \geq 0 \) and all \( x = (q_1, q_2, \ldots, q_3) \in \Omega \),

\[ \| P^t(x, \cdot) - \pi \|_{c^0_{\mathbb{R}}} \leq C (1 + p_2^2)^e^{-\beta \int \lambda t/2}. \]

**Remark 1.4.** If both heat baths are at the same temperature, say, \( T_1 = T_3 = T > 0 \), and the forces \( \tau_1 \) and \( \tau_3 \) are zero, then the system is at thermal equilibrium and the Gibbs measure with density proportional to \( e^{-H/T} \) is invariant. Indeed, one easily checks that this density verifies the stationary Fokker-Planck equation \( L^* e^{-H/T} = 0 \), where \( L^* \) is the formal adjoint of the generator \( L \) introduced below.

**Remark 1.5.** In fact, the results we prove here apply with hardly any modification to the ‘star’ configuration with one central rotor interacting with \( m \) external rotors, which in turn are coupled to heat baths (i.e. \( m + 1 \) rotors and \( m \) heat baths).

In addition, some studies (e.g. [9]) consider chains with fixed boundary conditions. For the left end of the chain, this corresponds to adding a ‘dummy’ rotor 0 which does not move but interacts with rotor 1. This is covered by our theory by adding some contribution to the pinning potential \( U_1 \). The same applies to the right end and \( U_3 \).
Chains of rotors provide toy models for the study of non-equilibrium statistical mechanics. In [9] long chains have been studied numerically, and it appears that even when the external temperatures are different and external forces are applied, local thermal equilibrium is satisfied in the stationary state in the limit of infinitely long chains. This stationary state may have some surprising features, like a large amount of energy in the bulk of the chain when the boundary conditions are properly chosen. In our case, of course, we are far from local thermal equilibrium, since we study only systems made of three rotors. We will present some numerical simulations of our system in section 6, highlighting some interesting properties of the stationary state.

What corresponds here to the breathers observed in other models is the situation where the energy of the system is very large and mostly concentrated in the middle rotor. The middle rotor then spins very rapidly, and the interaction forces oscillate so fast that they have very little net effect. In this case, the middle rotor effectively decouples from the rest of the system, and the main difficulty is to show that its energy eventually decreases with some well-controlled bounds.

The idea used in [7] for the chain of three pinned oscillators is to average the oscillatory forces, and exhibit a negative feedback in the regime where the breather dominates the dynamics. The proof of theorem 1.3 in the present paper is based on a systematisation of this idea, as explained in section 3.4.

The paper is structured as follows: in section 2 we introduce a sufficient condition for subgeometric ergodicity from [2]. In section 3 we study the behaviour of the middle rotor. In section 4 we show how to use the study of $p_2$ to get a Lyapunov function. In section 5 we provide the necessary technical input to the theorem of [2]. Finally, we illustrate numerically some properties of the non-equilibrium steady state in section 6.

2. Ergodicity and Lyapunov functions

The proof of theorem 1.3 relies on the results of [2] which in turn are based on the theory exposed in [10]. The theory of [10] shows that one can prove the ergodicity of an irreducible Markov process and estimate the rate of convergence toward its invariant measure if one has a good control of the return times of the process to particular sets, called petite sets. A set $K$ is petite if there exist a probability measure $\alpha$ on $[0, \infty)$ and a non-zero measure $\nu_\alpha$ on $\Omega$ such that for all $x \in K$ one has $\int_0^\infty P^t(x, \cdot)\alpha(dt) \geq \nu_\alpha(\cdot)$. In the case we are interested in, control arguments and the hypoellipticity of the generator imply that each compact set is petite (see section 5.1 for a proof of this property).

Let $L$ be the infinitesimal generator of the process, i.e. the second-order differential operator

$$L = \sum_{i=1}^3 \left( p_i \partial_{q_i} - u_i(q_i) \partial_{p_i} \right) + \sum_b \left[ w_b(q_2 - q_b)(\partial_{p_b} - \partial_{p_2}) + \tau_b \partial_{p_b} - \gamma_b p_b \partial_{p_b} + \gamma_b T_b \partial_{p_b}^2 \right].$$

Recall that for any sufficiently regular function $f$ we have $Lf(x) = \frac{d}{dt} \left[ \int f(y) P_t(x, dy) \right]_{t=0}$.

A classical way to control the return times to a petite set is to make use of Lyapunov functions. We call Lyapunov function a smooth function $V : \Omega \mapsto [1, \infty)$ with compact level sets (i.e. due to the structure of $\Omega$, a function such that $V(q, p) \to \infty$ when $\|p\| \to \infty$) such that for all $x \in \Omega$,

$$L V(x) \leq C \mathbf{1}_K(x) - \phi \circ V(x),$$

(2.1)
where $C$ is a constant, $\phi : [1, \infty) \to (0, \infty)$ is an increasing function, and $K$ is a petite set. If one can find such a function, and prove that some skeleton $P^\Delta(\Delta > 0)$ is $\mu$-irreducible for some measure $\mu$ (i.e. $\mu(A) > 0$ implies that for all $x \in \Omega$ there exists $k \in \mathbb{N}$ such that $P^\Delta(x, A) > 0$), then the Markov process is ergodic, with rate depending on $\phi$. In the case where $\phi(V) \propto V^\rho$, the convergence is geometric if $\rho = 1$ and polynomial if $\rho < 1$ (see [2,11]).

In this paper, we obtain $\phi(V) \sim V/\log V$. We rely on the work of Douc, Fort and Guillin [2], which gives a sufficient condition for subgeometric ergodicity of continuous-time Markov processes. We give here a simplified version of their result, adapted to our purpose. This statement is based on theorems 3.2 and 3.4 of [2].

**Theorem 2.1 (Douc-Fort-Guillin (2009)).** Assume that the process has an irreducible skeleton and that there exist a smooth function $V : \Omega \to [1, \infty)$ with $V(\cdot, \cdot) \to \infty$ when $\|p\| \to \infty$, an increasing, differentiable, concave function $\phi : [1, \infty) \to (0, \infty)$, a petite set $K$, and a constant $C$ such that (2.1) holds. Then the process admits a unique invariant measure $\pi$, and for each $z \in [0, 1]$, there exists a constant $C'$ such that for all $t \geq 0$ and all $x \in \Omega$,

$$
\|P_t(x, \cdot) - \pi\|_{(\phi \circ V)^z} \leq g(t)C'V(x),
$$

where $g(t) = (\phi \circ H^{-1}_\phi(t))^{z-1}$, with $H_\phi(u) = \int_1^u \frac{ds}{\phi(s)}$.

When $z = 0$, we retrieve the total variation norm $\|P_t^\Delta(x, \cdot) - \pi\|_{TV}$ and the rate is the fastest. Increasing $z$ strengthens the norm but slows the convergence rate down. When $z = 1$, the norm is the strongest, but no convergence is guaranteed since $g(t) \equiv 1$.

The core of the paper is devoted to the construction of a Lyapunov function such that (2.1) is satisfied with $\phi(s) \sim s/\log s$, and a set $K$ which is compact and therefore petite. This yields a stretched exponential convergence rate (see (2.4)). The existence of an irreducible skeleton required by theorem 2.1 and the fact that every compact set is petite are proved in section 5.

One might at first think that a Lyapunov function is simply given by the Hamiltonian $H$. Unfortunately, this is not the case, as

$$
LH = \sum_b \left(\tau_b p_b + \gamma_b (T_b - p_b^2)\right),
$$

where the right-hand side remains positive when $p_1, p_3$ are small and $p_2 \to \infty$. Thus, there is no bound of the form (2.1) for $H$. The same problem occurs if we take any function $f(H)$ of the energy.

In order to find a *bona fide* Lyapunov function, we will need more insight into how fast all three momenta decrease. The equality (2.2) suggests that $p_1$ and $p_3$ will not cause any problem. In fact, we have for $b = 1, 3$, that

$$
Lp_b = -\gamma_b p_b + w_b(q_2 - q_b) - u_b(q_b) + \tau_b.
$$

Since $w_b(q_2 - q_b) - u_b(q_b) + \tau_b$ is bounded, $|p_b|$ essentially decays at exponential rate when it is large. This is, of course, due to the friction terms that act on $p_1$ and $p_3$ directly. Such a result does not hold for $p_2$. In fact, the decay of $p_2$ is much slower. Our main insight is that in a sense

$$
Lp_2 \sim -cp_2^3.
$$

The proof of such a relation occupies a major part of this paper. As indicated earlier, this very slow damping of $p_2$ comes from the lack of effective interaction when the forces oscillate very rapidly. Once we have gained enough understanding of the dynamics of $p_2$, we will be able
to construct a Lyapunov function, whose properties are summarised in

**Proposition 2.2.** For all sufficiently small $\beta > 0$, there is a function $V : \Omega \to [1, \infty)$ satisfying the two following properties:

(i) There are positive constants $c_1, c_2$ such that
$$1 + c_1 e^{\beta H} \leq V \leq c_2 (1 + p_2^2) e^{\beta H}.$$  

(ii) There are positive constants $c_3, c_4$ and a compact set $K$ such that
$$LV \leq c_3 1_K - \phi(V),$$
where $\phi : [1, \infty) \to (0, \infty)$ is defined by
$$\phi(s) = \frac{c_4 s}{2 + \log(s)}. \quad (2.3)$$

The way we construct the Lyapunov function is somewhat different from that of [7]. There, it is obtained starting from some power of the Hamiltonian and then adding corrections by an averaging technique similar to ours (see remark 3.8). Here, we first average the dynamics of $p_2$ and then use the result to construct a Lyapunov function that essentially grows exponentially with the energy. This gives a stretched exponential rate of convergence instead of a polynomial rate as in [7]. The present method can in principle be applied to the model of [7] (see also [6]).

We now show how the main results follows.

**Proof of theorem 1.3.** The conclusions of theorem 1.3 immediately follow from theorem 2.1, proposition 2.2, the technical results stated in proposition 5.1, and the following two observations. Consider $0 \leq \beta' < \beta$ and choose $z \in (0, 1)$ such that $\beta' < z \beta$. First, the function $\phi$ defined in (2.3) yields, in the notation of theorem 2.1, a convergence rate
$$g(t) = (\phi \circ H_{\phi}^{-1}(t))^{z-1} \leq e^{-\lambda t}/2 \quad (2.4)$$
for some $c, \lambda > 0$. Indeed, we have $H_\phi(u) = \frac{1}{2} \int_1^u \frac{2 \log s}{s} ds = \frac{1}{2c_4} (\log u)^2 + \frac{2}{c_4} \log u$, so that $H_\phi^{-1}(t) = \exp((2c_4 t + 4)^{1/2} - 2)$ and $(\phi \circ H_{\phi}^{-1}(t)) = (2c_4 t + 4)^{-1/2} \exp((2c_4 t + 4)^{1/2} - 2) \geq Ce^{C' t/2}$ for some $C, C' > 0$. Thus, (2.4) holds with $\lambda = (1-z)C'$. Secondly, by proposition 2.2 (i), and since $\beta' < z \beta$, we observe that $e^{\beta H} \leq c (\phi \circ V)^z$ for some constant $c > 0$, so that $\| \cdot \|_{(\phi \circ V)^z} \leq c \| \cdot \|_{(\phi \circ V)^z}$. \hfill $\Box$

### 3. Effective dynamics for the middle rotor

The hardest and most interesting part of the problem is to determine how $p_2$ decreases when it is very large\(^5\). In this section, we obtain some asymptotic, effective dynamics for $p_2$ when $p_2 \to \infty$.

**3.1. Expected rate**

Before we start making any proof, we can get a hint of how $p_2$ decreases in the regime where $p_2$ is very large and both $p_1, p_3$ are small. Assume for simplicity that $u_t \equiv 0$ and that $W_b(s) = -\kappa \cos(s)$ so that $w_b(s) = \kappa \sin(s)$. In the regime of interest, we expect the middle rotor to decouple, so that $p_2$ will evolve very slowly. We will consider the system over times that are small enough for $p_2$ to remain almost constant (say equal to $\omega$), but large enough for

\(^4\) The 2 in the denominator ensures that $\phi$ is concave and increasing on $[1, \infty)$, as required in theorem 2.1.

\(^5\) To simplify notation, we say $p_2$ is large, but we always really mean that $|p_2|$ is large.
Let some ‘quasi-stationary’ regime to be reached. The reader can think of $\omega$ as being the ‘initial’ value of $p_2$. For $b = 1, 3$, we expect $p_b$ to be well approximated, at least qualitatively, by the equation

$$dp_b = \kappa \sin(\omega t) \, dt - \gamma_b p_b \, dt + \sqrt{2\gamma_b} \, dB^b_t,$$

whose solution is

$$p_b(t) = \kappa \frac{\gamma_b \sin(\omega t) - \omega \cos(\omega t)}{\sqrt{\gamma_b} + \omega^2} + \sqrt{2\gamma_b} \int_0^t e^{\gamma_b(t-s)} \, dB^b_s$$

$$\quad = -\kappa \frac{\cos(\omega t)}{\omega} + \sqrt{2\gamma_b} \int_0^t e^{\gamma_b(t-s)} \, dB^b_s + O\left(\frac{1}{\omega^2}\right).$$

We have neglected the exponentially decaying part $p_b(0)e^{-\gamma_s}$ since we assume that a quasi-stationary regime is reached. By (2.2), the rate of energy flowing into of the system at $b$ is $\gamma_b(T_b - p_b^2)$. Squaring $p_b$ and taking expectations, what remains is

$$\mathbb{E}p_b^2(t) = \kappa^2 \cos^2(\omega t) + 2\gamma_b T_b \mathbb{E}\left(\int_0^t e^{\gamma_b(t-s)} \, dB^b_s\right)^2 + O\left(\frac{1}{\omega^3}\right)$$

$$\quad = \kappa^2 \cos^2(\omega t) + (1 - e^{-2\gamma_s}) T_b + O\left(\frac{1}{\omega^3}\right),$$

where we have used the Itô isometry $\mathbb{E}(\int_0^t e^{\gamma_b(t-s)} \, dB^b_s)^2 = \int_0^t e^{2\gamma_b(t-s)} \, ds$. Neglecting again an exponentially decaying term, we obtain

$$\frac{d}{dt} \mathbb{E}H(t) = \sum_b \mathbb{E}\gamma_b(T_b - p_b^2(t)) \sim -\sum_b \gamma_b \kappa^2 \frac{\cos^2(\omega t)}{\omega^2}.$$

Since $\cos^2(\omega t)$ oscillates very rapidly around its average $1/2$, we expect to see an effective contribution $-\frac{\kappa^2}{\omega^2}$. This approximation was obtained by assuming that $p_2$ is almost constant and equal to $\omega$. Now, when $p_2$ is very large, the energy $H$ is dominated by the contribution $\frac{1}{2} p_2^2$, so that we expect to have $\frac{d}{dt} \mathbb{E}H \sim p_2^2 \frac{d}{dt} \mathbb{E}p_2$. Comparison with (3.1) leads to

$$\frac{d}{dt} \mathbb{E}p_2 \sim -\frac{1}{p_2^2} \sum_b \gamma_b \kappa^2.$$

We will obtain this result rigorously in proposition 3.4.

### 3.2. Notations

Let $\Omega^1 = \{(q, p) \in \Omega : p_2 \neq 0\}$. We denote throughout by $X_t = (q(t), p(t))$ the solution of the stochastic differential equation (1.1) with initial condition $X_0 = (q(0), p(0))$. For now, we restrict ourselves to $X_0 \in \Omega^1$ since we aim to obtain an effective dynamics for the middle rotor by performing an expansion in negative powers of $p_2$. Remark that since $\frac{d}{dt} p_2$ is bounded, there is for each initial condition $X_0 \in \Omega^1$ a deterministic time $t^* > 0$ (proportional to $|p_2(0)|$) such that $X_t \in \Omega^1$ for all $t \in [0, t^*)$ and all realisations of the random noises. To define a smooth Lyapunov function on the whole space $\Omega$ we will perform a regularisation in section 4.

**Definition 3.1.** We let $\mathcal{U}$ be the set of stochastic processes $u_t$, which are solutions of an SDE of the form

$$du_t = f_1(X_t)dt + f_2(X_t)dB^1_t + f_3(X_t)dB^3_t,$$

for some functions $f_j : \Omega \to \mathbb{R}$.
**Notation:** In the sequel, we write
\[ d\sigma_t = f_1 dt + f_2 dB_t^1 + f_3 dB_t^3 \]
instead of (3.2).

For any smooth function \( h \) on \( \Omega \), the stochastic process \( h(X_t) \) is in \( \mathcal{U} \) by the Itô formula (see below). Without further mention, we will see \( h \) both as a function on \( \Omega \) and as the stochastic process \( h(X_t) \). When referring to the stochastic process, we shall write simply \( dh \) instead of \( dh(X_t) \). Of course, only very few processes in \( \mathcal{U} \) can be written in the form \( h(X_t) \) for some function \( h \) on \( \Omega \).

The variables \( p_2 \) and \( q_2 \) will play a special role, as we are merely interested in the regime where \( p_2 \) is very large. For any function \( f \) over \( \Omega \), we call the quantity
\[
\langle f \rangle = \frac{1}{2\pi} \int_0^{2\pi} f dq_2
\]
the \( q_2 \)-average of \( f \) (or simply the average of \( f \)), which is a function of \( p, q_1 \) and \( q_3 \) only.

**Assumption 3.2.** We assume
\[
\langle U_2 \rangle = 0 \quad \text{and} \quad \langle W_b \rangle = 0, \quad b = 1, 3.
\]
This assumption merely fixes the additive constants of the potentials and therefore results in no loss of generality.

For conciseness, we shall omit the arguments of the potentials and forces, always assuming that
\[
W_b = W_b(q_2 - q_b), \quad w_b = w_b(q_2 - q_b), \quad b = 1, 3,
\]
\[
U_i = U_i(q_i), \quad u_i = u_i(q_i), \quad i = 1, 2, 3.
\]

To simplify the notations, we also introduce the potentials \( \Phi_1, \Phi_2, \Phi_3 \) associated to the three rotors, and the corresponding forces \( \phi_1, \phi_2, \phi_3 \) defined by
\[
\Phi_b = W_b + U_b, \quad \phi_b = -\phi_b, \quad \Phi_b = w_b - u_b, \quad b = 1, 3,
\]
\[
\Phi_2 = W_1 + W_3 + U_2, \quad \phi_2 = -\phi_2, \quad \Phi_2 = -w_1 - w_3 - u_2.
\]
(3.3)

Of course, \( \Phi_i \) and \( \phi_i \) are functions of \( q \) only. With these notations the dynamics reads more concisely
\[
dq_i = p_i dt, \quad i = 1, 2, 3,
\]
\[
dp_2 = \phi_2 dt,
\]
\[
dp_b = \left( \phi_b + \tau_b - \gamma_b p_b \right) dt + \sqrt{2\gamma_b T_b} dB_t^b, \quad b = 1, 3.
\]

We will mainly deal with functions of the form \( p_2^\ell p_i^n q(g(q)) \) and their linear combinations. We therefore introduce the notion of degree.

**Definition 3.3.** We say that a function \( f \) on \( \Omega^+ \) has degree \( \ell \in \mathbb{Z} \) if it can be written as a finite sum of elements of the kind \( p_2^\ell p_i^n q(g(q)) \) for some \( n, m \in \mathbb{N} \) and a smooth function \( g : \mathbb{T}^3 \to \mathbb{R} \). Moreover, we denote
\[
\hat{O}(p_2^\ell)
\]
a generic expression of order at most \( \ell \) (which can vary from line to line), i.e. a finite sum of functions of degree \( \ell, \ell - 1, \ell - 2, \ldots \).
We have by the Itô formula that for any smooth function $f$ on $\Omega$

$$
\begin{align*}
\frac{df}{dt} &= \sum_{i=1}^{n} \left( \frac{\partial f}{\partial q_i} dq_i + \frac{\partial f}{\partial p_i} dp_i \right) + \sum_{b} \gamma_b T_b \frac{\partial^2 f}{\partial p_b^2} dt \\
&= d^+ f + d^0 f + d^- f,
\end{align*}
$$

where

$$
\begin{align*}
d^+ f &= p_2 \frac{\partial f}{\partial q_2} dt, \\
d^- f &= \phi_2 \frac{\partial f}{\partial p_2} dt, \\
d^0 f &= \sum_{b} \left( p_b \frac{\partial f}{\partial q_b} + (\phi_b + \tau_b - \gamma_b p_b) \frac{\partial f}{\partial p_b} + \gamma_b T_b \frac{\partial^2 f}{\partial p_b^2} \right) dt \\
&\quad + \sum_{b} \sqrt{2\gamma_b T_b} \frac{\partial f}{\partial p_b} dB_b^b.
\end{align*}
$$

(By the discussion following definition 3.1, $f$, its partial derivatives, $p_2$ and the functions $\phi_i$ in this SDE are evaluated on the trajectory $X_t$.) Observe that when acting on a function of degree $\ell$, the contribution $d^+$ increases the degree of $p_2$ by one, while $d^0$ and $d^-$ respectively leave it unchanged and decrease it by one. In this sense, we will see $d^+$ as the ‘dominant’ part of $d$.

### 3.3. General idea

In this section we introduce the main idea, which consists of successively removing oscillatory terms order by order in the dynamics of $p_2$. We perform here the first step of the method in a somewhat naive, but pedestrian way. In the next two sections, we systematise the method and apply it.

We begin by looking at the equation

$$
\begin{align*}
\frac{dp_2}{dt} &= \phi_2 \frac{\partial f}{\partial p_2} dt,
\end{align*}
$$

(3.5)

When $p_2$ is large while $p_1$ and $p_3$ are small, the right-hand side is highly oscillatory and its time-average is almost zero, since $\langle \phi_2 \rangle = 0$. We will proceed to a change of variable in order to ‘see through’ this oscillatory term.

We first make the relation between the time-average and the $q_2$-average more precise. Consider some function $g$ on $\Omega$. In the regime where $p_2$ is very large and $p_1$, $p_3$ are small, the only fast variable is $q_2$. Now consider some interval of time $[0, T]$ short enough so that the other variables do not change significantly, but still large enough for $q_2$ to swipe through $[0, 2\pi)$ many times. We have in that case

$$
\begin{align*}
g(q(t), p(t)) &\sim g(q_1(0), q_2(0) + p_2(0)t, q_3(0), p(0)),
\end{align*}
$$

(3.6)

so that the time-average of $g$ is expected to be very close to the $q_2$-average $\langle g \rangle$.

Now, we want to estimate $p_2(t) = \int_0^t \phi_2(q(s)) ds$ in this situation. Approximating $\phi_2$ as in (3.6) and integrating formally with respect to time (remember that $\phi_2 = -\partial_{q_2} \Phi_2$) leads naturally to the decomposition

$$
\begin{align*}
p_2 &= \tilde{p}_2 - \frac{\Phi_2(q)}{p_2},
\end{align*}
$$

(3.7)
which consists in writing $p_2$ as sum of an oscillatory term $\Phi_2/p_2$ which is supposed to capture ‘most’ of the oscillatory dynamics, and some (hopefully) nicely behaved ‘slow’ process $\bar{p}_2$. And indeed, if we differentiate (3.7) we get
\[
\begin{align*}
d\bar{p}_2 &= d\left(p_2 + \frac{\Phi_2}{p_2}\right) \\
&= dp_2 + d^\Phi_2 + d^{\Phi_2} + d^\Phi_2 \\
&= \phi_2 \, dt - \phi_2 \, dt - \left(\frac{p_1 w_3}{p_2} + \frac{p_3 w_3}{p_2}\right) \, dt - \frac{\phi_2 \Phi_2}{p_2^2} \, dt \\
&= - \left(\frac{p_1 w_3}{p_2} + \frac{p_3 w_3}{p_2}\right) \, dt - \frac{\phi_2 \Phi_2}{p_2^2} \, dt.
\end{align*}
\]

As a result, we have a new process $\bar{p}_2$ which is asymptotically equal to $p_2$ in the regime of interest, and whose dynamics involves only terms that are small when $p_2$ is large, so that $\bar{p}_2$ is indeed a slow variable. Observe that the choice of adding $\Phi_2/p_2$ to $p_2$ has the effect that
\[
d^\Phi_2 = -\phi_2 \, dt,
\]
which precisely cancels the right-hand side of (3.5) while the remaining terms have negative powers of $p_2$. This observation is the starting point of the systematisation of the method.

Unfortunately, (3.8) is not good enough to understand how $\bar{p}_2$ (and therefore $p_2$) decreases in the long run, since the dynamics (3.8) of $\bar{p}_2$ still involves oscillatory terms. The idea is therefore to eliminate these oscillatory terms by absorbing them into a further change of variable $\bar{p}_2 = \bar{\bar{p}}_2 + G$ for some suitably chosen $G$. The result is that $d\bar{\bar{p}}_2$ is a sum of terms of degree $-2$ at most, which turn out to be still oscillatory. This procedure must then be iterated, successively eliminating oscillatory terms order by order, until we get some dynamics that has a non-zero average (which happens after finitely many steps). We will follow this idea, but in a way that does not require one to write the successive changes of variable explicitly. More precisely, we will prove

**Proposition 3.4.** There is a function $F = \frac{\Phi_2(q)}{p_2} + \hat{O}(p_2^{-3})$ such that whenever $p_2(t) \neq 0$ the process $\tilde{p}_2(t) = p_2(t) + F(X_t)$ satisfies
\[
\begin{align*}
d\tilde{p}_2(t) &= a(X_t) \, dt + \sum_b \sigma_b(X_t) dB^b_t, \quad (3.9)
\end{align*}
\]

with
\[
a(q, p) = -\gamma_1 \langle W_1^2 \rangle + \gamma_3 \langle W_3^2 \rangle + \hat{O}(p_2^{-4}) ,
\]
\[
\sigma_b(q, p) = \sqrt{2\gamma_b T_b W_b} \frac{p_2}{p_2^2} + \hat{O}(p_2^{-3}), \quad b = 1, 3.
\]

(By assumption 3.2, no arbitrary additive constant appears in $\langle W_1^2 \rangle$ and $\langle W_3^2 \rangle$.)

The next two sections are devoted to proving proposition 3.4.

### 3.4. Averaging

The crux of our analysis is to average oscillatory terms in the dynamics. This is a well known problem in differential equations. In classical averaging theory [13, 15], it is an external small parameter $\varepsilon$ that gives the time scale of the fast variables. Here, the role of $\varepsilon$ is played by $1/p_2$, which is a dynamical variable. We develop an averaging theory adapted to this case, and also to the stochastic nature of the problem.
The starting point is as follows. Imagine that for a function \( h \) on \( \Omega \) we find an expression of the kind
\[
dh = f \, dt + dr_t, \tag{3.10}
\]
for some function \( f = f(X_t) \) of degree \( \ell \) and some stochastic process \( r_t \in \mathcal{U} \) (see definition 3.1) which denotes the part of the dynamics that we do not want to interfere with. Thinking of \( f(X_t) \) as a highly oscillatory quantity when \( p_2 \) is very large, we would like to write
\[
h = \bar{h} + F
\]
for some small function \( F \) on \( \Omega \) such that
\[
d\bar{h} = \langle f \rangle \, dt + dr_t \quad \text{small corrections}, \tag{3.11}
\]
where the notion of small will be made precise in terms of powers of \( p_2 \). That is, we want to find some \( \bar{h} \) close to \( h \), such that its dynamics involves, instead of \( f \, dt \), the \( q_2 \)-average \( \langle f \rangle dt \) plus some smaller corrections. In other words, we are looking for some \( F \) such that
\[
dF = (h - \bar{h}) = (f - \langle f \rangle) \, dt + \text{small corrections}. \tag{3.12}
\]
Remembering that in terms of powers of \( p_2 \), \( d^+ \) is the dominant part of \( d \), the key is to find some \( F \) such that \( d^+ F = (f - \langle f \rangle) \, dt \). If we write \( L^+ = p_2 \partial_{q_2} \), we have \( d^+ F = L^+ F \, dt \). Thus, we really need to invert \( L^+ \) (which is in fact the dominant part of the generator \( L \) when \( p_2 \) is large).

We call here \( K \) the space of smooth functions \( \Omega^1 \to \mathbb{R} \), and we denote by \( K_0 \) the space of functions \( f \in K \) such that \( \langle f \rangle = 0 \). Note that \( L^+ \) maps \( K \) to \( K_0 \) since for all \( f \in K \), we have by periodicity
\[
\langle L^+ f \rangle = p_2 \langle \partial_{q_2} f \rangle = 0.
\]
We can define a right inverse \( (L^+)^{-1} : K_0 \to K_0 \) by letting for all \( g \in K_0 \)
\[
(L^+)^{-1} g = \frac{1}{p_2} \left( \int g \, dq_2 + c(p, q_1, q_3) \right),
\]
where the integration ‘constant’ \( c(p, q_1, q_3) \) is uniquely defined by requiring that \( \langle (L^+)^{-1} g \rangle = 0 \).

This leads naturally to the following

**Definition 3.5.** For any function \( f \in K \), we define the operator \( Q : K \to K_0 \) by
\[
Qf = (L^+)^{-1} (f - \langle f \rangle).
\]

**Remark 3.6.**

- If \( f \) is a function of degree \( \ell \), then \( Qf \) is of degree \( \ell - 1 \).
- By construction,
\[
d(Qf) = (f - \langle f \rangle) \, dt + d^0(Qf) + d^- (Qf). \tag{3.12}
\]
- Moreover, by definition, \( Qf \) is the only function such that
\[
\partial_{q_2} (Qf) = \frac{f - \langle f \rangle}{p_2} \quad \text{and} \quad \langle Qf \rangle = 0. \tag{3.13}
\]

Therefore, if (3.10) holds for some \( f \) of degree \( \ell \), then we obtain a quantitative expression for (3.11), namely
\[
d(h - Qf) = (f) \, dt + dr_t - d^0(Qf) - d^- (Qf),
\]
where the corrections are small in the sense that \( Qf \), \( d^0(Qf) \) and \( d^- (Qf) \) have degree respectively \( \ell - 1 \), \( \ell - 1 \) and \( \ell - 2 \).
Remark 3.7. Observe that (3.7) can be written now as $p_2 = \tilde{p}_2 + Q\phi_2$, since $Q\phi_2 = -\Phi_2/p_2$. Thus, the ‘naive’ correction we added in (3.7) also follows from the systematic method we have just introduced. This is no surprise: the naive correction in (3.7) was motivated by the approximation (3.6) in which only $q_2$ moves, which corresponds to considering only $d^+$. 

Remark 3.8. Our averaging procedure is inspired by techniques of [7]. There, the equivalent of $L^+$ is the generator $-q_2^2k^{-1}\partial_{p_2} + p_2\partial_{q_2}$ of the free dynamics of the middle oscillator, where $q_2^2/(2k)$ is the pinning potential. In their case, one cannot explicitly invert $L^+$, but one can show that $(L^+)^{-1}$ basically acts as a division by $E_2^{1/2}$, where $E_2$ is the energy of the middle oscillator. Again, taking formally the limit $k \to \infty$, one obtains that $(L^+)^{-1}$ acts as a division by $\sqrt{E_2}$, much like in our case where $(L^+)^{-1}$ acts as a division by $p_2 \sim \sqrt{E_2}$.

We now restate our averaging method as the following lemma, which follows from a trivial rearrangement of the terms in (3.12).

Lemma 3.9. (Averaging lemma) Consider some function $f = \hat{O}(p_2^{\ell})$ for some $\ell \in \mathbb{Z}$. Then

$$fdt = \langle f \rangle dt - d^0(Qf) - d^- (Qf) + d(Qf),$$

where $d^0(Qf)$ is of degree $\ell - 1$ at most and $d^- (Qf)$ is of degree $\ell - 2$ at most.

We now prove proposition 3.4 by using lemma 3.9 repeatedly.

3.5. Proof of proposition 3.4

We make the following observations, which we will use without reference. For any function $f$ on $\Omega^1$ that is smooth in $q_2$, we have by periodicity

$$\langle \partial_{q_2} f \rangle = 0. \quad (3.14)$$

Moreover, if $g$ is another such function, then we can integrate by parts to obtain

$$\langle (\partial_{q_2} f) g \rangle = -\langle f \partial_{q_2} g \rangle.$$

Furthermore, we have by assumption 3.2, (3.3) and (3.14) that

$$\langle W_b \rangle = \langle w_b \rangle = \langle \Phi_2 \rangle = \langle \phi_2 \rangle = 0.$$

We start by doing again the first step, which we did in section 3.3, but this time using the new toolset. In order to average the right-hand side of

$$dp_2 = \phi_2 dt,$$

we use lemma 3.9 with $f = \phi_2$, which is of order 0. We have $\langle f \rangle = 0$ and $Qf = -\Phi_2/p_2$ (by definition of $\phi_2$ and $\Phi_2$). We obtain

$$dp_2 = d^0\left(\frac{\Phi_2}{p_2}\right) + d^- \left(\frac{\Phi_2}{p_2}\right) - d\left(\frac{\Phi_2}{p_2}\right) = \frac{1}{p_2} \sum_b p_b \frac{\partial \Phi_2}{\partial q_b} dt - \frac{\phi_2 \Phi_2}{p_2^2} dt - d\left(\frac{\Phi_2}{p_2}\right), \quad (3.15)$$

This is exactly what we found in (3.8). We deal next with the terms $-p_b w_b / p_2$ in (3.15). Using lemma 3.9 with $f = p_b w_b / p_2$ (and therefore with $Qf = p_b W_b / p_2^2$), we find, since
\[ \langle f \rangle = \frac{p_b \langle w_b \rangle}{p_2^2} p_2 = 0, \text{ that for } b = 1, 3, \]
\[
\frac{p_b w_b}{p_2^2} \frac{dt}{d} = -\frac{1}{p_2^2} \left[ -p_b^2 w_b + (\phi_b + \tau_b - \gamma_b p_b) W_b \right] dt
\]
\[
- \frac{1}{p_2^2} \sqrt{2\gamma_b T_b} W_b dB^b_t + \frac{2}{p_2^2} p_b W_b \phi_2 \frac{dt}{d} + d\hat{O}(p_2^{-2}), \quad (3.16)
\]

where here and in the sequel, we denote by \(d\hat{O}(p_2^k)\) any generic expression of the kind \(dW(X_t)\) for some function \(w = \hat{O}(p_2^k)\) on \(\Omega\). Here \(d\hat{O}(p_2^{-2}) = d(p_2 W_b p_2^{-2})\). Substituting (3.16) into (3.15) leads to

\[
dp_2 = I \frac{dt}{d} + J \frac{dt}{d} + \frac{1}{p_2^2} \sum_b \sqrt{2\gamma_b T_b} W_b dB^b_t + \frac{\Phi_2}{p_2} + \hat{O}(p_2^{-2}), \quad (3.17)
\]

with
\[
I = -\sum_b \frac{p_b^2 w_b - (\phi_b + \tau_b - \gamma_b p_b) W_b}{p_2^2} - \frac{\Phi_2}{p_2},
\]
\[
J = \frac{2}{p_2^2} \sum_b p_b W_b \phi_2.
\]

We next deal with the terms \(I \frac{dt}{d} \) and \(J \frac{dt}{d} \).

First, we show that \(\langle I \rangle = 0\). It is immediate that \(\langle p_2^{-2} p_b^2 w_b \rangle\) and \(\langle p_2^{-2} (\tau_b - \gamma_b p_b) W_b \rangle\) are zero. Moreover, \(\langle p_2^{-2} \phi_2 \Phi_2 \rangle = -\frac{1}{2} p_2^{-2} \langle \phi_2 \Phi_2 \rangle = 0\). Thus,
\[
\langle I \rangle = \sum_b \left( \frac{1}{p_2^2} \phi_b W_b \right) = \sum_b \left( \frac{w_b - u_b W_b}{p_2^2} \right)
\]
\[
= -\sum_b \left( \frac{u_b (W_b)}{2 p_2^2} + u_b (W_b) \right) = 0.
\]

Since \(I\) is of order \(-2\) and \(\langle I \rangle = 0\), we find that \(Q I\) is of order \(-3\) and thus \(d^- (Q I) = \hat{O}(p_2^{-4}) \frac{dt}{d}\). Applying lemma 3.9 with \(f = I\), we find
\[
I \frac{dt}{d} = -d^0 (Q I) + \hat{O}(p_2^{-4}) \frac{dt}{d} + d\hat{O}(p_2^{-3}). \quad (3.18)
\]

Using that \(\langle Q I \rangle = 0\), the definition (3.4) of \(d^0\) leads, upon inspection, to
\[
d^0 (Q I) = \sum_b w_b \partial_{p_b} (Q I) \frac{dt}{d} + \mathcal{E} \frac{dt}{d} + \sum_b \hat{O}(p_2^{-3}) \frac{dB^b_t}{d},
\]

where \(\mathcal{E}\) is a sum of terms of order \(-3\) and \(\langle \mathcal{E} \rangle = 0\). Applying lemma 3.9 to \(w_b \partial_{p_b} (Q I) \frac{dt}{d}\) and \(\mathcal{E} \frac{dt}{d}\), we obtain
\[
d^0 (Q I) = \sum_b \left[ w_b \partial_{p_b} (Q I) \right] \frac{dt}{d} + \hat{O}(p_2^{-4}) \frac{dt}{d} + \sum_b \hat{O}(p_2^{-3}) \frac{dB^b_t}{d}. \quad (3.19)
\]

Using the definition of \(w_b\), integrating by parts once and using (3.13), we have for \(b = 1, 3, \)
\[
\langle w_b \partial_{p_b} (Q I) \rangle = \langle \partial_{q_b} (W_b) Q(\partial_{p_b} I) \rangle = -(W_b \partial_{q_b} Q(\partial_{p_b} I)) = -\frac{1}{p_2^2} (W_b \partial_{p_b} I).
\]

Since \(\partial_{p_b} I = -p_2^{-2} (2 p_b w_b + \gamma_b W_b)\), we get
\[
\langle w_b \partial_{p_b} (Q I) \rangle = \left( \frac{1}{p_2^2} W_b (2 p_b w_b + \gamma_b W_b) \right) = \frac{1}{p_2^2} \gamma_b \langle W_b^2 \rangle, \quad (3.20)
\]

\[\text{Nonlinearity 28 (2015) 2397} \quad \text{N Cuneo et al}\]
where again we have used that $\langle W_b w_b \rangle = \frac{1}{2} \langle \partial_{q_b} W_b^2 \rangle = 0$. Substituting (3.20) into (3.19) and then the result into (3.18) we finally get

$$I \, dt = -\frac{\alpha}{p_2^3} \, dt + \sum_b \dot{\partial} (p_2^{-3}) \, dB_b + \dot{\partial} (p_2^{-4}) \, dt + d\dot{\partial} (p_2^{-3}).$$

(3.21)

where

$$\alpha = \sum_b \gamma_b \langle W_b^2 \rangle.$$

We next deal with the term $J \, dt$ of (3.17). First, by lemma 3.9,

$$J \, dt = \langle J \rangle \, dt + \dot{\partial} (p_2^{-4}) \, dt + \sum_b \dot{\partial} (p_2^{-4}) \, dB_b + d\dot{\partial} (p_2^{-4}).$$

(3.22)

Unfortunately, $\langle J \rangle \neq 0$, and we will need some more subtle identifications. Integrating by parts, we have

$$\langle J \rangle = \frac{2p_b}{p_2^3} \sum_b \langle W_b \Phi_2 \rangle = -\frac{2p_b}{p_2^3} \sum_b \langle W_b \partial_q \Phi_2 \rangle$$

$$= \frac{2p_b}{p_2^3} \sum_b \langle \partial_{q_b} W_b \Phi_2 \rangle = \frac{2p_b}{p_2^3} \sum_b \langle w_b \Phi_2 \rangle$$

(3.23)

$$= -\frac{1}{p_2^3} \sum_b p_b \partial_{q_b} \langle \Phi_2^2 \rangle.$$

On the other hand, since $p_2^{-3} \langle \Phi_2^2 \rangle$ does not depend on $q_2$, we find $d^\ast (p_2^{-3} \langle \Phi_2^2 \rangle) = 0$, so that

$$d \left( \frac{\langle \Phi_2^2 \rangle}{p_2^3} \right) = d^0 \left( \frac{\langle \Phi_2^2 \rangle}{p_2^3} \right) + d^\ast \left( \frac{\langle \Phi_2^2 \rangle}{p_2^3} \right)$$

$$= \sum_b p_b \partial_{q_b} \left( \frac{\langle \Phi_2^2 \rangle}{p_2^3} \right) \, dt + \dot{\partial} (p_2^{-4}) \, dt.$$

(3.24)

Combining (3.23) and (3.24) we find

$$\langle J \rangle \, dt = \dot{\partial} (p_2^{-4}) \, dt + d(p_2^{-3} \langle \Phi_2^2 \rangle) = \dot{\partial} (p_2^{-4}) \, dt + d\dot{\partial} (p_2^{-3}),$$

so that from (3.22) we obtain

$$J \, dt = \dot{\partial} (p_2^{-4}) \, dt + \sum_b \dot{\partial} (p_2^{-4}) \, dB_b + d\dot{\partial} (p_2^{-3}).$$

This together with (3.17) and (3.21) finally shows that

$$d p_2 = -\left( \frac{\alpha + \dot{\partial} (p_2^{-4})}{p_2^3} \right) \, dt + \sum_b \left( \sqrt{2} \gamma_b T_b W_b \frac{p_b}{p_2^3} + \dot{\partial} (p_2^{-3}) \right) \, dB_b + d \left( -\frac{\Phi_2}{p_2^2} + \dot{\partial} (p_2^{-2}) \right).$$

which implies (3.9) and completes the proof of proposition 3.4.

**Remark 3.10.** We can argue (in a nonrigorous way) that when $|p_2|$ is very large, the dynamics of $\tilde{p}_2$ is approximately that of a particle interacting with two ‘effective’ heat baths at temperatures $T_1$ and $T_3$, but with some coupling of magnitude $p_2^{-4}$. Indeed, we can write (3.9) in the canonical ‘Langevin’ form

$$d \tilde{p}_2 (t) = \sum_b \left( -\gamma_b (X_b) \tilde{p}_2 (t) dt + \sigma_b (X_b) dB_b \right),$$

(3.9)

For example if $W_b = -\cos (q_2 - q_b)$, there are in $\langle J \rangle$ some terms of the kind $\langle p_3 \cos (q_2 - q_1) \sin (q_2 - q_1) \rangle$ and $\langle p_1 \sin (q_2 - q_1) \cos (q_2 - q_3) \rangle$ which are non-zero.
We now prove proposition 2.2. Throughout this section, we find is optimal.

**Remark 3.11.** The ergodicity of 1D Langevin processes is well understood: for any \( \sigma > 0 \), processes satisfying an SDE of the kind \( d\xi_t = \beta H \) with \( \beta \) large, \( \tau(t) \) is very close to \( \beta^{-1} \). This approximation indeed holds with effective temperature \( \beta T_b = \beta^{-1} \).

But then, by the Dambis–Dubins-Schwarz representation theorem, there is another Brownian motion \( \tilde{B}_b \) such that \( M(t) = \tilde{B}_b(t) \) with \( \tau(t) = \int_{s}^{t} W_b(s) ds \). Clearly, when \( |p_2| \) is very large, \( \tau(t) \approx t \) so that \( M(t) \) is very close to \( \tilde{B}_b \). In this sense, when \( |p_2| \to \infty \), it is reasonable to approximate \( \langle \sqrt{2\gamma_T} T_b W_b/p_2^2 \rangle d\tilde{B}_b \) with \( \langle \sqrt{2\gamma_T} T_b (W_b^2)/p_2^2 \rangle d\tilde{B}_b \), so that the Einstein-Smoluchowski relation indeed holds with effective temperature \( T_b = T_h \).

4. **Lyapunov function**

We now prove proposition 2.2. Throughout this section, \( \tilde{p}_2 \) is the function defined in proposition 3.4. The basic idea is to consider a Lyapunov function

\[
V \sim \rho(p) \tilde{p}_2 e^{\beta \tilde{p}_2} + e^{\beta H},
\]

where \( \rho(p) \) is non-zero only when \( |p_2| \) is much larger than \( |p_1| \) and \( |p_3| \). We will obtain that \( LV \lesssim -\phi(V) \), with \( \phi(s) \sim s/\log(s) \) as in proposition 2.2. The fact that we do not have a bound of the kind \( LV \lesssim -cV \) (which would yield exponential ergodicity) comes from the very slow decay of \( p_2 \). The basic idea is that, when \( p_2 \to \infty \) and \( p_1, p_3 \to 0 \),

\[
L \tilde{p}_2 \sim -p_2^{-3}, \quad \text{so that} \quad L \left( \tilde{p}_2 e^{\beta \tilde{p}_2} \right) \sim -e^{\beta \tilde{p}_2} \sim -\frac{V}{p_2^2} \sim -\frac{V}{\log V}.
\]

We now introduce the necessary tools to make this observation rigorous.

**Lemma 4.1.** For \( \beta > 0 \) small enough, there are constants \( C_1, C_2 > 0 \) such that

\[
L e^{\beta H} \leq (C_1 - C_2 (p_1^2 + p_3^2)) e^{\beta H}.
\]
Lemma 4.3. For terms of order less than or equal to $-1$. Using this, we obtain a constant. Therefore, any given nonlinearity

\[ d \hat{\beta} \leq C_3 + O(p_2^{-1})e^{\hat{\beta}}. \]  

(4.1)

Proof. Introducing $f(s) = s^2e^{\hat{s}^2}$, we have by the Itô formula and proposition 3.4 that

\[ d(\hat{p}_2^2 e^{\hat{s}^2}) = df(\hat{p}_2) = f'(\hat{p}_2)(a dt + \sum_b \sigma_b dB^b_{t}) + \frac{1}{2} f''(\hat{p}_2) \sum_b \sigma_b^2 dt \]

\[ = (2\hat{p}_2 + \beta \hat{p}_2^3)e^{\hat{s}^2}(a dt + \sum_b \sigma_b dB^b_{t}) + \frac{1}{2} (2 + 5\beta \hat{p}_2 + \beta^2 \hat{p}_2^3) e^{\hat{s}^2} \sum_b \sigma_b^2 dt. \]

Now since $a = -\alpha p_2^{-3} + \bar{\alpha}(p_2^{-4})$ with $\alpha = \sum_b \gamma_b(W_b^2)$, $\sigma_b = \sqrt{2}\gamma_b T_b W_b p_2^{-2} + \bar{\alpha}(p_2^{-4})$, and $\hat{p}_2^2 = p_2^2 + O(p_2^{-1})$ for all $k$, we find after taking the expectation value

\[ L(\hat{p}_2^2 e^{\hat{s}^2}) = (-\alpha\beta + \beta^2 \sum_b \gamma_b T_b W_b^2 + \bar{\alpha}(p_2^{-1})) e^{\hat{s}^2}, \]

which gives the desired bound if $\beta$ is small enough (recall that the $W_b^2$ are bounded). □

Convention: We fix $\beta > 0$ small enough so that the conclusions of lemma 4.1 and lemma 4.2 hold.

Let $k \geq 1$ be an integer and $R > 0$ be a constant (which we will fix later). We split $\Omega$ into three disjoint sets $\Omega_1$, $\Omega_2$, $\Omega_3$ defined by

- $\Omega_1 = \{x \in \Omega : |p_2| < (p_2^2 + p_3^2)^k + R\}$,
- $\Omega_2 = \{x \in \Omega : (p_2^2 + p_3^2)^k + R \leq |p_2| \leq 2(p_2^2 + p_3^2)^k + 2R\}$,
- $\Omega_3 = \{x \in \Omega : |p_2| > 2(p_2^2 + p_3^2)^k + 2R\}$.

Fix some $m, n \in \mathbb{N}$ and $\ell \geq 1$. On $\Omega_2 \cup \Omega_3$, we have by definition $|p_2| \geq (p_2^2 + p_3^2)^k + R$, so that

\[ \left| \frac{p_2^m p_3^n}{p_2^2} \right| \leq \frac{|p_2^m p_3^n|}{((p_2^2 + p_3^2)^k + R)^\ell} \quad \text{on } \Omega_2 \cup \Omega_3. \]

Clearly, if $k$ and $R$ are large enough, the right-hand side is bounded by an arbitrarily small constant. Therefore, any given $O(p_2^{-1})$ is also bounded by an arbitrarily small constant on $\Omega_2 \cup \Omega_3$ provided that $k$ and $R$ are large enough, since it is by definition a sum finitely many terms of order less than or equal to $-1$. Using this, we obtain

Lemma 4.3. For $k$ and $R$ large enough, there are constants $C_4, \ldots, C_7 > 0$ such that the following properties hold on $\Omega_2 \cup \Omega_3$:

\[ |\hat{p}_2^2 - p_2^2| < C_4, \]  

(4.2)

\[ L(\hat{p}_2^2 e^{\hat{s}^2}) \leq -C_5e^{\hat{s}^2}, \]  

(4.3)

\[ C_6 e^{-\frac{\beta}{2}(p_2^2 + p_3^2)}e^{\hat{s}^2} \leq e^{\hat{s}^2} \leq C_7 e^{-\frac{\beta}{2}(p_2^2 + p_3^2)}e^{\hat{s}^2}. \]  

(4.4)
Proof. Since \( \tilde{p}_2 = \frac{p_2 + \Phi_2(q)}{p_2 + \tilde{\mathcal{O}}(p_2^{-2})} \), we have \( \tilde{p}_2^2 = p_2^2 + 2\Phi_2(q) + \tilde{\mathcal{O}}(p_2^{-1}) \). By taking \( k \) large enough, the \( \tilde{\mathcal{O}}(p_2^{-1}) \) here is bounded by a constant on the set \( \Omega_2 \cup \Omega_3 \), which implies (4.2). Moreover, for large \( k \) and \( R \), the \( \tilde{\mathcal{O}}(p_2^{-1}) \) in (4.1) is also bounded on \( \Omega_2 \cup \Omega_3 \) by an arbitrarily small constant, which implies (4.3). To prove (4.4), observe that

\[
e^\frac{\xi}{k} \tilde{p}_2^2 = e^\frac{\xi}{k} (p_2^2 - p_2^2 - p_2 - U(q)) e^{\beta H},
\]

where \( U(q) \) contains all the potentials appearing in \( H \). This together with the boundedness of \( U \) and (4.2), implies (4.4). \( \square \)

Convention: We fix \( k \) and \( R \) such that the conclusions of lemma 4.3 hold.

Definition 4.4. Let \( \chi : \mathbb{R} \rightarrow [0, 1] \) be a smooth function such that \( \chi(s) = 0 \) when \( |s| < 1 \) and \( \chi(s) = 1 \) when \( |s| > 2 \). We introduce the cutoff function

\[
\rho(p) = \chi \left( \frac{p_2}{|p_2| + C_2^3 + R} \right),
\]

and the Lyapunov function

\[
V = 1 + A \rho(p) \tilde{p}_2^2 e^{\frac{\xi}{k} \tilde{p}_2^2} + e^{\beta H},
\]

with \( A > 0 \) (to be chosen later).

By construction \( \rho(p) \) is smooth, \( \rho(p) = 0 \) on \( \Omega_1 \) and \( \rho(p) = 1 \) on \( \Omega_3 \), with some transition on \( \Omega_2 \). Remember that \( \tilde{p}_2 \) is by construction smooth on \( \Omega^2 \), i.e. when \( p_2 \neq 0 \). In particular, since \( \Omega_2 \cup \Omega_3 \subset \Omega^2 \), the function \( \rho(p) \tilde{p}_2^2 e^{\frac{\xi}{k} \tilde{p}_2^2} \) is smooth on \( \Omega \), and so is \( V \). We can now finally give the

Proof of proposition 2.2. We show here that \( V \) satisfies the conditions enumerated in proposition 2.2 if \( A \) is large enough. Let us first prove the first statement, which is that there exist \( c_1, c_2 > 0 \) such that

\[
1 + c_1 e^{\beta H} \leq V \leq c_2 (1 + p_2^2) e^{\beta H}. \tag{4.5}
\]

Clearly the lower bound on \( V \) holds. We now prove the upper bound. Throughout the proof, we denote by \( c \) a generic positive constant which can be each time different. Since \( \rho \neq 0 \) only on \( \Omega_2 \cup \Omega_3 \), we have by (4.2) and (4.4),

\[
|A \rho(p) \tilde{p}_2^2 e^{\frac{\xi}{k} \tilde{p}_2^2}| \leq c(p_2 + 3^2 e^{-\frac{\xi}{k} (p_2^2 + p_2^4)}) e^{\beta H} \leq c(2 + 3^2 + C_4^3) e^{\beta H} \leq c(1 + p_2^2) e^{\beta H}.
\]

But then \( V \leq 1 + c(1 + p_2^2) e^{\beta H} \leq c(1 + p_2^2) e^{\beta H} \), where the last inequality holds because \( H \) is bounded below, so that \( e^{\beta H} \) is bounded away from zero.

Let us now move to the second statement of proposition 2.2, which is that for \( c_3, c_4 \) large enough and a compact set \( K \),

\[
LV \leq c_3 1_K - \phi(V) \quad \text{with} \quad \phi(s) = \frac{c_4 s}{2 + \log(s)}. \tag{4.6}
\]

We first show that

\[
LV \leq c_3 1_K - c e^{\beta H} \quad \text{with} \quad K = \{x \in \Omega_1 \cup \Omega_2 : p_1^2 + p_2^2 \leq M\}, \tag{4.7}
\]

for some large enough \( M \). Clearly \( K \) is compact, since \( \Omega_1 \cup \Omega_2 = \{x \in \Omega : |p_2| \leq 2(p_1^2 + p_2^4)^{\frac{1}{4}} + 2R\} \).

On \( \Omega_1 \) we simply have \( V = 1 + e^{\beta H} \). By lemma 4.1, we have \( LV \leq (C_1 - C_2(p_1^2 + p_2^2)) e^{\beta H} \). Since \( \Omega_1 \setminus K = \{x \in \Omega_1 : p_1^2 + p_2^2 > M\} \), we have for large enough \( M \) that \( LV \leq -c e^{\beta H} \) on \( \Omega_1 \setminus K \), and therefore (4.7) holds on \( \Omega_1 \).
On $\Omega_2$, the key is to observe that there is a polynomial $z(p_1, p_2, p_3)$ such that
\[
|L(A\rho(p)p_2^2e^{\frac{z}{p_1}})| \leq z(p)e^{-\frac{1}{2}(p_1^2 + p_2^2)}e^{\beta H},
\]
where the second inequality comes from (4.4). Now, since $p_1^2 + p_2^2 \approx |p_2|^{1/k}$ on $\Omega_2$, we have that $z(p)e^{-\frac{1}{2}(p_1^2 + p_2^2)}$ is bounded on $\Omega_2$. Therefore, by this and lemma 4.1, we have on $\Omega_2$,
\[
LV \leq \left( C_7 z(p)e^{-\frac{1}{2}(p_1^2 - p_2^2)} + C_1 - C_2(p_1^2 + p_2^2) \right) e^{\beta H}
\]
\[
\leq (c - C_2(p_1^2 + p_2^2))^5 e^{\beta H}.
\]
which, as in the previous case, implies that (4.7) holds on $\Omega_2$ if $M$ is large enough.

On $\Omega_3$, which is the critical region, we have $V = 1 + A \tilde{p}_2^2 e^{\frac{z}{p_1}} + e^{\beta H}$. By lemma 4.1 and (4.3), it holds in $\Omega_3$ that
\[
LV \leq (C_1 - C_2(p_1^2 + p_2^2))e^{\beta H} - C_5 A e^{\frac{z}{p_1}}. \tag{4.8}
\]
On the set $\{x \in \Omega_3 : C_1 - C_2(p_1^2 + p_2^2) < -1 \}$, we simply have $LV \leq -e^{\beta H}$, so that (4.7) holds trivially. On the other hand, on the set $\{x \in \Omega_3 : C_1 - C_2(p_1^2 + p_2^2) > -1 \}$ the quantity $p_1^2 + p_2^2$ is bounded, so that $e^{\frac{z}{p_1}} \geq ce^{\beta H}$ by (4.4), which with (4.8) implies that
\[
LV \leq (C_1 - C_2(p_1^2 + p_2^2))e^{\beta H} - c Ae^{\beta H} \leq (C_1 - cA)e^{\beta H}.
\]
By making $A$ large enough, we again find a bound $LV \leq -ce^{\beta H}$, so that (4.7) holds.

Therefore, (4.7) holds on all of $\Omega$. To obtain (4.6), we need only show that $e^{\beta H} \geq cV/(2 + \log V)$. By the boundedness of the potentials and the definition of $V$, we have $1 + p_2^2 \leq 2H + c \leq c \log(e^{\beta H}) + c \leq c \log V + c \leq c(\log V + 2)$. But then by (4.5) we indeed have that $e^{\beta H} \geq cV/(1 + p_2^2) \geq cV/(2 + \log V)$. This completes the proof of proposition 2.2. \hfill \Box

Remark 4.5. The external forces $r_b$ and the pinning potentials $U_i$ (if non-zero) do not play a central role in the properties of the Lyapunov function. On the contrary, the interaction potentials $W_b$ are very important, since we need $\alpha = \sum_b \gamma_b \{W_b^2\}$ to be strictly positive.

Remark 4.6. Although we assume throughout that $T_1$ and $T_3$ are strictly positive, the computations that lead to the Lyapunov function apply to zero temperatures as well (the temperatures only appear in some non-dominant terms in $V$ and $LV$). In that case, the existence of an invariant measure can still be obtained by compactness arguments (see e.g. proposition 5.1 of [7]). However, the smoothness, uniqueness and convergence assertions do not necessarily hold: when $T_1 = T_3 = 0$ the system is deterministic, the transition probabilities are not smooth, and there is at least one invariant measure concentrated at each stationary point of the system. The positive temperatures assumption is crucial in the next section.

5. Smoothness and irreducibility

This section is devoted to proving that the hypotheses of theorem 2.1 other than the existence of the Lyapunov function are satisfied. More precisely we will prove the following proposition.

Proposition 5.1. The following properties hold.

(i) The transition probabilities $P^t(x, \cdot)$ have a density $p_t(x, y)$ that is smooth in $(t, x, y)$ when $t > 0$. In particular, the process is strong Feller.
(ii) The time-1 skeleton \((X_n)_{n=0,1,2,\ldots}\) is irreducible, and the Lebesgue measure \(m\) on \((\Omega, \mathcal{B})\) is a maximal irreducibility measure.

(iii) Every compact set is petite.

In a sense, (i) shows that we have some effective diffusion in all directions at very short times, and (ii) shows that every part of the phase space is eventually reached with positive probability. Observe that (iii) follows from (i) and (ii). Indeed, by (i), (ii) and proposition 6.2.8 of [10], every compact set is petite for the time-1 skeleton. But then every compact set is also petite with respect to the process \(X_t\) (simply by choosing a sampling measure on \([0, \infty)\) that is concentrated on \(\mathbb{N}\)). Therefore, we need only prove (i) and (ii), which we do in the next two subsections.

5.1. Smoothness

We show here that the semigroup has a smoothing effect. More specifically, we show that a Hörmander bracket condition is satisfied, so that the transition probability \(P^t(x, dy)\) has a density \(p_t(x, y)\) that is smooth in \(t, x\) and \(y\), and every invariant measure has a smooth density [8].

We identify vector fields over \(\Omega\) and the corresponding first-order differential operators in the usual way (we identify the tangent space of \(\Omega\) with \(\mathbb{R}^6\)). This enables us to consider Lie algebras of vector fields over \(\Omega\) of the kind \(\sum_i (f_i(q, p)\partial q_i + g_i(q, p)\partial p_i)\), where the Lie bracket \([\cdot, \cdot]\) is the usual commutator of two operators.

**Definition 5.2.** We define \(\mathcal{M}\) as the smallest Lie algebra that

(i) contains the constant vector fields \(\partial p_1, \partial p_3\),

(ii) is closed under the operation \([\cdot, A_0]\), where

\[
A_0 = \sum_{i=1}^3 (p_i \partial q_i - u_i \partial p_i) + \sum_b (w_b(\partial p_b - \partial p_3) + \tau_b \partial p_3 - \gamma_b p_b \partial p_b)
\]

is the drift part of \(L\).

By the definition of a Lie algebra, \(\mathcal{M}\) is closed under linear combinations and Lie brackets.

**Lemma 5.3.** Hörmander’s bracket condition is satisfied. More precisely, for all \(x = (q, p)\), the set \(\{v(x) : v \in \mathcal{M}\}\) spans \(\mathbb{R}^6\).

**Proof.** By definition, the constant vector fields \(\partial p_1\) and \(\partial p_3\) belong to \(\mathcal{M}\). Moreover, for \(b = 1, 3\), \([\partial p_b, A_0] = \partial q_b - \gamma_b \partial p_b\). Since \(\mathcal{M}\) is closed under linear combinations and \(\partial p_b \in \mathcal{M}\), it follows that \(\partial q_b \in \mathcal{M}\) for \(b = 1, 3\). Thus it only remains to show that at each \(x \in \Omega\), we can span the directions of \(\partial p_1\) and \(\partial p_3\). In the following, \(f\) denotes a generic function on \(\Omega\) that can be each time different. We have \([\partial q_b, A_0] = w'_b(q_2 - q_b)\partial p_1 + f(q)\partial p_3\) so that commuting \(n - 1\) times with \(\partial q_b\) we get that for all \(n \geq 1\)

\[
u^{(n)}_b(q_2 - q_b)\partial p_1 + f(q)\partial p_3 \in \mathcal{M}. \tag{5.1} \]

Commuting the above with \(A_0\), we find that for all \(n \geq 1\),

\[
u^{(n)}_b(q_2 - q_b)\partial q_b + f(q, p)\partial p_1 + f(q)\partial p_3 + f(q, p)\partial p_b \in \mathcal{M}. \tag{5.2} \]

By assumption 1.2, there is some \(b \in \{1, 3\}\) such that for any fixed \(x \in \Omega\), there is an integer \(n \geq 1\) such that \(\nu^{(n)}_b(q_2 - q_b) \neq 0\). Thus, by (5.1) and (5.2) the proof is complete. \(\square\)

Thus, we have proved proposition 5.1 (i).
5.2. Irreducibility

We show in this section that the process has an irreducible skeleton. We give in fact two different proofs. The first one is given in a general and abstract framework, and works for chains of any lengths. The second one is more explicit, gives more than the irreducibility of a skeleton, but relies strongly on the fact that the chain is made of only three rotors.

5.2.1. Abstract version. Consider the transition probabilities \( \tilde{P}^t(\cdot, \cdot) \) of the system at equilibrium, i.e. with parameters \( \tau_1 = \tau_3 = 0 \) and \( T_1 = T_3 = T \) for some \( T > 0 \). For all \( x \) and \( t \), the measures \( P^t(x, \cdot) \) and \( \tilde{P}^t(x, \cdot) \) are equivalent. This equivalence holds because any change of the parameters \( \tau_1, \tau_3 \) (respectively \( T_1, T_3 \)) can be absorbed by shifting (respectively scaling) the Brownian motions appropriately. Therefore, it is enough to prove the irreducibility claim at equilibrium.

At equilibrium, the Gibbs measure \( \nu \) with density \( \frac{1}{Z} \exp\left(-\frac{H}{T}\right) \) is invariant (with some normalisation constant \( Z \)) as mentioned earlier. Note that we do not assume a priori that \( \nu \) is the unique invariant measure at equilibrium, nor that the system at equilibrium is irreducible. The only two properties that we need are invariance and (everywhere) positiveness of the density of \( \nu \).

**Lemma 5.4.** The equilibrium transition probabilities satisfy the following property: for every measurable set \( S \) one has for all \( t \)

\[
\int_S \tilde{P}^t(x, S^c) d\nu = \int_{S^c} \tilde{P}^t(x, S) d\nu.
\]

**Proof.** We have by the invariance of \( \nu \),

\[
\int_S \tilde{P}^t(x, S) d\nu - \int_S \tilde{P}^t(x, S^c) d\nu = \int_S \tilde{P}^t(x, S) d\nu + \int_S (\tilde{P}^t(x, S) - 1) d\nu
\]

\[= \int_\Omega \tilde{P}^t(x, S) d\nu - \int_S 1 d\nu = \nu(S) - \nu(S) = 0,
\]

which completes the proof. \( \square \)

**Lemma 5.5.** Let \( A \) be a closed set. If \( A \) is invariant under \( \tilde{P}^t \) (i.e. \( \tilde{P}^t(x, A) = 1 \) for all \( x \in A \)), then either \( A = \emptyset \) or \( A = \Omega \).

**Proof.** By lemma 5.4, \( \int_A \tilde{P}^t(x, A) d\nu = \int_A \tilde{P}^t(x, A^c) d\nu = 0 \) since \( \tilde{P}^t(x, A^c) = 0 \) for all \( x \in A \). This implies that \( \tilde{P}^t(x, A) = 0 \) for all \( x \in A^c \), since \( x \mapsto \tilde{P}^t(x, A) \) is continuous on the open set \( A^c \) and \( \nu \) has an everywhere positive density. But then \( \tilde{P}^t(x, A) = 1 \) when \( x \in A \) and 0 when \( x \in A^c \), so that by continuity we have \( \partial A = \emptyset \). Since \( \Omega \) is connected, the conclusion follows. \( \square \)

Note that same does not hold for non-closed sets: for example \( \Omega \) minus any set of zero Lebesgue measure is still an invariant set.

**Lemma 5.6.** The time-1 skeleton \( (X_n)_{n=0,1,2,\ldots} \) is irreducible, and the Lebesgue measure \( m \) is a maximal irreducibility measure.

**Proof.** As discussed above, it is enough to prove the result at equilibrium, i.e. with \( \tilde{P}^1(\cdot, \cdot) \). Let \( B \) be a set such that \( m(B) > 0 \). We need to show that the set \( A = \{ x \in \Omega : \sum_{n=1}^{\infty} \tilde{P}^n(x, B) = 0 \} \) is empty. By the smoothness of \( x \mapsto \tilde{P}^n(x, B) \), it is easy to see
that \( A' = \{ x \in \Omega : 3n > 0, \tilde{P}^n(x, B) > 0 \} \) is open, so that \( A \) is closed. Moreover, for all \( x \in A \) it holds that \( 0 = \sum_{n=1}^{\infty} \tilde{P}^n(x, B) \equiv \sum_{n=1}^{\infty} \tilde{P}^{n+1}(x, B) = \int_{\Omega} \tilde{P}^1(x, dy) \sum_{n=1}^{\infty} \tilde{P}^n(y, B). \)

But since by the definition of \( A \) we have \( \sum_{n=1}^{\infty} \tilde{P}^n(y, B) > 0 \) for all \( y \in A' \), we must have \( \tilde{P}^1(x, A') = 0 \) for all \( x \in A \), so that \( A \) is invariant. But then by lemma 5.5 either \( A = \emptyset \) or \( A = \Omega \). We need to eliminate the second possibility. Since \( m(B) > 0 \) and \( v \) has positive density, we have \( v(B) > 0 \). By the invariance of \( v \), we have \( \int_{\Omega} \tilde{P}^1(x, B) dv = v(B) > 0 \). But then there is some \( x \in \Omega \) such that \( \tilde{P}^1(x, B) > 0 \), so that \( x \in A' \). Therefore \( A \neq \Omega \), and thus \( A = \emptyset \) and the process is irreducible with measure \( m \). That \( m \) is a maximal irreducibility measure follows immediately from the fact that the transition probabilities are absolutely continuous with respect to \( m \). This completes the proof. \( \square \)

Thus, we have proved proposition 5.1(ii), so that the proof of proposition 5.1 is complete.

5.2.2. Direct control version. We give now an alternate proof of proposition 5.1(ii). We establish the irreducibility of our process by using controllability arguments. We aim to establish the controllability of (1.1), where the Brownian motions \( B^1_t \) and \( B^2_t \) are replaced with some deterministic, smooth controls \( f_b : \mathbb{R}^+ \to \mathbb{R} \). By absorbing some terms into the controls \( f_b \), this problem is obviously equivalent to controlling the differential equation

\[
\begin{align*}
\dot{q}_1(t) &= p_1(t), \\
\dot{p}_1(t) &= -\sum_b w_b (q_2(t) - q_b(t)), \\
\dot{p}_b(t) &= f_b(t).
\end{align*}
\]

In [4] the irreducibility of chains oscillators has been studied. The authors have proved that chains of any length are controllable in arbitrarily small times. This is of course not the case in our model: since the force applied to \( p_2 \) is bounded by some constants

\[
K^- = \sum_b \min_{x \in \mathbb{T}} w_b(x), \quad K^+ = \sum_b \max_{x \in \mathbb{T}} w_b(x),
\]

the minimal time we need to bring the system from \( x^i = (q^i, p^i) \) to \( x^f = (q^f, p^f) \) is at best proportional to \( |p^f_2 - p^i_2| \). On the other hand, \( q_1, p_1, q_3, p_3 \) can be put into any position in arbitrarily short time. Observe that due to assumption 1.2 and the fact that \( \langle w_b \rangle = 0 \), we have \( K^- < 0 < K^+ \). We will prove the following proposition (remember that the positions \( q_i \) are defined modulo \( 2\pi \)).

**Proposition 5.7.** The system (5.3) is approximately controllable in the sense that for all \( x^i = (q^i, p^i) \), \( x^f = (q^f, p^f) \) and all \( \varepsilon > 0 \), there is a time \( T^* > 0 \) satisfying \( T^* \leq c_1 + c_2 |p^f_2 - p^i_2| \) for some constants \( c_1 \) and \( c_2 \) such that for all \( T > T^* \) there are some smooth controls \( f_1, f_3 : [0, T] \to \mathbb{R} \) such that the solution of (5.3) with initial condition \( x^i \) satisfies \( \|x(T) - x^f\| < \varepsilon \).

This property implies the irreducibility of the chain, since the classical result of Stroock and Varadhan [14] links the support of the semigroup \( P^t \) and the accessible points for (5.3), and implies in particular that for all \( x^i = (q^i, p^i) \) and \( t > c_1 \) the subspace \( \{ x \in \Omega : |p_2 - p^i_2| \leq (t - c_1)/c_2 \} \) is included in the support of \( P^t(x^i, \cdot) \).

The idea is the following: in the next lemma, we show how the middle rotor can be forced into any configuration by applying some piecewise constant force \( g(t) \) to it, with \( g(t) \in [K^-, K^+] \). Then, we will argue that one can move \( q_1 \) and \( q_3 \) (on which we have good control) in such a way that the force exerted on the middle rotor is almost \( g(t) \).
Lemma 5.8. Consider the system
\begin{align}
\dot{q}_2(t) &= \ddot{p}_2(t), \\
\ddot{p}_2(t) &= g(t) - u_2(\ddot{q}_2(t)),
\end{align}
and fix some initial and terminal conditions \((q'_2, p'_2)\) and \((q''_2, p''_2)\). We claim that there is a \(T^*\) satisfying \(T^* \leq c_1 + c_2|p'_2 - p''_2|\) for some constants \(c_1\) and \(c_2\) such that for all \(T > T^*\) there is a piecewise constant control \(g(t) : \mathbb{R}^+ \to [K^-, K^+]\) (with finitely many constant pieces) such that the solution of (5.4) with initial data \((q'_2, p'_2)\) satisfies \(\ddot{p}_2(T) = p''_2\) and \(\ddot{q}_2(T) = q''_2\).

Proof. We prove this result only in the case \(u_2 \equiv 0\). If \(p'_2 \geq p''_2\), then let \(\Theta = (p'_2 - p''_2)/K^+\) and let \(g(t) = K^+\) for all \(t \in [0, \Theta]\). If \(p'_2 < p''_2\), let \(\Theta = (p'_2 - p''_2)/K^-\) and let \(g(t) = K^-\) for all \(t \in [0, \Theta]\). In both cases, \(\ddot{p}_2(\Theta) = p''_2\), while \(\ddot{q}_2(\Theta)\) might be anything. Let now \(K^* = \min(|K^*|, |K^-|)\), and consider some \(\Delta > 0\) and \(a \in [0, K^*]\). Assume that \(g(t) = a\) when \(t \in [\Theta, \Theta + \Delta]\) and \(g(t) = -a\) when \(t \in [\Theta + \Delta, \Theta + 2\Delta]\). Clearly \(\ddot{p}_2(\Theta + 2\Delta) = \ddot{p}_2(\Theta) = p''_2\) and
\[
\ddot{q}_2(\Theta + 2\Delta) = \ddot{q}_2(\Theta) + 2\Delta p''_2 + a\Delta^2.
\]
Observe that as soon as \(\Delta > \sqrt{2\pi/K^*}\), we can choose \(a \in [0, K^*]\) so that \(\ddot{q}_2(\Theta + 2\Delta)\) takes any value (modulo \(2\pi\)). In particular, we can choose it to be \(q''_2\), so that we have the advertised result with \(T^* = \Theta + \sqrt{2\pi/K^*}\). \(\square\)

Remark 5.9. We have given a proof only if \(u_2 \equiv 0\). However, the result remains true even if \(u_2 \neq 0\), although the proof is much more involved. Typically, if the pinning is stronger than the interaction forces \(w_b\), and the initial condition is such that \(p_2\) is small, we sometimes have to push the middle rotor several times back and forth to increase its energy enough to pass above the ‘potential barrier’ created by \(U_2\). Conversely, we sometimes have to brake the middle rotor with some non-trivial controls.

We now have some piecewise constant control \(g(t)\) that can bring the middle rotor to the final configuration of our choice. It remains to show that we can make the external rotors follow some trajectories that have the appropriate initial and terminal conditions, and such that the force exerted on the middle rotator closely approximates \(g(t)\). We do not prove this in detail, but we list here the main steps.

- Since \(K^- \leq g(t) \leq K^+\), it is possible to find piecewise smooth functions \(q_b^i(t), b = 1, 3\), such that \(\sum_b w_b(\ddot{q}_2(t) - q_b^i(t)) = g(t)\), where \(\ddot{q}_2(t)\) is the solution of (5.4).
- Let \(\delta > 0\) be small. We can find some smooth trajectories \(q_b(t)\) compatible with the boundary conditions \(x^i\) and \(x^f\), such that \(q_b(t) = q_b^i(t)\) for all \(t \in [0, T] \setminus A_3\), where \(A_3\) consists of a finite number of intervals of total length at most \(\delta\). We can choose the controls \(f_b\) so that the \(q_b(t)\) constructed here are solutions to (5.3) (when \(\delta\) is small, \(f_b(t)\) is typically very large for \(t \in A_3\)).
- Since the interaction forces \(w_b\) are bounded, their effect during the times \(t \in A_3\) is negligible when \(\delta\) is small. More precisely, it can be shown that the solution \(q_2(t)\) and \(p_2(t)\) of (5.3) converge uniformly on \([0, T]\) to the solutions \(\ddot{q}_2(t)\) and \(\ddot{p}_2(t)\) of (5.4) when \(\delta \to 0\). Therefore, the system is approximately controllable in the sense of proposition 5.7.
6. Numerical illustrations

In this section we illustrate some properties of the invariant measure in the case where $U_i \equiv 0$ and $W_1 = W_3 = -\cos$.

We use throughout the values $\gamma_1 = \gamma_3 = 1$ and $\tau_1 = 0$. We give examples of how the marginal distributions of $p_1$, $p_2$, $p_3$ depend on the temperatures $T_1$, $T_3$ and the external force $\tau_3$. We apply the numerical algorithm given in [9] with time-increment $h = 0.001$. The resulting graphs are quite independent of $h$. In order to obtain good statistics and smooth curves, the probability densities shown below are sampled over $10^8$ units of time and several hundred bins.

At equilibrium, i.e. when $T_1 = T_3$ and $\tau_3 = 0$ (remember that $\tau_1 = 0$ in this section), the marginal law of each $p_i$ has a density proportional to $\exp(-p_i^2 / 2T)$ for $i = 1, 2, 3$. This is obviously not the case out of equilibrium. Moreover, since we work with a finite number of rotors, we do not expect to see any form of local thermal equilibrium in the bulk of the chain (here the ‘bulk’ consists of only the middle rotor). Clearly, the distribution of $p_2$ can be quite far from Maxwellian (Gaussian).

In figure 2 we show the marginal distributions of $p_1$, $p_2$, $p_3$ for different temperatures and no external force. For each pair of temperatures, we show the distributions both in linear and logarithmic scale. At equilibrium, when $T_1 = T_3 = 10$, all three distributions coincide exactly and are Gaussian. However, when $T_1 \neq T_3$, we see that the distribution of $p_2$ is not Gaussian (clearly, the distribution is not a parabola in logarithmic scale).

We next consider the effect of the external force $\tau_3$ on the marginal distributions of the $p_i$, for $T_1 = 10$ and $T_3 = 15$. As illustrated in figure 3, the distributions of $p_1$ and $p_3$ are close to Gaussians with variance $T_1$ and $T_3$ and mean 0 and $\tau_3$. Note that when $\tau_3 \neq 0$, the distribution of $p_2$ has two maxima: one at 0 and one at $\tau_3$. The explanation for these two maxima can be found by looking at the trajectories $p_i(t)$ as shown in figure 4 (for $\tau_3 = 20$); $p_1$ fluctuates around 0, $p_3$ fluctuates around $\tau_3$, and $p_2$ switches between these two regimes. In the regime where $p_2$ fluctuates around zero, the rotor 2 interacts strongly with 1 and weakly

![Figure 2. Distribution of $p_1$, $p_2$, $p_3$, with no external force and several temperatures. (a) $T_1 = 1$, $T_3 = 10$, (b) $T_1 = 2$, $T_3 = 10$, (c) $T_1 = 10$, $T_3 = 10.$](image-url)
Figure 3. Distribution of $p_1$, $p_2$, $p_3$, with $T_1 = 10$, $T_3 = 15$ for 3 values of $\tau$. (a) $\tau = 0$, (b) $\tau = 10$, (c) $\tau = 20$.

Figure 4. Representation of the evolution of $p_1$, $p_2$, $p_3$ with $T_1 = 10$, $T_3 = 15$, $\tau = 20$.

with 3 (since then the force $w_3$ oscillates with ‘high frequency’ $p_3 - p_2 \sim \tau$). Inversely, in the regime where $p_2$ fluctuates around $\tau$, it interacts strongly with 3 and only weakly with 1. Other simulations (not shown here) show that, as expected, the larger $\tau$, the less frequent the switches between these two regimes. The asymmetry of the two maxima in figure 3 is explained by the inequality $T_1 < T_3$, which makes the fluctuations larger in the second regime, so that the mean sojourn time there is shorter.

Acknowledgments

This work has been partially supported by two ERC Advanced Grants (‘Bridges’ 290843 and ‘MALAdY’ 246953). CP thanks the Paris Dauphine University and the CEREMADE laboratory, where he held a post-doctoral position during most of this project. We thank M Hairer, F Huveneers, A Iacobucci, S Olla, and G Stolz for valuable advice and fruitful discussions.

Note added in proof. Based on the results of this paper, an extension to four rotors has been obtained more recently [16].

References

[1] Carmona P 2007 Existence and uniqueness of an invariant measure for a chain of oscillators in contact with two heat baths Stoch. Process. Appl. 117 1076–92
[2] Douc R, Fort G and Guillin A 2009 Subgeometric rates of convergence of $f$-ergodic strong Markov processes Stoch. Process. Appl. 119 897–923

2420
[3] Eckmann J-P and Hairer M 2000 Non-equilibrium statistical mechanics of strongly anharmonic chains of oscillators Commun. Math. Phys. 212 105–64
[4] Eckmann J-P, Pillet C-A and Rey-Bellet L 1999 Entropy production in nonlinear, thermally driven Hamiltonian systems J. Stat. Phys. 95 305–31
[5] Eckmann J-P, Pillet C-A and Rey-Bellet L 1999 Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures Commun. Math. Phys. 201 657–97
[6] Hairer M 2000 How hot can a heat bath get? Commun. Math. Phys. 292 131–77
[7] Hairer M and Mattingly J C 2009 Slow energy dissipation in anharmonic oscillator chains Commun. Pure Appl. Math. 62 999–1032
[8] Hörmander L 1967 Hypoelliptic second order differential equations Acta Math. 119 147–71
[9] Iacobucci A, Legoll F, Olla S and Stoltz G 2011 Negative thermal conductivity of chains of rotors with mechanical forcing Phys. Rev. E 84 061108
[10] Meyn S and Tweedie R L 2009 Markov Chains, Stochastic Stability 2nd edn (Cambridge: Cambridge University) (with a prologue by P W Glynn)
[11] Meyn S P and Tweedie R L 1993 Stability of Markovian processes. III. Foster–Lyapunov criteria for continuous-time processes Adv. Appl. Probab. 25 518–48
[12] Rey-Bellet L and Thomas L E 2002 Exponential convergence to non-equilibrium stationary states in classical statistical mechanics Commun. Math. Phys. 225 305–29
[13] Sanders J A, Verbiest F and Murdock J 2007 Averaging Methods in Nonlinear Dynamical Systems (Applied Mathematical Sciences vol 59) 2nd edn (New York: Springer)
[14] Stroock D W and Varadhan S R S 1972 On the support of diffusion processes with applications to the strong maximum principle Proc. of the Sixth Berkeley Symp. on Mathematical Statistics and Probability, vol III: Probability Theory (Berkeley, CA: University of California)
[15] Vela P A 2003 Averaging and control of nonlinear systems PhD Thesis California Institute of Technology
[16] Cuneo N and Eckmann J P 2015 Non-equilibrium steady states for chains of four rotors (arXiv:1504.04964)