UNIVERSITÉ DE GENÈVE
Section de Physique
Département de Physique théorique

FACULTÉ DES SCIENCES
Professeur Jean-Pierre ECKMANN
Professeur Yvan VELENIK

Non-equilibrium Steady States for Hamiltonian Chains and Networks

THÈSE

présentée à la Faculté des sciences de l'Université de Genève
pour obtenir le grade de
Docteur ès sciences, mention physique

par

Noé Cuneo
de
Lausanne (VD)

Thèse N° 4916

GENÈVE

2016

UNIVERSITÉ
DE GENÈVE

FACULTÉ DES SCIENCES

Doctorat ès sciences
Mention physique

Thèse de *Monsieur Noé CUNEO*

intitulée :

"Non-equilibrium Steady States for Hamiltonian Chains and Networks"

La Faculté des sciences, sur le préavis de Monsieur J.-P. ECKMANN, professeur honoraire et directeur de thèse (Département de physique théorique), Monsieur Y. VELENIK, professeur ordinaire et codirecteur de thèse (Section de mathématiques), Monsieur M. HAIRER, professeur (Mathematics Department, University of Warwick, Coventry, United Kingdom) et Monsieur A. SHIRIKYAN, professeur (Département de mathématiques, Université de Cergy-Pontoise, France), autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 11 avril 2016

Thèse - 4916 -

Le Doyen

N.B. - La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

Remerciements

Mon doctorat à Genève restera un excellent souvenir pour moi, et ce grâce au concours des nombreuses personnes qui m'ont supervisé, accompagné, conseillé, soutenu, et diverti.

Tout d'abord, mon directeur de thèse Jean-Pierre Eckmann entre définitivement dans les 5 catégories ci-dessus. Il m'a proposé un sujet qui m'a immédiatement captivé, et m'a offert un suivi exceptionnel. Au fil des traditionnels cafés, il a partagé avec le groupe son esprit amusant, bienveillant, intéressé à tout, et toujours capable de saisir l'essentiel. Jean-Pierre m'a permis de développer ma passion pour la physique mathématique, et m'a ouvert beaucoup de portes. Je suis aussi reconnaissant qu'honoré d'avoir pu apprendre de lui.

J'aimerais remercier aussi Yvan Velenik, qui a accepté de codiriger ma thèse, ainsi que les membres du comité, Martin Hairer et Armen Shirikyan, pour le temps accordé, leur intérêt et leurs questions. En particulier, ce travail n'aurait pas été une fraction de ce qu'il est sans l'inspiration de plusieurs articles de Martin Hairer.

J'ai eu beaucoup de plaisir à collaborer avec Christophe Poquet, et je me réjouis de continuer à le faire. J'aimerais remercier aussi Pierre Collet, Charles-Édouard Pfister et Clarence Eugene Wayne pour leurs nombreux conseils, ainsi que Vojkan Jakšić et Armen Shirikyan, qui m'ont invité à donner un mini-cours à Cergy-Pontoise, et avec qui j'aurai le plaisir de travailler à Montréal prochainement. Merci aussi à Andrey Dymov pour nos discussions sur le problème des résonances, et à Alexandre Boritchev d'avoir relu l'introduction de cette thèse.

Je suis redevable à Peter Wittwer pour tous nos échanges, et pour m'avoir permis d'être son assistant dans de nombreux cours. J'aimerais le remercier au nom des étudiants pour la qualité exceptionnelle de son enseignement. Ma reconnaissance va aussi au Département de Physique Théorique de l'Université de Genève, sa direction, ses secrétaires Francine Gennai-Nicole et Cécile Jaggi, et son responsable informatique Andreas Malaspinas. Leur efficacité et leur serviabilité sont sans comparaison.

Ces années auraient été infiniment plus ennuyeuses sans mes collègues, avec qui j'ai partagé bureau, cafés, repas, journées à ski, échanges linguistiques, et de nombreux moments mémorables. Je pense en particulier à Andrea Agazzi, Christoph Boeckle, Alexandre Boritchev, Rodrigo Cofré, David Dasenbrook, Neil Dobbs, Jimmy Dubuisson, Julien Guillod, Patrick Hofer, Tanya Yarmola et Maher Younan. Je remercie de même mes amis d'Yverdon, qui m'ont permis de me changer les idées et de passer du bon temps le week-end et pendant les vacances.

Je dois beaucoup aussi à mes parents et toute ma famille, y compris les parents de ma compagne Julitte, pour leurs encouragements, leur compagnie et les bons moments partagés. Après m'avoir si souvent demandé "à quoi ça sert", j'espère qu'ils auront trouvé un début de réponse lors de ma soutenance! Et merci à mon grand-père Roger pour ses conseils pour ma présentation.

Finalement, et plus que tout, je voudrais remercier Julitte pour tout ce qu'elle a fait pour moi, ainsi que pour son indispensable – et c'est peu dire – présence à mes côtés durant ces quatre ans.

Résumé

Dans cette thèse, nous étudions des chaînes et réseaux conduisant la chaleur, qui sont des prototypes de systèmes hors équilibre. Nous prouvons l'existence et l'unicité de la mesure invariante (appelée *état stationnaire hors équilibre*) pour des systèmes d'oscillateurs et de rotateurs classiques couplés à des thermostats stochastiques ayant (possiblement) des températures différentes. Nous présentons tout d'abord les différents modèles étudiés, leur contexte, leurs relations, ainsi que certaines propriétés démontrées dans la littérature. Nous prouvons ensuite les trois nouveaux résultats résumés ci-dessous.

Réseaux d'oscillateurs. Premièrement, nous étudions un réseau de particules (oscillateurs) connectées par des ressorts non linéaires. Certaines particules sont de plus couplées à des thermostats de type Langevin. La structure du réseau est arbitraire, mais le déplacement de chaque particule est 1D. Pour des interactions polynomiales, nous donnons des conditions suffisantes pour que le critère des "crochets de Hörmander" soit satisfait. Ceci implique l'unicité de l'état stationnaire (s'il existe), ainsi que la contrôlabilité du système associé en théorie du contrôle. Les conditions obtenues sont constructives; elles reposent sur l'inéquivalence des forces (modulo translation) et/ou la topologie du réseau. Elles s'appliquent récursivement: les particules "contrôlées" peuvent être utilisées pour contrôler leurs voisines, et ainsi de suite. Nous appliquons ensuite ce résultat à plusieurs types de réseaux physiques.

Chaînes de trois rotateurs. Deuxièmement, nous considérons une chaîne de trois rotateurs, dont les extrémités sont couplées à des thermostats. Sous une hypothèse de non-dégénérescence des potentiels d'interaction, nous montrons que le processus admet une unique mesure invariante, et obtenons une exponentielle étirée comme vitesse de relaxation. La partie cruciale du problème est d'estimer le taux de décroissance de l'énergie du rotateur central. Comme il n'interagit pas directement avec les thermostats, son énergie ne peut être dissipée qu'à travers les deux rotateurs externes. Mais lorsqu'il tourne très rapidement, le rotateur central se découple à cause des oscillations rapides des forces. En utilisant des méthodes de moyennage inspirées par des travaux de Hairer et Mattingly, nous obtenons une dynamique effective pour le rotateur central. Celle-ci nous permet de construire une fonction de Lyapunov qui, combinée à un argument d'irréductibilité, donne le résultat souhaité.

Chaînes de quatre rotateurs. Finalement, nous généralisons les résultats mentionnés ci-dessus à une chaîne de quatre rotateurs dans une configuration similaire. Plus précisément, nous obtenons à nouveau l'existence et l'unicité de la mesure invariante, et une convergence en exponentielle étirée. Le point central est maintenant d'étudier la dissipation de l'énergie des deux rotateurs centraux. Ceci se fait à nouveau par moyennage. La nouvelle difficulté avec quatre rotateurs est l'apparition de résonances quand les deux rotateurs centraux sont rapides et découplent de leurs voisins. Ces résonances ont un effet physique réel, comme illustré numériquement. Par des méthodes de moyennage plus complexes, qui reposent sur la thermalisation rapide des rotateurs externes, nous montrons que ces résonances n'empêchent pas la dissipation de l'énergie des rotateurs centraux. Ceci nous permet à nouveau de conclure en construisant une fonction de Lyapunov.

Abstract

In this thesis, we consider heat-conducting chains and networks, which are prototypical examples of non-equilibrium systems. We prove the existence and uniqueness of an invariant measure (called *non-equilibrium steady state*) for some classical systems made of interacting oscillators and rotors coupled to stochastic heat baths at (possibly) different temperatures. We start by introducing the models of interest in this work and putting them into context. We explain how the different models are related, and recall some earlier results. Then, we prove three new results, which are summarized below.

Networks of oscillators. First, we consider a network of particles (oscillators) connected by non-linear springs. Some particles are coupled to Langevin heat baths. The structure of the network is arbitrary, but the motion of each particle is 1D. For polynomial interactions, we give sufficient conditions for Hörmander's bracket condition to hold, which implies the uniqueness of the steady state (if it exists), as well as the controllability of the associated system in control theory. These conditions are constructive; they are formulated in terms of inequivalence of the forces (modulo translations) and/or conditions on the topology of the connections. The condition is recursive: we show that when some particles are "controlled", they can in turn be used to control some of their neighbors. We show that our criterion applies to several types of physical lattices.

Chains of three rotors. Secondly, we consider a chain of three rotors whose ends are coupled to heat baths. Under some non-degeneracy condition on the interaction potentials, we prove that the process admits a unique invariant probability measure, and relaxes to it at 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 neighbors due to the rapid oscillations of the forces. Using averaging techniques inspired by works of Hairer and Mattingly, 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.

Chains of four rotors. Thirdly, we generalize the results mentioned above to a chain of four rotors in a similar setup. Namely, we obtain again the existence, uniqueness, and stretched exponential convergence rate to the non-equilibrium steady state. The crucial point is now to estimate how fast the energy of the two central rotors is dissipated. This is again obtained by some averaging techniques. The main new difficulty with four rotors is the appearance of resonances when both central rotors are fast and decouple from their neighbors. These resonances have a physical meaning, as we illustrate numerically. By introducing some more involved averaging techniques, which rely on the rapid thermalization of the two external rotors, we show that the resonances do not play a disturbing role on the rate of energy dissipation. This allows us to construct a Lyapunov function and obtain the desired result.

Contents

1. Introduction

While there exists a well-established theory of equilibrium (and close to equilibrium) statistical mechanics, its non-equilibrium counterpart remains a work in progress. The field is attracting a lot of attention because many daily phenomena take place out of equilibrium, and because it raises very interesting mathematical challenges.

By definition, the "usual" statistical ensembles (microcanonical, canonical, grand canonical) describe systems in equilibrium, and are intrinsically time-reversible. These ensembles can therefore bear no current (of particles, energy, charges, ...). Non-equilibrium statistical mechanics attempts to describe systems which are maintained out of equilibrium by some external action. Typically, the external action can be

- an external (possibly time-dependent) force,
- letting different parts of the system interact with heat baths at different temperatures, or
- inserting particles in one region of the system and letting particles escape somewhere else.

In fact, most systems that we observe in everyday life are out of equilibrium (think for example of an electric conductor, a piece of metal heated at one end, the weather, a star, a cell, or even the human body). Nice introductions to non-equilibrium systems are given in [40, 48]. See also [2, 39].

For a given model of non-equilibrium system (which can be deterministic, stochastic, quantum, ...), the mathematical challenge is to extract statistical properties from the dynamics. In particular, one would like to compute the average of observable quantities, their fluctuations, etc.

One very important and long-studied class of non-equilibrium problems is that of heat-conducting systems. Consider a rod (typically an electrical insulator) in contact at both ends with heat reservoirs at different temperatures, say T_L at the left and T_R at the right, with $T_L > T_R$. According to Fourier's law, we expect to see a heat flux along the rod from left to right, which is proportional to the temperature difference $T_L - T_R$ (at least if the temperature difference is small), and inversely proportional to the length of the rod. A great challenge is to derive this law starting from some microscopical description of the medium (here, the rod). See [7, 42] for more information on this long-standing problem.

A model of a heat-conducting medium is typically a network of interacting sites (which can be thought of as "atoms") coupled to some "heat baths", and must specify the following.

- What each site is made of. We will consider rotors and oscillators undergoing some Hamiltonian dynamics.
- The notion of heat baths. While there are many ways to model heat baths, we will use (and introduce) the simplest for our purpose: Langevin thermostats.
- The interactions and the topology of the network. One typically considers chains with nearest-neighbor interactions, or regular (square, cubic, triangular, hexagonal, ...) lattices.

Different models can lead to very different results, and require quite different methods. Fourier's law has only been proved for a few models (see for example [5, 6, 24, 25] and the reviews [7, 42]). For networks with purely classical Hamiltonian interactions, proving Fourier's law seems currently out of reach. There are in fact questions which arise well before Fourier's law from an analytical point of view, and which are already surprisingly challenging: Is the system stable in a statistical sense? Does the system admit a steady state? Is this steady state unique?

In the current state of research, these "elementary" questions have been answered only in some specific cases. In this thesis, we give positive answers for some classes of models.

1.1. Hamiltonian systems interacting with Langevin thermostats

We consider a classical system of N particles, each of which only moves in one dimension. We denote by q_i, p_i the position and momentum of the particle i for $i \in \{1, \dots, N\}$, and we denote the phase space by Ω. We consider two cases, depending on the domain of q_i:

- **Oscillators**: $\Omega = \mathbb{R}^N \times \mathbb{R}^N$. That is, $q_i \in \mathbb{R}$ and $p_i \in \mathbb{R}$.
- **Rotors**: $\Omega = \mathbb{T}^N \times \mathbb{R}^N$, where $\mathbb{T}^N$ is the N-torus. We use the convention $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. We then have $q_i \in [0, 2\pi)$ and again $p_i \in \mathbb{R}$. We view each particle as a rotating disk (rotor).

We write $q = (q_1, \dots, q_N)$, $p = (p_1, \dots, p_N)$, and introduce the Hamiltonian

$$H(q,p) = \sum_{i=1}^{N} \left(\frac{p_i^2}{2} + U_i(q_i) \right) + \sum_{\substack{i,j=1 \\ i>j}}^{N} W_{i,j}(q_i - q_j) \ , \tag{1.1.1}$$

where the U_i are smooth *pinning potentials*, and the $W_{i,j}$ are smooth *interaction potentials*. Observe that all the masses are chosen to be 1, for simplicity. We view (1.1.1) as a network of interacting oscillators or rotors. The topology of the network depends on which interaction potentials are non-constant. Indeed, there is simply no interaction between two sites i and j ($i > j$) if $W_{i,j}$ is contant. We will always assume that H has compact level sets, *i.e.,* that the (total) potential is confining.

We now single out a set $\mathcal{B} \subset \{1, \dots, N\}$ of particles, which we let interact with some heat baths. For each $i \in \mathcal{B}$, we make the particle i interact with a Langevin thermostat at temperature $T_i > 0$, with coupling (or friction) constant $\gamma_i > 0$. More precisely, we consider the stochastic differential equation (SDE)

$$\begin{aligned}
\mathrm{d}q_i &= p_i \,\mathrm{d}t \ , && i = 1, \dots, N \ , \\
\mathrm{d}p_i &= -\partial_{q_i} H(q,p) \,\mathrm{d}t, && i \notin \mathcal{B}, \\
\mathrm{d}p_i &= -\partial_{q_i} H(q,p) \,\mathrm{d}t - \gamma_i p_i \,\mathrm{d}t + \sqrt{2\gamma_i T_i} \mathrm{d}B_i \ , && i \in \mathcal{B} \ ,
\end{aligned} \tag{1.1.2}$$

where the $B_i, i \in \mathcal{B}$, are independent, normalized Wiener processes[1]. That is, each Langevin heat bath consists of a stochastic term $\sqrt{2\gamma_i T_i}\mathrm{d}B_i$ and a friction term $-\gamma_i p_i \,\mathrm{d}t$, both acting on p_i. Langevin heat baths are a very simplistic choice of thermostats, but they have the advantage that the solutions to (1.1.2) form a Markov process. See for example §3 of [42] or §4 of [7] for reviews of other possible thermostats. We make the following remarks about Langevin heat baths.

- In the trivial case where $N = 1$ and $\mathcal{B} = \{1\}$, we retrieve the Langevin equation.

[1] In §3, we will drive the system further out of equilibrium by applying constant, external forces on some sites

- When all the temperatures are the same, say $T_i = \frac{1}{\beta}$ for all $i \in \mathcal{B}$ and some $\beta > 0$, the Gibbs-Boltzmann distribution $\frac{1}{Z}e^{-\beta H}$ is invariant (more on this below). This justifies the interpretation of T_i as a *temperature*.
- While no "physical" reservoir acts as a Langevin thermostat on the system, it is possible to retrieve such stochastic heat baths in some appropriate *weak coupling limit* [53].
- In [22, 23], the authors consider chains of oscillators interacting with "infinite" Hamiltonian heat reservoirs. These reservoirs take the form of free fields, the initial conditions of which are distributed according some Gibbs measures at (possibly) different temperatures. It is then shown that for specific choices of interaction between the chain and the fields, the latter can be "integrated out" and replaced with a finite number of auxiliary variables in such a way that the resulting process is Markovian. The equation (1.1.2) can then be retrieved by taking some appropriate limit (see the discussion above the equation (10) of [50], and also [28]).

We now introduce the questions that we will ask about (1.1.2). As mentioned above, the solutions to (1.1.2) form a Markov process, and we introduce the transition probabilities

$$P^t(x, A) \equiv \mathbb{P}\left(x_t \in A | x_0 = x\right), \tag{1.1.3}$$

where $t \geq 0$ is the time, x is the initial condition, and A is any Borel subset of Ω. This induces a semigroup $P^t, t \geq 0$ on the space of Borel probability measures on Ω. If ν is a probability measure (think of it as some distribution of initial conditions), we define a time-evolved probability measure $P^t\nu$ given by

$$(P^t\nu)(B) \equiv \int_\Omega P^t(x, B)\mathrm{d}\nu(x)\ ,$$

where B is any Borel subset of Ω. The semigroup makes a probability measure ν "evolve" with time: $P^t\nu$ is the probability distribution at time t, given that the probability distribution at time 0 was ν.

Now the question is whether there exists a probability measure μ which is invariant under the semigroup P^t, namely such that $P^t\mu = \mu$ for all $t \geq 0$. Such a μ is called an *invariant measure*, or a *stationary state*. For non-equilibrium systems, the term *non-equilibrium steady state* is often used.

If there exists such an invariant measure, then one can ask whether it is *unique*, and whether the transition probabilities converge to it, *i.e.,* whether

$$\lim_{t\to\infty} P^t(x, \,\cdot\,) = \mu \tag{1.1.4}$$

for every initial condition $x \in \Omega$. We will be interested in the norms with respect to which we have (1.1.4), and in the speed of convergence, which we call *relaxation rate*.

One can similarly consider a semigroup $T^t, t \geq 0$ acting on observables. For any measurable function $f : \Omega \to \mathbb{R}$ (with some conditions on its growth, see Chapter 3 and Chapter 4), we define

$$T^t f(x) = \int_\Omega f(y)P^t(x, \mathrm{d}y) = \mathbb{E}_x f(x_t)\ ,$$

where $\mathbb{E}_x$ is the expectation with respect to the process started at $x \in \Omega$.

We have the duality relation

$$\int_\Omega T^t f \mathrm{d}\nu = \int_\Omega f \mathrm{d}(P^t \nu) ,$$

provided that f is of sufficiently slow growth.

Now, if there is a unique invariant measure μ, the question dual to (1.1.4) is to determine a class of functions f for which

$$\lim_{t\to\infty} T^t f(x) = \int_\Omega f \mathrm{d}\mu , \tag{1.1.5}$$

for all $x \in \Omega$, and to determine the speed of convergence.

A central tool is the generator L of the semigroup T^t, which is the second-order differential operator

$$L = \sum_{i=1}^{N} (p_i \partial_{q_i} - (\partial_{q_i} H)\partial_{p_i}) + \sum_{i\in\mathcal{B}} \left(-\gamma_i p_i \partial_{p_i} + \gamma_i T_i \partial_{p_i}^2\right) . \tag{1.1.6}$$

We have formally $L = \frac{\mathrm{d}}{\mathrm{d}t} T^t|_{t=0}$. The first immediate observation is that

$$LH(q,p) = \sum_{i\in\mathcal{B}} (\gamma_i T_i - \gamma_i p_i^2) . \tag{1.1.7}$$

This reflects the fact that only the heat baths can change the energy of the system, since the Hamiltonian dynamics preserves it. Since the coefficients of the SDE (1.1.2) are locally Lipschitz, H is bounded below with compact level sets, and the right-hand side of (1.1.7) is bounded above, we know that the solutions to (1.1.2) are unique, and almost surely defined for all times and continuous (see for example [10, Theorem 3.5]). This also ensures that the transition kernel (1.1.3) is well-defined and enjoys the Feller property, namely that if f is continuous and bounded, then so is $T^t f$.

The generator of P^t is the formal adjoint of L, namely

$$L^* = \sum_{i=1}^{N} (-p_i \partial_{q_i} + (\partial_{q_i} H)\partial_{p_i}) + \sum_{i\in\mathcal{B}} \left(\gamma_i (1 + p_i \partial_{p_i}) + \gamma_i T_i \partial_{p_i}^2\right) ,$$

where we have used that $-\partial_{q_i} p_i + \partial_{p_i}(\partial_{q_i} H) = 0$ (which is in fact a manifestation of Liouville's theorem). At thermal equilibrium, namely when all the temperatures are equal, say $T_i = \frac{1}{\beta}$ for some $\beta > 0$ and all $i \in \mathcal{B}$, it is easy to check that

$$L^* e^{-\beta H(q,p)} = 0 \qquad (T_i = 1/\beta, i \in \mathcal{B}) .$$

This, together with the existence of solutions for all times discussed above, implies that in the equilibrium case, the (unnormalized) Gibbs-Boltzmann measure

$$\mathrm{d}\mu_\beta \equiv e^{-\beta H(q,p)} \mathrm{d}p \mathrm{d}q$$

is invariant. Moreover, if the potential is sufficiently confining (for example if it grows at least polynomially in all directions), this measure can be normalized to a probability measure. Thus, at thermal equilibrium, we know that there exists a steady state. Even in this case, however, we have at this point no information about whether it is unique, and whether (1.1.4) holds.

1.1.1. Earlier results

A class of systems which has been widely studied consists of chains of oscillators (see Figure 1.1). That is to say, we consider a Hamiltonian

$$H = \sum_{i=1}^{N} \Big(\frac{p_i^2}{2} + U_i(q_i)\Big) + \sum_{i=1}^{N-1} W_{i+1,i}(q_{i+1} - q_i) \ . \tag{1.1.8}$$

Such chains of oscillators have been studied numerically for many different potentials (and also many kinds of heat baths, including the Langevin thermostats that we consider here). See for example [1,8,31,41].

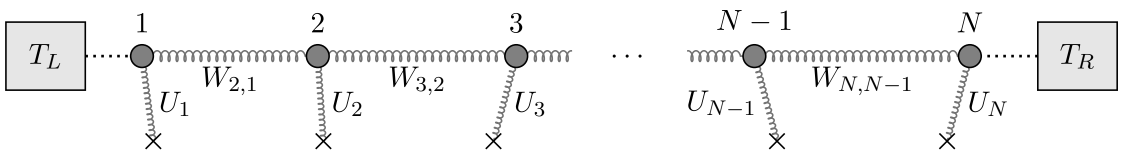


Figure 1.1 – A chain of oscillators.

From a rigorous point of view, however, only limited results are available. The existence of a steady state has only been obtained in some cases where the potentials are asymptotically polynomial. For the present discussion, we assume that

$$U_i(q_i) \sim |q_i|^\ell, \qquad \text{and} \qquad W_{i+1,i}(q_{i+1} - q_i) \sim |q_{i+1} - q_i|^k$$

for some exponents $k, \ell \geq 2$ (for the precise meaning of the relation $\sim$, see the references below). Depending on whether $k \geq \ell$ or $\ell > k$, the properties of the system are very different. We separate the two cases as follows.

- **Type-1 oscillators**: $k \geq \ell$. In this case, the interactions play an important role at high energy.
- **Type-2 oscillators**: $\ell > k$. In this case, the interactions vanish at high energy, which complicates the situation significantly (see below).

For type-1 oscillators, and under some non-degeneracy assumption on the interaction potentials, the existence of an invariant measure for chains of arbitrary length N has been proved, and the relaxation rate in (1.1.4) is *exponential*. This was first proved using functional analytic methods in [23], and then generalized in [19, 20, 22] (note that the heat baths are slightly different from ours). Later, a probabilistic proof (using discrete-time Lyapunov functions) was given in [50], and an adaptation to Langevin thermostats was provided in [9]. The basic idea is to understand the high-energy limit of the dynamics, by means of a scaling argument. The key is that because of the strong interactions, it never happens that the two external oscillators (sites 1 and N) remain at rest for a long time, and this ensures that the friction terms in (1.1.2) dissipate enough energy in any finite time interval. Further properties of chains of type-1 oscillators (including some asymptotic properties and entropy production fluctuations) have been studied in [22, 49, 51].

For type-2 oscillators, the situation is more complicated, and the existence of an invariant measure has been proved only for chains of length 3. The main difficulty is that when one of the oscillators has a lot of energy, it "sees" only its pinning potential, since it grows faster than the interaction potentials. The highly energetic site then oscillates very rapidly, which causes it to decouple from its neighbors.

Indeed, if one takes a particle of mass 1 in the 1D potential $U(x) = |x|^\ell$, then the period of oscillation depends on the energy E as follows (the turning points are $\pm E^{1/\ell}$):

$$\tau_E = 2\int_{-E^{1/\ell}}^{E^{1/\ell}} \frac{\mathrm{d}x}{\sqrt{2(E-|x|^\ell)}} \propto E^{\frac{1}{\ell}-\frac{1}{2}} \,. \tag{1.1.9}$$

For type-2 oscillators, we have $\ell > 2$ (since $k \geq 2$ and $\ell > k$), and therefore the frequency of the oscillations increases with the energy E. Now, assume that the site i has a very large energy E_i, and that all other sites have a small energy. The particle at site i then essentially oscillates at the natural frequency of the pinning potential, which grows like $E_i^{1/2-1/\ell}$. The interaction forces $W'_{i+1,i}(q_{i+1}-q_i)$ and $-W'_{i,i-1}(q_i - q_{i-1})$ then oscillate at this (high) frequency, and with average zero. They have therefore a vanishing effect. In other words, sites with a very large energy tend to decouple from their neighbors due to the very fast oscillations of the forces.

For a chain of length 3 with $k = 2$ (linear interactions) and $\ell \geq 3$, the existence of a steady state has been proved in [34], and it has been shown that the convergence in (1.1.4) is *not* exponential when ℓ is large enough. For longer chains of type-2 oscillators, the existence of an invariant measure remains an open problem. The results of [34] are obtained by some averaging/homogenization technique, which we will introduce in detail in Chapter 3, when we adapt it to the case of rotors.

We now argue that chains of rotors (see Figure 1.2) are very similar to chains of type-2 oscillators. Indeed, a crucial ingredient in [34] is the scaling of the period of the central oscillator with respect to its energy, which is given by (1.1.9). The larger the ℓ, the faster the oscillations at high energy, and the more the central oscillator decouples. If we take formally the limit $\ell \to \infty$ in (1.1.9), we obtain $\tau_E \sim E^{-1/2}$, *i.e.,* the frequency grows like $\sqrt{E}$. Also, when formally $\ell = \infty$, the pinning potential becomes an infinite well, so the position space is "compactified".

From this point of view, rotors can really be seen as a "worst-case" limit of type-2 oscillators. Indeed, for rotors the position space is compact and the frequency scales like $\sqrt{E}$ (the potentials are bounded, so at high energy the momentum is essentially proportional to $\sqrt{E}$). While the similarity between the two kinds of systems is purely formal at this point, it will become obvious in Chapter 3 when we deal with a chain of three rotors by adapting the methods of [34]. The advantage of rotors over oscillators is that all forces are bounded (since the position space $\mathbb{T}^N$ is compact), which is technically convenient.

Of course, in addition to the relation with chains of type-2 oscillators, chains of rotors have

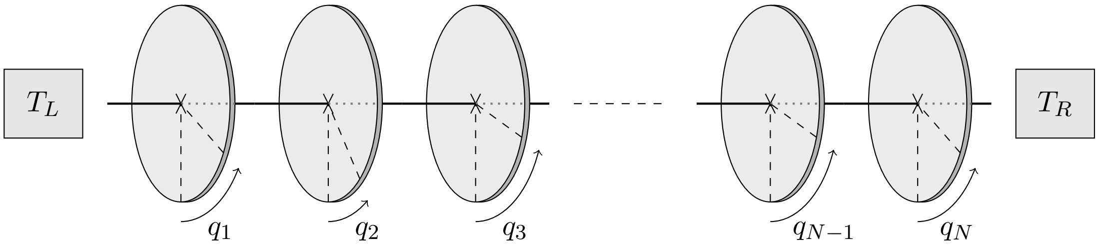


Figure 1.2 – A chain of rotors.

their intrinsic interest and many of their (sometimes very surprising) properties have been studied numerically and perturbatively [7, 15, 30, 37, 42].

1.2. Organization of the thesis

The remainder of this thesis is a reproduction of three papers.

In Chapter 2, which reproduces [11], we study the controllability/uniqueness problem for some networks of oscillators with polynomial interaction potentials. We give a sufficient condition for Hörmander's bracket condition to hold, which we formulate in terms of equivalence of the potentials. We also give some sufficient conditions that depend on the topology of the network alone, and which cover many physical lattices.

In Chapter 3, which was published in [13], we prove the existence of a steady state for a chain of 3 rotors, and show that the relaxation speed is a stretched exponential $e^{-c\sqrt{t}}$. We prove this result by adapting the method developed in [34] for 3 oscillators of type 2.

Finally, in Chapter 4, which reproduces [12], we prove a similar result for a chain of 4 rotors. The main new difficulty is the existence of resonances between the two central rotors. We argue that these resonances actually have a physical meaning. The main new method developed in Chapter 4 is some form of *stochastic* averaging (in addition to the averaging over the rapid oscillations). We show that the resonant terms can be approximated (up to a well-controlled error) by some stochastic averages, and that as a result, the resonances have no undesirable effect.

2. Networks of oscillators

We reproduce here the contents of [11]. The only modifications compared to the published version are a few supplementary references.

Controlling general polynomial networks

with Jean-Pierre Eckmann

Communications in Mathematical Physics **328** (2014), 1255–1274

Abstract

We consider networks of massive particles connected by non-linear springs. Some particles interact with heat baths at different temperatures, which are modeled as stochastic driving forces. The structure of the network is arbitrary, but the motion of each particle is 1D. For polynomial interactions, we give sufficient conditions for Hörmander's "bracket condition" to hold, which implies the uniqueness of the steady state (if it exists), as well as the controllability of the associated system in control theory. These conditions are constructive; they are formulated in terms of inequivalence of the forces (modulo translations) and/or conditions on the topology of the connections. We illustrate our results with examples, including "conducting chains" of variable cross-section. This then extends the results for a simple chain obtained in [23] .

2.1. Introduction

We consider a network of interacting particles described by an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with a set $\mathcal{V}$ of vertices and a set $\mathcal{E}$ of edges. Each vertex represents a particle, and each edge represents a spring connecting two particles. We single out a set $\mathcal{V}_* \subset \mathcal{V}$ of particles, each of which interacts with a heat bath. We address the question of when such a system has a unique stationary state. This question has been studied for several special cases: Starting from a linear chain [22, 23], results have become more refined in terms of the relation between the spring potentials and the pinning potentials which tie the masses to the laboratory frame [19, 49]. This problem is very delicate, as is apparent from the extensive study in [33] for the case of only 2 masses.

We provide conditions on the interaction potentials that imply Hörmander's "bracket condition," from which it follows that the semigroup associated to the process has a smoothing effect. This, together with some stability assumptions, implies the *uniqueness* of the stationary state. The *existence* is not discussed in this paper, but seems well understood in the case where the interaction potentials are stronger than the pinning potentials. This issue will be explained in a forthcoming paper [21].

Since the problem is known (see for example [32] and [3]) to be tightly related to the control problem where the stochastic driving forces are replaced with deterministic control forces, we shall use the terminology of control theory, and mention the implications of our results from the control-theoretic viewpoint.

We work with unit masses and interaction potentials that are polynomials of degree at least 3, and we say that two such potentials V_1 and V_2 have *equivalent second derivative* if there is a $\delta \in \mathbb{R}$ such that $V_1''(\,\cdot\,) = V_2''(\,\cdot\, + \delta)$.

We start with the set $\mathcal{V}_*$ of particles that interact with heat baths, and are therefore *controllable*. One of our results (Corollary 2.5.6) is formulated as a condition for some of the particles in the set of first neighbors $\mathcal{N}(\mathcal{V}_*)$ of $\mathcal{V}_*$ to be also controllable. Basically, the condition is that these particles must be "inequivalent" in a sense that involves both the topology of their connections to $\mathcal{V}_*$ and the corresponding interaction potentials. More precisely, a sufficient condition for a particle $v \in \mathcal{N}(\mathcal{V}_*)$ to be controllable is that for each other particle $w \in \mathcal{N}(\mathcal{V}_*)$ at least one of the two conditions holds:

(a) v and w are connected to $\mathcal{V}_*$ in a topologically different way,

(b) there is a particle c in $\mathcal{V}_*$ such that the interaction potential between c and v and that between c and w have inequivalent second derivative.

It is then possible to use this condition recursively, taking control of more and more masses at each step (Theorem 2.5.7). If by doing so we can control all the masses in the graph, then Hörmander's bracket condition holds.

In §2.6 we give examples of physically relevant networks whose controllability can be established using this method.

Our results imply in particular that connected graphs are controllable for "almost all" choices of the interaction potentials, provided that they are polynomials of degree at least 3 (Corollary 2.6.3).

2.2. The system

We define a Hamiltonian for the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ as follows. Each particle $v \in \mathcal{V}$ has position $q_v \in \mathbb{R}$ and momentum $p_v \in \mathbb{R}$ and is "pinned down" by a potential $U_v(q_v)$. Throughout, we assume the masses being 1, for simplicity of notation. See Remark 2.4.20 on how to adapt the results when the masses are not all equal.

We denote each edge $e \in \mathcal{E}$ by $\{u, v\}$ (or equivalently $\{v, u\}$) where u, v are the vertices adjacent to e.[1] To each edge $e = \{u, v\}$, we associate an interaction potential $V_{uv}(q_u - q_v)$, or equivalently $V_{vu}(q_v - q_u)$ with

$$V_{vu}(q_v - q_u) \equiv V_{uv}(q_u - q_v)\ . \tag{2.2.1}$$

Note that we do not require the potentials to be even functions; the condition (2.2.1) just makes sure that considering $e = \{u, v\}$ or $e = \{v, u\}$ leads to the same physical interaction, which is consistent with the fact that the edges are not oriented.

With the notation $q = (q_v)_{v\in\mathcal{V}}$ and $p = (p_v)_{v\in\mathcal{V}}$ the Hamiltonian is then

$$H(q,p) = \sum_{v\in\mathcal{V}} \bigl(p_v^2/2 + U_v(q_v)\bigr) + \sum_{e\in\mathcal{E}} V_e(\delta q_e)\ ,$$

where it is understood that $V_e(\delta q_e)$ denotes one of the two expressions in (2.2.1).

Finally, we make the following assumptions:

[1] Due to the physical nature of the problem, we assume that the graph has no self-edge.

Assumption 2.2.1.

1. *All the functions that we consider are smooth.*
2. *The level sets of H are compact, i.e., for each $K > 0$ the set $\{(q,p) \mid H(q,p) \leq K\}$ is compact.*
3. *The function $\exp(-\beta H)$ is integrable for some $\beta > 0$.*

Each particle $v \in \mathcal{V}_*$ is coupled to a heat bath at temperature $T_v > 0$ with coupling constant $\gamma_v > 0$. For convenience, we set $\gamma_v = 0$ when $v \notin \mathcal{V}_*$. The model is then described by the system of stochastic differential equations

$$\begin{aligned} dq_v &= p_v \, dt \, , \\ dp_v &= -U_v'(q_v)dt - \partial_{q_v}\Big(\sum_{e\in\mathcal{E}} V_e(\delta q_e)\Big)dt - \gamma_v p_v dt + \sqrt{2T_v\gamma_v}\, dW_v(t) \, , \end{aligned} \tag{2.2.2}$$

where the W_v are identical independent Wiener processes. The solutions to (2.2.2) form a Markov process. The generator of the associated semigroup is given by

$$L \equiv X_0 + \sum_{v\in\mathcal{V}_*} \gamma_v T_v \partial_{p_v}^2 \, ,$$

with

$$X_0 \equiv -\sum_{v\in\mathcal{V}_*}\gamma_v p_v \partial_{p_v} + \sum_{v\in\mathcal{V}}\big(p_v\partial_{q_v} - U_v'(q_v)\partial_{p_v}\big) - \sum_{\{u,v\}\in\mathcal{E}} V_{uv}'(q_u - q_v)\cdot(\partial_{p_u} - \partial_{p_v}) \, .$$

From now on, we assume that the interaction potentials V_e, $e \in \mathcal{E}$ are *polynomials of degree at least 3*. The condition on the degree means that we require throughout the presence of non-harmonicities. The fully-harmonic case has been described earlier [26], and the case where some but not all the potentials are harmonic is not covered here. We will show in a counter-example (Example 2.7.3) that the non-harmonicities are really essential for our results. We make no assumption about the pinning potentials U_v; we do not require them to be polynomials, and some of them may be identically zero.

We work in the space $\mathbb{R}^{2|\mathcal{V}|}$ with coordinates $x = (q,p)$. We identify the vector fields over $\mathbb{R}^{2|\mathcal{V}|}$ and the corresponding first-order differential operators in the usual way. This enables us to consider Lie algebras of vector fields over $\mathbb{R}^{2|\mathcal{V}|}$, where the Lie bracket $[\,\cdot\,,\,\cdot\,]$ is the usual commutator of two operators.

Definition 2.2.2. *We define $\mathcal{M}$ as the smallest Lie algebra that*

(i) contains ∂_{p_v} for all $v \in \mathcal{V}_*$,[2]

(ii) is closed under the operation $[\,\cdot\,, X_0]$,

(iii) is closed under multiplication by smooth scalar functions.

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

[2] Due to the identification mentioned above, we view here ∂_{p_v} as a constant vector field over $\mathbb{R}^{2|\mathcal{V}|}$.

Definition 2.2.3. *We say that a particle $v \in \mathcal{V}$ is* controllable *if we have $\partial_{q_v}, \partial_{p_v} \in \mathcal{M}$. We say that the network $\mathcal{G}$ is* controllable *if all the particles are controllable, i.e., if*

$$\partial_{q_v}, \partial_{p_v} \in \mathcal{M} \qquad \textit{for all } v \in \mathcal{V} \ . \tag{2.2.3}$$

The aim of this paper is to give sufficient conditions on $\mathcal{G}$ and the interaction potentials, which guarantee that the network is controllable.

If the network is controllable in the sense (2.2.3), then Hörmander's condition[3] [36] holds: for all x, the vector fields $F \in \mathcal{M}$ evaluated at x span all of $\mathbb{R}^{2|\mathcal{V}|}$, *i.e.,*

$$\{F(x) \mid F \in \mathcal{M}\} = \mathbb{R}^{2|\mathcal{V}|} \quad \text{ for all } x \in \mathbb{R}^{2|\mathcal{V}|} \ . \tag{2.2.4}$$

Hörmander's condition implies that the transition probabilities of the Markov process (2.2.2) are smooth, and that so is any invariant measure (see for example [3, Cor. 7.2]). We now briefly mention two implications of these smoothness properties. Proposition 2.2.4 and Proposition 2.2.5 below can be deduced from arguments similar to those exposed in [32]. The argument of §3.5.2 can also easily be adapted to the present case.

Proposition 2.2.4. *Under Assumption 2.2.1, if (2.2.4) holds, then the Markov process (2.2.2) has at most one invariant probability measure.*

The control-theoretic problem corresponding to (2.2.2) is the system of ordinary differential equations

$$\begin{aligned} \dot{q}_v &= p_v, \\ \dot{p}_v &= -U_v'(q_v) - \partial_{q_v}\Big(\sum_{e\in\mathcal{E}} V_e(\delta q_e)\Big) + (u_v(t) - \gamma_v p_v) \cdot \mathbf{1}_{v\in\mathcal{V}_*} \ , \end{aligned} \tag{2.2.5}$$

where for each $v \in \mathcal{V}_*$, $u_v : \mathbb{R} \to \mathbb{R}$ is a smooth *control function* (*i.e.,* the stochastic driving forces have been replaced with deterministic functions).[4]

Proposition 2.2.5. *Under the hypotheses of Proposition 2.2.4, the system (2.2.5) is* controllable *in the sense that for each $x^{(0)} = (q^{(0)}, p^{(0)})$ and $x^{(f)} = (q^{(f)}, p^{(f)})$, there are a time T and some smooth controls u_v, $v \in \mathcal{V}_*$, such that the solution $x(t)$ of (2.2.5) with $x(0) = x^{(0)}$ verifies $x(T) = x^{(f)}$.*

In fact, (2.2.4) is a well-known condition in control theory. See for example [38], which addresses the case of piecewise constant control functions. In particular, (2.2.4) implies by [38, Thm. 3.3] that for every initial condition $x^{(0)}$ and each time $T > 0$, the set $A(x^{(0)}, T)$ of all points that are accessible at time T (by choosing appropriate controls) is connected and full-dimensional.

[3]The condition (2.2.4) is slightly different, but equivalent to the usual statement of Hörmander's criterion. This can be checked easily. In particular, closing $\mathcal{M}$ under multiplication by smooth scalar functions does not alter the set in (2.2.4), and will be very convenient.

[4]Whether or not we keep the dissipative terms $-\gamma_v p_v$ in (2.2.5) makes no difference since they can always be absorbed in the control functions.

2.3. Strategy

We want to show that $\partial_{q_v}, \partial_{p_v} \in \mathcal{M}$ for all $v \in \mathcal{V}$. The next lemma shows that we only need to worry about the ∂_{p_v}.

Lemma 2.3.1. *Let A be a subset of $\mathcal{V}$.*

$$\text{If} \quad \sum_{v\in A} \partial_{p_v} \in \mathcal{M} \quad \text{then} \quad \sum_{v\in A} \partial_{q_v} \in \mathcal{M} \,.$$

Proof. Assuming $\sum_{v\in A} \partial_{p_v} \in \mathcal{M}$, we find that

$$\Big[\sum_{v\in A} \partial_{p_v}, X_0\Big] = \sum_{v\in A} \partial_{q_v} - \sum_{v\in\mathcal{V}_*\cap A} \gamma_v \partial_{p_v} \tag{2.3.1}$$

is in $\mathcal{M}$. But since $\partial_{p_v} \in \mathcal{M}$ for all $v \in \mathcal{V}_*$, the linear structure of $\mathcal{M}$ implies $\sum_{v\in\mathcal{V}_*\cap A} \gamma_v \partial_{p_v} \in \mathcal{M}$. Adding this to the vector field in (2.3.1) shows that $\sum_{v\in A} \partial_{q_v} \in \mathcal{M}$, as claimed. □

Definition 2.3.2. *We say that a set $A \subset \mathcal{V}$ is* jointly controllable *if $\sum_{v\in A} \partial_{p_v}$ is in $\mathcal{M}$ (and therefore, also $\sum_{v\in A} \partial_{q_v}$ by Lemma 2.3.1).*

Requiring all the particles in a set A to be (individually) controllable is stronger than asking the set A to be jointly controllable (indeed, if all the ∂_{p_v}, $v \in A$ are in $\mathcal{M}$, then so is their sum). We will obtain jointly controllable sets and then "refine" them until we control particles individually.

The strategy is as follows. In the next section, we start with a controllable particle c, and show that its first neighbors split into jointly controllable sets. Then, in §2.5, we consider several controllable particles, and basically intersect the jointly controllable sets obtained for each of them, in order to control "new" particles individually. Finally, we iterate this procedure, taking control of more particles at each step, until we establish (under some conditions) the controllability of the whole network.

Remark 2.3.3. Observe in the following that our results neither involve the pinning potentials U_v nor the coupling constants γ_v.

2.4. The neighbors of one controllable particle

We consider in this section a particle $c \in \mathcal{V}$, and denote by $\mathcal{T}^c$ the set of its first neighbors (the "targets"). The following notion of equivalence among polynomials enables us to split $\mathcal{T}^c$ into equivalence classes.

Definition 2.4.1. *Two polynomials f and g are called* equivalent *if there is a $\delta \in \mathbb{R}$ such that $f(\,\cdot\,) = g(\,\cdot\, + \delta)$.*

Definition 2.4.2. *We say that two particles $v, u \in \mathcal{T}^c$ are equivalent (with respect to c) if the two polynomials V''_{cv} and V''_{cu} are equivalent.*

Since this relation is symmetric and transitive, the set $\mathcal{T}^c$ is naturally decomposed into a disjoint union of equivalence classes:

$$\mathcal{T}^c = \cup_i \mathcal{T}_i^c \ .$$

An explanation of why we use the second derivative of the potentials instead of the first one (*i.e.,* the force) will be given in Example 2.7.2. The main result of this section is

Theorem 2.4.3. *Assume that c is controllable. Then, each equivalence class $\mathcal{T}_i^c$ is jointly controllable, i.e.,*

$$\sum_{v\in\mathcal{T}_i^c} \partial_{p_v} \in \mathcal{M} \qquad \textit{for all } i \ . \tag{2.4.1}$$

Furthermore, there are constants δ_{cv} such that for all i,

$$\sum_{v\in\mathcal{T}_i^c} (q_c - q_v + \delta_{cv})\partial_{p_v} \in \mathcal{M} \ . \tag{2.4.2}$$

The second part of the theorem will be used in the next section to intersect the equivalence classes $\mathcal{T}_i^c$ of several controllable particles c. We will now prepare the proof of Theorem 2.4.3. We assume in the remainder of this section that c is controllable. And since c is fixed, we write $\mathcal{T}$ and $\mathcal{T}_i$ instead of $\mathcal{T}^c$ and $\mathcal{T}_i^c$.

Lemma 2.4.4. *We have*

$$\sum_{v\in\mathcal{T}} V_{cv}''(q_c - q_v)\partial_{p_v} \in \mathcal{M} \ . \tag{2.4.3}$$

Proof. From Lemma 2.3.1 we conclude that $\partial_{q_c} \in \mathcal{M}$. Therefore, we find that

$$[\partial_{q_c}, X_0] = -U_c''(q_c)\partial_{p_c} - \sum_{v\in\mathcal{T}} V_{cv}''(q_c - q_v)\cdot(\partial_{p_c} - \partial_{p_v})$$

is in $\mathcal{M}$. Now, since $\partial_{p_c} \in \mathcal{M}$ and since $\mathcal{M}$ is closed under multiplication by scalar functions, we can subtract all the contributions that are along ∂_{p_c} and obtain (2.4.3). □

We need a bit of technology to deal with equivalent polynomials.

Definition 2.4.5. *Let $g(t) = \sum_{i=0}^{k} a_i t^i/i!$ be a polynomial of degree $k \geq 1$. If $a_{k-1} = 0$, we say that g is* adjusted. *As can be checked, the polynomial $\tilde{g}(\,\cdot\,) \equiv g(\,\cdot\, - a_{k-1}/a_k)$ is always adjusted, and is referred to as the* adjusted representation *of g.*

Observe that a polynomial and its adjusted representation are by construction equivalent and have the same degree and the same leading coefficient. In fact, given a polynomial g of degree $k \geq 1$, $\tilde{g}$ is the only polynomial equivalent to g that is adjusted. This adjusted representation will prove to be very useful thanks to the following obvious

Lemma 2.4.6. *Two polynomials g and h of degree at least 1 are equivalent iff $\tilde{g} = \tilde{h}$.*

Remark 2.4.7. If all the interaction potentials are *even*, then all the V''_{cv} are automatically adjusted, and some parts of the following discussion can be simplified.

We shift the argument of each V''_{cv} by a constant δ_{cv} so that they are all adjusted. We let $\tilde{f}_v$ be the adjusted representation of V''_{cv} and use the notation

$$x_v \equiv q_c - q_v + \delta_{cv}$$

so that

$$\tilde{f}_v(x_v) = V''_{cv}(q_c - q_v) \quad \text{for all} \quad q_c, q_v \in \mathbb{R}\ .$$

With this notation, (2.4.3) reads as

$$\sum_{v\in\mathcal{T}} \tilde{f}_v(x_v)\partial_{p_v} \in \mathcal{M}\ . \tag{2.4.4}$$

We will now mostly deal with "diagonal" vector fields, *i.e.,* vector fields of the kind (2.4.4), where the component along ∂_{p_v} depends only on x_v. When taking commutators, it is crucial to remember that x_v is only a notation for $q_c - q_v + \delta_{cv}$.

Remark 2.4.8. By the definition of equivalence and Lemma 2.4.6, two particles $v, w \in \mathcal{T}$ are equivalent iff $\tilde{f}_v$ and $\tilde{f}_w$ coincide.

Lemma 2.4.9. *Consider some functions g_v, $v \in \mathcal{T}$.*

$$\textit{If} \quad \sum_{v\in\mathcal{T}} g_v(x_v)\partial_{p_v} \in \mathcal{M} \qquad \textit{then} \qquad \sum_{v\in\mathcal{T}} g'_v(x_v)\partial_{p_v} \in \mathcal{M}\ . \tag{2.4.5}$$

Proof. This is immediate by commuting with ∂_{q_c} (which is in $\mathcal{M}$ by Lemma 2.3.1). □

We now introduce the main tool.

Definition 2.4.10. *Given two vector fields Y and Z, we define the* double commutator $[\![Y : Z]\!]$ *by*

$$[\![Y : Z]\!] \equiv [[X_0, Y], Z]\ .$$

Obviously, if the vector fields Y and Z are in $\mathcal{M}$, then so is $[\![Y : Z]\!]$.

Lemma 2.4.11. *Consider some functions g_v, h_v, $v \in \mathcal{T}$. Then*

$$[\![\sum_{v\in\mathcal{T}} g_v(x_v)\partial_{p_v} : \sum_{v'\in\mathcal{T}} h_{v'}(x_{v'})\partial_{p_{v'}}]\!] = \sum_{v\in\mathcal{T}} (g_v h_v)'(x_v)\,\partial_{p_v}\ . \tag{2.4.6}$$

Proof. Observe first that (omitting the arguments x_v)

$$[\,X_0, \sum_{v\in\mathcal{T}} g_v\partial_{p_v}\,] = \sum_{v\in\mathcal{T}} (p_c - p_v)g'_v\partial_{p_v} - \sum_{v\in\mathcal{T}} g_v\partial_{q_v} + \sum_{v\in\mathcal{T}\cap\mathcal{V}_*} \gamma_v g_v\partial_{p_v}\ .$$

Commuting with $\sum_{v'\in\mathcal{T}} h_{v'}(x_{v'})\partial_{p_{v'}}$ gives the desired result. □

We will prove Theorem 2.4.3 starting from (2.4.4) and using only (2.4.5) and double commutators of the kind (2.4.6).

Let d_v be the degree of $\tilde{f}_v$. Note that since the interaction potentials are of degree at least 3, we have $d_v \geq 1$. We define

$$d \equiv \max_{v \in \mathcal{T}} d_v \geq 1$$

as the maximal degree of the adjusted polynomials $\tilde{f}_v$ with $v \in \mathcal{T}$. We can then write

$$\tilde{f}_v(x) = \sum_{j=0}^{d} b_{vj}\, x^j/j! \ ,$$

for some real coefficients b_{vj}, $j = 0, \ldots, d$, with

$$b_{vj} = 0 \quad \text{if} \quad j > d_v \quad \text{and} \quad b_{v,d_v-1} = 0 \ ,$$

for all $v \in \mathcal{T}$.

Definition 2.4.12. *We define the set of particles $v \in \mathcal{T}$ corresponding to the maximal degree d:*

$$\mathcal{D}^{\,d} \equiv \{v \in \mathcal{T} \mid d_v = d\} \ .$$

For every ℓ, $0 \leq \ell \leq d$, we define the set

$$\mathcal{B}^{d}_{\ell} \equiv \{b_{v\ell} \mid v \in \mathcal{D}^{\,d}\}$$

of distinct *values taken by the coefficients of $x_v^{\ell}/\ell!$ in $\tilde{f}_v$, $v \in \mathcal{D}^{\,d}$.*

We begin with a technical lemma. Observe how it is expressed in terms of the x_v. In a sense, this shows that the x_v are really the “natural” variables for this problem. Thus, in addition to making the notion of equivalence trivial (Remark 2.4.8), working with adjusted representations will be very convenient from a technical point of view.

Lemma 2.4.13. *The following hold:*

(i) For each $b \in \mathcal{B}^{d}_{d}$, we have

$$\sum_{v \in \mathcal{D}^{\,d}\,:\,b_{vd}=b} x_v \partial_{p_v} \in \mathcal{M} \quad \textit{and} \quad \sum_{v \in \mathcal{D}^{\,d}\,:\,b_{vd}=b} \partial_{p_v} \in \mathcal{M} \ . \tag{2.4.7}$$

(ii) Furthermore,

$$\sum_{v \in \mathcal{D}^{\,d}} x_v \partial_{p_v} \in \mathcal{M} \quad \textit{and} \quad \sum_{v \in \mathcal{D}^{\,d}} \partial_{p_v} \in \mathcal{M} \ . \tag{2.4.8}$$

(iii) Let α_v, β_v, $v \in \mathcal{D}^{\,d}$ be real constants. If $d \geq 2$, we have the two implications

$$\textit{if} \ \sum_{v \in \mathcal{D}^{\,d}} \alpha_v\, \partial_{p_v} \in \mathcal{M} \ , \quad \textit{then} \ \sum_{v \in \mathcal{D}^{\,d}} \alpha_v\, x_v\, \partial_{p_v} \ \in \mathcal{M} \ , \tag{2.4.9}$$

$$\textit{if} \sum_{v\in\mathcal{D}^{\,d}} \alpha_v \,\partial_{p_v}, \quad \sum_{v\in\mathcal{D}^{\,d}} \beta_v \,\partial_{p_v} \in \mathcal{M}\,, \quad \textit{then} \quad \sum_{v\in\mathcal{D}^{\,d}} \alpha_v \,\beta_v \,\partial_{p_v} \ \in \mathcal{M}\,. \tag{2.4.10}$$

Remark 2.4.14. Observe that the assumption $d \geq 1$ is crucial in the proof of (i). Requiring the $\tilde{f}_v$ to be non-constant ensures that we can find non-trivial double commutators, which is the crux of our analysis. See Example 2.7.3 for what goes wrong for harmonic potentials.

Proof. (i). By (2.4.4) and using (2.4.5) recursively $d-1$ times, we find that

$$Y \equiv \sum_{v\in\mathcal{T}} \big(\partial^{\,d-1}\tilde{f}_v\big)(x_v)\,\partial_{p_v} = \sum_{v\in\mathcal{T}} (b_{v,d-1} + b_{vd}x_v)\,\partial_{p_v}$$

is in $\mathcal{M}$. But now, by (2.4.6),

$$[\![Y : Y/2]\!] = \sum_{v\in\mathcal{T}} b_{vd}(b_{v,d-1} + b_{vd}x_v)\partial_{p_v} \in \mathcal{M}\,.$$

Taking more double commutators with $Y/2$, we obtain for all $r \geq 1$:

$$\sum_{v\in\mathcal{T}} b_{vd}^{\,r}\,(b_{v,d-1} + b_{vd}x_v)\partial_{p_v} \in \mathcal{M}\,.$$

But the sum above is really only over $\mathcal{D}^{\,d}$ since $b_{vd} \neq 0$ only if $v \in \mathcal{D}^{\,d}$. Moreover, for these v, we have $b_{v,d-1} = 0$ since the polynomials are adjusted, so that for all $i \geq 2$,

$$\sum_{v\in\mathcal{D}^{\,d}} b_{vd}^{i} x_v \partial_{p_v} \in \mathcal{M}\,. \tag{2.4.11}$$

Let $b \in \mathcal{B}_d^{\,d}$. Using Lemma 2.9.1 (in the appendix below) with $s = 1$ and with the set of distinct and non-zero values $\{b_{vd} \mid v \in \mathcal{D}^{\,d}\} = \mathcal{B}_d^{\,d}$ we find real numbers $r_1, r_2, \ldots, r_n$ (with $n = \big|\mathcal{B}_d^{\,d}\big|$) such that $\sum_{i=1}^{n} r_i \, b_{vd}^{i+1}$ equals 1 if $b_{vd} = b$ and 0 when $b_{vd} \neq b$. Thus,

$$\sum_{i=1}^{n} r_i \sum_{v\in\mathcal{D}^{\,d}} b_{vd}^{i+1} x_v \partial_{p_v} = \sum_{v\in\mathcal{D}^{\,d}\,:\,b_{vd}=b} x_v \partial_{p_v}$$

is in $\mathcal{M}$ by (2.4.11). This together with (2.4.5) establishes the second inclusion of (2.4.7), so that we have shown (i).

The statement (ii) follows by summing (i) over all $b \in \mathcal{B}_d^{\,d}$.

Proof of (iii). Let us assume that $d \geq 2$ and that $\sum_{v\in\mathcal{D}^{\,d}} \alpha_v \,\partial_{p_v} \in \mathcal{M}$. By (2.4.7), we have for each $b \in \mathcal{B}_d^{\,d}$ that

$$\frac{1}{b} \sum_{v\in\mathcal{D}^{\,d}\,:\,b_{vd}=b} x_v \partial_{p_v} = \sum_{v\in\mathcal{D}^{\,d}\,:\,b_{vd}=b} \frac{x_v}{b_{vd}} \partial_{p_v} \in \mathcal{M}\,.$$

Taking the double commutator with $\sum_{v\in\mathcal{D}^d}\alpha_v\,\partial_{p_v}$ and summing over all $b\in\mathcal{B}_d^d$ shows that

$$U\equiv\sum_{v\in\mathcal{D}^d}\frac{\alpha_v}{b_{vd}}\partial_{p_v}\in\mathcal{M}\ .$$

Since we assume here $d\geq 2$, we have $Z\equiv\sum_{v\in\mathcal{T}}\bigl(\partial^{d-2}\tilde f_v\bigr)(x_v)\partial_{p_v}\in\mathcal{M}$. But then,

$$[\![U:Z]\!]=\sum_{v\in\mathcal{T}}\frac{\alpha_v}{b_{vd}}\bigl(\partial^{d-1}\tilde f_v\bigr)(x_v)\partial_{p_v}=\sum_{v\in\mathcal{D}^d}\frac{\alpha_v}{b_{vd}}\left(b_{v,d-1}+b_{vd}\,x_v\right)\partial_{p_v}$$

is also in $\mathcal{M}$. Recalling that $b_{v,d-1}=0$ for all $v\in\mathcal{D}^d$, we obtain (2.4.9). Finally (2.4.10) follows from (2.4.9) and the double commutator

$$[\![\sum_{v\in\mathcal{D}^d}\alpha_v\,x_v\,\partial_{p_v}:\sum_{v\in\mathcal{D}^d}\beta_v\,\partial_{p_v}]\!]=\sum_{v\in\mathcal{D}^d}\alpha_v\,\beta_v\,\partial_{p_v}\ .$$

This completes the proof. □

With these preparations, we can now prove Theorem 2.4.3.

Proof of Theorem 2.4.3. We distinguish the cases $d=1$ and $d\geq 2$.
Case $d=1$: This case is easy. Since all the $\tilde f_v$ have degree 1, we have that $\tilde f_v(x_v)=b_{v1}\,x_v$ for all $v\in\mathcal{T}$, with $b_{v1}\neq 0$. Consequently, the sets $\mathcal{T}_i$ consist of those v which have the same b_{v1} (see Remark 2.4.8). Thus, we have by (2.4.7) that for each $\mathcal{T}_i$:

$$\sum_{v\in\mathcal{T}_i}\partial_{p_v}\in\mathcal{M}\quad\text{and}\quad\sum_{v\in\mathcal{T}_i}x_v\partial_{p_v}\in\mathcal{M}\qquad(\text{if }d=1)\ .\tag{2.4.12}$$

This shows that the conclusion of Theorem 2.4.3 holds when $d=1$.
Case $d\geq 2$: In this case, (2.4.7) is not enough. First, (2.4.7) says nothing about the masses $v\in\mathcal{T}\backslash\mathcal{D}^d$, for which $b_{vd}=0$. And second, (2.4.7) provides us with no way to "split" the ∂_{p_v} corresponding to a common (non-zero) value b of b_{vd}, even though the corresponding v might be inequivalent due to some b_{vk} with $k<d$. To fully make use of these coefficients, we must develop some more advanced machinery.

Definition 2.4.15. *We denote by $\mathcal{P}_d$ the vector space of real polynomials in one variable of degree at most d. We consider the operator* $\mathrm{G}:\mathcal{P}_d\to\mathcal{P}_d$ *defined by*

$$(\mathrm{G}v)(x)\equiv(x\cdot v(x))'\ ,$$

and we introduce the set of operators

$$\mathcal{F}\equiv\mathrm{span}\{\mathrm{G},\mathrm{G}^2,\ldots,\mathrm{G}^{d+1}\}\ .$$

Observe that by (2.4.4) and (2.4.8) we have

$$[\![\sum_{v\in\mathcal{T}} \tilde f_v(x_v)\partial_{p_v} : \sum_{v\in\mathcal{D}^d} x_v\partial_{p_v}]\!] = \sum_{v\in\mathcal{D}^d} (\mathrm{G}\tilde f_v)(x_v)\partial_{p_v} \in \mathcal{M}\ .$$

Note that we obtain a sum over $\mathcal{D}^d$ only. By taking more double commutators with $\sum_{v\in\mathcal{D}^d} x_v\partial_{p_v}$, we find that $\sum_{v\in\mathcal{D}^d}(\mathrm{G}^k\tilde f_v)(x_v)\partial_{p_v}$ is in $\mathcal{M}$ for all $k\ge 1$. Thus, by the linear structure of $\mathcal{M}$, we obtain

Lemma 2.4.16. *For all* $\mathrm{P}\in\mathcal{F}$*, we have*

$$\sum_{v\in\mathcal{D}^d} (\mathrm{P}\tilde f_v)(x_v)\partial_{p_v} \in \mathcal{M}\ .$$

It is crucial to understand that it is the *same* operator P that is applied simultaneously to all the components, and that the components in $\mathcal{T}\setminus\mathcal{D}^d$ are “projected out”.

We now show that some very useful operators are in $\mathcal{F}$.

Proposition 2.4.17. *The following hold:*

(i) The projector

$$\mathrm{S}_\ell : \mathcal{P}_d \to \mathcal{P}_d\ , \quad \sum_{i=0}^d b_i x^i/i! \mapsto b_\ell x^\ell/\ell!$$

belongs to $\mathcal{F}$ *for all* $\ell = 0,\dots,d$.

(ii) The identity operator $\mathbf{1}$ *acting on* $\mathcal{P}_d$ *is in* $\mathcal{F}$.

Proof. Consider the basis $B=(e_0,e_1,\dots,e_d)$ of $\mathcal{P}_d$ where $e_j(x)=x^j/j!$. Observe that for all $j\ge 0$ we have $\mathrm{G}e_j=(j+1)e_j$, so that G is diagonal in the basis B. Thus, G^k is represented by the matrix $\mathrm{diag}(1^k,2^k,\dots,(d+1)^k)$ for all $k\ge 1$. Consequently, for each $\ell\in\{0,\dots,d\}$, there is by Lemma 2.9.1 with $s=0$ a linear combination of $\mathrm{G},\mathrm{G}^2,\dots,\mathrm{G}^{d+1}$ that is equal to S_ℓ. This proves (i). Moreover, we have that $\sum_{\ell=0}^d \mathrm{S}_\ell=\mathbf{1}$, so that the proof of (ii) is complete. $\square$

Lemma 2.4.18. *For all* $\ell=0,\dots,d$*, and for each* $b\in\mathcal{B}^d_\ell=\{b_{v\ell}\mid v\in\mathcal{D}^d\}$ *we have*

$$\sum_{v\in\mathcal{D}^d\,:\,b_{v\ell}=b} \partial_{p_v}\in\mathcal{M}\ . \tag{2.4.13}$$

Proof. Let $\ell\in\{0,1,\dots,d\}$. Using Lemma 2.4.16 and Proposition 2.4.17(i) we find that the vector field $\sum_{v\in\mathcal{D}^d}(b_{v\ell}x^\ell/\ell!)\partial_{p_v}$ is in $\mathcal{M}$. Using (2.4.5) repeatedly, we find that $\sum_{v\in\mathcal{D}^d} b_{v\ell}\partial_{p_v}$ is in $\mathcal{M}$. Thus, by (2.4.10),

$$\sum_{v\in\mathcal{D}^d} b^i_{v\ell}\partial_{p_v}\in\mathcal{M} \qquad \text{for all } i\ge 1.$$

Then, applying Lemma 2.9.1 to the set $\mathcal{B}_\ell^d \setminus \{0\}$ and with $s = 0$, we conclude that

$$\sum_{v\in\mathcal{D}^d\,:\,b_{v\ell}=b} \partial_{p_v} \in \mathcal{M} \quad \text{for all} \quad b \in \mathcal{B}_\ell^d \setminus \{0\}\ . \tag{2.4.14}$$

If $0 \notin \mathcal{B}_\ell^d$, we are done. Else, we obtain that (2.4.13) holds also for $b = 0$ by summing the vector field (2.4.14) over all $b \in \mathcal{B}_\ell^d \setminus \{0\}$ and subtracting the result from $\sum_{v\in\mathcal{D}^d} \partial_{p_v}$ (which is in $\mathcal{M}$ by (2.4.8)). This completes the proof. □

Remember that by Remark 2.4.8, a given equivalence class $\mathcal{T}_i$ is either a subset of $\mathcal{D}^d$ or completely disjoint from it.

Lemma 2.4.19. *Let $\mathcal{T}_i$ be an equivalence class such that $\mathcal{T}_i \subset \mathcal{D}^d$. Then*

$$\sum_{v\in\mathcal{T}_i} \partial_{p_v} \in \mathcal{M}\ , \quad \textit{and} \quad \sum_{v\in\mathcal{T}_i} x_v\, \partial_{p_v} \in \mathcal{M}\ . \tag{2.4.15}$$

Proof. All the polynomials $\tilde f_v$, $v \in \mathcal{T}_i$ are equal. Thus, there are coefficients $c_\ell \in \mathcal{B}_\ell^d$, $\ell = 0, 1, \ldots, d$ such that

$$\mathcal{T}_i = \bigcap_{\ell=0}^{d} \{v \in \mathcal{D}^d \mid b_{v\ell} = c_\ell\}\ . \tag{2.4.16}$$

By Lemma 2.4.18, we have for all $\ell = 0, \ldots, d$ that

$$\sum_{v\in\mathcal{D}^d\,:\,b_{v\ell}=c_\ell} \partial_{p_v} \in \mathcal{M}\ . \tag{2.4.17}$$

Now observe that whenever two sets $B, B' \subset \mathcal{D}^d$ are such that $\sum_{v\in B} \partial_{p_v} \in \mathcal{M}$ and $\sum_{v\in B'} \partial_{p_v} \in \mathcal{M}$, we have by (2.4.10) that $\sum_{v\in B\cap B'} \partial_{p_v} \in \mathcal{M}$. Applying this recursively to the intersection in (2.4.16) and using (2.4.17) shows that $\sum_{v\in\mathcal{T}_i} \partial_{p_v} \in \mathcal{M}$. Using now (2.4.9) implies that $\sum_{v\in\mathcal{T}_i} x_v \partial_{p_v} \in \mathcal{M}$, which completes the proof. □

With these tools, we are now ready to complete the proof of Theorem 2.4.3 (for the case $d \geq 2$). By Lemma 2.4.19, we are done if $\mathcal{D}^d = \mathcal{T}$ (*i.e.,* if all the $\tilde f_v$, $v \in \mathcal{T}$ have degree d). If this is not the case, we proceed as follows.

Observe that Lemma 2.4.16 and Proposition 2.4.17(ii) imply that $\sum_{v\in\mathcal{D}^d} \tilde f_v(x_v)\, \partial_{p_v}$ is in $\mathcal{M}$. Subtracting this from (2.4.4) shows that

$$\sum_{v\in\mathcal{T}\setminus\mathcal{D}^d} \tilde f_v(x_v)\partial_{p_v} \in \mathcal{M}\ .$$

Thus, we can start the above procedure again with this new "smaller" vector field, each component being a polynomial of degree at most

$$d' \equiv \max_{v\in\mathcal{T}\setminus\mathcal{D}^d} d_v\ ,$$

with obviously $d' < d$. Defining then $\mathcal{D}^{d'} = \{v \in \mathcal{T} \mid d_v = d'\}$, we get as in Lemma 2.4.19 that (2.4.15) holds for all $\mathcal{T}_i \subset \mathcal{D}^{d'}$. We then proceed like this inductively, dealing at each step with the components of highest degree and "removing" them, until all the remaining components have the same degree d^- (which is equal to $\min_{v\in\mathcal{T}} d_v$). If $d^- \geq 2$ we obtain again as in Lemma 2.4.19 that (2.4.15) holds for all $\mathcal{T}_i \subset \mathcal{D}^{d^-}$. And if $d^- = 1$, the conclusion follows from (2.4.12). Thus, (2.4.15) holds for every equivalence class $\mathcal{T}_i$, regardless of the degree of the polynomials involved. The proof of Theorem 2.4.3 is complete. □

Remark 2.4.20. Our method also covers the case where each particle $v \in \mathcal{V}$ can have an arbitrary positive mass m_v. The proofs work the same way, if we replace the functions $\tilde{f}_v$ with $\tilde{f}_v = V''_{cv}(x_v)/(m_c m_v)$. Thus, if for example all the V''_{cv}, $v \in \mathcal{T}$ are the same, but all the particles in $\mathcal{T}$ have distinct masses, then all the new $\tilde{f}_v$ are different, and the particles in $\mathcal{T}$ belong each to a separate $\mathcal{T}_i$.

2.5. Controlling a network

We now show how Theorem 2.4.3 can be used recursively to control a large class of networks. The idea is very simple: at each step of the recursion, we apply Theorem 2.4.3 to a controllable particle (or a set of such) in order to show that some neighboring vertices are also controllable. Starting this procedure with the particles in $\mathcal{V}_*$ (which are controllable by the definition of $\mathcal{M}$), we obtain under certain conditions that the whole network is controllable.

In order to make the distinction clear, we will say that a particle c is a *controller* if it is controllable and if we intend to use it as a starting point to control other particles.

Definition 2.5.1. *Let $\mathcal{J}$ be the collection of jointly controllable sets (i.e., of sets $A \subset \mathcal{V}$ such that $\sum_{v\in A} \partial_{p_v} \in \mathcal{M}$, and therefore also $\sum_{v\in A} \partial_{q_v}$ by Lemma 2.3.1).*

Obviously, a particle v is controllable iff $\{v\} \in \mathcal{J}$. The next lemma shows what we "gain" in $\mathcal{J}$ when we apply Theorem 2.4.3 to a controller c. Remember that the set $\mathcal{T}^c$ of first neighbors of c is partitioned into equivalence classes $\mathcal{T}_i^c$, as discussed in §2.4.

Lemma 2.5.2. *Let $c \in \mathcal{V}$ be a controller. Then,*

(i) for all i,

$$\mathcal{T}_i^c \in \mathcal{J} \ ,$$

(ii) for all i and for all $A \in \mathcal{J}$ the sets

$$A \cap \mathcal{T}_i^c, \qquad A \setminus \mathcal{T}_i^c \quad \text{and} \quad \mathcal{T}_i^c \setminus A \tag{2.5.1}$$

are in $\mathcal{J}$.

We illustrate some possibilities in Figure 2.1.

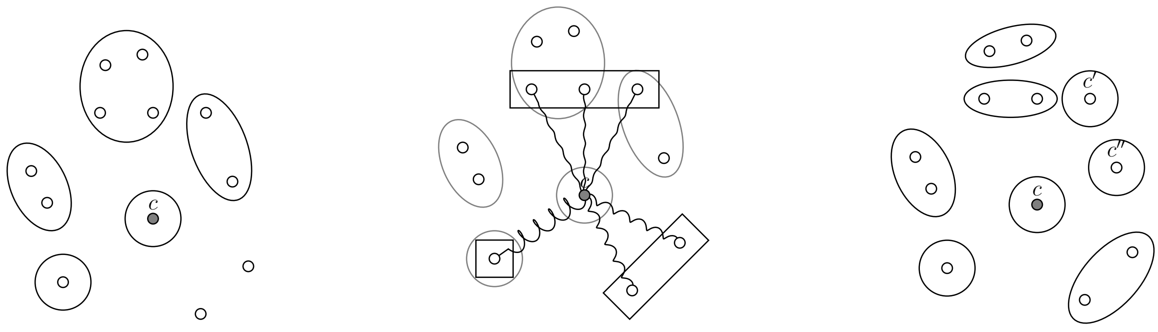


Figure 2.1 – a: A controller c and a few sets in $\mathcal{J}$, shown as ovals. b: The equivalence classes $\mathcal{T}_i^c$ are shown as rectangles. (Only the edges incident to c are shown.) c: New sets "appear" in $\mathcal{J}$. In particular, c' and c'' are controllable.

Proof. (i). This is an immediate consequence of (2.4.1) and the definition of $\mathcal{J}$.
(ii). We consider an equivalence class $\mathcal{T}_i^c$ and a set $A \in \mathcal{J}$. By (2.4.2) we find that

$$[\sum_{v\in\mathcal{T}_i^c}(q_c - q_v + \delta_{cv})\partial_{p_v}, \sum_{v\in A}\partial_{q_v}] = \sum_{v\in A\cap\mathcal{T}_i^c}\partial_{p_v} - \mathbf{1}_{c\in A}\cdot\sum_{v\in\mathcal{T}_i^c}\partial_{p_v}$$

is in $\mathcal{M}$. By the linear structure of $\mathcal{M}$ and since $\sum_{v\in\mathcal{T}_i^c}\partial_{p_v}$ is in $\mathcal{M}$ by (2.4.1), we can discard the second term and we find $\sum_{v\in A\cap\mathcal{T}_i^c}\partial_{p_v} \in \mathcal{M}$. This proves that $A\cap\mathcal{T}_i^c$ is in $\mathcal{J}$. Then, subtracting $\sum_{v\in A\cap\mathcal{T}_i^c}\partial_{p_v}$ from $\sum_{v\in A}\partial_{p_v}$ (resp. from $\sum_{v\in\mathcal{T}_i^c}\partial_{p_v}$) shows that $\sum_{v\in A\setminus\mathcal{T}_i^c}\partial_{p_v}$ (resp. $\sum_{v\in\mathcal{T}_i^c\setminus A}\partial_{p_v}$) is in $\mathcal{M}$, which completes the proof of (ii). □

We can now give an algorithm that applies Lemma 2.5.2 recursively, and that can be used to show that a large variety of networks is controllable.

Proposition 2.5.3. *Consider the following algorithm that builds step by step a collection of sets $\mathcal{W}\subset\mathcal{J}$ and a list of controllable particles K.*

Start with $\mathcal{W} = \{\{v\} \mid v\in\mathcal{V}_*\}$ and put (in any order) the vertices of $\mathcal{V}_*$ in K.

1. *Take the first unused controller $c\in K$.*
2. *Add each equivalence class $\mathcal{T}_i^c$ to $\mathcal{W}$.*
3. *For each $\mathcal{T}_i^c$ and each $A\in\mathcal{W}$ add the sets of (2.5.1) to $\mathcal{W}$.*
4. *If in 2. or 3. new singletons appear in $\mathcal{W}$, add the corresponding vertices (in any order) at the end of K.*
5. *Consider c as used. If there is an unused controller in K, start again at 1. Else, stop.*

We have the following result: if in the end K contains all the vertices of $\mathcal{V}$, then the network is controllable.

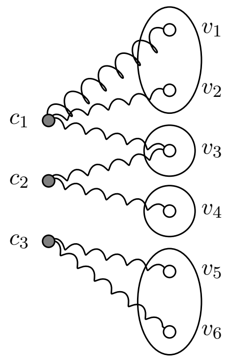
Proof. By Lemma 2.5.2, the collection $\mathcal{W}$ remains at each step a subset of $\mathcal{J}$, and K contains only controllers. Thus, the result holds by construction. □

The algorithm stops after at most $|\mathcal{V}|$ iterations, and one can show that the result does not depend on the order in which we use the controllers. This algorithm is probably the easiest to implement, but does not give much insight into what really makes a network controllable with our criteria. For this reason, we now formulate a similar result in terms of equivalence with respect to a *set* of controllers, which underlines the role of the "cooperation" of several controllers.

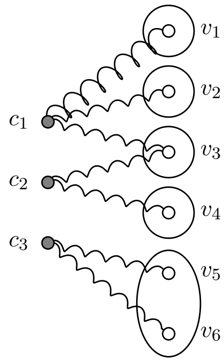
Definition 2.5.4. *We consider a set C of controllers and denote by $\mathcal{N}(C)$ the set of first neighbors of C that are not themselves in C. We say that two particles $v, w \in \mathcal{N}(C)$ are C-siblings if v and w are connected to C in exactly the same way, i.e., if for every $c \in C$ the edges $\{c, v\}$ and $\{c, w\}$ are either both present or both absent.*

Moreover, we say that v and w are C-equivalent if they are C-siblings, and if in addition, for each $c \in C$ that is linked to v and w, we have that v and w are equivalent with respect to c (i.e., there is a $\delta \in \mathbb{R}$ such that $V''_{cv}(\,\cdot\,) = V''_{cw}(\,\cdot\, + \delta)$).

The C-equivalence classes form a refinement of the sets of C-siblings, see Figure 2.2.



(a) C-siblings



(b) C-equivalence classes

Figure 2.2 – Illustration of Definition 2.5.4. We assume that all the springs are identical, except for the edge $\{c_1, v_1\}$. The particles $v_1, \dots, v_6$ form 4 sets of C-siblings, with $C = \{c_1, c_2, c_3\}$. The one containing v_1 and v_2 is further split into two C-equivalence classes, since v_1 and v_2 are by assumption inequivalent with respect to c_1.

Proposition 2.5.5. *Let C be a set of controllers. Then, for each C-equivalence class $U \subset \mathcal{N}(C)$, we have $U \in \mathcal{J}$.*

Proof. See Figure 2.3. Let $U = \{v_1, \dots, v_n\} \subset \mathcal{N}(C)$ be a C-equivalence class. We denote by $c_1, \dots, c_k$ the controllers in C that are linked to v_1, and therefore also to $v_2, \dots, v_n$, since the elements of U are C-siblings. For each $j \in \{1, \dots, k\}$, there is a $\mathcal{T}_i^{c_j}$ with $v_1, \dots, v_n \in \mathcal{T}_i^{c_j}$, and we define

$\mathcal{S}_j \equiv \mathcal{T}_i^{c_j} \setminus C$. We consider the set

$$\widehat{U} \equiv \bigcap_{j=1}^{k} \mathcal{S}_j \ .$$

Clearly, $U \subset \widehat{U}$, and $\widehat{U} \in \mathcal{J}$ by Lemma 2.5.2. We have $\widehat{U} = \{v_1, \dots, v_n, v_1^*, \dots, v_r^*\}$, where the v_j^* are those particles that are equivalent to $v_1, \dots, v_n$ from the point of view of $c_1, \dots, c_k$, but that are also connected to some controller(s) in $C \setminus \{c_1, \dots, c_k\}$. In particular, for each $j \in \{1, \dots, r\}$, there is a $c_j^* \in C \setminus \{c_1, \dots, c_k\}$ and an i such that v_j^* is in $\mathcal{S}_j^* \equiv \mathcal{T}_i^{c_j^*}$. By construction, $\mathcal{S}_j^* \cap U = \emptyset$. Thus,

$$\widehat{U} \setminus \bigcup_{j=1}^{r} \mathcal{S}_j^* = U \ .$$

Starting from $\widehat{U} \in \mathcal{J}$ and removing one by one the $\mathcal{S}_j^*$, we find by Lemma 2.5.2(ii) that U is indeed in $\mathcal{J}$, as we claim. □

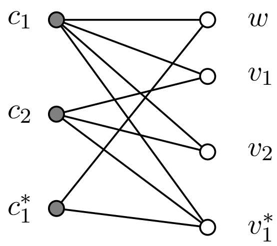


Figure 2.3 – Illustration of the proof of Proposition 2.5.5 for identical springs. We consider the C-equivalence class $U = \{v_1, v_2\}$, where C contains c_1, c_2, c_1^* and possibly other particles (not shown) that are not linked to v_1, v_2. With the notation of the proof, we have $\mathcal{S}_1 = \{w, v_1, v_2, v_1^*\}$ and $\mathcal{S}_2 = \{v_1, v_2, v_1^*\}$ so that $\widehat{U} = \mathcal{S}_1 \cap \mathcal{S}_2 = \{v_1, v_2, v_1^*\}$. Since v_1^* belongs to $\mathcal{S}_1^* = \{w, v_1^*\}$, we find $\widehat{U} \setminus \mathcal{S}_1^* = U$.

An immediate consequence is

Corollary 2.5.6. *Let C be a set of controllers. If a vertex $v \in \mathcal{N}(C)$ is alone in its C-equivalence class, then it is controllable.*

Applying this recursively, we obtain

Theorem 2.5.7. *We start with $C_0 \equiv \mathcal{V}_*$. For each $k \geq 0$, let*

$$C_{k+1} \equiv C_k \cup \{v \in \mathcal{N}(C_k) \mid \text{ no other vertex in } \mathcal{N}(C_k) \text{ is } C_k\text{-equivalent to } v\} \ .$$

Then, if $C_k = \mathcal{V}$ for some $k \geq 0$, the network is controllable.

Proof. By Corollary 2.5.6 we have that each C_k contains only controllers (remember also that $\mathcal{V}_*$ contains only controllers by the definition of $\mathcal{M}$). Thus if $C_k = \mathcal{V}$ for some $k \geq 0$ we find that all vertices are controllers, which is what we claim. □

2.6. Examples

In this section we illustrate by several examples the range of our controllability criteria.

Example 2.6.1. *A single controller c can control* several *particles if the interaction potentials between c and its neighbors have pairwise inequivalent second derivative. See Figure 2.4.*

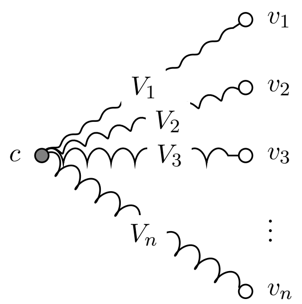


Figure 2.4 – If no two springs are equivalent, the v_i are controllable. Springs from the v_i to other particles or from one v_i to another may exist but are not shown. They do not change the conclusion.

The example above does not use the topology of the network (*i.e.,* the notion of siblings), but only the inequivalence due to the second derivative of the potentials. We have the following immediate generalization, which we formulate as

Theorem 2.6.2. *Assume that $\mathcal{G}$ is connected, that $\mathcal{V}_*$ is not empty, and that for each $v \in \mathcal{V}$, the first neighbors of v are all pairwise inequivalent with respect to v (i.e., no two distinct neighbors u, w of v are such that $V''_{vu}(\,\cdot\,) = V''_{vw}(\,\cdot\, + \delta)$ for some constant $\delta \in \mathbb{R}$). Then, the network is controllable.*

Proof. We use Theorem 2.5.7. Observe that under these assumptions, we have at each step $C_{k+1} = C_k \cup \mathcal{N}(C_k)$. Thus, since the network is connected, there is indeed a $k \geq 0$ such that $C_k = \mathcal{V}$. $\square$

One can restate Theorem 2.6.2 as a genericity condition:

Corollary 2.6.3. *Assume that $\mathcal{G}$ is connected, that $\mathcal{V}_*$ is not empty, and that for each $e \in \mathcal{E}$ the degree of the polynomial V_e is fixed (and is at least 3). Then, $\mathcal{G}$ is almost surely controllable if we pick the coefficients of each V_e at random according to a probability law that is absolutely continuous w.r.t. Lebesgue.*

Example 2.6.4. *The 1D chain (shown in Figure 2.5) is controllable. Our theory applies when the interactions are polynomials of degree at least 3; for a somewhat different variant, see [23]. To apply our criteria, we start with $C = \{c\}$. Clearly, v_1 is alone in its C-equivalence class, and is therefore controllable by Corollary 2.5.6. We then take $C' = \{c, v_1\}$. Since v_2 is alone in its C'-equivalence class, it is also controllable. Continuing like this, we find that the whole chain is controllable.*

Observe that the chain described in the example above is controllable whether some pairs of springs are equivalent or not. There are in fact many networks that are controllable thanks to their topology alone, regardless of the potentials. In particular, we have

Figure 2.5 – A one-dimensional chain.

Example 2.6.5 (Physically relevant networks)**.** *We consider the network in Figure 2.6(a) and we start with $C = \{c_1, \dots, c_4\}$ (i.e., we assume that the vertices in the first column are controllers). Since no two vertices in the second column are C-siblings, they each belong to a distinct C-equivalence class, and therefore by Corollary 2.5.6 they are controllable (regardless of the potentials). Let us now denote by C' the set of all vertices in the first two columns, which are controllable as we have just seen. Repeating the argument above, we obtain that the vertices in the third column are controllable. Continuing like this, we gain control of the whole network. In the same way, one also easily obtains that the networks in Figure 2.6(b-d) are controllable thanks to their topology alone.*

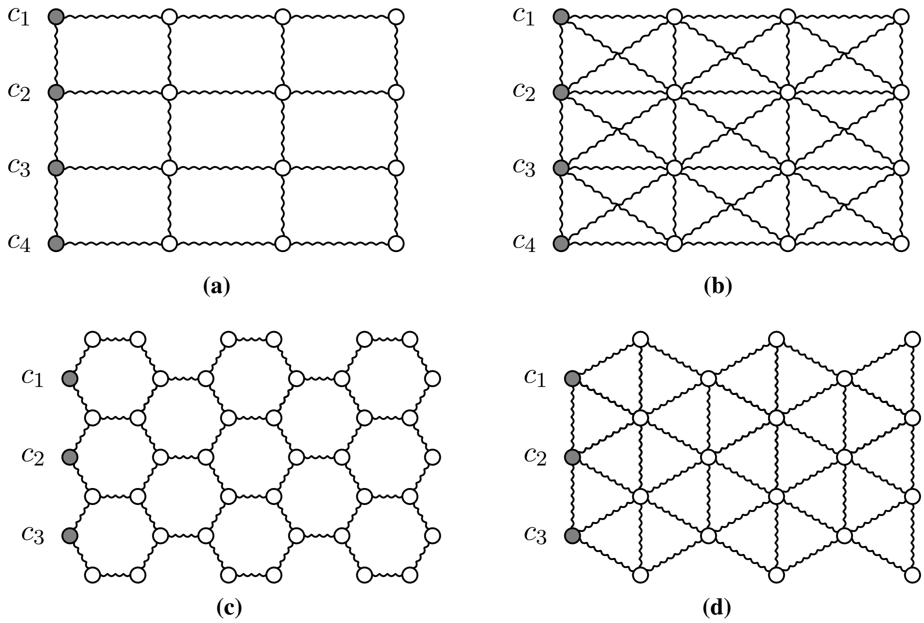


Figure 2.6 – Four networks that are controllable by their topology alone, regardless of the potentials (as long as they are polynomials of degree at least 3).

2.7. Limitations and extensions

Our theory is local in the sense that the central tool (Theorem 2.4.3) involves only a controller and its first neighbors. When we "walk through the graph," starting from $\mathcal{V}_*$ and taking at each step control of more particles, we only look at the interaction potentials that involve the particles we already control and their first neighbors. We never look "farther" in the graph. This makes our criteria quite easy to apply, but this is also the main limitation of our theory, as illustrated in

Example 2.7.1. *We consider the network shown in Figure 2.7, where c is a controller. If V''_{cv_1} and V''_{cv_3} are equivalent, then our theory fails to say anything about the controllability of the network. In order to draw any conclusion, one has to look at "what comes next" in the network. Of course, if the lower branch is an exact copy of the upper one (i.e., if the interaction and pinning potentials are the same), then the network is truly uncontrollable, and this is obvious for symmetry reasons. However, without such an "unfortunate" symmetry, the network may still be controllable. Indeed, by the study above, we know that the vector field $Y \equiv \partial_{q_{v_1}} + \partial_{q_{v_3}}$ is in $\mathcal{M}$. By commuting with X_0 and subtracting some contributions already in $\mathcal{M}$, one easily obtains that the vector field*

$$U''_{v_1}\partial_{p_{v_1}} + U''_{v_3}\partial_{p_{v_3}} + V''_{v_1v_2} \cdot (\partial_{p_{v_1}} - \partial_{p_{v_2}}) + V''_{v_3v_4} \cdot (\partial_{p_{v_3}} - \partial_{p_{v_4}})$$

is in $\mathcal{M}$. Observe that now the pinning potentials U_{v_1} and U_{v_3} as well as the interaction potentials $V_{v_1v_2}$ and $V_{v_3v_4}$ come into play. Taking first commutators with Y and then taking double commutators among the obtained vector fields, one obtains further vector fields of the form $\sum_{i=1}^4 g_i(q_{v_1}, q_{v_2}, q_{v_3}, q_{v_4})\partial_{p_{v_i}}$, where the g_i involve derivatives and products of the potentials mentioned above. In many cases, these are enough to prove that the network in Figure 2.7 is controllable, even though our theory fails to say so.

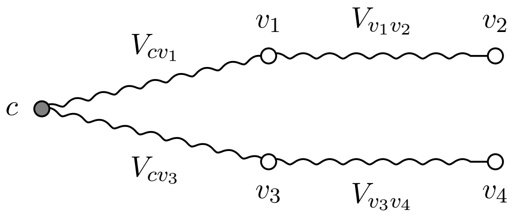


Figure 2.7 – The network used in Example 2.7.1. If V_{cv_3} is equivalent to V_{cv_1} our theory does not allow to conclude, but the network might still be controllable.

One question that might arise is: why does only the *second* derivative of the interaction potentials enter the theory? The next example shows that this issue is related to the notion of locality mentioned above.

Example 2.7.2. *We consider the network in Figure 2.8, where c is a controller. We study the case where*

$$\begin{aligned} V_{cv}(q_c - q_v) &= (q_c - q_v)^4, & U_v(q_v) &= q_v^6\ , \\ V_{cw}(q_c - q_w) &= (q_c - q_w)^4 + a \cdot (q_c - q_w), & U_w(q_w) &= q_w^6 + b \cdot q_w\ , \end{aligned}$$

for some constants a and b. The terms in a and b act as constant forces on c and w. Since $V''_{cv} \sim V''_{cw}$, the particles v and w are equivalent with respect to c by our definition. Thus, our theory fails to say anything. We seem to be missing the fact that when $a \neq 0$, the particles v and w can be told apart due to the first derivative of the potentials. However, having $a \neq 0$ is not enough; the controllability of the network also depends on b. Indeed, if $a = b$, the vector field X_0 is symmetric in v and w, and *therefore the network is genuinely uncontrollable. If now $a \neq b$, we have checked, by following a different strategy of taking commutators, that the network is controllable. Consequently, when two potentials have equivalent second derivative, but inequivalent first derivative, no conclusion can be drawn in general without knowing more about the network (here, it is one of the pinning potentials, but in more complex situations, it can be some subsequent springs).*

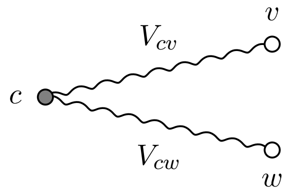


Figure 2.8 – The network discussed in Example 2.7.2.

Our theory applies only to strictly anharmonic systems, since we assume that the interaction potentials have degree at least 3. The next example shows what can go wrong if we drop this assumption. Again, this is related to the locality of our criteria.

Example 2.7.3. *We consider the harmonic system shown in Figure 2.9. The vertex c is a controller, and all the pinning potentials are equal and harmonic, i.e., of the form $\lambda x^2/2$. The interaction potentials are also harmonic. The spring $\{c, v_1\}$ has coupling constant 2, the springs $\{c, v_2\}$ and $\{v_2, v_3\}$ have coupling constant 1 and the spring $\{v_3, v_4\}$ has coupling $k > 0$. Since $V''_{cv_1} \equiv 2$ and $V''_{cv_2} \equiv 1$, the particles v_1 and v_2 are inequivalent with respect to c. Yet, this is not enough to obtain that they are controllable (unlike in the strictly anharmonic case covered by our theory). With standard methods for harmonic systems, it can be shown that the network is controllable iff $k \neq 2$. When $k = 2$, one of the eigenmodes decouples from the controller c, and no particle except c is controllable. Thus, one cannot obtain that v_1 and v_2 are controllable without knowing more about the potentials that come farther in the graph.*

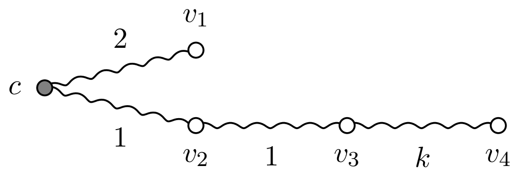


Figure 2.9 – A harmonic network that may or may not be controllable depending on the value of the coupling constant k.

Remark 2.7.4. As presented here, our method only works when the motion of each particle is 1D. To some extent, our results can be generalized to higher dimensions. For example, one can check that in any dimension $r \geq 1$, the network of Figure 2.4 with potentials $V_k(\mathbf{x}_k) = a_k \, (x_{k,1}^2 + \cdots + x_{k,r}^2)^2$, $k = 1, \ldots, n$, is controllable when the a_k are all distinct and non-zero. But for generic polynomial potentials, the situation is more complicated: taking multiple commutators does not always lead to tractable expressions (in particular, we do not have the nice form (2.4.6) for double commutators anymore). Further research is needed to find an adequate method for general higher dimensional problems. For some networks with special topology (such as the one in Figure 2.6(a) but not the ones in Figure 2.6(b-d)), simple conditions can be found for controllability, even for non-polynomial potentials (see [21]).

2.8. Comparison with other commutator techniques

It is perhaps useful to compare the techniques used in this paper to those used elsewhere: To unify notation, we consider the hypoellipticity problem in the classical form

$$L = X_0 + \sum_{i>0} X_i^2 \ .$$

In [23], the authors considered a chain, so that $\mathcal{V}_*$ is just the first and the last particle in the chain. Starting with ∂_{p_1} (the left end of the chain) one then forms (with simplified notation, which glosses over details which can be found in that paper)

$$\partial_{q_1} = [\partial_{p_1}, X_0] \ , \quad \partial_{p_2} = (M_{1,2})^{-1}[\partial_{q_1}, X_0] \ , \quad \partial_{q_2} = [\partial_{p_2}, X_0] \ ,$$

and so on, going through the chain. Here, the particles are allowed to move in several dimensions, and $M_{j,j+1}$ is basically the Hessian matrix of $V_{j,j+1}$. This technique requires that $M_{j,j+1}$ be invertible, which implies some restrictions on the potentials.

Villani [56] uses another sequence of commutators:

$$C_0 = \{X_i\}_{i>0} \ , \quad C_{j+1} = [C_j, X_0] + \text{ remainder}_j \ .$$

With this superficial notation, the current paper uses again a walk through the network, but the basic step involves double commutators of the form

$$[\![Z_1 : Z_2]\!]$$

with Z_i typically of the form $\sum g_v(x_v)\partial_{p_v}$, where we use abundantly that the V_e are polynomials. This allows for the “fanning out” of Figure 2.4 and is at the basis of our ability to control very general networks. In particular, this shows that networks with variable cross-section can be controlled.

2.9. Appendix: Vandermonde determinants

Lemma 2.9.1. *Let $c_1, \ldots, c_n \in \mathbb{R}$ be distinct and non-zero, and let $s \geq 0$. Then, for all $k \in \{1, \ldots, n\}$ there are constants $r_1, \ldots, r_n \in \mathbb{R}$ such that for all $j = 1, \ldots, n$,*

$$\sum_{i=1}^{n} r_i \, c_j^{i+s} = \delta_{jk} \ .$$

Proof. We have that the Vandermonde determinant

$$\begin{vmatrix} c_1^{s+1} & c_1^{s+2} & \cdots & c_1^{s+n} \\ c_2^{s+1} & c_2^{s+2} & \cdots & c_2^{s+n} \\ & \vdots & & \vdots \\ c_n^{s+1} & c_n^{s+2} & \cdots & c_n^{s+n} \end{vmatrix} = \left(\prod_{i=1}^{n} c_i^{s+1}\right) \begin{vmatrix} 1 & c_1 & c_1^2 & \cdots & c_1^{n-1} \\ 1 & c_2 & c_2^2 & \cdots & c_2^{n-1} \\ & \vdots & \vdots & & \vdots \\ 1 & c_n & c_n^2 & \cdots & c_n^{n-1} \end{vmatrix} = \prod_{i=1}^{n} c_i^{s+1} \prod_{j=i+1}^{n} (c_j - c_i)$$

is non-zero under our assumptions. Thus, the columns of this matrix form a basis of $\mathbb{R}^n$, which proves the lemma. $\square$

3. Chains of three rotors

We reproduce here [13]. The main result was obtained independently by Christophe Poquet and myself. Fortunately, we were made aware of this early enough, and decided to write [13] together. The only addition in this chapter compared to [13] is the discussion about the lower bound in §3.7.

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

with Jean-Pierre Eckmann and Christophe Poquet

Nonlinearity **28** (2015), 2397–2421

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.

3.1. 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 [23] 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.,* [19], 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 [50]: First, the authors used a more probabilistic approach, based on Harris recurrence as developed by Meyn and Tweedie [43]. 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 [9], 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 [34] in the case of a chain of 3 oscillators with strong pinning.

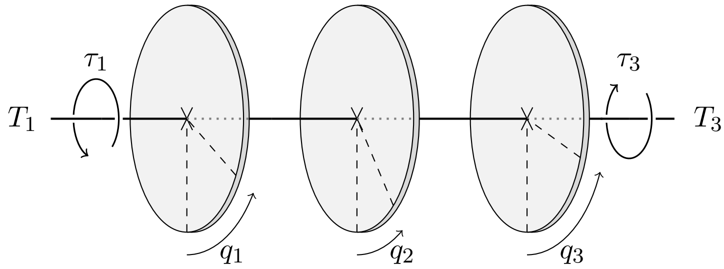


Figure 3.1 – A chain of three rotors with two external torques τ_1 and τ_3 and two heat baths at temperatures T_1 and T_3.

In this paper, we discuss a model with three rotors (see Figure 3.1), 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 $(\Omega, \mathcal{B})$, where $\mathcal{B}$ is the Borel σ-field over Ω. We will denote the points of Ω 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 τ_1 and τ_3 to the two outer rotors. Introducing $w_b = W_b'$ and $u_i = U_i'$, we obtain the system of SDE:

$$\begin{aligned}
\mathrm{d}q_i(t) &= p_i(t)\,\mathrm{d}t\,, \qquad i = 1, 2, 3\,, \\
\mathrm{d}p_2(t) &= -\sum_{b=1,3} w_b\big(q_2(t) - q_b(t)\big)\,\mathrm{d}t - u_2(q_2(t))\,\mathrm{d}t\,, \\
\mathrm{d}p_b(t) &= \Big(w_b\big(q_2(t) - q_b(t)\big) + \tau_b - u_b\big(q_b(t)\big) - \gamma_b p_b(t)\Big)\,\mathrm{d}t + \sqrt{2\gamma_b T_b}\,\mathrm{d}B_t^b\,, \qquad b = 1, 3\,,
\end{aligned} \tag{3.1.1}$$

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 3.1.1. Our model can be viewed as an extreme case of that studied in [34]. 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 $\frac{1}{2} - \frac{1}{2k}$. 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 3.1.2. *There is at least one $b \in \{1,3\}$ such that for each $s \in \mathbb{T}$, at least one of the derivatives $w_b^{(k)}(s), k \geq 1$ 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 (3.1.1). Since the coefficients of the SDE (3.1.1) 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

$$\|\nu\|_f = \sup_{|h|\leq f} \int_\Omega h \mathrm{d}\nu$$

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

Theorem 3.1.3. *Under Assumption 3.1.2, the following holds for the Markov process defined by (3.1.1):*

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

(ii) The process admits a unique invariant measure π, 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 $x = (q_1, q_2, \dots, p_3) \in \Omega$,

$$\|P^t(x,\,\cdot\,) - \pi\|_{e^{\beta' H}} \leq C(1 + p_2^2) e^{\beta H(x)} e^{-\lambda t^{1/2}} \,.$$

Remark 3.1.4. If both heat baths are at the same temperature, say, $T_1 = T_3 = T > 0$, and the forces τ_1 and τ_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 3.1.5. 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.,* [37]) 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 [37] 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 only study systems made of three rotors. We will present some numerical simulations of our system in §3.6, 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 [34] 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 3.1.3 in the present paper is based on a systematisation of this idea, as explained in §3.3.4.

The paper is structured as follows: In §3.2 we introduce a sufficient condition for subgeometric ergodicity from [16]. In §3.3 we study the behaviour of the middle rotor. In §3.4 we show how to use the study of p_2 to get a Lyapunov function. In §3.5 we provide the necessary technical input to the theorem of [16]. Finally, we illustrate numerically some properties of the non-equilibrium steady state in §3.6.

3.2. Ergodicity and Lyapunov functions

The proof of Theorem 3.1.3 relies on the results of [16] which in turn are based on the theory exposed in [43]. The theory of [43] 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 a on $[0,\infty)$ and a non-zero measure ν_a on Ω such that for all $x \in K$ one has $\int_0^\infty P^t(x,\,\cdot\,)a(\mathrm{d}t) \geq \nu_a(\,\cdot\,)$. In the case we are interested in, control arguments and the hypoellipticity of the generator imply that each compact set is petite (see §3.5.1 for a proof of this property).

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

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

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

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 Ω, a function such that $V(q,p) \to \infty$ when $\|p\| \to \infty$) such that for all $x \in \Omega$,

$$(LV)(x) \leq C\mathbf{1}_K(x) - \varphi \circ V(x) \,, \tag{3.2.1}$$

where C is a constant, $\varphi : [1, \infty) \to (0, \infty)$ 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 μ-irreducible for some measure μ (*i.e.,* $\mu(A) > 0$ implies that for all $x \in \Omega$ there exists $k \in \mathbb{N}$ such that $P^{k\Delta}(x, A) > 0$), then the Markov process is ergodic, with rate depending on φ. In the case where $\varphi(V) \propto V^{\rho}$, the convergence is geometric if $\rho = 1$ and polynomial if $\rho < 1$ (see [16,44]). In this paper, we obtain $\varphi(V) \sim V/\log V$.

We rely on the work of Douc, Fort and Guillin [16], 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 Theorem 3.2 and Theorem 3.4 of [16].

Theorem 3.2.1 (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 $V(q, p) \to \infty$ when $\|p\| \to \infty$, an increasing, differentiable, concave function $\varphi : [1, \infty) \to (0, \infty)$, a petite set K, and a constant C such that (3.2.1) holds. Then the process admits a unique invariant measure π, and for each $z \in [0, 1]$, there exists a constant C' such that for all $t \geq 0$ and all $x \in \Omega$,*

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

where $g(t) = (\varphi \circ H_{\varphi}^{-1}(t))^{z-1}$, with $H_{\varphi}(u) = \int_1^u \frac{\mathrm{d}s}{\varphi(s)}$.

When $z = 0$, we retrieve the total variation norm $\|P^t(x, \cdot\,) - \pi\|_{\mathrm{TV}}$ 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 (3.2.1) is satisfied with $\varphi(s) \sim s/\log s$, and a set K which is compact and therefore petite. This yields a stretched exponential convergence rate (see (3.2.4)). The existence of an irreducible skeleton required by Theorem 3.2.1 and the fact that every compact set is petite are proved in §3.5.

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

$$LH = \sum_b \left(\tau_b p_b + \gamma_b (T_b - p_b^2)\right)\ , \tag{3.2.2}$$

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 (3.2.1) 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 (3.2.2) 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 3.2.2. *For all sufficiently small $\beta > 0$, there is a function $V : \Omega \to [1, \infty)$ satisfying the two following properties:*

1. *There are positive constants c_1, c_2 such that*

$$1 + c_1 e^{\beta H} \leq V \leq c_2(1 + p_2^2)e^{\beta H} .$$

2. *There are positive constants c_3, c_4 and a compact set K such that*

$$LV \leq c_3 \mathbf{1}_K - \varphi(V) ,$$

where $\varphi : [1, \infty) \to (0, \infty)$ is defined by[1]

$$\varphi(s) = \frac{c_4\, s}{2 + \log(s)} . \tag{3.2.3}$$

The way we construct the Lyapunov function is somewhat different from that of [34]. There, it is obtained starting from some power of the Hamiltonian and then adding corrections by an averaging technique similar to ours (see Remark 3.3.8). 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 [34]. The present method can in principle be applied to the model of [34] (see also [33]).

We now show how the main results follows.

Proof of Theorem 3.1.3. The conclusions of Theorem 3.1.3 immediately follow from Theorem 3.2.1, Proposition 3.2.2, the technical results stated in Proposition 3.5.1, and the following two observations. Consider $0 \leq \beta' < \beta$ and choose $z \in (0, 1)$ such that $\beta' < z\beta$. First, the function φ defined in (3.2.3) yields, in the notation of Theorem 3.2.1, a convergence rate

$$g(t) = (\varphi \circ H_\varphi^{-1}(t))^{z-1} \leq ce^{-\lambda t^{1/2}} \tag{3.2.4}$$

for some $c, \lambda > 0$. Indeed, we have $H_\varphi(u) = \frac{1}{c_4}\int_1^u \frac{2+\log s}{s} \mathrm{d}s = \frac{1}{2c_4}(\log u)^2 + \frac{2}{c_4}\log u$, so that $H_\varphi^{-1}(t) = \exp((2c_4 t + 4)^{1/2} - 2)$ and $(\varphi \circ H_\varphi^{-1}(t)) = (2c_4 t + 4)^{-1/2} \exp((2c_4 t + 4)^{1/2} - 2) \geq Ce^{C' t^{1/2}}$ for some $C, C' > 0$. Thus, (3.2.4) holds with $\lambda = (1 - z)C'$. Secondly, by Proposition 3.2.2

[1] The 2 in the denominator ensures that φ is concave and increasing on $[1, \infty)$, as required in Theorem 3.2.1.

(i), and since $\beta' < z\beta$, we observe that $e^{\beta' H} \leq c(\varphi \circ V)^z$ for some constant $c > 0$, so that $\|\cdot\|_{e^{\beta' H}} \leq c\|\cdot\|_{(\varphi\circ V)^z}$. $\square$

3.3. 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.[2] In this section, we obtain some asymptotic, effective dynamics for p_2 when $p_2 \to \infty$.

3.3.1. 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 $u_i \equiv 0$ and that $W_b(s) = -\varkappa \cos(s)$ so that $w_b(s) = \varkappa \sin(s)$. 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 ω), but large enough for some "quasi-stationary" regime to be reached. The reader can think of ω 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

$$\mathrm{d}p_b = \varkappa \sin(\omega t)\,\mathrm{d}t - \gamma_b p_b\,\mathrm{d}t + \sqrt{2\gamma_b T_b}\,\mathrm{d}B_t^b\,,$$

whose solution is

$$\begin{aligned} p_b(t) &= \varkappa \frac{\gamma_b \sin(\omega t) - \omega\cos(\omega t)}{\gamma_b^2 + \omega^2} + \sqrt{2\gamma_b T_b}\int_0^t e^{\gamma_b(s-t)}\mathrm{d}B_s^b \\ &= -\varkappa\frac{\cos(\omega t)}{\omega} + \sqrt{2\gamma_b T_b}\int_0^t e^{\gamma_b(s-t)}\mathrm{d}B_s^b + \mathcal{O}\left(\frac{1}{\omega^2}\right)\,. \end{aligned}$$

We have neglected the exponentially decaying part $p_b(0)e^{-\gamma_b t}$ since we assume that a quasi-stationary regime is reached. By (3.2.2), 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

$$\begin{aligned} \mathbb{E}p_b^2(t) &= \varkappa^2\frac{\cos^2(\omega t)}{\omega^2} + 2\gamma_b T_b \mathbb{E}\left(\int_0^t e^{\gamma_b(s-t)}\mathrm{d}B_s^b\right)^2 + \mathcal{O}\left(\frac{1}{\omega^3}\right) \\ &= \varkappa^2\frac{\cos^2(\omega t)}{\omega^2} + (1 - e^{-2\gamma_b t})T_b + \mathcal{O}\left(\frac{1}{\omega^3}\right)\,, \end{aligned}$$

where we have used the Itō isometry $\mathbb{E}(\int_0^t e^{\gamma_b(s-t)}\mathrm{d}B_s^b)^2 = \int_0^t e^{2\gamma_b(s-t)}\mathrm{d}s$. Neglecting again an exponentially decaying term, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}H(t) = \sum_b \mathbb{E}\left(\gamma_b(T_b - p_b^2(t))\right) \sim -\sum_b \gamma_b \varkappa^2\frac{\cos^2(\omega t)}{\omega^2}\,. \tag{3.3.1}$$

[2] To simplify notation, we say p_2 is large, but we always really mean that $|p_2|$ is large.

Since $\cos^2(\omega t)$ oscillates very rapidly around its average $1/2$, we expect to see an effective contribution $-\frac{\gamma\varkappa^2}{2\omega^2}$. This approximation was obtained by assuming that p_2 is almost constant and equal to ω. 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 $\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}H \sim p_2\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}p_2$. Comparison with (3.3.1) leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}p_2 \sim -\frac{1}{p_2^3}\sum_b \frac{\gamma_b\varkappa^2}{2} .$$

We will obtain this result rigorously in Proposition 3.3.4.

3.3.2. Notations

Let $\Omega^\dagger = \{(q,p) \in \Omega : p_2 \neq 0\}$. We denote throughout by $X_t = (q(t), p(t))$ the solution of the stochastic differential equation (3.1.1) with initial condition $X_0 = (q(0), p(0))$. For now, we restrict ourselves to $X_0 \in \Omega^\dagger$ 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 $\frac{\mathrm{d}}{\mathrm{d}t}p_2$ is bounded, there is for each initial condition $X_0 \in \Omega^\dagger$ a deterministic time $t^* > 0$ (proportional to $|p_2(0)|$) such that $X_t \in \Omega^\dagger$ for all $t \in [0, t^*)$ and all realisations of the random noises. To define a smooth Lyapunov function on the whole space Ω we will perform a regularisation in §3.4.

Definition 3.3.1. *We let $\mathcal{U}$ be the set of stochastic processes u_t which are solutions of an SDE of the form*

$$\mathrm{d}u_t = f_1(X_t)\mathrm{d}t + f_2(X_t)\mathrm{d}B_t^1 + f_3(X_t)\mathrm{d}B_t^3 , \tag{3.3.2}$$

for some functions $f_i : \Omega \to \mathbb{R}$.

Notation: In the sequel, we write

$$\mathrm{d}u_t = f_1\mathrm{d}t + f_2\mathrm{d}B_t^1 + f_3\mathrm{d}B_t^3$$

instead of (3.3.2).

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

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 Ω we call the quantity

$$\langle f \rangle = \frac{1}{2\pi}\int_0^{2\pi} f\,\mathrm{d}q_2$$

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 3.3.2. *We assume*

$$\langle U_2 \rangle = 0 \qquad \textit{and} \qquad \langle W_b \rangle = 0, \qquad b = 1, 3 .$$

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

$$\begin{aligned} W_b &= W_b(q_2 - q_b)\ , & w_b &= w_b(q_2 - q_b)\ , & b &= 1, 3\ , \\ U_i &= U_i(q_i)\ , & u_i &= u_i(q_i)\ , & i &= 1, 2, 3\ . \end{aligned}$$

To simplify the notations, we also introduce the potentials Φ_1, Φ_2, Φ_3 associated to the three rotors, and the corresponding forces $\varphi_1, \varphi_2, \varphi_3$ defined by

$$\begin{aligned} \Phi_b &= W_b + U_b\ , & \varphi_b &= -\partial_{q_b}\Phi_b = w_b - u_b\ , & b &= 1, 3\ , \\ \Phi_2 &= W_1 + W_3 + U_2\ , & \varphi_2 &= -\partial_{q_2}\Phi_2 = -w_1 - w_3 - u_2\ . \end{aligned} \tag{3.3.3}$$

Of course, Φ_i and φ_i are functions of q only. With these notations the dynamics reads more concisely

$$\begin{aligned} \mathrm{d}q_i &= p_i\,\mathrm{d}t\ , & i &= 1, 2, 3\ , \\ \mathrm{d}p_2 &= \varphi_2\,\mathrm{d}t\ , & & \\ \mathrm{d}p_b &= \Big(\varphi_b + \tau_b - \gamma_b p_b\Big)\,\mathrm{d}t + \sqrt{2\gamma_b T_b}\,\mathrm{d}B_t^b, & b &= 1, 3\ . \end{aligned}$$

We will mainly deal with functions of the form $p_2^\ell p_1^n p_3^m g(q)$ and their linear combinations. We therefore introduce the notion of degree.

Definition 3.3.3. *We say that a function f on $\Omega^\dagger$ has* degree *$\ell \in \mathbb{Z}$ if it can be written as a* finite *sum of elements of the kind $p_2^\ell p_1^n p_3^m g(q)$ for some $n, m \in \mathbb{N}$ and a smooth function $g : \mathbb{T}^3 \to \mathbb{R}$. Moreover, we denote*

$$\hat{\mathcal{O}}(p_2^\ell)$$

a generic expression of order at most ℓ (which can vary from line to line), i.e., a finite sum of functions of degree ℓ, $\ell - 1$, $\ell - 2$,. . . .

We have by the Itō formula that for any smooth function f on Ω

$$\begin{aligned} \mathrm{d}f &= \sum_{i=1}^{3} \left(\frac{\partial f}{\partial q_i}\mathrm{d}q_i + \frac{\partial f}{\partial p_i}\mathrm{d}p_i\right) + \sum_b \gamma_b T_b \frac{\partial^2 f}{\partial p_b^2}\,\mathrm{d}t \\ &= \mathrm{d}^+ f + \mathrm{d}^0 f + \mathrm{d}^- f\ , \end{aligned}$$

where

$$\begin{aligned} \mathrm{d}^+ f &= p_2 \frac{\partial f}{\partial q_2}\,\mathrm{d}t\ , \\ \mathrm{d}^- f &= \varphi_2 \frac{\partial f}{\partial p_2}\,\mathrm{d}t\ , \\ \mathrm{d}^0 f &= \sum_b \left(p_b \frac{\partial f}{\partial q_b} + (\varphi_b + \tau_b - \gamma_b p_b)\frac{\partial f}{\partial p_b} + \gamma_b T_b \frac{\partial^2 f}{\partial p_b^2}\right)\,\mathrm{d}t + \sum_b \sqrt{2\gamma_b T_b}\frac{\partial f}{\partial p_b}\mathrm{d}B_t^b\ . \end{aligned} \tag{3.3.4}$$

(By the discussion following Definition 3.3.1, f, its partial derivatives, p_2 and the functions φ_i in this SDE are evaluated on the trajectory X_t.) Observe that when acting on a function of degree ℓ, 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.

3.3.3. General idea

In this section we introduce the main idea, which consists in 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

$$\mathrm{d}p_2 = \varphi_2 \, \mathrm{d}t \, . \tag{3.3.5}$$

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 $\langle \varphi_2 \rangle = 0$. 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 Ω. 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 $q_2(t) \sim q_2(0) + p_2(0)t$ (remember that q_2 is defined modulo 2π) and

$$g(q(t), p(t)) \sim g\big(q_1(0), q_2(0) + p_2(0)t, q_3(0), p(0)\big) \, , \tag{3.3.6}$$

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 $p_2(t) = \int_0^t \varphi_2(q(s))\mathrm{d}s$ in this situation. Approximating φ_2 as in (3.3.6) and integrating formally with respect to time (remember that $\varphi_2 = -\partial_{q_2} \Phi_2$) leads naturally to the decomposition

$$p_2 = \bar{p}_2 - \frac{\Phi_2(q)}{p_2} \, , \tag{3.3.7}$$

which consists in writing p_2 as sum of an oscillatory term Φ_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 (3.3.7) we get

$$\begin{aligned}
\mathrm{d}\bar{p}_2 &= \mathrm{d}\left(p_2 + \frac{\Phi_2}{p_2}\right) \\
&= \mathrm{d}p_2 + \mathrm{d}^+ \frac{\Phi_2}{p_2} + \mathrm{d}^0 \frac{\Phi_2}{p_2} + \mathrm{d}^- \frac{\Phi_2}{p_2} \\
&= \varphi_2 \, \mathrm{d}t - \varphi_2 \, \mathrm{d}t - \left(\frac{p_1 w_1}{p_2} + \frac{p_3 w_3}{p_2}\right) \mathrm{d}t - \frac{\varphi_2 \Phi_2}{p_2^2} \, \mathrm{d}t \\
&= \qquad\qquad\quad - \left(\frac{p_1 w_1}{p_2} + \frac{p_3 w_3}{p_2}\right) \mathrm{d}t - \frac{\varphi_2 \Phi_2}{p_2^2} \, \mathrm{d}t \, .
\end{aligned} \tag{3.3.8}$$

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 Φ_2/p_2 to p_2 has the effect that $\mathrm{d}^+(\Phi_2/p_2) = -\varphi_2\,\mathrm{d}t$, which precisely cancels the right-hand side of (3.3.5) while the remaining terms have negative powers of p_2. This observation is the starting point of the systematisation of the method.

Unfortunately, (3.3.8) is not good enough to understand how $\bar{p}_2$ (and therefore p_2) decreases in the long run, since the dynamics (3.3.8) 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 $\mathrm{d}\bar{\bar{p}}_2$ 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 to write the successive changes of variable explicitly. More precisely, we will prove

Proposition 3.3.4. *There is a function $F = \frac{\Phi_2(q)}{p_2} + \hat{\mathcal{O}}(p_2^{-2})$ such that whenever $p_2(t) \neq 0$ the process $\tilde{p}_2(t) = p_2(t) + F(X_t)$ satisfies*

$$\mathrm{d}\tilde{p}_2(t) = a(X_t)\,\mathrm{d}t + \sum_b \sigma_b(X_t)\mathrm{d}B_t^b\ , \tag{3.3.9}$$

with

$$\begin{aligned} a(q,p) &= -\frac{\gamma_1\left\langle W_1^2\right\rangle + \gamma_3\left\langle W_3^2\right\rangle}{p_2^3} + \hat{\mathcal{O}}\big(p_2^{-4}\big)\ , \\ \sigma_b(q,p) &= \frac{\sqrt{2\gamma_b T_b}W_b}{p_2^2} + \hat{\mathcal{O}}\big(p_2^{-3}\big)\ , \qquad b = 1,3\ . \end{aligned}$$

(By Assumption 3.3.2, no arbitrary additive constant appears in $\left\langle W_1^2\right\rangle$ and $\left\langle W_3^2\right\rangle$.)

The next two sections are devoted to proving Proposition 3.3.4.

3.3.4. 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 [52, 55], it is an *external* small parameter ε that gives the time scale of the fast variables. Here, the role of ε 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 Ω we find an expression of the kind

$$\mathrm{d}h = f\,\mathrm{d}t + \mathrm{d}r_t\ , \tag{3.3.10}$$

for some function $f = f(X_t)$ of degree ℓ and some stochastic process $r_t \in \mathcal{U}$ (see Definition 3.3.1) 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 Ω such that

$$\mathrm{d}\bar{h} = \langle f \rangle \, \mathrm{d}t + \mathrm{d}r_t + \text{small corrections}\,, \tag{3.3.11}$$

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 $f\,\mathrm{d}t$, the q_2-average $\langle f \rangle\,\mathrm{d}t$ plus some smaller corrections. In other words, we are looking for some F such that

$$\mathrm{d}F = \mathrm{d}(h - \bar{h}) = (f - \langle f \rangle)\,\mathrm{d}t + \text{small corrections}\,.$$

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 $\mathrm{d}^+F = (f - \langle f \rangle)\,\mathrm{d}t$. If we write $L^+ = p_2\partial_{q_2}$, we have $\mathrm{d}^+F = L^+F\mathrm{d}t$. 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 $\mathcal{K}$ the space of smooth functions $\Omega^\dagger \to \mathbb{R}$, and we denote by $\mathcal{K}_0$ the space of functions $f \in \mathcal{K}$ such that $\langle f \rangle = 0$. Note that L^+ maps $\mathcal{K}$ to $\mathcal{K}_0$ since for all $f \in \mathcal{K}$, we have by periodicity

$$\left\langle L^+ f \right\rangle = p_2 \left\langle \partial_{q_2} f \right\rangle = 0\,.$$

We can define a right inverse $(L^+)^{-1} : \mathcal{K}_0 \to \mathcal{K}_0$ by letting for all $g \in \mathcal{K}_0$

$$(L^+)^{-1} g = \frac{1}{p_2}\left(\int g\,\mathrm{d}q_2 + c(p, q_1, q_3)\right)\,,$$

where the integration "constant" $c(p, q_1, q_3)$ is uniquely defined by requiring that $\left\langle (L^+)^{-1} g\right\rangle = 0$.

This leads naturally to the following

Definition 3.3.5. *For any function $f \in \mathcal{K}$, we define the operator $Q : \mathcal{K} \to \mathcal{K}_0$ by*

$$Qf = (L^+)^{-1}(f - \langle f \rangle)\,.$$

Remark 3.3.6.

- If f is a function of degree ℓ, then Qf is of degree $\ell - 1$.
- By construction,
$$\mathrm{d}(Qf) = (f - \langle f \rangle)\,\mathrm{d}t + \mathrm{d}^0(Qf) + \mathrm{d}^-(Qf)\,. \tag{3.3.12}$$
- Moreover, by definition, Qf is the only function such that
$$\partial_{q_2}(Qf) = \frac{f - \langle f \rangle}{p_2} \quad \text{and} \quad \langle Qf \rangle = 0\,. \tag{3.3.13}$$

Therefore, if (3.3.10) holds for some f of degree ℓ, then we obtain a quantitative expression for (3.3.11), namely

$$\mathrm{d}(h - Qf) = \langle f \rangle \, \mathrm{d}t + \mathrm{d}r_t - \mathrm{d}^0(Qf) - \mathrm{d}^-(Qf)\,,$$

where the corrections are small in the sense that Qf, $\mathrm{d}^0(Qf)$ and $\mathrm{d}^-(Qf)$ have degree respectively $\ell - 1$, $\ell - 1$ and $\ell - 2$.

Remark 3.3.7. Observe that (3.3.7) can be written now as $p_2 = \bar{p}_2 + Q\varphi_2$, since $Q\varphi_2 = -\Phi_2/p_2$. Thus, the "naive" correction we added in (3.3.7) also follows from the systematic method we have just introduced. This is no surprise: the naive correction in (3.3.7) was motivated by the approximation (3.3.6) in which only q_2 moves, which corresponds to considering only d^+.

Remark 3.3.8. Our averaging procedure is inspired by techniques of [34]. 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 $q_2^{2k}/(2k)$ 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 $E_2^{\frac{1}{2}-\frac{1}{2k}}$, 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 (3.3.12).

Lemma 3.3.9. *(Averaging lemma) Consider some function $f = \hat{\mathcal{O}}(p_2^\ell)$ for some $\ell \in \mathbb{Z}$. Then*

$$f\,\mathrm{d}t = \langle f\rangle\,\mathrm{d}t - \mathrm{d}^0\,(Qf) - \mathrm{d}^-\,(Qf) + \mathrm{d}(Qf)\;,$$

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

We now prove Proposition 3.3.4 by using Lemma 3.3.9 repeatedly.

3.3.5. Proof of Proposition 3.3.4

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

$$\langle \partial_{q_2} f\rangle = 0\;. \tag{3.3.14}$$

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 3.3.2, (3.3.3) and (3.3.14) that

$$\langle W_b\rangle = \langle w_b\rangle = \langle\Phi_2\rangle = \langle\varphi_2\rangle = 0\;.$$

We start by doing again the first step, which we did in §3.3.3, but this time using the new toolset. In order to average the right-hand side of

$$\mathrm{d}p_2 = \varphi_2\,\mathrm{d}t\;,$$

we use Lemma 3.3.9 with $f = \varphi_2$, which is of order 0. We have $\langle f\rangle = 0$ and $Qf = -\Phi_2/p_2$ (by definition of φ_2 and Φ_2). We obtain

$$\begin{aligned}\mathrm{d}p_2 &= \mathrm{d}^0\left(\frac{\Phi_2}{p_2}\right) + \mathrm{d}^-\left(\frac{\Phi_2}{p_2}\right) - \mathrm{d}\left(\frac{\Phi_2}{p_2}\right)\\ &= \frac{1}{p_2}\sum_b p_b\frac{\partial\Phi_2}{\partial q_b}\,\mathrm{d}t - \frac{\varphi_2\Phi_2}{p_2^2}\,\mathrm{d}t - \mathrm{d}\left(\frac{\Phi_2}{p_2}\right)\\ &= -\frac{1}{p_2}\sum_b p_b w_b\,\mathrm{d}t - \frac{\varphi_2\Phi_2}{p_2^2}\,\mathrm{d}t - \mathrm{d}\left(\frac{\Phi_2}{p_2}\right)\ .\end{aligned} \tag{3.3.15}$$

This is exactly what we found in (3.3.8).

We deal next with the terms $-p_b w_b/p_2\ \mathrm{d}t$ in (3.3.15). Using Lemma 3.3.9 with $f = p_b w_b/p_2$ (and therefore with $Qf = p_b W_b/p_2^2$), we find, since $\langle f\rangle = p_b\langle w_b\rangle/p_2 = 0$, that for $b = 1, 3$,

$$\begin{aligned}\frac{p_b w_b}{p_2}\,\mathrm{d}t = &-\frac{1}{p_2^2}\left[-p_b^2 w_b + (\varphi_b + \tau_b - \gamma_b p_b)W_b\right]\,\mathrm{d}t\\ &-\frac{1}{p_2^2}\sqrt{2\gamma_b T_b}W_b\mathrm{d}B_t^b + \frac{2}{p_2^3}p_b W_b\varphi_2\,\mathrm{d}t + \mathrm{d}\hat{\mathcal{O}}\big(p_2^{-2}\big)\ ,\end{aligned} \tag{3.3.16}$$

where here and in the sequel, we denote by $\mathrm{d}\hat{\mathcal{O}}(p_2^k)$ any generic expression of the kind $\mathrm{d}w(X_t)$ for some function $w = \hat{\mathcal{O}}(p_2^k)$ on Ω. Here $\mathrm{d}\hat{\mathcal{O}}\big(p_2^{-2}\big) = \mathrm{d}(p_b W_b p_2^{-2})$. Substituting (3.3.16) into (3.3.15) leads to

$$\mathrm{d}p_2 = I\,\mathrm{d}t + J\,\mathrm{d}t + \frac{1}{p_2^2}\sum_b\sqrt{2\gamma_b T_b}W_b\mathrm{d}B_t^b + \mathrm{d}\left(-\frac{\Phi_2}{p_2} + \hat{\mathcal{O}}\big(p_2^{-2}\big)\right)\ , \tag{3.3.17}$$

with

$$\begin{aligned}I &= -\sum_b\frac{p_b^2 w_b - (\varphi_b + \tau_b - \gamma_b p_b)W_b}{p_2^2} - \frac{\varphi_2\Phi_2}{p_2^2}\ ,\\ J &= \frac{2}{p_2^3}\sum_b p_b W_b\varphi_2\ .\end{aligned}$$

We next deal with the terms $I\,\mathrm{d}t$ and $J\,\mathrm{d}t$.

First, we show that $\langle I\rangle = 0$. It is immediate that $\left\langle p_2^{-2}p_b^2 w_b\right\rangle$ and $\left\langle p_2^{-2}(\tau_b - \gamma_b p_b)W_b\right\rangle$ are zero. Moreover, $\left\langle p_2^{-2}\varphi_2\Phi_2\right\rangle = -\frac{1}{2}p_2^{-2}\left\langle\partial_{q_2}\Phi_2^2\right\rangle = 0$. Thus,

$$\begin{aligned}\langle I\rangle &= \sum_b\left\langle\frac{1}{p_2^2}\varphi_b W_b\right\rangle = \sum_b\left\langle\frac{w_b - u_b}{p_2^2}W_b\right\rangle\\ &= -\sum_b\left(\frac{\left\langle\partial_{q_2}W_b^2\right\rangle}{2p_2^2} + \frac{u_b\langle W_b\rangle}{p_2^2}\right) = 0\ .\end{aligned}$$

Since I is of order -2 and $\langle I\rangle = 0$, we find that QI is of order -3 and thus $\mathrm{d}^-(QI) = \hat{\mathcal{O}}\big(p_2^{-4}\big)\,\mathrm{d}t$.

Applying Lemma 3.3.9 with $f = I$, we find

$$I\,\mathrm{d}t = -\mathrm{d}^0\,(QI) + \hat{\mathcal{O}}\big(p_2^{-4}\big)\,\mathrm{d}t + \mathrm{d}\hat{\mathcal{O}}\big(p_2^{-3}\big)\ . \tag{3.3.18}$$

Using that $\langle QI\rangle = 0$, the definition (3.3.4) of d^0 leads, upon inspection, to

$$\mathrm{d}^0\,(QI) = \sum_b w_b\partial_{p_b}(QI)\mathrm{d}t + \mathcal{E}\mathrm{d}t + \sum_b \hat{\mathcal{O}}\big(p_2^{-3}\big)\,\mathrm{d}B_t^b\ ,$$

where $\mathcal{E}$ is a sum of terms of order -3 and $\langle\mathcal{E}\rangle = 0$.

Applying Lemma 3.3.9 to $w_b\partial_{p_b}(QI)\mathrm{d}t$ and $\mathcal{E}\,\mathrm{d}t$, we obtain

$$\mathrm{d}^0\,(QI) = \sum_b \langle w_b\partial_{p_b}(QI)\rangle\,\mathrm{d}t + \hat{\mathcal{O}}\big(p_2^{-4}\big)\,\mathrm{d}t + \sum_b \hat{\mathcal{O}}\big(p_2^{-3}\big)\,\mathrm{d}B_t^b\ . \tag{3.3.19}$$

Using the definition of w_b, integrating by parts once and using (3.3.13), we have for $b = 1, 3$,

$$\langle w_b\partial_{p_b}(QI)\rangle = \langle\partial_{q_2}(W_b)Q(\partial_{p_b}I)\rangle = -\langle W_b\partial_{q_2}Q(\partial_{p_b}I)\rangle = -\frac{1}{p_2}\langle W_b\partial_{p_b}I\rangle\ .$$

Since $\partial_{p_b}I = -p_2^{-2}\,(2p_bw_b + \gamma_bW_b)$, we get

$$\langle w_b\partial_{p_b}(QI)\rangle = \left\langle\frac{1}{p_2^3}W_b\,(2p_bw_b + \gamma_bW_b)\right\rangle = \frac{1}{p_2^3}\gamma_b\left\langle W_b^2\right\rangle\ , \tag{3.3.20}$$

where again we have used that $\langle W_bw_b\rangle = \frac{1}{2}\left\langle\partial_{q_2}W_b^2\right\rangle = 0$.

Substituting (3.3.20) into (3.3.19) and then the result into (3.3.18) we finally get

$$I\,\mathrm{d}t = -\frac{\alpha}{p_2^3}\mathrm{d}t + \sum_b \hat{\mathcal{O}}\big(p_2^{-3}\big)\,\mathrm{d}B_t^b + \hat{\mathcal{O}}\big(p_2^{-4}\big)\,\mathrm{d}t + \mathrm{d}\hat{\mathcal{O}}\big(p_2^{-3}\big)\ , \tag{3.3.21}$$

where

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

We next deal with the term $J\,\mathrm{d}t$ of (3.3.17). First, by Lemma 3.3.9,

$$J\,\mathrm{d}t = \langle J\rangle\,\mathrm{d}t + \hat{\mathcal{O}}\big(p_2^{-4}\big)\ \mathrm{d}t + \sum_b \hat{\mathcal{O}}\big(p_2^{-4}\big)\,\mathrm{d}B_t^b + \mathrm{d}\hat{\mathcal{O}}\big(p_2^{-4}\big)\ . \tag{3.3.22}$$

Unfortunately, $\langle J\rangle \neq 0$,[3] and we will need some more subtle identifications. Integrating by parts, we

[3]For example if $W_b = -\cos(q_2 - q_b)$, there are in $\langle J\rangle$ some terms of the kind $\langle p_3\cos(q_2 - q_1)\sin(q_2 - q_3)\rangle$ and $\langle p_1\sin(q_2 - q_1)\cos(q_2 - q_3)\rangle$ which are non-zero.

have

$$\begin{aligned}\langle J\rangle &= \frac{2p_b}{p_2^3}\sum_b \langle W_b\varphi_2\rangle = -\frac{2p_b}{p_2^3}\sum_b \langle W_b\partial_{q_2}\Phi_2\rangle \\ &= \frac{2p_b}{p_2^3}\sum_b \langle(\partial_{q_2}W_b)\Phi_2\rangle = \frac{2p_b}{p_2^3}\sum_b \langle w_b\Phi_2\rangle \\ &= -\frac{1}{p_2^3}\sum_b p_b\partial_{q_b}\langle\Phi_2^2\rangle\ .\end{aligned} \tag{3.3.23}$$

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

$$\begin{aligned}\mathrm{d}\left(\frac{\langle\Phi_2^2\rangle}{p_2^3}\right) &= \mathrm{d}^0\left(\frac{\langle\Phi_2^2\rangle}{p_2^3}\right) + \mathrm{d}^-\left(\frac{\langle\Phi_2^2\rangle}{p_2^3}\right) \\ &= \sum_b p_b\partial_{q_b}\left(\frac{\langle\Phi_2^2\rangle}{p_2^3}\right)\,\mathrm{d}t + \hat{\mathcal{O}}\big(p_2^{-4}\big)\,\mathrm{d}t\ .\end{aligned} \tag{3.3.24}$$

Combining (3.3.23) and (3.3.24) we find

$$\langle J\rangle\,\mathrm{d}t = \hat{\mathcal{O}}(p_2^{-4})\,\mathrm{d}t + \mathrm{d}(p_2^{-3}\,\langle\Phi\rangle_2^2) = \hat{\mathcal{O}}(p_2^{-4})\,\mathrm{d}t + \mathrm{d}\hat{\mathcal{O}}(p_2^{-3})\ ,$$

so that from (3.3.22) we obtain

$$J\,\mathrm{d}t = \hat{\mathcal{O}}\big(p_2^{-4}\big)\,\mathrm{d}t + \sum_b \hat{\mathcal{O}}\big(p_2^{-4}\big)\,\mathrm{d}B_t^b + \mathrm{d}\hat{\mathcal{O}}\big(p_2^{-3}\big)\ .$$

This together with (3.3.17) and (3.3.21) finally shows that

$$\mathrm{d}p_2 = -\left(\frac{\alpha}{p_2^3} + \hat{\mathcal{O}}\big(p_2^{-4}\big)\right)\mathrm{d}t + \sum_b\left(\frac{\sqrt{2\gamma_b T_b}W_b}{p_2^2} + \hat{\mathcal{O}}\big(p_2^{-3}\big)\right)\mathrm{d}B_t^b + \mathrm{d}\left(-\frac{\Phi_2}{p_2} + \hat{\mathcal{O}}\big(p_2^{-2}\big)\right)\ ,$$

which implies (3.3.9) and completes the proof of Proposition 3.3.4.

Remark 3.3.10. 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 (3.3.9) in the canonical "Langevin" form

$$\mathrm{d}\tilde{p}_2(t) = \sum_b\big(-\tilde{\gamma}_b(X_t)\tilde{p}_2(t)\mathrm{d}t + \sigma_b(X_t)\mathrm{d}B_t^b\big)\ ,$$

with $\sigma_b(q,p) = \sqrt{2\gamma_b T_b}W_b/p_2^2 + \hat{\mathcal{O}}(p_2^{-3})$ as in Proposition 3.3.4 and $\tilde{\gamma}_b(q,p) = \gamma_b\left\langle W_b^2\right\rangle/p_2^4 + \hat{\mathcal{O}}(p_2^{-5})$. We would like to introduce an effective temperature $\tilde{T}_b$ by some Einstein-Smoluchowski relation of the kind $\sigma_b^2/(2\tilde{\gamma}_b) = \tilde{T}_b$ in the limit $|p_2|\to\infty$. Unfortunately,

$$\lim_{|p_2|\to\infty}\frac{\sigma_b^2}{2\tilde{\gamma}_b} = \frac{W_b^2}{\left\langle W_b^2\right\rangle}T_b\ ,$$

which instead of a constant is an oscillatory quantity (with mean T_b). Now observe that these oscillations disappear if we approximate the oscillatory term W_b in σ_b by its quadratic mean $\langle W_b^2 \rangle^{1/2}$. This approximation is reasonable in the following sense: for small t and large $|p_2|$, we have that $p_2(s) \approx p_2(0)$ for $s \leq t$, so that

$$\int_0^t \frac{\sqrt{2\gamma_b T_b} W_b}{p_2^2(s)} \mathrm{d}B_s^b \approx \frac{\sqrt{2\gamma_b T_b}}{p_2^2(0)} \langle W_b^2 \rangle^{1/2} M(t) \quad \text{with } M(t) = \int_0^t \frac{W_b}{\langle W_b^2 \rangle^{1/2}} \mathrm{d}B_s^b \,.$$

But then, by the Dambis-Dubins-Schwarz representation theorem, there is another Brownian motion $\tilde{B}^b$ such that $M(t) = \tilde{B}^b_{\tau(t)}$ with $\tau(t) = \int_0^t W_b^2 / \langle W_b^2 \rangle \mathrm{d}s$. Clearly, when $|p_2|$ is very large, $\tau(t) \approx t$ so that $M(t)$ is very close to $\tilde{B}_t^b$. In this sense, when $|p_2| \to \infty$, it is reasonable to approximate $(\sqrt{2\gamma_b T_b} W_b / p_2^2) \mathrm{d}B_s^b$ with $(\sqrt{2\gamma_b T_b} \langle W_b^2 \rangle^{1/2} / p_2^2) \mathrm{d}\tilde{B}_s^b$, so that the Einstein-Smoluchowski relation indeed holds with effective temperature $\tilde{T}_b = T_b$.

Remark 3.3.11. The ergodicity of 1D Langevin processes is well understood: for any $\delta \in (-1, 0)$, processes satisfying an SDE of the kind

$$\mathrm{d}p \sim -C_1 p^\delta \,\mathrm{d}t + C_2 \mathrm{d}B_t$$

asymptotically (when $|p| \to \infty$) are typically ergodic with a rate bounded above and below by $\exp(-c_\pm t^{(1+\delta)/(1-\delta)})$ for some constants $c_+, c_- > 0$ (see [16,33] and references therein, in particular [33] for the lower bound). As argued in Remark 3.3.10, the variable $\tilde{p}_2$ (which is expected to be the component of the system that limits the convergence rate) essentially obeys an equation of the kind $\mathrm{d}p \sim -C_1 p^{-3} \,\mathrm{d}t + C_2 p^{-2} \mathrm{d}B_t$ asymptotically. It is easy to check that a change of variable $y = p^3$ yields the asymptotic dynamics $\mathrm{d}y \sim -C_1' y^{-1/3} \,\mathrm{d}t + C_2' \mathrm{d}B_t$ so that with $\delta = -1/3$, we expect a rate $\exp(-ct^{1/2})$. This suggests that the rate of convergence we find is optimal.

3.4. Lyapunov function

We now prove Proposition 3.2.2. Throughout this section, $\tilde{p}_2$ is the function defined in Proposition 3.3.4. The basic idea is to consider a Lyapunov function

$$V \sim \rho(p) \tilde{p}_2^2 e^{\frac{\beta}{2}\tilde{p}_2^2} + e^{\beta H} \,,$$

where $\rho(p)$ is non-zero only when $|p_2|$ is much larger than $|p_1|$ and $|p_3|$. We will obtain that $LV \lesssim -\varphi(V)$, with $\varphi(s) \sim s / \log(s)$ as in Proposition 3.2.2. 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 $p_1, p_3 \sim 0$,

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

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

Lemma 3.4.1. *For $\beta > 0$ small enough, there are constants $C_1, C_2 > 0$ such that*

$$Le^{\beta H} \leq (C_1 - C_2(p_1^2 + p_3^2))e^{\beta H} .$$

Proof. We have $Le^{\beta H} = \sum_b \big(-\gamma_b\beta(1-\beta T_b)p_b^2 + \beta\tau_b p_b + \gamma_b\beta T_b\big)\, e^{\beta H}$. If $\beta < 1/\max(T_1, T_3)$, then $\gamma_b\beta(1-\beta T_b) > 0$. Moreover, since $\beta\tau_b p_b < \frac{1}{2}\gamma_b\beta(1-\beta T_b)p_b^2 + C$ for C large enough, we find the desired bound. □

Lemma 3.4.2. *For $\beta > 0$ small enough, there is a constant $C_3 > 0$ such that on $\Omega^\dagger = \{x \in \Omega : p_2 \neq 0\}$,*

$$L\big(\tilde{p}_2^2 e^{\frac{\beta}{2}\tilde{p}_2^2}\big) \leq \big(-C_3 + \hat{\mathcal{O}}\big(p_2^{-1}\big)\big)e^{\frac{\beta}{2}\tilde{p}_2^2} . \tag{3.4.1}$$

Proof. Introducing $f(s) = s^2 e^{\frac{\beta}{2}s^2}$, we have by the Itō formula and Proposition 3.3.4 that

$$\begin{aligned}
\mathrm{d}\big(\tilde{p}_2^2 e^{\frac{\beta}{2}\tilde{p}_2^2}\big) &= \mathrm{d}f(\tilde{p}_2) = f'(\tilde{p}_2)(a\,\mathrm{d}t + \sum_b \sigma_b \mathrm{d}B_t^b) + \frac{1}{2}f''(\tilde{p}_2)\sum_b \sigma_b^2\,\mathrm{d}t \\
&= (2\tilde{p}_2 + \beta\tilde{p}_2^3)e^{\frac{\beta}{2}\tilde{p}_2^2}(a\,\mathrm{d}t + \sum_b \sigma_b \mathrm{d}B_t^b) + \frac{1}{2}(2 + 5\beta\tilde{p}_2^2 + \beta^2\tilde{p}_2^4)e^{\frac{\beta}{2}\tilde{p}_2^2}\sum_b \sigma_b^2\,\mathrm{d}t .
\end{aligned}$$

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

$$L\big(\tilde{p}_2^2 e^{\frac{\beta}{2}\tilde{p}_2^2}\big) = \big(-\alpha\beta + \beta^2\sum_b \gamma_b T_b W_b^2 + \hat{\mathcal{O}}\big(p_2^{-1}\big)\big)e^{\frac{\beta}{2}\tilde{p}_2^2} ,$$

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

Convention: We fix $\beta > 0$ small enough so that the conclusions of Lemma 3.4.1 and Lemma 3.4.2 hold.

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

- $\Omega_1 = \{x \in \Omega : |p_2| < (p_1^2 + p_3^2)^k + R\}$,
- $\Omega_2 = \{x \in \Omega : (p_1^2 + p_3^2)^k + R \leq |p_2| \leq 2(p_1^2 + p_3^2)^k + 2R\}$,
- $\Omega_3 = \{x \in \Omega : |p_2| > 2(p_1^2 + p_3^2)^k + 2R\}$.

Fix some $m, n \in \mathbb{N}$ and $\ell \geq 1$. On $\Omega_2 \cup \Omega_3$, we have by definition $|p_2| \geq (p_1^2 + p_3^2)^k + R$, so that

$$\left|\frac{p_1^n p_3^m}{p_2^\ell}\right| \leq \frac{|p_1^n p_3^m|}{((p_1^2 + p_3^2)^k + R)^\ell} \qquad (\text{on } \Omega_2 \cup \Omega_3) .$$

Clearly, if k and R are large enough, the right-hand side is bounded by an arbitrarily small constant. Therefore, any given $\hat{\mathcal{O}}(p_2^{-1})$ 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 or equal to -1. Using this, we obtain

Lemma 3.4.3. *For k and R large enough, there are constants $C_4, \dots, C_7 > 0$ such that the following properties hold on $\Omega_2 \cup \Omega_3$:*

$$|\tilde{p}_2^2 - p_2^2| < C_4 \,, \tag{3.4.2}$$

$$L\big(\tilde{p}_2^2 e^{\frac{\beta}{2}\tilde{p}_2^2}\big) \le -C_5 e^{\frac{\beta}{2}\tilde{p}_2^2} \,, \tag{3.4.3}$$

$$C_6 e^{-\frac{\beta}{2}(p_1^2+p_3^2)} e^{\beta H} \le e^{\frac{\beta}{2}\tilde{p}_2^2} \le C_7 e^{-\frac{\beta}{2}(p_1^2+p_3^2)} e^{\beta H} \,. \tag{3.4.4}$$

Proof. Since $\tilde{p}_2 = p_2 + \Phi_2(q)/p_2 + \hat{\mathcal{O}}(p_2^{-2})$, we have $\tilde{p}_2^2 = p_2^2 + 2\Phi_2(q) + \hat{\mathcal{O}}(p_2^{-1})$. By taking k large enough, the $\hat{\mathcal{O}}(p_2^{-1})$ here is bounded by a constant on the set $\Omega_2 \cup \Omega_3$, which implies (3.4.2). Moreover, for large k and R, the $\hat{\mathcal{O}}\big(p_2^{-1}\big)$ in (3.4.1) is also bounded on $\Omega_2 \cup \Omega_3$ by an arbitrarily small constant, which implies (3.4.3). To prove (3.4.4), observe that

$$e^{\frac{\beta}{2}\tilde{p}_2^2} = e^{\frac{\beta}{2}\left(\tilde{p}_2^2 - p_2^2 - p_1^2 - p_3^2 - U(q)\right)} e^{\beta H} \,,$$

where $U(q)$ contains all the potentials appearing in H. This together with the boundedness of U and (3.4.2), implies (3.4.4). □

Convention: We fix k and R such that the conclusions of Lemma 3.4.3 hold.

Definition 3.4.4. *Let $\chi : \mathbb{R} \to [0,1]$ be a smooth function such that $\chi(s) = 0$ when $|s| < 1$ and $\chi(s) = 1$ when $|s| > 2$. We introduce the cutoff function*

$$\rho(p) = \chi\left(\frac{p_2}{(p_1^2 + p_3^2)^k + R}\right) \,,$$

and the Lyapunov function

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

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

By construction $\rho(p)$ is smooth, $\rho(p) = 0$ on Ω_1 and $\rho(p) = 1$ on Ω_3, with some transition on Ω_2. Remember that $\tilde{p}_2$ is by construction smooth on $\Omega^\dagger$, *i.e.,* when $p_2 \neq 0$. In particular, since $\Omega_2 \cup \Omega_3 \subset \Omega^\dagger$, the function $\rho(p)\tilde{p}_2^2 e^{\frac{\beta}{2}\tilde{p}_2^2}$ is smooth on Ω, and so is V. We can now finally give the

Proof of Proposition 3.2.2. We show here that V satisfies the conditions enumerated in Proposition 3.2.2 if A is large enough. Let us first prove the first statement, which is that there exist $c_1, c_2 > 0$ such that

$$1 + c_1 e^{\beta H} \le V \le c_2(1 + p_2^2)e^{\beta H} \,. \tag{3.4.5}$$

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 (3.4.2) and (3.4.4),

$$\begin{aligned}|A\rho(p)\tilde{p}_2^2 e^{\frac{\beta}{2}\tilde{p}_2^2}| &\le c(p_2 + C_4)^2 e^{-\frac{\beta}{2}(p_1^2+p_3^2)} e^{\beta H} \\ &\le c(p_2^2 + 2C_4 p_2 + C_4^2)e^{\beta H} \le c(1 + p_2^2)e^{\beta H} \,.\end{aligned}$$

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 3.2.2, which is that for c_3, c_4 large enough and a compact set K,

$$LV \leq c_3 \mathbf{1}_K - \varphi(V) \quad \text{with } \varphi(s) = \frac{c_4\, s}{2 + \log(s)} \,. \tag{3.4.6}$$

We first show that

$$LV \leq c\mathbf{1}_K - ce^{\beta H} \qquad \text{with} \quad K = \{x \in \Omega_1 \cup \Omega_2 : p_1^2 + p_3^2 \leq M\} \,, \tag{3.4.7}$$

for some large enough M. Clearly K is compact, since $\Omega_1 \cup \Omega_2 = \{x \in \Omega : |p_2| \leq 2(p_1^2+p_3^2)^k + 2R\}$.

- On Ω_1 we simply have $V = 1 + e^{\beta H}$. By Lemma 3.4.1, we have $LV \leq (C_1 - C_2(p_1^2+p_3^2))e^{\beta H}$. Since $\Omega_1 \setminus K = \{x \in \Omega_1 : p_1^2 + p_3^2 > M\}$, we have for large enough M that $LV \leq -ce^{\beta H}$ on $\Omega_1 \setminus K$, and therefore (3.4.7) holds on Ω_1.

- On Ω_2, the key is to observe that there is a polynomial $z(p_1, p_2, p_3)$ such that

$$|L(A\rho(p)\tilde{p}_2^2 e^{\frac{\beta}{2}\tilde{p}_2^2})| \leq z(p)e^{\frac{\beta}{2}\tilde{p}_2^2} \leq C_7 z(p) e^{-\frac{\beta}{2}(p_1^2+p_3^2)} e^{\beta H} \,,$$

  where the second inequality comes from (3.4.4). Now, since $p_1^2 + p_3^2 \sim |p_2|^{1/k}$ on Ω_2, we have that $z(p)e^{-\frac{\beta}{2}(p_1^2+p_3^2)}$ is bounded on Ω_2. Therefore, by this and Lemma 3.4.1, we have on Ω_2,

$$\begin{aligned} LV &\leq \Big(C_7 z(p) e^{\frac{\beta}{2}(-p_1^2-p_3^2)} + C_1 - C_2(p_1^2+p_3^2)\Big) e^{\beta H} \\ &\leq \big(c - C_2(p_1^2+p_3^2)\big) e^{\beta H} \,. \end{aligned}$$

  which, as in the previous case, implies that (3.4.7) holds on Ω_2 if M is large enough.

- On Ω_3, which is the critical region, we have $V = 1 + A\tilde{p}_2^2 e^{\frac{\beta}{2}\tilde{p}_2^2} + e^{\beta H}$. By Lemma 3.4.1 and (3.4.3), it holds in Ω_3 that

$$LV \leq (C_1 - C_2(p_1^2+p_3^2))e^{\beta H} - C_5 A e^{\frac{\beta}{2}\tilde{p}_2^2} \,. \tag{3.4.8}$$

  On the set $\{x \in \Omega_3 : C_1 - C_2(p_1^2+p_3^2) \leq -1\}$, we simply have $LV \leq -e^{\beta H}$, so that (3.4.7) holds trivially. On the other hand, on the set $\{x \in \Omega_3 : C_1 - C_2(p_1^2+p_3^2) > -1\}$ the quantity $p_1^2 + p_3^2$ is bounded, so that $e^{\frac{\beta}{2}\tilde{p}_2^2} \geq ce^{\beta H}$ by (3.4.4), which with (3.4.8) implies that

$$LV \leq (C_1 - C_2(p_1^2+p_3^2))e^{\beta H} - cAe^{\beta H} \leq (C_1 - cA)e^{\beta H} \,.$$

  By making A large enough, we again find a bound $LV \leq -ce^{\beta H}$, so that (3.4.7) holds.

Therefore, (3.4.7) holds on all of Ω. To obtain (3.4.6), 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 (3.4.5) we indeed have that $e^{\beta H} \geq cV/(1+p_2^2) \geq cV/(2+\log V)$. This completes the proof of Proposition 3.2.2. □

Remark 3.4.5. The external forces τ_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 $\alpha = \sum_b \gamma_b \left\langle W_b^2 \right\rangle$ to be strictly positive.

Remark 3.4.6. 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 [34]). 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.

3.5. Smoothness and irreducibility

This section is devoted to proving that the hypotheses of Theorem 3.2.1 other than the existence of the Lyapunov function are satisfied. More precisely we will prove the following proposition.

Proposition 3.5.1. *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 $(X_n)_{n=0,1,2,\dots}$ 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 [43], 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.

3.5.1. 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 $P^t(x, \mathrm{d}y)$ has a density $p_t(x,y)$ that is smooth in t, x and y, and every invariant measure has a smooth density [35].

We identify vector fields over Ω and the corresponding first-order differential operators in the usual way (we identify the tangent space of Ω with $\mathbb{R}^6$). This enables us to consider Lie algebras of vector fields over Ω 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 3.5.2. *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

$$A_0 = \sum_{i=1}^{3} (p_i \partial_{q_i} - u_i \partial_{p_i}) + \sum_b (w_b(\partial_{p_b} - \partial_{p_2}) + \tau_b \partial_{p_b} - \gamma_b p_b \partial_{p_b})$$

is the drift part of L.

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

Lemma 3.5.3. *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 ∂_{p_1} and ∂_{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 ∂_{q_2} and ∂_{p_2}. In the following, f denotes a generic function on Ω that can be each time different. We have $[\partial_{q_b}, A_0] = w_b'(q_2 - q_b)\partial_{p_2} + f(q)\partial_{p_b}$ so that commuting $n - 1$ times with ∂_{q_b} we get that for all $n \geq 1$

$$w_b^{(n)}(q_2 - q_b)\partial_{p_2} + f(q)\partial_{p_b} \in \mathcal{M} \,. \tag{3.5.1}$$

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

$$w_b^{(n)}(q_2 - q_b)\partial_{q_2} + f(q,p)\partial_{p_2} + f(q)\partial_{p_b} + f(q,p)\partial_{q_b} \in \mathcal{M} \,. \tag{3.5.2}$$

By Assumption 3.1.2, there is some $b \in \{1, 3\}$ such that for any fixed $x \in \Omega$, there is an integer $n \geq 1$ such that $w_b^{(n)}(q_2 - q_b) \neq 0$. Thus, by (3.5.1) and (3.5.2) the proof is complete. $\square$

Thus, we have proved Proposition 3.5.1 (i).

3.5.2. 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.

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 τ_1, τ_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 ν with density $\frac{1}{Z}\exp(-H/T)$ is invariant (with some normalisation constant Z) as mentioned earlier. Note that we do not assume *a priori* that ν 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 ν .

Lemma 3.5.4. *The equilibrium transition probabilities satisfy the following property: for every measurable set S one has for all t*

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

Proof. We have by the invariance of ν,

$$\begin{aligned}\int_{S^c} \tilde{P}^t(x, S)\mathrm{d}\nu - \int_S \tilde{P}^t(x, S^c)\mathrm{d}\nu &= \int_{S^c} \tilde{P}^t(x, S)\mathrm{d}\nu + \int_S (\tilde{P}^t(x, S) - 1)\mathrm{d}\nu \\ &= \int_\Omega \tilde{P}^t(x, S)\mathrm{d}\nu - \int_S 1\mathrm{d}\nu = \nu(S) - \nu(S) = 0 \ ,\end{aligned}$$

which completes the proof. □

Lemma 3.5.5. *Let A be a closed set. If A is invariant under $\tilde{P}^1$ (i.e., $\tilde{P}^1(x, A) = 1$ for all $x \in A$), then either $A = \emptyset$ or $A = \Omega$.*

Proof. By Lemma 3.5.4, $\int_{A^c} \tilde{P}^1(x, A)\mathrm{d}\nu = \int_A \tilde{P}^1(x, A^c)\mathrm{d}\nu = 0$ since $\tilde{P}^1(x, A^c) = 0$ for all $x \in A$. This implies that $\tilde{P}^1(x, A) = 0$ for all $x \in A^c$, since $x \mapsto \tilde{P}^1(x, A)$ is continuous on the open set A^c and ν has an everywhere positive density. But then $\tilde{P}^t(x, A)$ is 1 when $x \in A$ and 0 when $x \in A^c$, so that by continuity we have $\partial A = \emptyset$. Since Ω is connected, the conclusion follows. □

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

Lemma 3.5.6. *The time-1 skeleton $(X_n)_{n=0,1,2,\cdots}$ 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 $A^c = \{x \in \Omega : \exists n > 0, \tilde{P}^n(x, B) > 0\}$ is open, so that A is closed. Moreover, for all $x \in A$ it holds that $0 = \sum_{n=1}^\infty \tilde{P}^n(x, B) \geq \sum_{n=1}^\infty \tilde{P}^{n+1}(x, B) = \int_\Omega \tilde{P}^1(x, dy) \sum_{n=1}^\infty \tilde{P}^n(y, B)$. But since by the definition of A we have $\sum_{n=1}^\infty \tilde{P}^n(y, B) > 0$ for all $y \in A^c$, we must have $\tilde{P}^1(x, A^c) = 0$ for all $x \in A$, so that A is invariant. But then by Lemma 3.5.5 either $A = \emptyset$ or $A = \Omega$. We need to eliminate the second possibility. Since $m(B) > 0$ and ν has positive density, we have $\nu(B) > 0$. By the invariance of ν, we have $\int_\Omega \tilde{P}^1(x, B)\mathrm{d}\nu = \nu(B) > 0$. But then there is some $x \in \Omega$ such that $\tilde{P}^1(x, B) > 0$, so that $x \in A^c$.

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. □

Thus, we have proved Proposition 3.5.1(ii), so that the proof of Proposition 3.5.1 is complete.

Direct control version

We give now an alternate proof of Proposition 3.5.1(ii). We establish the irreducibility of our process by using controllability arguments. We aim to establish the controllability of (3.1.1), where the Brownian motions B_t^1 and B_t^3 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

$$\begin{aligned} \dot{q}_i(t) &= p_i(t) , \\ \dot{p}_2(t) &= -\sum_b w_b\big(q_2(t) - q_b(t)\big) , \\ \dot{p}_b(t) &= f_b(t) . \end{aligned} \tag{3.5.3}$$

In [22] 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

$$K^- = \sum_b \min_{s\in\mathbb{T}} w_b(s), \quad K^+ = \sum_b \max_{s\in\mathbb{T}} w_b(s) ,$$

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 $|p_2^f - p_2^i|$. 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 3.1.2 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 2π).

Proposition 3.5.7. *The system (3.5.3) 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 $T^* \leq c_1 + c_2|p_2^f - p_2^i|$ 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 (3.5.3) 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 [54] links the support of the semigroup P^t and the accessible points for (3.5.3), and implies in particular that for all $x^i = (q^i, p^i)$ and $t > c_1$ the subspace $\{x \in \Omega : |p_2 - p_2^i| \leq (t - c_1)/c_2\}$ 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 3.5.8. *Consider the system*

$$\begin{aligned} \dot{\bar{q}}_2(t) &= \bar{p}_2(t) , \\ \dot{\bar{p}}_2(t) &= g(t) - u_2(\bar{q}_2(t)) , \end{aligned} \tag{3.5.4}$$

and fix some initial and terminal conditions (q_2^i, p_2^i) *and* (q_2^f, p_2^f)*. We claim that there is a* T^* *satisfying* $T^* \le c_1 + c_2|p_2^f - p_2^i|$ *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 (3.5.4) with initial data* (q_2^i, p_2^i) *satisfies* $\bar{p}_2(T) = p_2^f$ *and* $\bar{q}_2(T) = q_2^f$*.*

Proof. We prove this result only in the case $u_2 \equiv 0$. If $p_2^f \ge p_2^i$, then let $\Theta = (p_2^f - p_2^i)/K^+$ and let $g(t) = K^+$ for all $t \in [0, \Theta)$. If $p_2^f < p_2^i$ let $\Theta = (p_2^f - p_2^i)/K^-$ and let $g(t) = K^-$ for all $t \in [0, \Theta)$. In both cases, $\bar{p}_2(\Theta) = p_2^f$, while $\bar{q}_2(\Theta)$ 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 $\bar{p}_2(\Theta + 2\Delta) = \bar{p}_2(\Theta) = p_2^f$ and

$$\bar{q}_2(\Theta + 2\Delta) = \bar{q}_2(\Theta) + 2\Delta p_2^f + a\Delta^2 .$$

Observe that as soon as $\Delta > \sqrt{2\pi/K^*}$, we can choose $a \in [0, K^*]$ so that $\bar{q}_2(\Theta + 2\Delta)$ takes any value (modulo 2π). In particular, we can choose it to be q_2^f, so that we have the advertised result with $T^* = \Theta + \sqrt{2\pi/K^*}$. □

Remark 3.5.9. 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 $K^- \le g(t) \le K^+$, it is possible to find piecewise smooth functions $q_b^*(t)$, $b = 1, 3$, such that $\sum_b w_b(\bar{q}_2(t) - q_b^*(t)) \equiv g(t)$, where $\bar{q}_2(t)$ is the solution of (3.5.4).
- 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 $q_b(t) = q_b^*(t)$ for all $t \in [0, T] \setminus A_\delta$, where A_δ consists of a finite number of intervals of total length at most δ. We can choose the controls f_b so that the $q_b(t)$ constructed here are solutions to (3.5.3) (when δ is small, $f_b(t)$ is typically very large for $t \in A_\delta$).
- Since the interaction forces w_b are bounded, their effect during the times $t \in A_\delta$ is negligible when δ is small. More precisely, it can be shown that the solution $q_2(t)$ and $p_2(t)$ of (3.5.3)

converge uniformly on $[0, T]$ to the solutions $\bar{q}_2(t)$ and $\bar{p}_2(t)$ of (3.5.4) when $\delta \to 0$. Therefore, the system is approximately controllable in the sense of Proposition 3.5.7.

3.6. Numerical illustrations

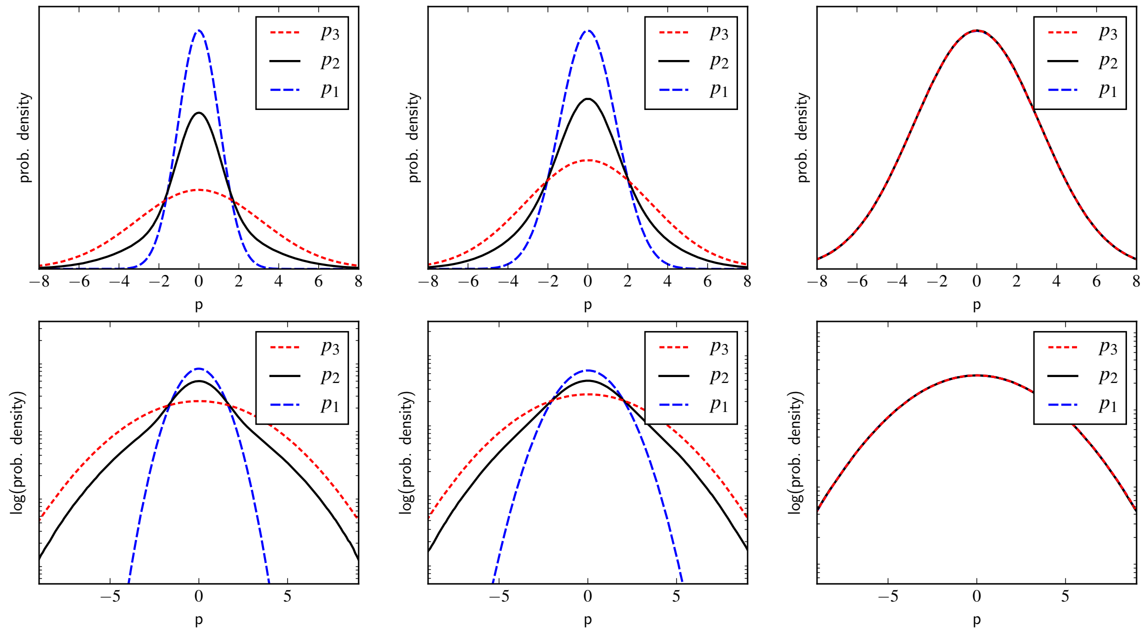
In this section we illustrate some properties of the invariant measure in the case where $U_i \equiv 0$ and $W_1 = W_3 = -\cos$.

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 τ_3. We apply the numerical algorithm given in [37] 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 $T_1 = T_3 = T$ 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).

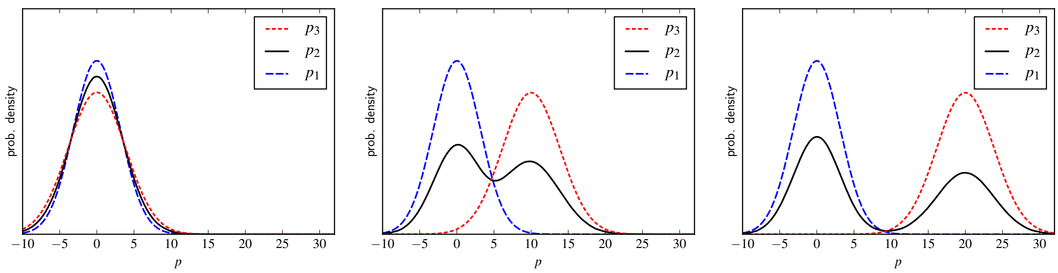
In Figure 3.2 we show the marginal distributions of p_1, p_2, p_3 for different temperatures and no external force. For each pair of temperatures, we show the distributions both in linear and logarithmic scale. At equilibrium, when $T_1 = T_3 = 10$, all three distributions coincide exactly and are Gaussian. 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 τ_3 on the marginal distributions of the p_i, for $T_1 = 10$ and $T_3 = 15$. As illustrated in Figure 3.3, the distributions of p_1 and p_3 are close to Gaussians with variance T_1 and T_3 and mean 0 and τ_3. Note that when $\tau_3 \neq 0$, the distribution of p_2 has two maxima: one at 0 and one at τ_3. The explanation for these two maxima can be found by looking at the trajectories $p_i(t)$ as shown in Figure 3.4 (for $\tau_3 = 20$); p_1 fluctuates around 0, p_3 fluctuates around τ_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 with 3 (since then the force w_3 oscillates with "high frequency" $p_3 - p_2 \sim \tau_3$). Inversely, in the regime where p_2 fluctuates around τ_3, it interacts strongly with 3 and only weakly with 1. Other simulations (not shown here) show that, as expected, the larger τ_3, the less frequent the switches between these two regimes. The asymmetry of the two maxima in Figure 3.3 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.



(a) $T_1 = 1, T_3 = 10$ **(b)** $T_1 = 2, T_3 = 10$ **(c)** $T_1 = 10, T_3 = 10$

Figure 3.2 – Distribution of p_1, p_2, p_3, with no external force and several temperatures.



(a) $\tau_3 = 0$ **(b)** $\tau_3 = 10$ **(c)** $\tau_3 = 20$

Figure 3.3 – Distribution of p_1, p_2, p_3, with $T_1 = 10, T_3 = 15$ for 3 values of τ_3.

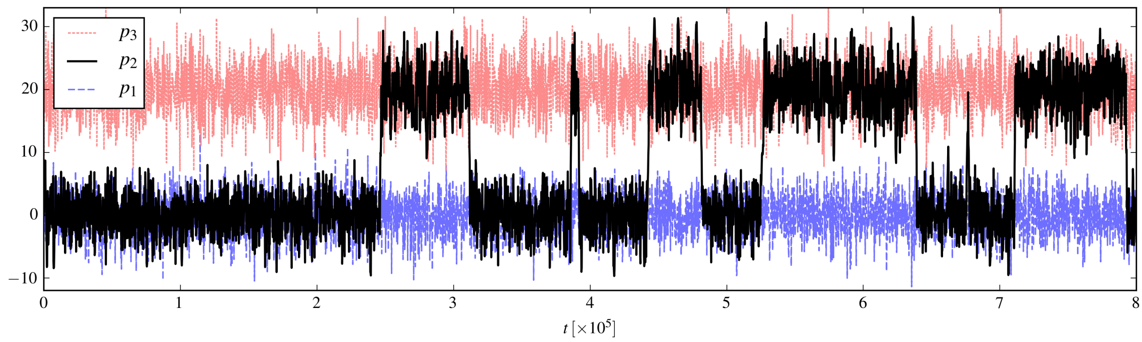


Figure 3.4 – Representation of the evolution of p_1, p_2, p_3 with $T_1 = 10, T_3 = 15, \tau_3 = 20$.

3.7. Supplement: lower bound on the convergence

In this supplement[4], we show that the stretched exponential bound obtained above is optimal in a sense. For simplicity, we consider here only a chain of two rotors (see Figure 3.5).

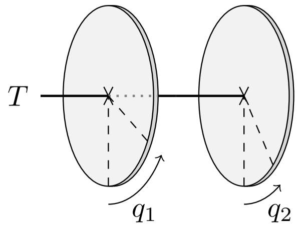


Figure 3.5 – A simplified system with two rotors.

The first rotor is coupled to a heat bath at temperature $T > 0$, with a coupling constant $\gamma > 0$. The second rotor interacts with the first one through a smooth potential $W(q_2 - q_1)$, and is not coupled to any heat bath. We apply no other external force to the system, and we take no pinning potential. The phase space is now $\Omega = \mathbb{T}^2 \times \mathbb{R}^2$, and we write $x = (q, p) = (q_1, q_2, p_1, p_2)$. The Hamiltonian is

$$H(q, p) = \frac{p_1^2 + p_2^2}{2} + W(q_2 - q_1) ,$$

and we consider the stochastic differential equation

$$\begin{aligned} \mathrm{d}q_i(t) &= p_i(t)\,\mathrm{d}t , & i = 1, 2 , \\ \mathrm{d}p_1(t) &= w\big(q_2(t) - q_1(t)\big)\,\mathrm{d}t - \gamma p_1(t)\,\mathrm{d}t + \sqrt{2\gamma T}\,\mathrm{d}B_t , \\ \mathrm{d}p_2(t) &= -w\big(q_2(t) - q_1(t)\big)\,\mathrm{d}t , \end{aligned} \tag{3.7.1}$$

where we have introduced the derivative w of W, and where B_t is a standard Wiener process. Without loss of generality, we choose the additive constant in W such that $\int_0^{2\pi} W(s)\mathrm{d}s = 0$. We denote again by $P^t(x, \cdot\,)$ the transition probabilities, by $\mathbb{E}_x$ the expectation with respect to the process started at x, and we introduce the generator

$$L = p_1\partial_{q_1} + p_2\partial_{q_2} + w(q_2 - q_1)(\partial_{p_1} - \partial_{p_2}) - \gamma p_1\partial_{p_1} + \gamma T\partial_{p_1}^2 .$$

The physical picture is essentially the same as with three rotors: the second rotor decouples when its energy is large, and the crux is to obtain an effective dynamics in this regime. Since there is only one temperature at hand, the invariant measure π is simply the Gibbs measure, *i.e.,*

$$\mathrm{d}\pi(q, p) = \frac{1}{Z}e^{-\frac{H}{T}}\mathrm{d}p\mathrm{d}q ,$$

where Z is a normalization constant.

[4] which is available as an independent note in [14]

We now formulate the main result of this section, which says that for the chain of two rotors considered here, the convergence to the invariant measure happens no faster than a stretched exponential with exponent $1/2$.

Theorem 3.7.1. *There is a constant $c_* > 0$ such that for each initial condition $x \in \Omega$, there exist a constant $C > 0$ and a sequence $(t_n)_{n\geq 0}$ increasing to infinity such that*

$$\|P^{t_n}(x, \,\cdot\,) - \pi\|_{\mathrm{TV}} \geq Ce^{-c_*\sqrt{t_n}}\,.$$

Proof. In Proposition 3.7.4 below, we will construct a test function $F : \Omega \to [1, \infty)$ such that $\pi(F^{1-\varepsilon}) = \infty$ for some $\varepsilon \in (0, 1)$, and such that for some $A > 0$ and all $x \in \Omega$,

$$\mathbb{E}_x F(x_t) \leq F(x)e^{\sqrt{2At}}\,. \tag{3.7.2}$$

The desired result then follows from [33, Theorem 3.6 and Corollary 3.7]. For completeness, we give here an explicit adaptation of the proof to the present case.

We fix $x \in \Omega$ and write $\nu_t = P^t(x, \,\cdot\,)$. The result follows from comparing an upper bound on the tail of ν_t with a lower bound on the tail of π.

- By (3.7.2) and Markov's inequality, we have for all $w > 0$ the upper bound

$$\nu_t(F > w) \leq \frac{F(x)e^{\sqrt{2At}}}{w}\,. \tag{3.7.3}$$

- Since $(1-\varepsilon)\int_1^\infty \pi(F > w)w^{-\varepsilon}\mathrm{d}w = \pi(F^{1-\varepsilon}) = \infty$, there is a sequence $(w_n)_{n\geq 0}$ increasing to infinity such that $\pi(F > w_n)w_n^{-\varepsilon} \geq w_n^{-1-\varepsilon/2}$. As a consequence, we have for each $n \geq 0$ the inequality

$$\pi(F > w_n) \geq \frac{1}{w_n^{1-\varepsilon/2}}\,. \tag{3.7.4}$$

By (3.7.3), (3.7.4) and the definition of the total variation norm, we have for all n that

$$\|\nu_t - \pi\|_{\mathrm{TV}} \geq \pi(F > w_n) - \nu_t(F > w_n) \geq \frac{1}{w_n^{1-\varepsilon/2}} - \frac{F(x)e^{\sqrt{2At}}}{w_n}\,.$$

Picking now t_n such that $F(x)e^{\sqrt{2At_n}} = \frac{1}{2}w_n^{\varepsilon/2}$, we obtain

$$\|\nu_{t_n} - \pi\|_{\mathrm{TV}} \geq \frac{1}{2w_n^{1-\varepsilon/2}} = Ce^{-c_*\sqrt{t_n}}\,,$$

with $c_* = (\frac{2}{\varepsilon} - 1)\sqrt{2A}$ and $C = \frac{1}{2}(2F(x))^{1-\frac{2}{\varepsilon}}$. This completes the proof. □

We now construct a function F that has the properties needed in the proof of Theorem 3.7.1. This function F will grow fast enough along the p_2-axis so that $\pi(F^{1-\varepsilon}) = \infty$ for all small enough ε. Moreover, F will satisfy a relation of the kind $LF \lesssim F/\log F$, which implies (3.7.2) as we will show.

We start by approximating the dynamics of p_2 by an "averaged" variable $\bar{p}_2$ in the regime where p_2 is very large. As we will need some control also when p_1 scales linearly with p_2 (see Proposition 3.7.4), we cannot simply use an expansion in negative powers of p_2 as we did in §3.3.4. We instead need to consider negative powers of $p_2 - p_1$, with the following more refined notion of *order* (see also Remark 3.7.5).

Definition 3.7.2. *For any continuous function $f : \mathbb{T}^2 \to \mathbb{R}$ and any choice of integers $k, \ell \geq 0$ and $m \in \mathbb{Z}$, we say that*

$$\frac{f(q)p_1^k p_2^m}{(p_2 - p_1)^\ell}$$

is of order *$k + m - \ell$. We denote by $\mathcal{R}(j)$ a generic remainder of order at most j, i.e., a finite sum of terms of order up to j.*

The usual rules apply, in particular $\mathcal{R}(j) + \mathcal{R}(k) = \mathcal{R}(\max(j, k))$, and $\mathcal{R}(j)\mathcal{R}(k) = \mathcal{R}(j + k)$. The aim now is to introduce a new variable $\bar{p}_2 = p_2 + \mathcal{R}(-1)$, which is defined when $p_2 \neq p_1$, and which satisfies

$$\mathrm{d}\bar{p}_2 = \mathcal{R}(-3)\mathrm{d}t + \mathcal{R}(-2)\mathrm{d}B_t \ . \tag{3.7.5}$$

In contrast to §3.3.4, it will not be necessary to estimate the remainders $\mathcal{R}(-3)$ and $\mathcal{R}(-2)$ here. For this reason (and the fact that there are only two rotors), the computations are short enough to simply proceed explicitly. We have

$$\mathrm{d}p_2 = -w(q_2 - q_1)\,\mathrm{d}t \ .$$

We then introduce a first correction

$$p_2^{(1)} = p_2 + \frac{W(q_2 - q_1)}{p_2 - p_1} \ ,$$

and obtain by Itô's formula

$$\mathrm{d}p_2^{(1)} = \frac{W(q_2 - q_1)}{(p_2 - p_1)^2}(2w(q_2 - q_1) - \gamma p_1)\,\mathrm{d}t + \mathcal{R}(-2)\mathrm{d}B_t + \mathcal{R}(-3)\mathrm{d}t \ .$$

Since $\int_0^{2\pi} W(s)\mathrm{d}s = 0$, there exists an indefinite integral $W^{[1]}$ of W on $\mathbb{T}$, which we choose so that $\int_0^{2\pi} W^{[1]}(s)\mathrm{d}s = 0$. In turn, we introduce an indefinite integral $W^{[2]}$ of $W^{[1]}$. By construction, we have $\big(W^{[1]}\big)' = W$ and $\big(W^{[2]}\big)' = W^{[1]}$.

We then set

$$p_2^{(2)} = p_2^{(1)} + \frac{\gamma p_1 W^{[1]}(q_2 - q_1) - (W(q_2 - q_1))^2}{(p_2 - p_1)^3} \ ,$$

and obtain

$$\mathrm{d}p_2^{(2)} = -\frac{\gamma^2 p_1 W^{[1]}(q_2-q_1)}{(p_2-p_1)^3} - \frac{3\gamma^2 p_1^2 W^{[1]}(q_2-q_1)}{(p_2-p_1)^4} + \mathcal{R}(-2)\mathrm{d}B_t + \mathcal{R}(-3)\mathrm{d}t \,.$$

We finally obtain (3.7.5) by letting

$$\bar{p}_2 = p_2^{(2)} + \frac{\gamma^2 p_1 W^{[2]}(q_2-q_1)}{(p_2-p_1)^4} + \frac{3\gamma^2 p_1^2 W^{[2]}(q_2-q_1)}{(p_2-p_1)^5} \,.$$

In order to construct the test function F, we now introduce some positive parameters β_-, β_+ and δ satisfying

$$\beta_- < \frac{1}{T} < \beta_+ < \left(1 + \frac{1}{(1+2\delta)^2}\right)\beta_- \,, \tag{3.7.6}$$

and consider the partition of Ω (see Figure 3.6) given by

- $\Omega_0 = \{x \in \Omega : p_1^2 + p_2^2 < 1\}$,
- $\Omega_1 = \{x \in \Omega : |p_2| \leq (1+\delta)|p_1|\} \setminus \Omega_0$,
- $\Omega_2 = \{x \in \Omega : (1+\delta)|p_1| < |p_2| \leq (1+2\delta)|p_1|\} \setminus \Omega_0$,
- $\Omega_3 = \{x \in \Omega : |p_2| > (1+2\delta)|p_1|\} \setminus \Omega_0$.

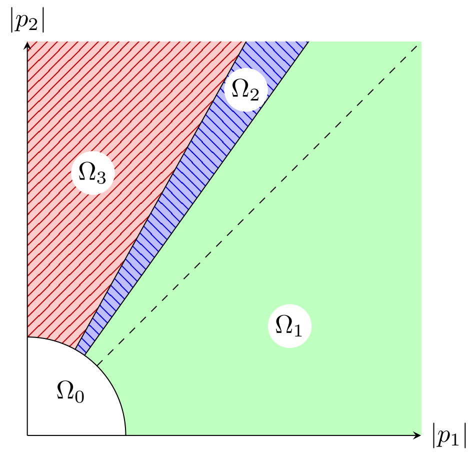


Figure 3.6 – Partition of Ω (in momentum space).

We immediately have

Lemma 3.7.3. *There are constants C_1 and C_2 such that on the set $\Omega_2 \cup \Omega_3$, we have the two inequalities*

$$|\bar{p}_2^2 - p_2^2| < C_1 \,, \tag{3.7.7}$$

$$Le^{\frac{\beta_+}{2}\bar{p}_2^2} \leq C_2 p_2^{-2} e^{\frac{\beta_+}{2}\bar{p}_2^2} \; . \tag{3.7.8}$$

Proof. Observe that for all $k, \ell \geq 0$ and $m \in \mathbb{Z}$, there is a constant C such that on the set $\Omega_2 \cup \Omega_3 = \{x \in \Omega : |p_2| > (1+\delta)|p_1|,\ p_1^2 + p_2^2 \geq 1\}$, we have

$$\left| \frac{p_1^k p_2^m}{(p_2 - p_1)^\ell} \right| \leq C |p_2|^{k+m-\ell} \; .$$

This implies that any remainder $\mathcal{R}(j)$ is bounded in absolute value by some constant times $|p_2|^j$ on $\Omega_2 \cup \Omega_3$. In particular, since $\bar{p}_2^2 = (p_2 + \mathcal{R}(-1))^2 = p_2^2 + \mathcal{R}(0)$, we obtain that (3.7.7) holds on $\Omega_2 \cup \Omega_3$ for some appropriate C_1.

In order to prove (3.7.8), we write $f(s) = e^{\frac{\beta_+}{2}s^2}$ and obtain by Itô's formula

$$\mathrm{d}\big(e^{\frac{\beta_+}{2}\bar{p}_2^2}\big) = \mathrm{d}f(\bar{p}_2) = f'(\bar{p}_2)(\mathcal{R}(-3)\,\mathrm{d}t + \mathcal{R}(-2)\mathrm{d}B_t) + \frac{1}{2}f''(\bar{p}_2)\mathcal{R}(-4)\,\mathrm{d}t \; .$$

We thus find, since $f'(\bar{p}_2) = \mathcal{R}(1)e^{\frac{\beta_+}{2}\bar{p}_2^2}$ and $f''(\bar{p}_2) = \mathcal{R}(2)e^{\frac{\beta_+}{2}\bar{p}_2^2}$, that

$$Le^{\frac{\beta_+}{2}\bar{p}_2^2} = f'(\bar{p}_2)\mathcal{R}(-3) + \frac{1}{2}f''(\bar{p}_2)\mathcal{R}(-4) = \mathcal{R}(-2)e^{\frac{\beta_+}{2}\bar{p}_2^2} \; .$$

Now, on the set $\Omega_2 \cup \Omega_3$, the $\mathcal{R}(-2)$ above is bounded by $C_2 p_2^{-2}$ for some $C_2 > 0$, and thus (3.7.8) holds. $\square$

We next introduce a smooth cutoff function $\rho : \mathbb{R}^2 \to [0,1]$ such that $\rho(p) = 1$ on Ω_3 and $\rho(p) = 0$ on Ω_1, with some transition on Ω_2. More precisely, let $\chi : [0,\infty] \to \mathbb{R}$ be a smooth function such that $\chi(s) = 1$ when $s \geq 1 + 2\delta$, and $\chi(s) = 0$ when $s \leq 1 + \delta$. On $\Omega \setminus \Omega_0$, we let

$$\rho(p) = \chi\left(\left| \frac{p_2}{p_1} \right| \right) \; ,$$

and we freely choose ρ on Ω_0 so that it is smooth on all of Ω.

We now define the function $F : \Omega \to [1, \infty)$ by

$$F(x) = 1 + e^{\beta_- H(x)} + \rho(p)e^{\frac{\beta_+}{2}\bar{p}_2^2} \; , \tag{3.7.9}$$

for some β_-, β_+, δ satisfying (3.7.6). Observe that while F resembles the Lyapunov function V of §3.4, it grows much faster along the p_2-axis.

Proposition 3.7.4. *Let F be as defined in (3.7.9). Then, $\pi(F^{1-\varepsilon}) = \infty$ for small enough ε, and (3.7.2) holds for large enough A.*

Proof. We let c be a generic positive constant which may vary from occurrence to occurrence. This constant may depend on the parameters at hand, but *not* on the point in Ω.

Let $\Gamma = \{x : |p_1| \le 1\} \cap \Omega_3$. Using (3.7.7) and the definition of H, we find

$$\begin{aligned}\pi(F^{1-\varepsilon}) &\ge \int_\Gamma \exp\left(\frac{\beta_+(1-\varepsilon)\bar{p}_2^2}{2}\right)\frac{\exp\left(-\frac{H}{T}\right)}{Z}\mathrm{d}p\mathrm{d}q \\ &\ge c\int_\Gamma \exp\left(\frac{\beta_+(1-\varepsilon)p_2^2}{2} - \frac{p_2^2}{2T}\right)\mathrm{d}p\mathrm{d}q\;.\end{aligned} \tag{3.7.10}$$

Provided that we pick ε small enough so that $\frac{1}{T} < \beta_+(1-\varepsilon)$, which is possible by (3.7.6), the last integral in (3.7.10) is infinite, and thus $\pi(F^{1-\varepsilon}) = \infty$.

We now prove the second assertion. As in §3.2, we introduce the concave and increasing function $\varphi : [1,\infty) \to (0,\infty)$ defined by

$$\varphi(s) = \frac{As}{2+\log s}$$

for some $A > 0$. We will show that if A is large enough,

$$LF \le \varphi(F)\;, \tag{3.7.11}$$

which implies the desired result. Indeed, assume that (3.7.11) holds. Let $(C_n)_{n\ge 0}$ be an increasing sequence of compact sets such that $C_n \uparrow \Omega$, and consider the corresponding first exit times $\tau_n = \inf\{t \ge 0 : x_t \notin C_n\}$. We have $\tau_n \to \infty$ almost surely, since the process is non-explosive. By Dynkin's formula, we find

$$\begin{aligned}\mathbb{E}_x F(x_{t\wedge\tau_n}) - F(x) &= \mathbb{E}_x \int_0^{t\wedge\tau_n} LF(x_s)\mathrm{d}s \le \mathbb{E}_x \int_0^{t\wedge\tau_n} \varphi(F(x_s))\mathrm{d}s \\ &\le \mathbb{E}_x \int_0^t \varphi(F(x_{s\wedge\tau_n}))\mathrm{d}s \;\le \int_0^t \varphi(\mathbb{E}_x F(x_{s\wedge\tau_n}))\mathrm{d}s,\end{aligned}$$

where the last inequality comes from Fubini's theorem and Jensen's inequality (since φ is concave). In other words, $g(t) \equiv \mathbb{E}_x F(x_{t\wedge\tau_n})$ satisfies the integral inequality $g(t) \le g(0) + \int_0^t \varphi(g(s))\mathrm{d}s$. The solution of the ordinary differential equation $y'(t) = \varphi(y(t))$ with $y(0) = y_0 \ge 1$ is

$$y(t) = \exp\left(\sqrt{(\ln(y_0)+2)^2 + 2At} - 2\right) \le y_0 e^{\sqrt{2At}}\;, \tag{3.7.12}$$

where we have used that $\sqrt{\cdot}$ is subadditive. By comparison, we thus obtain that $\mathbb{E}_x F(x_{t\wedge\tau_n}) \le F(x)\exp(\sqrt{2At})$. Taking the limit $n \to \infty$ and using Fatou's lemma gives (3.7.2).

Thus, it only remains to prove (3.7.11). First, observe that there is a polynomial $z(p_1, p_2)$ such that

$$\begin{aligned}L\big(\rho(p)e^{\frac{\beta_+}{2}\bar{p}_2^2}\big) &\le c + \mathbf{1}_{\Omega_3} L e^{\frac{\beta_+}{2}\bar{p}_2^2} + \mathbf{1}_{\Omega_2} z(p) e^{\frac{\beta_+}{2}\bar{p}_2^2} \\ &\le c + \mathbf{1}_{\Omega_3} c p_2^{-2} e^{\frac{\beta_+}{2}\bar{p}_2^2} + \mathbf{1}_{\Omega_2} z(p) e^{\frac{\beta_+}{2}\bar{p}_2^2}\;,\end{aligned} \tag{3.7.13}$$

where we have used (3.7.8).

Moreover, since $\beta_- < 1/T$, there is a set G of the form

$$G = \{x \in \Omega : |p_1| < c\}$$

such that

$$\begin{aligned} Le^{\beta_- H} &= \big((\beta_- T - 1)p_1^2 + T\big)\,\gamma\beta_- e^{\beta_- H} \le (c - cp_1^2)e^{\beta_- H} \\ &\le c\mathbf{1}_G e^{\beta_- H} - ce^{\beta_- H} \le c\mathbf{1}_G e^{\beta_- \frac{p_2^2}{2}} - ce^{\beta_- H}\,. \end{aligned} \tag{3.7.14}$$

Combining (3.7.13) and (3.7.14), we find

$$LF \le c + \mathbf{1}_{\Omega_3} c p_2^{-2} e^{\frac{\beta_+}{2}\bar{p}_2^2} + \mathbf{1}_{\Omega_2} z(p) e^{\frac{\beta_+}{2}\bar{p}_2^2} + c\mathbf{1}_G e^{\beta_- \frac{p_2^2}{2}} - ce^{\beta_- H}\,. \tag{3.7.15}$$

We now make two observations. First, on Ω_2, we have by (3.7.7), the definition of H, and the definition of Ω_2 that

$$z(p)e^{\frac{\beta_+}{2}\bar{p}_2^2} e^{-\beta_- H} \le cz(p)e^{\frac{\beta_+}{2}p_2^2 - \frac{\beta_-}{2}(p_1^2+p_2^2)} \le cz(p)e^{\frac{p_2^2}{2}\left(\beta_+ - \beta_-\left(1+\frac{1}{(1+2\delta)^2}\right)\right)}\,.$$

By (3.7.6), the above goes to zero when $\|p\| \to \infty$ in Ω_2, so we have

$$\mathbf{1}_{\Omega_2} z(p) e^{\frac{\beta_+}{2}\bar{p}_2^2} - ce^{\beta_- H} \le c\,. \tag{3.7.16}$$

In a similar way, since $\beta_+ > \beta_-$ and $G \subset \Omega_3 \cup K$ for some compact set K (on which $\exp(\beta_- \frac{p_2^2}{2})$ is bounded), we have

$$\mathbf{1}_G e^{\beta_- \frac{p_2^2}{2}} \le c + \mathbf{1}_{\Omega_3} c p_2^{-2} e^{\frac{\beta_+}{2}\bar{p}_2^2}\,, \tag{3.7.17}$$

where we have also used (3.7.7). Combining now (3.7.15), (3.7.16) and (3.7.17), we obtain

$$LF \le c + \mathbf{1}_{\Omega_3} c p_2^{-2} e^{\frac{\beta_+}{2}\bar{p}_2^2} \le c + \mathbf{1}_{\Omega_3} \frac{ce^{\frac{\beta_+}{2}\bar{p}_2^2}}{2 + \log\left(e^{\frac{\beta_+}{2}\bar{p}_2^2}\right)}\,,$$

where the second inequality uses once more (3.7.7). Since $\mathbf{1}_{\Omega_3}\exp(\frac{\beta_+}{2}\bar{p}_2^2) \le F$, and since the function $s \mapsto s/(2 + \log s)$ is increasing, we obtain

$$LF \le c + \frac{cF}{2 + \log F} \le \frac{cF}{2 + \log F}\,,$$

where the second inequality holds because $F \ge 1$. Thus, we indeed have (3.7.11) for large enough A, which completes the proof. □

Remark 3.7.5. The fact that we have to work with two different constants β_+ and β_- seems to force us to take a "transition region" Ω_2 where p_1 scales linearly with p_2 (all other attempts have resulted in some troublesome terms coming from the cutoffs). Unlike in §3.3.4, we are therefore not allowed to assume that p_1 is small when we compute $\bar{p}_2$, which forces us to work with negative powers of $(p_2 - p_1)$ instead of simply p_2. While this causes no trouble here with only two rotors, technical complications arise if we try to generalize the computations above to chains of three rotors. Indeed, terms involving $(p_2 - p_1)^{-j}$ and $(p_2 - p_3)^{-j}$ have to be assembled, and this leads to troublesome error terms. We are currently able to provide such a generalization only if the potential W consists of finitely many Fourier modes. In addition, with three rotors and two different temperatures, the invariant measure π is not known explicitly, and some supplementary work would have to be done to prove that the function F satisfies $\pi(F^{1-\varepsilon}) = \infty$ for some ε.

4. Chains of four rotors

The contents of this chapter is as published in [12], except for some references to [13] that have been replaced with the corresponding references to Chapter 3 in the present thesis.

Non-equilibrium steady states for chains of four rotors
with Jean-Pierre Eckmann
Communications in Mathematical Physics
2016, first online, DOI:10.1007/s00220-015-2550-2

Abstract

We study a chain of four interacting rotors (rotators) connected at both ends to stochastic heat baths at different temperatures. We show that for non-degenerate interaction potentials the system relaxes, at a stretched exponential rate, to a non-equilibrium steady state (NESS). Rotors with high energy tend to decouple from their neighbors due to fast oscillation of the forces. Because of this, the energy of the central two rotors, which interact with the heat baths only through the external rotors, can take a very long time to dissipate. By appropriately averaging the oscillatory forces, we estimate the dissipation rate and construct a Lyapunov function. Compared to the chain of length three (considered previously by C. Poquet and the current authors), the new difficulty with four rotors is the appearance of resonances when both central rotors are fast. We deal with these resonances using the rapid thermalization of the two external rotors.

4.1. Introduction

We consider a chain of 4 classical rotors interacting at both ends with stochastic heat baths at different temperatures. Under the action of such heat baths, many Hamiltonian systems are known to relax to an invariant probability measure called non-equilibrium steady state (NESS). In general, the explicit expression for this invariant measure is unknown, and the convergence rate depends on the nature of the system. For the model under consideration, we obtain a stretched exponential rate.

For several examples of Hamiltonian chains, properties of the NESS (*e.g.,* thermal conductivity, validity of the Fourier law, temperature profile, . . .) have been studied numerically, perturbatively, or via some effective theories. See for example [4, 7, 15, 17, 27, 30, 37, 42] for chains of rotors and [1, 4, 7, 8, 18, 31, 41, 42] for chains of oscillators. From a rigorous point of view however, the mere existence of an invariant measure is not evident, and has been proved only in special cases.

A lot of attention has been devoted to chains of classical oscillators with (nonlinear) nearest neighbor interactions. In such models, each oscillator has a position $q_i \in \mathbb{R}$ (we take one dimension for simplicity), is attached to the reference frame with a *pinning* potential $U(q_i)$, and interacts with its neighbors via some *interaction* potentials $W(q_{i+1} - q_i)$ and $W(q_i - q_{i-1})$.

It turns out that the properties of the chain depend crucially on the relative growth of W and U at high energy. In the case of (asymptotically) polynomial potentials, and for Markovian heat baths, it has been shown [9, 19, 22–24, 50] that if W grows faster than U, the system typically relaxes exponentially fast to a NESS. The convergence is fast because, thanks to the strong interactions, the sites in the bulk of the chain "feel" the heat baths effectively even though they are separated from them by other sites.

In the strongly pinned case, *i.e.,* when U grows faster than W, the situation is more complicated. When a given site has a lot of energy, the corresponding oscillator essentially feels only its pinning potential $U(q_i)$ and not the interaction. Assume $U(q) \propto q^{2k}$ with $k > 1$. An isolated oscillator pinned with a potential U and with an energy E oscillates with a frequency that grows like $E^{1/2-1/2k}$. This scaling plays a central role, since the larger the energy at a site, the faster the corresponding q_i oscillates. But then, the interaction forces with the sites $i+1$ and $i-1$ oscillate very rapidly and become ineffective at high energy. Therefore, a site (or a set of sites) with high energy tends to decouple from the rest of the chain, so that energy can be "trapped" in the bulk. This mechanism not only makes the convergence to the invariant measure slower, but it also makes the proof of its existence harder. The case where W is quadratic is considered in [34]. There, Hairer and Mattingly show that if $U(q) \propto q^{2k}$ with k sufficiently large, no exponential convergence to an invariant measure (if there is one) can take place. Moreover, they show that an invariant measure exists in the case of 3 oscillators when $k > 3/2$. The existence of a NESS for longer chains of oscillators remains an open problem when the pinning dominates the interactions.

Chains of rotors are in fact closely related to strongly pinned oscillator chains: The frequency of a rotor scales as $E^{1/2}$, where E is its energy. This scaling corresponds to that of an oscillator in the limit $k \to \infty$, for the pinning $U(q) \propto q^{2k}$ discussed above. In this sense, our chain of rotors behaves as a chain of oscillators in the limit of "infinite pinning", which is some kind of worst-case scenario from the point of view of the asymptotic decoupling at high energy. On the other hand, the compactness of the position-space (it is a torus) in the rotor case is technically very convenient. The problems appearing with chains of strongly pinned oscillators are very similar to those faced with chains of rotors, and so are the ideas involved to solve them.

The existence of an invariant measure for the chain of 3 rotors has been proved in [13] (reproduced here in Chapter 3), as well as a stretched exponential upper bound of the kind $e^{-c\sqrt{t}}$ on the convergence rate. The methods, which involve averaging the rapid oscillations of the central rotor, are inspired by those of [34] for the chain of 3 oscillators.

In the present paper, we generalize the result of [13] to the case of 4 rotors, and obtain again a bound $e^{-c\sqrt{t}}$ on the convergence rate. The main new difficulty in this generalization is the presence of resonances among the two central rotors. When they both have a large energy, there are two fast variables and some resonant terms make the averaging technique of [13] insufficient. A large portion of the present paper is devoted to eliminating such resonant terms by using the rapid thermalization of the external rotors.

It would be of course desirable to be able to work with a larger number of rotors. The present paper uses explicit methods to deal with the averaging phenomena. We hope that by crystallizing the essentials of our methods, longer chains can be handled in the same spirit. We expect that for longer chains, the convergence rate is of the form $\exp(-ct^k)$, for some exponent $k \in (0, 1)$ which depends on the length of the chain. We formulate a conjecture and explain the main difficulties for longer chains in Remark 4.5.3.

We now introduce the model and state the main results. In §4.2, we study the behavior of the system when one of the two central rotors is fast, and construct a Lyapunov function in this region. In §4.3, we do the same in the regime where both central rotors are fast. In §4.4 we construct a Lyapunov function that is valid across all regimes, and in §4.5 we provide the technicalities necessary to obtain the main result.

4.1.1. The model

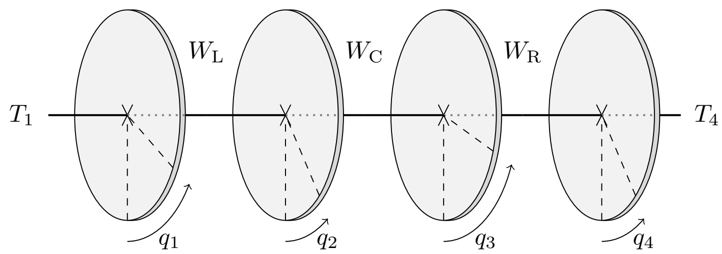


Figure 4.1 – A chain of four rotors with two heat baths at temperatures T_1 and T_4.

We study a model of 4 rotors, each given by a momentum $p_i \in \mathbb{R}$ and an angle $q_i \in \mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$, $i = 1, \dots, 4$. We write in the sequel $q = (q_1, \dots, q_4) \in \mathbb{T}^4$, $p = (p_1, \dots, p_4) \in \mathbb{R}^4$, and $x = (q, p) \in \Omega$, where $\Omega = \mathbb{T}^4 \times \mathbb{R}^4$ is the phase space of the system. We consider the Hamiltonian

$$H(x) = \sum_{i=1}^{4} \frac{p_i^2}{2} + W_\mathrm{L}(q_2 - q_1) + W_\mathrm{C}(q_3 - q_2) + W_\mathrm{R}(q_3 - q_4) \,, \tag{4.1.1}$$

where $W_\mathrm{I} : \mathbb{T} \to \mathbb{R}$, $\mathrm{I} = \mathrm{L}, \mathrm{C}, \mathrm{R}$ (standing for left, center and right) are smooth 2π-periodic interaction potentials (see Figure 3.1).

Convention: Unless specified otherwise, the arguments of the potentials are always as above, namely $W_\mathrm{L} = W_\mathrm{L}(q_2 - q_1)$, $W_\mathrm{C} = W_\mathrm{C}(q_3 - q_2)$ and $W_\mathrm{R} = W_\mathrm{R}(q_3 - q_4)$. The same applies to any function with index L, C and R. Note that the argument for R is $q_3 - q_4$ (and not $q_4 - q_3$) since this choice will lead to more symmetrical expressions between the sites 1 and 4.

To model the interaction with two heat baths, we add at each end of the chain a Langevin thermostat at temperature $T_b > 0$, with dissipation constant $\gamma_b > 0$, $b = 1, 4$. Introducing the derivative of the potentials $w_\mathrm{I} = W_\mathrm{I}'$, $\mathrm{I} = \mathrm{L}, \mathrm{C}, \mathrm{R}$, the main object of our study is the SDE:

$$\begin{aligned}
\mathrm{d}q_i &= p_i \,\mathrm{d}t \,, \qquad i = 1, \dots, 4 \,, \\
\mathrm{d}p_1 &= (w_\mathrm{L} - \gamma_1 p_1) \,\mathrm{d}t + \sqrt{2\gamma_1 T_1} \mathrm{d}B_t^1 \,, \\
\mathrm{d}p_2 &= (w_\mathrm{C} - w_\mathrm{L}) \,\mathrm{d}t \,, \\
\mathrm{d}p_3 &= -(w_\mathrm{C} + w_\mathrm{R}) \,\mathrm{d}t \,, \\
\mathrm{d}p_4 &= (w_\mathrm{R} - \gamma_4 p_4) \,\mathrm{d}t + \sqrt{2\gamma_4 T_4} \mathrm{d}B_t^4 \,,
\end{aligned} \tag{4.1.2}$$

where B_t^1, B_t^4 are independent standard Brownian motions. The generator of the semigroup associated to (4.1.2) reads

$$
\begin{aligned}
L = \sum_{i=1}^{4} p_i \partial_{q_i} + w_{\mathrm{L}} \cdot (\partial_{p_1} - \partial_{p_2}) + w_{\mathrm{C}} \cdot (\partial_{p_2} - \partial_{p_3}) + w_{\mathrm{R}} \cdot (\partial_{p_4} - \partial_{p_3}) \\
+ \sum_{b=1,4} \left(-\gamma_b p_b \partial_{p_b} + \gamma_b T_b \partial_{p_b}^2\right) \ .
\end{aligned} \tag{4.1.3}
$$

Remark 4.1.1. In contrast to Chapter 3, we do not allow for the presence of pinning potentials $U(q_i)$ and of constant forces at the ends of the chain, although we believe that the main result still holds with such modifications. While constant forces would be easy to handle, the addition of a pinning potential would require some generalization of a technical result (Proposition 4.3.12) which we are currently unable to provide (see Remark 4.3.13).

We consider the measure space $(\Omega, \mathcal{B})$, with the Borel σ-field $\mathcal{B}$ over Ω. The coefficients in (4.1.2) are globally Lipschitz, and therefore the solutions are almost surely defined for all times and all initial conditions. We denote the transition probability of the corresponding Markov process by $P^t(x, \cdot)$, for all $x \in \Omega$ and $t \geq 0$.

4.1.2. Main results

We will often refer to the sites 1 and 4 as the *outer* (or *external*) rotors, and the sites 2 and 3 as the *central* rotors. We require the interactions from the inner rotors to the outer rotors to be non-degenerate in the following sense:

Assumption 4.1.2. *We assume that for* $\mathrm{I} = \mathrm{L}, \mathrm{R}$ *and for each* $s \in \mathbb{T}$*, at least one of the derivatives* $w_{\mathrm{I}}^{(k)}(s)$*,* $k \geq 1$ *is non-zero.*

This assumption is not very restrictive. In particular, it holds if all the potentials consist of finitely many nonconstant Fourier modes.

Our main result is a statement about the speed of convergence to a unique stationary state of the system (4.1.2). In order to state it, we introduce for each continuous function $f : \Omega \to (0, \infty)$ the norm $\|\cdot\|_f$ on the space of signed Borel measures on Ω:

$$
\|\nu\|_f = \sup_{|g| \leq f} \int_{\Omega} g \mathrm{d}\nu \ .
$$

If $f \equiv 1$, we retrieve the total variation norm.

Theorem 4.1.3 (NESS and rate of convergence). *Under Assumption 4.1.2, the Markov process defined by (4.1.2) satisfies:*

(i) The transition probabilities $P^t(x, \mathrm{d}y)$ *have a* $\mathcal{C}^\infty((0,\infty) \times \Omega \times \Omega)$ *density* $p_t(x, y)$*.*

(ii) There is a unique invariant measure π*, and it has a smooth density.*

(iii) *For all* $0 \leq \theta_1 < \min(1/T_1, 1/T_4)$ *and all* $\theta_2 > \theta_1$, *there exist constants* $C, \lambda > 0$ *such that for all* $x = (q_1, q_2, \dots, p_4) \in \Omega$ *and all* $t \geq 0$,

$$\|P^t(x, \cdot\,) - \pi\|_{e^{\theta_1 H}} \leq C e^{\theta_2 H(x)} e^{-\lambda t^{1/2}} . \tag{4.1.4}$$

At thermal equilibrium, namely when $T_1 = T_4 = 1/\beta$ for some $\beta > 0$, the invariant measure is the Gibbs distribution with density $e^{-\beta H(x)}/Z$, where Z is a normalization constant.

Theorem 4.1.3 will be proved in §4.5 with help of results of [16] and the existence of a Lyapunov function, the properties of which are stated in

Theorem 4.1.4 (Lyapunov function). *Let* $0 < \theta < \min(1/T_1, 1/T_4)$. *Under Assumption 4.1.2, there is a function* $V : \Omega \to [1, \infty)$ *satisfying:*

(i) *There are constants* $c_1 > 0$ *and* $a \in (0, 1)$ *such that*

$$1 + e^{\theta H} \leq V \leq c_1 (e^{|p_2|^a} + e^{|p_3|^a}) e^{\theta H} . \tag{4.1.5}$$

(ii) *There are a compact set* K *and constants* $c_2, c_3 > 0$ *such that*

$$LV \leq c_2 \mathbf{1}_K - \varphi(V) , \tag{4.1.6}$$

with $\varphi : [1, \infty) \to (0, \infty)$ *the increasing, concave function defined by*

$$\varphi(s) = \frac{c_3\, s}{2 + \log(s)} . \tag{4.1.7}$$

Most of the paper will be devoted to proving the existence of such a Lyapunov function.

Remark 4.1.5. We assume throughout that T_1 and T_4 are strictly positive. While the conclusions of Theorem 4.1.4 remain true for $T_1 = T_4 = 0$ (with any $\theta > 0$), part of the argument has to be changed in this case, as sketched in Remark 4.3.17. The positivity of the temperatures is, however, essential for Theorem 4.1.3; at zero temperature, the system is not irreducible, and none of the conclusions of Theorem 4.1.3 hold.

4.1.3. Overview of the dynamics

To gain some insight into the strategy of the proof, we illustrate some essential features of the dynamics (4.1.2). Since the exterior rotors (at sites 1 and 4) are directly damped by the $-\gamma_b p_b \,\mathrm{d}t$ terms in (4.1.2), we expect their energy to decrease rapidly with large probability. More specifically, for $b = 1, 4$, we find that Lp_b is equal to $-\gamma_b p_b$ plus some bounded terms, and thus we expect p_b to decay exponentially (in expectation value) when it is large. Therefore, the external rotors recover very fast from thermal fluctuations, and will not be hard to deal with.

On the other hand, the central rotors are not damped directly, and feel the dissipative terms of (4.1.2) only indirectly, by interacting with the outer rotors. The interesting issue appears when the energy of the system is very large and mostly concentrated in one or both of the central rotors. If most of the energy is at site 2 (meaning that $|p_2|$ is much larger than all other momenta), the corresponding rotor spins very rapidly, *i.e.,* q_2 moves very rapidly on $\mathbb{T}$. But then, the interaction forces $w_{\rm L}(q_2 - q_1)$ and $w_{\rm C}(q_3 - q_2)$ oscillate rapidly, which causes the site to essentially decouple from its neighbors. The same happens when most of the energy is at site 3, when $w_{\rm C}(q_3 - q_2)$ and $w_{\rm R}(q_3 - q_4)$ oscillate rapidly. And when both $|p_2|$ and $|p_3|$ are large and much larger than $|p_1|$ and $|p_4|$, the forces $w_{\rm L}(q_2 - q_1)$ and $w_{\rm R}(q_3 - q_4)$ are highly oscillatory, so that the central two rotors almost decouple from the outer ones (the force $w_{\rm C}$ might or might not oscillate depending on p_2 and p_3).

This asymptotic decoupling is the interesting feature of the model: in principle, if the central rotors do not recover sufficiently fast from thermal fluctuations, the energy of the chain could grow (in expectation value) without bounds. On the other hand, when their energy is large, the decoupling phenomenon should make the central rotors less affected by the fluctuations of the heat baths. Our results imply that both effects combine in a way that prevents overheating. See Remark 3.3.10 for a quantitative discussion of these two effects for a chain of three rotors. See also [33] for a clear exposition of the overheating problem in a related model.

Figure 4.2 illustrates the evolution[1] of the momenta at two different time scales, starting with $p(0) = (50, 20, 30, 40)$. The upper graph shows that indeed p_1 and p_4 decrease very fast, and the lower graph indicates that p_2 and p_3 remain large for a significantly longer time, but eventually also decrease. Since for this initial condition p_3 is larger than p_2, the force $w_{\rm R}$ oscillates faster than $w_{\rm L}$. Therefore, p_3 couples less effectively to the outer rotors (where the dissipation happens) than p_2, and hence p_3 decreases more slowly.

If one were to look at these trajectories for much longer times, one would eventually observe some fluctuations of arbitrary magnitude, followed by new recovery phases. But large fluctuations are very rare.

Since the system is rapidly driven to small p_1, p_4, it is really the dynamics of (p_2, p_3) that plays the most important role. We will often argue in terms of the 8-dimensional dynamics projected onto the p_2p_3-plane. We illustrate some trajectories in this plane for several initial conditions in Figure 4.3. To make the illustration readable, we used a very small temperature, so that the picture is dominated by the deterministic dynamics.

The typical trajectory is as follows. Starting with some large $|p_2|$ and $|p_3|$, the slower of the two central rotors is damped faster than the other, so that the projection drifts rapidly towards one of the axes. This leads to a regime where only one of the central rotors is fast, while the other is essentially thermalized. The energy in this fast rotor is gradually dissipated, so that the orbit follows the axis towards the origin.

The behavior that we observe in Figure 4.3 around the diagonal $p_2 = p_3$ far enough from the origin is easily explained: in the “center of mass frame” of the two central rotors, we simply see two interacting rotors that oscillate slowly in opposition, while being almost decoupled from the outer rotors. More precisely, introducing $Q = q_3 - q_2 \in \mathbb{T}$ and $P = p_3 - p_2 \in \mathbb{R}$, we see that (Q, P) acts approximately as a mathematical pendulum with potential $2W_{\rm C}$, plus some rapidly oscillating (and

[1] The numerical algorithm used in this paper is based on the one described in [37]. The time step is either 10^{-2} or 10^{-3} depending on the situation.

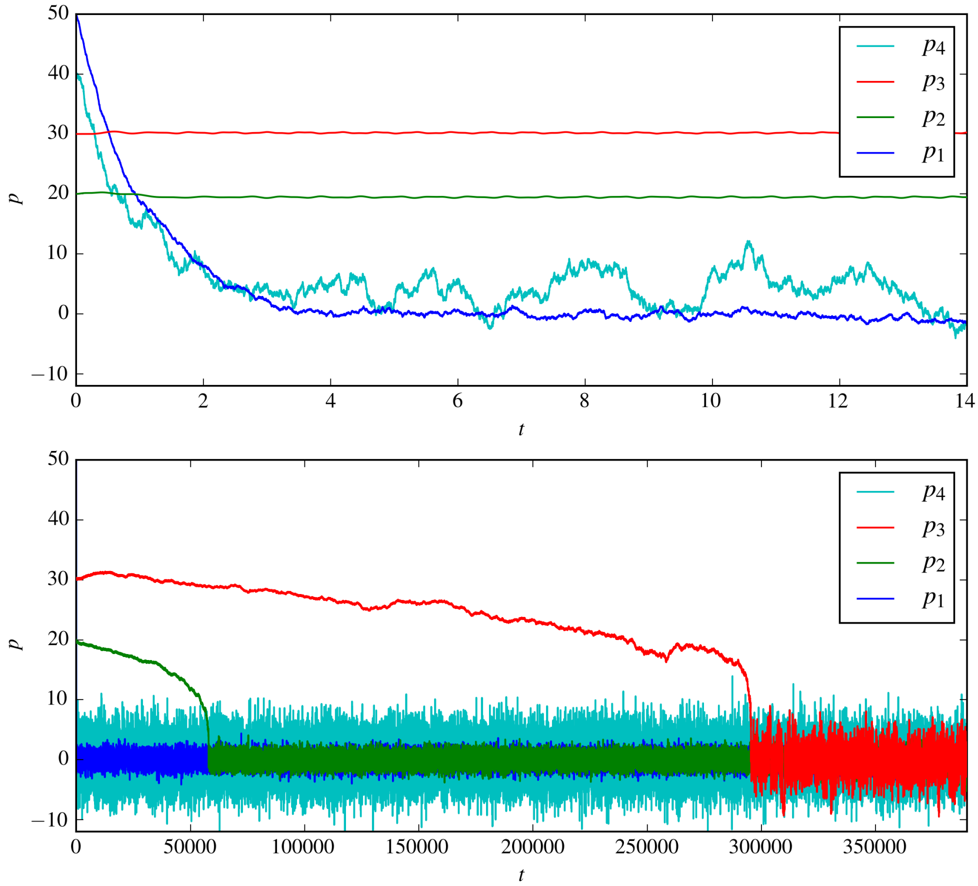


Figure 4.2 – Evolution of the momenta $p_1, \ldots, p_4$, for $\gamma_1 = \gamma_4 = 1$, $T_1 = 1$, $T_4 = 10$, $q(0) = (0, 0, 0, 0)$ and $p(0) = (50, 20, 30, 40)$. The interaction potentials in the simulations here are $W_{\mathrm{I}} = -\cos$ so that the forces are $w_{\mathrm{I}} = \sin$, $\mathrm{I} = \mathrm{L}, \mathrm{C}, \mathrm{R}$.

therefore weak) interactions with the outer rotors:

$$\dot{Q} = P, \qquad \dot{P} = -2w_{\mathrm{C}}(Q) + \text{weak interactions.}$$

Typically, if at first the energy in the center of mass frame is not large enough to make a "full turn," Q oscillates slowly around a minimum of W_{C}, which corresponds to a back-and-forth exchange of momentum between 2 and 3, and explains the strips that we observe around the diagonal. The two central rotors are then gradually slowed down, until at some point the interaction with the external rotors tears them apart.

The picture in the absence of noise (that is, when $T_1 = T_4 = 0$, which is not covered by our assumptions) is quite different, due to some resonances. We discuss their nature in Appendix 4.6. These resonances are washed away by the noise, and are therefore not visible here. They nevertheless play an important role in our computations, as we will see.

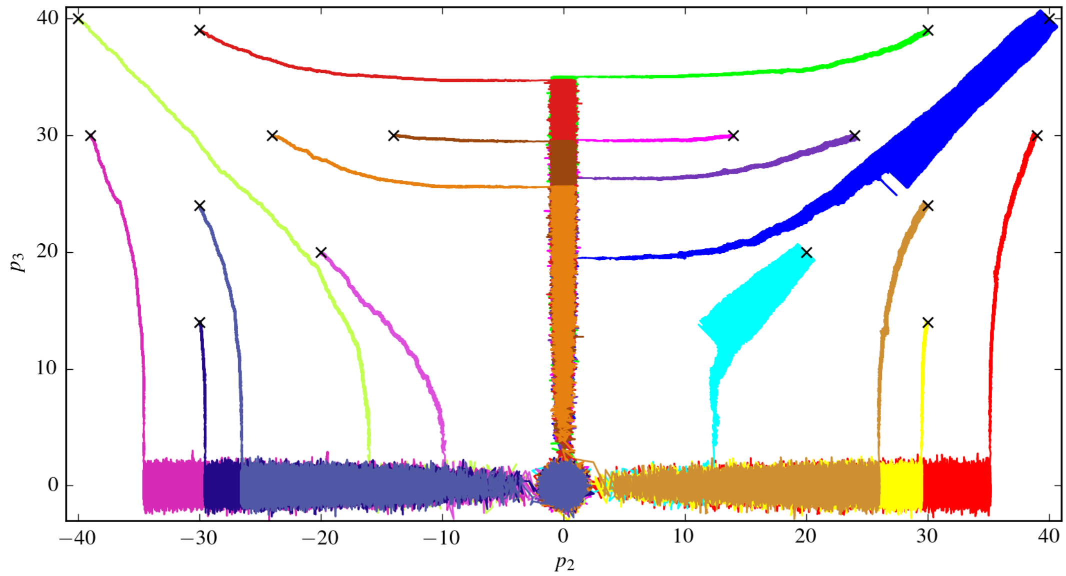


Figure 4.3 – The evolution of p_2 and p_3 for several initial conditions. The potentials are $W_{\mathrm{I}} = -\cos$, $\mathrm{I} = \mathrm{L}, \mathrm{C}, \mathrm{R}$. Furthermore, $\gamma_1 = \gamma_4 = 1$, $T_1 = 0.1$, and $T_4 = 0.4$. Each "×" sign indicates the beginning of a trajectory.

4.1.4. Strategy

In order to obtain rigorous results about the dynamics and construct a Lyapunov function, we will apply specific methods to each regime described above. We present them here in increasing order of difficulty.

- When a significant part of the energy is contained in the outer rotors, then as discussed above, the momenta of the two outer rotors essentially decrease exponentially fast. In this region, the Lyapunov function will be $e^{\theta H}$, and we will show that when $p_1^2 + p_4^2$ is large enough and $\theta < \min(1/T_1, 1/T_4)$, then $Le^{\theta H} \lesssim -e^{\theta H}$ (Lemma 4.4.1).

- When most of the energy is contained at just one of the central sites, namely at site $j = 2$ or $j = 3$, we will show that $Lp_j \sim -p_j^{-3}$ when averaged appropriately (Proposition 4.2.2). This corresponds to the neighborhood of the axes in Figure 4.3. This case is essentially treated as in Chapter 3. In this region, we use a Lyapunov function $V_j \sim e^{|p_j|^a + \frac{\theta}{2}p_j^2}$ (with $a \in (0,1)$) such that $LV_j \lesssim -V_j/p_j^2$ (Proposition 4.2.4).

- When both $|p_2|$ and $|p_3|$ are large and hold most of the energy, we do not approximate the dynamics of p_2 and p_3 separately, but we consider instead the "central" Hamiltonian $H_c = \frac{p_2^2}{2} + \frac{p_3^2}{2} + W_{\mathrm{L}} + W_{\mathrm{C}} + W_{\mathrm{R}}$. We show that when averaged properly, $LH_c \sim -p_2^{-2} - p_3^{-2}$ (Proposition 4.3.2). The Lyapunov function in this region is $V_c \sim H_c e^{\theta H_c}$, and we show (Proposition 4.3.5) that $LV_c \lesssim -V_c/H_c$. Showing that $LH_c \sim -p_2^{-2} - p_3^{-2}$ is the most difficult part of our proof. The averaging of the rapidly oscillating forces will prove to be insufficient due to some resonances, which manifest themselves for some rational values of p_3/p_2. We will consider separately the vicinity of the $p_2 = p_3$ diagonal, which is easy to deal with (Lemma 4.3.7), and the case where $|p_3 - p_2|$ is large, which requires substantially more work (§4.3.3). In the latter case, we will use the rapid thermalization of the external rotors in order to eliminate the resonant terms.

The factors $1/p_2^2$ and $1/H_c$ in $LV_j \lesssim -V_j/p_j^2$ and $LV_c \lesssim -V_c/H_c$ are the cause of the logarithmic contribution in (4.1.7), which leads to the subexponential convergence rate.

The final step (§4.4) is to combine $e^{\theta H}$, V_2, V_3 and V_c (which each behave nicely in a given regime) to obtain a Lyapunov function V that behaves nicely everywhere and satisfies the conclusions of Theorem 4.1.4.

4.1.5. The domains

Following the discussion above, we decompose Ω into several sub-regions. This decomposition only involves the momenta, and not the positions. All the sets in the decomposition are defined in the complement of a ball B_R of (large) radius R in p-space:

$$B_R = \mathbb{T}^4 \times \left\{p \in \mathbb{R}^4 : \sum_{i=1}^{4} p_i^2 \leq R^2\right\} .$$

For convenience, we consider only $R \geq \sqrt{2}$ (see Remark 4.1.6). We also use (large) integers k, ℓ, and m which will be fixed in §4.4, and we assume throughout that

$$1 \leq k < \ell < m . \tag{4.1.8}$$

The first regions we consider are along the p_2 and p_3 axes:

$$\begin{aligned} \Omega_2 = \Omega_2(k, R) &= \left\{x \in \Omega : p_2^2 > (p_1^2 + p_3^2 + p_4^2)^k\right\} \setminus B_R , \\ \Omega_3 = \Omega_3(k, R) &= \left\{x \in \Omega : p_3^2 > (p_1^2 + p_2^2 + p_4^2)^k\right\} \setminus B_R . \end{aligned} \tag{4.1.9}$$

The region Ω_2 (resp. Ω_3) corresponds to the configurations where most of the energy is concentrated at site 2 (resp. 3). The next region corresponds to the configurations where most of the energy is shared among the sites 2 and 3:

$$\Omega_c = \Omega_c(\ell, m, R) = \left\{x \in \Omega : p_2^2 + p_3^2 > (p_1^2 + p_4^2)^m,\ p_3^{2\ell} > p_2^2,\ p_2^{2\ell} > p_3^2\right\} \setminus B_R \tag{4.1.10}$$

(the conditions $p_3^{2\ell} > p_2^2$ and $p_2^{2\ell} > p_3^2$ ensure that both $|p_2|$ and $|p_3|$ diverge sufficiently fast when $\|p\| \to \infty$ in Ω_c). These regions are illustrated in Figure 4.4 and Figure 4.5. Note that $\Omega_2, \Omega_3, \Omega_c$ *do* intersect and *do not* cover Ω. However, for R large enough, the set $\Omega_2 \cup \Omega_3 \cup \Omega_c \cup B_R$ contains the p_2p_3-plane (more precisely, the product of $\mathbb{T}^4$ and some neighborhood of the p_2p_3-plane in momentum space), which is where the determining part of the dynamics lies, as discussed above.

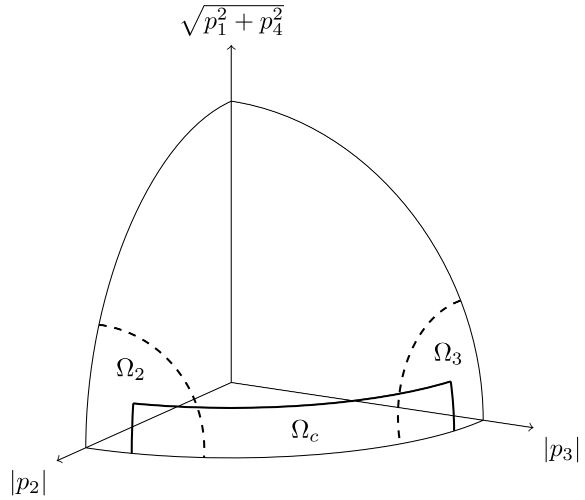


Figure 4.4 – A projection of the domains $\Omega_2, \Omega_3, \Omega_c$. The spherical surface represents $\sum_{i=1}^4 p_i^2 = C^2$ for some $C > R$.

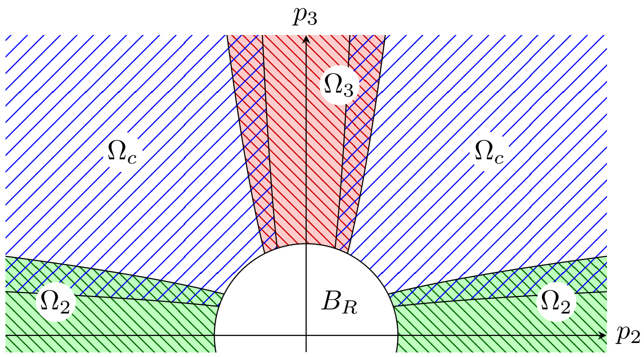


Figure 4.5 – The intersection of the sets $\Omega_2, \Omega_3, \Omega_c$ with the p_2p_3-plane (the lower half-plane is obtained by axial symmetry).

Remark 4.1.6. As a consequence of the restriction $R \geq \sqrt{2}$, we have $\Omega_j(k', R') \subset \Omega_j(k, R)$ for all $k' \geq k$, $R' \geq R$, and $j = 2, 3$. Therefore, if a bound holds for all $x \in \Omega_j(k, R)$, it also holds for all $x \in \Omega_j(k', R')$. Similarly, at fixed ℓ, we have $\Omega_c(\ell, m', R') \subset \Omega_c(\ell, m, R)$ for all $m' \geq m$ and $R' \geq R$. This allows us to increase k, m and R as needed (but not ℓ). We also observe immediately that for all k, ℓ, m, and for $j = 2, 3$,

$$\lim_{R\to\infty} \inf_{x\in\Omega_j(k,R)} |p_j| = \infty \,, \tag{4.1.11}$$

$$\lim_{R\to\infty} \inf_{x\in\Omega_c(\ell,m,R)} |p_j| = \infty \,. \tag{4.1.12}$$

4.1.6. Notations

Since averaging functions that rapidly oscillate in time will play an important role, we introduce the q_i-average $\langle f\rangle_i = \frac{1}{2\pi}\int_0^{2\pi} f \mathrm{d}q_i$ of a function $f : \Omega \to \mathbb{R}$ over one period of q_i. The result is a function of p and $\{q_j : j \neq i\}$. In the presence of a generic function $f : \mathbb{T} \to \mathbb{R}$ of one variable, we write simply $\langle f\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(s)\mathrm{d}s$, which is a constant.

For any function $f : \mathbb{T} \to \mathbb{R}$ satisfying $\langle f\rangle = 0$, one can find a unique integral $F : \mathbb{T} \to \mathbb{R}$ such that $F' = f$ and $\langle F\rangle = 0$. More generally, we write $f^{[j]}$ for the j^{th} integral of f that averages to zero.

Without loss of generality, we fix the additive constants of the potentials so that

$$\langle W_{\mathrm{I}}\rangle = 0, \qquad \mathrm{I} = \mathrm{L}, \mathrm{C}, \mathrm{R}\ . \tag{4.1.13}$$

We also introduce two "effective dissipation constants":

$$\alpha_2 = \gamma_1 \left\langle W_{\mathrm{L}}^2\right\rangle > 0\ , \qquad \alpha_3 = \gamma_4 \left\langle W_{\mathrm{R}}^2\right\rangle > 0\ , \tag{4.1.14}$$

where the positivity follows from Assumption 4.1.2. Note also that because of (4.1.13), there is no indeterminate additive constant in the α_j.

Finally, throughout the proofs, c denotes a generic positive constant that can be each time different. These constants are allowed to depend on the parameters and functions at hand, but *not* on the position x. We sometimes also use c' to emphasize that the constant has changed.

4.2. When only one of the central rotors is fast

We consider the regime where either $|p_2|$ or $|p_3|$ (but not both) is much larger than all other momenta. The estimates for this regime are simple adaptations from §3.3.4, but we recall here the main ideas.

We start with some formal computations, thinking in terms of powers of p_2 (resp. p_3) only. Then, we will restrict ourselves to the set $\Omega_2(k, R)$ (resp. $\Omega_3(k, R)$) for some large enough k and R, so that the other momenta are indeed "negligible" (see Lemma 4.2.3) compared to p_2 (resp. p_3).

4.2.1. Averaging with one fast variable

Assume that $|p_2|$ is much larger than the other momenta. We think in terms of the following fast-slow decomposition: the variables q_1, q_3, q_4 and p evolve slowly, while q_2 evolves rapidly, since $\dot{q}_2 = p_2$, and p_2 is large. In this regime, the variable q_2 swipes through $\mathbb{T}$ many times before any other variable changes significantly. The dynamics for short times is

$$\begin{aligned}
&p(t) \approx p(0)\ ,\\
&q_i(t) \approx q_i(0)\ , \qquad\qquad i = 1, 3, 4\ ,\\
&q_2(t) \approx q_2(0) + p_2(0)t \pmod{2\pi}\ .
\end{aligned} \tag{4.2.1}$$

We consider an observable $f : \Omega \to \mathbb{R}$ and let g be defined by

$$Lf = g\ . \tag{4.2.2}$$

Under the approximation (4.2.1), the quantity $g(x(t))$ oscillates very rapidly around its q_2-average $\langle g\rangle_2$, which is a function of the slow variables q_1, q_3, q_4 and p. We therefore expect the effective equation $Lf \approx \langle g\rangle_2$ to describe the evolution of f over several periods of oscillations, and we now show how to give a precise meaning to this approximation.

Although the stochastic terms (which appear as the second-order part of the differential operator L) appear in the computations, they do not play an important conceptual role in this discussion; the rapid oscillations that we average are of dynamical nature and are present regardless of the stochastic forcing exerted by the heat baths.

The generator of the dynamics (4.2.1) is simply

$$L_2 = p_2 \partial_{q_2} .$$

Decomposing the generator L defined in (4.1.3) as $L = L_2 + (L - L_2)$ and considering powers of p_2, we view L_2 as large, and $L - L_2$ as small. Note that for all smooth $h : \Omega \to \mathbb{R}$, we have $\langle L_2 h\rangle_2 = p_2 \langle \partial_{q_2} h\rangle_2 = 0$ by periodicity, so that the image of L_2 contains only functions with zero q_2-average. Consider next the indefinite integral $G = \int (g - \langle g\rangle_2)\mathrm{d}q_2$ (we choose the integration constant $C(q_1, q_3, q_4, p)$ to our convenience). By construction, we have $L_2(G/p_2) = g - \langle g\rangle_2$, so that

$$L\left(f - \frac{G}{p_2}\right) = \langle g\rangle_2 + (L_2 - L)\frac{G}{p_2} . \tag{4.2.3}$$

By subtracting the "small" counterterm G/p_2 from f, we have managed to replace g with its q_2-average in the right-hand side, plus some "small" correction. This procedure is what we refer to as averaging with respect to q_2, and it makes sense only in the regime where $|p_2|$ is very large. If $\langle g\rangle_2 = 0$ and $(L_2 - L)(G/p_2)$ is still oscillatory, the procedure must be repeated.

4.2.2. Application to the central momenta

We now apply this averaging method to the observable p_2, in the regime where $|p_2|$ is very large. By the definition of L, we find

$$Lp_2 = w_\mathrm{C} - w_\mathrm{L} . \tag{4.2.4}$$

We have $\langle w_\mathrm{C}\rangle_2 = \langle w_\mathrm{L}\rangle_2 = 0$. Moreover, $\partial_{q_2} W_\mathrm{C}(q_3 - q_2) = -w_\mathrm{C}(q_3 - q_2)$ and $\partial_{q_2} W_\mathrm{L}(q_2 - q_1) = w_\mathrm{L}(q_2 - q_1)$. Thus, in the notation above, $G = \int (w_\mathrm{C} - w_\mathrm{L})\mathrm{d}q_2 = -W_\mathrm{C} - W_\mathrm{L}$, and we introduce the new variable

$$p_2^{(1)} = p_2 - \frac{G}{p_2} = p_2 + \frac{W_\mathrm{C} + W_\mathrm{L}}{p_2} . \tag{4.2.5}$$

By (4.2.3), we obtain

$$\begin{aligned} Lp_2^{(1)} &= -(L_2 - L)\left(\frac{W_\mathrm{C} + W_\mathrm{L}}{p_2}\right) \\ &= \frac{p_3 w_\mathrm{C} - p_1 w_\mathrm{L}}{p_2} + \frac{W_\mathrm{C} w_\mathrm{L} - W_\mathrm{L} w_\mathrm{C} + W_\mathrm{L} w_\mathrm{L} - W_\mathrm{C} w_\mathrm{C}}{p_2^2} . \end{aligned} \tag{4.2.6}$$

Observe that the right-hand side of (4.2.6) is still oscillatory, but now with an amplitude of order $1/p_2$, which is much smaller than the amplitude of (4.2.4) when $|p_2|$ is large. Furthermore, the right-hand side of (4.2.6) has zero mean, since $\langle w_{\mathrm{C}}\rangle_2 = \langle w_{\mathrm{L}}\rangle_2 = 0$ and

$$\langle W_{\mathrm{C}}w_{\mathrm{L}} - W_{\mathrm{L}}w_{\mathrm{C}} + W_{\mathrm{L}}w_{\mathrm{L}} - W_{\mathrm{C}}w_{\mathrm{C}}\rangle_2 = \frac{1}{2}\left\langle \partial_{q_2}(W_{\mathrm{C}} + W_{\mathrm{L}})^2\right\rangle_2 = 0\;,$$

by periodicity. In order to see a net effect, we need to average again. We consider now the observable $f = p_2^{(1)}$, and apply the same procedure. Instead of averaging the right-hand side of (4.2.6) in one step, we first deal only with the terms of order -1 in p_2, by introducing

$$p_2^{(2)} = p_2^{(1)} + \frac{p_1 W_{\mathrm{L}} + p_3 W_{\mathrm{C}}}{p_2^2}\;. \tag{4.2.7}$$

We postpone further computations to the proof of Proposition 4.2.2 below, and explain here the main steps. We will see that $Lp_2^{(2)}$ consists of terms of order -2 and -3 (by construction, the contribution of order -1 disappears). The terms of order -2 have mean zero, and will be removed by introducing a new variable $p_2^{(3)}$. We will then find that $Lp_2^{(3)}$ contains terms of order -3 and -4. To replace the terms of order -3 with their average (which is finally non-zero), we will introduce a function $p_2^{(4)}$. This will complete the averaging procedure.

We illustrate in Figure 4.6 the time-dependence of $p_2, p_2^{(1)}$ and $p_2^{(2)}$ (slightly shifted for better readability)[2]. Clearly, the oscillations of $p_2^{(1)}$ are much smaller than those of p_2, and we barely perceive the oscillations of $p_2^{(2)}$, since they are smaller than the random fluctuations.

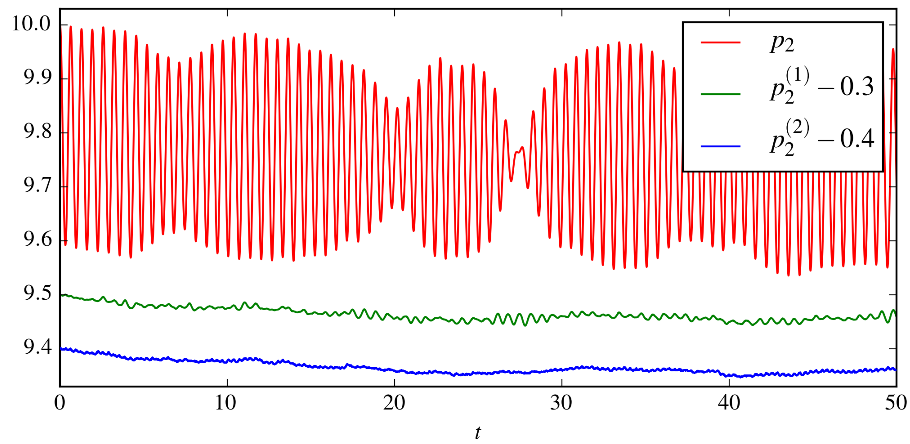


Figure 4.6 – The effect of the coordinate changes (4.2.5) and (4.2.7) on the effective oscillations. Note that two of the curves are shifted vertically for easier readability.

Before we state the result of this averaging process, we introduce a convenient notation for the

[2]The irregularity of the envelope of p_2 in Figure 4.6 is due to the randomness of the phases of the two oscillatory forces w_{L} and w_{C}: they sometimes add up, and sometimes compensate each other. Note also that the trajectory of $p_2^{(2)}$ is rougher than the other two, since the definition of $p_2^{(2)}$ involves p_1, which is directly affected by the stochastic force.

remainders.

Definition 4.2.1. *Let f, g be two functions defined on the set $\{x \in \Omega : p_2 \neq 0\}$. We say that f is $\mathcal{O}_2(g)$ if there is a polynomial z such that when $|p_2|$ is large enough,*

$$|f(x)| \leq z(p_1, p_3, p_4)|g(x)| \ . \tag{4.2.8}$$

The analogous notation $\mathcal{O}_3$ will be used when $|p_3|$ is large, and with a polynomial $z(p_1, p_2, p_4)$.

This notation reflects the fact that when most of the energy is at site 2 (resp. 3), one can forget about the dependence on p_1, p_3, p_4 (resp. p_1, p_2, p_4), provided that it is at most polynomial (by the compactness of $\mathbb{T}^4$, the position q is irrelevant). For example, the term $(p_1 W_{\mathrm{L}} + p_3 W_{\mathrm{C}})/p_2^2$ in (4.2.7) is $\mathcal{O}_2(p_2^{-2})$.

It is easy to realize that the $\mathcal{O}_j$, $j = 2, 3$, follow the same basic rules as the usual $\mathcal{O}$. In particular, $\mathcal{O}_j(g_1) + \mathcal{O}_j(g_2) = \mathcal{O}_j(|g_1| + |g_2|)$ and $\mathcal{O}_j(g_1)\mathcal{O}_j(g_2) = \mathcal{O}_j(g_1 g_2)$.

Proposition 4.2.2. *There are functions $\tilde{p}_2$ and $\tilde{p}_3$ of the form*

$$\tilde{p}_2 = p_2 + \frac{W_{\mathrm{L}} + W_{\mathrm{C}}}{p_2} + \frac{p_1 W_{\mathrm{L}} + p_3 W_{\mathrm{C}}}{p_2^2} + \mathcal{O}_2(p_2^{-3}) \ , \tag{4.2.9}$$

$$\tilde{p}_3 = p_3 + \frac{W_{\mathrm{R}} + W_{\mathrm{C}}}{p_3} + \frac{p_4 W_{\mathrm{R}} + p_2 W_{\mathrm{C}}}{p_3^2} + \mathcal{O}_3(p_3^{-3}) \ , \tag{4.2.10}$$

such that for $j = 2, 3$,

$$L\tilde{p}_j = -\alpha_j p_j^{-3} + \mathcal{O}_j(p_j^{-4}) \ , \tag{4.2.11}$$

where $\alpha_2 > 0$, $\alpha_3 > 0$ are defined in (4.1.14). Furthermore,

$$\begin{aligned} \partial_{p_1}\tilde{p}_2 &= \frac{W_{\mathrm{L}}}{p_2^2} + \mathcal{O}_2(p_2^{-3}) \ , & \partial_{p_4}\tilde{p}_2 &= \mathcal{O}_2(p_2^{-4}) \ , \\ \partial_{p_1}\tilde{p}_3 &= \mathcal{O}_3(p_3^{-4}) \ , & \partial_{p_4}\tilde{p}_3 &= \frac{W_{\mathrm{R}}}{p_3^2} + \mathcal{O}_3(p_3^{-3}) \ . \end{aligned} \tag{4.2.12}$$

Proof. It suffices to consider the case $j = 2$. The variable $\tilde{p}_2$ is constructed as in §3.3.4. We continue the averaging procedure started above. It is easy to check that $Lp_2^{(2)}$ can be written as

$$Lp_2^{(2)} = \frac{1}{p_2^2}\partial_{q_2} R_1 + \frac{2p_1\left(W_{\mathrm{L}} w_{\mathrm{L}} - W_{\mathrm{L}} w_{\mathrm{C}}\right) + 2p_3\left(W_{\mathrm{C}} w_{\mathrm{L}} - W_{\mathrm{C}} w_{\mathrm{C}}\right)}{p_2^3} \ ,$$

with

$$R_1 = -p_1^2 W_{\mathrm{L}} - p_3^2 W_{\mathrm{C}} - \gamma_1 p_1 W_{\mathrm{L}}^{[1]} + W_{\mathrm{L}}^2 + W_{\mathrm{L}} W_{\mathrm{C}} + W_{\mathrm{C}}^2 + W_{\mathrm{C}}^{[1]} w_{\mathrm{R}} \ .$$

Since it is a total derivative, the term of order -2 has zero q_2-average, and by introducing $p_2^{(3)} =$

$p_2^{(2)} - \frac{R_1}{p_2^3}$, we find

$$Lp_2^{(3)} = \frac{\partial_{q_2} R_2 + w_\mathrm{L} W_\mathrm{L}^{[1]}\gamma_1 + p_1\left(W_\mathrm{C} w_\mathrm{L} - 2W_\mathrm{L} w_\mathrm{C}\right) + p_3\left(2W_\mathrm{C} w_\mathrm{L} - W_\mathrm{L} w_\mathrm{C}\right)}{p_2^3} + \mathcal{O}_2(p_2^{-4}) \ , \tag{4.2.13}$$

with

$$\begin{aligned} R_2 = &-p_1^3 W_\mathrm{L} - p_3^3 W_\mathrm{C} - 3\gamma_1 p_1^2 W_\mathrm{L}^{[1]} - \gamma_1^2 p_1 W_\mathrm{L}^{[2]} + 3p_1 W_\mathrm{L}^2 \\ &+ 2\gamma_1 T_1 W_\mathrm{L}^{[1]} + (p_3 - p_4) w_\mathrm{R}' W_\mathrm{C}^{[2]} + 3p_3 W_\mathrm{C}^2 + 3p_3 W_\mathrm{C}^{[1]} w_\mathrm{R} \ . \end{aligned}$$

One can then average the terms of order -3 in (4.2.13). We have again $\langle \partial_{q_2} R_2 \rangle_2 = 0$ by periodicity, and after integration by parts we find

$$\langle W_\mathrm{C} w_\mathrm{L} \rangle_2 = \langle w_\mathrm{C} W_\mathrm{L} \rangle_2 \quad \text{and} \quad \left\langle \gamma_1 w_\mathrm{L} W_\mathrm{L}^{[1]} \right\rangle_2 = -\gamma_1 \left\langle W_\mathrm{L}^2 \right\rangle_2 = -\alpha_2$$

(for the signs, recall that $W_\mathrm{L} = W_\mathrm{L}(q_2 - q_1)$ and $W_\mathrm{C} = W_\mathrm{C}(q_3 - q_2)$). By adding appropriate counterterms (not written explicitly), we obtain a function $p_2^{(4)} = p_2^{(3)} + \mathcal{O}_2(p_2^{-4})$ such that

$$Lp_2^{(4)} = -\frac{\alpha_2}{p_2^3} + \frac{\langle p_3 W_\mathrm{L} w_\mathrm{C} - p_1 W_\mathrm{C} w_\mathrm{L} \rangle_2}{p_2^3} + \mathcal{O}_2(p_2^{-4}) \ .$$

The first term in the right-hand side is the one we are looking for, and we deal with the other term of order -3 (which is non-zero) as follows. We observe that

$$\langle p_3 W_\mathrm{L} w_\mathrm{C} - p_1 W_\mathrm{C} w_\mathrm{L} \rangle_2 = (p_1 \partial_{q_1} + p_3 \partial_{q_3}) \langle W_\mathrm{L} W_\mathrm{C} \rangle_2 = L \langle W_\mathrm{L} W_\mathrm{C} \rangle_2 \ ,$$

since $\langle W_\mathrm{L} W_\mathrm{C} \rangle_2$ is a function of q_1, q_3 only. We then set

$$\tilde{p}_2 = p_2^{(4)} - \frac{\langle W_\mathrm{L} W_\mathrm{C} \rangle_2}{p_2^3}$$

and obtain (4.2.11). It is immediate by the construction of $\tilde{p}_2$ that (4.2.12) holds. □

We now introduce a lemma, which says that remainders of the kind $\mathcal{O}_j(|p_j|^{-r})$, $j = 2, 3$, can be made very small on $\Omega_j(k, R)$, provided that the parameters k, R are large enough.

Lemma 4.2.3. *Let $j \in \{2, 3\}$ and $r > 0$. Fix $\varepsilon > 0$ and a function $f = \mathcal{O}_j(|p_j|^{-r})$. Then, for all sufficiently large k and R, we have*

$$\sup_{x \in \Omega_j(k,R)} |f(x)| \le \varepsilon \ .$$

Proof. We prove the result for $j = 2$. By Definition 4.2.1 and (4.1.11), there is a polynomial z such that for all large enough R and all k, we have $|f| \le z(p_1, p_3, p_4)|p_2|^{-r}$ on $\Omega_2(k, R)$. But then, we have on the same set

$$|f| \leq \frac{z(p_1,p_3,p_4)}{|p_2|^r} \leq \frac{c+c(p_1^2+p_3^2+p_4^2)^N}{|p_2|^r} \leq \frac{c+c|p_2|^{\frac{2N}{k}}}{|p_2|^r} \leq c|p_2|^{\frac{2N}{k}-r} \ ,$$

where the second inequality is immediate for sufficiently large N, the third inequality comes from the definition of Ω_2, and the fourth inequality holds because $|p_2|$ is bounded away from zero on $\Omega_2(k,R)$. Recalling (4.1.11), we obtain the desired result when k is large enough so that $\frac{2N}{k} - r < -\frac{r}{2}$. □

We now construct partial Lyapunov functions in the regions Ω_2 and Ω_3.

Proposition 4.2.4. *Let $0 < \theta < \min(1/T_1, 1/T_4)$ and $a \in (0,1)$. Consider the functions*[3]

$$\begin{aligned} V_2 &= e^{|\tilde p_2|^a + \frac{\theta}{2}\tilde p_2^2}\left(1 + F_2(q_2-q_1)/p_2^3\right) \ , \\ V_3 &= e^{|\tilde p_3|^a + \frac{\theta}{2}\tilde p_3^2}\left(1 + F_3(q_3-q_4)/p_3^3\right) \ , \end{aligned} \tag{4.2.14}$$

with the $\tilde p_j$ of Proposition 4.2.2, and $F_2, F_3 : \mathbb{T} \to \mathbb{R}$ such that respectively $F_2'(s) = \theta^2\gamma_1 T_1(\langle W_{\rm L}^2\rangle - W_{\rm L}^2(s))$ and $F_3'(s) = \theta^2\gamma_4 T_4(\langle W_{\rm R}^2\rangle - W_{\rm R}^2(s))$. Then, there are constants $C_1, C_2, C_3 > 0$, independent of $a \in (0,1)$, such that for all sufficiently large k and R, we have for $j = 2,3$ the following inequalities on Ω_j:

$$C_1 e^{|p_j|^a + \frac{\theta}{2}p_j^2} < V_j < C_2 e^{|p_j|^a + \frac{\theta}{2}p_j^2} \ , \tag{4.2.15}$$

$$LV_j \leq -C_3 p_j^{-2} e^{|p_j|^a + \frac{\theta}{2}p_j^2} \ . \tag{4.2.16}$$

Proof. By symmetry, it suffices to prove the result for $j = 2$. In this proof, we do not allow the $\mathcal{O}_2$ to depend on $a \in (0,1)$ (that is, we want the bound (4.2.8) to hold uniformly in $a \in (0,1)$). We start by proving (4.2.15). For large enough R, we have that $|p_2| > 2$ on Ω_2. Moreover, since $\tilde p_2 = p_2 + \mathcal{O}_2(p_2^{-1})$ and $F_2(q_2-q_1)/p_2^3 = \mathcal{O}_2(p_2^{-3})$, we have by Lemma 4.2.3 that for large enough k, R, it holds on Ω_2 that

$$|\tilde p_2| > 1 \quad \text{and} \quad \left|\frac{F_2(q_2-q_1)}{p_2^3}\right| < \frac{1}{2} \ . \tag{4.2.17}$$

Moreover, since both $|\tilde p_2|$ and $|p_2|$ are > 1, (4.2.9) implies, for all $a \in (0,1)$,

$$\left||\tilde p_2|^a - |p_2|^a\right| \leq |\tilde p_2^2 - p_2^2| = \left|2(W_{\rm L} + W_{\rm C}) + \mathcal{O}_2(p_2^{-1})\right| \ .$$

Since $W_{\rm L}$ and $W_{\rm C}$ are bounded, it follows from Lemma 4.2.3 that we can bound the right-hand side by a constant, so that we find

$$c e^{|p_2|^a + \frac{\theta}{2}p_2^2} < e^{|\tilde p_2|^a + \frac{\theta}{2}\tilde p_2^2} < c' e^{|p_2|^a + \frac{\theta}{2}p_2^2} \ , \tag{4.2.18}$$

[3] The role of the contribution $|\tilde p_j|^a$ is to facilitate the patchwork that will lead to a global Lyapunov function in §4.4. The corrections involving F_2 and F_3 help average some $W_{\rm L}^2$ and $W_{\rm R}^2$ that appear in the computations. Without this correction, we would need a condition on θ that is more restrictive than the natural condition $\theta < \min(1/T_1, 1/T_4)$.

uniformly in a. By this, by the definition of V_2, and by (4.2.17), we obtain (4.2.15). We now prove (4.2.16). Let $f(s) = e^{|s|^a + \frac{\theta}{2}s^2}$ and note that

$$L\big(e^{|\tilde{p}_2|^a + \frac{\theta}{2}\tilde{p}_2^2}\big) = Lf(\tilde{p}_2) = f'(\tilde{p}_2)L\tilde{p}_2 + f''(\tilde{p}_2)\sum_{b=1,4}\gamma_b T_b(\partial_{p_b}\tilde{p}_2)^2 \,. \tag{4.2.19}$$

By Proposition 4.2.2, we have on Ω_2 that

$$\begin{aligned} f'(\tilde{p}_2)L\tilde{p}_2 &= e^{|\tilde{p}_2|^a + \frac{\theta}{2}\tilde{p}_2^2}\left(\frac{a|\tilde{p}_2|^a}{\tilde{p}_2} + \theta\tilde{p}_2\right)\left(-\alpha_2 p_2^{-3} + \mathcal{O}_2(p_2^{-4})\right) \\ &= e^{|\tilde{p}_2|^a + \frac{\theta}{2}\tilde{p}_2^2}\left(-\alpha_2\theta p_2^{-2} + \mathcal{O}_2(p_2^{-3})\right) \ , \end{aligned} \tag{4.2.20}$$

where we have used that $\tilde{p}_2 = p_2 + \mathcal{O}_2(p_2^{-1})$, and that $a|\tilde{p}_2|^{a-1} < 1$ (since $|\tilde{p}_2| > 1$), so that on Ω_2, the $\mathcal{O}_2(p_2^{-3})$ obtained is indeed uniform in a. Next, one can verify that uniformly in $a \in (0,1)$ and $|\tilde{p}_2| > 1$,

$$f''(\tilde{p}_2) \le f(\tilde{p}_2)\left(\theta^2\tilde{p}_2^2 + 2\theta|\tilde{p}_2| + c\right) \ .$$

Moreover, by (4.2.12) we have $\sum_{b=1,4}\gamma_b T_b(\partial_{p_b}\tilde{p}_2)^2 = \gamma_1 T_1 W_{\mathrm{L}}^2/p_2^4 + \mathcal{O}_2(p_2^{-5})$, so that on Ω_2,

$$\begin{aligned} f''(\tilde{p}_2)\sum_{b=1,4}\gamma_b T_b(\partial_{p_b}\tilde{p}_2)^2 &\le e^{|\tilde{p}_2|^a + \frac{\theta}{2}\tilde{p}_2^2}\left(\theta^2\tilde{p}_2^2 + 2\theta|\tilde{p}_2| + c\right)\left(\gamma_1 T_1\frac{W_{\mathrm{L}}^2}{p_2^4} + \mathcal{O}_2(p_2^{-5})\right) \\ &= e^{|\tilde{p}_2|^a + \frac{\theta}{2}\tilde{p}_2^2}\left(\theta^2\gamma_1 T_1\frac{W_{\mathrm{L}}^2}{p_2^2} + \mathcal{O}_2(p_2^{-3})\right) \ . \end{aligned} \tag{4.2.21}$$

Therefore, by (4.2.19), (4.2.20) and (4.2.21),

$$L\big(e^{|\tilde{p}_2|^a + \frac{\theta}{2}\tilde{p}_2^2}\big) \le \frac{1}{p_2^2}e^{|\tilde{p}_2|^a + \frac{\theta}{2}\tilde{p}_2^2}\left(-\alpha_2\theta + \theta^2\gamma_1 T_1 W_{\mathrm{L}}^2 + \mathcal{O}_2(p_2^{-1})\right) \ .$$

But then

$$\begin{aligned} LV_2 &= \left(1 + \frac{F_2(q_2 - q_1)}{p_2^3}\right)L\left(e^{|\tilde{p}_2|^a + \frac{\theta}{2}\tilde{p}_2^2}\right) + e^{|\tilde{p}_2|^a + \frac{\theta}{2}\tilde{p}_2^2}L\left(\frac{F_2(q_2 - q_1)}{p_2^3}\right) \\ &\le \left(1 + \frac{F_2(q_2 - q_1)}{p_2^3}\right)\frac{1}{p_2^2}e^{|\tilde{p}_2|^a + \frac{\theta}{2}\tilde{p}_2^2}\left(-\alpha_2\theta + \theta^2\gamma_1 T_1 W_{\mathrm{L}}^2 + \mathcal{O}_2(p_2^{-1})\right) \\ &\quad + e^{|\tilde{p}_2|^a + \frac{\theta}{2}\tilde{p}_2^2}\left(\frac{\theta^2\gamma_1 T_1\big(\langle W_{\mathrm{L}}^2\rangle_2 - W_{\mathrm{L}}^2\big)}{p_2^2} + \mathcal{O}_2(p_2^{-3})\right) \\ &= \frac{1}{p_2^2}e^{|\tilde{p}_2|^a + \frac{\theta}{2}\tilde{p}_2^2}\left(-\alpha_2\theta + \theta^2\gamma_1 T_1\langle W_{\mathrm{L}}^2\rangle + \mathcal{O}_2(p_2^{-1})\right) \ . \end{aligned}$$

Using the definition of α_2 in (4.1.14) and the condition on θ, we find that $-\alpha_2\theta + \theta^2\gamma_1 T_1\langle W_{\mathrm{L}}^2\rangle$ is negative. Using then Lemma 4.2.3 to make the $\mathcal{O}_2(p_2^{-1})$ very small, and combining the result with (4.2.18) completes the proof. $\square$

4.3. When both central rotors are fast

We now study the regime where *both* $|p_2|$ and $|p_3|$ are large (not necessarily of the same order of magnitude), and $|p_1|$ and $|p_4|$ are much smaller. We then have two fast variables: q_2 and q_3. As we will see, this will lead to some trouble related to resonances, and averaging the rapid oscillations will not be enough. We start with some formal computations thinking in terms of powers of p_2 and p_3, and then restrict ourselves to the set $\Omega_c(\ell, m, R)$ for some appropriate parameters.

4.3.1. Averaging with two fast variables: resonances

Now the fast-slow decomposition is as follows: q_1, q_4 and p are the slow variables, and q_2, q_3 are the fast variables, with the approximate dynamics (for short times)

$$\begin{aligned} p(t) &\approx p(0)\ , \\ q_i(t) &\approx q_i(0)\ , & i &= 1,4\ , \\ q_2(t) &\approx q_2(0) + p_2(0)t & &\pmod{2\pi}\ , \\ q_3(t) &\approx q_3(0) + p_3(0)t & &\pmod{2\pi}\ , \end{aligned} \tag{4.3.1}$$

generated by $L_2 + L_3 = p_2\partial_{q_2} + p_3\partial_{q_3}$, which we see as the most important contribution in L. Let again $f, g : \Omega \to \mathbb{R}$ and assume that

$$Lf = g\ .$$

We would like, as above, to add a correction to f in the left-hand side in order to replace g with its average in the right-hand side. However, since the fast motion of (q_2, q_3) on $\mathbb{T}^2$ (in the dynamics (4.3.1)) follows orbits that are open or closed depending on whether p_2 and p_3 are commensurable or not, there seems to be no natural notion of "average of g" that is continuous with respect to the slow variables.

Consider for example $g(x) = \sin(2q_2 - q_3)$. In our approximation, $\sin(2q_2(t) - q_3(t))$ oscillates with frequency $(2p_2 - p_3)/2\pi$. The average is zero when $p_3 \neq 2p_2$, and $\sin(2q_2(t) - q_3(t))$ remains constant when $p_3 = 2p_2$. When p_3 is close to $2p_2$, the oscillations are slow, and one cannot simply average $\sin(2q_2(t) - q_3(t))$. More generally, any smooth function g on Ω can be written as $\sum_{n,m\in\mathbb{Z}} a_{n,m} \sin(nq_2 + mq_3 + \varphi_{n,m})$ for some coefficients $a_{n,m}$ and $\varphi_{n,m}$ which depend on the slow variables q_1, q_4 and p. Each such term gives rise to problems close to the line $p_3/p_2 = -n/m$ in the p_2p_3-plane.

However, if g depends on q_2 but not on q_3, then no problem appears. In the approximation (4.3.1), the quantity $g(x(t))$ then oscillates rapidly around $\langle g\rangle_2$, which is then a function of the slow variables q_1, q_4 and p. Then, as in §4.2.1, we use $G = \int(g - \langle g\rangle_2)\mathrm{d}q_2$ (we choose the integration constant independent of q_3), so that $(L_2 + L_3)(G/p_2) = L_2(G/p_2) = g - \langle g\rangle_2$. Thus, $L(f - G/p_2) = \langle g\rangle_2 + (L - L_2 - L_3)(f - G/p_2)$, which has the desired form. Similarly, if g depends on q_3 but not on q_2, we use the counterterm G/p_3 with $G = \int(g - \langle g\rangle_3)\mathrm{d}q_3$. And of course, if g can be decomposed as the sum of a function not involving q_3 and a function not involving q_2, then we can average each part separately and sum the two counterterms.

It turns out that we will mostly encounter terms that depend only on one of the fast variables, and are therefore easy to average. We will go as far as possible averaging such terms, and then introduce a method to deal with the resonant terms (involving both q_2 and q_3) that appear.

4.3.2. Application to the central energy

As a starting point, we use the central energy

$$H_c = \frac{p_2^2}{2} + \frac{p_3^2}{2} + W_{\rm L} + W_{\rm C} + W_{\rm R} \; .$$

Definition 4.3.1. *Let $A_* \equiv \{x \in \Omega : p_2 \neq 0, p_3 \neq 0\}$ and let f, g be two functions defined on a set $A \subset A_*$. We say that f is $\mathcal{O}_c(g)$ (on the set A) if there is a polynomial z such that for all $x \in A$ with $\min(|p_2|, |p_3|)$ large enough, we have*

$$|f(x)| \leq z(p_1, p_4)|g(x)| \; .$$

Unless explicitly stated otherwise, we take $A = A_*$.

We state the main result of this section.

Proposition 4.3.2. *There is a function of the form*

$$\widetilde{H}_c = H_c + \frac{p_1 W_{\rm L}}{p_2} + \frac{p_4 W_{\rm R}}{p_3} + \mathcal{O}_c(|p_2|^{-2} + |p_3|^{-2}) \; , \tag{4.3.2}$$

such that

$$L\widetilde{H}_c = -\frac{\alpha_2}{p_2^2} - \frac{\alpha_3}{p_3^2} + \mathcal{O}_c(|p_2|^{-5/2} + |p_3|^{-5/2}) \; , \tag{4.3.3}$$

with α_j as defined in (4.1.14). Furthermore,

$$\partial_{p_1}\widetilde{H}_c = \frac{W_{\rm L}}{p_2} + \mathcal{O}_c(p_2^{-2}) \; , \qquad \partial_{p_4}\widetilde{H}_c = \frac{W_{\rm R}}{p_3} + \mathcal{O}_c(p_3^{-2}) \; . \tag{4.3.4}$$

In order to reduce the length of some symmetric formulae, we use the notation "$+ \Leftrightarrow$" as a shorthand for the other half of the terms with the indices exchanged as follows: $1 \Leftrightarrow 4$, $2 \Leftrightarrow 3$, $\rm L \Leftrightarrow R$, and the sign of $w_{\rm C}$ changed (due to the asymmetry of the argument $q_3 - q_2$ of $W_{\rm C}$).

In order to prepare the proof of Proposition 4.3.2, we proceed as follows. We first see that

$$L(H_c) = -p_1 w_{\rm L} - p_4 w_{\rm R} \; .$$

Since $w_{\rm L}$ does not involve q_3 and $w_{\rm R}$ does not involve q_2, it easy to find appropriate counterterms: we introduce

$$H_c^{(1)} = H_c + \frac{p_1 W_{\rm L}}{p_2} + \frac{p_4 W_{\rm R}}{p_3} \; , \tag{4.3.5}$$

and obtain

$$LH_c^{(1)} = \frac{-\gamma_1 p_1 W_\mathrm{L} - p_1^2 w_\mathrm{L} + w_\mathrm{L} W_\mathrm{L}}{p_2} + \frac{p_1 W_\mathrm{L}(w_\mathrm{L} - w_\mathrm{C})}{p_2^2} + \Leftrightarrow .$$

The terms of order $1/p_2$ do not depend on q_3 and have mean zero with respect to q_2 (again $w_\mathrm{L} W_\mathrm{L} = \partial_{q_2} W_\mathrm{L}^2/2$ has zero q_2-average by periodicity). Similarly, the terms in $1/p_3$ do not involve q_2 and average to zero with respect to q_3. Therefore, we introduce a next round of counterterms:

$$H_c^{(2)} = H_c^{(1)} + \left(\frac{\gamma_1 p_1 W_\mathrm{L}^{[1]} + p_1^2 W_\mathrm{L} - W_\mathrm{L}^2/2}{p_2^2} + \Leftrightarrow \right) ,$$

and obtain

$$\begin{aligned} LH_c^{(2)} = &\left(\frac{-p_1^3 w_\mathrm{L} - 3\gamma_1 p_1^2 W_\mathrm{L} - \gamma_1^2 p_1 W_L^{[1]} + 4 p_1 W_\mathrm{L} w_\mathrm{L} + 2\gamma_1 T_1 W_\mathrm{L}}{p_2^2} + \Leftrightarrow \right) \\ &+ \frac{\gamma_1 w_\mathrm{L} W_\mathrm{L}^{[1]}}{p_2^2} + \frac{\gamma_4 w_\mathrm{R} W_\mathrm{R}^{[1]}}{p_3^2} - \frac{p_1 W_\mathrm{L} w_\mathrm{C}}{p_2^2} + \frac{p_4 W_\mathrm{R} w_\mathrm{C}}{p_3^2} + \mathcal{O}_c(|p_2|^{-3} + |p_3|^{-3}) . \end{aligned} \tag{4.3.6}$$

The terms in the first line are easy to eliminate, since each one depends on only one of the fast variables and averages to zero. The terms $\gamma_1 w_\mathrm{L} W_\mathrm{L}^{[1]}/p_2^2$ and $\gamma_4 w_\mathrm{R} W_\mathrm{R}^{[1]}/p_3^2$ are the ones we are looking for, since after integrating by parts, we find $\langle \gamma_1 w_\mathrm{L} W_\mathrm{L}^{[1]} \rangle_2 = -\gamma_1 \langle W_\mathrm{L}^2 \rangle_2 = -\alpha_2$ and $\langle \gamma_4 w_\mathrm{R} W_\mathrm{R}^{[1]} \rangle_3 = -\gamma_4 \langle W_\mathrm{R}^2 \rangle_3 = -\alpha_3$. The two "resonant" terms involving $W_\mathrm{L} w_\mathrm{C}$ and $W_\mathrm{R} w_\mathrm{C}$ are more problematic and we leave them untouched for now. By introducing the appropriate counterterms (which we do not write explicitly), we obtain a function $H_c^{(3)} = H_c^{(2)} + \mathcal{O}_c(|p_2|^{-3} + |p_3|^{-3})$ such that

$$LH_c^{(3)} = -\frac{\alpha_2}{p_2^2} - \frac{\alpha_3}{p_3^2} - \frac{p_1 W_\mathrm{L} w_\mathrm{C}}{p_2^2} + \frac{p_4 W_\mathrm{R} w_\mathrm{C}}{p_3^2} + \mathcal{O}_c(|p_2|^{-3} + |p_3|^{-3}) . \tag{4.3.7}$$

In order to obtain (4.3.3), we must get rid of the two "mixed" terms involving $W_\mathrm{L} w_\mathrm{C}$ and $W_\mathrm{R} w_\mathrm{C}$, which are of the same order as the dissipative contributions involving α_2 and α_3. Since they each depend on both q_2 and q_3, these terms are not easy to get rid of, due to the resonance phenomenon discussed above. In fact, as discussed in Appendix 4.6, these resonances have a physical meaning. Their effect becomes clearly visible when $T_1 = T_4 = 0$ (which is not covered by our assumptions): they alter the dynamics in the $p_2 p_3$-plane, but do not prevent H_c from decreasing in average. We postpone to §4.3.3 the construction of the counterterms that will eliminate these resonant terms.

We introduce next two technical lemmata and an application of Proposition 4.3.2. The following lemma is analogous to Lemma 4.2.3.

Lemma 4.3.3. *Let $j \in \{2, 3\}$ and $r > 0$. Fix an integer $\ell > 0$, and an $\varepsilon > 0$. Let f be some $\mathcal{O}_c(|p_j|^{-r})$ on the set $A_* = \{x \in \Omega : p_2 \neq 0, p_3 \neq 0\}$. Then, for all sufficiently large m and R, we have*

$$\sup_{x \in \Omega_c(\ell, m, R)} |f(x)| < \varepsilon .$$

Proof. We prove the result for $f = \mathcal{O}_c(|p_2|^{-r})$ and proceed as in Lemma 4.2.3. By Definition 4.3.1 and (4.1.12), there is a polynomial z such that for all m and all sufficiently large R, we have on $\Omega_c(\ell, m, R)$,

$$\begin{aligned}|f| &\le \frac{z(p_1,p_4)}{|p_2|^r} \le \frac{c + c(p_1^2+p_4^2)^N}{|p_2|^r} \le \frac{c + c(p_2^2+p_3^2)^{\frac{N}{m}}}{|p_2|^r} \\ &\le \frac{c + (p_2^2+p_2^{2\ell})^{\frac{N}{m}}}{|p_2|^r} \le \frac{c + c|p_2|^{\frac{2\ell N}{m}}}{|p_2|^r} \le c|p_2|^{\frac{2\ell N}{m}-r} ,\end{aligned}$$

where we choose N large enough and use the definition of Ω_c. By (4.1.12), we conclude that the desired result holds for m large enough so that $\frac{2\ell N}{m} - r < -\frac{r}{2}$. $\square$

Lemma 4.3.4. *Let $f = \mathcal{O}_c(p_2^{z_1} p_3^{z_2})$ for some $z_1, z_2 \in \mathbb{R}$ such that z_1, z_2 have the same sign. Then, $f = \mathcal{O}_c(|p_2|^{z_1+z_2} + |p_3|^{z_1+z_2})$.*

Proof. We apply Young's inequality in the form $xy \le x^a + y^b$ with $a = \frac{z_1+z_2}{z_1} > 1$, $b = \frac{z_1+z_2}{z_2} > 1$, and $x = |p_2|^{z_1}$, $y = |p_3|^{z_2}$. We obtain $|p_2|^{z_1}|p_3|^{z_2} \le |p_2|^{z_1+z_2} + |p_3|^{z_1+z_2}$. This, and the definition of $\mathcal{O}_c$, complete the proof. $\square$

As a consequence of Proposition 4.3.2 we have:

Proposition 4.3.5. *Let $0 < \theta < \min(1/T_1, 1/T_4)$ and define*

$$V_c = \widetilde{H}_c e^{\theta \widetilde{H}_c} \left(1 + \frac{F_2(q_2-q_1)}{p_2^3} + \frac{F_3(q_3-q_4)}{p_3^3}\right) ,$$

with the $\widetilde{H}_c$ of Proposition 4.3.2 and F_2, F_3 as in Proposition 4.2.4. Let $\ell > 1$ be a fixed integer. Then, there are constants $C_4, C_5, C_6 > 0$ such that for all large enough m and R, the following inequalities hold on $\Omega_c(m, \ell, R)$:

$$C_4(p_2^2+p_3^2)e^{\frac{\theta}{2}(p_2^2+p_3^2)} < V_c < C_5(p_2^2+p_3^2)e^{\frac{\theta}{2}(p_2^2+p_3^2)} , \tag{4.3.8}$$

$$LV_c \le -C_6 e^{\frac{\theta}{2}(p_2^2+p_3^2)} . \tag{4.3.9}$$

Proof. We first prove (4.3.8). By (4.3.2), the boundedness of the potentials, and Lemma 4.3.3, we have for m, R large enough that on Ω_c,

$$\left|\widetilde{H}_c - \frac{p_2^2}{2} - \frac{p_3^2}{2}\right| < c \qquad \text{and} \qquad \left|\frac{F_2(q_2-q_1)}{p_2^3} + \frac{F_3(q_3-q_4)}{p_3^3}\right| < \frac{1}{2} . \tag{4.3.10}$$

In addition, if m, R are large enough, $p_2^2 + p_3^2$ is large on Ω_c, so that the first part of (4.3.10) implies that $c(p_2^2+p_3^2) < \widetilde{H}_c < c'(p_2^2+p_3^2)$. This and (4.3.10) imply (4.3.8).

We next prove (4.3.9). Define $f(s) = se^{\theta s}$. By Proposition 4.3.2,

$$
\begin{aligned}
L\big(\widetilde{H}_c e^{\theta \widetilde{H}_c}\big) &= Lf(\widetilde{H}_c) = f'(\widetilde{H}_c)L\widetilde{H}_c + f''(\widetilde{H}_c)\sum_{b=1,4}\gamma_b T_b(\partial_{p_b}\widetilde{H}_c)^2 \\
&= e^{\theta\widetilde{H}_c}\left(\theta\widetilde{H}_c + 1\right)\left(-\frac{\alpha_2}{p_2^2} - \frac{\alpha_3}{p_3^2} + \mathcal{O}_c(|p_2|^{-5/2} + |p_3|^{-5/2})\right) \\
&\quad + e^{\theta\widetilde{H}_c}\left(\theta^2\tilde{H}_c + 2\theta\right)\left(\gamma_1 T_1\frac{W_{\mathrm{L}}^2}{p_2^2} + \gamma_4 T_4\frac{W_{\mathrm{R}}^2}{p_3^2} + \mathcal{O}_c(|p_2|^{-3} + |p_3|^{-3})\right)\ .
\end{aligned}
\tag{4.3.11}
$$

Now observe that for any $C \in \mathbb{R}$, we have

$$
\widetilde{H}_c + C = \frac{p_2^2 + p_3^2}{2} + \mathcal{O}_c(1) = \frac{p_2^2 + p_3^2}{2}\left(1 + \mathcal{O}_c(p_2^{-2} + p_3^{-2})\right),
$$

since trivially $(p_2^2 + p_3^2)^{-1} \leq p_2^{-2} + p_3^{-2}$. But then, by (4.3.11) and Lemma 4.3.4, we find that

$$
L\big(\widetilde{H}_c e^{\theta\widetilde{H}_c}\big) \leq e^{\theta\widetilde{H}_c}\frac{p_2^2 + p_3^2}{2}\left(\frac{\theta^2\gamma_1 T_1 W_{\mathrm{L}}^2 - \theta\alpha_2 + \mathcal{O}_c(|p_2|^{-1/2})}{p_2^2} + \Leftrightarrow\right)\ .
$$

As in the proof of Proposition 4.2.4, the corrections involving F_2 and F_3 replace the oscillatory terms W_{L}^2 and W_{R}^2 with their averages:

$$
\begin{aligned}
LV_c &= L\left(\widetilde{H}_c e^{\theta\widetilde{H}_c}\right)\left(1 + \frac{F_2}{p_2^3} + \frac{F_3}{p_3^3}\right) + \widetilde{H}_c e^{\theta\widetilde{H}_c} L\left(\frac{F_2}{p_2^3} + \frac{F_3}{p_3^3}\right) \\
&\leq e^{\theta\widetilde{H}_c}\frac{p_2^2 + p_3^2}{2}\left(\frac{\theta^2\gamma_1 T_1\left\langle W_{\mathrm{L}}^2\right\rangle - \theta\alpha_2 + \mathcal{O}_c(|p_2|^{-1/2})}{p_2^2} + \Leftrightarrow\right)\ .
\end{aligned}
$$

Therefore, by the definition (4.1.14) of α_j and the condition on θ, we have

$$
LV_c \leq e^{\theta\widetilde{H}_c}\frac{p_2^2 + p_3^2}{2}\left(\frac{-c + \mathcal{O}_c(|p_2|^{-1/2})}{p_2^2} + \frac{-c + \mathcal{O}_c(|p_3|^{-1/2})}{p_3^2}\right)\ .
$$

Finally, by Lemma 4.3.3, and using that $(p_2^2 + p_3^2)(p_2^{-2} + p_3^{-2}) > 2$, we indeed obtain (4.3.9). □

We now return to the proof of Proposition 4.3.2. We need to find some counterterms to eliminate the mixed terms in (4.3.7). For this, we use a subdivision of $A_* = \{x \in \Omega : p_2 \neq 0, p_3 \neq 0\}$ into 3 disjoint pieces, as shown in Figure 4.7:

$$
\begin{aligned}
A_1 &= \{x \in A_* : |p_2 + p_3| \geq (p_2 - p_3)^2\}\ , \\
A_2 &= \big\{x \in A_* : (p_2 - p_3)^2 > |p_2 + p_3| > (p_2 - p_3)^2/2\big\}\ , \\
A_3 &= \big\{x \in A_* : (p_2 - p_3)^2 \geq 2|p_2 + p_3|\big\}\ .
\end{aligned}
\tag{4.3.12}
$$

By construction, A_1 is close to the diagonal $p_2 = p_3$, A_3 is far from it, and A_2 is some transition region.

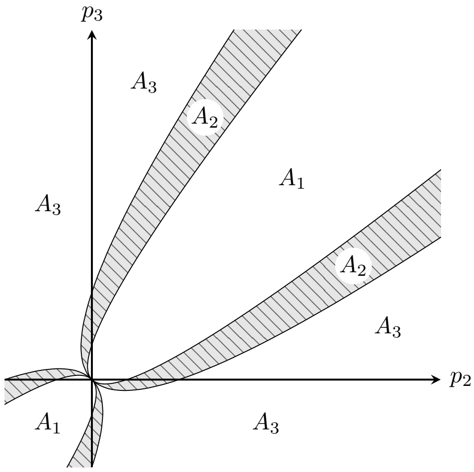


Figure 4.7 – Projection of the partition $A_* = A_1 \cup A_2 \cup A_3$ onto the p_2p_3-plane. Note that the sets A_1, A_2 and A_3 do not include the p_2 and p_3 axes.

Lemma 4.3.6. *The following holds:*

(i) On $A_1 \cup A_2$, the quantity $|p_2 - p_3|$ is both $\mathcal{O}_c(\sqrt{|p_2|})$ and $\mathcal{O}_c(\sqrt{|p_3|})$.

(ii) On $A_2 \cup A_3$, the quantity $|p_2 - p_3|^{-1}$ is both $\mathcal{O}_c(|p_2|^{-1/2})$ and $\mathcal{O}_c(|p_3|^{-1/2})$.

Proof. Trivially, (i) holds because on $A_1 \cup A_2$, we have the scaling $|p_3 - p_2| \lesssim \sqrt{|p_2+p_3|} \sim \sqrt{p_2} \sim \sqrt{p_3}$. To obtain (ii), observe that either p_2 and p_3 have the same sign and by the definition of $A_2 \cup A_3$, $|p_2 - p_3| \gtrsim \sqrt{|p_2+p_3|} = \sqrt{|p_2|+|p_3|} \geq \max(\sqrt{|p_2|}, \sqrt{|p_3|})$, or they have a different sign and $|p_2 - p_3| = |p_2| + |p_3| \geq \max(|p_2|, |p_3|) \gtrsim \max(\sqrt{|p_2|}, \sqrt{|p_3|})$. In both cases, we have the desired bound. □

We first work on $A_1 \cup A_2$. In this region, p_2 and p_3 are close to each other, and are both large in absolute value. It is then easy to find a counterterm for $p_1 W_{\mathrm{L}} w_{\mathrm{C}}/p_2^2$ and $p_4 W_{\mathrm{R}} w_{\mathrm{C}}/p_3^2$. Indeed, W_{L} and W_{R} oscillate very rapidly (the respective frequencies are approximately $p_2/2\pi$ and $p_3/2\pi$), while w_{C} oscillates only "moderately", with frequency $(p_3 - p_2)/2\pi$. One can then simply average the rapidly oscillating part, and obtain

Lemma 4.3.7. *Let $R_{12} = p_1 W_{\mathrm{L}}^{[1]} w_{\mathrm{C}}/p_2^3 - p_4 W_{\mathrm{R}}^{[1]} w_{\mathrm{C}}/p_3^3$. Then,*

$$LR_{12} = \frac{p_1 W_{\mathrm{L}} w_{\mathrm{C}}}{p_2^2} - \frac{p_4 W_{\mathrm{R}} w_{\mathrm{C}}}{p_3^2} + \mathcal{O}_c(|p_2|^{-5/2} + |p_3|^{-5/2}) \qquad (\textit{on } A_1 \cup A_2)\,.$$

Proof. We have for the first term:

$$\begin{aligned} L\frac{p_1 W_{\mathrm{L}}^{[1]} w_{\mathrm{C}}}{p_2^3} &= \frac{p_1 W_{\mathrm{L}} w_{\mathrm{C}}}{p_2^2} + \frac{p_1 W_{\mathrm{L}}^{[1]} w_{\mathrm{C}}' \cdot (p_3 - p_2)}{p_2^3} + \mathcal{O}_c(p_2^{-3}) \\ &= \frac{p_1 W_{\mathrm{L}} w_{\mathrm{C}}}{p_2^2} + \mathcal{O}_c(|p_2|^{-5/2})\,, \end{aligned}$$

where the last equality uses Lemma 4.3.6 (i). A similar computation for the second term completes the proof. □

The counterterm R_{12} works well on $A_1 \cup A_2$ because $|p_3 - p_2|$ is small compared to p_2 and p_3. We now have to find a counterterm R_{23} that works on $A_2 \cup A_3$ and then patch the two counterterms together on A_2. We state the properties of the counterterm R_{23} in the following lemma, but postpone its construction to §4.3.3.

Lemma 4.3.8. *There is a function $R_{23} = \mathcal{O}_c(|p_2|^{-2} + |p_3|^{-2})$ defined on $A_2 \cup A_3$ such that*

$$LR_{23} = \frac{p_1 W_{\rm L} w_{\rm C}}{p_2^2} - \frac{p_4 W_{\rm R} w_{\rm C}}{p_3^2} + \mathcal{O}_c(|p_2|^{-5/2} + |p_3|^{-5/2}) \qquad (\textit{on } A_2 \cup A_3)$$

and

$$\partial_{p_1} R_{23} = \mathcal{O}_c(p_2^{-2}) \qquad \textit{and} \qquad \partial_{p_4} R_{23} = \mathcal{O}_c(p_3^{-2})\,. \tag{4.3.13}$$

Assuming that Lemma 4.3.8 is proved, we next join the two counterterms R_{12} and R_{23} by a smooth interpolation on A_2 in order to prove Proposition 4.3.2.

Proof of Proposition 4.3.2. We introduce a smooth function $\rho : \mathbb{R} \cup \{-\infty, \infty\} \to [0,1]$ such that $\rho(x) = 1$ when $|x| \leq 1$ and $\rho(x) = 0$ when $|x| \geq 2$. We then consider the function

$$\rho\left(\frac{(p_3 - p_2)^2}{p_2 + p_3}\right)\,, \tag{4.3.14}$$

which is well-defined and smooth on the set $A_* = A_1 \cup A_2 \cup A_3 = \{x \in \Omega : p_2 \neq 0, p_3 \neq 0\}$. Moreover, it is equal to 1 on A_1, and 0 on A_3. We now omit the arguments and simply write ρ instead of (4.3.14). Using Lemma 4.3.7 and Lemma 4.3.8, we obtain

$$\begin{aligned} L\left(\rho R_{12} + (1-\rho) R_{23}\right) &= \rho L R_{12} + (1-\rho) L R_{23} + (R_{12} - R_{23}) L\rho \\ &= \frac{p_1 W_{\rm L} w_{\rm C}}{p_2^2} - \frac{p_4 W_{\rm R} w_{\rm C}}{p_3^2} + \mathcal{O}_c(|p_2|^{-5/2} + |p_3|^{-5/2}) + (R_{12} - R_{23}) L\rho\,. \end{aligned} \tag{4.3.15}$$

Observe next that

$$L\rho = \rho' \cdot \left(2\frac{(p_3 - p_2)}{p_2 + p_3}(w_{\rm L} - w_{\rm R} - 2w_{\rm C}) + \frac{(p_3 - p_2)^2}{(p_2 + p_3)^2}(w_{\rm L} + w_{\rm R})\right)\,. \tag{4.3.16}$$

Since ρ' has support in A_2, where $|p_3 - p_2| \sim |p_2 + p_3|^{\frac{1}{2}} \sim |p_2|^{1/2} \sim |p_3|^{1/2}$, we see that $L\rho$ is simultaneously $\mathcal{O}_c(|p_2 + p_3|^{-1/2})$, $\mathcal{O}_c(|p_2|^{-1/2})$ and $\mathcal{O}_c(|p_3|^{-1/2})$. But then, by (4.3.15) and using that $R_{12} - R_{23} = \mathcal{O}_c(|p_2|^{-2} + |p_3|^{-2})$, we find

$$L\left(\rho R_{12} + (1-\rho) R_{23}\right) = \frac{p_1 W_{\rm L} w_{\rm C}}{p_2^2} - \frac{p_4 W_{\rm R} w_{\rm C}}{p_3^2} + \mathcal{O}_c(|p_2|^{-5/2} + |p_3|^{-5/2})\,. \tag{4.3.17}$$

We set now

$$\widetilde{H}_c = H_c^{(3)} + \rho R_{12} + (1-\rho) R_{23}\,.$$

From (4.3.17) and (4.3.7), we deduce immediately that (4.3.3) holds. Moreover, (4.3.4) follows from (4.3.13), the expressions for $H_c^{(3)}$ and R_{12}, and the fact that ρ does not depend on p_1 and p_4. This completes the proof of Proposition 4.3.2. □

4.3.3. Fully decoupled dynamics approximation

We construct here the counterterm R_{23} of Lemma 4.3.8, which eliminates the two resonant terms $-p_1 W_{\mathrm{L}} w_{\mathrm{C}}/p_2^2$ and $p_4 W_{\mathrm{R}} w_{\mathrm{C}}/p_3^2$ on $A_2 \cup A_3$ when both $|p_2|$ and $|p_3|$ are large. In this regime, all three interaction forces $w_{\mathrm{L}}, w_{\mathrm{C}}, w_{\mathrm{R}}$ oscillate rapidly (since $|p_2|, |p_3 - p_2|$ and $|p_3|$ are all large) and we expect the dynamics to be well approximated by the following *decoupled dynamics*, where all the interaction forces are removed.

Definition 4.3.9. *We call* decoupled dynamics *the SDE*

$$\begin{aligned}
\mathrm{d}q_i &= p_i \,\mathrm{d}t \,, & i &= 1, \dots, 4 \,, \\
\mathrm{d}p_b &= -\gamma_b p_b \,\mathrm{d}t + \sqrt{2\gamma_b T_b}\mathrm{d}B_t^b \,, & b &= 1, 4 \,, \\
\mathrm{d}p_j &= 0 \,\mathrm{d}t \,, & j &= 2, 3 \,,
\end{aligned} \tag{4.3.18}$$

with generator

$$\bar{L} = \sum_{i=1}^{4} p_i \partial_{q_i} + \sum_{b=1,4} \left(-\gamma_b p_b \partial_{p_b} + \gamma_b T_b \partial_{p_b}^2 \right) \,, \tag{4.3.19}$$

and denote by $\bar{\mathbb{E}}_x$ the corresponding expectation value with initial condition $x \in \Omega$.

We will construct two functions U_1, U_4 such that $\bar{L}U_1 = p_1 W_{\mathrm{L}} w_{\mathrm{C}}$ and $\bar{L}U_4 = -p_4 W_{\mathrm{R}} w_{\mathrm{C}}$. Then, we will introduce a change of variable $x \mapsto \bar{x}(x)$ such that $\bar{x}$ approximately obeys the decoupled dynamics, so that $L(U_1(\bar{x})) \approx p_1 W_{\mathrm{L}} w_{\mathrm{C}}$ and $L(U_4(\bar{x})) \approx -p_4 W_{\mathrm{R}} w_{\mathrm{C}}$ in the regime of interest. Finally, we will show that the choice $R_{23}(x) = U_1(\bar{x})/p_2^2 + U_4(\bar{x})/p_3^2$ satisfies the conclusions of Lemma 4.3.8.

The decoupled dynamics can be integrated explicitly for any initial condition $x = (q_1, \dots, p_4) \in \Omega$. For the outer rotors $b = 1, 4$, we have

$$\begin{aligned}
p_b(t) &= e^{-\gamma_b t} p_b + \sqrt{2\gamma_b T_b} \int_0^t e^{-(t-s)\gamma_b} \mathrm{d}B_s^b \,, \\
q_b(t) &= q_b + \frac{1 - e^{-\gamma_b t}}{\gamma_b} p_b + \sqrt{2\gamma_b T_b} \int_0^t \Big(\int_0^s e^{-(s-s')\gamma_b} \mathrm{d}B_{s'}^b \Big) \mathrm{d}s \,,
\end{aligned} \tag{4.3.20}$$

and for the central ones ($j = 2, 3$) we simply have

$$\begin{aligned}
p_j(t) &= p_j \,, \\
q_j(t) &= q_j + p_j t \pmod{2\pi} \,,
\end{aligned} \tag{4.3.21}$$

which is deterministic. We decompose the variables between the central and external rotors as

$$x = (x_e, x_c) \quad \text{with } x_e = (q_1, p_1, q_4, p_4) \quad \text{and } x_c = (q_2, p_2, q_3, p_3) \,.$$

Under the decoupled dynamics, the two processes $x_e(t)$ and $x_c(t)$ are independent and $x_c(t)$ is deterministic. Moreover, under the decoupled dynamics, $x_e(t)$ has the generator

$$\bar{L}_e = \sum_{b=1,4} (p_b \partial_{q_b} - \gamma_b p_b \partial_{p_b} + \gamma_b T_b \partial_{p_b}^2) ,$$

and admits the invariant probability measure $\bar{\pi}_e$ on $(\mathbb{T}\times\mathbb{R})^2$ given by

$$\mathrm{d}\bar{\pi}_e(x_e) = \frac{1}{Z} e^{-\frac{p_1^2}{2T_1} - \frac{p_4^2}{2T_4}} \mathrm{d}q_1 \mathrm{d}p_1 \mathrm{d}q_4 \mathrm{d}p_4 ,$$

where Z is a normalization constant (recall that $T_1, T_4 > 0$ by assumption).

Definition 4.3.10. *We denote by $\mathcal{S}$ the set of functions $f \in \mathcal{C}^\infty(\Omega, \mathbb{R})$ for which the norm*

$$|\!|\!| f |\!|\!| = \sup_{x\in\Omega} \frac{|f(x)|}{1 + p_1^2 + p_4^2} \tag{4.3.22}$$

is finite. We denote by $\mathcal{S}_0$ the subspace of functions $f \in \mathcal{S}$ for which

$$\int_{(\mathbb{T}\times\mathbb{R})^2} f(x_e, x_c) \mathrm{d}\bar{\pi}_e(x_e) = 0 \qquad \textit{for all } x_c \in (\mathbb{T}\times\mathbb{R})^2 .$$

We will later consider $f = p_1 W_{\mathrm{L}} w_{\mathrm{C}}$ and $f = -p_4 W_{\mathrm{R}} w_{\mathrm{C}}$, which are manifestly in $\mathcal{S}_0$.

Lemma 4.3.11. *There are constants $C_*, c_* > 0$ such that for all $f \in \mathcal{S}_0$, all $x \in \Omega$, and all $t \geq 0$,*

$$\left|\bar{\mathbb{E}}_x f(x(t))\right| \leq C_* e^{-c_* t} |\!|\!| f |\!|\!| \left(1 + p_1^2 + p_4^2\right) . \tag{4.3.23}$$

Proof. As mentioned, $x_e(t)$ and $x_c(t)$ are independent under the decoupled dynamics. Introducing the expectation value $\bar{\mathbb{E}}^e$ with respect to the process $x_e(t)$ under the decoupled dynamics, we obtain that for any function f on Ω,

$$\bar{\mathbb{E}}_x f(x(t)) = \bar{\mathbb{E}}^e_{x_e} f(x_e(t), x_c(t)) , \tag{4.3.24}$$

where $x_c(t)$ is (deterministically) given by (4.3.21).

The process $x_e(t)$ under the decoupled dynamics is exponentially ergodic, with the unique invariant measure $\bar{\pi}_e$ defined above. Indeed, one can check explicitly that this measure is invariant, and introducing the Lyapunov function $V_e(x_e) = 1 + p_1^2 + p_4^2$, we easily obtain that $\bar{L}_e V_e \leq c - cV_e$. It follows from [44, Theorem 6.1][4] that there are two constants $C_*, c_* > 0$ such that for any function $g : (\mathbb{T}\times\mathbb{R})^2 \to \mathbb{R}$ such that g/V_e is bounded,

$$\sup_{x_e} \frac{|\bar{\mathbb{E}}^e_{x_e} g(x_e(t)) - \bar{\pi}_e(g)|}{1 + p_1^2 + p_4^2} \leq C_* e^{-c_* t} \sup_{x_e} \frac{|g(x_e) - \bar{\pi}_e(g)|}{1 + p_1^2 + p_4^2} . \tag{4.3.25}$$

[4] One should also check that there is a skeleton with respect to which every compact set is petite. This is obvious, but can be proved with methods similar to those of §3.5.2.

Let now $f \in \mathcal{S}_0$. For any fixed $v \in (\mathbb{T} \times \mathbb{R})^2$, we apply (4.3.25) to the function $g_v(x_e) = f(x_e, v)$. Since $f \in \mathcal{S}_0$, we have $\bar{\pi}_e(g_v) = 0$. Therefore, for any $t \geq 0$,

$$\sup_{x_e} \frac{|\bar{\mathbb{E}}^e_{x_e} f(x_e(t), v)|}{1 + p_1^2 + p_4^2} \leq C_* e^{-c_* t} \sup_{x_e} \frac{|f(x_e, v)|}{1 + p_1^2 + p_4^2} \leq C_* e^{-c_* t} |\!|\!| f |\!|\!| \ . \tag{4.3.26}$$

This holds for all v, and in particular for $v = x_c(t)$. Therefore, by (4.3.24), we have the desired result. □

The next proposition constructs a right inverse of $\bar{L}$ on $\mathcal{S}_0$ [45–47]. We use here the notation

$$x = (x_1, \ldots, x_8) = (q_1, \ldots, q_4, p_1, \ldots, p_4) \ . \tag{4.3.27}$$

Proposition 4.3.12. *Let $f \in \mathcal{S}_0$ be a function such that for all multi-indices $\underline{a}$, we have $\partial^{\underline{a}} f \in \mathcal{S}_0$, and let*

$$(\bar{K} f)(x) = -\int_0^\infty \bar{\mathbb{E}}_x f(x(t))\, \mathrm{d}t \ . \tag{4.3.28}$$

Then:

(i) $\bar{K} f$ and its derivatives of all orders are in $\mathcal{S}$.

(ii) We have

$$\bar{L} \bar{K} f = f \ .$$

Proof. By Lemma 4.3.11, the integral (4.3.28) converges absolutely for all x and we have $\bar{K} f \in \mathcal{S}$. We now prove the result about the derivatives. By (4.3.20) and (4.3.21), we can write

$$\frac{\partial x_i(t)}{\partial x_j} = h_{ij}(t) \ , \tag{4.3.29}$$

where the h_{ij} are deterministic functions of t only that grow at most linearly (namely 0, 1, $e^{-\gamma_b t}$, $(1 - e^{-\gamma_b t})/\gamma_b$ and t). We then have

$$\frac{\partial}{\partial x_j} \left(\bar{\mathbb{E}}_x f(x(t)) \right) = \sum_{i=1}^{8} h_{ij}(t) \bar{\mathbb{E}}_x [(\partial_i f)(x(t))] \ .$$

For the derivatives of order n, we find by induction

$$\partial_{j_1, \ldots, j_n} \left(\bar{\mathbb{E}}_x f(x(t)) \right) = \sum_{i_1, \ldots, i_n} \Big(\prod_{k=1}^{n} h_{i_k j_k}(t) \Big) \bar{\mathbb{E}}_x [(\partial_{i_1, \ldots, i_n} f)(x(t))] \ , \tag{4.3.30}$$

where the sum is taken over all $(i_1, \ldots, i_n) \in \{1, 2, \ldots, 8\}^n$. Since by assumption $\partial_{i_1, \ldots, i_n} f \in \mathcal{S}_0$, we have by Lemma 4.3.11 that

$$\left| \bar{\mathbb{E}}_x [(\partial_{i_1, \ldots, i_n} f)(x(t))] \right| \leq c e^{-c_* t} \left(1 + p_1^2 + p_4^2 \right) \ .$$

But then, by (4.3.30), we have

$$\begin{aligned}\left|\partial_{i_1,\dots,i_n}\bar{K}f(x)\right| &= \left|\int_0^\infty \partial_{i_1,\dots,i_n}\left(\bar{\mathbb{E}}_x f(x(t))\right)\mathrm{d}t\right| \\ &\le c\left(1+p_1^2+p_4^2\right)\sum_{i_1,\dots,i_n}\left|\int_0^\infty \Big(\prod_{k=1}^n h_{i_k j_k}(t)\Big)e^{-c_* t}\,\mathrm{d}t\right| .\end{aligned}$$

Since the h_{ij} grow at most linearly, the time-integrals in the right-hand side converge. Therefore, $\bar{K}f$ is $\mathcal{C}^\infty$ and (i) holds.

For the second statement, we observe that

$$\bar{L}\bar{K}f = -\int_0^\infty \bar{L}\bar{\mathbb{E}}_x f(x(t))\mathrm{d}t = -\int_0^\infty \frac{\mathrm{d}}{\mathrm{d}t}\bar{\mathbb{E}}_x f(x(t))\mathrm{d}t = \bar{\mathbb{E}}_x f(x(0)) = f(x) ,$$

where we have used that $\lim_{t\to\infty}\mathbb{E}_x f(x(t)) = 0$ by (4.3.26). □

Remark 4.3.13. The proof of Proposition 4.3.12, and in particular (4.3.29), relies on the linear nature of the decoupled dynamics. If we add constant forces τ_1 and τ_4 at the ends of the chain (as in Chapter 3), the method above applies with little modification, and with the replacements $p_b \to p_b - \tau_b/\gamma_b$, $b = 1,4$, in the invariant measure $\bar{\pi}_e$. However, if we add pinning potentials of the kind $U(q_i)$, the decoupled dynamics cannot be solved explicitly, and we do not have (4.3.29) for some deterministic functions $h_{ij}(t)$. Although we believe there exists an analog of Proposition 4.3.12 in that case, we are currently unable to provide it. The situation is even worse in the simultaneous presence of constant forces and pinning potentials. In that case, the expression of $\bar{\pi}_e$ is not known [29], which makes it difficult to decide whether a given function is in $\mathcal{S}_0$. (Of course, although there is no difficulty there, the averaging of p_2, p_3 and H_c also needs to be adapted to accommodate for such modifications of the model.)

We now have an inverse of $\bar{L}$ on a given class of functions. We next use it to find an approximate inverse of L. The key is to introduce a change of variables $\bar{x} = (\bar{q}_1, \bar{p}_1, \dots, \bar{q}_4, \bar{p}_4)$ such that for nice enough functions f, it holds that $L(f(\bar{x})) \approx (\bar{L}f)(\bar{x})$ in the regime of interest. Here and in the sequel, it is always understood that $\bar{x}$ is viewed as a function of x. We compare the actions of L and $\bar{L}$ in Lemma 4.3.14. We state this lemma with the notation (4.3.27), and write generically

$$L = \sum_i (b_i(x)\partial_i + \sigma_i\partial_i^2) \qquad \text{and} \qquad \bar{L} = \sum_i (\bar{b}_i(x)\partial_i + \sigma_i\partial_i^2) . \tag{4.3.31}$$

In our case, only σ_5 and σ_8, which correspond to the variables p_1 and p_4, are non-zero.

Lemma 4.3.14. *Consider a change of coordinates $x \mapsto \bar{x}(x) = x + s(x)$, defined on some set $\Omega_0 \subset \Omega$. Assume that for all j,*

$$L(\bar{x}_j) = \bar{b}_j(\bar{x}) + \varepsilon_j(x)$$

for some ε_j. Then, for any smooth function h, we have for all $x \in \Omega_0$ that

$$L(h(\bar{x})) = (\bar{L}h)(\bar{x}) + \zeta(x) ,$$

where

$$\zeta(x) = \sum_j (\partial_j h)(\bar{x})\varepsilon_j(x) + 2\sum_{i,k} \sigma_i (\partial_{ik} h)(\bar{x})\partial_i s_k(x) + \sum_{i,j,k} \sigma_i (\partial_{jk} h)(\bar{x})\partial_i s_j(x)\partial_i s_k(x)\ . \quad (4.3.32)$$

Proof. We do the computation for the case of just one variable $x \in \mathbb{R}$. Let $g(x) = \bar{x}(x) = x + s(x)$. From the definition of L and $\bar{L}$, and since by assumption $Lg = \bar{b} \circ g + \varepsilon$, we find

$$\begin{aligned} L(h \circ g) &= (h' \circ g) \cdot Lg + \sigma \cdot (h'' \circ g) \cdot {g'}^2 \\ &= (h' \circ g) \cdot (\bar{b} \circ g + \varepsilon) + \sigma \cdot (h'' \circ g) \cdot {g'}^2 \\ &= (\bar{L}h) \circ g + (h' \circ g) \cdot \varepsilon + \sigma \cdot (h'' \circ g) \cdot ({g'}^2 - 1) \\ &= (\bar{L}h) \circ g + (h' \circ g) \cdot \varepsilon + \sigma \cdot (h'' \circ g) \cdot (2s' + s'^2)\ . \end{aligned}$$

The desired result follows from generalizing to the multivariate case. $\square$

We consider now the following change of variables defined on $A_2 \cup A_3$:

$$\begin{aligned} \bar{q}_1 &= q_1\ , & \bar{p}_1 &= p_1 - \frac{W_{\mathrm{L}}(q_2 - q_1)}{p_2} = p_1 + \mathcal{O}_c(p_2^{-1})\ , \\ \bar{q}_2 &= q_2\ , & \bar{p}_2 &= p_2 + \frac{W_{\mathrm{L}}(q_2 - q_1)}{p_2} - \frac{W_{\mathrm{C}}(q_3 - q_2)}{p_3 - p_2} = p_2 + \mathcal{O}_c(|p_2|^{-1/2})\ , \end{aligned} \quad (4.3.33)$$

with analogous expressions for the indices 3, 4. Here, we have used Lemma 4.3.6 (ii) to replace $W_{\mathrm{L}}/(p_3 - p_2)$ with $\mathcal{O}_c(|p_2|^{-1/2})$. Straightforward computations show that, on $A_2 \cup A_3$,

$$\begin{aligned} L(\bar{p}_1) &= -\gamma_1 p_1 + \mathcal{O}_c(p_2^{-1}) = -\gamma_1 \bar{p}_1 + \mathcal{O}_c(p_2^{-1})\ , \\ L(\bar{q}_1) &= p_1 = \bar{p}_1 + \mathcal{O}_c(p_2^{-1})\ , \\ L(\bar{p}_2) &= \mathcal{O}_c(|p_2|^{-1} + (p_2 - p_3)^{-2}) = \mathcal{O}_c(p_2^{-1})\ , \\ L(\bar{q}_2) &= p_2 = \bar{p}_2 + \mathcal{O}_c(|p_2|^{-1/2})\ , \end{aligned} \quad (4.3.34)$$

with similar expressions for the indices $3, 4$ (we have again used Lemma 4.3.6 (ii)).

While one could choose a more refined change of variables by going to higher orders, the change (4.3.33) is good enough for our purpose.

Lemma 4.3.15. *Let $f \in \mathcal{S}$. Then f is $\mathcal{O}_c(1)$. Moreover, given any function $\xi : \Omega \to [0, 1]$, we have that $f\big(x + \xi(x)(\bar{x} - x)\big) = \mathcal{O}_c(1)$ on $A_2 \cup A_3$. In particular, $f(\bar{x}) = \mathcal{O}_c(1)$ on $A_2 \cup A_3$.*

Proof. By assumption, $|f(x)| \le |||f|||(1 + p_1^2 + p_4^2)$ on Ω, so that immediately $f = \mathcal{O}_c(1)$. Moreover, $f\big(x + \xi(x)(\bar{x} - x)\big)$ is well-defined on $A_2 \cup A_3$, and

$$\big|f\big(x + \xi(x)(\bar{x} - x)\big)\big| \le |||f||| \left(1 + \left(p_1 - \xi(x)\frac{W_{\mathrm{L}}}{p_2}\right)^2 + \left(p_4 - \xi(x)\frac{W_{\mathrm{R}}}{p_3}\right)^2\right)\ ,$$

which is indeed a $\mathcal{O}_c(1)$ on this set. The claim about $f(\bar{x})$ follows from the choice $\xi \equiv 1$. $\square$

Proposition 4.3.16. *Let f satisfy the assumptions of Proposition 4.3.12, and consider the change of coordinates (4.3.33). Let $h = \bar{K}f$. Then, on the set $A_2 \cup A_3$ (meaning that we take $x \in A_2 \cup A_3$, and not necessarily $\bar{x} \in A_2 \cup A_3$), we have $h(\bar{x}) = \mathcal{O}_c(1)$ and*

$$L(h(\bar{x})) = f(x) + \mathcal{O}_c(|p_2|^{-1/2} + |p_3|^{-1/2}) .$$

Proof. We use again the notations $x = (x_1, \ldots, x_8) = (q_1, \ldots, q_4, p_1, \ldots, p_4)$ and (4.3.31). We apply Lemma 4.3.14 with the coordinate change $\bar{x} = x + s(x)$ defined by (4.3.33). Then, the s_j are given by (4.3.33), and the ε_j are given by (4.3.34). Observe then that on $A_2 \cup A_3$, all the s_j and ε_j and are at most $\mathcal{O}_c(|p_2|^{-1/2})$ or $\mathcal{O}_c(|p_3|^{-1/2})$. The only non-zero σ_i are $\sigma_5 = \gamma_1 T_1$ and $\sigma_8 = \gamma_4 T_4$. Moreover, $\partial_{x_5} s_j = \partial_{p_1} s_j = 0$ for all $j \in \{1, 2, \ldots, 8\}$, and similarly $\partial_{x_8} s_j = \partial_{p_4} s_j = 0$. Therefore, from (4.3.32) we are left with $\zeta(x) = \sum_j (\partial_j h)(\bar{x}) \varepsilon_j(x)$. We now apply this to the function $h = \bar{K}f$. By Proposition 4.3.12, we have $\bar{L}h = f$, so that

$$L(h(\bar{x})) = f(\bar{x}) + \sum_j (\partial_j h)(\bar{x}) \varepsilon_j(x) . \tag{4.3.35}$$

To obtain the desired results, it remains to make the following two observations. First, by the mean value theorem, there is for each x some $\xi(x) \in [0, 1]$ such that on $A_2 \cup A_3$,

$$f(\bar{x}) - f(x) = \sum_j s_j(x)(\partial_j f)(x + \xi(x)s(x)) = \mathcal{O}_c(|p_2|^{-1/2} + |p_3|^{-1/2}) , \tag{4.3.36}$$

where we have applied Lemma 4.3.15 to $\partial_j f$, which is in $\mathcal{S}$ by assumption. Secondly, using Lemma 4.3.15 and the fact that $\partial_j h \in \mathcal{S}$ by Proposition 4.3.12, we find

$$\sum_j (\partial_j h)(\bar{x}) \varepsilon_j(x) = \mathcal{O}_c(|p_2|^{-1/2} + |p_3|^{-1/2}) ,$$

which, together with (4.3.35) and (4.3.36), completes the proof. $\square$

We are now ready for the

Proof of Lemma 4.3.8. Let

$$\begin{aligned} U_1(q_1, \ldots, q_3, p_1, \ldots, p_3) &= \bar{K}(p_1 W_{\mathrm{L}}(q_2 - q_1) w_{\mathrm{C}}(q_3 - q_2)) , \\ U_4(q_2, \ldots, q_4, p_2, \ldots, p_4) &= \bar{K}(-p_4 W_{\mathrm{R}}(q_3 - q_4) w_{\mathrm{C}}(q_3 - q_2)) , \end{aligned}$$

and

$$R_{23}(x) = \frac{U_1(\bar{x})}{p_2^2} + \frac{U_4(\bar{x})}{p_3^2} .$$

That U_1 depends only on $(q_1, \ldots, q_3, p_1, \ldots, p_3)$ follows from the independence of the four rotors under the decoupled dynamics. Similarly for U_4. It is easy to check that $f = p_1 W_{\mathrm{L}} w_{\mathrm{C}}$ satisfies the assumptions of Proposition 4.3.12: Since $\langle f \rangle_1 = 0$, we also have $\langle \partial^{\underline{a}} f \rangle_1 = 0$ for each multi-index $\underline{a}$. From this it follows that $\bar{\pi}_e(f) = 0$ and that $\bar{\pi}_e(\partial^{\underline{a}} f) = 0$, since $\bar{\pi}_e$ is uniform with respect to q_1.

Since no powers of p_1 or p_4 appear upon differentiation, we indeed obtain that f and all its derivatives are in $\mathcal{S}_0$. A similar argument applies to $f = -p_4 W_{\rm L} w_{\rm C}$. Therefore, applying Proposition 4.3.16, we find that on the set $A_2 \cup A_3$, the functions $U_1(\bar{x})$ and $U_4(\bar{x})$ are $\mathcal{O}_c(1)$, and that

$$\begin{aligned} L(U_1(\bar{x})) &= p_1 W_{\rm L} w_{\rm C} + \mathcal{O}_c(|p_2|^{-1/2} + |p_3|^{-1/2}) ,\\ L(U_4(\bar{x})) &= -p_4 W_{\rm R} w_{\rm C} + \mathcal{O}_c(|p_2|^{-1/2} + |p_3|^{-1/2}) . \end{aligned} \tag{4.3.37}$$

In (4.3.37), the arguments of $W_{\rm L}, W_{\rm R}$ and $W_{\rm C}$ are indeed x and not $\bar{x}$. Finally, we have

$$LR_{23} = \frac{L(U_1(\bar{x}))}{p_2^2} + \frac{L(U_4(\bar{x}))}{p_3^2} + \mathcal{O}_c(p_2^{-3}) + \mathcal{O}_c(p_3^{-3}) . \tag{4.3.38}$$

The main assertion of the lemma then follows from this, (4.3.37), and Lemma 4.3.4. The assertion (4.3.13) follows from the definition of R_{23} and the following observation: using the explicit expression for $\bar{x}$, Proposition 4.3.12 (i) and Lemma 4.3.15, we obtain $\partial_{p_1}(U_1(\bar{x})) = (\partial_{p_1}U_1)(\bar{x}) = \mathcal{O}_c(1)$, and $\partial_{p_4}(U_1(\bar{x})) = (\partial_{p_4}U_1)(\bar{x}) = 0$ (and similarly for U_4). □

Remark 4.3.17. The construction above relies on the strict positivity of the temperatures (which we assume throughout). Nonetheless, it can be adapted to the case $T_1 = T_4 = 0$. In this case, the external rotors are not ergodic under the decoupled dynamics: they deterministically slow down and asymptotically reach a given position that depends on the initial condition. Therefore, the conclusion of Lemma 4.3.11 does not hold. However, the counterterm R_{23} that we obtained still produces the desired effect. Indeed, at zero temperature, the definition of U_1 becomes

$$U_1(x) = -\int_0^\infty p_1(t) W_{\rm L}(q_2(t) - q_1(t)) w_{\rm C}(q_3(t) - q_2(t)) ,$$

where $x(t)$ is the deterministic solution given in (4.3.20) and (4.3.21) with initial condition x and $T_1 = T_4 = 0$. Since $p_1(t)$ decreases exponentially fast and $W_{\rm L} w_{\rm C}$ is bounded, this integral still converges. A similar argument applies to U_4.

4.4. Constructing a global Lyapunov function

We construct here the Lyapunov function of Theorem 4.1.4. We start by fixing the parameters defining the sets $\Omega_2, \Omega_3, \Omega_c$ and the functions V_2, V_3, V_c.

We assume throughout this section that θ is fixed and satisfies

$$0 < \theta < \min\left(\frac{1}{T_1}, \frac{1}{T_4}\right) . \tag{4.4.1}$$

This condition is necessary to apply Proposition 4.2.4 and Proposition 4.3.5. In addition, it guarantees that when $p_1^2 + p_4^2$ is large, $\exp(\theta H)$ decreases very fast:

Lemma 4.4.1. *There are constants $C_7, C_8 > 0$ such that*

$$Le^{\theta H} \le (C_7 - C_8(p_1^2 + p_4^2))e^{\theta H} .$$

Proof. Since $Le^{\theta H} = \sum_{b=1,4} \left(-\gamma_b\theta(1-\theta T_b)p_b^2 + \gamma_b\theta T_b\right) e^{\theta H}$, the result follows from the condition on θ. $\square$

We next choose the constants k, ℓ, a, m, and finally R. First, we fix k large enough, and require a lower bound R_0 on R so that the conclusions of Proposition 4.2.4 hold on $\Omega_j(k, R)$, $j = 2, 3$. We then fix the parameters a (appearing in V_2, V_3) and ℓ such that

$$\frac{2}{\ell} < a < \frac{2}{k} \ . \tag{4.4.2}$$

As a consequence, $\Omega_c(\ell, m, R)$ now depends only on m and R, which we fix large enough so that Proposition 4.3.5 applies, and so that $m > \ell$ and $R \geq R_0$.

This choice satisfies the condition $1 \leq k < \ell < m$ imposed in (4.1.8). This ensures that the sets Ω_j ($j = 2, 3$) and Ω_c have "large" intersections, and that they indeed look as shown in Figure 4.4 and Figure 4.5. Moreover, condition (4.4.2) ensures that for large $|p_j|$, $j = 2, 3$,

$$|p_j|^{2/\ell} \ll |p_j|^a \ll |p_j|^{2/k} \ ,$$

which will be crucial.

We next introduce smooth cutoff functions for the sets $\Omega_2, \Omega_3, \Omega_c$. For this, we consider for each set a thin "boundary layer" included in the set itself.

Definition 4.4.2. *Let $\mathcal{P}$ be a subset of the momentum space $\mathbb{R}^4$. We define $\mathcal{B}(\mathcal{P}) = \{p \in \mathcal{P} : \mathrm{dist}(p, \mathcal{P}^c) < 1\}$.*

Lemma 4.4.3. *Let $\mathcal{P} \subset \mathbb{R}^4$. Then, there is a smooth function $\psi : \mathbb{R}^4 \to [0, 1]$ with the following properties. First, $\psi(p) = 1$ on $\mathcal{P} \setminus \mathcal{B}(\mathcal{P})$ and $\psi(p) = 0$ on $\mathcal{P}^c$, with some interpolation on $\mathcal{B}(\mathcal{P})$. Secondly, $\partial^{\underline{a}}\psi$ is bounded on $\mathbb{R}^4$ for each multi-index $\underline{a}$.*

Proof. Such a function is obtained by appropriately regularizing the characteristic function of the set $\{p \in \mathcal{P} : \mathrm{dist}(p, \mathcal{P}^c) > 1/2\} \subset \mathbb{R}^4$. $\square$

Since the definition of sets Ω_c and Ω_j, $j = 2, 3$, involves only the momenta, we can write $\Omega_c = \mathbb{T}^4 \times \mathcal{P}_c$ and $\Omega_j = \mathbb{T}^4 \times \mathcal{P}_j$ for some sets $\mathcal{P}_c, \mathcal{P}_j \subset \mathbb{R}^4$. We apply Lemma 4.4.3 to $\mathcal{P}_c, \mathcal{P}_2$ and $\mathcal{P}_3$, and denote by ψ_c, ψ_2, and ψ_3 the functions obtained. We introduce also the sets

$$\mathcal{B}(\Omega_c) = \mathbb{T}^4 \times \mathcal{B}(\mathcal{P}_c), \quad \mathcal{B}(\Omega_2) = \mathbb{T}^4 \times \mathcal{B}(\mathcal{P}_2), \quad \mathcal{B}(\Omega_3) = \mathbb{T}^4 \times \mathcal{B}(\mathcal{P}_3) \ .$$

Obviously, $\mathcal{B}(\Omega_c) \subset \Omega_c$ and $\mathcal{B}(\Omega_j) \subset \Omega_j$.

Proof of Theorem 4.1.4. We show that the Lyapunov function

$$V = 1 + e^{\theta H} + \sum_{j=2,3} \psi_j(p)V_j + M\psi_c(p)V_c$$

has the necessary properties, provided that the constant M is large enough. We start by proving (4.1.5). From (4.2.15) and (4.3.8), we immediately obtain the bound

$$1 + e^{\theta H} \le V \le c(\psi_2 e^{|p_2|^a} + \psi_3 e^{|p_3|^a} + \psi_c \cdot (p_2^2 + p_3^2))e^{\theta H} , \tag{4.4.3}$$

which is slightly sharper than (4.1.5).

We next turn to the bound on LV. We introduce the set

$$G = \{x \in \Omega : p_1^2 + p_4^2 < (1 + C_7)/C_8\} , \tag{4.4.4}$$

with C_7, C_8 as in Lemma 4.4.1, so that

$$Le^{\theta H} \le -e^{\theta H} + (1 + C_7)\mathbf{1}_G e^{\theta H} \le -e^{\theta H} + C_9 \mathbf{1}_G e^{\frac{\theta}{2}(p_2^2+p_3^2)} \tag{4.4.5}$$

for some $C_9 > 0$, where we have used that $H \le c + \frac{p_2^2}{2} + \frac{p_3^2}{2}$ on G. Moreover, observe that for $j = 2, 3$, there is a polynomial $z_j(p)$ such that

$$\begin{aligned} L(\psi_j V_j) &= \psi_j L V_j + V_j L\psi_j + 2\sum_{b=1,4} \gamma_b T_b (\partial_{p_b}\psi_j)(\partial_{p_b} V_j) \\ &\le -C_3 \psi_j p_j^{-2} e^{|p_j|^a + \frac{\theta}{2}p_j^2} + V_j L\psi_j + 2\sum_{b=1,4} \gamma_b T_b (\partial_{p_b}\psi_j)(\partial_{p_b} V_j) \\ &\le (-\mathbf{1}_{\Omega_j} C_3 p_j^{-2} + \mathbf{1}_{\mathcal{B}(\Omega_j)} z_j(p)) e^{|p_j|^a + \frac{\theta}{2}p_j^2} , \end{aligned} \tag{4.4.6}$$

where the first inequality follows from (4.2.16) and the second inequality holds because the derivatives of $\psi_j(p)$ have support on $\mathcal{B}(\Omega_j)$, because $|\psi_j - \mathbf{1}_{\Omega_j}| \le \mathbf{1}_{\mathcal{B}(\Omega_j)}$, and because of (4.2.15).

Similarly, using Proposition 4.3.5, we obtain a polynomial $z_c(p)$ such that on Ω

$$L(\psi_c(p)V_c) \le (-\mathbf{1}_{\Omega_c} C_6 + \mathbf{1}_{\mathcal{B}(\Omega_c)} z_c(p)) e^{\frac{\theta}{2}(p_2^2+p_3^2)} . \tag{4.4.7}$$

Combining (4.4.5), (4.4.6) and (4.4.7), we find

$$\begin{aligned} LV \le &-e^{\theta H} - \sum_{j=2,3} \mathbf{1}_{\Omega_j} C_3 p_j^{-2} e^{|p_j|^a + \frac{\theta}{2}p_j^2} - \mathbf{1}_{\Omega_c} M C_6 e^{\frac{\theta}{2}(p_2^2+p_3^2)} \\ &+ C_9 \mathbf{1}_G e^{\frac{\theta}{2}(p_2^2+p_3^2)} + \sum_{j=2,3} \mathbf{1}_{\mathcal{B}(\Omega_j)} z_j(p) e^{|p_j|^a + \frac{\theta}{2}p_j^2} + M\mathbf{1}_{\mathcal{B}(\Omega_c)} z_c(p) e^{\frac{\theta}{2}(p_2^2+p_3^2)} . \end{aligned} \tag{4.4.8}$$

The first line contains the "good" terms. We next show that these terms dominate the others. Let $\varepsilon > 0$. We claim that there is a (large) compact set K (which depends on ε) such that

$$\mathbf{1}_{\mathcal{B}(\Omega_j)} z_j(p) e^{|p_j|^a + \frac{\theta}{2}p_j^2} \le \varepsilon e^{\theta H} + c\mathbf{1}_K , \quad j = 2, 3 , \tag{4.4.9}$$

$$\mathbf{1}_{\mathcal{B}(\Omega_c)} z_c(p) e^{\frac{\theta}{2}(p_2^2+p_3^2)} \le \varepsilon e^{\theta H} + \varepsilon \sum_{j=2,3} \mathbf{1}_{\Omega_j} p_j^{-2} e^{|p_j|^a + \frac{\theta}{2}p_j^2} + c\mathbf{1}_K , \tag{4.4.10}$$

$$\mathbf{1}_G e^{\frac{\theta}{2}(p_2^2+p_3^2)} \le \mathbf{1}_{\Omega_c} e^{\frac{\theta}{2}(p_2^2+p_3^2)} + \varepsilon \sum_{j=2,3} \mathbf{1}_{\Omega_j} p_j^{-2} e^{|p_j|^a + \frac{\theta}{2}p_j^2} + c\mathbf{1}_K \ . \tag{4.4.11}$$

We prove these bounds one by one.

- Proof of (4.4.9). We prove the bound for $j = 2$. First observe that

$$z_2(p)e^{|p_2|^a+\frac{\theta}{2}p_2^2-\theta H} < cz_2(p)e^{|p_2|^a-\frac{\theta}{2}(p_1^2+p_3^2+p_4^2)} \ . \tag{4.4.12}$$

By the definition of Ω_2, when $\|p\| \to \infty$ in $\mathcal{B}(\Omega_2)$, we find $p_1^2 + p_3^2 + p_4^2 \sim |p_2|^{2/k} \gg |p_2|^a$ (recalling that $2/k > a$). Thus, the right-hand side of (4.4.12) vanishes in this limit, since z_2 is only a polynomial. This implies (4.4.9) if K is large enough.

- Proof of (4.4.10). By inspection of the definition (4.1.10) of Ω_c, there are three regions $B(\Omega_c)_i$, $i = 1, 2, 3$, such that if the compact set K is large enough,

$$\mathcal{B}(\Omega_c) \subset K \cup \mathcal{B}(\Omega_c)_1 \cup \mathcal{B}(\Omega_c)_2 \cup \mathcal{B}(\Omega_c)_3 \ ,$$

where $\mathcal{B}(\Omega_c)_1$ is such that $p_2^2 + p_3^2 \sim (p_1^2 + p_4^2)^m$, where $\mathcal{B}(\Omega_c)_2 \subset \Omega_2$ is such that $p_3^{2\ell} \sim p_2^2$, and where $\mathcal{B}(\Omega_c)_3 \subset \Omega_3$ is such that $p_2^{2\ell} \sim p_3^2$ (see Figure 4.4 and Figure 4.5).

Since $z_c(p)e^{\frac{\theta}{2}(p_2^2+p_3^2)}$ is bounded on compact sets, (4.4.10) trivially holds on K. We next turn to $\mathcal{B}(\Omega_c)_1$. We have

$$z_c(p)e^{\frac{\theta}{2}(p_2^2+p_3^2)-\theta H} < cz_c(p)e^{-\frac{\theta}{2}(p_1^2+p_4^2)} \ .$$

The right-hand side vanishes when $\|p\| \to \infty$ in $\mathcal{B}(\Omega_c)_1$, and thus by enlarging K if necessary, we find $z_c(p)e^{\frac{\theta}{2}(p_2^2+p_3^2)} \le \varepsilon e^{\theta H} + c\mathbf{1}_K$ on $\mathcal{B}(\Omega_c)_1$, which implies (4.4.10).

Now, consider $\mathcal{B}(\Omega_c)_2$. We have

$$\frac{z_c(p)e^{\frac{\theta}{2}(p_2^2+p_3^2)}}{p_2^{-2}e^{|p_2|^a+\frac{\theta}{2}p_2^2}} = z_c(p)p_2^2 e^{\frac{\theta}{2}p_3^2-|p_2|^a} \ .$$

As $\|p\| \to \infty$ in $\mathcal{B}(\Omega_c)_2$, the right-hand side vanishes, since $p_3^2 \sim |p_2|^{2/\ell}$ and $a > 2/\ell$. Therefore, $z_c(p)e^{\frac{\theta}{2}(p_2^2+p_3^2)} \le \varepsilon p_2^{-2}e^{|p_2|^a+\frac{\theta}{2}p_2^2} + c\mathbf{1}_K$ on $\mathcal{B}(\Omega_c)_2$ for large enough K, and thus (4.4.10) holds on $\mathcal{B}(\Omega_c)_2$ since $\mathcal{B}(\Omega_c)_2 \subset \Omega_2$.

A similar argument applies for $\mathcal{B}(\Omega_c)_3$, which completes the proof of (4.4.10).

- Proof of (4.4.11). Observe that for K large enough, the set G defined in (4.4.4) verifies

$$G \subset K \cup \Omega_2 \cup \Omega_c \cup \Omega_3 \ .$$

On K and Ω_c, (4.4.11) holds trivially. On $G \cap \Omega_2 \setminus \Omega_c$ and for large enough $\|p\|$, we have $p_3^2 \le |p_2|^{2/\ell}$ (otherwise we would have $x \in \Omega_c$), and therefore

$$\frac{e^{\frac{\theta}{2}(p_2^2+p_3^2)}}{p_2^{-2}e^{|p_2|^a+\frac{\theta}{2}p_2^2}} = p_2^2 e^{\frac{\theta}{2}p_3^2-|p_2|^a} \le p_2^2 e^{\frac{\theta}{2}|p_2|^{2/\ell}-|p_2|^a} \ .$$

Since $a > 2/\ell$, and by enlarging K if necessary, we have $e^{\frac{\theta}{2}(p_2^2+p_3^2)} \leq \varepsilon p_2^{-2} e^{|p_2|^a+\frac{\theta}{2}p_2^2} + c\mathbf{1}_K$ on $G \cap \Omega_2 \setminus \Omega_c$, so that (4.4.11) holds on this set. Since a similar argument applies in $G \cap \Omega_3 \setminus \Omega_c$, the proof of (4.4.11) is complete.

Substituting (4.4.9), (4.4.10), and (4.4.11) into (4.4.8), we find

$$\begin{aligned} LV \leq &-(1-2\varepsilon-M\varepsilon)e^{\theta H} - \sum_{j=2,3} \mathbf{1}_{\Omega_j}(C_3 - C_9\varepsilon - M\varepsilon)p_j^{-2}e^{|p_j|^a+\frac{\theta}{2}p_j^2} \\ &- \mathbf{1}_{\Omega_c}(MC_6 - C_9)e^{\frac{\theta}{2}(p_2^2+p_3^2)} + c\mathbf{1}_K \ . \end{aligned}$$

Since the constants C_i do not depend on ε and M, we can make the three parentheses $(1-2\varepsilon-M\varepsilon)$, $(C_3 - C_9\varepsilon - M\varepsilon)$ and $(MC_6 - C_9)$ positive by choosing M large enough and then ε small enough. Using again (4.2.15) and (4.3.8), we finally obtain

$$LV \leq -ce^{\theta H} - c\sum_{j=2,3}\mathbf{1}_{\Omega_j}\frac{V_j}{p_j^2} - c\mathbf{1}_{\Omega_c}\frac{V_c}{p_2^2+p_3^2} + c\mathbf{1}_K \ . \tag{4.4.13}$$

We now show that this implies (4.1.6). Observe that since $V \geq e^{\theta H}$, we have $\log V \geq \theta H$, and therefore, for $j = 2, 3$,

$$p_j^2 \leq p_2^2 + p_3^2 \leq cH + c \leq c\log V + c \leq c(\log V + 2) \ .$$

Since also $-e^{\theta H} \leq -e^{\theta H}/(2+\log V)$, we obtain by (4.4.13) that

$$LV \leq -c\frac{e^{\theta H} + \mathbf{1}_{\Omega_2}V_2 + \mathbf{1}_{\Omega_3}V_3 + \mathbf{1}_{\Omega_c}V_c}{2+\log V} + c\mathbf{1}_K = -\frac{cV}{2+\log(V)} + c\mathbf{1}_K \ ,$$

which, by the definition (4.1.7) of φ, proves (4.1.6). □

4.5. Proof of the main theorem

Now that we have a Lyapunov function (Theorem 4.1.4), we can prove Theorem 4.1.3 in the spirit of of Chapter 3. In addition to Theorem 4.1.4, we need a few other ingredients.

We first use the result of [16] about subgeometric ergodicity. We state it here in a simplified form. For a definition of "irreducible skeleton" and "petite set", see the introduction of [16] or §3.2 in the present thesis.

Theorem 4.5.1 (Douc-Fort-Guillin (2009))**.** *Assume that a skeleton of the process (4.1.2) is irreducible and let $V : \Omega \to [1, \infty)$ be a smooth function with $\lim_{\|p\|\to\infty} V(q, p) = +\infty$. If there are a petite set K and a constant C such that $LV \leq C\mathbf{1}_K - \varphi(V)$ for some differentiable, concave and increasing function $\varphi : [1, \infty) \to (0, \infty)$, then the process admits a unique invariant measure π, and for any $z \in [0, 1]$, there exists a constant C' such that for all $t \geq 0$ and all $x \in \Omega$,*

$$\|P^t(x, \cdot\,) - \pi\|_{(\varphi\circ V)^z} \leq g(t)C'V(x) \ , \tag{4.5.1}$$

where $g(t) = (\varphi \circ H_\varphi^{-1}(t))^{z-1}$, *with* $H_\varphi(u) = \int_1^u \frac{\mathrm{d}s}{\varphi(s)}$.

Proof. This is a combination of [16, Theorems 3.2 and 3.4] for the following "inverse Young's functions" (in the language of [16]): $\Psi_1(s) \propto s^{z-1}$ and $\Psi_2(s) \propto s^z$. $\square$

Theorem 4.1.4 provides most of the input to Theorem 4.5.1, but we still need to check that there is an irreducible skeleton, and that the set K of Theorem 4.1.4 is petite. To this end, we introduce, as in Chapter 3,

Proposition 4.5.2. *The following holds.*

(i) The transition probabilities $P^t(x, \mathrm{d}y)$ *have a* $\mathcal{C}^\infty((0,\infty) \times \Omega \times \Omega)$ *density* $p_t(x,y)$ *and the process is strong Feller.*

(ii) The time-1 skeleton chain $(x_n)_{n=0,1,2,\dots}$ *admits the Lebesgue measure on* $(\Omega, \mathcal{B})$ *as a maximal irreducibility measure.*

(iii) All compact subsets of Ω *are petite.*

Proof. (i) follows from Hörmander's condition. The proof that Hörmander's condition holds, which relies on Assumption 4.1.2, is very similar to that of Lemma 3.5.3 and is left to the reader. The proof of (ii) is exactly as Lemma 3.5.6, and (iii) follows from (i), (ii), and Proposition 6.2.8 of [43]. $\square$

We can now finally give the

Proof of Theorem 4.1.3. Let $0 \le \theta_1 < \min(1/T_1, 1/T_4)$ and $\theta_2 > \theta_1$. Choose now $\theta \in (\theta_1, \theta_2)$ such that $\theta < \min(1/T_1, 1/T_4)$. By Theorem 4.1.4, we have a Lyapunov function $1 + e^{\theta H} \le V \le c(e^{|p_2|^a} + e^{|p_3|^a})e^{\theta H}$ with $a \in (0,1)$ such that $LV \le c\mathbf{1}_K - \varphi(V)$, where $\varphi(s) = c_3\, s/(2 + \log(s))$, and where K is a compact (and therefore petite) set. Let now $z \in (0,1)$ be such that $z\theta > \theta_1$. By Theorem 4.5.1, we obtain the existence of a unique invariant measure π such that

$$\|P^t(x,\,\cdot\,) - \pi\|_{(\varphi\circ V)^z} \le c e^{-\lambda t^{1/2}} V(x)\;, \tag{4.5.2}$$

where we have used that with the notation of Theorem 4.5.1,

$$g(t) = (\varphi \circ H_\varphi^{-1}(t))^{z-1} \le c e^{-\lambda t^{1/2}} \tag{4.5.3}$$

for some $\lambda > 0$. Indeed, $H_\varphi(u) = \frac{1}{c_3}\int_1^u \frac{2+\log s}{s}\mathrm{d}s = \frac{1}{2c_3}(\log u)^2 + \frac{2}{c_3}\log u$, so that $H_\varphi^{-1}(t) = \exp((2c_3 t + 4)^{1/2} - 2)$ and $(\varphi \circ H_\varphi^{-1}(t)) = (2c_3 t + 4)^{-1/2}\exp((2c_3 t + 4)^{1/2} - 2) \ge c e^{ct^{1/2}}$, which implies (4.5.3).

Then, (4.1.4) follows from (4.5.2) and the following two observations. First, we have $V \le c e^{\theta_2 H}$ since $\theta < \theta_2$. Secondly, by our choice of z, we have $e^{\theta_1 H} \le c(\varphi \circ V)^z$, so that $\|P^t(x,\,\cdot\,) - \pi\|_{e^{\theta_1 H}} \le c\,\|P^t(x,\,\cdot\,) - \pi\|_{(\varphi\circ V)^z}$.

Thus, we have proved (iii). Since (i) and the smoothness assertion in (ii) follow from Proposition 4.5.2, the proof is complete. $\square$

Remark 4.5.3. It would of course be desirable to generalize Theorem 4.1.3 to longer chains of rotors. The proof of Proposition 4.5.2 carries on unchanged to chains of arbitrary length. Therefore, in order to prove the existence of a steady state and obtain a convergence rate (with Theorem 4.5.1), it "suffices" to find an appropriate Lyapunov function. We expect the convergence rate to be limited by the central rotor (if the length of the chain is odd) or the two central rotors (if the length is even). Preliminary studies indicate that for chains of length n, a convergence rate $\exp(-ct^k)$ with $k = 1/(2\lceil n/2\rceil - 2)$ is to be expected. Obtaining such a result raises some major technical difficulties. First, the averaging procedure has to be carried to much higher orders, which quickly becomes intractable if we proceed explicitly, as we do here. Moreover, the number of regimes to consider grows very rapidly with n. And finally, some generalization of Proposition 4.3.12 to more general (nonlinear) decoupled systems will be needed, with the difficulties mentioned in Remark 4.3.13. We are trying to solve these issues by developing a inductive method which requires fewer explicit calculations, but much work remains to be done.

4.6. Appendix: Resonances in the deterministic case

In §4.2, two resonant terms appeared, namely $p_1 W_{\mathrm{L}} w_{\mathrm{C}}/p_2^2$ and $-p_4 W_{\mathrm{R}} w_{\mathrm{C}}/p_3^2$. These terms have a physical meaning. We start with the case where $W_{\mathrm{I}}(s) = -\cos(s)$, $\mathrm{I} = \mathrm{L}, \mathrm{C}, \mathrm{R}$. Then,

$$W_{\mathrm{L}} w_{\mathrm{C}} = -\cos(q_2 - q_1)\sin(q_3 - q_2) = \frac{\sin(q_1 - q_3)}{2} + \frac{\sin(2q_2 - q_3 - q_1)}{2} . \tag{4.6.1}$$

Consider now the regime where most of the energy is concentrated at sites 2 and 3. In the approximate dynamics (4.3.1), we see that $\sin(q_1 - q_3)$ oscillates with frequency $p_3/2\pi$ and mean zero, while $\sin(2q_2 - q_3 - q_1)$ oscillates with frequency $(2p_2 - p_3)/2\pi$. When $p_3 = 2p_2$, the second term does not oscillate.

In Figure 4.8, we represent some trajectories projected onto the p_2p_3-plane in the deterministic case (*i.e.,* $T_1 = T_4 = 0$). We observe that some trajectories are "trapped" by the line $p_3/p_2 = 2$, while some others just cross it. By symmetry, the same happens when $p_3/p_2 = 1/2$ because of the term $-p_4 W_{\mathrm{R}} w_{\mathrm{C}}/p_3^2$. This phenomenon does not occur when the same conditions are used with positive temperatures (see Figure 4.3). A finer analysis (not detailed here) shows that in the resonant regime $p_3/p_2 = 2$, a net momentum flux from p_3 to p_2 appears, and similarly for $p_3/p_2 = 1/2$ with a flux from p_2 to p_3. These fluxes stabilize the resonant regimes.

If we take $W_{\mathrm{I}}(s) = -\cos(n_{\mathrm{I}} s)$ for some $n_{\mathrm{I}} \in \mathbb{Z}_*$, $\mathrm{I} = \mathrm{L}, \mathrm{C}, \mathrm{R}$, we find by a decomposition similar to (4.6.1) some resonances at

$$\frac{p_3}{p_2} \in \left\{ \frac{n_{\mathrm{C}} + n_{\mathrm{L}}}{n_{\mathrm{C}}}, \frac{n_{\mathrm{C}} - n_{\mathrm{L}}}{n_{\mathrm{C}}}, \frac{n_{\mathrm{C}}}{n_{\mathrm{C}} + n_{\mathrm{R}}}, \frac{n_{\mathrm{C}}}{n_{\mathrm{C}} - n_{\mathrm{R}}} \right\} .$$

(If some of these values are 0 or ∞, we exclude them since our approximation is reasonable when both $|p_2|$ and $|p_3|$ are very large.) For example, if we choose $(n_{\mathrm{L}}, n_{\mathrm{C}}, n_{\mathrm{R}}) = (3, 1, 3)$, we obtain the ratios $p_3/p_2 = 4, 1/4, -2, -1/2$, which we indeed observe in Figure 4.9.

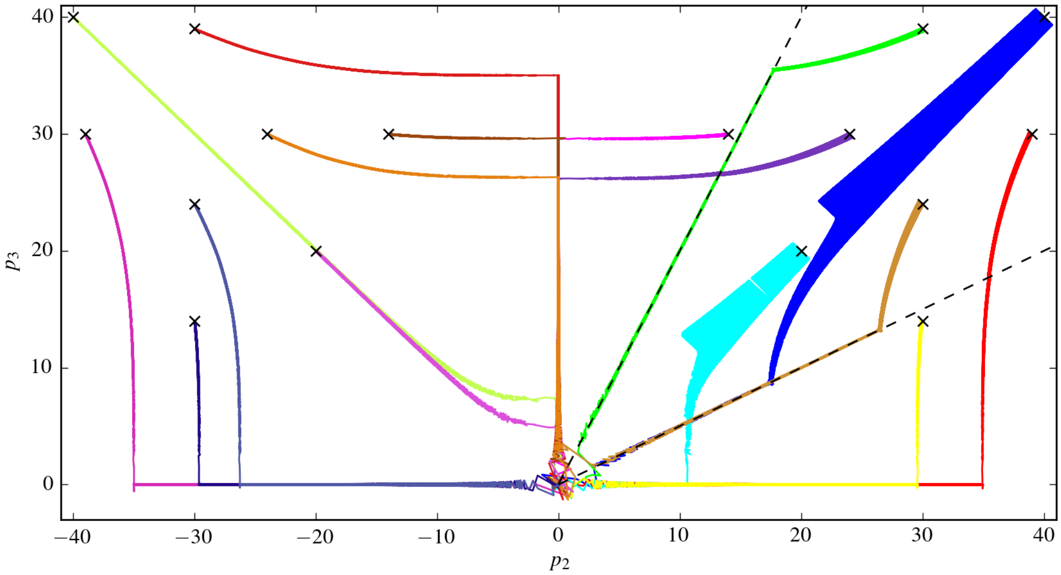


Figure 4.8 – Projection of a few orbits on the p_2p_3-plane, with $W_{\rm L} = W_{\rm C} = W_{\rm R} = -\cos$, $\gamma_1 = \gamma_4 = 1$, $T_1 = T_4 = 0$. The resonances are depicted as dashed lines.

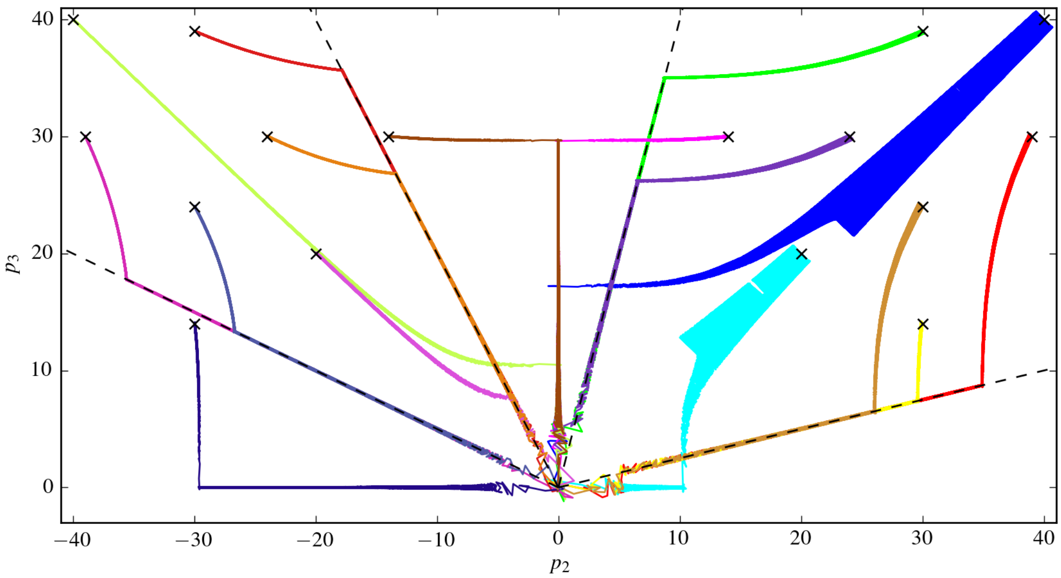


Figure 4.9 – Projection of a few orbits on the p_2p_3-plane, with $\gamma_1 = \gamma_4 = 1$, $T_1 = T_4 = 0$, and $(n_{\rm L}, n_{\rm C}, n_{\rm R}) = (3, 1, 3)$. The resonances are depicted as dashed lines.

Of course, a similar analysis applies to more general interaction potentials by taking their Fourier series and treating the (products of) modes separately.

Bibliography

[1] K. Aoki, J. Lukkarinen, and H. Spohn. Energy transport in weakly anharmonic chains. *J. Stat. Phys.* **124** (2006), 1105–1129.

[2] R. Balescu. Equilibrium and nonequilibrium statistical mechanics. *NASA STI/Recon Technical Report A* **76** (1975), 32809.

[3] L. Bellet. Ergodic properties of markov processes. *Open Quantum Systems II* (2006), 1–39.

[4] C. Bernardin, F. Huveneers, J. L. Lebowitz, C. Liverani, and S. Olla. Green-kubo formula for weakly coupled systems with noise. *Commun. Math. Phys.* **334** (2015), 1377–1412.

[5] C. Bernardin and S. Olla. Fourier's law for a microscopic model of heat conduction. *J. Stat. Phys.* **121** (2005), 271–289.

[6] F. Bonetto, J. L. Lebowitz, and J. Lukkarinen. Fourier's law for a harmonic crystal with self-consistent stochastic reservoirs. *J. Stat. Phys.* **116** (2004), 783–813.

[7] F. Bonetto, J. L. Lebowitz, and L. Rey-Bellet. Fourier's law: a challenge to theorists. In: *Mathematical physics 2000* (London: Imp. Coll. Press, 2000), pp. 128–150.

[8] J. Bricmont and A. Kupiainen. Towards a derivation of Fourier's law for coupled anharmonic oscillators. *Commun. Math. Phys.* **274** (2007), 555–626.

[9] P. Carmona. Existence and uniqueness of an invariant measure for a chain of oscillators in contact with two heat baths. *Stochastic Process. Appl.* **117** (2007), 1076–1092.

[10] R. Chas'minskij. *Stochastic stability of differential equations* (Springer, 2012).

[11] N. Cuneo and J.-P. Eckmann. Controlling general polynomial networks. *Commun. Math. Phys.* **328** (2014), 1255–1274.

[12] N. Cuneo and J.-P. Eckmann. Non-equilibrium steady states for chains of four rotors. *Commun. Math. Phys.* **345** (2016), 185–221.

[13] N. Cuneo, J.-P. Eckmann, and C. Poquet. Non-equilibrium steady state and subgeometric ergodicity for a chain of three coupled rotors. *Nonlinearity* **28** (2015), 2397–2421.

[14] N. Cuneo and C. Poquet. On the relaxation rate of short chains of rotors interacting with langevin thermostats (2016), arXiv:1604.03408 [math–ph].

[15] W. De Roeck and F. Huveneers. Asymptotic localization of energy in nondisordered oscillator chains. *Comm. Pure Appl. Math.* **68** (2015), 1532–1568.

[16] R. Douc, G. Fort, and A. Guillin. Subgeometric rates of convergence of f-ergodic strong Markov processes. *Stochastic Process. Appl.* **119** (2009), 897–923.

[17] A. Dymov. Nonequilibrium statistical mechanics of Hamiltonian rotators with alternated spins. *J. Stat. Phys.* **158** (2015), 968–1006.

[18] A. Dymov. Nonequilibrium statistical mechanics of weakly stochastically perturbed system of oscillators. *Annales Henri Poincaré* (2015), first online, DOI:10.1007/s00023–015–0441–x.

[19] J.-P. Eckmann and M. Hairer. Non-equilibrium statistical mechanics of strongly anharmonic chains of oscillators. *Commun. Math. Phys.* **212** (2000), 105–164.

[20] J.-P. Eckmann and M. Hairer. Spectral properties of hypoelliptic operators. *Commun. Math. Phys.* **235** (2003), 233–253.

[21] J.-P. Eckmann, M. Hairer, and L. Rey-Bellet. Non-equilibrium steady states for networks of springs, in preparation.

[22] J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet. Entropy production in nonlinear, thermally driven Hamiltonian systems. *J. Stat. Phys.* **95** (1999), 305–331.

[23] J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet. Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures. *Commun. Math. Phys.* **201** (1999), 657–697.

[24] J.-P. Eckmann and L.-S. Young. Temperature profiles in Hamiltonian heat conduction. *Europhys. Lett.* **68** (2004), 790–796.

[25] J.-P. Eckmann and L.-S. Young. Nonequilibrium energy profiles for a class of 1-D models. *Commun. Math. Phys.* **262** (2006), 237–267.

[26] J.-P. Eckmann and E. Zabey. Strange heat flux in (an)harmonic networks. *J. Stat. Phys.* **114** (2004), 515–523.

[27] M. Eleftheriou, S. Lepri, R. Livi, and F. Piazza. Stretched-exponential relaxation in arrays of coupled rotators. *Phys. D* **204** (2005), 230–239.

[28] G. W. Ford, M. Kac, and P. Mazur. Statistical mechanics of assemblies of coupled oscillators. *J. Mathematical Phys.* **6** (1965), 504–515.

[29] G. Gallavotti, A. Iacobucci, and S. Olla. Nonequilibrium stationary state for a damped rotator (2013), arXiv:1310.5379 [cond–mat.stat–mech].

[30] O. V. Gendelman and A. V. Savin. Normal heat conductivity of the one-dimensional lattice with periodic potential of nearest-neighbor interaction. *Phys. Rev. Lett.* **84** (2000), 2381–2384.

[31] C. Giardinà, R. Livi, A. Politi, and M. Vassalli. Finite thermal conductivity in 1D lattices. *Physical review letters* **84** (2000), 2144.

[32] M. Hairer. A probabilistic argument for the controllability of conservative systems (2005), arXiv:math–ph/0506064.

[33] M. Hairer. How hot can a heat bath get? *Commun. Math. Phys.* **292** (2009), 131–177.

[34] M. Hairer and J. C. Mattingly. Slow energy dissipation in anharmonic oscillator chains. *Comm. Pure Appl. Math.* **62** (2009), 999–1032.

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

[36] L. Hörmander. *The Analysis of Linear Partial Differential Operators I–IV* (New York: Springer, 1985).

[37] A. Iacobucci, F. Legoll, S. Olla, and G. Stoltz. Negative thermal conductivity of chains of rotors with mechanical forcing. *Phys. Rev. E* **84** (2011), 061108.

[38] V. Jurdjevic. *Geometric control theory* (Cambridge University Press, 1997).

[39] R. Kubo, M. Toda, and N. Hashitsume. *Statistical physics II: nonequilibrium statistical mechanics*, volume 31 (Springer Science & Business Media, 2012).

[40] M. Le Bellac. Non equilibrium statistical mechanics. Lecture (2007).

[41] R. Lefevere and A. Schenkel. Normal heat conductivity in a strongly pinned chain of anharmonic oscillators. *Journal of Statistical Mechanics: Theory and Experiment* **2006** (2006), L02001.

[42] S. Lepri, R. Livi, and A. Politi. Thermal conduction in classical low-dimensional lattices. *Phys. Rep.* **377** (2003), 1–80.

[43] S. Meyn and R. L. Tweedie. *Markov chains and stochastic stability* (Cambridge University Press, 2009), second edition.

[44] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes III: Foster-Lyapunov criteria for continuous-time processes. *Adv. Appl. Probab.* **25** (1993), 518–548.

[45] E. Pardoux and A. Y. Veretennikov. On the Poisson equation and diffusion approximation. I. *Ann. Probab.* **29** (2001), 1061–1085.

[46] E. Pardoux and A. Y. Veretennikov. On Poisson equation and diffusion approximation. II. *Ann. Probab.* **31** (2003), 1166–1192.

[47] E. Pardoux and A. Y. Veretennikov. On the Poisson equation and diffusion approximation. III. *Ann. Probab.* **33** (2005), 1111–1133.

[48] L. Rey-Bellet. Fluctuations of the entropy production in nonequilibrium steady states (2011), lecture slides.

[49] L. Rey-Bellet and L. E. Thomas. Asymptotic behavior of thermal nonequilibrium steady states for a driven chain of anharmonic oscillators. *Commun. Math. Phys.* **215** (2000), 1–24.

[50] L. Rey-Bellet and L. E. Thomas. Exponential convergence to non-equilibrium stationary states in classical statistical mechanics. *Commun. Math. Phys.* **225** (2002), 305–329.

[51] L. Rey-Bellet and L. E. Thomas. Fluctuations of the entropy production in anharmonic chains. *Ann. Henri Poincaré* **3** (2002), 483–502.

[52] J. A. Sanders, F. Verhulst, and J. Murdock. *Averaging methods in nonlinear dynamical systems*, volume 59 of *Applied Mathematical Sciences* (Springer, New York, 2007), second edition.

[53] H. Spohn. Kinetic equations from Hamiltonian dynamics: Markovian limits. *Rev. Modern Phys.* **52** (1980), 569–615.

[54] D. W. Stroock and S. R. S. Varadhan. On the support of diffusion processes with applications to the strong maximum principle. In: *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. III: Probability theory* (Univ. California Press, Berkeley, Calif., 1972).

[55] P. A. Vela. Averaging and control of nonlinear systems. Ph.D. thesis, California Institute of Technology (2003).

[56] C. Villani. Hypocoercivity. *Memoirs A.M.S.* **202** (2009), 1–184.