Exponential Convergence to Non-Equilibrium Stationary States in Classical Statistical Mechanics

Luc Rey-Bellet[1], Lawrence E. Thomas[2]

Department of Mathematics, University of Virginia
Kerchof Hall, Charlottesville VA 22903, USA

Abstract

We continue the study of a model for heat conduction [6] consisting of a chain of non-linear oscillators coupled to two Hamiltonian heat reservoirs at different temperatures. We establish existence of a Liapunov function for the chain dynamics and use it to show exponentially fast convergence of the dynamics to a unique stationary state. Ingredients of the proof are the reduction of the infinite dimensional dynamics to a finite-dimensional stochastic process as well as a bound on the propagation of energy in chains of anharmonic oscillators.

1 Introduction

In its present state, non-equilibrium statistical mechanics is lacking the firm theoretical foundations that equilibrium statistical mechanics has. This is due, perhaps, to the extremely great variety of physical phenomena that non-equilibrium statistical mechanics describes. We will concentrate here on a system which is maintained, by suitable forces, in a state far from equilibrium. In such an idealization, the non-equilibrium phenomena, can be described by stationary non-equilibrium states (SNS), which are the analog of canonical or microcanonical states of equilibrium.

Recently many works have been devoted to the rigorous study of SNS. Two main streams are emerging. In the first approach, for *open systems*, a system is driven out of equilibrium by interacting with several reservoirs at different temperatures. In the second approach, for *thermostated systems* , a system is driven out of equilibrium by non-Hamiltonian forces and constrained to a compact energy surface by Gaussian (or others) thermostats [9, 23]. One should view both approaches as two different idealizations of

[1]Email: lr7q@virginia.edu.

[2]Email: let@virginia.edu.

Supported in part by NSF Grant 980139

the same physical situation, in the same spirit as the equivalence of ensembles in equilibrium statistical mechanics. But for the moment, the extent to which both approaches are equivalent remains a largely open problem.

We consider here an open system, a model of heat conduction consisting of a finite-dimensional classical Hamiltonian model, a one-dimensional finite lattice of anharmonic oscillators (referred to as the chain), coupled, at the boundaries only, to two reservoirs of classical non-interacting phonons at positive and different temperatures. We believe this model to be quite realistic, in particular it is completely Hamiltonian and non-linear.

This model goes back (in the linear case) to [8] (see also [22, 25]). First rigorous results for anharmonic models appear [6] and further in [7, 5]. Similar models in classical and quantum mechanics have attracted attention in the last few years, mostly for systems coupled to a single reservoir at zero or positive temperature, i.e., for systems near thermal equilibrium (see e.g. [12, 13, 3, 15, 24]. In our case, with two reservoirs, no Gibbs Ansatz is available and in general, even the very *existence* of a (non-equilibrium) stationary state is a mathematically challenging question which requires a sufficiently deep understanding of the dynamics. For the model at hand, conditions for the *existence* of the SNS have been given in [6] and generalized in [5]. The *uniqueness* of the SNS as well as the strict positivity of *entropy production* (or heat flux) have been proved in [7]. The leading asymptotics of the invariant measure (for low temperatures) are studied in [21] and shown to be described by a variational principle.

Under suitable assumptions on the chain interactions and its interactions with the reservoirs, we establish the existence of a Liapunov function for the chain dynamics. We then use this Liapunov function to establish that the relaxation to the SNS occurs at an *exponential* rate, and finally we prove that the system has a *spectral gap* (using probabilistic techniques developed by Meyn and Tweedie in [18]).

The Hamiltonian of the model has the form

$$H = H_B + H_S + H_I \,. \tag{1}$$

The two reservoirs of free phonons are described by wave equations in $\mathbf{R}^d$ with Hamiltonian

$$\begin{aligned} H_B &= H(\varphi_L, \pi_L) + H(\varphi_R, \pi_R) \,, \\ H(\varphi, \pi) &= \frac{1}{2} \int dx \, (|\nabla \varphi(x)|^2 + |\pi(x)|^2) \,, \end{aligned}$$

where L and R stand for the "left" and "right" reservoirs, respectively. The

Hamiltonian describing the chain of length n is given by

$$\begin{aligned} H_S(p,q) &= \sum_{i=1}^{n} \frac{p_i^2}{2} + V(q_1, \cdots, q_n)\,, \\ V(q) &= \sum_{i=1}^{n} U^{(1)}(q_i) + \sum_{i=1}^{n-1} U^{(2)}(q_i - q_{i+1})\,. \end{aligned}$$

where $(p_i, q_i) \in \mathbf{R}^d \times \mathbf{R}^d$ are the coordinates and momenta of the i^{th} particle of the chain. The phase space of the chain is $\mathbf{R}^{2dn}$. The interaction between the chain and the reservoirs occurs at the boundaries only and is of dipole-type

$$H_I = q_1 \cdot \int dx\, \nabla\varphi_L(x)\rho_L(x) + q_n \cdot \int dx\, \nabla\varphi_R(x)\rho_R(x)\,,$$

where ρ_L and ρ_R are coupling functions ("charge densities") which we will assume spherically symmetric.

Our assumptions on the anharmonic lattice described by $H_S(p,q)$ are the following:

- **H1 Growth at infinity**: The potentials $U^{(1)}(x)$ and $U^{(2)}(x)$ are $\mathcal{C}^\infty$ and grow at infinity like $\|x\|^{k_1}$ and $\|x\|^{k_2}$: There exist constants C_i, D_i, $i = 1, 2$ such that

$$\lim_{\lambda\to\infty} \lambda^{-k_i} U^{(i)}(\lambda x) = a^{(i)} \|x\|^{k_i}\,, \tag{2}$$

$$\lim_{\lambda\to\infty} \lambda^{-k_i+1} \nabla U^{(i)}(\lambda x) = a^{(i)} k_i \|x\|^{k_i-2} x\,, \tag{3}$$

$$\|\partial^2 U^{(i)}(x)\| \le (C_i + D_i V(x))^{1-\frac{2}{k_i}}\,. \tag{4}$$

  where $\|\cdot\|$ in Eq. (4) denotes some matrix-norm.

  Moreover we will assume that

$$k_2 \ge k_1 \ge 2\,,$$

  so that, for large $\|x\|$ the interaction potential $U^{(2)}$ is "stiffer" than the one-body potential $U^{(1)}$. It follows from Eqs. (2) and (3) that the critical set of $V(q)$, i.e, the set $\{q : \nabla V(q) = 0\}$ is a compact set.

- **H2 Non-degeneracy**: The coupling potential between nearest neighbors $U^{(2)}$ is non-degenerate in the following sense. For $x \in \mathbf{R}^d$ and $m = 1, 2, \cdots$, let $A^{(m)}(x) : \mathbf{R}^d \to \mathbf{R}^{d^m}$ denote the linear maps given by

$$(A^{(m)}(x)v)_{l_1 l_2 \cdots l_m} = \sum_{l=1}^{d} \frac{\partial^{m+1} U^{(2)}}{\partial x^{(l_1)} \cdots \partial x^{(l_m)} \partial x^{(l)}}(x) v_l\,.$$

We assume that for each $x \in \mathbf{R}^d$ there exists m_0 such that

$$\mathrm{Rank}(A^{(1)}(x), \cdots A^{(m_0)}(x)) = d\,.$$

The class of coupling functions ρ_i, $i \in \{L, R\}$ we can allow is relatively restrictive:

- **H3 Rationality of the coupling**: Let $\hat{\rho}_i$ denote the Fourier transform of ρ_i. We assume that

  $$|\hat{\rho}_i(k)|^2 \;=\; \frac{1}{Q_i(k^2)}\,,$$

  where Q_i, $i \in \{L, R\}$ are polynomials with real coefficients and no roots on the real axis. In particular, if k_0 is a root of Q_i, then so are $-k_0$, $\overline{k}_0$ and $-\overline{k}_0$.

Under these conditions we have the following result (a more detailed and precise statement will be given in the next section). Let $F(p, q)$ be an observable on the phase space of the chain, for example any function with at most polynomial growth (no smoothness is required). We denote as $(p(t), q(t))$ the solution of the Hamiltonian equation of motion with Hamiltonian (1) and initial conditions (p, q). Of course $(p(t), q(t))$ depends also on the variable of the reservoirs, though only through their initial conditions $(\pi_L, \varphi_L, \pi_R, \varphi_R)$. We introduce the temperature by making the assumption that the initial conditions of the reservoirs are distributed according to thermal equilibrium at temperature T_R and T_L respectively and we denote $\langle \cdot \rangle_{LR}$ as the corresponding average.

Theorem 1.1 *Under Conditions* **H1–H3***, there is a measure $\nu(dp, dq)$ with a smooth everywhere positive density such that the Law of Large Numbers holds:*

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T F(p(t), q(t))dt \;=\; \int F\, d\nu$$

for almost all initial conditions $(\pi_L, \varphi_L, \pi_R, \varphi_R)$ of the reservoirs and for all *initial conditions (p, q) of the chain. Moreover there exist a constant $r > 1$ and a function $C(p, q)$ with $\int C\, d\nu < \infty$ such that*

$$\left| \langle F(p(t), q(t)) \rangle_{LR} - \int F d\nu \right| \;\leq\; C(p, q) r^{-t}\,.$$

for all initial conditions (p, q). That is, if we average over the initial conditions of the reservoirs the convergence is exponential.

Note that the ergodic properties stated in Theorem 1.1 hold not only for ν-almost every initial condition (p,q), but in fact for every (p,q).

The existence of a (unique) stationary state was proved for (exactly solvable) quadratic harmonic potentials $V(q)$ in [25], for $k_1 = k_2 = 2$ (i.e., for potential which are quadratic at infinity) in [6, 7] and generalized to the case $k_2 > k_1 \geq 2$ in [5]. What is really new here is that we prove that the convergence occurs *exponentially fast* and we also weaken slightly the conditions on the potential (in particular the case $k_1 = k_2$ is allowed and our condition **H2** on $U^{(2)}$ is weaker than the one used in [6, 7, 5]). Our methods also differ notably from those used in [6, 5]; in fact we reprove the existence of the SNS (with a shorter and more constructive proof than in [6, 5]) and, at the same time, we prove much stronger ergodic properties.

We devote the rest of this section to a brief discussion of the assumptions **H1** $-$ **H3**. Since the reservoirs are free phonons gases and since we make a statistical assumption on the initial condition of the reservoirs, one can integrate out the variables of the reservoirs yielding random integro-differential equations for the variable (p,q). Our assumption **H3** of rational coupling is, in effect, a Markovian assumption: with such coupling one can eliminate the memory terms by adding a finite number of auxiliary variables to obtain a system of Markovian *stochastic* differential equations on the extended phase space consisting of the dynamical variables (p,q) together with the auxiliary variables. The main (new) ingredient in our proof is then the construction of a *Liapunov function* for the system, which implies, using probabilistic methods developed in [1, 20, 18], the exponential convergence.

To explain the construction of a Liapunov function, note that the dynamics of the chain in the bulk is simply Hamiltonian, while at the boundaries the action of the reservoirs results into two distinct forces. There are *dissipative* forces which correspond to the fact that the energy of the chain dissipates into the reservoirs. This force is *independent* of the temperature. On the other hand since the reservoirs are infinite and at positive temperatures, they exert (random) forces at the boundaries of the chain and these forces turn out to be proportional to the temperatures of the reservoirs.

The construction of the Liapunov function proceeds in two steps. In a first step we neglect completely the random force, only dissipation acts. This corresponds to dynamics at temperature zero, and one can prove that the energy decreases and that the system relaxes to a (local) equilibrium of the Hamiltonian $H(p,q)$. We establish the *rate* at which this relaxation takes place (at sufficiently high energies). In the second step we consider the complete dynamics and we show that for energies which are much higher than the temperatures of the reservoirs, the random force is essentially negligible with respect to the dissipation. This means that except for (exponentially) rare excursions the system spends most of its time in a compact neighborhood of the equilibrium points. On the other hand, in this compact set, i.e., at energies of order of the temperatures of the reservoirs, the dynamics is essentially determined by the fluctuations and to prove exponential convergence to a SNS one has to show that the fluctuations are such that every part of the phase space is visited by the dynamics. To summarize, we control the dynamics at any temperature by the dynamics at zero temperature.

This allows one to understand the meaning of our assumptions on the potential $V(q)$. If we suppose that the energy has an infinite number of local minima tending to infinity, the zero temperature (long-time) dynamics is *not* confined to a compact energy domain and our argument fails. With regard to the condition $k_2 \geq k_1$ in **H1** on the exponents of the potentials, since the results of [26] and the rigorous proofs of [17, 2], it is known that stable (in the sense of Nekhoroshev) localized states exist in non-linear lattices. Consider, for example, an infinite chain of oscillators (without reservoirs). Numerically and in certain cases rigorously [17], one can show the existence of *breathers*, i.e., of solutions which are spatially (exponentially) localized and time-periodic. Although the breathers occurs both for $k_1 > k_2$ and $k_2 \geq k_1$ they behave differently at high energies. For $k_1 > k_2$, the higher the energy, the more localized the breather get (hard breathers), while for $k_2 \geq k_1$, as the the energy gets bigger the breather become less and less localized (soft breathers). In fact a key point of our analysis is to show that at high energy, if the energy E of the initial condition is localized away from the boundary, then after a time of order one, the oscillators at the boundaries carry at least an energy of order E^{2/k_2} so that the chain system energy can relax into the reservoirs.

Although we believe that the existence of a SNS probably may not depend too much on these localization phenomena, the rate of convergence to the SNS presumably does. Our approach of controlling the dynamics by the the zero-temperature dynamics may not be adequate if the condition **H1** fails to hold and so more refined estimates on the dynamics are needed to show that these localized states might be in fact destroyed by the coupling to the reservoirs.

As regards the organization of this paper, Sec. 2 presents the effective stochastic differential equations for the chain, a discussion of allowable interactions between the reservoirs and the chain and a concise statement, Theorem 2.1, of the exponential convergence. In Sec. 3 we discuss the dissipative deterministic system (corresponding to reservoirs at temperature 0), Theorem 3.3, and then we show the extent to which the random paths follow the deterministic ones, Proposition 3.7. We give a lower bound on the random energy dissipation, Corollary 3.8. We then conclude Sec. 3 by providing the Liapunov function, Theorem 3.10, and bounds on the exponential hitting times on (sufficiently large) compact sets, Theorem 3.11. In Sec. 4 we prove the random process has a smooth law and at most one ergodic component, improving slightly results of [6, 7, 5]. Finally in Sec. 5 we conclude the proof of Theorem 2.1 by invoking results of [18] on the ergodic theory of Markov process.

2 Effective Equations

We first give a precise description of the reservoirs and of their coupling to the system and derive the stochastic equations which we will study. A free phonon gas is described by a linear wave equation in $\mathbf{R}^d$, i.e., by the pair of real fields $\phi(x) = (\varphi(x), \pi(x))$, $x \in \mathbf{R}^d$. We define the norm $\|\phi\|$ by $\|\phi\|^2 \equiv \int dx\, |\nabla\phi(x)|^2 + |\pi(x)|^2$ and denote $\langle\cdot,\cdot\rangle$ the corresponding scalar product. The phase space of the reservoirs at finite energy is the real Hilbert space of functions $\phi(x)$ such that the energy $H_B(\phi) = \|\phi\|^2/2$ is finite and the equations of motion are

$$\dot{\phi}(t,x) \;=\; \mathcal{L}\phi(t,x)\,, \quad \mathcal{L} \;=\; \begin{pmatrix} 0 & 1 \\ -\Delta & 0 \end{pmatrix}.$$

In order to describe the coupling of the reservoir to the system, let us consider first a single confined particle in $\mathbf{R}^d$ with Hamiltonian $H_S(p,q) = p^2/2 + V(q)$. As the Hamiltonian for the coupled system particle plus reservoirs, we have

$$\begin{aligned} H(\phi,p,q) &= \frac{1}{2}\|\phi\|^2 + p^2 + V(q) + q \cdot \int dx\, \nabla\varphi(x)\rho(x) \\ &= H_B(\phi) + H_S(p,q) + q\cdot\langle\phi,\alpha\rangle\,, \end{aligned}$$

where $\rho(x)$ is a real rotation invariant function and $\alpha = (\alpha^{(1)},\cdots,\alpha^{(d)})$ is, in Fourier space, given by

$$\hat{\alpha}^{(i)} \;=\; \begin{pmatrix} -ik^{(i)}\hat{\rho}(k)/k^2 \\ 0 \end{pmatrix}.$$

We introduce the covariance matrix $C^{(ij)}(t) = \langle \exp(\mathcal{L}t)\alpha^{(i)}\,,\, \alpha^{(j)}\rangle$. A simple computation shows that

$$C^{(ij)}(t) \;=\; \frac{1}{d}\delta_{ij}\int dk\, |\rho(k)|^2 e^{i|k|(t-s)}\,,$$

and we define a coupling constant λ by putting $\lambda^2 = C^{(ij)}(0) = \int dk|\rho(k)|^2$. The equations of motion of the coupled system are

$$\begin{aligned} \dot{q}(t) &= p(t)\,, \\ \dot{p}(t) &= -\nabla V(q(t)) - \langle \phi, \alpha \rangle\,, \\ \dot{\phi}(t,k) &= \mathcal{L}\left(\phi(t,k) + q(t)\cdot\alpha(k)\right)\,. \end{aligned} \tag{5}$$

With the change of variables $\psi(k) = \phi(k) + q\cdot\alpha(k)$, Eqs. (5) become

$$\begin{aligned} \dot{q}(t) &= p(t)\,, \\ \dot{p}(t) &= -\nabla V_{\text{eff}}(q(t)) - \langle \psi, \alpha \rangle\,, \\ \dot{\psi}(t,k) &= \mathcal{L}\psi(t,k) + p(t)\cdot\alpha(k)\,, \end{aligned} \tag{6}$$

where $V_{\text{eff}}(q) = V(q) - \lambda^2 q^2/2$. Integrating the last of Eqs. (6) with initial condition $\psi_0(k)$ one finds

$$\psi(t,k) \;=\; e^{\mathcal{L}t}\psi_0(k) + \int_0^t ds\, e^{\mathcal{L}(t-s)}\alpha(k)\cdot p(s)\,.$$

and inserting into the second of Eqs. (6) gives

$$\begin{aligned} \dot{q}(t) &= p(t)\,, \\ \dot{p}(t) &= -\nabla V_{\text{eff}}(q(t)) - \int_0^t dsC(t-s)p(s) - \langle \psi_0, e^{-\mathcal{L}t}\alpha\rangle\,. \end{aligned} \tag{7}$$

If we now assume that, at time $t = 0$, the reservoir is at temperature T, then ψ_0 is distributed according to the Gaussian measure with covariance $T\langle\cdot,\cdot\rangle$ and then $\xi(t) \equiv \langle\psi_0, e^{-\mathcal{L}t}\alpha\rangle$ is a d-dimensional stationary Gaussian process with mean 0 and covariance $C(t-s)$. Note that the covariance itself appears in the deterministic memory term on the r.h.s. of Eq.(7) (fluctuation-dissipation relation).

By assumption **H3** there is a polynomial $p(u)$ which is a real function of iu and which has its roots in the lower half plane such that

$$\int dk\,|\rho(k)|^2 e^{i|k|(t-s)} \;=\; \int_{-\infty}^{\infty} du\,\frac{1}{|p(u)|^2}e^{iu(t-s)}\,.$$

Note that this is a Markovian assumption [4]: $\xi(t)$ is Markovian in the sense that we have the identity $p(-id/dt)\xi(t) = \dot{\omega}(t)$, where $\dot{\omega}(t)$ is a white noise, i.e, the joint motion of $d^m\xi(t)/dt^m$, $0 \le m \le \deg p - 1$ is a (Gaussian) Markov process. This assumption together with the fluctuation-dissipation relation permits, by extending the phase space with a finite number of variables, to rewrite the integro-differential equations (7) as a Markov process. Note that $\xi(t)$ can be written as [4]

$$\xi(t) \;=\; \int_{-\infty}^{\infty} k(t-t')d\omega(t')\,, \quad k(t) \;=\; \int du\, e^{iut}p(u)^{-1}\,.$$

with $k(t)=0$ for $t \leq 0$. For example if $p(u) \propto iu+\gamma$ then $C(t)=\lambda^2 e^{-\gamma|t|}$. Introducing the variable r defined by

$$\lambda r(t) \;=\; \int_0^t ds\, C(t-s)p(s) + \int_{-\infty}^t k(t-t')d\omega(t')\,,$$

then we obtain from Eqs.(7) the set of Markovian differential equations:

$$\begin{aligned} \dot{q}(t) &= p(t)\,, \\ \dot{p}(t) &= -\nabla V_{\mathrm{eff}}(q(t)) - \lambda r(t)\,, \\ dr(t) &= (-\gamma r(t) + \lambda p(t))\,dt + (2T\gamma)^{1/2} d\omega(t)\,. \end{aligned} \tag{8}$$

If $p(u) \propto (iu+\gamma+i\sigma)(iu+\gamma-i\sigma)$ then $C(t) \;=\; \lambda^2\cos(\sigma t)e^{-\gamma|t|}$ and introducing the two auxiliary variables r and s defined by

$$\begin{aligned} \lambda r(t) &= \lambda^2\int_0^t ds\, \cos(\sigma(t-s))e^{-\gamma|t-s|}p(s) \\ &\quad +(T\lambda^2\gamma)^{1/2}\int_{-\infty}^t \cos(\sigma(t-s))e^{-\gamma|t-s|}d\omega(s)\,, \\ \lambda s(t) &= \lambda^2\int_0^t dt\, \sin(\sigma(t-s))e^{-\gamma|t-s|}p(s) \\ &\quad +(T\lambda^2\gamma)^{1/2}\int_{-\infty}^t dt\, \sin(\sigma(t-s))e^{-\gamma|t-s|}d\omega(s)\,, \end{aligned}$$

we obtain then the set of Markovian differential equations:

$$\begin{aligned} \dot{q}(t) &= p(t)\,, \\ \dot{p}(t) &= -\nabla V_{\mathrm{eff}}(q(t)) - \lambda r(t)\,, \\ dr(t) &= (-\gamma r(t) - \sigma s(t) + \lambda p(t))\,dt + (2T\gamma)^{1/2} d\omega(t)\,, \\ \dot{s}(t) &= -\gamma s(t) + \sigma r(t)\,. \end{aligned} \tag{9}$$

Obviously others similar set of equations can be derived for arbitrary polynomial $p(u)$.

Another coupling which we could easily handle with our methods occurs in the following limiting case, see [8]. Formally one wants to take

$C(t) = \eta^2\delta(t)$. Note that this corresponds to a coupling function with $|\rho(k)|^2 = 1$ in which case $\lambda^2 = \infty$. A possible limiting procedure consists in taking a sequence of covariances tending to a delta function and at the same time suitably rescaling the coupling (see [8]). In this case one obtains the Langevin equations which serve as commonly-used model system with reservoir in the physics literature,

$$\begin{aligned}
\dot{q}(t) &= p(t)\,, \\
dp(t) &= (-\nabla V_{\text{eff}}(q(t)) - \eta^2 p(t))\,dt + (2T\eta^2)^{1/2} d\omega(t)\,.
\end{aligned} \tag{10}$$

The derivation of the effective equations for the chain is a straightforward generalization of the above computations. Our techniques apply equally well to any of the couplings above. However, for simplicity, we will only consider the case where the couplings to both reservoirs satisfy $|\rho_i(k)|^2 \propto k^2 + \gamma^2$, $i = L, R$. For notational simplicity we set $T_1 = T_L$ and $T_n = T_R$, we denote r_1 and r_n as the two auxiliary variables and we will use the notations $r = (r_1, r_n)$, and $x = (p, q, r) \in X = \mathbf{R}^{2d(n+1)}$. In this case we obtain the set of Markovian stochastic differential equations given by

$$\begin{aligned}
\dot{q}_1 &= p_1\,, \\
\dot{p}_1 &= -\nabla_{q_1} V_{\text{eff}}(q) - \lambda r_1\,, \\
dr_1 &= (-\gamma r_1 + \lambda p_1)\,dt + (2T_1\gamma)^{1/2} d\omega_1\,, \\
\dot{q}_j &= p_j\,, \qquad\qquad j = 2, \ldots, n-1\,, \\
\dot{p}_j &= -\nabla_{q_j} V_{\text{eff}}(q) \quad j = 2, \ldots, n-1\,, \\
\dot{q}_n &= p_n\,, \\
\dot{p}_n &= -\nabla_{q_n} V_{\text{eff}}(q) - \lambda r_n\,, \\
dr_n &= (-\gamma r_n + \lambda p_n)\,dt + (2T_n\gamma)^{1/2} d\omega_n\,,
\end{aligned} \tag{11}$$

where $V_{\text{eff}}(q) = V(q) - \lambda^2 q_1^2/2 - \lambda^2 q_n^2/2$. From now on, for notational simplicity we will suppress the index "eff" and consider $V = V_{\text{eff}}$ as our potential energy.

It will be useful to introduce the following notation. We define the linear maps $\Lambda : \mathbf{R}^{dn} \to \mathbf{R}^{2d}$ by $\Lambda(x_1, \ldots, x_n) = (\lambda x_1, \lambda x_n)$ and $T : \mathbf{R}^{2d} \to \mathbf{R}^{2d}$ by $T(x, y) = (T_1 x, T_n y)$. With this we can rewrite Eq.(11) in the compact form

$$\begin{aligned}
\dot{q} &= p\,, \\
\dot{p} &= -\nabla_q V - \Lambda^T r\,, \\
dr &= (-\gamma r + \Lambda p)\,dt + (2\gamma T)^{1/2} d\omega\,.
\end{aligned} \tag{12}$$

The solution $x(t)$ of Eq.(12) is a Markov process. We denote T^t as the associated semigroup,

$$T^t f(x) \;=\; \mathbf{E}_x[f(x(t))]\,,$$

with generator

$$L \;=\; \gamma\,(\nabla_r T \nabla_r - r\nabla_r) + (\Lambda p \nabla_r - r\Lambda\nabla_p) + (p\nabla_q - (\nabla_q V(q))\nabla_p)\;, \quad (13)$$

and $P_t(x, dy)$ as the transition probability of the Markov process $x(t)$. There is a natural energy function which is associated to Eq.(12), given by

$$G(p,q,r) \;=\; \frac{r^2}{2} + H(p,q)\,.$$

A straightforward computation shows that in the special case $T_1 = T_n = T$

$$Z^{-1} e^{-G(p,q,r)/T}$$

is an invariant measure for the Markov process $x(t)$.

Given a function $W : X \to \mathbf{R}$ satisfying $W \geq 1$ we consider the following weighted total variation norm $\|\cdot\|_W$ given by

$$\|\pi\|_W \;=\; \sup_{|f|\leq W} |\int f d\pi|\,, \quad (14)$$

for any (signed) measure π. We introduce norms $\|\cdot\|_\theta$ and Banach spaces $L^\infty_\theta(X)$ given by

$$\|f\|_\theta \;=\; \sup_{x\in X} \frac{|f(x)|}{e^{\theta G(x)}}\,, \quad L^\infty_\theta(X) \;=\; \{f \,:\, \|f\|_\theta < \infty\}\,, \quad (15)$$

and write $\|K\|_\theta$ for the norm of an operator $K : L^\infty_\theta(X) \to L^\infty_\theta(X)$.

Theorem 1.1 is a direct consequence of the following result

Theorem 2.1 *Assume that* **H1** *and* **H2** *hold. The Markov process $x(t)$ which solves (12) has smooth transition probability densities, $P_t(x, dy) = p_t(x,y)dy$, with $p_t(x,y) \in \mathcal{C}^\infty((0,\infty)\times X \times X)$. The Markov process $x(t)$ has a unique invariant measure μ, and μ has a $\mathcal{C}^\infty$ everywhere positive density. For any θ with $0 < \theta < (\max\{T_1, T_n\})^{-1}$ there exist constants $r = r(\theta) > 1$ and $R = R(\theta) < \infty$ such that*

$$\|P_t(x,\cdot) - \mu\|_{\exp(\theta G)} \;\leq\; R r^{-t} \exp(\theta G(x))\,, \quad (16)$$

for all $x \in X$, (exponential convergence to the SNS) or equivalently

$$\|T^t - \mu\|_\theta \le Rr^{-t} ,$$

(spectral gap). Furthermore for all functions f, g with f^2, $g^2 \in L^\infty_\theta(X)$ and all $t > 0$ we have

$$\left| \int gT^t f \, d\mu - \int f \, d\mu \int g \, d\mu \right| \le Rr^{-t} \|f^2\|_\theta^{1/2} \|g^2\|_\theta^{1/2} ,$$

(exponential decay of correlations in the SNS).

The convergence in the weighted variation norm, Eq. (16), implies that the Law of large Numbers holds [10, 18].

Corollary 2.2 *Under assumptions* **H1** *and* **H2** *$x(t)$ satisfies the Law of Large Numbers: For all initial conditions $x \in X$ and all $f \in L^1(X, d\mu)$*

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T f(x(t)) \, dt = \int f \, d\mu$$

almost surely.

The convergence of the transition probabilities as given in (16) is shown in [18] to follow from the following properties:

- **Strong Feller property** The diffusion process is strong Feller, i.e., the semigroup T^t maps bounded measurable functions into continuous functions.

This is a consequence of the hypoellipticity of the diffusion $x(t)$, which follows from Condition **H2**, see Section 4.

- **Small-time open set accessibility**. For all $t > 0$, all $x \in X$ and all open set $A \subset X$ we have $P_t(x, A) > 0$.

This means that the Markov process is "strongly aperiodic". In particular, combined with the strong Feller property it implies uniqueness of the invariant measure. This property is discussed in Section 4 using the support theorem of [27] and explicit computations. This generalizes (slightly) the result obtained in [7].

- **Liapunov function and hitting times** Fix $s > 0$ arbitrary. Set $W = \exp(\theta G)$ and choose θ with $0 < \theta < (\max\{T_1, T_n\})^{-1}$. Then W is a Liapunov function for the Markov chain $\{x(ns)\}_{n\geq 0}$: $W > 1$, W has compact level sets and there is a compact set U, (depending on s and θ) and constants $\kappa < 1$ and $b < \infty$, (both depending on U, s and θ) such that
$$T^s W(x) \leq \kappa W(x) + b\mathbf{1}_U(x)\,, \tag{17}$$
where $\mathbf{1}_U$ denotes the indicator function of the set U.

The existence of a Liapunov function is the main technical result of this paper (see Section 3) and the condition **H1** is crucial to obtain it. Note that the time derivative of the (averaged) energy

$$\frac{d}{dt}\mathbf{E}_x[G(x(t))] = \gamma \mathbf{E}_x[\mathrm{Tr(T) - r^2(t)}]\,,$$

is not necessarily negative. But it is the case, as follows from our analysis below that, that for $t > 0$, $\mathbf{E}_x[G(x(t)) - G(x)] < -cG(x)^{2/k_2}$ for x sufficiently large.

A nice interpretation of a Liapunov bound of the form (17) is in terms of hitting times. Let τ_U denote the first time the diffusion $x(t)$ hits the set U; then Eq. (17) implies that τ_U is exponentially bounded. We will show that for any $a > 0$, *no matter how large*, we can find a compact set $U = U(a)$ such that

$$\mathbf{E}_x[e^{a\tau_U}] < \infty\,,$$

for all $x \in X$. So except for exponentially rare excursions the Markov process $x(t)$ lives on the compact set U. Combined with the fact that the process has a smooth law, this provides an intuitive picture of the exponential convergence result of Theorem 2.1.

3 Liapunov Function and Hitting Times

3.1 Scaling and Deterministic Energy Dissipation

We first consider the question of energy dissipation for the following deterministic equations

$$\begin{aligned} \dot{q} &= p\,, \\ \dot{p} &= -\nabla_q V(q) - \Lambda^T r\,, \\ \dot{r} &= -\gamma r + \Lambda p\,, \end{aligned} \tag{18}$$

obtained from Eq.(12) by setting $T_1 = T_n = 0$, corresponding an initial condition of the reservoirs with energy 0. A simple computation shows that the energy $G(p, q, r)$ is non-increasing along the flow $x(t) = (p(t), q(t), r(t))$ given by Eq.(18):

$$\frac{d}{dt} G(p(t), q(t), r(t)) \;=\; -\gamma r^2(t) \;\leq\; 0\,.$$

We now show by a scaling argument that for any initial condition with sufficiently high energy, after a small time, a substantial amount of energy is dissipated.

At high energy, the two-body interaction $U^{(2)}$ in the potential dominates the term $U^{(1)}$ since $k_2 \geq k_1$ and so for an initial condition with energy $G(x) = E$, the natural time scale – essentially the period of a single one-dimensional oscillator in the potential $|q|^{k_2}$ – is $E^{1/k_2 - 1/2}$. We scale a solution of Eq.(18) with initial energy E as follows

$$\begin{aligned} \tilde{p}(t) &= E^{-\frac{1}{2}} p(E^{\frac{1}{k_2}-\frac{1}{2}} t)\,, \\ \tilde{q}(t) &= E^{-\frac{1}{k_2}} q(E^{\frac{1}{k_2}-\frac{1}{2}} t)\,, \\ \tilde{r}(t) &= E^{-\frac{1}{k_2}} r(E^{\frac{1}{k_2}-\frac{1}{2}} t)\,. \end{aligned} \tag{19}$$

Accordingly the energy scales as $G(p, q, r) = E\tilde{G}_E(\tilde{p}, \tilde{q}, \tilde{r})$, where

$$\begin{aligned} \tilde{G}_E(\tilde{p}, \tilde{q}, \tilde{r}) &= E^{\frac{2}{k_2}-1} \frac{\tilde{r}^2}{2} + \frac{\tilde{p}^2}{2} + \tilde{V}_E(\tilde{q})\,, \\ \tilde{V}_E(\tilde{q}) &= \sum_{i=1}^{n} \tilde{U}^{(1)}(\tilde{q}_i) + \sum_{i=1}^{n-1} \tilde{U}^{(2)}(\tilde{q}_i - \tilde{q}_{i+1})\,, \\ \tilde{U}^{(i)}(\tilde{x}) &= E^{-1} \tilde{U}^{(i)}(E^{\frac{1}{k_2}} x)\,, \quad i = 1, 2\,. \end{aligned}$$

The equations of motion for the rescaled variables are

$$\begin{aligned} \dot{\tilde{q}} &= \tilde{p}\,, \\ \dot{\tilde{p}} &= -\nabla_{\tilde{q}} \tilde{V}_E(\tilde{q}) - E^{\frac{2}{k_2}-1} \Lambda^T r\,, \\ \dot{\tilde{r}} &= -E^{\frac{1}{k_2}-\frac{1}{2}} \gamma \tilde{r} + \Lambda \tilde{p}\,. \end{aligned} \tag{20}$$

By assumption **H1**, as $E \to \infty$ the rescaled energy becomes

$$\begin{aligned} \tilde{G}_\infty(\tilde{p}, \tilde{q}, \tilde{r}) &\equiv \lim_{E \to \infty} \tilde{G}_E(\tilde{p}, \tilde{q}, \tilde{r}) \\ &= \begin{cases} \tilde{p}^2/2 + \tilde{V}_\infty(\tilde{q}) & k_1 = k_2 > 2 \text{ or } k_2 > k_1 \geq 2 \\ \tilde{r}^2/2 + \tilde{p}^2/2 + \tilde{V}_\infty(\tilde{q}) & k_1 = k_2 = 2 \end{cases}\,, \end{aligned}$$

where

$$V_\infty(\tilde q) \;=\; \left\{ \begin{array}{ll} \sum a^{(1)}\|\tilde q_i\|^{k_2} + \sum a^{(2)}\|\tilde q_i - \tilde q_{i+1}\|^{k_2} & k_1 = k_2 \geq 2 \\ \sum a^{(2)}\|\tilde q_i - \tilde q_{i+1}\|^{k_2} & k_2 > k_1 \geq 2 \end{array} \right. .$$

The equations of motion scale in this limit to

$$\begin{aligned} \dot{\tilde q} &= \tilde p\,, \\ \dot{\tilde p} &= -\nabla_{\tilde q}\tilde V_\infty(\tilde q)\,, \\ \dot{\tilde r} &= \Lambda \tilde p\,, \end{aligned} \tag{21}$$

in the case $k_2 > 2$, while they scale to

$$\begin{aligned} \dot{\tilde q} &= \tilde p\,, \\ \dot{\tilde p} &= -\nabla_{\tilde q}\tilde V_\infty(\tilde q) - \Lambda^T r\,, \\ \dot{\tilde r} &= -\gamma r + \Lambda \tilde p\,, \end{aligned} \tag{22}$$

in the case $k_1 = k_2 = 2$.

Remark 3.1 The scaling for the p and q is natural due to the Hamiltonian nature of the problem, but the scaling of r has a certain amount of arbitrariness. Since G is quadratic in r, it might appear natural to scale r with a factor $E^{-1/2}$ instead of E^{-1/k_2} as we do. On the other hand, the very definition of r as an integral of p suggests that r should scale as q, as we have chosen.

Remark 3.2 Had we supposed, instead of **H1**, that $k_1 > k_2$, then the natural time scale at high energy would be $E^{1/k_1-1/2}$. Scaling the variables (with k_2 replaced by k_1 would yield the limiting Hamiltonian $\tilde p^2/2 + \sum a^{(1)}\|\tilde q_i\|^{k_1}$, i.e., the Hamiltonian of n *uncoupled* oscillators. So in this case, at high energy, essentially no energy is transmitted through the chain. While this does not necessary preclude the existence of an invariant measure, we expect in this case the convergence to a SNS to be much slower. In any case even the existence of the SNS in this case remains an open problem.

Theorem 3.3 *Given $\tau > 0$ fixed there are constants $c > 0$ and $E_0 < \infty$ such that for any x with $G(x) = E > E_0$ and any solution $x(t)$ of Eq.(18) with $x(0) = x$ we have the estimate, for $t_E = E^{1/k_2-1/2}\tau$,*

$$G(x(t_E)) - E \;\leq -cE^{\frac{3}{k_2}-\frac{1}{2}}\,. \tag{23}$$

Remark 3.4 In view of Eq. (23), this shows that r is at least typically $O(E^{1/k_2})$ on the time interval $[0, E^{1/k_2-1/2}\tau]$.

Proof: Given a solution of Eq.(18) with initial condition x of energy $G(x) = E$, we use the scaling given by Eq.(19) and we obtain

$$G(x(t_E)) - E \;=\; -\gamma \int_0^{t_E} dt\, r^2(t) \;=\; -\gamma E^{\frac{3}{k_2}-\frac{1}{2}} \int_0^{\tau} dt\, \tilde{r}^2(t)\,, \qquad (24)$$

where $\tilde{r}(t)$ is the solution of Eq.(20) with initial condition $\tilde{x}$ of (rescaled) energy $\tilde{G}_E(\tilde{x}) = 1$. By Assumption **H2** we may choose E_0 so large that for $E > E_0$ the critical points of $\tilde{G}_E$ are contained in, say, the set $\{\tilde{G}_E \leq 1/2\}$.

For a fixed E and x with $G(x) = E$, we show that there is a constant $c_{x,E} > 0$ such that

$$\int_0^{\tau} dt\, \tilde{r}^2(t) \;\geq\; c_{\tilde{x},E}\,. \qquad (25)$$

The proof is by contradiction, c.f. [21]. Suppose that $\int_0^\tau dt\, \tilde{r}^2(t) = 0$, then we have $\tilde{r}(t) = 0$, for all $t \in [0,\tau]$. From the third equation in (20) we conclude that $\tilde{p}_1(t) = \tilde{p}_n(t) = 0$ for all $t \in [0,\tau]$, and so from the first equation in (20) we see that $\tilde{q}_1(t)$ and $\tilde{q}_n(t)$ are constant on $[0,\tau]$. The second equation in (20) gives then

$$0 \;=\; \dot{\tilde{p}}_1(t) \;=\; -\nabla_{\tilde{q}_1}\tilde{V}(\tilde{q}(t)) \;=\; -\nabla_{\tilde{q}_1}\tilde{U}^{(1)}(\tilde{q}_1(t)) - \nabla_{\tilde{q}_1}\tilde{U}^{(2)}(\tilde{q}_1(t) - \tilde{q}_2(t))\,,$$

together with a similar equation for $\dot{p}_n$. By our assumption **H1** the map $\nabla\tilde{U}^{(2)}$ has a right inverse g locally bounded and measurable and thus we obtain

$$\tilde{q}_2(t) \;=\; \tilde{q}_1(t) - g(\tilde{U}^{(1)}(\tilde{q}_1(t)))\,.$$

Since $\tilde{q}_1$ is constant, this implies that $\tilde{q}_2$ is also constant on $[0,\tau]$. Similarly we see that $\tilde{q}_{n-1}$ is constant on $[0,\tau]$. Using again the first equation in (20) we obtain now $\tilde{p}_2(t) = \tilde{p}_{n-1}(t) = 0$ for all $t \in [0,\tau]$. Inductively one concludes that $\tilde{r} = 0$ implies $\tilde{p} = 0$ and $\nabla_{\tilde{q}}\tilde{V} = 0$ and thus the initial condition $\tilde{x}$ is a critical point of $\tilde{G}_E$. This contradicts our assumption and Eq. (25) follows.

Now for given E, the energy surface $\tilde{G}_E$ is compact. Using the continuity of the solutions of O.D.E with respect to initial conditions we conclude that there is a constant $c_E > 0$ such that

$$\inf_{\tilde{x} \in \{\tilde{G}_E = 1\}} \int_0^{\tau} dt\, \tilde{r}^2(t) \;\geq\; c_E\,.$$

Finally we investigate the dependence on E of c_E. We note that for $E=\infty$, $\tilde{G}_\infty$ has a well-defined limit given by Eq.(21) and the rescaled equations of motion, in the limit $E \to \infty$, are given by Eqs. (21) in the case $k_2 > 2$ and by Eq. (22) in the case $k_1 = k_2 = 2$. Except in the case $k_1 = k_2 = 2$ the energy surface $\{\tilde{G}_\infty = 1\}$ is *not* compact. However, in the case $k_1 = k_2 > 2$, the Hamiltonian $\tilde{G}_\infty$ and the equation of motion are invariant under the translation $r \mapsto r + a$, for any $a \in \mathbf{R}^{2d}$. And in the case $k_2 > k_1 > 2$ the Hamiltonian $\tilde{G}_\infty$ and the equation of motion are invariant under the translation $r \mapsto r + a$ $q \mapsto q + b$, for any $a \in \mathbf{R}^{2d}$ and $b \in \mathbf{R}^{dn}$. The quotient of the energy surface $\{\tilde{G}_\infty = 1\}$ by these translation, is compact.

Note that for a given $\tilde{x} \in \{\tilde{G}_\infty = 1\}$ a similar argument as above show that $\int_0^\tau dt(\tilde{r}+a)^2 > 0$, for any $a > 0$ and since this integral clearly goes to ∞ as $a \to \infty$ there exists a constant $c_\infty > 0$ such that

$$\inf_{\tilde{x} \in \{\tilde{G}_\infty = 1\}} \int_0^\tau \tilde{r}^2(t)\, dt > c_\infty \,.$$

Using again that the solution of O.D.E depends smoothly on its parameters, we obtain

$$\inf_{E > E_0} \inf_{\tilde{x} \in \{\tilde{G}_E = 1\}} \int_0^\tau dt\, \tilde{r}^2(t) \; > \; c \,.$$

This estimate, together with Eq. (24) gives the conclusion of Theorem 3.3. ■

3.2 Approximate Deterministic Behavior of Random Paths

In this section we show, that at sufficiently high energies, the overwhelming majority of the random paths $x(t) = x(t,\omega)$ solving Eqs.(12) follows very closely the deterministic paths $x_{\rm det}$ solving Eqs.(18). As a consequence, for most random paths the same amount of energy is dissipated into the reservoirs as for the corresponding deterministic ones. We need the following *a priori* "no-runaway" bound on the growth of $G(x(t))$.

Lemma 3.5 *Let* $\theta \le (\max\{T_1, T_n\})^{-1}$. *Then* $\mathbf{E}_x[\exp(\theta G(x(t)))]$ *is well-defined and satisfies the bound*

$$\mathbf{E}_x[\exp(\theta G(x(t)))] \; \le \; \exp(\gamma {\rm Tr}(T)\theta t) \exp(\theta G(x)) \,. \tag{26}$$

Moreover for any x *with* $G(x) = E$ *and any* $\delta > 0$ *we have the estimate*

$$\mathbf{P}_x \left\{ \sup_{0 \le s \le t} G(x(s)) \ge (1+\delta)E \right\} \; \le \; \exp(\gamma {\rm Tr}(T)\theta t) \exp(-\delta\theta E) \,. \tag{27}$$

Remark 3.6 The lemma shows that for E sufficiently large, with very high probability, $G(x(t)) = O(E)$ if $G(x) = E$. The assumption on θ here arises naturally in the proof, where we need $(1-\theta T) \geq 0$, cf. Eq. (28).

Proof: For $\theta \leq (\max\{T_1, T_n\})^{-1}$ we have the bound (the generator L is given by Eq. (13))

$$\begin{aligned} L\exp(\theta G(x)) &= \gamma\theta\exp(\theta G(x))\,(\mathrm{Tr}(T) - r(1-\theta T)r) \\ &\leq \gamma\theta\mathrm{Tr}(T)\exp(\theta G(x))\,, \end{aligned} \tag{28}$$

so that for the function $W(t,x) = \exp(-\gamma\theta\mathrm{Tr}(T)t)\exp(\theta G(x))$ we have the inequality $(\partial_t + L)W(t,x) \leq 0$. We denote σ_R as the exit time from the set $\{G(x) < R\}$, i.e., $\sigma_R = \inf\{t \geq 0, G(x(t)) \geq R\}$. If the initial condition x satisfies $G(x) = E < R$, we denote $x_R(t)$ the process which is stopped when it exits $\{G(x) < R\}$, i.e., $x_R(t) = x(t)$ for $t < \sigma_R$ and $x_R(t) = x(\sigma_R)$ for $t \geq \sigma_R$. We set $\sigma_R(t) = \min\{\sigma_R, t\}$ and applying Ito's formula with stopping time to the function $W(t,x)$ we obtain

$$\mathbf{E}_x\left[\exp(\theta G(x(\sigma_R(t))))\exp(-\gamma\theta\mathrm{Tr}(T)\sigma_R(t))\right] - \exp(\theta G(x)) \leq 0\,,$$

thus

$$\mathbf{E}_x\left[\exp(\theta G(x(\sigma_R(t))))\right] \leq \exp(\gamma\theta\mathrm{Tr}(T)t)\exp(\theta G(x))\,. \tag{29}$$

Since

$$\begin{aligned} \mathbf{E}_x\left[\exp(\theta G(x(\sigma_R(t))))\right] &\geq \mathbf{E}_x\left[\exp(\theta G(x(\sigma_R(t))))\mathbf{1}_{\sigma_R<t}\right] \\ &= \mathbf{P}_x\{\sigma_R < t\}\exp(\theta R)\,, \end{aligned}$$

we obtain the bound

$$\mathbf{P}_x\{\sigma_R < t\} \leq \exp(\gamma\theta\mathrm{Tr}(T)t)\exp(\theta(E-R))\,.$$

As a consequence $\mathbf{P}_x\{\sigma_R < t\} \to 0$ as $R \to \infty$ and thus the Markov process $x(t)$ is non-explosive.

It follows that $G(x_R(t)) \to G(x(t))$ almost surely as $R \to \infty$, so by the Fatou lemma we obtain from Eq. (29) the bound Eq. (26). The bound Eq. (27) is obtained by noting that the left side is equal to

$$\mathbf{P}_x\{\sigma_{E(1+\delta)} < t\} \leq \exp(\gamma\theta\mathrm{Tr}(T)t)\exp(-\delta\theta E)\,,$$

and this concludes the proof of Lemma 3.5.

We have the following "tracking" estimates to the effect that the random path closely follows the deterministic one at least up to time t_E for a set of paths which have nearly full measure. We set $\Delta x(t) \equiv x(t,\omega) - x_{\rm det}(t) = (\Delta r(t), \Delta p(t), \Delta q(t))$ with both $x(t)$ and $x_{\rm det}(t)$ having initial condition x. Let

$$S(x,E,t) \;=\; \{x(\cdot)\,;\, G(x) = E \text{ and } \sup_{0\le s\le t} G(x(s)) < 2E\}\,.$$

By Lemma 3.5, $\mathbf{P}\{S(x,E,t)\} \ge 1 - \exp\left(\gamma\theta{\rm Tr}(T)t - \theta E\right)$.

Proposition 3.7 *There exist constants $E_0 < \infty$ and $c > 0$ such that for paths $x(t,\omega) \in S(x,E,t_E)$ with $t_E = E^{1/k_2/-1/2}\tau$ and $E > E_0$ we have*

$$\sup_{0\le t\le t_E} \left(\begin{array}{c} \|\Delta q(t)\| \\ \|\Delta p(t)\| \\ \|\Delta r(t)\| \end{array}\right) \le c \sup_{0\le t\le t_E} \|\sqrt{2\gamma T}\omega(t)\| \left(\begin{array}{c} E^{\frac{2}{k_2}-1} \\ E^{\frac{1}{k_2}-\frac{1}{2}} \\ 1 \end{array}\right). \qquad (30)$$

Proof: We write differential equations for $\Delta x(t)$ again assuming both the random and deterministic paths start at the same point x with energy $G(x) = E$. These equations can be written in the somewhat symbolic form:

$$\begin{aligned} d\Delta q &= \Delta p dt\,, \\ d\Delta p &= \Big(O(E^{1-2/k_2})\Delta q - \Lambda^T\Delta r\Big)\,dt\,, \\ d\Delta r &= \left(-\gamma\Delta r + \Lambda\Delta p\right)dt + \sqrt{2\gamma T}d\omega \end{aligned} \qquad (31)$$

The $O(E^{1-2/k_2})$ coefficient refers to the difference between forces, $-\nabla_q V(\cdot)$ evaluated at $x(t)$ and $x_{\rm det}(t)$; we have that $G(x(t)) \le 2E$, so that $\nabla_q V(q) - \nabla_q V(q_{\rm det}) = O(\partial^2 V)\Delta q = O(E^{1-2/k_2})\Delta q$. For later purposes we pick a constant c' so large that

$$\rho = \rho(x) = c' E^{1-\frac{2}{k_2}} \ge \sup_i \sum_j \sup_{\{q: V(q)\le 2E\}} \left|\frac{\partial^2 V(q)}{\partial q_i \partial q_j}\right|$$

for all sufficiently large E.

In order to estimate the solutions of Eqs. (31), we consider the 3×3 matrix which bounds the coefficients in this system, and which is given by

$$M \;=\; \left(\begin{array}{ccc} 0 & 1 & 0 \\ \rho & 0 & \lambda \\ 0 & \lambda & \gamma \end{array}\right) \qquad (32)$$

We have the following estimate on powers of M; For $\Delta X^{(0)} = (0,0,1)^T$, we set $\Delta X^{(m)} \equiv M^m \Delta X^{(0)}$. For $\alpha = \max(1, \gamma+\lambda)$, we obtain $\Delta X^{(1)} \le$

$\alpha(0,1,1)^T$ $\Delta X^{(2)} \leq \alpha^2(1,1,1)^T$ and for $m \geq 3$,

$$\Delta X^{(m)} \equiv \begin{pmatrix} u^{(m)} \\ v^{(m)} \\ w^{(m)} \end{pmatrix} \leq \alpha^m 2^{m-2} \begin{pmatrix} \rho^{\frac{m-2}{2}} \\ \rho^{\frac{m-1}{2}} \\ \rho^{\frac{m-2}{2}} \end{pmatrix},$$

where the inequalities are componentwise. From this we obtain the bound

$$e^{tM} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \leq \begin{pmatrix} \frac{1}{2}(\alpha t)^2 e^{\sqrt{\rho}2\alpha t} \\ \alpha t e^{\sqrt{\rho}2\alpha t} \\ 1 + \alpha t + \frac{1}{2}(\alpha t)^2 e^{\sqrt{\rho}2\alpha t} \end{pmatrix}. \tag{33}$$

If $0 \leq t \leq t_E$ we have $\sqrt{\rho}t < \sqrt{c'}$. Then the exponentials in the above equation are bounded, and

$$e^{tM} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \leq c \begin{pmatrix} 1/\rho \\ 1/\sqrt{\rho} \\ 1 \end{pmatrix}, \tag{34}$$

for some constant c.

Returning now to the original differential equation system Eq.(31), we write this equation in the usual integral equation form:

$$\begin{pmatrix} \Delta q(t) \\ \Delta p(t) \\ \Delta r(t) \end{pmatrix} = \int_0^t \begin{pmatrix} \Delta p(s) \\ -\nabla_q V(q(s,\omega))\, ds + \nabla_q V(q_{\text{det}}(s)) - \Gamma^T \Delta r(s) \\ -\gamma \Delta r(s) + \Lambda \Delta p(s) \end{pmatrix}$$
$$+ \begin{pmatrix} 0 \\ 0 \\ \sqrt{2\gamma T}\omega(t) \end{pmatrix}. \tag{35}$$

From this we obtain the bound

$$\begin{pmatrix} \|\Delta q(t)\| \\ \|\Delta p(t)\| \\ \|\Delta r(t)\| \end{pmatrix} \leq \int_0^t M \begin{pmatrix} \|\Delta q(t)\| \\ \|\Delta p(t)\| \\ \|\Delta r(t)\| \end{pmatrix} ds + \begin{pmatrix} 0 \\ 0 \\ \omega_{\max} \end{pmatrix},$$

where M is the matrix given by Eq.(32), and $\omega_{\max} = \sup_{t \leq t_E} \|\sqrt{2\gamma T}\omega(t)\|$. Note that the solution of the integral equation

$$\Delta X(t) = \int_0^t ds\, M \Delta X(s) + \begin{pmatrix} 0 \\ 0 \\ \omega_{\max} \end{pmatrix}, \tag{36}$$

is $\Delta X(t) = \exp{(tM)}(0, 0, \omega_{\max})^T$. We can solve both Eq.(35) and Eq.(36) by iteration. Let $\Delta x_m(s)$, $\Delta X_m(s)$ denote the respective m^{th} iterates (with $\Delta x_0(s) = (0, 0, \sqrt{2\gamma T}\omega(s))^T$, and $\Delta X_0(s) = (0, 0, \omega_{\max})^T$, $0 \leq s \leq t_E$). The ΔX_m's are monotone increasing in m. Then it is easy to see that

$$\begin{pmatrix} \|\Delta q_m(t)\| \\ \|\Delta p_m(t)\| \\ \|\Delta r_m(t)\| \end{pmatrix} \leq \Delta X_m(t) \leq \Delta X(t) ,$$

for each iterate. By Eqs.(33), (34), and the definition of ρ the conclusion Eq. (30) follows. ■

As a consequence of Theorem 3.3 and Proposition 3.7 we obtain

Corollary 3.8 *Let $\Omega(E) = E^\alpha$ with $\alpha < 1/k_2$ and assume that $w(t)$ is such that $\sup_{0\leq t\leq t_E} \|\sqrt{2\gamma T}\omega(t)\| \leq \Omega(E)$ and $x(\cdot, \omega) \in S(x, E, t_E)$. Then there are constants $c > 0$ and $E_0 < \infty$ such that all paths $x(t, w)$ with initial condition x with $G(x) = E > E_0$ satisfy the bound*

$$\int_0^{t_E} r^2(s)ds \geq cE^{\frac{3}{k_2}-\frac{1}{2}} . \tag{37}$$

Remark 3.9 For large energy E, paths *not* satisfying the hypotheses of the corollary have measure bounded by

$$\begin{aligned} &\mathbf{P}_x\{\sup_{0\leq s\leq t_E} \|\sqrt{2\gamma T}\omega\| > \Omega(E)\} + \mathbf{P}\{S(x, E, t_E)^C\} \\ &\leq \frac{a}{2}\exp\left(-\frac{\Omega(E)^2}{b\gamma T_{\max}t_E}\right) + \exp{(\theta(\gamma \mathrm{Tr}(T)t_E - E))} \\ &\leq a\exp\left(-\frac{\Omega(E)^2}{b\gamma T_{\max}t_E}\right) , \end{aligned} \tag{38}$$

where a and b are constants which depend only on the dimension of ω. Here we have used the reflection principle to estimate the first probability and Eq. (27) and the definition of S to estimate the second probability. For E large enough, the second term is small relative to the first.

Proof: It is convenient to introduce the L^2-norm on functions on $[0, t]$, $\|f\|_t \equiv \left(\int_0^t \|f(s)\|^2 ds\right)^{1/2}$. By Theorem 3.3, there are constants E_1 and c_1 such that for $E > E_1$ the deterministic paths $x_{\mathrm{det}}(s)$ satisfy the bound

$$\|r_{\mathrm{det}}\|_{t_E}^2 = \int_0^{t_E} r_{\mathrm{det}}^2(s)ds \geq c_1E^{\frac{3}{k_2}-\frac{1}{2}} .$$

By Proposition 3.7, there are constants E_2 and c_2 such that $\|\Delta r(s)\| \leq c_2\Omega(E)$, uniformly in s, $0 \leq s \leq t_E$, and uniformly in x with $G(x) > E_2$. So we have

$$\|r\|_{t_E} \geq \|r_{\rm det}\|_{t_E} - \|\Delta r\|_{t_E} \geq \left(c_1 E^{\frac{3}{k_2}-\frac{1}{2}}\right)^{1/2} - c_2\Omega(E)\left(E^{\frac{1}{k_2}-\frac{1}{2}}\right)^{1/2}.$$

But the last term is $O(E^{\alpha-1/4+1/2k_2})$, which is of lower order than the first since $\alpha < 1/k_2$, so the corollary follows, for an appropriate constant c and E sufficiently large. ■

3.3 Liapunov Function and Exponential Hitting Times

With the estimates we prove now our main technical result.

Theorem 3.10 *Let $s > 0$ and $\theta < \theta_0 \equiv (\max\{T_1, T_n\})^{-1}$. Then there are a compact set $U = U(s,\theta)$ and constants $\kappa = \kappa(U,s,\theta) < 1$ and $L = L(U,s,\theta) < \infty$ such that*

$$T^s \exp(\theta G)(x) \;\leq\; \kappa \exp(\theta G)(x) + L\mathbf{1}_U(x)\,. \tag{39}$$

where $\mathbf{1}_U$ is the indicator function of the set U. The constant κ can be made arbitrarily small by choosing U large enough.

Proof: For any compact set U and for any t, $T^s \exp(\theta G)(x)$ is a bounded function, uniformly on $[0,t]$. So, in order to prove Eq.(39), we only have to prove that there exist a compact set U and $\kappa < 1$ such that

$$\sup_{x \in U^{\rm C}} \mathbf{E}_x\left[\exp\left(\theta(G(x(s)) - G(x))\right)\right] \;\leq\; \kappa \;<\; 1\,.$$

Using Ito's Formula to compute $G(x(s) - G(x)$ in terms of a stochastic integral we obtain

$$\begin{aligned}&\mathbf{E}_x\left[\exp\left(\theta(G(x(s)) - G(x))\right)\right]\\ &= \exp\left(\theta\gamma{\rm Tr}(T)s\right)\mathbf{E}_x\left[\exp\left(-\theta\int_0^s \gamma r^2\,dt + \theta\int_0^s \sqrt{2\gamma T}r\,d\omega(t)\right)\right].\end{aligned}\tag{40}$$

For any $\theta < \theta_0$, we choose $p > 1$ such that $\theta p < \theta_0$. Using Hölder inequality we obtain,

$$\mathbf{E}_x\left[\exp\left(-\theta\int_0^s \gamma r^2\,dt + \theta\int_0^s \sqrt{2\gamma T}r\,d\omega(t)\right)\right]$$

$$
\begin{aligned}
&= \mathbf{E}_x \left[\exp\left(-\theta \int_0^s \gamma r^2 \, dt + \frac{p\theta^2}{2} \int_0^s (\sqrt{2\gamma T} r)^2 \, dt \right) \times \right. \\
&\qquad \left. \times \exp\left(-\frac{p\theta^2}{2} \int_0^s (\sqrt{2\gamma T} r)^2 \, dt + \theta \int_0^s \sqrt{2\gamma T} r d\omega(t) \right) \right] \\
&\le \mathbf{E}_x \left[\exp\left(-q\theta \int_0^s \gamma r^2 \, dt + \frac{qp\theta^2}{2} \int_0^s (\sqrt{2\gamma T} r)^2 \, dt \right) \right]^{1/q} \times \\
&\quad \times \mathbf{E}_x \left[\exp\left(-\frac{p^2\theta^2}{2} \int_0^s (\sqrt{2\gamma T} r)^2 \, dt + \theta p \int_0^s \sqrt{2\gamma T} r d\omega(t) \right) \right]^{1/p} \\
&= \mathbf{E}_x \left[\exp\left(-q\theta \int_0^s dt \, \gamma r^2 + \frac{qp\theta^2}{2} \int_0^s dt \, (\sqrt{2\gamma T} r)^2 \right) \right]^{1/q}
\end{aligned}
$$

Here, in the next to last line, we have used the fact that the second factor is the expectation of a martingale (the integrand is non-anticipating) with expectation 1. Finally we obtain the bound

$$
\begin{aligned}
& \mathbf{E}_x \left[\exp\left(\theta (G(x(s)) - G(x)) \right) \right] \\
\le \quad & \exp\left(\theta \gamma \mathrm{Tr}(T) s \right) \mathbf{E}_x \left[\exp\left(-q\theta (1 - p\theta T_{\max}) \int_0^s dt \, \gamma r^2 \right) \right]^{1/q} \qquad (41)
\end{aligned}
$$

In order to proceed we need to distinguish two cases according if $3/k_2 - 1/2 > 0$ or $3/k_2 - 1/2 \le 0$ (see Corollary 3.8). In the first case we let E_0 be defined by $s = E_0^{1/k_2 - 1/2} \tau$. For $E > E_0$ we break the expectation Eq. (41) into two parts according to whether the paths satisfy the hypotheses of Corollary 3.8 or not. For the first part we use Corollary 3.8 and that $\int_0^s r^2(s) ds \ge \int_0^{t_E} r^2(s) \ge c E^{3/k_2 - 1/2}$; for the second part we use estimate (38) in Remark 3.9 on the probability of unlikely paths together with the fact that the exponential under the expectation in Eq. (41) is bounded by 1. We obtain for all x with $G(x) = E > E_0$ the bound

$$
\begin{aligned}
& \mathbf{E}_x \left[\exp\left(\theta (G(x(s)) - G(x)) \right) \right] \le \exp\left(\theta \gamma \mathrm{Tr}(T) t_{E_0} \right) \times \\
& \quad \times \left[\exp\left(-q\theta (1 - p\theta T_{\max}) c E^{\frac{3}{k_2} - \frac{1}{2}} \right) + a \exp\left(-\frac{\Omega(E)^2 \theta_0}{b \gamma t_E} \right) \right]^{1/q} \qquad (42)
\end{aligned}
$$

Choosing the set $U = \{x \, ; \, G(x) \le E_1\}$ with E_1 large enough we can make the term in Eq. (42) as small as we want.

If $3/k_2 - 1/2 \le 0$, for a given s and a given x with $G(x) = E$ we split the time interval $[0, s]$ into $E^{1/2 - 1/k_2}$ pieces $[t_j, t_{j+1}]$, each one of size of order $E^{1/k_2 - 1/2} s$. For the "good" paths, i.e., for the paths $x(t)$ which satisfy the hypotheses of Corollary 3.8 on each time interval $[t_j, t_{j+1}]$, the tracking estimates of Proposition 3.7 imply that $G(x(t)) = O(E)$ for t in each interval. Applying Corollary 3.8 and using that $G(x(t_j)) = O(E)$ we conclude that $\int_0^s r^2(s)ds$ is at least of order $E^{3/k_2-1/2} \times E^{1/2-1/k_2} = E^{2/k_2}$. The probability of the remaining paths can be estimated, using Eq. (38), not to exceed

$$1 - \left(1 - a\exp\left(-\frac{\Omega_{\max}^2 \theta_0}{b\gamma t_E}\right)\right)^{E^{\frac{1}{2}-\frac{1}{k_2}}} .$$

The remainder of the argument is essentially as above, Eq. (42) and this concludes the proof of Theorem 3.10. ∎

The existence of the Liapunov function given by Eq.(39) can be interpreted in terms of hitting times. Let τ_U be the time for the diffusion $x(t)$ to hit the set U.

Theorem 3.11 *Assume that* $\theta < (\max\{T_1, T_n\})^{-1}$. *For any (arbitrarily large)* $a > 0$ *there exist a constant* $E_0 = E_0(a) > 0$ *such that for* $U = \{x\,;\, G(x) \leq E_0\}$ *and* $x \in U^C$ *we have*

$$\mathbf{E}_x\left[e^{a\tau_U}\right] < e^a + (e^a - 1)\exp\left(\theta(G(x) - E_0)\right). \tag{43}$$

Proof: Let $s = 1$ and $\theta < \theta_0$ be given, we set $\kappa = \exp(-a)/2$ and take U to be the set given by Thm. 3.10. Let X_n be the Markov chain defined by $X_n = x(n)$ and N_U be the least integer such that $X_{N_U} \in U$. Then

$$\mathbf{E}_x[e^{a\tau_U}] \quad \leq \quad \mathbf{E}_x[e^{aN_U}], \tag{44}$$

so that to estimate the exponential hitting time, it suffices to estimate the exponential "step number".

Using Chernov's inequality we obtain

$$\begin{aligned}
\mathbf{P}_x\{N_U > n\} \quad &= \quad \mathbf{P}_x\{-\sum_{j=1}^{n}(G(X_j) - G(X_{j-1}) < G(x) - E_0, X_j \in U^c\} \\
&\leq \quad e^{\theta(G(x)-E_0)}\mathbf{E}_x\left[\prod_{j=1}^{n} e^{\theta(G(X_j)-G(X_{j-1}))}, X_j \in U^c\right] \\
&\leq \quad e^{\theta(G(x)-E_0)}\mathbf{E}_x\left[\prod_{j=1}^{n-1} e^{\theta(G(X_j)-G(X_{j-1}))} \times\right.
\end{aligned}$$

$$
\begin{aligned}
& \qquad\qquad \times \mathbf{E}_{X_{n-1}}\left[e^{\theta(G(X_n)-G(X_{n-1})}\right],\ X_j \in U^c\Bigg] \\
\leq\ & e^{\theta(G(x)-E_0)} \sup_{y\in U^c} \mathbf{E}_y[e^{\theta(G(X_1)-G(y))}] \times \\
& \qquad\qquad \times \mathbf{E}_x\left[\prod_{j=1}^{n-1} e^{\theta(G(X_j)-G(X_{j-1}))}, X_j \in U^c\right] \\
\leq\ & \cdots \leq e^{\theta(G(x)-E_0)} \left(\sup_{y\in U^c} \mathbf{E}_y[e^{\theta(G(X_1)-G(y))}]\right)^n .
\end{aligned}
$$

By Thm. 3.10 we have

$$
\sup_{x\in U^c} \mathbf{E}_x[e^{\theta(G(X_1)-G(x))}] \,<\, \kappa\,,
$$

and therefore we have geometric decay of $P_{>n} \equiv \mathbf{P}_x\{N_U > n\}$ in n, $P_{>n} \leq \kappa^n \exp\left(\theta G(x) - E_0\right)$. Summing by parts we obtain

$$
\begin{aligned}
\mathbf{E}_x\left[e^{aN_U}\right] &= \sum_{n=1}^{\infty} e^{an}\mathbf{P}_x\{\tau_U = n\} \\
&= \lim_{M\to\infty}\left[\sum_{n=1}^{M} P_{>n}(e^{a(n+1)} - e^{an}) + e^a P_{>0} - e^{a(M+1)}P_{>M}\right]
\end{aligned}
$$

which, together with Eq. (44) gives Eq. (43). ■

4 Accessibility and Strong Feller Property

In this section we prove that the Markov process is strong Feller and moreover we show that it is strongly aperiodic in the sense that for all $t > 0$, all $x \in X$ and all open sets $A \subset X$ we have $P_t(x, A) > 0$. Both results imply immediately that $x(t)$ has at most one invariant measure: Since the process is strong Feller the invariant measure (if it exists) has a smooth density which is everywhere positive by the property of aperiodicity. Obviously no two different such measures can exist.

The strong Feller property is an immediate consequence of the hypoelliptic properties of the generator L of the diffusion. The result is an easy consequence of the estimates in [7, 5], since there much stronger global hypoelliptic estimates are proven (under stronger conditions on the potential $U^{(2)}$). We present here the argument for completeness.

The generator of the Markov process $x(t)$ can be written in the form

$$L = \sum_{i=1}^{2d} X_i^2 + X_0 .$$

If the Lie algebra generated by the set of commutators

$$\{X_i\}_{i=1}^{2d}, \quad \{[X_i, X_i]\}_{i,j=0}^{2d}, \quad \{[[X_i, X_j], X_k]\}_{i,j,k=0}^{2d}, \quad \cdots \tag{45}$$

has rank $\dim(X)$ at every point $x \in X$, then the Markov process has a $\mathcal{C}^\infty$ law. In particular it is strong Feller. This is a consequence of Hörmander Theorem [11, 16] or it can be proved directly using Malliavin Calculus developed by Malliavin, Bismut, Stroock and others (see e.g. [19]).

Proposition 4.1 *If* **H2** *holds then the generator L given by Eq. (13) satisfies the rank condition (45).*

Proof: This is a straightforward computation. The vector fields X_i, $i = 1, \cdots 2d$ gives $\partial_{r_i^{(j)}}$, $i = 1, n$, $j = 1, \cdots, d$. The commutators

$$\begin{aligned} \left[\partial_{r_1^{(j)}}, X_0\right] &= \gamma\partial_{r_1^{(j)}} - \lambda\partial_{p_1^{(j)}}, \\ \left[\left[\partial_{r_1^{(j)}}, X_0\right], X_0\right] &= \gamma^2\partial_{r_1^{(j)}} - \gamma\lambda\partial_{p_1^{(j)}} - \lambda\partial_{q_1^{(j)}}, \end{aligned}$$

yield the vector fields $\partial_{p_1^{(j)}}$ and $\partial_{q_1^{(j)}}$. Further

$$\left[\partial_{q_1^{(j)}}, X_0\right] = \sum_{l=1}^{d} \frac{\partial^2 V}{\partial_{q_1^{(j)}}\partial_{q_1^{(l)}}}(q)\partial_{p_1^{(l)}} + \sum_{l=1}^{d} \frac{\partial^2 U^{(2)}}{\partial_{q_1^{(j)}}\partial_{q_2^{(l)}}}(q_1 - q_2)\partial_{p_2^{(l)}} .$$

If $U^{(2)}$ is strictly convex, this yields $\partial_{p_2^{(j)}}$ while in the general case we need to consider further the commutators

$$\begin{aligned} &\left[\partial_{q_1^{(j_1)}}, \left[\cdots, \left[\partial_{q_1^{(j_{m-1})}}, \sum_{l=1}^{d} \frac{\partial^2 U^{(2)}}{\partial_{q_1^{(j_m)}}\partial_{q_2^{(l)}}}(q_1 - q_2)\partial_{p_2^{(l)}}\right]\right]\right] \\ &= \sum_{l=1}^{d} \frac{\partial^{m+1} U^{(2)}}{\partial_{q_1^{(j_1)}}\cdots\partial_{q_1^{(j_m)}}\partial_{q_1^{(l)}}}(q_1 - q_2)\partial_{p_2^{(l)}} . \end{aligned}$$

The condition **H3** means that we can write $\partial_{p_2^{(j)}}$ as a linear combination of these commutators for every $x \in X$. The other basis elements of the tangent space are obtained inductively following the same procedure. ∎

We now prove the strong aperiodicity of the process $x(t)$. This is based on the support theorem of Stroock and Varadhan [27]. The support of the diffusion process $x(t)$ with initial condition x on the time interval $[0,t]$, is by definition the smallest closed subset $S_{x,t}$ of $\mathcal{C}([0,t])$ such that $\mathbf{P}_x[x(t,\omega) \in S_{x,t}] = 1$. The support can be studied using the associated control system, i.e., the ordinary differential equation where the white noise $\dot{\omega}(t)$ is replaced by a control $u(t) \in L^1([0,T])$: For our problem we have the control system

$$\begin{aligned} \dot{q} &= p\,, \\ \dot{p} &= -\nabla_q V + \Lambda^T r\,, \\ \dot{r} &= (-\gamma r + \Lambda p) + u\,, \end{aligned} \tag{46}$$

and we denote $x_u(t)$ the solution of this control system with initial condition x and control u. The support theorem asserts that the support of the diffusion $S_{x,t}$ is the closure of the set $\{x_u\,;\, u \in L^1([0,t])\}$. As a consequence $\operatorname{supp} P_t(x,\cdot)$, the support of the transition probabilities is equal to the closure of the set of accessible points $\{y\,;\, \exists u \in L^1([0,t]) \text{ s.t. } x_u(t) = y\}$.

Proposition 4.2 *If condition* **H1** *holds then for all $t > 0$, all $x \in X$*

$$\operatorname{supp} P_t(x,\cdot) \;=\; X\,. \tag{47}$$

Proof: This result is proved in [7] under the additional condition that the interaction potential $U^{(2)}$ is strictly convex, in particular $\nabla U^{(2)}$ is a diffeomorphism. Our condition **H1** implies that $\nabla U^{(2)}$ is surjective. We can choose an inverse $g : \mathbf{R}^d \to \mathbf{R}^d$ which is locally bounded. From this point the proof proceeds exactly as in Theorem 3.2 of [7] and we will not repeat it here.

5 Proof of Theorem 2.1

The proof of Theorem 2.1 is a consequence of the theory linking the ergodic properties of Markov process with existence of Liapunov functions, a theory which has been developed over the past twenty years. The proof of these ergodic properties relies on the intuition that the compact set U together with a Liapunov function plays much the same role as an atom in, say, a countable state space Markov chain. The technical device to implement this idea was invented in [1, 20], and is called *splitting* It consists in constructing a new Markov chain with state space $X_0 \cup X_1$, where X_i are two copies of the original state space X. The new chain possesses an atom and has a projection being the original chain. The ergodic properties of a chain with an atom are then analyzed by means of *renewal theory* and a *coupling argument* applied to the return times to the atom. A complete account of this theory for a discrete time Markov process is developed in the book of Meyn and Tweedie [18], from where the result needed here is taken (Chapter 15).

For a given $s > 0$ consider the discrete time Markov chain $X_j = x(js)$ with transition probabilities $P(x, dy) \equiv P_s(x, dy)$ and semigroup $P^j \equiv T^{js}$. By the results of Section 4, the Markov chain is strongly aperiodic, i.e., $P(x, A) > 0$ for any open set A and for any x and it is strong Feller. The exponential bound on the hitting time given in Theorem 3.11 implies in particular that $\mathbf{E}_x[\tau_U]$ is finite for all $x \in X$ and thus we have an invariant measure μ (for hypoelliptic diffusions this is established in [14]). By aperiodicity and the strong Feller property, this invariant measure is unique.

The following Theorem is proved in [18]:

Theorem 5.1 *If the Markov chain $\{X_j\}$ is strong Feller and strongly aperiodic and if there is a function $W > 1$, a compact set U and $\kappa < 1$ and $L < \infty$ such that*

$$PW(x) \le \kappa W(x) + L\mathbf{1}_U(x)\,, \tag{48}$$

then there exist constants $r > 1$ and $R < \infty$ such that, for any x,

$$\sum_n r^n \|P(x, \cdot) - \mu\|_W \le RW(x)\,,$$

where the weighted variation norm $\|\cdot\|_W$ is defined in Eq. (14).

By Theorem 3.10 the assumptions of Theorem 5.1 are satisfied with $W = \exp(\theta G)$ and $\theta < (\max\{T_1, T_n\})^{-1}$. For the semigroup T^t we note that we have the apriori estimate $T^t \exp(\theta G)(x) \le \exp(\gamma\theta \mathrm{Tr}(T)t)\exp(\theta G)(x)$, cf. Lemma 3.5 which shows that T^t is a bounded operator on $L^\infty_\theta(X)$ defined in Eq. (15). Setting $t = ns + u$ with $0 \le u < s$, and using the invariance of μ one obtains

$$\|T^t - \mu\|_\theta \le \|T^{n\tau} - \mu\|_\theta \|T^s\|_\theta \le \tilde{R}\tilde{r}^{-t}\,, \tag{49}$$

for some $\tilde{r} > 1$ and $\tilde{R} < \infty$ or equivalently

$$\int_0^\infty \tilde{r}^t \|P_t(x, \cdot) - \mu\|_{\exp(\theta G)} \le \tilde{R}\exp(\theta G(x))\,.$$

As a consequence, for any $s > 0$, T^s has 1 as a simple eigenvalue and the rest of the spectrum is contained in a disk of radius $\rho < 1$. The exponential decay of correlations in the stationary states follows from this.

Corollary 5.2 *There exist constants $R < \infty$ and $r > 1$ such that for all f, g with f^2, $g^2 \in L^\infty_\theta(X)$, we have*

$$\left| \int f T^t g \, d\mu - \int f \, d\mu \int g \, d\mu \right| \leq R \|f^2\|_\theta^{1/2} \|g^2\|_\theta^{1/2} r^{-t} .$$

Proof: If $f^2 \in L^\infty_\theta$, we have $|f(x)| \leq \|f^2\|_\theta^{1/2} \exp(\theta G(x)/2)$ and similarly for g. Further if Eq.(49) holds with $W = \exp(\theta G)$ it also holds for $\exp(\theta G/2)$ and thus for some $R_1 < \infty$ and $r_1 > 1$ we have

$$\left| T^t g(x) - \int g \, d\mu \right| \;\leq\; R_1 r_1^{-t} \|g^2\|_\theta^{1/2} \exp\left(\frac{\theta G(x)}{2} \right) .$$

Therefore we obtain

$$\begin{aligned} \left| \int f T^t g \, d\mu - \int f \, d\mu \int g \, d\mu \right| \;&\leq\; \int |f(x)| \left| T^t g(x) - \int g \, d\mu \right| d\mu \\ &\leq\; \left(\int \exp(\theta G) d\mu \right) R_1 r_1^{-t} \|f^2\|_\theta^{1/2} \|g^2\|_\theta^{1/2} . \end{aligned}$$

To conclude we need to show that $\int \exp(\theta G) d\mu < \infty$. This follows from Eq.(48) which we rewrite as

$$\epsilon \exp(\theta G(x)) \;\leq\; \exp(\theta G(x)) - P \exp(\theta G(x)) + L \mathbf{1}_U(x) ,$$

with $\epsilon = 1 - \kappa$. From this we obtain

$$\epsilon \frac{1}{N} \sum_{k=1}^{N} \exp(\theta G(X_k)) \leq \frac{1}{N} \exp(\theta G(x)) + L \frac{1}{N} \sum_{k=1}^{N} \mathbf{1}_U(X_k) . \tag{50}$$

By the Law of Large Numbers the r.h.s of Eq.(50) converges to $L\mu(U)$ which is finite, and thus $\int \exp(\theta G) \, d\mu$ is finite, too. ∎

This concludes the proof of Theorem 2.1.

Acknowledgements We would like to thank Pierre Collet, Jean-Pierre Eckmann, Servet Martinez and Claude-Alain Pillet for their comments and suggestions as well as Martin Hairer for useful comments on the controllability issues discussed in Section 4.

References

[1] Athreya, K.B. and Ney, P.: A new approach to the limit theory of recurrent Markov chains. Trans. Am. Math. Soc. **245**, 493–501 (1978)

[2] Bambusi, D.: Exponential stability of breathers in Hamiltonian networks of weakly coupled oscillators. Nonlinearity **9**, 433-457 (1996)

[3] Bach, V., Fröhlich, J., and Sigal, I. M.: Quantum electrodynamics of confined nonrelativistic particles. Adv. Math. **137**, 299–395 (1998)

[4] Dym H. and McKean, H.P.: *Gaussian processes, function theory, and the inverse spectral problem.* Probability and Mathematical Statistics, Vol. **31**, New York–London: Academic Press, 1976

[5] Eckmann, J.P. and Hairer, M.: Non-equilibrium statistical mechanics of strongly anharmonic chains of oscillators. Commun. Math. Phys. **212**, 105–164 (2000)

[6] Eckmann, J.-P., Pillet C.-A., and Rey-Bellet, L.: Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures. Commun. Math. Phys. **201**, 657–697 (1999)

[7] Eckmann, J.-P., Pillet, C.-A., and Rey-Bellet, L.: Entropy production in non-linear, thermally driven Hamiltonian systems. J. Stat. Phys. **95**, 305–331 (1999)

[8] Ford, G.W., Kac, M. and Mazur, P.: Statistical mechanics of assemblies of coupled oscillators. J. Math. Phys. **6**, 504–515 (1965)

[9] Gallavotti, G. and Cohen E.G.D.: Dynamical ensembles in stationary states. J. Stat. Phys. **80**, 931–970 (1995)

[10] Has'minskii, R.Z.: *Stochastic stability of differential equations.* Alphen aan den Rijn—Germantown: Sijthoff and Noordhoff, 1980

[11] Hörmander, L.: *The Analysis of linear partial differential operators.* Vol **III**, Berlin: Springer, 1985

[12] Jakšić, V. and Pillet, C.-A.: Ergodic properties of classical dissipative systems. I. Acta Math. **181**, 245–282 (1998)

[13] Jakšić, V. and Pillet, C.-A.,: On a model for quantum friction. III. Ergodic properties of the spin-boson system. Commun. Math. Phys. **178**, 627–651 (1996)

[14] Kliemann, W.: Recurrence and invariant measures for degenerate diffusions. Ann. of Prob. **15**, 690–702 (1987)

[15] Komech, A., Spohn, H., and Kunze, M.: Long-time asymptotics for a classical particle interacting with a scalar wave field. Comm. Partial Differential Equations **22**, 307–335 (1997)

[16] Kunita, H. Supports of diffusion processes and controllability problems. In *Proc. Intern. Symp. SDE Kyoto 1976*, New York: Wiley, 1978, pp. 163–185

[17] MacKay, R.S. and Aubry, S.: Proof of existence of breathers for time-reversible or Hamiltonian networks of weakly coupled oscillators. Nonlinearity **7**, 1623–1643 (1994)

[18] Meyn, S.P. and Tweedie, R.L.: *Markov Chains and Stochastic Stability.* Communication and Control Engineering Series, London: Springer-Verlag London, 1993

[19] Norriss, J.: Simplified Malliavin Calculus. In *Séminaire de probabilités XX*, Lectures Note in Math. **1204**, 0 Berlin: Springer, 1986, pp. 101–130

[20] Nummelin, E.: A splitting technique for stationary Markov Chains. Z. Wahrscheinlichkeitstheorie Verw. Geb. **43**, 309–318 (1978)

[21] Rey-Bellet, L. and Thomas, L.E.: Asymptotic behavior of bhermal bon-equilibrium steady states for a driven chain of anharmonic oscillators. Commun. Math. Phys. **215**, 1–24 (2000)

[22] Rieder, Z., Lebowitz, J.L., and Lieb, E.: Properties of a harmonic crystal in a stationary non-equilibrium state. J. Math. Phys. **8**, 1073–1085 (1967)

[23] Ruelle, D.: Smooth dynamics and new theoretical ideas in non-equilibrium statistical mechanics. J. Stat. Phys. **95**, 393–468 (1999)

[24] Ruelle, D.: Natural non-equilibrium states in quantum statistical mechanics. J. Stat. Phys. **98**, 57–75 (2000)

[25] Spohn, H. and Lebowitz, J.L.: Stationary non-equilibrium states of infinite harmonic systems. Commun. Math. Phys. **54**, 97–120 (1977)

[26] Sievers, A.J. and Takeno, S.: Intrinsic localized modes in anharmonic crystals. Phys. Rev. Lett. **61** 970–973 (1988)

[27] Stroock, D.W. and Varadhan, S.R.S.: On the support of diffusion processes with applications to the strong maximum principle. In *Proc. 6-th Berkeley Symp. Math. Stat. Prob.*, Vol **III**, Berkeley: Univ. California Press, 1972, pp. 361–368