MIXING RATES OF PARTICLE SYSTEMS WITH ENERGY EXCHANGE

ALEXANDER GRIGO, KONSTANTIN KHANIN, AND DOMOKOS SZASZ

1. Introduction

1.1. **Motivation and related works.** A fundamental problem of non-equilibrium statistical mechanics is the derivation of effective equations in the hydrodynamic limit. Often these are hydrodynamic equations (Euler, Navier-Stokes), or related transport equations (Burgers equation, heat equation). There are very few models for which rigorous results exist. They include particle models like simple exclusion, zero range processes, see [KL99] and references therein, and continuous systems like the Ginzburg-Landau equation [Fri87b, Fri87a, GPV88] and the model of [KMP82].

The rigorous study of hydrodynamic limits requires detailed information about rates of convergence to equilibrium for finite sized systems, especially if the system is of non-gradient type. In particular, the scaling of the spectral gap of the generator with the system size N is of crucial importance. Such information is typically obtained by analyzing the Dirichlet form, corresponding to the explicitly known stationary distributions.

Obtaining good estimates (in terms of the system size) on the spectral gap of the generator is highly non-trivial. For example, to obtain the corresponding results for the Kac model [Kac56] it took almost half a century [Jan01] (using Yau's martingale method [Yau96, Yau97]) and [CCL03, CGL08].

Recently there has been a growing interest in and hope for establishing hydrodynamic limits for systems that are either purely deterministic or originate (somehow) from deterministic, in particular mechanical, models. A program to obtain information about the stationary distributions under the influence of stochastic boundary conditions was proposed in [EY06]. Another approach was suggested in the recent series of papers [GG08a, GG09], where the analysis of the hydrodynamic limit of a billiards lattice model was outlined by following a two-step procedure. In the first step the deterministic dynamics is rescaled in order to obtain a mesoscopic stochastic model (also referred to as master equation). In a second step the hydrodynamic behavior of the mesoscopic stochastic model should be derived.

For neither of the two steps proposed in [GG08a, GG09] rigorous results are available. Deriving master equations from interacting mechanical models is a very

---

Date: August 15, 2011.

A.G. gratefully acknowledges the generous support and warm hospitality of the Fields Institute and the Mathematical Institute of the Budapest University of Technology and Economics.

D. Sz. is grateful to Hungarian National Foundation for Scientific Research grants No. T 046187, K 71693, NK 63066 and TS 049835. D. Sz. expresses his sincere thanks to Department of Mathematics of University of Toronto and to the Fields Institute, Toronto where part of this work has been done (Fall term 2009 and June 2011, resp.).

difficult problem. Only recently some rigorous results in this direction were obtained in [DL10], where the weak interaction limit is considered opposed to the rare interaction limit of [GG08a, GG09]. As a matter of fact, the second step, i.e. deriving the hydrodynamic limit from the master equation, seems a much more tractable mathematical problem. The present paper is an attempt to make a first step in this direction by providing information about the spectral gap of the generator of an entire class of models, which are of similar type as the master equation of the billiard lattice model considered in [GG08a, GG09]. In particular, the model [GG09] belongs to the class of models we are considering, the obtained spectral bound is exactly the necessary one for which the derivation of the hydrodynamics limit is feasible.

I changed the wording of Doma's addition a bit; Aug 15.

1.2. **Description of the model.** The model we consider in this paper is as follows. Let $N \geq 2$ be an integer, and consider the lattice $\{1, 2, \ldots, N\}$. To every site i of this lattice we associate an energy x_i, which is a positive real number. The collection of all the energies will be denoted by $x = (x_1, \ldots, x_N) \in \mathbb{R}_+^N$. To each nearest neighbor pair of the lattice we associate an independent exponential clock with a rate Λ that depends on the total energy of this pair. As soon as one of the $N-1$ clocks rings, say for the pair $(i, i+1)$, then a number $0 \leq \alpha \leq 1$ is drawn, independently of everything else, according to a distribution P, that only depends on the two energies x_i, x_{i+1}. The update of the energies is then such that the new energy at site i is $\alpha\,(x_i + x_{i+1})$, the new energy at site $i+1$ is $(1-\alpha)\,(x_i + x_{i+1})$, and all other energies remain unchanged.

This procedure defines a continuous time Markov jump process $\mathsf{X}(t)$ on $\mathbb{R}_+^N$. More formally, we define the process $\mathsf{X}(t)$ by its infinitesimal generator $\mathcal{L}$, acting on bounded [1] functions $A\colon \mathbb{R}_+^N \to \mathbb{R}$ as

$$\mathcal{L}A(x) = \sum_{i=1}^{N-1} \Lambda(x_i, x_{i+1}) \int P(x_i, x_{i+1}, d\alpha)\,[A(T_{i,\alpha}x) - A(x)] \tag{1}$$

where $\Lambda\colon \mathbb{R}_+^2 \to \mathbb{R}_+$ is continuous, and $P(x_i, x_{i+1}, d\alpha)$ is a probability measure on $[0,1]$, which depends continuously on $(x_i, x_{i+1}) \in \mathbb{R}_+^2$. The maps $T_{i,\alpha}$ model the energy exchange between the neighboring sites i and $i+1$, and are defined by

$$T_{i,\alpha}(x) = x + [\alpha\, x_{i+1} - (1-\alpha)\, x_i]\,[\mathbf{e}_i - \mathbf{e}_{i+1}] \tag{2}$$

where $\mathbf{e}_i$ denotes the i-th unit vector of $\mathbb{R}^N$.

In particular, the process $\mathsf{X}(t)$ preserves the total energy, i.e. for any two times t_1 and t_2 the identity $\sum_{i=1}^N \mathsf{X}_i(t_1) = \sum_{i=1}^N \mathsf{X}_i(t_2)$ holds. Therefore, we introduce for any $\epsilon > 0$ [2] the sets

$$\mathcal{S}_{\epsilon,N} = \left\{ x \in \mathbb{R}_+^N : \sum_{i=1}^N \frac{1}{N}\, x_i = \epsilon \right\}$$

which are invariant for the process $\mathsf{X}(t)$. The value of ϵ represents the mean energy per site.

[1] Throughout this paper we will always assume that the various functions are Borel measurable without stating this assumption explicitly. This will not lead to confusion, since higher regularity assumptions (like continuity or integrability) are stated explicitly.

[2] The parameter ϵ denotes the average energy per site and should not be thought of as a necessarily small number. We hope that this does not cause any confusion, even though it is a common practice to reserve the use of the symbol ϵ to denote a small number.

Since $\mathcal{S}_{\epsilon,N}$ is compact and invariant the assumed continuity of Λ and P guarantees the existence of at least one stationary distribution $\pi_{\epsilon,N}$ for $\mathsf{X}(t)$ on each $\mathcal{S}_{\epsilon,N}$. As we pointed out, the scaling of the rate of convergence towards the stationary distribution in terms of the lattice size N is of crucial importance in studying the hydrodynamic limit of this model rigorously.

1.3. **Outline of the paper.** The purpose of this paper is to present a dynamical and geometric approach to establish the scaling of the spectral gap of the generator (1) under rather general assumptions on the rate function Λ and transition kernel P. The strategy we adopt is as follows. In Section 3 we show that for a large class of rates Λ and transition operators P the scaling of the spectral gap of the corresponding generator (1) can be obtained by considering only the special case of a constant rate Λ and a state independent transition kernel P. The precise statement is formulated in Theorem 3.1, which we prove under the two key assumptions: the reversibility of the process $\mathsf{X}(t)$, and the existence of a lower bound on the rate function Λ. The requirement of a lower bound on the rate function seems to be a technical condition, but it cannot be removed at present.

In Section 5 we show that (a slight modification of) the three-dimensional stochastic billiard lattice model of [GG09] is a special case of the general model considered in the present paper, provided that one introduces a lower cut-off for the rate function originally considered in [GG09]. In particular, we show that it then follows that the spectral gap scales as $\mathcal{O}(N^{-2})$.

Since we assume reversibility of the stationary distribution to derive the spectral properties, we provide in Section 4 a classification of reversible stationary distributions of product type. Such measures are of particular interest in the hydrodynamic limit, and appear naturally in mechanical models and statistical mechanics in form of Gibbs measures. We show in Theorem 4.3 that if a model of the class (1) considered in this paper admits a reversible product distributions, then this measure must necessarily be a product Gamma-distributions (or a single atom). This is precisely the type of product measures considered in statistical mechanics for mechanical models.

The main part of the paper deals with establishing the scaling of the spectral gap of the generator for the process with constant rates Λ and state independent transition kernel P. This case is studied in Section 2. The key difference of our analysis when compared to the above mentioned related works is that instead of focusing directly on L^2 convergence, for example by analyzing the associated Dirichlet form, we first establish weak convergence towards a stationary distribution. For the later part it is crucial that this weak convergence is made quantitative in a sufficiently strong metric for the weak topology. For this purpose we use the Vaserstein distance and prove in Theorem 2.9 that there is an exponential rate of convergence of $\mathsf{X}(t)$ to equilibrium, which scales as $\mathcal{O}(N^{-2})$ in the system size N. The key step in the proof is to construct an adapted metric on the state space of $\mathsf{X}(t)$, for which the contraction property can be established. This requires special coordinates and a coupling argument, which is presented in Proposition 2.6 and Proposition 2.8.

The advantage of first establishing exponential convergence in the weak sense is that it allows to include very general transition kernels P (for example, non-absolutely continuous kernels), and does make reference to the invariant measure. Instead it relies on a very natural geometric property of the interaction mechanism of $\mathsf{X}(t)$.

In a second step we assume reversibility of the constructed unique invariant measure, and show that the L^2 convergence occurs at an exponential rate, which is explicitly related to the rate of convergence in Vaserstein metric. In particular, this shows that the spectral gap scales as $\mathcal{O}(N^{-2})$ in the lattice size N. The precise statement is given in Theorem 2.12, whose prove relies on the Kantorovich-Rubinstein duality property of the Vaserstein metric, see Lemma 2.11. This is another manifestation of the usefulness of the weak convergence in Vaserstein distance in the study of the spectral gap for interacting particle systems.

Section 6 contains final comments and conclusions.

2. Analysis of a special case

In this section we consider a special case of the class of processes defined by generators of the form (1). Namely we consider the case where the rate function Λ is constant, and the transition kernel P is state independent. In other words we consider a process $\mathsf{X}(t)$ with infinitesimal generator

$$\mathcal{L}A(x) = \Lambda \sum_{i=1}^{N-1} \int P(d\alpha)\,[A(T_{i,\alpha}x) - A(x)] \tag{3}$$

acting on the space of bounded observables $A\colon \mathbb{R}_+^N \to \mathbb{R}$.

As was already mentioned the process $\mathsf{X}(t)$ preserves the total energy. This implies that the process cannot have a unique stationary state on all of $\mathbb{R}_+^N$. However, we will show below that the restriction of the process to any of the invariant sets $\mathcal{S}_{\epsilon,N}$ has a unique stationary distribution.

The first step in this direction is to introduce more convenient coordinates on $\mathcal{S}_{\epsilon,N}$, which is the purpose of the next result.

Lemma 2.1 (x in terms of u)**.** *Let N and ϵ be fixed. Then any $x \in \mathcal{S}_{\epsilon,N}$ can be uniquely written as*

$$x = \epsilon\,\mathbf{1} + \sum_{i=1}^{N-1} u_i\,[\mathbf{e}_i - \mathbf{e}_{i+1}]$$

for some $u \in \mathbb{R}^{N-1}$, where $\mathbf{1}$ denote the vector $(1,\ldots,1)$. Furthermore, via this change of coordinates the set $\mathcal{S}_{\epsilon,N} \subset \mathbb{R}_+^N$ is in one-to-one correspondence with the set

$$\hat{\mathcal{S}}_{\epsilon,N} = \{u \in \mathbb{R}^{N-1} \colon -\epsilon \le u_1,\ u_{i-1} \le \epsilon + u_i,\ u_{N-1} \le \epsilon\}\ .$$

I chose a slightly different way to change this. **OK?**

Note that the vectors $\mathbf{e}_i - \mathbf{e}_{i+1}$ for $i = 1,\ldots,N-1$ span the simplex $\mathcal{S}_{\epsilon,N}$, but they are not mutually orthogonal. However, they almost are in the sense that any two of them are perpendicular as soon as they correspond to two values of i, which differ by at least 2.

In the following we will also need the inverse coordinate transformation, which expresses u in terms of x.

Lemma 2.2 (u in terms of x)**.** *Let $x \in \mathbb{R}_+^N$ be given. Then the corresponding ϵ is given by $\epsilon = \sum_{i=1}^N \frac{1}{N} x_i$, and the corresponding u is the solution to the discrete Poisson equation with Dirichlet boundary conditions*

$$u_{i-1} - 2\,u_i + u_{i+1} = x_{i+1} - x_i \qquad \text{for} \qquad i = 1,\ldots,N-1$$

where we formally set $u_0 \equiv u_N \equiv 0$*. More explicitly*

$$u_i = \sum_{k=1}^{i}(x_k - \epsilon) = \Big[1 - \frac{i}{N}\Big]\sum_{k=1}^{i} x_k - \frac{i}{N}\sum_{k=i+1}^{N} x_k \qquad \textit{for all} \qquad 1 \le i \le N-1$$

is the expression for the corresponding $u \in \mathbb{R}^{N-1}$*.*

Proof. Clearly, $x \in \mathcal{S}_{\epsilon,N}$ if and only if ϵ is given by the claimed formula. Furthermore, it follows immediately from the definition of the coordinates u, that $x_i = \epsilon + u_i - u_{i-1}$ for all i, where we use the convention $u_0 \equiv u_N \equiv 0$. This implies that u must solve the discrete Poisson equation with zero Dirichlet boundary conditions.

On the other hand we can sum up the expression for x_i in terms of u and obtain a telescoping sum, which yields

$$u_i = \sum_{k=1}^{i}(u_k - u_{k-1}) = \sum_{k=1}^{i}(x_k - \epsilon)$$

for all $i = 1, \dots, N-1$.

And since $\epsilon\, N = \sum_{i=1}^{N} x_i$ we can replace ϵ in terms of this sum, and thus obtain the second expression for u_i. □

The point of the change of coordinates from x to ϵ and u is to separate out the conserved quantity ϵ, and consider only the evolution of the nontrivial part $\mathsf{U}(t)$ of the process $\mathsf{X}(t)$

$$\mathsf{X}(t) = \epsilon\,\mathbf{1} + \sum_{i=1}^{N-1} \mathsf{U}_i(t)\,[\mathbf{e}_i - \mathbf{e}_{i+1}]\,, \tag{4}$$

namely the u-coordinate vector corresponding to $\mathsf{X}(t)$. Since ϵ is conserved it follows that $\mathsf{U}(t)$ is also a homogeneous Markov process (for each ϵ separately). Using the results of Lemma 2.1 and Lemma 2.2 we can now derive the infinitesimal generator of $\mathsf{U}(t)$.

Lemma 2.3 (The generator of $\mathsf{U}(t)$)**.** *Let N and ϵ be fixed. Then the process* $\mathsf{U}(t)$ *is a homogeneous Markov process on* $\hat{\mathcal{S}}_{\epsilon,N}$*, whose infinitesimal generator* $\hat{\mathcal{L}}_{\epsilon,N}$ *is given by*

$$\hat{\mathcal{L}}_{\epsilon,N} A(u) = \Lambda \sum_{i=1}^{N-1} \int P(d\alpha)\,[A(\hat{T}^{\epsilon}_{i,\alpha} u) - A(u)]\,,$$

where

$$\hat{T}^{\epsilon}_{i,\alpha} u - u = [(1-\alpha)\,u_{i-1} + \alpha\,u_{i+1} + (2\,\alpha - 1)\,\epsilon - u_i]\,\mathbf{e}_i \in \mathbb{R}^{N-1}$$

Here $\mathbf{e}_i$ is in $\mathbb{R}^{N-1}$, before it was in $\mathbb{R}^N$. **OK?**

with the convention $u_0 \equiv u_N \equiv 0$*.*

Proof. From its definition (2) we have $T_{i,\alpha}(x) = x + [\alpha\, x_{i+1} - (1-\alpha)\, x_i]\,[\mathbf{e}_i - \mathbf{e}_{i+1}]$. Note that $[T_{i,\alpha} x]_k$ agrees with x_k for all k different from i and $i+1$, and $[T_{i,\alpha} x]_i + [T_{i,\alpha} x]_{i+1}$ equals $x_i + x_{i+1}$ (local energy conservation). Therefore, $[\hat{T}^{\epsilon}_{i,\alpha} u]_k$ equals u_k for all $k \neq i$, because by Lemma 2.2 we have $u_i = \sum_{k=1}^{i}(x_k - \epsilon)$.

So it remains to consider $[\hat{T}^{\epsilon}_{i,\alpha} u]_i$. Using the above two expressions for u and $T_{i,\alpha}(x)$ we obtain

$$\begin{aligned}[\hat{T}^{\epsilon}_{i,\alpha} u]_i - u_i &= \sum_{k=1}^{i} ([T_{i,\alpha}(x)]_k - \epsilon) - \sum_{k=1}^{i} (x_k - \epsilon) = [T_{i,\alpha}(x)]_i - x_i \\ &= \alpha\, x_{i+1} - (1-\alpha)\, x_i \ .\end{aligned}$$

Using Lemma 2.1 we can express x in terms of u as $x_i = \epsilon + u_i - u_{i-1}$, where we used the convention $u_0 \equiv u_N \equiv 0$. Substituting this expression in the previous formula yields the claimed expression for $\hat{T}^{\epsilon}_{i,\alpha} u - u$. Furthermore, this (trivially) also shows the claimed expression for the infinitesimal generator of $\mathsf{U}(t)$. ☐

2.1. **Weak convergence.** Fix again the values of ϵ and N. To study the existence of and rate of convergence to a stationary distribution we consider a bivariate Markov process $(\mathsf{U}(t), \mathsf{U}'(t))$ on $\hat{\mathcal{S}}_{\epsilon,N} \times \hat{\mathcal{S}}_{\epsilon,N}$, whose infinitesimal generator

$$\bar{\mathcal{L}} A(u,u') = \Lambda \sum_{i=1}^{N-1} \int P(d\alpha)\, [A(\hat{T}^{\epsilon}_{i,\alpha} u, \hat{T}^{\epsilon}_{i,\alpha} u') - A(u,u')] \tag{5}$$

for any (bounded) observable A on $\hat{\mathcal{S}}_{\epsilon,N} \times \hat{\mathcal{S}}_{\epsilon,N}$. Note that this is a special Markov coupling of two copies of the Markov chains generated by $\hat{\mathcal{L}}$.

In order to analyze the weak convergence of the process $\mathsf{X}(t)$ towards a stationary distribution we consider the Vaserstein metric on the probability measures on $\mathcal{S}_{\epsilon,N}$. This requires, however, a metric $\mathrm{d}(.,.)$ on $\mathcal{S}_{\epsilon,N}$. We equip $\hat{\mathcal{S}}_{\epsilon,N}$ with the Euclidean metric

$$\hat{\mathrm{d}}(u,u') := \Big[\sum_{i=1}^{N-1} (u_i - u'_i)^2 \Big]^{\frac{1}{2}} \tag{6a}$$

which corresponds to the metric

$$\mathrm{d}(x,x') = \Big[\sum_{i=1}^{N-1} \Big(\sum_{k=1}^{i} [x_k - x'_k] \Big)^2 \Big]^{\frac{1}{2}} \equiv \hat{\mathrm{d}}(u,u') \tag{6b}$$

on $\mathcal{S}_{\epsilon,N}$. In particular, we the have following estimate on the diameter of $\mathcal{S}_{\epsilon,N}$.

Lemma 2.4 (Diameter of $\mathcal{S}_{\epsilon,N}$)**.** *Let ϵ and N be fixed. Then*

$$\max_{x,x' \in \mathcal{S}_{\epsilon,N}} \mathrm{d}(x,x') = \max_{u,u' \in \hat{\mathcal{S}}_{\epsilon,N}} \hat{\mathrm{d}}(u,u') \le \epsilon\, N \sqrt{N-1}$$

holds.

Proof. By Lemma 2.1 it follows that for any $u \in \hat{\mathcal{S}}_{\epsilon,N}$ the inequality $-i\,\epsilon \le u_i \le \epsilon\,(N-i)$ holds for all $i = 1, \ldots, N-1$. Therefore,

$$\hat{\mathrm{d}}(u,u')^2 = \sum_{i=1}^{N-1} (u_i - u'_i)^2 \le \sum_{i=1}^{N-1} (N\,\epsilon)^2 = \epsilon^2\, N^2\, (N-1)$$

for any two u and u', which implies the claim. ☐

May be we should call this C^{N+1}, so that it will always C^N below instead of C^{N-1}.

The following Proposition 2.6 provides the first step to estimate $\hat{\mathrm{d}}(\mathsf{U}(t), \mathsf{U}'(t))$. A particular role will be played by the matrix

$$\mathcal{C}^{(N)} = \left(\begin{array}{c|ccccc|c} 2 & 0 & -1 & 0 & 0 & & \\ \hline 0 & 2 & 0 & -1 & 0 & & \\ -1 & 0 & 2 & 0 & -1 & & \\ & & & \ddots & & & \\ & -1 & 0 & 2 & 0 & & -1 \\ & 0 & -1 & 0 & 2 & & 0 \\ \hline & 0 & 0 & -1 & 0 & & 2 \end{array}\right) \in \mathbb{R}^{N\times N} . \tag{7}$$

The spectral properties of $\mathcal{C}^{(N)}$ are provided by the following Lemma 2.5.

Lemma 2.5 (Spectrum of $\mathcal{C}^{(N-1)}$). *If N is odd, then the eigenvalues of $\mathcal{C}^{(N-1)}$ are given by*

$$4\,\sin^2\Big[\frac{\pi\,k}{N+1}\Big] \qquad \textit{for} \qquad k = 1, \dots, \frac{N-1}{2}$$

where each has multiplicity two. If N is even, then the eigenvalues of $\mathcal{C}^{(N-1)}$ are given by

$$4\,\sin^2\Big[\frac{\pi\,k}{N}\Big] \qquad \textit{for} \qquad k = 1, \dots, \frac{N}{2} - 1$$
$$4\,\sin^2\Big[\frac{\pi\,k}{N+2}\Big] \qquad \textit{for} \qquad k = 1, \dots, \frac{N}{2}$$

each of multiplicity one.

Proof. By the definition of $\mathcal{C}^{(N)}$ we see that the even and odd indices separate. In fact, it is readily seen that the action of $\mathcal{C}^{(N)}$ on the odd indexed $(u_1, u_3, \dots)$ and the even indexed $(u_2, u_4, \dots)$ entries of u is given by the action of the matrix

$$A = \left(\begin{array}{c|ccc|c} 2 & -1 & 0 & 0 & \\ \hline -1 & 2 & -1 & 0 & \\ & & \ddots & & \\ & 0 & -1 & 2 & -1 \\ \hline & 0 & 0 & -1 & 2 \end{array}\right) .$$

It is readily verified that if $A \in \mathbb{R}^{m\times m}$, then for $k = 1, \dots, m$ the vectors $(\sin[\pi\,k\,\frac{1}{m+1}], \dots, \sin[\pi\,k\,\frac{m}{m+1}])$ are eigenvectors of A corresponding to the eigenvalues

$$4\,\sin^2\Big[\frac{\pi\,k}{2\,(m+1)}\Big] \qquad \text{for} \qquad k = 1, \dots, m\,.$$

If N is odd, say $N = 2\,m+1$ for some $m \geq 1$, then there are m odd and m even indexed entries in $u \in \mathbb{R}^{N-1}$. Therefore, the eigenvalues of $\mathcal{C}^{(2\,m)}$ are given by

$$4\,\sin^2\Big[\frac{\pi\,k}{2\,(m+1)}\Big] \qquad \text{for} \qquad k = 1, \dots, m$$

where each has multiplicity two.

If N is even, say $N = 2\,m+2$ for some $m \geq 1$, then there are $m+1$ odd and m even indexed entries in $u \in \mathbb{R}^{N-1}$. Therefore, the eigenvalues of $\mathcal{C}^{(2\,m+1)}$ are given by

$$4\,\sin^2\left[\frac{\pi\,k}{2\,(m+1)}\right] \qquad \text{for} \qquad k=1,\ldots,m$$
$$4\,\sin^2\left[\frac{\pi\,k}{2\,(m+2)}\right] \qquad \text{for} \qquad k=1,\ldots,m+1$$

where each has multiplicity one. □

Proposition 2.6 (Average contraction rate)**.** *Assume that the transition kernel P satisfies $\int P(d\alpha)\,\alpha = \frac{1}{2}$. Then*

$$\bar{\mathcal{L}}[\hat{\mathrm{d}}(u,u')^2] \le -\Lambda\,[1-4\,\sigma_P^2]\,\sin^2\left[\frac{\pi}{N+2}\right]\hat{\mathrm{d}}(u,u')^2$$

holds for any two states u and u', where σ_P^2 denotes the variance of P.

Remark 2.7. *Since P is supported on $[0,1]$ and is assumed to have mean $\int P(d\alpha)\,\alpha = \frac{1}{2}$ it follows that the variance of P satisfies $0 \le 1-4\,\sigma_P^2 \le 1$.*

Proof of Proposition 2.6. From the definition of the generator $\bar{\mathcal{L}}$ and the distance $\hat{\mathrm{d}}(.,.)$ it follows

$$\bar{\mathcal{L}}\hat{\mathrm{d}}(u,u')^2 = \Lambda\sum_{i=1}^{N-1}\int P(d\alpha)\,[\hat{\mathrm{d}}(\hat{T}^{\epsilon}_{i,\alpha}u,\hat{T}^{\epsilon}_{i,\alpha}u')^2 - \hat{\mathrm{d}}(u,u')^2]$$

and

$$\begin{aligned}\hat{\mathrm{d}}(\hat{T}^{\epsilon}_{i,\alpha}u,\hat{T}^{\epsilon}_{i,\alpha}u')^2 - \hat{\mathrm{d}}(u,u')^2 &= \sum_{k=1}^{N-1}\Big[([\hat{T}^{\epsilon}_{i,\alpha}u]_k - [\hat{T}^{\epsilon}_{i,\alpha}u']_k)^2 - (u_k-u'_k)^2\Big]\\ &= \sum_{k=1}^{N-1}\Big[[\hat{T}^{\epsilon}_{i,\alpha}u-u]_k - [\hat{T}^{\epsilon}_{i,\alpha}u'-u']_k\Big]\cdot\\ &\qquad\cdot\Big[[\hat{T}^{\epsilon}_{i,\alpha}u-u]_k - [\hat{T}^{\epsilon}_{i,\alpha}u'-u']_k + 2\,(u_k-u'_k)\Big]\,.\end{aligned}$$

Making use of the explicit expression for $\hat{T}^{\epsilon}_{i,\alpha}u-u$ provided by Lemma 2.3

$$\hat{T}^{\epsilon}_{i,\alpha}u-u = [(1-\alpha)\,u_{i-1}+\alpha\,u_{i+1}+(2\,\alpha-1)\,\epsilon-u_i]\,\mathbf{e}_i$$

the above sum simplifies to

$$\begin{aligned}&\hat{\mathrm{d}}(\hat{T}^{\epsilon}_{i,\alpha}u,\hat{T}^{\epsilon}_{i,\alpha}u')^2 - \hat{\mathrm{d}}(u,u')^2\\ &\qquad= \Big[(1-\alpha)\,[u_{i-1}-u'_{i-1}]+\alpha\,[u_{i+1}-u'_{i+1}]-[u_i-u'_i]\Big]\cdot\\ &\qquad\qquad\cdot\Big[[\hat{T}^{\epsilon}_{i,\alpha}u-u]_i-[\hat{T}^{\epsilon}_{i,\alpha}u'-u']_i+2\,(u_i-u'_i)\Big]\\ &\qquad= \Big[(1-\alpha)\,[u_{i-1}-u'_{i-1}]+\alpha\,[u_{i+1}-u'_{i+1}]-[u_i-u'_i]\Big]\cdot\\ &\qquad\qquad\cdot\Big[(1-\alpha)\,[u_{i-1}-u'_{i-1}]+\alpha\,[u_{i+1}-u'_{i+1}]+[u_i-u'_i]\Big]\\ &\qquad= \Big[(1-\alpha)\,[u_{i-1}-u'_{i-1}]+\alpha\,[u_{i+1}-u'_{i+1}]\Big]^2-[u_i-u'_i]^2\\ &\qquad= (1-\alpha)^2\,[u_{i-1}-u'_{i-1}]^2+\alpha^2\,[u_{i+1}-u'_{i+1}]^2\\ &\qquad\quad+2\,\alpha\,(1-\alpha)\,[u_{i-1}-u'_{i-1}]\,[u_{i+1}-u'_{i+1}]-[u_i-u'_i]^2\end{aligned}$$

which in particular shows that the above expression depends only on the difference vector $u - u'$.

Performing now the sum over i yields

$$\sum_{i=1}^{N-1} [\hat{\mathrm{d}}(\hat{T}^{\epsilon}_{i,\alpha} u, \hat{T}^{\epsilon}_{i,\alpha} u')^2 - \hat{\mathrm{d}}(u,u')^2] = (1-\alpha)^2 \sum_{i=1}^{N-2} [u_i - u'_i]^2 + \alpha^2 \sum_{i=2}^{N-1} [u_i - u'_i]^2 + \alpha\,(1-\alpha) \sum_{i=2}^{N-2} 2\,[u_{i-1} - u'_{i-1}]\,[u_{i+1} - u'_{i+1}] - \sum_{i=1}^{N-1} [u_i - u'_i]^2$$

where we made use of the convention $u_0 \equiv u_N \equiv u'_0 \equiv u'_N \equiv 0$.

Note now that the assumption $\int P(d\alpha)\,\alpha = \frac{1}{2}$ implies

$$\int P(d\alpha)\,\alpha^2 = \int P(d\alpha)\,(1-\alpha)^2 = \sigma_P^2 + \frac{1}{4}\,, \qquad \int P(d\alpha)\,\alpha\,(1-\alpha) = \frac{1}{4} - \sigma_P^2$$

and hence

$$\begin{aligned} \frac{1}{\Lambda}\,\bar{\mathcal{L}}[\hat{\mathrm{d}}(u,u')^2] = {} & \int P(d\alpha)(1-\alpha)^2 \sum_{i=1}^{N-2} [u_i - u'_i]^2 + \int P(d\alpha)\alpha^2 \sum_{i=2}^{N-1} [u_i - u'_i]^2 \\ & + \int P(d\alpha)\alpha\,(1-\alpha) \sum_{i=2}^{N-2} 2\,[u_{i-1} - u'_{i-1}]\,[u_{i+1} - u'_{i+1}] \\ & - \sum_{i=1}^{N-1} [u_i - u'_i]^2 \\ = {} & -\frac{1-4\,\sigma_P^2}{4} \Big[\sum_{i=1}^{N-1} 2\,[u_i - u'_i]^2 - \sum_{i=2}^{N-2} 2\,[u_{i-1} - u'_{i-1}]\,[u_{i+1} - u'_{i+1}] \Big] \\ & - \frac{1+4\,\sigma_P^2}{4} \Big[[u_1 - u'_1]^2 + [u_{N-1} - u'_{N-1}]^2 \Big]\,. \end{aligned}$$

It is now straightforward to verify that

$$\begin{aligned} \bar{\mathcal{L}}[\hat{\mathrm{d}}(u,u')^2] = {} & -\Lambda\,\frac{1-4\,\sigma_P^2}{4}\,[u-u']^T\,\mathcal{C}^{(N-1)}\,[u-u'] \\ & -\Lambda\,\frac{1+4\,\sigma_P^2}{4} \Big[[u_1 - u'_1]^2 + [u_{N-1} - u'_{N-1}]^2 \Big]\,, \end{aligned}$$

where the matrix $\mathcal{C}^{(N-1)}$ was defined in (7) above.

Observe that by Lemma 2.5 the smallest eigenvalue of $\mathcal{C}^{(N-1)}$ equals $4\,\sin^2[\frac{\pi}{N+1}]$ if N is odd, and $4\,\sin^2[\frac{\pi}{N+2}]$ if N is even. Therefore,

$$\begin{aligned} \bar{\mathcal{L}}[\hat{\mathrm{d}}(u,u')^2] & \le -\Lambda\,\frac{1-4\,\sigma_P^2}{4}\,[u-u']^T\,\mathcal{C}^{(N-1)}\,[u-u'] \\ & \le -\Lambda\,[1-4\,\sigma_P^2]\,\sin^2\Big[\frac{\pi}{N+2}\Big]\,\hat{\mathrm{d}}(u,u')^2 \end{aligned}$$

follows from the fact that $\mathcal{C}^{(N-1)}$ is a symmetric matrix, and $0 \le 1 - 4\,\sigma_P^2$. □

Let U and U' be any two random variables on $\hat{\mathcal{S}}_{\epsilon,N}$ with distribution denoted by μ and μ', respectively. Recall that for $p \ge 1$ the Vaserstein-p distance is defined by

$$\rho_p(\mathsf{U},\mathsf{U}') \equiv \rho_p(\mu,\mu') = \inf_{\Gamma} \Big[\int_{\hat{\mathcal{S}}_{\epsilon,N} \times \hat{\mathcal{S}}_{\epsilon,N}} \Gamma(du,du')\,\hat{\mathrm{d}}(u,u')^p \Big]^{\frac{1}{p}}\,,$$

where the infimum is taken over all couplings Γ of μ and μ'. To shorten the notation we set $\rho(\mu,\mu') \equiv \rho_1(\mu,\mu')$ in the special case $p=1$.

Proposition 2.8 (Rate of convergence in Vaserstein-2 distance)**.** *Assume that the transition kernel P satisfies $\int P(d\alpha)\,\alpha = \frac{1}{2}$. Let $\mathsf{U}(t)$ and $\mathsf{U}'(t)$ be two Markov chains generated by $\hat{\mathcal{L}}$ on $\hat{\mathcal{S}}_{\epsilon,N}$. Then for all $t \geq 0$*

$$\begin{aligned}\rho_2(\mathsf{U}(t),\mathsf{U}'(t)) &\leq \rho_2(\mathsf{U}(0),\mathsf{U}'(0))\,\exp\Big(-\frac{1}{2}\,\Lambda\,[1-4\,\sigma_P^2]\,\sin^2\Big[\frac{\pi}{N+2}\Big]\,t\Big)\\ &\leq \epsilon\,N\,\sqrt{N-1}\,\exp\Big(-\frac{1}{2}\,\Lambda\,[1-4\,\sigma_P^2]\,\sin^2\Big[\frac{\pi}{N+2}\Big]\,t\Big)\,.\end{aligned}$$

Proof. Denote the distribution of the bivariate Markov process $(\mathsf{U}(t),\mathsf{U}'(t))$ with generator $\bar{\mathcal{L}}$ by $\Gamma_t(du,du')$, and denote by $\mu_t(du)$ and $\mu'_t(du')$ the two marginals.

Observe that the generator $\bar{\mathcal{L}}$ of this bivariate process $(\mathsf{U}(t),\mathsf{U}'(t))$ is constructed in such a way that $\mathsf{U}(t)$ and $\mathsf{U}'(t)$ are Markov chains with generator $\hat{\mathcal{L}}$ whose distributions are given by $\mu_t(du)$ and $\mu'_t(du')$, respectively.

Therefore, $\Gamma_t(du,du')$ is a coupling of the two distributions $\mu_t(du)$ and $\mu'_t(du')$ for all $t\geq 0$. In particular,

$$\rho_2(\mathsf{U}(t),\mathsf{U}'(t))^2 \leq \int_{\hat{\mathcal{S}}_{\epsilon,N}\times\hat{\mathcal{S}}_{\epsilon,N}} \Gamma_t(du,du')\,\hat{\mathrm{d}}(u,u')^2$$

follows from the very definition of the Vaserstein distance.

By the Markov property of the bivariate chain

$$\hat{\mathrm{d}}(\mathsf{U}(t),\mathsf{U}'(t))^2 - \hat{\mathrm{d}}(\mathsf{U}(0),\mathsf{U}'(0))^2 - \int_0^t \bar{\mathcal{L}}\hat{\mathrm{d}}(\mathsf{U}(s),\mathsf{U}'(s))^2\,ds$$

is a centered martingale. Hence for all $t\geq 0$

$$\mathbb{E}\,\hat{\mathrm{d}}(\mathsf{U}(t),\mathsf{U}'(t))^2 = \mathbb{E}\,\hat{\mathrm{d}}(\mathsf{U}(0),\mathsf{U}'(0))^2 + \int_0^t \mathbb{E}\,\bar{\mathcal{L}}\hat{\mathrm{d}}(\mathsf{U}(s),\mathsf{U}'(s))^2\,ds\,.$$

Differentiating with respect to t and applying the estimate of Proposition 2.6 yields

$$\frac{d}{dt}\,\mathbb{E}[\hat{\mathrm{d}}(\mathsf{U}(t),\mathsf{U}'(t))^2] \leq -\Lambda\,[1-4\,\sigma_P^2]\,\sin^2\Big[\frac{\pi}{N+2}\Big]\;\mathbb{E}[\hat{\mathrm{d}}(\mathsf{U}(t),\mathsf{U}'(t))^2]\,.$$

Gronwall's inequality shows that

$$\begin{aligned}\rho_2(\mathsf{U}(t),\mathsf{U}'(t))^2 &\leq \mathbb{E}[\hat{\mathrm{d}}(\mathsf{U}(t),\mathsf{U}'(t))^2]\\ &\leq \exp\Big(-\Lambda\,[1-4\,\sigma_P^2]\,\sin^2\Big[\frac{\pi}{N+2}\Big]\,t\Big)\,\mathbb{E}[\hat{\mathrm{d}}(\mathsf{U}(0),\mathsf{U}'(0))^2]\end{aligned}$$

for any initial distribution Γ_0 of the bivariate chain.

Taking in the infimum over all couplings Γ_0 of μ_0 and μ'_0 yields

$$\begin{aligned}\rho_2(\mathsf{U}(t),\mathsf{U}'(t))^2 &\leq \exp\Big(-\Lambda\,[1-4\,\sigma_P^2]\,\sin^2\Big[\frac{\pi}{N+2}\Big]\,t\Big)\,\rho_2(\mathsf{U}(0),\mathsf{U}'(0))^2\\ &\leq \epsilon^2\,N^2\,(N-1)\,\exp\Big(-\Lambda\,[1-4\,\sigma_P^2]\,\sin^2\Big[\frac{\pi}{N+2}\Big]\,t\Big)\end{aligned}$$

where the second inequality is due to the estimate on the diameter of $\hat{\mathcal{S}}_{\epsilon,N}$ provided in Lemma 2.4. □

By definition of the metric $\mathrm{d}(.,.)$ on $\mathcal{S}_{\epsilon,N}$ in terms of $\hat{\mathrm{d}}(.,.)$ it follows immediately from Proposition 2.8 that there is at most one stationary distribution for $\mathsf{X}(t)$ on each $\mathcal{S}_{\epsilon,N}$, and that the rate of convergence in the associated Vaserstein distance is the same as the rate of convergence for $\mathsf{U}(t)$.

Furthermore, by assumption the process $\mathsf{X}(t)$ on $\mathcal{S}_{\epsilon,N}$ generated by $\mathcal{L}$ is stochastically continuous. Hence the compactness of $\mathcal{S}_{\epsilon,N}$ (in the topology induced by the chosen metric) allows us to apply the Bogolyubov-Krylov argument to show that there is at least one stationary distribution. This proves the following Theorem 2.9.

Theorem 2.9 (Ergodicity and mixing rate of $\mathsf{X}(t)$ on each $\mathcal{S}_{\epsilon,N}$.)**.** *If the transition kernel P satisfies $\int P(d\alpha)\,\alpha = \frac{1}{2}$, and $\sigma_P^2 < \frac{1}{4}$, then there exists a unique stationary distribution $\pi_{\epsilon,N}$ on $\mathcal{S}_{\epsilon,N}$. Furthermore,*

$$\begin{aligned}\rho_2(\mathsf{X}(t),\pi_{\epsilon,N}) &\le \rho_2(\mathsf{X}(0),\pi_{\epsilon,N})\,\exp\Big(-\frac{1}{2}\,\Lambda\,[1-4\,\sigma_P^2]\,\sin^2\Big[\frac{\pi}{N+2}\Big]\,t\Big)\\ &\le \epsilon\,N\,\sqrt{N-1}\,\exp\Big(-\frac{1}{2}\,\Lambda\,[1-4\,\sigma_P^2]\,\sin^2\Big[\frac{\pi}{N+2}\Big]\,t\Big)\end{aligned}$$

holds for all t, and any initial distribution of $\mathsf{X}(0)$ on $\mathcal{S}_{\epsilon,N}$.

2.2. $L^2_{\pi_{\epsilon,N}}$–**Spectral gap.** In order to analyse the spectrum of $\mathcal{L}$ in $L^2_{\pi_{\epsilon,N}}$ we will make an extra assumption on the invariant measure $\pi_{\epsilon,N}$. Recall that a measure μ is called reversible under $\mathcal{L}$ if for all bounded $f\colon \mathcal{S}_{\epsilon,N}\times\mathcal{S}_{\epsilon,N}\to\mathbb{R}$

$$\int \mu(dx)\,[\mathcal{L}f(.,x)](x) = \int \mu(dx)\,[\mathcal{L}f(x,.)](x) \tag{8}$$

holds. In particular, considering functions f of the form $f(x,x') = F(x)$ for some bounded $F\colon\mathcal{S}_{\epsilon,N}\to\mathbb{R}$ shows that μ must be invariant under $\mathcal{L}$.

Furthermore, $\mathcal{L}$ acts on L^2_μ as a bounded, self-adjoint negative semi-definite operator. An estimate on the size of its spectral gap is provided in Theorem 2.12 below. Because the result of the following Lemma 2.10 will play a central role in the proof of Theorem 2.12 we include the details of this well-known result for completeness.

Lemma 2.10 (Auxiliary estimate on the spectrum of a self-adjoint operator)**.** *Let H be a real (or complex) Hilbert space and $T\colon H\to H$ a bounded, self-adjoint linear operator. Suppose there exists a constant $0\le\gamma$ and a dense subspace $G\subset H$ on which for all $g\in G$ and $f\in H$ there exists a constant $C_{f,g}>0$ such that $|\langle f, T^n g\rangle| \le C_{f,g}\,\gamma^n$ for all $n\ge 1$. Then the spectrum of T is contained in $[-\gamma,\gamma]$.*

Proof. The classical spectral theory of bounded self-adjoint linear operators [Con90] states that the spectrum $\sigma(T)$ of T is a compact interval in $[-\|T\|,\|T\|]$, and there exists a unique spectral measure $E(d\lambda)$ such that for any $f,g\in H$

$$1 = \int_{\mathbb{R}} E(d\lambda)\,, \qquad T^n = \int_{\mathbb{R}} \lambda^n\,E(d\lambda)\,, \qquad \langle T^n f, g\rangle = \int_{\mathbb{R}} \lambda^n\,\langle E(d\lambda)f, g\rangle$$

where $E(d\lambda)$ is supported on $\sigma(T)$, and $m_{f,g}(d\lambda)\equiv\langle E(d\lambda)f,g\rangle$ is a finite signed measure on $\sigma(T)$, whose total variation norm satisfies $|m_{f,g}|_{\mathrm{TV}}\le\|f\|\,\|g\|$.

Suppose that the spectrum $\sigma(T)$ of T is not contained in $[-\gamma,\gamma]$. Then there exists $s>\gamma$ such that for $S_s=(-\infty,-s)\cup(s,\infty)$ the projection $E(S_s)$ is nonzero. Hence there exists a nonzero $f_s\in H$ with $E(S_s)f_s=f_s$. In particular,

Can we shorten this entire argument?

$$\| f_s \|^2 = \int_{\sigma(T)} m_{f_s,f_s}(d\lambda) = \int_{S_s} m_{f_s,f_s}(d\lambda) > 0 \,,$$

because the support of the measure $m_{f_s,f_s}(d\lambda)$ is contained in S_s by choice of f_s. In particular, $m_{f_s,f_s} \neq 0$.

For any $g \in G$, and all $n \geq 0$ we have

$$\frac{1}{\gamma^{2\,n}} \langle\, f_s, T^{2\,n} g \,\rangle = \frac{1}{\gamma^{2\,n}} \langle\, T^{2\,n} f_s, g \,\rangle = \int_{S_s} \left|\frac{\lambda}{\gamma}\right|^{2\,n} m_{f_s,g}(d\lambda) \,.$$

Due to the assumption on G we also have that

$$\left| \frac{1}{\gamma^{2\,n}} \langle\, f_s, T^{2\,n} g \,\rangle \right| \leq C_{f_s,g}$$

Since $m_{f_s,g}$ is a finite measure, and $|\frac{\lambda}{\gamma}| \geq \frac{s}{\gamma} > 1$ on its support, the boundedness of the above expression for all n can only be satisfied if in fact $m_{f_s,g} = 0$.

Thus we have shown that $m_{f_s,f_s} \neq 0$, but $m_{f_s,g} = 0$ for all $g \in G$. Since $m_{f_s,g}$ is continuous in g (in fact linear and bounded) the denseness of G implies that there exists a sequence $(g_n)_{n\geq 1} \subset G$ such that $g_n \to f_s$ in H, and hence $0 = m_{f_s,g_n} \to m_{f_s,f_s} \neq 0$. This is a contradiction to continuity. Therefore the assumption on s must have been wrong, so that for all $s > \gamma$ the projection $E(S_s)$ must be zero. And since $\lambda \in \mathbb{R}$ is in the resolvent set of T if and only if there exists an open neighborhood S of λ such that $E(S) = 0$ it follows that $\sigma(T) \subset [-\gamma, \gamma]$. □

Lemma 2.11 (Lipschitz contraction)**.** *Let $A\colon \mathcal{S}_{\epsilon,N} \to \mathbb{R}$ be a Lipschitz continuous function with respect to the distance* $\mathrm{d}(.,.)$*, and set $A_t(x) = \mathbb{E}[A(\mathsf{X}(t)) \,|\, \mathsf{X}(t) = x]$ for all $t \geq 0$ and $x \in \mathcal{S}_{\epsilon,N}$. Then A_t is Lipschitz continuous with Lipschitz constant*

$$\mathrm{Lip}(A_t) \leq \mathrm{Lip}(A)\, \exp\Big(-\frac{1}{2}\,\Lambda\,[1 - 4\,\sigma_P^2]\, \sin^2\Big[\frac{\pi}{N+2}\Big]\, t\Big)$$

for all $t \geq 0$.

Proof. By Jensen's inequality it follows immediately from the very definition of the Vaserstein distance that $\rho_{p_1}(\mathsf{X}(t), \mathsf{X}'(t)) \leq \rho_{p_2}(\mathsf{X}(t), \mathsf{X}'(t))$ for all $1 \leq p_1 \leq p_2$. Therefore it follows from Proposition 2.8 that

$$\rho_1(\mathsf{X}(t), \mathsf{X}'(t)) \leq \rho_2(\mathsf{X}(0), \mathsf{X}'(0))\, \exp\Big(-\frac{1}{2}\,\Lambda\,[1 - 4\,\sigma_P^2]\, \sin^2\Big[\frac{\pi}{N+2}\Big]\, t\Big)$$

for any joint distribution of $(\mathsf{X}(0), \mathsf{X}'(0))$ on $\mathcal{S}_{\epsilon,N} \times \mathcal{S}_{\epsilon,N}$.

Note that $\mathcal{S}_{\epsilon,N}$ is compact, and hence

$$\sup_{\mathrm{Lip}(A)\leq 1} |\, \mathbb{E}\, A(\mathsf{X}(t)) - \mathbb{E}\, A(\mathsf{X}'(t))| = \rho_1(\mathsf{X}(t), \mathsf{X}'(t))$$

which is the well-know Kantorovich-Rubinstein duality theorem for the Vaserstein-1 metric.

Using the specific initial distribution $(\mathsf{X}(0), \mathsf{X}'(0)) = (x, x')$ on $\mathcal{S}_{\epsilon,N} \times \mathcal{S}_{\epsilon,N}$ we obtain

$$\begin{aligned} |A_t(x) - A_t(x')| &\leq \mathrm{Lip}(A)\, \rho_1(\mathsf{X}(t), \mathsf{X}'(t)) \\ &\leq \mathrm{Lip}(A)\, \mathrm{d}(x, x')\, \exp\Big(-\frac{1}{2}\,\Lambda\,[1 - 4\,\sigma_P^2]\, \sin^2\Big[\frac{\pi}{N+2}\Big]\, t\Big) \end{aligned}$$

because in this case $\rho_2(\mathsf{X}(0), \mathsf{X}'(0)) = \mathrm{d}(x, x')$. And since $x, x' \in \mathcal{S}_{\epsilon,N}$ are arbitrary we see that A_t is Lipschitz continuous with the claimed estimate on its Lipschitz constant. □

Combining now the result of Lemma 2.11 with that of Lemma 2.10 we are in a position to estimate the spectral gap of $\mathcal{L}$ acting on $L^2_{\pi_{\epsilon,N}}$, provided we assume that the stationary distribution $\pi_{\epsilon,N}$ is reversible. In this case $\mathcal{L}$ is a self-adjoint, bounded, negative semi-definite operator on $L^2_{\pi_{\epsilon,N}}$.

Theorem 2.12 ($L^2_{\pi_{\epsilon,N}}$–spectral gap for reversible $\pi_{\epsilon,N}$)**.** *Suppose that P satisfies $\int P(d\alpha)\,\alpha = \frac{1}{2}$ and $\sigma_P^2 < \frac{1}{4}$. If the stationary distribution $\pi_{\epsilon,N}$ of $\mathsf{X}(t)$ on $\mathcal{S}_{\epsilon,N}$ is reversible, then*

$$\sigma(\mathcal{L}) \subset \Big(-\infty, -\frac{1}{2}\,\Lambda\,[1-4\,\sigma_P^2]\,\sin^2\Big[\frac{\pi}{N+2}\Big]\Big] \cup \{0\}\,,$$

where 0 is a simple eigenvalue corresponding to the constant eigenfunction.

Proof. By assumption $\mathcal{L}$ generates a self-adjoint, positive semi-definite contraction semigroup $e^{t\,\mathcal{L}}$ on $L^2_{\pi_{\epsilon,N}}$, which satisfies $e^{t\,\mathcal{L}}1 = 1$. Therefore, the subspace H of $L^2_{\pi_{\epsilon,N}}$ consisting of functions perpendicular to the constant functions is invariant. Hence, the decomposition $L^2_{\pi_{\epsilon,N}} = H \oplus \operatorname{span}\{1\}$ is invariant under $e^{t\,\mathcal{L}}$, and $e^{t\,\mathcal{L}}$ may be restricted to H.

Furthermore, it is a consequence of Lusin's theorem [Rud87] that the set of Lipschitz continuous functions on $\mathcal{S}_{\epsilon,N}$ is dense in $L^2_{\pi_{\epsilon,N}}$. Hence the set G of Lipschitz continuous functions A on $\mathcal{S}_{\epsilon,N}$ with $\int \pi_{\epsilon,N}(dx)\,A(x) = 0$ is dense in H.

By Lemma 2.4 and the mean value theorem, for any $f \in H$ and $g \in G$

$$|\langle\, f, g\,\rangle| \le \|\, f\,\|\;\|\, g\,\| \le \|\, f\,\|\;\operatorname{diam}\mathcal{S}_{\epsilon,N}\,\operatorname{Lip}(g) \le \|\, f\,\|\;\epsilon\,N\,\sqrt{N-1}\,\operatorname{Lip}(g)$$

and hence

$$|\langle\, f, e^{n\,t\,\mathcal{L}} g\,\rangle| \le \|\, f\,\|\;\epsilon\,N\,\sqrt{N-1}\,\operatorname{Lip}(g)\exp\Big(-\frac{1}{2}\,\Lambda\,[1-4\,\sigma_P^2]\,\sin^2\Big[\frac{\pi}{N+2}\Big]\,t\Big)^n$$

follows from Lemma 2.11 for all $n \ge 0$.

Since $e^{t\,\mathcal{L}}$ is a positive operator the result of Lemma 2.10 yields

$$\sigma(e^{t\,\mathcal{L}}|_H) \subset \Big(0, \exp\Big(-\frac{1}{2}\,\Lambda\,[1-4\,\sigma_P^2]\,\sin^2\Big[\frac{\pi}{N+2}\Big]\,t\Big)\Big]\,.$$

This implies

$$\sigma(\mathcal{L}|_H) = \frac{1}{t}\,\log\sigma(e^{t\,\mathcal{L}}) \subset \Big(-\infty, -\frac{1}{2}\,\Lambda\,[1-4\,\sigma_P^2]\,\sin^2\Big[\frac{\pi}{N+2}\Big]\Big]\,,$$

which finishes the proof. □

Remark 2.13. *From the proof of Theorem 2.12 it is clear that the abstract result Lemma 2.10 shows that an estimate on the exponential rate of weak convergence of $\mathsf{X}(t)$ in Vaserstein-1 distance automatically yields an estimate on the spectral gap of $\mathcal{L}$ on L^2_π, provided that the stationary distribution π is reversible. And since convergence in Vaserstein-1 distance can be controlled by two different approaches (recall the Kantorovich-Rubinstein duality theorem) we expect this general result to be also useful in other settings to prove estimates on L^2 spectral gaps.*

I rewrote this completely.

Remark 2.14. *All results of this section are essentially consequences of Proposition 2.6 and Lemma 2.10. And since the statement of Proposition 2.6 is readily rephrased for the embedded discrete time Markov chain with transition operator*

$$\mathcal{P}A(x) = A(x) + \frac{1}{N-1}\,\frac{1}{\Lambda}\,\mathcal{L}A(x) = \sum_{i=1}^{N-1}\frac{1}{N-1}\int P(d\alpha)\,A(T_{i,\alpha}x)$$

the results of this section all carry over (essentially verbatim) to the discrete time setting. One only has to multiply the rate of convergence (and hence the spectral gap) by $\frac{1}{N-1}\frac{1}{\Lambda}$ *in the results for continuous time to obtain the corresponding results for the discrete time setting.*

3. Spectral gap in $L^2_{\pi_{\epsilon,N}}$ for the general case

Now we consider the general situation where the continuous-time Markov process $\mathsf{X}(t)$ is generated by the infinitesimal generator $\mathcal{L}$ given in (1). Suppose that $\pi_{\epsilon,N}$ is a reversible measure for $\mathcal{L}$. Then the associated Dirichlet form

$$\mathcal{D}_{\epsilon,N}(A) = \int \pi_{\epsilon,N}(dx)\, A(x)\, [-\mathcal{L}A](x) \tag{9a}$$

is defined for all $A \in L^2_{\pi_{\epsilon,N}}$, and has the representation

$$\mathcal{D}_{\epsilon,N}(A) = \frac{1}{2}\sum_{i=1}^{N-1}\int \pi_{\epsilon,N}(dx)\,\Lambda(x_i,x_{i+1})\int P(x_i,x_{i+1},d\alpha)\,[A(T_{i,\alpha}x)-A(x)]^2\ . \tag{9b}$$

The basic idea to prove convergence rates for $\mathsf{X}(t)$ is to compare the spectral gap of its generator $\mathcal{L}$ to a suitably chosen reference process of the type (3) considered in Section 2. In order to distinguish these two generators we use a superscript $\star$

$$\mathcal{L}^\star A(x) = \Lambda^\star \sum_{i=1}^{N-1}\int P^\star(d\alpha)\,[A(T_{i,\alpha}x)-A(x)]$$

$$\mathcal{D}^\star_{\epsilon,N}(A) = \frac{1}{2}\int \pi^\star_{\epsilon,N}(dx)\sum_{i=1}^{N-1}\Lambda^\star\int P^\star(d\alpha)\,[A(T_{i,\alpha}x)-A(x)]^2$$

to denote the invariant measure, the generator and the corresponding Dirichlet form of the reference process.

Theorem 3.1 (Spectral gap for $\mathcal{L}$)**.** *Fix $\epsilon>0$ and N, and let $\pi_{\epsilon,N}$ be a reversible stationary distribution of $\mathcal{L}$ on $\mathcal{S}_{\epsilon,N}$. Suppose that there exist a constant $\Lambda^\star>0$ and a probability measure $P^\star$ on $[0,1]$ with mean $\int P^\star(d\alpha)\,\alpha = \frac{1}{2}$ and variance $\sigma^2_{P^\star} < \frac{1}{4}$ such that the following are satisfied:*

I rewrote the conditions.

- *(i) The rate function Λ satisfies $\Lambda(x_i,x_{i+1}) \geq \Lambda^\star$ for $\pi_{\epsilon,N}$–almost all $x\in\mathcal{S}_{\epsilon,N}$, and all $1\leq i\leq N-1$.*
- *(ii) There exists a constant $\beta>0$ such that P satisfies the minorization condition $P(x_i,x_{i+1},.)\geq \beta\,P^\star(.)$ for $\pi_{\epsilon,N}$–almost all $x\in\mathcal{S}_{\epsilon,N}$, and all $1\leq i\leq N-1$.*
- *(iii) The unique (recall Theorem 2.9) stationary distribution $\pi^\star_{\epsilon,N}$ of $\mathcal{L}^\star$ on $\mathcal{S}_{\epsilon,N}$ (corresponding to $\Lambda^\star$ and $P^\star$) is reversible.*
- *(iv) The measures $\pi_{\epsilon,N}$ and $\pi^\star_{\epsilon,N}$ are uniformly equivalent, i.e. there exist two constants $0 < C^-_\epsilon \leq C^+_\epsilon < \infty$ such that their Radon-Nikodym derivative satisfies $C^-_\epsilon \leq \frac{\pi_{\epsilon,N}(dx)}{\pi^\star_{\epsilon,N}(dx)} \leq C^+_\epsilon$ for all N.*

I dropped the N-dependence to avoid confusion about the size of the gap, and what exactly "uniformly equivalent" means. This also cleans up the notation; Aug 15.

Then the spectrum of $\mathcal{L}$ in $L^2_{\pi_{\epsilon,N}}$ satisfies

$$\sigma(\mathcal{L}) \subset \Big(-\infty, -\beta\,\frac{C^-_\epsilon}{C^+_\epsilon}\,\Lambda^\star\,\frac{1}{2}\,[1-4\,\sigma^2_{P^\star}]\,\sin^2\Big[\frac{\pi}{N+2}\Big]\Big] \cup \{0\}\ ,$$

where 0 is a simple eigenvalue.

This paragraph is new.
I changed the wording a bit and gave the remark a number. Was that actually intended by you? Aug 15.

Remark 3.2. *Later we will see that - apart from condition* (i) *- the conditions of Theorem 3.1 are fulfilled in a wide range of models of mechanical origin, interesting to us. Indeed, in Theorem 4.3 we will prove a characterization of reversible measures of a particular type. Among others, it will provide the existence of reversible stationary measures for a large class rate functions* Λ *and transition kernels* P*. This result, in particular, addresses conditions* (iii) *and* (iv) *in the above Theorem 3.1 in quite satisfactory generality. Also, in Section 5, we show that* (ii) *is satisfied, for instance, in the Gaspard-Gilbert model with three-dimensional balls. Finally,* (i) *is the consequence of our method. Nevertheless establishing hydrodynamical limit transition is a great challenge even under our conditions and for doing it our theorems serve as an excellent background. Finally we note that the applicability of the statement in its present form seems to be restricted to models where* $\pi_{\epsilon,N} = \pi^\star_{\epsilon,N}$ *therefore the weakening of condition* (iv) *would also be desirable.*

Proof. Since we assume reversibility the generator is self-adjoint, and hence we have the following variational characterization

$$\gamma = \inf\left\{\frac{\mathcal{D}_{\epsilon,N}(A)}{\mathrm{Var}_{\epsilon,N}(A)} : A \in L^2_{\pi_{\epsilon,N}},\ \mathrm{Var}_{\epsilon,N}(A) \neq 0\right\}$$

of the spectral gap γ of $\mathcal{L}$ acting on $L^2_{\pi_{\epsilon,N}}$, where $\mathrm{Var}_{\epsilon,N}(A)$ denotes the variance of A with respect to $\pi_{\epsilon,N}$.

By assumption we can compare the measures $\pi_{\epsilon,N}$, $\pi^\star_{\epsilon,N}$, and P, $P^\star$, so that for the Dirichlet form, recall (9), we obtain the estimate

$$\begin{aligned}\mathcal{D}_{\epsilon,N}(A) &= \frac{1}{2}\sum_{i=1}^{N-1}\int \pi_{\epsilon,N}(dx)\,\Lambda(x_i,x_{i+1})\int P(x_i,x_{i+1},d\alpha)\,[A(T_{i,\alpha}x)-A(x)]^2\\ &\geq \beta\, C_\epsilon^-\,\frac{1}{2}\sum_{i=1}^{N-1}\int \pi^\star_{\epsilon,N}(dx)\,\Lambda^\star\int P^\star(d\alpha)\,[A(T_{i,\alpha}x)-A(x)]^2\end{aligned}$$

which is nothing else but $\mathcal{D}_{\epsilon,N}(A) \geq \beta\, C_\epsilon^-\, \mathcal{D}^\star_{\epsilon,N}(A)$ for all $A \in L^2_{\pi_{\epsilon,N}}$.

Furthermore, the variational characterization of the variance yields the estimate

$$\begin{aligned}\mathrm{Var}_{\epsilon,N}(A) &= \inf_{c\in\mathbb{R}}\int \pi_{\epsilon,N}(dx)\,[A(x)-c]^2 = \inf_{c\in\mathbb{R}}\int \pi^\star_{\epsilon,N}(dx)\,\frac{\pi_{\epsilon,N}(dx)}{\pi^\star_{\epsilon,N}(dx)}\,[A(x)-c]^2\\ &\leq C_\epsilon^+\inf_{c\in\mathbb{R}}\int \pi^\star_{\epsilon,N}(dx)\,[A(x)-c]^2 = C_\epsilon^+\ \mathrm{Var}^\star_{\epsilon,N}(A)\end{aligned}$$

for all $A \in L^2_{\pi_{\epsilon,N}}$.

Combining both of the above estimates shows

$$\frac{\mathcal{D}_{\epsilon,N}(A)}{\mathrm{Var}_{\epsilon,N}(A)} \geq \beta\,\frac{C_\epsilon^-}{C_\epsilon^+}\,\frac{\mathcal{D}^\star_{\epsilon,N}(A)}{\mathrm{Var}^\star_{\epsilon,N}(A)}$$

for any $A \in L^2_{\pi_{\epsilon,N}}$ with $\mathrm{Var}_{\epsilon,N}(A) \neq 0$. In other words, the spectral gap γ of $\mathcal{L}$ admits the estimate

$$\gamma \geq \beta\,\frac{C_\epsilon^-}{C_\epsilon^+}\,\inf\left\{\frac{\mathcal{D}^\star_{\epsilon,N}(A)}{\mathrm{Var}^\star_{\epsilon,N}(A)} : A \in L^2_{\pi_{\epsilon,N}},\ \mathrm{Var}_{\epsilon,N}(A) \neq 0\right\}.$$

Finally, note that the assumed bounds $C_\epsilon^- \leq \frac{\pi_{\epsilon,N}(dx)}{\pi^\star_{\epsilon,N}(dx)} \leq C_\epsilon^+$ imply that $L^2_{\pi_{\epsilon,N}} = L^2_{\pi^\star_{\epsilon,N}}$ so that the above estimate can be rewritten as

$$\gamma \geq \beta \, \frac{C_\epsilon^-}{C_\epsilon^+} \, \gamma^\star$$

where $\gamma^\star$ denotes the spectral gap of $\mathcal{L}^\star$ in $L^2_{\pi^\star_{\epsilon,N}}$.

Now recall that by Theorem 2.12

$$\gamma^\star \geq \frac{1}{2} \, \Lambda^\star \, [1 - 4\,\sigma^2_{P^\star}] \, \sin^2 \left[\frac{\pi}{N+2}\right]$$

which in turn shows for the spectral gap γ of $\mathcal{L}$

$$\gamma \geq \beta \, \frac{C_\epsilon^-}{C_\epsilon^+} \, \Lambda^\star \, \frac{1}{2} \, [1 - 4\,\sigma^2_{P^\star}] \, \sin^2 \left[\frac{\pi}{N+2}\right] \, ,$$

which finishes the proof. □

4. Classification of reversible product measures

Doma's comment added on Aug 15.

In this section we will characterize reversible product measures of $\mathsf{X}(t)$. It is worth recalling at this point that for any N fixed the sets $\mathcal{S}_{\epsilon,N} \subset \mathbb{R}_+^N$ are invariant for the process for any choice of $\epsilon > 0$. And since these are simplexes there are no (non-trivial) product measures $\mu(dx) = \nu(dx_1) \cdots \nu(dx_N)$ supported by a single $\mathcal{S}_{\epsilon,N}$. However, conditioning an invariant product measure on all of $\mathbb{R}_+^N$ to any $\mathcal{S}_{\epsilon,N}$ yields an invariant measure on $\mathcal{S}_{\epsilon,N}$. Therefore, we will consider product measures on all of $\mathbb{R}_+^N$ (canonical measures) instead on the ergodic components $\mathcal{S}_{\epsilon,N}$ (micro-canonical measures). And since our main convergence result Theorem 3.1 is for reversible invariant measures, we consider here only reversible product measures.

The first step in classifying all of them is provided by Lemma 4.1, which says that it suffices to consider $N = 2$.

Lemma 4.1 (Reversible product measures and system size)**.** *Let ν be a probability measure on $\mathbb{R}_+$. Then the product (probability) measure $\mu(dx) = \nu(dx_1) \cdots \nu(dx_N)$ on $\mathbb{R}_+^N$ is reversible for $\mathsf{X}(t)$ (with generator* (1)*) for some N if and only if it is reversible for $N = 2$.*

I added the explanation of the notation $[\mathcal{L}A(x,.)]$. **OK**?

Proof. Let $A\colon \mathbb{R}_+^N \times \mathbb{R}_+^N \to \mathbb{R}$ be bounded. To shorten the notation we use $[\mathcal{L}A(x,.)]$ to denote the function obtained by the action of the generator $\mathcal{L}$ on second variable of the function $A(x,x')$, while treating the first variable as a parameter. Further we use $[\mathcal{L}A(x,.)](x')$ to denote the evaluation of the function $[\mathcal{L}A(x,.)]$ at the point x'. Correspondingly, in $[\mathcal{L}A(.,x')]$ the second variable is treated as a parameter.

By definition (1) of the generator $\mathcal{L}$ we have

$$\int_{\mathbb{R}_+^N} \mu(dx)\,[\mathcal{L}A(.,x)](x) = \sum_{i=1}^{N-1} \int_{\mathbb{R}_+^N} \nu(dx_1)\cdots\nu(dx_N)\,\Lambda(x_i,x_{i+1})\cdot \\ \cdot \int P(x_i,x_{i+1},d\alpha)\,[A(T_{i,\alpha}x,x) - A(x,x)] \\ = \sum_{i=1}^{N-1} \int_{\mathbb{R}_+^2} \nu(dx_i)\,\nu(dx_{i+1})\,\Lambda(x_i,x_{i+1}) \int P(x_i,x_{i+1},d\alpha)\cdot \\ \cdot\Big[A_{i,i+1}(\alpha\,[x_i+x_{i+1}],(1-\alpha)\,[x_i+x_{i+1}],x_i,x_{i+1}) \\ - A_{i,i+1}(x_i,x_{i+1},x_i,x_{i+1})\Big]\ ,$$

where we used the short hand notation

Made the formula look nicer; Aug 15.

$$A_{i,i+1}(x_i,x_{i+1},x_i',x_{i+1}') = \int_{\mathbb{R}_+^{N-2}} \nu(dx_1)\cdots\nu(dx_{i-1})\,\nu(dx_{i+2})\cdots\nu(dx_N)\,A(x,z_i)\ , \\ z_i \equiv (x_1,\ldots,x_{i-1},x_i',x_{i+1}',x_{i+2},\ldots,x_N)\ .$$

Recall that reversibility means $\int_{\mathbb{R}_+^N} \mu(dx)\,[\mathcal{L}A(.,x)](x) = \int_{\mathbb{R}_+^N} \mu(dx)\,[\mathcal{L}A(x,.)](x)$, so that reversibility holds if and only if

$$\sum_{i=1}^{N-1} \int_{\mathbb{R}_+^2} \nu(dx_i)\,\nu(dx_{i+1})\,\Lambda(x_i,x_{i+1}) \int P(x_i,x_{i+1},d\alpha)\cdot \\ \cdot A_{i,i+1}(\alpha\,[x_i+x_{i+1}],(1-\alpha)\,[x_i+x_{i+1}],x_i,x_{i+1}) \\ = \sum_{i=1}^{N-1} \int_{\mathbb{R}_+^2} \nu(dx_i)\,\nu(dx_{i+1})\,\Lambda(x_i,x_{i+1}) \int P(x_i,x_{i+1},d\alpha) \\ \cdot A_{i,i+1}(x_i,x_{i+1},\alpha\,[x_i+x_{i+1}],(1-\alpha)\,[x_i+x_{i+1}])$$

for any bounded $A\colon \mathbb{R}_+^N \times \mathbb{R}_+^N \to \mathbb{R}$.

In the particular case where $A(x,x') = \phi(x_1,x_1')$ for some bounded $\phi\colon \mathbb{R}_+\times\mathbb{R}_+ \to \mathbb{R}$

$$A_{1,2}(x_1,x_2,x_1',x_2') = \phi(x_1,x_1') \\ A_{i,i+1}(x_i,x_{i+1},x_i',x_{i+1}') = \int_{\mathbb{R}_+} \nu(dx_1)\,\phi(x_1,x_1) \equiv \text{const}$$

for all $i = 2,\ldots,N-1$. Hence reversibility requires

$$\int_{\mathbb{R}_+^2} \nu(dx_1)\,\nu(dx_2)\,\Lambda(x_1,x_2) \int P(x_1,x_2,d\alpha)\,\phi(\alpha\,[x_1+x_2],x_1) \\ = \int_{\mathbb{R}_+^2} \nu(dx_1)\,\nu(dx_2)\,\Lambda(x_1,x_2) \int P(x_1,x_2,d\alpha)\,\phi(x_1,\alpha\,[x_1+x_2])\ .$$

Consider now $A(x,x') = \psi(x_1,x_2,x_1',x_2')$ for some bounded $\psi\colon \mathbb{R}_+^2 \times \mathbb{R}_+^2 \to \mathbb{R}$. Then

$$\begin{aligned} A_{1,2}(x_1,x_2,x_1',x_2') &= \psi(x_1,x_2,x_1',x_2') \\ A_{2,3}(x_2,x_3,x_2',x_3') &= \int_{\mathbb{R}_+} \nu(dx_1)\,\psi(x_1,x_2,x_1,x_2') \equiv \hat\psi(x_2,x_2') \\ A_{i,i+1}(x_i,x_{i+1},x_i',x_{i+1}') &= \int_{\mathbb{R}_+^2} \nu(dx_1)\,\nu(dx_2)\,\psi(x_1,x_2,x_1,x_2) \equiv \text{const} \end{aligned}$$

for all $i = 3,\dots,N-1$. Combining this with the previous special case (applied to $\phi = \hat\psi$) shows that reversibility requires

$$\begin{aligned} &\int_{\mathbb{R}_+^2} \nu(dx_1)\,\nu(dx_2)\,\Lambda(x_1,x_2) \int P(x_1,x_2,d\alpha)\cdot \\ &\qquad\cdot \psi(\alpha\,[x_1+x_2],(1-\alpha)\,[x_1+x_2],x_1,x_2) \\ &= \int_{\mathbb{R}_+^2} \nu(dx_1)\,\nu(dx_2)\,\Lambda(x_1,x_2) \int P(x_1,x_2,d\alpha)\cdot \\ &\qquad\cdot \psi(x_1,x_2,\alpha\,[x_1+x_2],(1-\alpha)\,[x_1+x_2]) \end{aligned}$$

for any bounded test function $\psi\colon \mathbb{R}_+^2 \times \mathbb{R}_+^2 \to \mathbb{R}$. And since this is also sufficient for reversibility, it follows that reversibility of the product measure holds if and only if the above equality holds for all ψ.

Finally, observe that this last expression is precisely the reversibility condition for $N = 2$, which finishes the proof. □

The final expression in the above proof actually shows that reversibility of the product measure is equivalent to a slightly stronger statement than the one stated in Lemma 4.1. Namely, because both integrands agree at $(0,0)$ the reversibility of the product measure is equivalent to

$$\begin{aligned} &\int_{\mathbb{R}_+^2\setminus\{(0,0)\}} \nu(dx_1)\,\nu(dx_2)\,\Lambda(x_1,x_2) \int P(x_1,x_2,d\alpha)\cdot \\ &\qquad\cdot \psi(\alpha\,[x_1+x_2],(1-\alpha)\,[x_1+x_2],x_1,x_2) \\ &= \int_{\mathbb{R}_+^2\setminus\{(0,0)\}} \nu(dx_1)\,\nu(dx_2)\,\Lambda(x_1,x_2) \int P(x_1,x_2,d\alpha)\cdot \\ &\qquad\cdot \psi(x_1,x_2,\alpha\,[x_1+x_2],(1-\alpha)\,[x_1+x_2])\ . \end{aligned} \tag{10}$$

This simplification is relevant, because so far we have not ruled out yet the possibility of ν having an atom at 0.

For the further analysis we will need to assume that the rate function Λ and the transition kernel P are of the form

$$\begin{aligned} \Lambda(x_i,x_{i+1}) &= \Lambda_s(x_i+x_{i+1})\,\Lambda_r\Big(\frac{x_i}{x_i+x_{i+1}}\Big) \\ P(x_i,x_{i+1},d\alpha) &= P\Big(\frac{x_i}{x_i+x_{i+1}},d\alpha\Big)\ . \end{aligned} \tag{11}$$

Alternatively, we could say that the ratio is just a formal way to say the function is homogeneous of degree zero.

Explain the subscripts; Aug 15.

Here the subscripts s and r stand for "sum" and "ratio", respectively. Note that $\frac{x_i}{x_i+x_{i+1}}$ makes sense everywhere on $\mathbb{R}_+^2 \setminus \{(0,0)\}$, and by the above this set is all that we need to consider. In Section 5 below we will see that the representation (11) naturally occurs in models originating from mechanical systems.

We have already shown that in order to classify reversible product measures for arbitrary N it is enough to study the case $N = 2$. This, however, is still not a completely straightforward problem, since the answer might depend on the rate functions Λ_s and Λ_r. The next Corollary 4.2 simplifies this issue.

Added Kostya's modification here; Aug 15.

Corollary 4.2 (Reversible product measures and rate functions)**.** *If $\Lambda_s(\eta) > 0$ for all $0 < \eta < \infty$, then the process has a reversible stationary product measure μ (as in Lemma 4.1) if and only if*

$$\int_{\mathbb{R}_+^2\setminus\{(0,0)\}} \nu(dx_1)\,\nu(dx_2)\,\Lambda_r\Big(\frac{x_1}{x_1+x_2}\Big)\int P\Big(\frac{x_1}{x_1+x_2},d\alpha\Big)\cdot \cdot\,\eta\Big(x_1+x_2,\alpha,\frac{x_1}{x_1+x_2}\Big)$$
$$= \int_{\mathbb{R}_+^2\setminus\{(0,0)\}} \nu(dx_1)\,\nu(dx_2)\,\Lambda_r\Big(\frac{x_1}{x_1+x_2}\Big)\int P\Big(\frac{x_1}{x_1+x_2},d\alpha\Big)\cdot \cdot\,\eta\Big(x_1+x_2,\frac{x_1}{x_1+x_2},\alpha\Big)$$

holds for all bounded $\eta\colon \mathbb{R}_+\setminus\{0\}\times[0,1]^2\to\mathbb{R}$.

Proof. By (10) reversibility of the product measure is equivalent to

$$\int_{\mathbb{R}_+^2\setminus\{(0,0)\}} \nu(dx_1)\,\nu(dx_2)\,\Lambda_s(x_1+x_2)\,\Lambda_r\Big(\frac{x_1}{x_1+x_2}\Big)\int P\Big(\frac{x_1}{x_1+x_2},d\alpha\Big)\cdot \cdot\,\psi(\alpha\,[x_1+x_2],(1-\alpha)\,[x_1+x_2],x_1,x_2)$$
$$= \int_{\mathbb{R}_+^2\setminus\{(0,0)\}} \nu(dx_1)\,\nu(dx_2)\,\Lambda_s(x_1+x_2)\,\Lambda_r\Big(\frac{x_1}{x_1+x_2}\Big)\int P\Big(\frac{x_1}{x_1+x_2},d\alpha\Big)\cdot \cdot\,\psi(x_1,x_2,\alpha\,[x_1+x_2],(1-\alpha)\,[x_1+x_2])$$

for any (non-negative) test function $\psi\colon \mathbb{R}^2\setminus\{(0,0)\}\times\mathbb{R}_+^2\to\mathbb{R}$. On $\mathbb{R}_+^2\setminus\{(0,0)\}$ the change of coordinates $(x_1,x_2)\mapsto(x_1+x_2,\frac{x_1}{x_1+x_2})$ is one-to-one, hence any such function ψ may be recast as

$$\psi(x_1,x_2,x_1',x_2') \equiv \eta\Big(x_1+x_2,\frac{x_1}{x_1+x_2},x_1'+x_2',\frac{x_1'}{x_1'+x_2'}\Big)$$

for some function $\eta\colon(\mathbb{R}_+\times[0,1])^2\to\mathbb{R}$. Therefore reversibility holds if and only if

$$\int_{\mathbb{R}_+^2\setminus\{(0,0)\}} \nu(dx_1)\,\nu(dx_2)\,\Lambda_s(x_1+x_2)\,\Lambda_r\Big(\frac{x_1}{x_1+x_2}\Big)\int P\Big(\frac{x_1}{x_1+x_2},d\alpha\Big)\cdot \cdot\,\eta\Big(x_1+x_2,\alpha,\frac{x_1}{x_1+x_2}\Big)$$
$$= \int_{\mathbb{R}_+^2\setminus\{(0,0)\}} \nu(dx_1)\,\nu(dx_2)\,\Lambda_s(x_1+x_2)\,\Lambda_r\Big(\frac{x_1}{x_1+x_2}\Big)\int P\Big(\frac{x_1}{x_1+x_2},d\alpha\Big)\cdot \cdot\,\eta\Big(x_1+x_2,\frac{x_1}{x_1+x_2},\alpha\Big)$$

holds for all $\eta\colon\mathbb{R}_+\setminus\{0\}\times[0,1]^2\to\mathbb{R}$.

And since $x_1+x_2>0$ our assumption on Λ_s implies that Λ_s is strictly positive, and hence may as well be combined with η, because η is arbitrary. This finishes the proof. □

With Lemma 4.1, and Corollary 4.2 we are now in a position to classify all reversible product measures, which is the content of the following Theorem 4.3. This classification relies on a well-known fact [Luk55] about Gamma distributions. Namely, suppose that X_1 and X_2 are two non-constant, independent, positive random variables. Then $\mathsf{X}_1 + \mathsf{X}_2$ and $\frac{\mathsf{X}_1}{\mathsf{X}_1+\mathsf{X}_2}$ are independent if and only if X_1 and X_2 are independent, identically Gamma-distributed random variables.

Added this to explained the notation of the delta measure; Aug 15.

In the theorem below we use the following notation: For $\epsilon > 0$ we denote by $\delta(\epsilon, d\alpha)$ the Dirac measure concentrated at ϵ.

We switched the order of the two cases; Aug 15.

Theorem 4.3 (Reversible product measures)**.** *Suppose that the Markov chain on* $[0,1]$ *with transition kernel* $P(\beta, d\alpha)$ *has a unique invariant distribution, say* $p(.)$. *Let* N *be arbitrary, and suppose further that* Λ_s *is such that* $\Lambda_s(\sigma) > 0$ *for all* $\sigma > 0$, *and* $\Lambda_r(\beta) > 0$ *for all* $0 < \beta < 1$. *Then the product measure* $\mu(dx) = \nu(dx_1) \cdots \nu(dx_N)$ *is reversible for* $\mathsf{X}(t)$ *if and only if* p *is a reversible measure for the Markov chain generated by* P, *and either of the following two holds:*

(1) There exists $\epsilon > 0$ *and* $d > 0$ *such that*

$$\nu(dx_1) = \frac{dx_1}{\epsilon} \left[\frac{x_1}{\epsilon}\right]^{\frac{d}{2}-1} \frac{e^{-\frac{x_1}{\epsilon}}}{\Gamma(\frac{d}{2})}$$

$$p(d\beta) = d\beta \, [\beta\,(1-\beta)]^{\frac{d}{2}-1} \, \frac{\Gamma(d)}{\Gamma(\frac{d}{2})^2} \, \Lambda_r(\beta) \, \frac{1}{Z}$$

where Z *is the normalizing constant.*

(2) There exists $\epsilon > 0$ *such that* $\nu(dx_1) = \delta(\epsilon, dx_1)$, $p(d\alpha) = \delta(\frac{1}{2}, d\alpha)$, *and* $P(\frac{1}{2}, d\alpha) = \delta(\frac{1}{2}, d\alpha)$.

Changed s to σ in all the formulas below to avoid notation like $\Lambda_s(s)$; Aug 15.

Proof. From Lemma 4.1 we know that it suffices to consider $N = 2$, and Corollary 4.2 shows - as it is also clear intuitively - that the choice of Λ_s is irrelevant, and that we only need to consider the process on $\mathbb{R}_+^2 \setminus \{(0,0)\}$.

Using the change of variables $\sigma = x_1 + x_2$, $\beta = \frac{x_1}{x_1+x_2}$ on $\mathbb{R}_+^2 \setminus \{(0,0)\}$ we can disintegrate the product measure $\nu(dx_1)\,\nu(dx_2)$ such that for any (bounded) $\eta\colon \mathbb{R}_+^2 \to \mathbb{R}$ we have

$$\int_{\mathbb{R}_+^2 \setminus \{(0,0)\}} \nu(dx_1)\,\nu(dx_2)\,\eta(x_1,x_2) = \int_{\mathbb{R}_+ \setminus \{0\}} \nu_s(d\sigma) \int_{[0,1]} \nu_r(\sigma, d\beta)\,\eta(\beta\,\sigma, (1-\beta)\,\sigma)$$

where $\nu_s(.)$ is the distribution of the sum $x_1 + x_2$ and $\nu_r(\sigma, .)$ is the conditional distribution of the ratio $\frac{x_1}{x_1+x_2}$ given that $x_1 + x_2 = \sigma$.

Using this notation the condition for the reversibility of the product measure of Corollary 4.2 takes on the form

$$\int \nu_s(d\sigma) \int \nu_r(\sigma, d\beta)\,\Lambda_r(\beta) \int P(\beta, d\alpha)\,\eta(\sigma, \alpha, \beta)$$
$$= \int \nu_s(d\sigma) \int \nu_r(\sigma, d\beta)\,\Lambda_r(\beta) \int P(\beta, d\alpha)\,\eta(\sigma, \beta, \alpha)\ .$$

This holds if and only if for ν_s–almost every σ

$$\begin{aligned} &\int \nu_r(\sigma, d\beta)\,\Lambda_r(\beta) \int P(\beta, d\alpha)\,\tilde\eta(\alpha, \beta) \\ &\qquad = \int \nu_r(\sigma, d\beta)\,\Lambda_r(\beta) \int P(\beta, d\alpha)\,\tilde\eta(\beta, \alpha) \end{aligned} \tag{12}$$

for all bounded $\tilde{\eta}\colon [0,1]\times[0,1]\to\mathbb{R}$.

Suppose now that the product measure is reversible. The special choice $\eta(\alpha,\beta) = \psi(\alpha)$ for some $\psi\colon [0,1]\to\mathbb{R}$ thus shows that

$$\int \nu_r(\sigma, d\beta)\,\Lambda_r(\beta)\int P(\beta, d\alpha)\,\psi(\alpha) = \int \nu_r(\sigma, d\beta)\,\Lambda_r(\beta)\,\psi(\beta)$$

for all ψ. In other words, the (not normalized) non-negative measure $\nu_r(\sigma, d\beta)\,\Lambda_r(\beta)$ must be invariant under P. And since by assumption P has a unique invariant distribution, denote it by p, it thus follows that

$$\frac{1}{Z}\,\nu_r(\sigma, d\beta)\,\Lambda_r(\beta) = p(d\beta)\ , \qquad Z = \int \nu_r(\sigma, d\beta)\,\Lambda_r(\beta)$$

for ν_s–almost every σ, where $Z > 0$ by assumption on Λ_r.

In particular, this means that the conditional distribution $\nu_r(\sigma, .)$ of the ratio $\frac{x_1}{x_1+x_2}$ given that $\sigma = x_1 + x_2$ actually is the same for all values of σ. In other words the sum $x_1 + x_2$ and the ratio $\frac{x_1}{x_1+x_2}$ are independent. And since also x_1 and x_2 are independent (by assumption) we conclude [Luk55] that either ν is a point mass, i.e. $\nu(dx_1) = \delta(\epsilon, dx_1)$ for some $\epsilon > 0$, or ν is a Gamma distribution, i.e.

$$\nu(dx_1) = \frac{dx_1}{\epsilon}\left[\frac{x_1}{\epsilon}\right]^{\frac{d}{2}-1}\frac{e^{-\frac{x_1}{\epsilon}}}{\Gamma(\frac{d}{2})} \qquad (0 < x_1 < \infty)$$

for some $\epsilon > 0$ and $d > 0$

In the former case it follows

$$\nu_s(d\sigma) = \delta(2\,\epsilon, d\sigma)\ , \qquad p(d\beta) = \nu_r(d\beta) = \delta(\frac{1}{2}, d\beta)$$

for ν_s, ν_r. Hence the reversibility condition (12) becomes $\int P(\frac{1}{2}, d\alpha)\,\eta(\alpha, \frac{1}{2}) = \int P(\frac{1}{2}, d\alpha)\,\eta(\frac{1}{2}, \alpha)$ for all η, which is equivalent to

$$P\Big(\frac{1}{2}, d\alpha\Big) = \delta\Big(\frac{1}{2}, d\alpha\Big)\ .$$

Similarly, in the latter case

$$\nu_s(d\sigma) = \frac{d\sigma}{\epsilon}\left[\frac{\sigma}{\epsilon}\right]^{d-1}\frac{e^{-\frac{\sigma}{\epsilon}}}{\Gamma(d)}\ , \qquad \nu_r(d\beta) = d\beta\,[\beta\,(1-\beta)]^{\frac{d}{2}-1}\,\frac{\Gamma(d)}{\Gamma(\frac{d}{2})^2}$$

follows for ν_s, ν_r, where we used the well-known properties of Gamma and Beta distributions. The reversibility condition (12) becomes

$$\int_0^1 d\beta\,[\beta\,(1-\beta)]^{\frac{d}{2}-1}\,\frac{\Gamma(d)}{\Gamma(\frac{d}{2})^2}\,\Lambda_r(\beta)\int P(\beta, d\alpha)\,\eta(\alpha,\beta)$$
$$= \int_0^1 d\beta\,[\beta\,(1-\beta)]^{\frac{d}{2}-1}\,\frac{\Gamma(d)}{\Gamma(\frac{d}{2})^2}\,\Lambda_r(\beta)\int P(\beta, d\alpha)\,\eta(\beta,\alpha)$$

for all η, and

$$p(d\beta) = d\beta\,[\beta\,(1-\beta)]^{\frac{d}{2}-1}\,\frac{\Gamma(d)}{\Gamma(\frac{d}{2})^2}\,\Lambda_r(\beta)\,\frac{1}{Z}$$

must be the expression for the unique stationary distribution of P.

This proves that if the product measure is reversible, then ν is either constant, or a Gamma distribution, and the transition kernel must have the claimed stationary distribution.

To finish the proof it remains to consider the converse. Assume either of the two possible distributions for ν and also the corresponding assumption on P. For these special distributions it is well-known (and easily verified) that the sum and the ratio are independent with the distributions as considered above. Hence we see that the reversibility condition (12) is indeed satisfied. □

We finish the discussion of reversible product measures with the following remark. Note that in the statement of Theorem 4.3 there is an assumption on the kernel P that appears in the generator of the process $\mathsf{X}(t)$. By Lemma 4.1 and Corollary 4.2 it suffices to consider the reversibility of the product measure for $N = 2$ and constant rates. Upon restricting this process to any of the invariant sets $\mathcal{S}_{\epsilon,2}$, the embedded discrete time Markov chain is precisely the Markov chain on $[0, 1]$ with transition kernel $P(\beta, d\alpha)$. Therefore, the assumption in Theorem 4.3 on the kernel P is equivalent to saying that for $N = 2$ and constant rates the process $\mathsf{X}(t)$ has a unique stationary distribution on any of the $\mathcal{S}_{\epsilon,2}$.

A sufficient condition for this uniqueness is to assume that P satisfies a uniform minorization condition, i.e. there exists a constant $\gamma > 0$ and a probability measure $P^\star$ on $[0, 1]$ with $\int P^\star(d\alpha)\, \alpha = \frac{1}{2}$ and $\sigma^2_{P^\star} < \frac{1}{4}$ such that $P(\beta, .) \geq \gamma\, P^\star(.)$ for all $\beta \in [0, 1]$. Recall that this is also the type of condition P assumed in Theorem 3.1.

5. Example: the rarely interacting billiard lattice

Here we illustrate the use of Theorem 3.1 and Theorem 4.3 with the billiard lattice model studied in [GG09], which was one of the main motivations for our work presented in this paper. It was argued in [GG09] that in the limit of rare collisions the dynamics of a billiard lattice becomes a Markov jump process. The notation used in [GG09] differs from ours in that we separate the rate of interaction $\Lambda_s\, \Lambda_r$ from the transition probability kernel P, whereas in [GG09] the product $\Lambda_s\, \Lambda_r\, P$ is denoted by W, and the rate function $\Lambda_s\, \Lambda_r$ is denoted by ν. Changing equations (61) and (62) of [GG09] to our notation yields

$$\frac{P(\beta, d\alpha)}{d\alpha} = \frac{3}{2}\, \frac{1 \wedge \sqrt{\frac{\alpha \wedge (1-\alpha)}{\beta \wedge (1-\beta)}}}{\frac{1}{2} + \beta \vee (1-\beta)}\ , \quad \Lambda_r(\beta) = \frac{\sqrt{2\pi}}{6}\, \frac{\frac{1}{2} + \beta \vee (1-\beta)}{\sqrt{\beta \vee (1-\beta)}}\ , \quad \Lambda_s(s) = \sqrt{s}$$

for the transition kernel P and the rate functions Λ_s and Λ_r, respectively. The symbol $\vee$ denotes the maximum and $\wedge$ denotes the minimum.

changed s to σ to avoid things like $\nu_s(ds)$.

Since the underlying mechanical model has a three-dimensional configuration space for each of the constituent particles it follows that

$$d = 3\ , \qquad \nu(dx_1) = \frac{dx_1}{\epsilon} \sqrt{\frac{x_1}{\epsilon}}\, \frac{2\, e^{-\frac{x_1}{\epsilon}}}{\sqrt{\pi}}$$

$$\nu_s(d\sigma) = \frac{d\sigma}{\epsilon} \left[\frac{\sigma}{\epsilon}\right]^2 \frac{e^{-\frac{\sigma}{\epsilon}}}{2}\ , \qquad \nu_r(d\beta) = d\beta\, \sqrt{\beta\,(1-\beta)}\, \frac{8}{\pi}$$

$$p(d\alpha) = d\alpha\, \sqrt{\alpha\,(1-\alpha)}\, \frac{8}{\pi}\, \Lambda_r(\alpha)\, \frac{1}{Z}$$

should be used in Theorem 4.3. In fact, this measure is the (canonical) Gibbs measure for the mechanical model, and thus must also be invariant for the limiting jump process.

Another general property that the jump process inherits from the underlying mechanical model is that the rate function Λ is proportional to the square root of the total energy of the two sites that interact, i.e. $\Lambda_s(\sigma) = \sqrt{\sigma}$ as mentioned above. This cannot be avoided when taking scaling limits of interacting mechanical models, because it corresponds to the kinematic scaling relation between the energy and the velocity (and hence the time scale). However, a rate function without a uniform lower bound leads to serious technical complications at various levels. See, for example, [DL10] for how this issue seriously complicates the rigorous derivation of the weak interaction limit of a related deterministic model.

Should we keep this reference?

Furthermore, such a rate function also complicates the rigorous analysis of the rate of convergence to equilibrium. In fact, in order to apply the results established in this paper we need to have Λ_s bounded from below. Recall that we showed in Lemma 4.1 that the above reversible product measure is also a reversible stationary distribution for the process generated by the infinitesimal generator corresponding to any other function Λ_s (while keeping Λ_r and P unchanged). And since Λ_s represents the kinematic scaling, and not the nature of the energy exchange during an interaction, we will change the model of [GG09] in that we change Λ_s. In fact, our next Lemma is most useful exactly under the setup of the aforementioned work.

Lemma 5.1. *If Λ_s is replaced by any non-negative continuous function, which is bounded away from zero, then the following hold for any N and ϵ.*

(1) The product measure $\mu(dx) = \nu(dx_1)\cdots\nu(dx_N)$ with $\nu(dx_1) = \frac{dx_1}{\epsilon}\sqrt{\frac{x_1}{\epsilon}}\,\frac{2\,e^{-\frac{x_1}{\epsilon}}}{\sqrt{\pi}}$ is the unique reversible product measure for $\mathsf{X}(t)$.

(2) On every $\mathcal{S}_{\epsilon,N}$ there exists a unique stationary distribution $\pi_{\epsilon,N}$. This measure is obtained by conditioning $\mu(dx)$.

(3) The spectrum $\sigma(\mathcal{L})$ of the generator $\mathcal{L}$ acting on $L^2_{\pi_{\epsilon,N}}$ satisfies

$$\sigma(\mathcal{L}) \subset \Big(-\infty, -C\,\sin^2\Big[\frac{\pi}{N+2}\Big]\Big] \cup \{0\}$$

for some constant C, which may depend on the choice of Λ_s.

Proof. The explicit expressions for the transition kernel and the rate functions allow us to show

$$\Lambda_r(\beta)\,\frac{P(\beta,d\alpha)}{d\alpha} = \frac{\sqrt{2\pi}}{4}\,\frac{1\wedge\sqrt{\frac{\alpha\wedge(1-\alpha)}{\beta\wedge(1-\beta)}}}{\sqrt{\beta\vee(1-\beta)}} = \frac{\sqrt{2\pi}}{4}\,\frac{\sqrt{\beta\wedge(1-\beta)}\wedge\sqrt{\alpha\wedge(1-\alpha)}}{\sqrt{\beta\,(1-\beta)}}$$

for all $\alpha,\beta\in[0,1]$. Hence $p(d\beta)\,P(\beta,d\alpha) = p(d\alpha)\,P(\alpha,d\beta)$, i.e.

$$\int p(d\beta)\int P(\beta,d\alpha)\,\psi(\alpha,\beta) = \int p(d\beta)\int P(\beta,d\alpha)\,\psi(\beta,\alpha)$$

holds for all $\psi\colon[0,1]^2\to\mathbb{R}$.

Furthermore, the estimate

$$\begin{aligned}\frac{P(\beta,d\alpha)}{\nu_r(d\alpha)} &= \frac{3\pi}{16}\,\frac{1\wedge\sqrt{\frac{\alpha\wedge(1-\alpha)}{\beta\wedge(1-\beta)}}}{\frac{1}{2}+\beta\vee(1-\beta)}\,\frac{1}{\sqrt{\alpha\,(1-\alpha)}}\\ &= \frac{3\pi}{16}\Big[\frac{\frac{1}{\sqrt{\alpha\,(1-\alpha)}}}{\frac{1}{2}+\beta\vee(1-\beta)}\wedge\frac{\frac{1}{\sqrt{\alpha\vee(1-\alpha)}}}{[\frac{1}{2}+\beta\vee(1-\beta)]\,\sqrt{\beta\wedge(1-\beta)}}\Big]\\ &\geq \frac{3\pi}{16}\Big[\frac{4}{3}\wedge\sqrt{2}\Big] = \frac{\pi}{4}\end{aligned}$$

provides the minorization condition $P(\beta, d\alpha) \geq \frac{\pi}{4}\, \nu_r(d\alpha)$. In particular, this implies that the Markov chain on $[0,1]$ with transition kernel P has a unique invariant measure.

Therefore, it follows then from Theorem 4.3 that $\mu(dx)$ is a reversible product measure, and must be unique.

Observe that by Theorem 4.3 the infinitesimal generator corresponding to $P^\star(d\alpha) = \nu_r(d\alpha)$ and a constant rate function also has $\pi_{\epsilon,N}$ as a stationary reversible distribution. Combining this with the above minorization condition for P and $\frac{\sqrt{\pi}}{3} \leq \Lambda_r(\beta) \leq \frac{\sqrt{2\pi}}{4}$ we see that under the assumption that Λ_s is bounded from below all assumptions of Theorem 3.1 are satisfied. □

This paragraph is new

The important result of Lemma 5.1 is that it provides an interesting model that fits the conditions of Theorem 3.1. We would like to point out that previous to [GG09] the analogous two-dimensional billiard network was studied in [GG08b]. However, in this case the uniform mixing condition (ii) of Theorem 3.1 fails to hold, which is why we restricted our attention in the above to the three-dimensional setting.

6. Conclusion

The authors of [GG09] suggested a two step strategy for deriving the heat equation from a mechanical model. Motivated by that we have introduced in this work a class of stochastic models with the aim to implement the second step of their strategy: the derivation of the heat equation from a mesoscopic stochastic model.

At present it is widely understood that a necessary ingredient to rigorously establish the hydrodynamical limit is a sharp bound on the dependence of the spectral gap of the generator on the system size. Such a bound is one of the main results of the present paper.

Besides the importance of this bound for the hydrodynamic limit, an additional value of our result is that for systems with continuous state space such bounds are hard and scarce, e.g. [Jan01, CCL03]. As in those works, our method requires to assume that the rates are bounded away from zero.

In more detail: according to our main result the spectral gap of the infinitesimal generator of the process scales as $\mathcal{O}(N^{-2})$ in terms of the systems size N. This is precisely the kind of scaling which allows for a diffusive scaling limit, and hence the study of the hydrodynamic limit. However, we do not study the hydrodynamic limit in this paper, because it requires different ideas and techniques, and results on the spectral gap are of interest in their own right.

To keep the model as close to the mechanical ones as possible (cf. Section 5 and [GG09]) it is desirable to remove the assumption on existence of a uniform lower bound of the rate function. Numerical simulations suggest that the $\mathcal{O}(N^{-2})$ scaling of the spectral gap remains true also for rate functions that can approach zero. In particular for the square root of the total energy of the interacting pair, which is the rate function that appears in mechanical models due to the kinematic scaling of the velocity with the energy. However, we do not have a rigorous proof of such a statement available at present.

References

[CCL03] E. A. Carlen, M. C. Carvalho, and M. Loss. Determination of the spectral gap for Kac's master equation and related stochastic evolution. Acta Math., 191(1):1–54, 2003.

[CGL08] Eric A. Carlen, Jeffrey S. Geronimo, and Michael Loss. Determination of the spectral gap in the Kac model for physical momentum and energy-conserving collisions. SIAM J. Math. Anal., 40(1):327–364, 2008.

[Con90] John B. Conway. A course in functional analysis, volume 96 of Graduate Texts in Mathematics. Springer-Verlag, New York, second edition, 1990.

[DL10] D. Dolgopyat and C. Liverani. Energy transfer in a fast-slow Hamiltonian system. ArXiv e-prints, October 2010.

[EY06] J.-P. Eckmann and L.-S. Young. Nonequilibrium energy profiles for a class of 1-D models. Comm. Math. Phys., 262(1):237–267, 2006.

[Fri87a] J. Fritz. On the hydrodynamic limit of a one-dimensional Ginzburg-Landau lattice model. The a priori bounds. J. Statist. Phys., 47(3-4):551–572, 1987.

[Fri87b] J. Fritz. On the hydrodynamic limit of a scalar Ginzburg-Landau lattice model: the resolvent approach. In Hydrodynamic behavior and interacting particle systems (Minneapolis, Minn., 1986), volume 9 of IMA Vol. Math. Appl., pages 75–97. Springer, New York, 1987.

[GG08a] P. Gaspard and T. Gilbert. Heat conduction and fourier's law in a class of many particle dispersing billiards. New Journal of Physics, 10(10):30, 2008.

[GG08b] P. Gaspard and T. Gilbert. On the derivation of fouriers law in stochastic energy exchange systems. Journal of Statistical Mechanics: Theory and Experiment, 2008(11):20, 2008.

[GG09] P. Gaspard and T. Gilbert. Heat transport in stochastic energy exchange models of locally confined hard spheres. Journal of Statistical Mechanics: Theory and Experiment, 2009(08):24, 2009.

[GPV88] M. Z. Guo, G. C. Papanicolaou, and S. R. S. Varadhan. Nonlinear diffusion limit for a system with nearest neighbor interactions. Comm. Math. Phys., 118(1):31–59, 1988.

[Jan01] Elise Janvresse. Spectral gap for Kac's model of Boltzmann equation. Ann. Probab., 29(1):288–304, 2001.

[Kac56] M. Kac. Foundations of kinetic theory. In Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III, pages 171–197, Berkeley and Los Angeles, 1956. University of California Press.

[KL99] Claude Kipnis and Claudio Landim. Scaling limits of interacting particle systems, volume 320 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1999.

[KMP82] C. Kipnis, C. Marchioro, and E. Presutti. Heat flow in an exactly solvable model. J. Statist. Phys., 27(1):65–74, 1982.

[Luk55] Eugene Lukacs. A characterization of the gamma distribution. Ann. Math. Statist., 26:319–324, 1955.

[Rud87] Walter Rudin. Real and complex analysis. McGraw-Hill Book Co., New York, third edition, 1987.

[Yau96] Horng-Tzer Yau. Logarithmic Sobolev inequality for lattice gases with mixing conditions. Comm. Math. Phys., 181(2):367–408, 1996.

[Yau97] Horng-Tzer Yau. Logarithmic Sobolev inequality for generalized simple exclusion processes. Probab. Theory Related Fields, 109(4):507–538, 1997.

Department of Mathematics, University of Toronto, Toronto, Canada
E-mail address: `alexander.grigo@utoronto.ca`

Department of Mathematics, University of Toronto, Toronto, Canada
E-mail address: `khanin@math.utoronto.ca`

Budapest University of Technology and Economics, Mathematical Institute, Egry J. u. 1, 1111 Budapest, Hungary
E-mail address: `szasz@math.bme.hu`