A THERMODYNAMIC FORMALISM FOR CONTINUOUS TIME MARKOV CHAINS WITH VALUES ON THE BERNOULLI SPACE: ENTROPY, PRESSURE AND LARGE DEVIATIONS

ARTUR LOPES, ADRIANA NEUMANN, AND PHILIPPE THIEULLEN

ABSTRACT. Through this paper we analyze the ergodic properties of continuous time Markov chains with values on the one-dimensional spin lattice $\{1,\dots,d\}^{\mathbb{N}}$ (also known as the Bernoulli space). Initially, we consider as the infinitesimal generator the operator $L = \mathscr{L}_A - I$, where $\mathscr{L}_A$ is a discrete time Ruelle operator (transfer operator), and $A : \{1,\dots,d\}^{\mathbb{N}} \to \mathbb{R}$ is a given fixed Lipschitz function. The associated continuous time stationary Markov chain will define the *a priori* probability.

Given a Lipschitz interaction $V : \{1,\dots,d\}^{\mathbb{N}} \to \mathbb{R}$, we are interested in Gibbs (equilibrium) state for such V. This will be another continuous time stationary Markov chain. In order to analyze this problem we will use a continuous time Ruelle operator (transfer operator) naturally associated to V. Among other things we will show that a continuous time Perron-Frobenius Theorem is true in the case V is a Lipschitz function.

We also introduce an entropy, which is negative (see also [28]), and we consider a variational principle of pressure. Finally, we analyze large deviations properties for the empirical measure in the continuous time setting using results by Y. Kifer (see [20]). In the last appendix of the paper we explain why the techniques we develop here have the capability to be applied to the analysis of convergence of a certain version of the Metropolis algorithm.

1. INTRODUCTION

In this paper we will consider thermodynamic formalism in a continuous time setting in a similar way as in [3] and [28], where the time is discrete. In order to be able to work in this new context (continuous time) we need to consider first a stationary continuous time Markov chain, and this will define the *a priori probability*, on the space of trajectories. The infinitesimal generator of this continuous time Markov chain will be associated to a discrete time Ruelle operator. Namely, we consider as the infinitesimal generator the operator $L = \mathscr{L}_A - I$, where $\mathscr{L}_A$ is a discrete time Ruelle operator.

In the continuous time setting we will be able to define a new Ruelle operator, in a similar fashion as in [28]. The continuous time setting requires some extra effort to get results, as can be seen in [2] and [22]. However, we will be able to get here the analogous properties of the Ruelle operator which appear in the discrete time setting (transfer operators). Based on the theory of stochastic processes we can define the continuous time Ruelle operator, as well as the entropy and the pressure in this new context.

The Heat-Bath Glauber dynamics is a continuous time Markov chain as described in [6]. Questions related to the Ising model on a regular tree are consider in this mentioned work. The infinitesimal generator we consider here is a generalization of (1) in this paper. Our setting is a general one where several possible models of Statistical Mechanics can fit well (see for instance [40]).

In a future work we will apply the techniques we developed here to the analysis of a special version of the Metropolis algorithm (see [35], [14], [15] and [23]) which will be suitable for applications in problems where the state space is the one-dimensional spin lattice. Suppose A is fixed for good (in this way we fix an a priori probability). Given a certain function $V : \{1,\dots,d\}^{\mathbb{N}} \to \mathbb{R}$ we would like to find the point $x \in \{1,\dots,d\}^{\mathbb{N}}$ which maximize this function. For each value $\beta > 0$ one can consider the potential βV and the associated Gibbs state $\mathbb{P}^{\beta V}$ which is a probability over the set of continuous time paths (a new continuous time Markov chain). Now, from ergodicity, if we choose at random a continuous time sample path we get a good approximation for the occupation time probability on $\{1,\dots,d\}^{\mathbb{N}}$ (Monte-Carlo method). This path can be seen as a random algorithm which is exploring the configuration space $\{1,\dots,d\}^{\mathbb{N}}$. In Appendix F we show that if we take β more and more large, then, the sample path we choose will stay more an more time close to the maximum of V. For large and fixed β it is important, from the point of view of the

Date: July 28, 2013.

algorithm, to understand the large deviation properties of the associated empirical probability of the path on $\{1,\ldots,d\}^{\mathbb{N}}$. This is related to the second part of our paper. This will be carefully explained in the end of Appendix F.

We point out that some of the results we obtain in our paper are due to the good properties already known for the classical Ruelle operators $\mathscr{L}_A$ on discrete time (transfer operators). So we begin by recalling some important topics of this subject.

Consider the shift σ acting on the one-dimensional spin lattice $\{1,\ldots,d\}^{\mathbb{N}}$. We denote by $P(B)$ the pressure of the potential $B:\{1,\ldots,d\}^{\mathbb{N}}\to\mathbb{R}$ (see [8], [32] and [33]). The value $P(B)$ is the supremum of $h(\mu)+\int B\mathrm{d}\mu$, among all σ-invariant probabilities on $\{1,\ldots,d\}^{\mathbb{N}}$, where $h(\mu)$ is the Kolmogorov entropy of the invariant probability μ. If B is Lipschitz there exists a unique μ_B such that $P(B)=h(\mu_B)+\int B\mathrm{d}\mu_B$. We call μ_B the (discrete time) equilibrium state for B (see [32] and [33]). Each point $x\in\{1,\ldots,d\}^{\mathbb{N}}$ has a finite number of preimages $y\in\{1,\ldots,d\}^{\mathbb{N}}$ by σ. For a Lipschitz potential B we define the Ruelle operator by

$$\mathscr{L}_B(f)\,(x)=\sum_{\sigma(y)=x}e^{B(y)}\,f(y)\,,$$

for any continuous function $f:\{1,\ldots,d\}^{\mathbb{N}}\to\mathbb{R}$ and $x\in\{1,\ldots,d\}^{\mathbb{N}}$. We say a Lipschitz potential $A:\{1,\ldots,d\}^{\mathbb{N}}\to\mathbb{R}$ is normalized if for any $x\in\{1,\ldots,d\}^{\mathbb{N}}$ we have

$$\sum_{\sigma(y)=x}e^{A(y)}=1\,.$$

To assume that all the potentials which we consider are Lipschitz is an essential issue (but, it could be relaxed to Holder). Nice references in thermodynamic formalism are [4] and [37].

The dual of $\mathscr{L}_A$ is the operator $\mathscr{L}_A^*$, which acts on probabilities on $\{1,\ldots,d\}^{\mathbb{N}}$ in the following way:

$$\int g\,\mathrm{d}\mathscr{L}_A^*(\nu)=\int\mathscr{L}_A(g)\,\mathrm{d}\nu\,,$$

for any continuous function $g:\{1,\ldots,d\}^{\mathbb{N}}\to\mathbb{R}$. The probability ν such that $\mathscr{L}_A^*(\nu)=\nu$ is called the (discrete time) Gibbs probability. If A is a Lipschitz normalized potential, we have $P(A)=0$, and, one can show that $\mathscr{L}_A^*(\mu_A)=\mu_A$. There is a unique fixed point probability for $\mathscr{L}_A^*$. In this case the Gibbs state for A is the equilibrium state for A (see [33]). Equilibrium states describe the probabilities that naturally appear in problems in Statistical Mechanics over the one-dimensional lattice $\{1,\ldots,d\}^{\mathbb{N}}$.

After this brief introduction on discrete time dynamics, we consider now the setting in which we will get our main results. Let $\mathscr{D}:=\mathscr{D}\big([0,+\infty),\{1,\ldots,d\}^{\mathbb{N}}\big)$ be the path space of *càdlàg* (right continuous with left limits) trajectories taking values in $\{1,\ldots,d\}^{\mathbb{N}}$ (see [27] and [34]). This space is usually endowed with the Skorohod metric (for more details about this metric see [18]), and it is called the Skorohod space. A typical element of $\mathscr{D}$ is a function $\omega:[0,\infty)\to\{1,\ldots,d\}^{\mathbb{N}}$ which is right continuous and has left limit in all points. This space is complete and has a countable dense set, in other words, it is a Polish space, but it is not compact (see [18]). The continuous time dynamics that we consider here will be given by the action of the continuous time shift $\Theta_t:\mathscr{D}\to\mathscr{D}, t\geq 0$. Given $t_0>0$ and a path $\omega\in\mathscr{D}$ on the Skorohod space, then, $\Theta_{t_0}(\omega)$ is the path η such that $\eta(t)=\omega(t+t_0)$, for all $t\geq 0$. We consider here the dynamics associated to such semiflow, $\{\Theta_t,\,t\geq 0\}$. Notice that the transformation Θ_t is not injective, because for a fixed t and for each $\eta\in\mathscr{D}$ there exists an uncountable number of preimages $\omega\in\mathscr{D}$ such that $\Theta_t(\omega)=\eta$.

We said that the probability $\tilde{\mathbb{P}}$ on the Skorohod space is invariant if it is invariant for the semiflow $\{\Theta_t, t\geq 0\}$; that is, for any Borel set $\mathscr{K}$ in $\mathscr{D}$ and $t>0$, we have $\tilde{\mathbb{P}}[\Theta_t^{-1}(\mathscr{K})]=\tilde{\mathbb{P}}[\mathscr{K}]$. In order to find invariant probabilities on the Skorohod space, it is natural to consider a continuous time Markov chain taking values on the one-dimensional spin lattice (we point out that not all invariant probabilities on the Skorohod space appear on this way). In this direction, we will use a Ruelle operator (transfer operator) with Lipschitz normalized potencial $A:\{1,\ldots,d\}^{\mathbb{N}}\to\mathbb{R}$ for defining the infinitesimal generator of a continuous time Markov chain in the form

$$(\mathscr{L}_A-I)(f)(x)=\sum_{\sigma(y)=x}e^{A(y)}\,[f(y)-f(x)]\,,$$

for all bounded measurable function $f:\{1,\ldots,d\}^{\mathbb{N}}\to\mathbb{R}$ and $x\in\{1,\ldots,d\}^{\mathbb{N}}$.

Denote by $L := \mathscr{L}_A - I$ this infinitesimal generator. For $x \in \{1,\dots,d\}^{\mathbb{N}}$, consider an initial probability measure δ_x on $\{1,\dots,d\}^{\mathbb{N}}$, and denote by $\mathbb{P}_x$ the probability measure on $\mathscr{D}$, which is induced by the infinitesimal generator L and the initial probability δ_x. It defines a Markov process $\{X_t; t \geq 0\}$ with values on the state space $\{1,\dots,d\}^{\mathbb{N}}$ (see [12], [16], and [22]). As usual, when necessary, we will consider the canonical version of the process, i.e., $X_s(\omega) = \omega(s) := \omega_s$, for any $\omega \in \mathscr{D}$ and $s \geq 0$. The stochastic semigroup generated by L is $\{P_t := e^{tL}, t \geq 0\}$ (the operator L is bounded and $L(1) \equiv 0$). The expectation concerning $\mathbb{P}_x$ is denoted by $\mathbb{E}_x$. Given μ an initial probability on $\{1,\dots,d\}^{\mathbb{N}}$, we can define the probability $\mathbb{P}_\mu$ on $\mathscr{D}$ as

$$\mathbb{P}_\mu[\mathscr{K}] = \int_{\{1,\dots,d\}^{\mathbb{N}}} \mathbb{P}_x[\mathscr{K}]\, \mathrm{d}\mu(x)\,,$$

for all Borel set $\mathscr{K} \subset \mathscr{D}$.

The above process describes the behavior of a particle, such that when located at $x \in \{1,\dots,d\}^{\mathbb{N}}$, jumps to one of its σ-preimages y, with probabilities described by $e^{A(y)}$ and after an exponential time of parameter 1. Notice that for almost every trajectory ω beginning in $x = \omega_0$, all the values $\omega_t, t \geq 0$, which are possibly attained belong to the *total pre-orbit set*, by the shift σ, of the initial point x, that is, the set of y *such that for some* $n \in \mathbb{N}$ *we have* $\sigma^n(y) = x$. The space $\{1,\dots,d\}^{\mathbb{N}}$ is not countable. We point out that in most of the papers in the literature the state space is finite (or, countable). In this last situation the infinitesimal generator is a matrix which satisfies the condition of line sum zero. Here this matrix is replaced by an operator described by the expression $L = \mathscr{L}_A - I$, where A is normalized.

The discrete Gibbs state probability μ_A over $\{1,\dots,d\}^{\mathbb{N}}$ (see [33]) for the potential $A : \{1,\dots,d\}^{\mathbb{N}} \to \mathbb{R}$ clearly satisfies that

$$\int L(f)\, \mathrm{d}\mu_A = 0\,,$$

for all f continuous function, where $L = \mathscr{L}_A - I$. This is the condition for stationarity of the initial probability of the continuous time Markov chain generated by L (see [39]). Using that $P_t = e^{tL}$, we get $\int f \mathrm{d}\mu_A = \int P_t f \mathrm{d}\mu_A$, for all f and $t \geq 0$. Therefore, μ_A is a stationary initial measure for the continuous time Markov chain associated to the stochastic semigroup $\{P_t, t \geq 0\}$. Notice that there is a unique probability such that $\mathscr{L}_A^*(\mu_A) = \mu_A$ (see [33]). This shows that the initial stationary probability for the Markov semigroup P_t is unique. The associated probability $\mathbb{P}_{\mu_A}$ on the Skorohod space is invariant for the semiflow $\{\Theta_t, t \geq 0\}$. In this way by taking different potentials A we can get a large number of invariant probabilities for the continuous time semiflow. In appendix G we show that the stationary probability $\mathbb{P}_{\mu_A}$ is ergodic for the continuous time shift $\{\Theta_t, t \geq 0\}$.

One can also ask if this process $\{X_t = X_t^{\mu_A},\ t \geq 0\}$, with initial condition μ_A, is ergodic for the stochastic semigroup, that is, if the following is true: if for a given measurable f we have that $L(f) = 0$, then, f is constant μ_A - a.s. This is indeed the case and it will be proved in the beginning of next section. We point out that now the meaning of the word ergodic for μ_A (a probability on the state space $\{1,\dots,d\}^{\mathbb{N}}$) is for the continuous time evolution of the stochastic semigroup.

The probability $\mathbb{P}_{\mu_A}$ induced on $\mathscr{D}$ by L and the initial probability μ_A will be called the *a priori probability*. The process $\{X_t = X_t^{\mu_A},\ t \geq 0\}$ is called the *a priori process*. We will need all of the above in order to define the continuous time Ruelle operator.

One can ask if $L = \mathscr{L}_A - I$ acting on the Hilbert space $\mathbb{L}^2(\mu_A)$ is symmetric. The answer to this question is no, because $L^* = \mathscr{K} - I$, where $\mathscr{K}$ is the Koopman operator, $g \to \mathscr{K}(g) = g \circ \sigma$ (according to [33]). Therefore, in our setting the process is not reversible. In order to make the system reversible we could consider, as usual, the generator $\frac{1}{2}(L + L^*)$. For this new process the particle can jump either way: forward or backward (for the action of σ). We briefly consider such process in the end of the paper (see Appendix F).

In our reasoning we will consider a fixed choice of A and this defines an *a priori* probability. After this is settled, we want to analyze the disturbed system by the intervention of an external Lipschitz potential $V : \{1,\dots,d\}^{\mathbb{N}} \to \mathbb{R}$. More precisely, we would like to obtain a new continuous time Markov chain $\{Y_T^V, T \geq 0\}$, with state space $\{1,\dots,d\}^{\mathbb{N}}$, which plays the role of the *continuous time* Gibbs state for V. In order to obtain this new process $\{Y_T^V, T \geq 0\}$, we need to define the continuous time Ruelle operator acting on functions defined in the Bernoulli space, based in Feynman-Kac theory (see, for example, [22] and [39]).

We will also need to show the existence of an eigenfunction $F:\{1,\dots,d\}^{\mathbb{N}}\to\mathbb{R}$ in the case that V is a Lipschitz function.

We will show that given a Lipschitz potential V there exists $\lambda=\lambda_V$ and a positive function $F=F_\lambda$ such that for any $T\geq 0$,

$$e^{T\,(L+V)}(F)=e^{\lambda_V\,T}F\,.$$

One can consider alternatively a continuous time Markov chain associated to a discrete time Ruelle operator in a more general setting. In fact, this will naturally occur as we will see in the analysis of the *continuous time* Gibbs state for V. When we defined the initial Markov process $\{X_t,t\geq 0\}$, we could have chosen another parameter for the exponential clock (not constant equal to 1). Below we briefly present how to proceed in this situations.

Let γ a continuous positive function and B a Lipschitz normalized potential, one could also consider a more general operator

$$L_{\gamma,B}(f)(x)=\gamma(x)\sum_{\sigma(y)=x}e^{B(y)}\,[f(y)-f(x)]\,,$$

acting on bounded measurable functions $f:\{1,\dots,d\}^{\mathbb{N}}\to\mathbb{R}$. Notice that $L_{\gamma,B}=\gamma(\mathscr{L}_B-I)$. We point out that most of the results we will prove in this paper are also true if the *a priori* probability is defined via the stochastic semigroup $\{e^{tL_{\gamma,B}},t\geq 0\}$ (and the associated stationary initial probability), instead of $\{e^{tL},t\geq 0\}$. In this case, if we denote $\mu_{B,\gamma}=\frac{1}{\gamma}\frac{\mu_B}{\int\frac{1}{\gamma}\,\mathrm{d}\mu_B}$, then, for any continuous function $f:\{1,\dots,d\}^{\mathbb{N}}\to\mathbb{R}$

$$\int L_{\gamma,B}(f)\,\mathrm{d}\mu_{B,\gamma}=0\,,$$

where μ_B is the *discrete time* equilibrium state for B. Then, $\mu_{B,\gamma}$ is the initial stationary probability for the continuous time Markov process with infinitesimal generator $L_{\gamma,B}$. It is also stationary for the flow $\{\Theta_t,t\geq 0\}$. Notice that $\mu_{B,\gamma}$ is not invariant for the discrete time action of the shift σ. The probability μ_B is invariant for the discrete time shift σ.

We denote by $\{Z_t,t\geq 0\}$ the continuous time Markov chain taking values in the one-dimensional spin lattice $\{1,\dots,d\}^{\mathbb{N}}$ generated by such $L_{\gamma,B}$ and a given initial measure (not necessarily the process needs to begin on a stationary probability). The process $\{Z_t,t\geq 0\}$ with infinitesimal generator $L_{\gamma,B}$ can be described in the following: if the particle is located at $x\in\{1,\dots,d\}^{\mathbb{N}}$, then it waits an exponential time of parameter $\gamma(x)$, and, then it jumps to a σ-preimage y with probability $e^{B(y)}$. As we will see in the third section of this paper, there exist γ and B which naturally appear when we have to describe properties of what we will call the *continuous time* Gibbs state for V.

Let's come back to the original setting where the a priory probability on the Skorohod space was defined by the process defined by the infinitesimal generator $L=\mathscr{L}_A-I$. In order to present in advance the final solution, we can say that the *continuous time* Gibbs state for V is the process $\{Y^V_T,T\geq 0\}$, which has the infinitesimal generator acting on bounded mensurable functions $f:\{1,\dots,d\}^{\mathbb{N}}\to\mathbb{R}$ given by

$$L^V(f)(x)=\gamma_V(x)\sum_{\sigma(y)=x}e^{B_V(y)}\big[f(y)-f(x)\big]\,,$$

where $B_V(y):=A(y)-\log\gamma_V(\sigma(y))+\log F_V(y)-\log F_V(\sigma(y))$, $\gamma_V(x):=1-V(x)+\lambda_V$ and the function F_V is such that

$$\frac{\mathscr{L}_A(F_V)(x)}{F_V(x)}=\sum_{\sigma(y)=x}\frac{e^{A(y)}\,F_V(y)}{F_V(x)}=1-V(x)+\lambda_V\,.$$

The appearance of the term γ_V in the infinitesimal generator L^V introduce a new element which was not present in the classical discrete time setting. This continuous time stationary Markov chain describes the solution one naturally get, from the point of view of Statistical Mechanics, for a system under the influence of an external potential V.

Now, we can ask: "Is there a maximizing pressure principle on this setting?" and "Can we talk about entropy in this setting?" In other words: is this stationary Gibbs probability an equilibrium measure in some sense? These questions appear naturally for the discrete time Ruelle operator setting (thermodynamic formalism). Answering these questions is one of the purposes of the present work. Given an a priory probability (associated to A) we will define an entropy for a class of continuous time Markov chains. It will be a non-positive number.

Lastly, we study the large deviation principle for the empirical measure associated to the a priori *process*. So that one can consider, for each $t \geq 0$ and each $\omega \in \mathscr{D}$, the empirical probability L_t^ω defined by the occupational time of the process $\{X_t, t \geq 0\}$ on a set, that is, for any Borel $\Gamma \subset \{1,\dots,d\}^{\mathbb{N}}$, we have

$$L_t^\omega(\Gamma) = \frac{1}{t}\int_0^t \mathbf{1}_\Gamma(X_s(\omega))\,\mathrm{d}s\,.$$

Then, under ergodicity, we have $\lim_{t\to\infty} L_t^\omega = \mu_A$, $\mathbb{P}_{\mu_A}$-almost surely ω (see page 108 in [39]). The Ergodic Theorem says little or nothing about the rate of convergence. Since L_t^ω is random, it is almost unavoidable to ask oneself about deviations from the stationary measure μ_A.

Let $\mathscr{M}(\{1,\dots,d\}^{\mathbb{N}})$ be the set of all measures on $\{1,\dots,d\}^{\mathbb{N}}$. The large deviation rate function $I : \mathscr{M}(\{1,\dots,d\}^{\mathbb{N}}) \to \mathbb{R}$, associated to this continuous time process $\{X_t, t \geq 0\}$, helps to estimate the exponential decay of the asymptotic empirical probability of deviations from the stationary measure μ_A, when the time parameter t goes to infinity. Thus, we are naturally led to the investigation and identification of the large deviations rate function in the set of the measures on Bernoulli space. We will analyze large deviation properties of the empirical probability (as we mentioned before the system we consider is not reversible). This is also known as level two large deviation theory (see [16], [17] and [20]). The level one large deviation principle follows by standard procedures: Orey's contraction principle (see for instance [30]).

It is important to remark that the understanding of previous results which were obtained for a general potential V plays a fundamental role in the large deviation properties of the unperturbed system (with infinitesimal generator $L = \mathscr{L}_A - I$). This follows the general philosophy of [10], [12] [20] and [21].

Suppose λ_V is the main eigenvalue we get from the continuous time Ruelle-Perron Operator for V. We denote by $\mathscr{C}$ the set of continuous functions and by $\mathscr{C}^+$ the set of strictly positive continuous functions.

Our main result in the second part of the paper is:

Theorem A: A large deviation principle at level two for the *a priori process* $\{X_t = X_t^{\mu_A},\ t \geq 0\}$ generated by $L = \mathscr{L}_A - I$ is true with the deviation function I given

$$I(\nu) = \sup_{V\in\mathscr{C}} \Big(\int V\,\mathrm{d}\nu - Q(V) \Big),$$

where $Q(V)$ is a function which is equal to the main eigenvalue λ_V when V is Lipschitz.

Moreover,

$$I(\nu) = -\inf_{u\in\mathscr{C}^+} \int \frac{L(u)}{u}\,\mathrm{d}\nu\,.$$

We point out that the above Theorem 25 in [20] (see also [21]) is presented in a different setting: the state space is a Riemannian manifold and it is considered a certain class of differential operators as infinitesimal generators. We do not consider here such differentiable structure.

The paper is divide in sections as follows: in Section 2, we present the continuous time Ruelle operator and we prove the continuous time Perron-Frobenius Theorem. In Section 3, we present the continuous time Gibbs state for V. This is a continuous time stationary process. In Section 4, we define relative entropy, pressure and equilibrium state for V, and we also prove a variational principle for the Gibbs state. In Section 5, the main result that we will get is the large deviation principle for the empirical measure associated to the *a priori process*. Finally, in the Appendix we show many technical results using basic tools of continuous time Markov chains. Among them: we present a Radon-Nikodim derivative result, we briefly comment on the spectrum of $\mathscr{L}_A - I + V$ on $\mathbb{L}^2(\mu)$, where μ is a natural probability on the Bernoulli space $\{1,\dots,d\}^{\mathbb{N}}$, and, finally, some remarks on the associated symmetric process. In this last section we consider a fixed potential V and we ask about the limit of the invariant probability (invariant for the continuous time equilibrium Gibbs state for βV, when β is large) over $\{1,\dots,d\}^{\mathbb{N}}$ when temperature goes to zero.

2. DISTURBING THE SYSTEM BY AN EXTERNAL LIPSCHITZ POTENTIAL V: THE CONTINUOUS TIME PERRON-FROBENIUS THEOREM.

First of all we recall the definition of the *a priori process*. A Lipschitz normalized potential A will be considered fixed through the whole paper. We denote by $\{P_t, t \geq 0\}$, the stochastic semigroup generated by $L = \mathscr{L}_A - I$. We need an *a priori* continuous time stationary probability for our reasoning, for this reason we are considering $\mathbb{P}_{\mu_A}$ the probability obtained from the semigroup $\{P_t, t \geq 0\}$ and the initial probability μ_A. As we have said, this probability $\mathbb{P}_{\mu_A}$ plays the role of the *a priori* measure (see [3] and [28]). The associated stochastic process will be denoted by $\{X_t = X_t^{\mu_A},\ t \geq 0\}$.

Given a continuous time stochastic semigroup with compact state space and an initial stationary probability we get a continuous time invariant probability on the Skhorohod space. The continuous time Birkhoff Theorem associated to the continuous time stochastic semigroup (for not necessarily ergodic probabilities) is true (see Remark 1 on page 382 in [41] or Theorem 17 page 708 and Exercise 19 page 721 in [12]).

Therefore, given a continuous function $g : \{1,\dots,d\}^{\mathbb{N}} \to \mathbb{R}$ we get an integrable measurable function $f : \{1,\dots,d\}^{\mathbb{N}} \to \mathbb{R}$ which describes the possible mean continuous time limits for g. This function f is invariant for the action of the stochastic semigroup. Therefore, $L(f) = 0$. In the case f is constant μ_A - a.e.w. then the mean continuous time limits for g are all the same μ_A - a.e.w. and equal to the μ_A space average on $\{1,\dots,d\}^{\mathbb{N}}$.

The probability μ_A is ergodic for the continuous time action, that is, the following is true: if for a given f we have that $L(f) = 0$, then, f is constant μ_A - a.e.w. This follows from the following simple argument suggested by D. Smania: suppose $\mathscr{L}_A(f) = f$ for a μ_A-integrable f, then, for a given ε we can write $f = g + w$, where w is integrable with $L^1(\mu_A)$ norm smaller than ε and g is Lipchitz. Then, $f = \mathscr{L}_A^n(f) = \mathscr{L}_A^n(g) + \mathscr{L}_A^n(w)$.

Note that $\mathscr{L}_A^n(w)$ has L^1 norm smaller then ε. Moreover, $\mathscr{L}_A^n(g)$ converges to a constant a_g, where a_g is $\int f d\mu_A$ up to ε (see Theorem 2.2 (iv) [33]). Therefore, taking the limit in n we get that $f - a_g$ has norm smaller than ε. Now, taking $\varepsilon \to 0$, we get that a_g converges to $\int f d\mu_A$. Therefore, for all x, μ_A - a.e.w, we have that $f(x) = \int f d\mu_A$.

In the same spirit of Classical thermodynamic formalism (see [33]), given a potential V (an interaction), we want to get here another continuous time Markov process which will be the equilibrium stationary process for the system under the influence of the potential V.

Let $V : \{1,\dots,d\}^{\mathbb{N}} \to \mathbb{R}$ a Lipschitz function and consider the operator $L + V = \mathscr{L}_A - I + V$, which acts on mensurable and bounded functions $f : \{1,\dots,d\}^{\mathbb{N}} \to \mathbb{R}$ by the expression

$$(L+V)(f)(x) \,=\, (\mathscr{L}_A - I)(f)(x) \,+\, V(x)f(x)\,,$$

for all $x \in \{1,\dots,d\}^{\mathbb{N}}$. For $T \geq 0$, we consider

$$P_T^V(f)(x) \,:=\, \mathbb{E}_x\big[e^{\int_0^T V(X_r)\,dr} f(X_T)\big]\,, \tag{1}$$

for all continuous function $f : \{1,\dots,d\}^{\mathbb{N}} \to \mathbb{R}$ and $x \in \{1,\dots,d\}^{\mathbb{N}}$. By Feynman-Kac, $\{P_T^V, T \geq 0\}$ defines a semigroup associated to the infinitesimal operator $L + V = \mathscr{L}_A - I + V$ (see Appendix 1.7 in [22]).

Let $\mathscr{C}$ be the space of continuous functions from $\{1,\dots,d\}^{\mathbb{N}}$ to $\mathbb{R}$ endowed with uniform topology. Denote by $\mathscr{C}^+$ the subspace of functions of $\mathscr{C}$ which are strictly positive. Let $\mathscr{P}(\{1,\dots,d\}^{\mathbb{N}})$ be the space of probabilities on the Borel sigma-algebra of the one-dimensional spin lattice $\{1,\dots,d\}^{\mathbb{N}}$. Define $\mathscr{M}(\{1,\dots,d\}^{\mathbb{N}})$ as the space of measures on the Borel sigma-algebra of the Bernoulli space $\{1,\dots,d\}^{\mathbb{N}}$.

Notice that, in general, this semigroup in not stochastic, because $P_T^V(1)(x) \neq 1$. We want to associate to this semigroup, another one which is also stochastic, this will be only possible due to the next result, which we consider the main one in this section.

Theorem 1 (Continuous Time Perron-Frobenius Theorem)**.** *Suppose that V is a Lipschitz function. Then, there exists a strictly positive Lipschitz eigenfunction $F : \{1,\dots,d\}^{\mathbb{N}} \to (0,+\infty)$ for the family of operators $P_T^V : \mathscr{C} \to \mathscr{C}$, $T \geq 0$, associated to an eigenvalue $e^{\lambda_V T}$, where $\lambda = \lambda_V$ depends only on V. By this we mean: for any $T \geq 0$,*

$$P_T^V(F) = e^{\lambda_V T} F\,.$$

The eigenvalue λ_V is simple and it is equal to the spectral radius (maximal). Moreover, there exists a eigenprobability ν_V in $\mathscr{P}(\{1,\dots,d\}^{\mathbb{N}})$ such that

$$(P_T^V)^*(\nu_V) = e^{\lambda_V T}\nu_V\,, \quad \forall T \geq 0\,.$$

The proof of this theorem we will present in the Subsections 2.1 and 2.2.

As a consequence of this theorem we will be able to normalize the semigroup $\{P_T^V,\, T \geq 0\}$ in order to get another stochastic semigroup, and, then we will finally obtain what we call the Gibbs state in the continuous time setting.

A quite simple version of this result was presented in [2]. In this paper, V depends just on X_0 and the state space is $\{1,2,\dots,d\}$.

Example 2. *To clarify ideas, we present a simple example where is easy to verify the validity of the above theorem. Given $0 < p_1 < 1$, $0 < p_2 < 1$, the stochastic matrix*

$$\begin{pmatrix} 1-p_1 & p_1 \\ p_2 & 1-p_2 \end{pmatrix},$$

defines a Ruelle operator $\mathscr{L}_A$ acting on the one-dimensional spin lattice $\{1,2\}^{\mathbb{N}}$ such that, $\mathscr{L}_A(1) = 1$. More precisely, $e^{A(1,1,x_2,\dots)} = 1-p_1$, $e^{A(2,1,x_2,\dots)} = p_1$ and $e^{A(1,2,x_2,\dots)} = p_2$, $e^{A(2,2,x_2,\dots)} = 1-p_2$. Notice that

$$L = \begin{pmatrix} 1-p_1 & p_1 \\ p_2 & 1-p_2 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -p_1 & p_1 \\ p_2 & -p_2 \end{pmatrix}$$

defines a line sum zero matrix. One can consider a potential V such that is constant in the cylinders of size one, i.e., $V(1,x_1,x_2,\dots) = V_1$, and $V(2,x_1,x_2,\dots) = V_2$ In this case $L+V$ is the matrix

$$\begin{pmatrix} -p_1+V_1 & p_1 \\ p_2 & -p_2+V_2 \end{pmatrix}$$

If V_1,V_2 are positive and large then the positive cone goes inside the positive cone. Then, there is a positive eigenvalue and a positive eigenfunction. One can add a constant to V in order to get an eigenvector with just positive entries.

We will consider on the Bernoulli space the usual metric d. Let $0 < \theta < 1$, then for all $x = \{x_i\}, y = \{y_i\} \in \{1,\dots,d\}^{\mathbb{N}}$

$$d(x,y) := \theta^N,$$

where N is such that $x_i = y_i$, $\forall i \leq N$ and $x_{N+1} \neq y_{N+1}$. In the following, when $a \in \{1,\dots,d\}$ and $x \in \{1,\dots,d\}^{\mathbb{N}}$ the notation ax means $(a,x_1,x_2,\dots) \in \{1,\dots,d\}^{\mathbb{N}}$, i.e., ax is a preimage of x by shift operator.

We point out that $d(ax,ay) \leq \theta d(x,y)$, for all $x,y \in \{1,\dots,d\}^{\mathbb{N}}$ and $a \in \{1,\dots,d\}$, this is a central idea in Lemma 5, when we estimate the ratio $\frac{P_T^V(f)(x)}{P_T^V(f)(y)}$. First, we will characterize the operator P_T^V, in Lemma 3. This characterization allow us to conclude that the family of operators $\{P_T^V,\, T \geq 0\}$ describes a natural generalization of the discrete time Ruelle operator (see [2]).

Lemma 3. *Let $f \in \mathscr{C}$, $T \geq 0$, and $x \in \{1,\dots,d\}^{\mathbb{N}}$. Consequently, $P_T^V(f)(x) = \mathbb{E}_x\big[e^{\int_0^T V(X_r)\,dr} f(X_T)\big]$ can be rewritten as*

$$e^{TV(x)}f(x)e^{-T} + \sum_{n=1}^{+\infty}\sum_{a_1=1}^{d}\cdots\sum_{a_n=1}^{d} e^{A(a_1x)}\dots e^{A(a_n\dots a_1x)} f(a_n\dots a_1x)\,\mathscr{I}_V^T(a_n\dots a_1x)\,,$$

where

$$\mathscr{I}_V^T(a_n\dots a_1x) = \int_0^\infty dt_n\dots\int_0^\infty dt_0\, e^{t_0V(x)+\cdots+(T-\sum_{i=0}^{n-1}t_i)V(a_n\dots a_1x)}\,\mathbf{1}_{[\sum_{i=0}^{n-1}t_i \leq T < \sum_{i=0}^{n}t_i]}\, e^{-t_0}\dots e^{-t_n}\,.$$

As the proof of this lemma is very technical we present it in Appendix B.

Observe that, if one consider $V \equiv 0$, the previous lemma says that

$$\begin{aligned} P_T(f)(x) &= \mathbb{E}_x\big[f(X_T)\big] \\ &= e^{-T}\Big\{f(x) + \sum_{n=1}^{+\infty}\frac{T^n}{n!}\sum_{a_1=1}^{d}\cdots\sum_{a_n=1}^{d} e^{A(a_1x)}\dots e^{A(a_n\dots a_1x)}f(a_n\dots a_1x)\Big\} \\ &= e^{-T}\Big\{f(x) + \sum_{n=1}^{+\infty}\frac{T^n}{n!}(\mathscr{L}_A^n(f))(x)\Big\}, \end{aligned}$$

because

$$\int_0^\infty dt_n\dots\int_0^\infty dt_0\,\mathbf{1}_{[\sum_{i=0}^{n-1}t_i\le T<\sum_{i=0}^{n}t_i]}e^{-t_0}\dots e^{-t_n} = e^{-T}\,\frac{T^n}{n!}\,.$$

Thus, $P_T(f)(x) = \frac{1}{e^T}e^{T\mathscr{L}_A}(f)(x)$, which is in accordance with the fact that $\{P_T,\,T\ge 0\}$ is the semigroup associated to the generator $L = \mathscr{L}_A - I$.

Lemma 4. *For any non-negative continuous function f such that there exist $x\in\{1,\dots,d\}^{\mathbb{N}}$ and $T>0$ satisfying $P_T(f)(x) = 0$, we have that $f\equiv 0$.*

Proof. By the Lemma 3, $f(a_n,\dots,a_1x) = 0$, for all $a_i\in\{1,2,\dots,d\}$, $i=1,\dots,n$, for any $n\in\mathbb{N}$. Then $f(z)=0$, for any $z\in\{y;$ there exists n such that $\sigma^n(y)=x\}$. But this set is dense in $\{1,\dots,d\}^{\mathbb{N}}$ and f is continuous, thus $f(z)=0$, for any $z\in\{1,\dots,d\}^{\mathbb{N}}$. □

Lemma 5. *If the function f satisfies $f(x)\le e^{C_f d(x,y)}f(y)$, for all $x,y\in\{1,\dots,d\}^{\mathbb{N}}$, where C_f is a constant depending only on f, then*

$$P_T^V(f)(x) \le \exp\Big\{\big[(C_A\theta + TC_V)(1-\theta)^{-1} + C_f\theta\big]\,d(x,y)\Big\}\,P_T^V(f)(y)\,,$$

for all $T\ge 0$.

The proof of this lemma is in Appendix B (it is similar to the proof of the Lemma 3).

2.1. **Eigenprobability.** In this subsection we will present the proof of existence of eigenprobability. Without loss of generality, we will assume that the perturbation V is positive and its minimum is large enough (just add a large constant to the initial V). We will find an eigenprobability for $\mathscr{L}_A - I + V$. The constant we eventually add to the in initial potential will not harm our argument.

First we need to analyze the dual of $\mathscr{L}_A - I + V$ acting on signed measures.

As we know $(\mathscr{L}_A - I + V)^*$ acts on measures on the Bernoulli space via the expression: given ν, then

$$\big\langle f, (\mathscr{L}_A - I + V)^*(\nu)\big\rangle = \big\langle(\mathscr{L}_A - I + V)(f),\nu\big\rangle,$$

for any $f\in\mathscr{C}$. This leads us to consider the operator G on probabilities of the one-dimensional spin lattice. Given ν probability on $\{1,\dots,d\}^{\mathbb{N}}$, G acts on ν as

$$\big\langle f, G(\nu)\big\rangle = \frac{\big\langle(\mathscr{L}_A - I + V)(f),\nu\big\rangle}{\big\langle(\mathscr{L}_A - I + V)(1),\nu\big\rangle} = \frac{\big\langle(\mathscr{L}_A - I + V)(f),\nu\big\rangle}{\big\langle V,\nu\big\rangle}\,,$$

for any $f\in\mathscr{C}$. The function G is well defined by the hypothesis on V. This G is continuous, because it is the ratio of two continuous functions. From Schauder-Tychonoff Theorem, we get the existence of a fixed point probability ν_V for G. Therefore, there exists $\lambda_V = \int V\,\mathrm{d}\nu_V$ such that

$$\int(\mathscr{L}_A - I + V)(f)\,\mathrm{d}\nu_V = \big\langle(\mathscr{L}_A - I + V)(f),\nu_V\big\rangle = \lambda_V\,\big\langle f,\nu_V\big\rangle = \lambda_V\int f\,\mathrm{d}\nu_V\,,$$

for any $f\in\mathscr{C}$. Since $L = \mathscr{L}_A - I$, we have

$$\int(L + V - \lambda_V)(f)\,\mathrm{d}\nu_V = 0\,, \tag{2}$$

for any $f\in\mathscr{C}$. By Feynman-Kac, the semigroup associated to operator $L+V-\lambda_V$ is $\frac{P_T^V}{e^{\lambda_V T}}$. Using the Trotter-Kato Theorem (see chapter IX section 12 in [41]), we get

$$\frac{P_T^V(f)}{e^{\lambda_V T}} = \lim_{n\to\infty}\Big(I - \frac{T}{n}(L + V - \lambda_V)\Big)^n(f)\,.$$

Observe that is true

$$\int \Big(I - \tfrac{T}{n}(L+V-\lambda_V)\Big)^n (f)\,\mathrm{d}\nu_V = \int f\,\mathrm{d}\nu_V\,, \quad \forall n,$$

and, this is a consequence of two properties: the first one is that when the operator $L+V-\lambda_V$ acts on $\mathscr{C}$ its image is contained $\mathscr{C}$ too; the second one is the equality (2). By Dominated Convergence Theorem, we get

$$\int \frac{P_T^V(f)}{e^{\lambda_V T}}\,\mathrm{d}\nu_V \,=\, \int f\,\mathrm{d}\nu_V\,, \tag{3}$$

for any $f \in \mathscr{C}$. Consequently,

$$\int f\,\mathrm{d}\big[(P_T^V)^*(\nu_V)\big] \,=\, e^{\lambda_V T}\int f\,\mathrm{d}\nu_V\,,$$

for any $f \in \mathscr{C}$.

2.2. **Eigenfunction.** Here, we present the existence of an eigenprobability.

Suppose that $\theta \le 1/2$. Let

$$\Lambda = \big\{f \in \mathscr{C};\, 0 \le f \le 1 \text{ and } f(x) \le \exp\{\tfrac{C_A+C_V}{1-\theta}d(x,y)\}f(y), \forall x,y \in \{1,\dots,d\}^{\mathbb{N}}\big\}\,.$$

The set Λ is convex, because for all $f,g \in \Lambda$ and $t \in (0,1)$

$$tf(x)+(1-t)g(x) \le \exp\{\tfrac{C_A+C_V}{1-\theta}d(x,y)\}\big(tf(y)+(1-t)g(y)\big)\,.$$

Let $\{f_n\} \subset \Lambda$, then $\|f_n\|_\infty \le 1$ and

$$\begin{aligned}|f_n(x)-f_n(y)| &\le \|f_n\|_\infty\Big(\exp\{\tfrac{C_A+C_V}{1-\theta}d(x,y)\}-1\Big)\\ &\le \tfrac{C_A+C_V}{1-\theta}d(x,y)\exp\{\tfrac{C_A+C_V}{1-\theta}\}\,,\end{aligned}$$

for all $n \in \mathbb{N}$. By Arzelà-Ascoli Theorem, the sequence $\{f_n\}$ has a limit point. Therefore, Λ is a compact set.

By the Lemma 5, for all $f \in \Lambda$, we have

$$P_T^V(f)(x) \le \exp\Big\{\big[\tfrac{C_A\theta+TC_V}{1-\theta}+\tfrac{C_A+C_V}{1-\theta}\theta\big]\,\mathrm{d}(x,y)\Big\}P_T^V(f)(y)\,, \quad \forall T \ge 0\,.$$

Take $T \le \theta$, then

$$P_T^V(f)(x) \le \exp\Big\{2\tfrac{C_A+C_V}{1-\theta}\theta\,\mathrm{d}(x,y)\Big\}P_T^V(f)(y) \le \exp\Big\{\tfrac{C_A+C_V}{1-\theta}\,\mathrm{d}(x,y)\Big\}P_T^V(f)(y)\,.$$

The last inequality is due to the assumption about θ. Unfortunately, $P_T^V(f)$ can be greater than one, then we need to define for all $n \in \mathbb{N}$, the operator Q_T^n that acts on $g \in \Lambda$ as

$$Q_T^n(g) := \frac{P_T^V(g+1/n)}{\big\|P_T^V(g+1/n)\big\|_\infty}\,.$$

Notice that, for all $n \in \mathbb{N}$, the function constant equal to $1/n$ belongs to Λ, then

$$P_T^V(1/n)(x) \le \exp\{\tfrac{C_A+C_V}{1-\theta}\,\mathrm{d}(x,y)\}P_T^V(1/n)(y)\,,$$

for all $T \in [0,\theta]$. This allows us to show that $Q_T^n : \Lambda \to \Lambda$, for all $n \in \mathbb{N}$.

Since Λ is convex and a compact set, we can apply the Schauder-Tychonoff Fixed Point Theorem to each $Q_T^n : \Lambda \to \Lambda$ and see that there exists $h_n^T \in \Lambda$ such that

$$\frac{P_T^V(h_n^T+1/n)}{\big\|P_T^V(h_n^T+1/n)\big\|_\infty} = h_n^T\,,\ \forall n\,,\ \forall T \in [0,\theta]\,. \tag{4}$$

Now, for fixed $T \in [0,\theta]$, there exists $F_T \in \Lambda$ a limit point of the sequence $\{h_n^T\}_n \subset \Lambda$, because Λ is compact. By the continuity of the operator P_T^V, the expression above becomes

$$P_T^V(F_T) = \big\|P_T^V(F_T)\big\|_\infty F_T\,,\ \forall T \in [0,\theta]\,. \tag{5}$$

First of all, we would like to prove that $F_T > 0$. Hence, we begin to analyze the norm $\big\|P_T^V(F_T)\big\|_\infty$. By the equation (4), we have

$$\big\|P_T^V(h_n^T+1/n)\big\|_\infty h_n^T(x) = \mathbb{E}_x\big[e^{\int_0^T V(X_r)\,dr}(h_n^T+1/n)(X_T)\big] \ge \Big[\big(\inf h_n^T\big)+1/n\Big]\,e^{-T\|V\|_\infty}\,,$$

for all x. Then,

$$\left(\left\|P_T^V(h_n^T+1/n)\right\|_\infty - e^{-T\|V\|_\infty}\right)\inf h_n^T \geq (1/n)\,e^{-T\|V\|_\infty} > 0\,.$$

This implies that

$$\left\|P_T^V(h_n^T+1/n)\right\|_\infty > e^{-T\|V\|_\infty}\,, \quad \forall n\,.$$

Recalling that F_T is a limit point of $\{h_n^T\}_n$, the last inequality is transformed in

$$\left\|P_T^V(F_T)\right\|_\infty \geq e^{-T\|V\|_\infty}\,. \tag{6}$$

Finally, suppose that $F_T(x_0)=0$, for some $x_0\in\{1,\dots,d\}^{\mathbb{N}}$. Due to the fact that F_T is eigenfunction of the operator P_T^V, we have $P_T^V(F_T)(x_0)=0$. Using the Lemma 4, we get that $F_T\equiv 0$. But it is a contraction in relation to (6), because P_T^V is linear. As a result $F_T>0$.

Now, we will characterize the eigenvalue, in order to do this we use the eigenprobability ν_V. The equations (5) and (3) together imply that

$$\left\|P_T^V(F_T)\right\|_\infty \int F_T\,d\nu_V = \int P_T^V(F_T)\,d\nu_V = e^{\lambda_V T}\int F_T\,d\nu_V\,.$$

Since $F_T\geq 0$, we get $\|P_T^V(F_T)\|_\infty = e^{\lambda_V T}$, and using (5) one can conclude $P_T^V(F_T)=e^{\lambda_V T}F_T$, $\forall T\in[0,\theta]$.

The next step is to prove that $e^{\lambda_V T}$ is a simple eigenvalue for P_T^V. We suppose that for each $T\in[0,\theta]$ there exists G_T such that $P_T^V(G_T)=e^{\lambda_V T}G_T$. Define $\alpha_0^T:=\inf_x\frac{G_T(x)}{F_T(x)}$. Since the Bernoulli space is compact, there exist $x_0\in\{1,\dots,d\}^{\mathbb{N}}$ such that $G_T(x_0)-\alpha_0^T F_T(x_0)=0$. Observe that $H_T:=G_T(x)-\alpha_0^T F_T(x)$ is a non-negative eigenfunction of P_T^V. Then $P_T^V(H_T)(x_0)=0$. By Lemma 4, $H_T\equiv 0$. Thus, G_T is a scalar multiple of F_T. This shows that $e^{\lambda_V T}$ is a simple eigenvalue.

We will try to eliminate the dependence on $T\in[0,\theta]$ in the functions F_T. Recall that $\theta\leq 1/2$. Let $n_0:=\min\{n;\,2^{-n}\leq\theta\}$. Denote by $F:=F_{2^{-n_0}}$. We claim that $P_{2^{-n}}^V(F)=e^{\lambda_V 2^{-n}}F$, $\forall n\geq n_0$. To prove this note that by the semigroup property we have that $P_{2^{-n_0}}^V(F_{2^{-n}})$ can be rewritten as $P_{2^{-n}}^V\dots P_{2^{-n}}^V(F_{2^{-n}})$, $\forall n\geq n_0$. Applying 2^{n-n_0} times the fact that $F_{2^{-n}}$ is eigenfunction of the operator $P_{2^{-n}}^V$, we have $P_{2^{-n_0}}^V(F_{2^{-n}})=e^{\lambda_V 2^{-n_0}}F_{2^{-n}}$, $\forall n\geq n_0$. Since $e^{\lambda_V 2^{-n_0}}$ is simple eigenvalue to the operator $P_{2^{-n_0}}^V$, we get $F_{2^{-n}}=F$, $\forall n\geq n_0$. This finishes the claim.

The last claim and the fact that the semigroup $\{P_T^V,\,T\geq 0\}$ is associated to the operator $L+V$ imply that

$$(L+V)(F)=\lim_{n\to\infty}\frac{P_{2^{-n}}^V(F)-F}{2^{-n}}=\lim_{n\to\infty}\frac{e^{\lambda_V 2^{-n}}-1}{2^{-n}}F=\lambda_V F\,.$$

Since the operator $L+V=\mathscr{L}_A-I+V$ is a bounded operator, using the equality above we get

$$P_T^V(F)(x)=e^{T(L+V)}(F)(x)=\sum_{n=0}^{\infty}\frac{T^n}{n!}(L+V)^n(F)(x)=\sum_{n=0}^{\infty}\frac{T^n}{n!}\lambda_V^n F(x)=e^{\lambda_V T}F(x)\,,$$

for any $T\geq 0$.

Therefore, with these final considerations, we finished the proof of one of our main results, which is Theorem 1 (Perron-Frobenius). Notice that λ_V is both eigenvalue for the eigenfunction (see Section 2.2) and also eigenvalue for the dual operator (section 2.1).

In terms of discrete time dynamics we just showed the following result:

Corollary 6. *Given a normalized Lipschitz potential $A:\{1,\dots,d\}^{\mathbb{N}}\to\mathbb{R}$ and a Lipschitz function $V:\{1,\dots,d\}^{\mathbb{N}}\to\mathbb{R}$, there exists Lipschitz function $F=F_V:\{1,\dots,d\}^{\mathbb{N}}\to\mathbb{R}$ and $\lambda=\lambda_V$ such that, for all $x\in\{1,\dots,d\}^{\mathbb{N}}$*

$$\frac{\mathscr{L}_A(F)(x)}{F(x)}=\sum_{\sigma(y)=x}\frac{e^{A(y)}F(y)}{F(x)}=1-V(x)+\lambda\,. \tag{7}$$

Notice that the addition of a constant to V produces an additive change in the eventual eigenvalue λ.

3. THE CONTINUOUS TIME GIBBS STATE FOR V

From the Perron-Frobenius Theorem associated to V, we can define a new continuous time Markov chain which will be the Gibbs state for V. Remember that $L+V=\mathscr{L}_A-I+V$ generates the semigroup $\{P_T^V, T\geq 0\}$.

For $T\geq 0$, if one defines

$$\mathscr{P}_T^V(f)(x)=\mathbb{E}_x\Big[e^{\int_0^T V(X_r)dr}\,\frac{F(X_T)}{e^{\lambda_V T}F(x)}\,f(X_T)\Big]=\frac{P_T^V(Ff)(x)}{e^{\lambda_V T}F(x)}\,, \tag{8}$$

where F and λ_V are the eigenfunction and the eigenvalue, respectivelly. Then $\mathscr{P}_T^V(1)(x)=1$, $\forall x\in\{1,\dots,d\}^{\mathbb{N}}$. This will define the stochastic semigroup we were looking for. From this we will get a new continuous time Markov chain which will help to define the Gibbs state for V.

We point out that $\frac{\mathscr{L}_A(F)}{F}(y)=1-V(y)+\lambda_V=\gamma_V(x)>c>0$, for some positive c. We can say that because F and $\mathscr{L}_A(F)$ are continuous strictly positive functions and the state space is compact.

From the above, it is natural to consider a new normalized Lipschitz potential B_V and a function γ_V defined by

$$\begin{aligned}&B_V(y):=A(y)-\log\left(1-V(\sigma(y))+\lambda_V\right)+\log F(y)-\log F(\sigma(y))\,,\quad \forall y\in\{1,\dots,d\}^{\mathbb{N}}\\ &\text{and}\qquad \gamma_V(x):=1-V(x)+\lambda_V\,,\quad \forall x\in\{1,\dots,d\}^{\mathbb{N}}\,,\end{aligned} \tag{9}$$

where V, F and λ_V were introduced before.

Proposition 7. *If V is a Lipschitz function we define the operator L^V acting on bounded mensurable functions $f:\{1,\dots,d\}^{\mathbb{N}}\to\mathbb{R}$ as*

$$L^V(f)(x)=\gamma_V(x)\sum_{\sigma(y)=x}e^{B_V(y)}\left[f(y)-f(x)\right], \tag{10}$$

where $B_V(y)$ and γ_V are defined in (9)*. Then, this operator, L^V is the infinitesimal generator associated to a semigroup $\{\mathscr{P}_T^V, T\geq 0\}$ defined in* (8)*.*

Proof. We begin proving that the $\{\mathscr{P}_T^V, T\geq 0\}$ is a semigroup. Recalling its definition, we get

$$\mathscr{P}_t^V(\mathscr{P}_s^V(f))(x)=\frac{P_t^V(F\mathscr{P}_s^V(f))(x)}{e^{\lambda_V t}F(x)}\,,$$

we need to analyze $P_t^V(F\mathscr{P}_s^V(f))(x)$. In this way,

$$\begin{aligned}P_t^V(F\mathscr{P}_s^V(f))(x)&=\mathbb{E}_x\Big[e^{\int_0^t V(X_r)dr}F(X_t)\mathscr{P}_s^V(f)(X_t)\Big]\\ &=\mathbb{E}_x\Big[e^{\int_0^t V(X_r)dr}\,\frac{F(X_t)}{e^{\lambda_V s}F(X_t)}P_s^V(Ff)(X_t)\Big]=\tfrac{1}{e^{\lambda_V s}}P_{t+s}^V(Ff)(x)\,.\end{aligned}$$

One can conclude that $\{\mathscr{P}_T^V, T\geq 0\}$ is a semigroup.

To prove that the infinitesimal generator (10) is associated to this semigroup, we need to observe that

$$\frac{\mathscr{P}_t^V(f)(x)-f(x)}{t}=\frac{1}{e^{\lambda_V t}F(x)}\left(\frac{P_t^V(Ff)(x)-(Ff)(x)}{t}\right)+f(x)\left(\frac{e^{-\lambda_V t}-1}{t}\right).$$

Taking the limit as t goes to zero the expression above converges to

$$\frac{1}{F(x)}(L+V)(Ff)(x)-f(x)\lambda=-\lambda f(x)+V(x)f(x)+\tfrac{1}{F(x)}L(Ff)(x)\,, \tag{11}$$

which we denote by $L^V(f)(x)$. Using the hypotheses about V and equation (7) of the Lemma 6, we get that $L^V(f)(x)$ is equal to

$$\sum_{\sigma(y)=x}\tfrac{e^{A(y)}F(y)}{F(x)}f(y)-(1-V(x)+\lambda)f(x)=\sum_{\sigma(y)=x}\tfrac{e^{A(y)}F(y)}{F(x)}\left[f(y)-f(x)\right].$$

Again, we use the Lemma 6 to obtain $\gamma_V(x)F(x) = \mathscr{L}_A(F)(x)$. Thus, the expression above can be rewritten as

$$\gamma_V(x) \sum_{\sigma(y)=x} \tfrac{e^{A(y)}F(y)}{\mathscr{L}_A(F)(x)} \big[f(y) - f(x)\big] = \gamma_V(x) \sum_{\sigma(y)=x} e^{B_V(y)} \big[f(y) - f(x)\big] .$$

□

Corollary 8. *For all $f \in \mathscr{C}^+$, $x \in \{1,\ldots,d\}^{\mathbb{N}}$ and $t > 0$ small*

$$\log\Big(\frac{\mathscr{P}_t^V(f)(x)}{f(x)}\Big) \sim \frac{tL^V(f)(x)}{f(x)} ,$$

where $a_n \sim b_n$ means that $a_n/b_n \to 1$, as $n \to \infty$.

Proof. In the proof above we obtained that

$$\lim_{t\to 0} \frac{\mathscr{P}_t^V(f)(x) - f(x)}{t} = L^V(f)(x) ,$$

where $L^V(f)(x) = -\lambda f(x) + V(x)f(x) + \frac{1}{F(x)}L(Ff)(x)$. Then, for t small

$$\frac{\mathscr{P}_t^V(f)(x)}{f(x)} - 1 \sim \frac{tL^V(f)(x)}{f(x)} ,$$

for all $f \in \mathscr{C}^+$. Since for all x fixed and t small we get

$$\log\Big(\frac{\mathscr{P}_t^V(f)(x)}{f(x)}\Big) \sim \frac{\mathscr{P}_t^V(f)(x)}{f(x)} - 1 ,$$

we finished the proof.

□

We will elaborate now on the initial stationary probability μ_{B_V,γ_V}. Notice that all of the above depends on the choice of the initial a priori probability (which, in our case, is associated to the generator $L = \mathscr{L}_A - I$). The stationary measure for the continuous time process generated by L^V (with exponential time of jump equal to $\gamma(x) = \gamma_V(x) = 1 - V(x) + \lambda_V$) is

$$d\mu_{B_V,\gamma_V}(x) = \frac{1}{\gamma_V(x)} \frac{d\mu_{B_V}(x)}{\int \frac{1}{\gamma_V} d\mu_{B_V}} , \tag{12}$$

where μ_{B_V} is discrete time equilibrium for the normalized Lipschitz potential $B_V(y) = A(y) + \log F(y) - \log F(\sigma(y)) - \log\gamma_V(\sigma(y))$. In other words, for any $f \in \mathscr{C}$, we have

$$\int L^V(f) \, d\mu_{B_V,\gamma_V} = 0 .$$

As we said before, the appearance of the term $\frac{1}{\gamma_V}$ introduce a new element, which was not present in the classical discrete time setting.

Definition 9. *Given a Lipschitz function V, we define a continuous time Markov process $\{Y_T^V, T \geq 0\}$ with state space $\{1,\ldots,d\}^{\mathbb{N}}$ whose infinitesimal generator L^V acts on bounded mensurable functions $f : \{1,\ldots,d\}^{\mathbb{N}} \to \mathbb{R}$ by the expression*

$$L^V(f)(x) = \gamma_V(x) \sum_{\sigma(y)=x} e^{B_V(y)} \big[f(y) - f(x)\big] , \tag{13}$$

where B_V and γ_V are defined in (9). *Now, we consider the initial stationary probability μ_{B_V,γ_V} defined in* (12). *We call this process $\{Y_T^V, T \geq 0\}$* ***the continuous time Gibbs state for the potential*** *V. This defines a probability $\mathbb{P}^V := \mathbb{P}^V_{\mu_{B_V,\gamma_V}}$ on the Skorohod space $\mathscr{D}$ which we call* ***the Gibbs probability for the interaction*** *V.*

Notice that for $\{Y_T^V, T \geq 0\}$, the exponential time of jumping tends to be larger when we are close to the maximum of V. For a generic continuous time path, the particle stays more time on this region.

If V is of the form $-\frac{L(u)}{u}$, for some $u \in \mathscr{C}^+$, then, $\lambda = 0$, and $\mu_A = \mu_{B_V}$. In this case $\gamma = \frac{\mathscr{L}_A(u)}{u}$.

4. Relative Entropy, Pressure and the equilibrium state for V

One can ask: "Did the Gibbs state of the last section satisfy a variational principle?" We will address this question in the present section.

Definition 10. *The probability $\tilde{\mathbb{P}}_\mu = \tilde{\mathbb{P}}_\mu^{\tilde{\gamma},\tilde{A}}$ on $\mathscr{D}$ is called admissible, if it is generated by the initial measure μ and the continuous time Markov chain with infinitesimal generator $\tilde{L}$, which acts on bounded mensurable functions $f : \{1,\dots,d\}^{\mathbb{N}} \to \mathbb{R}$ by*

$$\tilde{L}(f)(x) = \tilde{\gamma}(x) \sum_{\sigma(y)=x} e^{\tilde{A}(y)} \left[f(y) - f(x)\right], \tag{14}$$

where $\tilde{\gamma}$ is a strictly positive continuous function, and, $\tilde{A}$ is a normalized Lipschitz potential. We point out that μ do not have to be stationary for this chain.

Notice that according to the last section all the Gibbs Markov chains $\mathbb{P}^V_{\mu_{B_V},\gamma_V}$ one gets from a generic V are admissible. If we take any μ on $\{1,\dots,d\}^{\mathbb{N}}$, and we denote by $\mathbb{P}_\mu$ the one we get when $\tilde{A} = A$ and $\tilde{\gamma} = 1$, i.e., the one we get from the unperturbed system with the initial measure μ, then $\mathbb{P}_\mu$ is also admissible.

In the same way as in (12), the stationary measure for the continuous time process with generator (14) is

$$\mathrm{d}\mu_{\tilde{A},\tilde{\gamma}}(x) = \frac{1}{\tilde{\gamma}(x)} \frac{\mathrm{d}\mu_{\tilde{A}}(x)}{\int \frac{1}{\tilde{\gamma}}\,\mathrm{d}\mu_{\tilde{A}}}, \tag{15}$$

where $\mu_{\tilde{A}}$ is discrete time equilibrium for $\tilde{A}$.

From now on, we will consider a certain Lipschitz potential V fixed until the end of this section. The different probabilities $\tilde{\mathbb{P}}^{\tilde{\gamma},\tilde{A}}_{\mu_{\tilde{A},\tilde{\gamma}}}$ on $\mathscr{D}$ will describe the possible candidates for being the *stationary equilibrium continuous time Markov chain for V* as we will explain later in our reasoning.

Given V we will consider a *variational problem in the continuous time setting* which is analogous to the pressure problem in the discrete time setting (thermodynamic formalism). This requires a meaning for *entropy*. A continuous time stationary Markov chain, which maximizes our variational problem, will be the *continuous time equilibrium state for V*. By changing $\tilde{\gamma}$ and $\tilde{A}$, we get a set of different infinitesimal generators that are candidates to define *the continuous time equilibrium state* for the given potential V. Nevertheless, it just makes sense to look for candidates among the admissible ones. We will show in the end that the continuous time equilibrium state for V is indeed the Gibbs state $\mathbb{P}^V_{\mu_{B_V},\gamma_V}$ of the last section.

We will fix a certain μ on $\mathscr{P}(\{1,\dots,d\}^{\mathbb{N}})$ (no restrictions about it). First, we want to give a meaning for the relative entropy of any admissible probability $\tilde{\mathbb{P}}_\mu$ concerning $\mathbb{P}_\mu$. The reason why we use the same initial measure μ for both processes is that we need that the associated probabilities, $\tilde{\mathbb{P}}_\mu$ and $\mathbb{P}_\mu$, on $\mathscr{D}$ are absolutely continuous with respect to each other. Anyway, the final numerical result for the value of entropy will not depend on the common μ we chose as the initial probability, as can be seen in Lemma 13. The common μ could de eventually μ_A. For a fixed $T \geq 0$, we consider the relative entropy of the $\tilde{\mathbb{P}}_\mu = \tilde{\mathbb{P}}_\mu^{\tilde{\gamma},\tilde{A}}$, for some $\tilde{\gamma}$,$\tilde{A}$, concerning $\mathbb{P}_\mu$ up to time $T \geq 0$ by

$$H_T(\tilde{\mathbb{P}}_\mu|\mathbb{P}_\mu) = -\int_{\mathscr{D}} \log\left(\frac{\mathrm{d}\tilde{\mathbb{P}}_\mu}{\mathrm{d}\mathbb{P}_\mu}\Big|_{\mathscr{F}_T}\right)(\omega)\,\mathrm{d}\tilde{\mathbb{P}}_\mu(\omega). \tag{16}$$

Using the property that the logarithm is a concave function and Jensen's inequality, we obtain that for any g we have $\int \log g\,\mathrm{d}\mu \leq \log \int g\,\mathrm{d}\mu$. Then $H_T(\tilde{\mathbb{P}}_\mu|\mathbb{P}_\mu) \leq 0$. Negative entropies appear in a natural way when one analyzes a dynamical system with the property that each point has an uncountable number of preimages (see [28] and [31]).

By Proposition 27 in Appendix C, the logarithm of the Radon-Nikodym derivative described above can be written as

$$\log\left(\frac{\mathrm{d}\tilde{\mathbb{P}}_\mu}{\mathrm{d}\mathbb{P}_\mu}\Big|_{\mathscr{F}_T}\right)(\omega)$$

$$= \int_0^T [1 - \tilde{\gamma}(\omega_s)]\,\mathrm{d}s + \sum_{s\leq T} \mathbf{1}_{\{\sigma(\omega_s)=\omega_{s^-}\}} \left[\tilde{A}(\omega_s) - A(\omega_s) + \log\left(\tilde{\gamma}(\sigma(\omega_s))\right)\right]. \tag{17}$$

Lemma 11. *For all $G \in \mathscr{C}$, it is true that*

$$\int_{\mathscr{D}} \sum_{s \le T} \mathbf{1}_{\{\sigma(\omega_s) = \omega_{s^-}\}} G(\omega_s)\, d\tilde{\mathbb{P}}_\mu(\omega) = \int_{\mathscr{D}} \int_0^T \tilde{\gamma}(\omega_s) G(\omega_s)\, ds\, d\tilde{\mathbb{P}}_\mu(\omega)$$
$$= \int_0^T \int_{\{1,\dots,d\}^{\mathbb{N}}} \tilde{P}_s(\tilde{\gamma} G)(x)\, d\mu(x)\, ds\,,$$

where $\{\tilde{P}_s,\, s \ge 0\}$ is the semigroup associated to the Markov chain that it was generated by $\tilde{L}$, see (14).

The proof of this lemma is in Appendix D.

Now, from (16), (17) and the lemma above we obtain

$$\begin{aligned} H_T(\tilde{\mathbb{P}}_\mu|\mathbb{P}_\mu) = & \int_0^T \int_{\{1,\dots,d\}^{\mathbb{N}}} \tilde{P}_s(\tilde{\gamma} - 1)(x)\, \mathrm{d}\mu(x)\, \mathrm{d}s \\ & + \int_0^T \int_{\{1,\dots,d\}^{\mathbb{N}}} \tilde{P}_s(\tilde{\gamma}[A - \tilde{A} - \log \tilde{\gamma} \circ \sigma])(x)\, \mathrm{d}\mu(x)\, \mathrm{d}s\,. \end{aligned} \tag{18}$$

From the previous expression and ergodicity we get that there exists the limit $\lim_{T\to\infty} \frac{1}{T} H_T(\tilde{\mathbb{P}}_\mu|\mathbb{P}_\mu)$.

Definition 12. *For a fixed initial probability μ on $\mathscr{P}(\{1,\dots,d\}^{\mathbb{N}})$, the limit*

$$\lim_{T\to\infty} \frac{1}{T} H_T(\tilde{\mathbb{P}}_\mu|\mathbb{P}_\mu)$$

is called the relative entropy *of the measure $\tilde{\mathbb{P}}_\mu$ concerning the measure $\mathbb{P}_\mu$ (recall that $\mathbb{P}_\mu$ is associated to the initial fixed potential A). Moreover, we denote this limit by $H(\tilde{\mathbb{P}}_\mu|\mathbb{P}_\mu)$.*

The goal of the next result is characterize the *relative entropy* of the measure $\tilde{\mathbb{P}}_\mu$ concerning $\mathbb{P}_\mu$.

Lemma 13. *The relative entropy $H(\tilde{\mathbb{P}}_\mu|\mathbb{P}_\mu)$ can be written as*

$$\begin{aligned} & \int_{\{1,\dots,d\}^{\mathbb{N}}} (\tilde{\gamma}(x) - 1)\, d\mu_{\tilde{A},\tilde{\gamma}}(x) \\ & + \int_{\{1,\dots,d\}^{\mathbb{N}}} \tilde{\gamma}(x) \left[A(x) - \tilde{A}(x) - \log(\tilde{\gamma}\circ\sigma)(x)\right] d\mu_{\tilde{A},\tilde{\gamma}}(x)\,. \end{aligned}$$

Proof. This proof follows by Definition 12, expression (18) and Ergodic Theorem. □

Definition 14. *For A fixed, and a given Lipschitz potential V, we denote the Pressure (or, Free Energy) of V as the value*

$$\boldsymbol{P}(V) := \sup_{\substack{\tilde{\mathbb{P}}_\mu \\ \text{admissible}}} H(\tilde{\mathbb{P}}_\mu|\mathbb{P}_\mu) + \int_{\{1,\dots,d\}^{\mathbb{N}}} V(x)\, d\mu_{\tilde{A},\tilde{\gamma}}(x)\,,$$

where $\mu_{\tilde{A},\tilde{\gamma}}$ is the initial stationary probability for the infinitesimal generator $\tilde{L}$, defined in (14)*. Moreover, any admissible element which maximizes $\boldsymbol{P}(V)$ is called a continuous time equilibrium state for V.*

Finally, we can state the main result of this section:

Proposition 15. *The pressure of the potential V is given by*

$$\boldsymbol{P}(V) = H(\mathbb{P}^V_\mu|\mathbb{P}_\mu) + \int_{\{1,\dots,d\}^{\mathbb{N}}} V(x)\, d\mu_{B_V,\gamma_V}(x) = \lambda_V\,.$$

Therefore, the equilibrium state for V is the Gibbs state for V.

Proof. Recalling the definition of the measure $\mu_{\tilde{A},\tilde{\gamma}}$ in (15) and the fact that the measure $\mu_{\tilde{A}}$ is invariant for the shift, we get that the second term in (18) can be rewritten as

$$\left[\int \tfrac{1}{\tilde{\gamma}}\, \mathrm{d}\mu_{\tilde{A}}\right]^{-1} \int_{\{1,\dots,d\}^{\mathbb{N}}} \big(A(x) - \tilde{A}(x)\big) \mathrm{d}\mu_{\tilde{A}}(x) - \int_{\{1,\dots,d\}^{\mathbb{N}}} \tilde{\gamma}(x)\, \log \tilde{\gamma}(x)\, \mathrm{d}\mu_{\tilde{A},\tilde{\gamma}}(x)\,.$$

Let V be a Lipschtz function. Thus,

$$\begin{aligned} &H(\tilde{\mathbb{P}}_\mu|\mathbb{P}_\mu)+\int_{\{1,\dots,d\}^{\mathbb{N}}} V(x)\,\mathrm{d}\mu_{\tilde{A},\tilde{\gamma}}(x) \\ &= \int_{\{1,\dots,d\}^{\mathbb{N}}} \big(\tilde{\gamma}(x)-\tilde{\gamma}(x)\,\log\tilde{\gamma}(x)-1+V(x)\big)\,\mathrm{d}\mu_{\tilde{A},\tilde{\gamma}}(x) \\ &\qquad + \Big[\int \tfrac{1}{\tilde{\gamma}}\,\mathrm{d}\mu_{\tilde{A}}\Big]^{-1}\int_{\{1,\dots,d\}^{\mathbb{N}}} \big(A(x)-\tilde{A}(x)\big)\mathrm{d}\mu_{\tilde{A}}(x)\,. \end{aligned}$$

From equation (7), we can express the function V as $\lambda_V+1-\gamma_V(x)$. Then the expression above becomes

$$\begin{aligned} &H(\tilde{\mathbb{P}}_\mu|\mathbb{P}_\mu)+\int_{\{1,\dots,d\}^{\mathbb{N}}} V(x)\,\mathrm{d}\mu_{\tilde{A},\tilde{\gamma}}(x) \\ &= \lambda_V + \Big[\int \tfrac{1}{\tilde{\gamma}}\,\mathrm{d}\mu_{\tilde{A}}\Big]^{-1}\int_{\{1,\dots,d\}^{\mathbb{N}}} \Big(1-\log\tilde{\gamma}(x)-\frac{\gamma_V(x)}{\tilde{\gamma}(x)}\Big)\,\mathrm{d}\mu_{\tilde{A}}(x) \\ &\qquad + \Big[\int \tfrac{1}{\tilde{\gamma}}\,\mathrm{d}\mu_{\tilde{A}}\Big]^{-1}\int_{\{1,\dots,d\}^{\mathbb{N}}} \big(A(x)-\tilde{A}(x)\big)\,\mathrm{d}\mu_{\tilde{A}}(x)\,. \end{aligned} \tag{19}$$

The last integral above is equal to $\int A\,\mathrm{d}\mu_{\tilde{A}}+h(\mu_{\tilde{A}})$. In order to analyze the second term in (19), we add and subtract $\log\gamma_V(x)$ in the integrand and we use $1+\log y-y\le 0$, for all $y\in(0,\infty)$. Thus,

$$1-\log\tilde{\gamma}(x)-\frac{\gamma_V(x)}{\tilde{\gamma}(x)} \le -\log\tfrac{\mathscr{L}_A(F)(x)}{F(x)}\,,$$

because $\gamma_V(x)=\frac{\mathscr{L}_A(F)(x)}{F(x)}$, for any $x\in\{1,\dots,d\}^{\mathbb{N}}$. This implies that

$$\begin{aligned} &H(\tilde{\mathbb{P}}_\mu|\mathbb{P}_\mu)+\int_{\{1,\dots,d\}^{\mathbb{N}}} V(x)\,\mathrm{d}\mu_{\tilde{A},\tilde{\gamma}}(x) \\ &\le \lambda_V + \Big[\int \tfrac{1}{\tilde{\gamma}}\,\mathrm{d}\mu_{\tilde{A}}\Big]^{-1}\Big[-\int_{\{1,\dots,d\}^{\mathbb{N}}} \log\tfrac{\mathscr{L}_A(F)(x)}{F(x)}\,\mathrm{d}\mu_{\tilde{A}}(x)+\int_{\{1,\dots,d\}^{\mathbb{N}}} A\,\mathrm{d}\mu_{\tilde{A}}+h(\mu_{\tilde{A}})\Big]\,. \end{aligned}$$

By [29] (see Theorem 4) and [20], we have

$$\int_{\{1,\dots,d\}^{\mathbb{N}}} A\,\mathrm{d}\mu_{\tilde{A}}+h(\mu_{\tilde{A}}) \;=\; \inf_{u\in\mathscr{C}^+}\int_{\{1,\dots,d\}^{\mathbb{N}}} \log\tfrac{\mathscr{L}_A(u)(x)}{u(x)}\,\mathrm{d}\mu_{\tilde{A}}(x)\,.$$

Since $F\in\mathscr{C}^+$ and $\int\frac{1}{\tilde{\gamma}}\,\mathrm{d}\mu_{\tilde{A}}>0$, we obtain

$$H(\tilde{\mathbb{P}}_\mu|\mathbb{P}_\mu)+\int_{\{1,\dots,d\}^{\mathbb{N}}} V(x)\,\mathrm{d}\mu_{\tilde{A},\tilde{\gamma}}(x)\;\le\;\lambda_V\,.$$

One special case is when the measure $\tilde{\mathbb{P}}_\mu$ is $\mathbb{P}^V_\mu$, i.e.,

$$\tilde{\gamma}(x)=\gamma_V(x)=1-V(x)+\lambda_V=\frac{\mathscr{L}_A(F)(x)}{F(x)},$$

and

$$\tilde{A}(x)=B_V(x)=A(x)+\log F(x)-\log\mathscr{L}_A(F)(\sigma(x)).$$

In this case, the expression (19) becomes

$$\begin{aligned} &H(\mathbb{P}^V_\mu|\mathbb{P}_\mu)+\int_{\{1,\dots,d\}^{\mathbb{N}}} V(x)\,\mathrm{d}\mu_{B_V,\gamma_V}(x) \\ &=\lambda_V + \Big[\int \tfrac{1}{\gamma_V}\,\mathrm{d}\mu_{B_V}\Big]^{-1}\int_{\{1,\dots,d\}^{\mathbb{N}}} \Big[-\log\Big(\tfrac{\mathscr{L}_A(F)(x)}{F(x)}\Big) \\ &\qquad -\log F(x)+\log\mathscr{L}_A(F)(\sigma(x))\Big]\,\mathrm{d}\mu_{B_V}(x)\,. \end{aligned}$$

Due to the fact that μ_{B_V} is an invariant measure for the shift, we finally get

$$H(\mathbb{P}^V_\mu|\mathbb{P}_\mu)+\int_{\{1,\dots,d\}^{\mathbb{N}}} V(x)\,\mathrm{d}\mu_{B_V,\gamma_V}(x)\;=\;\lambda_V\,.$$

□

5. A LARGE DEVIATION PRINCIPLE FOR THE EMPIRICAL MEASURE

Nice general references on this topic are [12] and [20]. We point out that the process we consider **is not** reversible differently from [16].

This section is divided on two subsections. The first one deals with the existence and the uniqueness of equilibrium states and the second one is about large deviation properties.

5.1. **Existence and uniqueness of equilibrium states.** As before, we considered a fixed normalized Lipschitz potential A and the corresponding infinitesimal generator $L=\mathscr{L}_A-I$. In this subsection we will assume that the perturbation V is a Lipschitz function. As we mentioned before (see Subsection 2.1), for the given potential V, one can find an eigenprobability ν_V. This means that there exists $\lambda_V=\int V\,\mathrm{d}\nu_V$ such that

$$\int(\mathscr{L}_A-I+V)(f)\,\mathrm{d}\nu_V=\lambda_V\int f\,\mathrm{d}\nu_V\,,$$

for any $f\in\mathscr{C}$. As usual, we denote $\gamma_V(x)=1-V(x)+\lambda_V$. Notice that $\int\gamma_V(x)\,\mathrm{d}\nu_V(x)=\int(1-V(x)+\lambda_V)\,\mathrm{d}\nu_V=1=\int\mathscr{L}_A(1)\,\mathrm{d}\nu_V$.

Remember that $L^V(f)(x)=\gamma_V(x)\sum_{\sigma(y)=x}e^{B_V(y)}\big[f(y)-f(x)\big]$, where

$$B_V(y)=A(y)+\log F(y)-\log F(\sigma(y))-\log(1-V(\sigma(x))+\lambda_V),$$

is the infinitesimal generator associated to a semigroup $\{\mathscr{P}^V_T,\,T\geq0\}$.

Moreover, for all $u\in\mathscr{C}$ it is true that

$$\int\mathscr{P}^V_t(u)\,\mathrm{d}\mu_{B_V,\gamma_V}=\int u\,\mathrm{d}\mu_{B_V,\gamma_V}\,,$$

where $\mathrm{d}\mu_{B_V,\gamma_V}(x)=\frac{1}{\gamma_V(x)}\frac{\mathrm{d}\mu_{B_V}(x)}{\int\frac{1}{\gamma_V}\,\mathrm{d}\mu_{B_V}}$.

Lemma 16. *Suppose $F=F_V>0$ is the main eigenfunction of the operator $\mathscr{L}_A-I+V$ with eigenvalue λ_V, then $d\tilde\nu_V(x):=\frac{1}{F(x)}d\mu_{B_V,\gamma_V}(x)=\frac{1}{F(x)}\frac{1}{\gamma_V(x)}\frac{d\mu_{B_V}(x)}{\int\frac{1}{\gamma_V}d\mu_{B_V}}$ satisfies, for all $g\in\mathscr{C}$,*

$$\int(\mathscr{L}_A-I+V)(g)\,d\tilde\nu_V=\lambda_V\int g\,d\tilde\nu_V\,.$$

Therefore, $\tilde\nu_V$ is an eigenprobability for $(\mathscr{L}_A-I+V)^*$. Moreover, if we know that the initial stationary probability for $\{\mathscr{P}^V_t=e^{tL^V},\,t\geq0\}$ is unique, then the eigenprobability is unique.

Proof. It is known that

$$\int L^V(f)\,\mathrm{d}\mu_{B_V,\gamma_V}=0,\quad\forall f\in\mathscr{C}\,.$$

We can consider an equivalent expression for $L^V(f)$, which is in (11), then for any $f\in\mathscr{C}$, we have

$$\int\tfrac{1}{F}(L+V)(Ff)\,\mathrm{d}\mu_{B_V,\gamma_V}=\lambda\int f\,\mathrm{d}\mu_{B_V,\gamma_V}\,.$$

Denote by $\tilde\nu_V$ the measure $\frac{1}{F}\mu_{B_V,\gamma_V}$. Given a $g\in\mathscr{C}$, take $f=g/F$, thus,

$$\int(L+V)(g)\,\mathrm{d}\tilde\nu_V=\lambda\int g\,\mathrm{d}\tilde\nu_V\,.$$

This shows the first claim, that is, $\tilde\nu_V$ is the eigenprobability. Suppose that the initial stationary probability, μ_{B_V,γ_V}, for $\{e^{tL^V},\,t\geq0\}$ is unique and $\tilde\nu_V$ is the eigenprobability. By hypothesis $F=F_V$ is the unique main eigenfunction for $L+V$. We can reverse the above argument for the measure $F_V\,\mathrm{d}\tilde\nu_V$. Notice that each step is an equivalence. Therefore, one can show that

$$\int L^V(f)\,F_V\,\mathrm{d}\tilde\nu_V=0,\quad\forall f\in\mathscr{C}\,.$$

From the uniqueness we assumed above, we get $\frac{d\mu_{B_V,\gamma_V}}{d\tilde\nu_V}=F_V$. The final conclusion is that if the initial stationary probability for the continuous time Markov chain associated to V satisfies $\mu_{B_V,\gamma_V}=F_V\,\nu_V$, then $\tilde\nu_V$ is unique. □

Lemma 17. *If there exists a function $F \in \mathscr{C}^+$ such that $(L+V)F = \lambda_V F$, then the functional acting on $\mathscr{P}(\{1,\dots,d\}^{\mathbb{N}})$ given by*

$$I(\nu) := -\inf_{u\in\mathscr{C}^+} \int \frac{L(u)}{u}\, d\nu \geq 0, \tag{20}$$

satisfies

$$\lambda_V = \sup_{\nu\in\mathscr{P}(\{1,\dots,d\}^{\mathbb{N}})} \Big(\int V d\nu - I(\nu)\Big).$$

The supremum value above is achieved on the probability μ_{B_V,γ_V}. Moreover, if for any Lipschitz V all the above is true, then, using the Legendre Transform, we obtain

$$I(\nu) = \sup_{V\in\mathscr{C}} \Big(\int V\, d\nu - \lambda_V\Big) = \sup_{\substack{V\in\mathscr{C}\\ \text{and } V \text{ is Lipschitz}}} \Big(\int V\, d\nu - \lambda_V\Big),$$

for all ν probability on $\{1,\dots,d\}^{\mathbb{N}}$ and $I(\nu) = \infty$ in any other case.

Proof. We follow the reasoning described in Section 4 of [20] adapted to the present case. First, we show that

$$\lambda_V \geq \sup_{\nu\in\mathscr{P}(\{1,\dots,d\}^{\mathbb{N}})} \Big(\int V \mathrm{d}\nu - I(\nu)\Big). \tag{21}$$

Let $\nu \in \mathscr{P}(\{1,\dots,d\}^{\mathbb{N}})$, by definition of the functional I, we get

$$\int V\,\mathrm{d}\nu - I(\nu) \leq \int V\,\mathrm{d}\nu + \int \frac{L(u)}{u}\,\mathrm{d}\nu\,, \quad \forall u\in\mathscr{C}^+.$$

We will take $u = F$, where F is the eigenfunction of the P_T^V, then we obtain

$$\int V\,\mathrm{d}\nu - I(\nu) \leq \int V\,\mathrm{d}\nu + \int \frac{L(F)}{F}\,\mathrm{d}\nu\,.$$

Using the equation (7), we can rewrite $\frac{L(F)}{F}$ as $-V+\lambda_V$, then the inequality follows.

Now, we will show that

$$\lambda_V \leq \sup_{\nu\in\mathscr{P}(\{1,\dots,d\}^{\mathbb{N}})} \Big(\int V \mathrm{d}\nu - I(\nu)\Big). \tag{22}$$

Actually, we will prove that

$$\lambda_V \leq \int V\,\mathrm{d}\mu_{B_V,\gamma_V} - I(\mu_{B_V,\gamma_V})\,,$$

and this implies the inequality (22).

In order to show the above, we consider a general $u\in\mathscr{C}^+$. Recalling the expression of L^V, which is in Proposition 7, we get

$$\frac{L^V(u/F)}{u/F} = \frac{L(u)}{u} + V - \lambda_V\,.$$

As the infinitesimal generator of $\mathscr{P}_t^V$ is L^V, from Corollary 8, we have for all $u\in\mathscr{C}^+$ and t small

$$\frac{L^V(u/F)}{u/F} \sim \frac{1}{t}\log\Big(\frac{\mathscr{P}_t^V(u/F)}{u/F}\Big).$$

Using these two last expressions we get for any $u\in\mathscr{C}^+$

$$\int \Big[\frac{L(u)}{u} + V - \lambda_V\Big]\,\mathrm{d}\mu_{B_V,\gamma_V} \sim \frac{1}{t}\int \log\Big(\frac{\mathscr{P}_t^V(u/F)}{u/F}\Big)\,\mathrm{d}\mu_{B_V,\gamma_V}\,.$$

By Jensen's inequality for any $u\in\mathscr{C}^+$ and small $t>0$, we have the right-hand side in the last expression is bounded from below by

$$\frac{1}{t}\int \Big[\mathscr{P}_t^V\Big(\log\big(u/F\big)\Big) - \log\big(u/F\big)\Big]\,\mathrm{d}\mu_{B_V,\gamma_V} = 0\,.$$

The last equality is due to fact that μ_{B_V,γ_V} is the invariant measure. Therefore, we take the infimum among all $u\in\mathscr{C}^+$ in the above expression, and we get

$$\inf_{u\in\mathscr{C}^+}\int \Big[\frac{L(u)}{u} + V\Big]\,\mathrm{d}\mu_{B_V,\gamma_V} \geq \lambda_V\,.$$

Thus, we finish the proof of the inequality (22). Consequently, using (21) and (22) one can conclude the statement of the lemma. The last claim follows from a standard procedure via the classical Legendre transform.

□

We point out that indeed is true that for any Lipchitz V there exist F and λ as above. Therefore, the conclusion of last result is true in our case (for the corresponding I).

In the future we will need the property that for each Lipchitz V the probability which attains the maximal value $\sup_{\nu\in\mathscr{P}(\{1,\dots,d\}^{\mathbb{N}})}\left(\int V d\nu - I(\nu)\right)$ is unique. In this direction we consider first the following lemma.

Lemma 18. *For a fixed Lipschitz V , if ρ realizes*

$$\lambda_V = \int V\, d\rho - I(\rho),$$

then $(\mathscr{P}_t^V)^*(\rho) = \rho$, for all $t \geq 0$.

Proof. By hypothesis, we get

$$\inf_{u\in\mathscr{C}^+}\int \frac{(L+V-\lambda_V)(u)}{u}\, d\rho = 0\,.$$

We have to show that for any $f\in\mathscr{C}$ it is true $\int L^V(f)\, d\rho = 0$. The inspiration for the main idea of this proof comes from the reasoning of Sections 2 and 3 in [10] (a little bit different from the last paragraph of the proof of Proposition 3.1 in [21]).

Recalling the definition of the operator L^V given in Proposition 7, and using the Corollary 6, we obtain the next equality

$$L^V(f)(x) = \frac{1}{F(x)}\sum_{\sigma(y)=x} e^{A(y)}\, F(y)\,[f(y)-f(x)] = \frac{\mathscr{L}_A(F\,f)(x)}{F(x)} - \mathscr{L}_A(F)(x)\,\frac{f(x)}{F(x)}\,. \tag{23}$$

We point out that as $F\in\mathscr{C}^+$ is the main eigenfunction of $L+V)$ with eigenvalue λ_V, then, F realizes the infimum

$$\inf_{u\in\mathscr{C}^+}\int \frac{(L+V-\lambda_V)(u)}{u}\, d\rho = 0\,.$$

Given any $f\in\mathscr{C}$, take $\varepsilon>0$ such that $\varepsilon<\frac{1}{c}$, where $\|f\|_\infty\leq c$. For this choice of ε, observe that $F\,(1+\varepsilon f)\in\mathscr{C}^+$. Denoting

$$G(\varepsilon) := \int \frac{(L+V-\lambda_V)\,(F\,(1+\varepsilon f))}{F\,(1+\varepsilon f)}\, d\rho \geq 0,$$

we note that $G(\varepsilon)$ takes its minimal value at $\varepsilon=0$. Now, taking derivative of G with respect to ε and applying to the value $\varepsilon=0$, we get from (23)

$$\begin{aligned} 0 = G'(0) &= \int \frac{(L+V-\lambda_V)\,(F\,f)\;F}{F^2} - \frac{(F\,f)\;(L+V-\lambda_V)\,(F)}{F^2}\, d\rho \\ &= \int \frac{((\mathscr{L}_A - I)+V-\lambda_V)\,(F\,f)}{F} - \frac{f\,((\mathscr{L}_A - I)+V-\lambda_V)\,(F)}{F}\, d\rho \\ &= \int [\frac{\mathscr{L}_A(F\,f)}{F} - \frac{f\,\mathscr{L}_A(F)}{F}]\, d\rho = \int L^V(f)\, d\rho\,. \end{aligned}$$

□

Uniqueness will follow from the next result.

Proposition 19. *When V is Lipschitz function, there is only one ρ which is the initial stationary probability for the stochastic semigroup $\{\mathscr{P}_t^V, t\geq 0\}$ generated by L^V.*

Proof. Suppose ρ is such that for all $t\geq 0$, we have $(\mathscr{P}_t^V)^*(\rho)=\rho$. This means that for any $f\in\mathscr{C}$, we have

$$\int \mathscr{P}_t^V(f)\, d\rho = \int f\, d\rho\,.$$

This implies that $\int L^V(f)\,\mathrm{d}\rho = 0$, for any $f \in \mathscr{C}$. Using the expression (11) for L^V, the last integral becomes

$$\int \Big[\tfrac{1}{F}(\mathscr{L}_A - I + V)(Ff) - \lambda_V f\Big]\,\mathrm{d}\rho \;=\; 0\,,$$

for any $f \in \mathscr{C}$, which is equivalent to

$$\int f\,(1 - V + \lambda_V)\,\mathrm{d}\rho \;=\; \int \frac{\mathscr{L}_A(Ff)}{F}\,\mathrm{d}\rho = \int \frac{\mathscr{L}_A(Ff)}{F}\,\frac{1}{1 - V + \lambda_V}\,(1 - V + \lambda_V)\,\mathrm{d}\rho\,.$$

We point out that it is known that $1 - V + \lambda_V$ is strictly positive. Consider $B_V(y) = A(y) - \log[1 - V(\sigma(y)) + \lambda] + \log F(y) - \log F(\sigma(y))$ and consider the following Ruelle operator

$$f \;\to\; \mathscr{L}_{B_V}(f)(x) \;=\; \sum_{\sigma(y)=x} e^{B_V(y)}\,f(y)\,,$$

which satisfies $\mathscr{L}_{B_V}(1) = 1$. From classical results in thermodynamic formalism there is a unique $\tilde{\mu}$ such that $\mathscr{L}^*_{B_V}(\tilde{\mu}) = \tilde{\mu}$. We will show that $\mathrm{d}\tilde{\mu} = (1 - V + \lambda_V)\,\mathrm{d}\rho$. Indeed, $\mathscr{L}^*_{B_V}(\tilde{\mu}) = \tilde{\mu}$ means that for any f, we have

$$\int f\,\mathrm{d}\tilde{\mu} = \int \sum_{\sigma(y)=x} e^{A(y) - \log[1 - V(\sigma(y)) + \lambda] + \log F(y) - \log F(\sigma(y))}\,f(y)\,\mathrm{d}\tilde{\mu}(x) = \int \frac{\mathscr{L}_A(Ff)}{F}\,\frac{1}{1 - V + \lambda_V}\,\mathrm{d}\tilde{\mu}\,,$$

for all $f \in \mathscr{C}$. Since $\tilde{\mu}$ is unique, we get that ρ is unique. □

The next theorem follows easily from the last two results.

Theorem 20. *For a fixed Lipschitz function V, there is a unique ρ which realizes*

$$\lambda_V = \int V\,d\rho - I(\rho)\,.$$

Moreover, $\rho = \mu_{B_V,\gamma_V}$, which is the initial stationary probability for L^V, and the measure $\mathbb{P}^V_{\mu_{B_V,\gamma_V}}$ is invariant for the continuous time semiflow $\{\Theta_t,\, t \geq 0\}$ on the Skorohod space.

We consider now some general statements that will be necessary in the next section.

In the case that there exists the eigenfunction F, it is possible to show that

$$\lim_{T\to\infty} \tfrac{1}{T}\,\log \int_{\mathscr{D}} e^{\int_0^T V(\omega_r)\,dr}\,\mathrm{d}\mathbb{P}_x(\omega) = \lambda_V\,,$$

for all $x \in \{1,\ldots,d\}^{\mathbb{N}}$. Indeed, $\log \int_{\mathscr{D}} e^{\int_0^T V(\omega_r)\,dr}\,\mathrm{d}\mathbb{P}_x(\omega)$ can be written as $\lambda_V T + \log\Big(F(x)\frac{P^V_T(1)(x)}{P^V_T(F)(x)}\Big)$. Since the eigenfunction F is strictly positive on a compact set, the second term in the last sum is bounded above and below by constants that depend only on F. This proves the desired limit.

In a similar way as above, we obtain

$$\lim_{T\to\infty} \tfrac{1}{T}\,\log \int_{\mathscr{D}} e^{\int_0^T V(\omega_r)\,dr}\,\mathrm{d}\mathbb{P}_{\mu_A}(\omega) \;=\; \lambda_V\,. \tag{24}$$

Remember that the value λ_V was obtained from V as the one such that $(\mathscr{L}_A - I + V)^*\nu_V = \lambda_V\,\nu_V$, with $\lambda_V = \int V\,\mathrm{d}\nu_V$, see the Subsection 2.1.

We consider below a general continuous potential V.

Lemma 21. *For all continuous function $V : \{1,\ldots,d\}^{\mathbb{N}} \to \mathbb{R}$, there exists the limit*

$$\lim_{T\to\infty} \tfrac{1}{T}\,\log \int\int_{\mathscr{D}} e^{\int_0^T V(\omega_r)\,dr}\,d\mathbb{P}_x(\omega)\,d\mu_A(x) \;=\; \lim_{T\to\infty} \tfrac{1}{T}\,\log \int P^V_T(1)(x)\,d\mu_A(x)\,.$$

We will denote this limit by $Q(V)$.

Proof. Notice that, for all $x \in \{1,\ldots,d\}^{\mathbb{N}}$ and $T, S \geq 0$, it is true that

$$\begin{aligned} P^V_{T+S}(1)(x) \;&=\; P^V_T\big(P^V_S(1)\big)(x) \;=\; \mathbb{E}_x\Big[e^{\int_0^T V(X_r)\,dr} P^V_S(1)(X_T)\Big] \\ &\leq \int P^V_S(1)(x)\,\mathrm{d}\mu_A(x)\;\; \mathbb{E}_x\Big[e^{\int_0^T V(X_r)\,dr}\Big] \leq \int P^V_S(1)(x)\,\mathrm{d}\mu_A(x)\;\; P^V_T(1)(x)\,, \end{aligned}$$

μ_A-a.s. in x. Then the limit in the statement of this lemma follows by subadditivity. □

The above result is related to questions raised in (4.3) in [20] and (4.2.21) in [12].

From the above we get the next lemma.

Lemma 22. *For any Lipschitz function V, we have $Q(V) = \lambda_V$.*

We will show several properties of $Q(V)$ in Appendix E. More precisely, we show that in our setting the expressions (2.1) and (2.2) in [20] are true.

5.2. **Large deviations.** In this subsection we will apply to our setting the general results stated in [20]. The purpose of this subsection is to show that the large deviation principle at the level two (see (1.1) and (1.2) in [20]) is true for the a priori *process*. We will have to show that the hypothesis of Theorem 2.1 in [20] is true in our setting. General references for large deviations are [9], [12], [13], [17], [19] and [24].

First, we will present the sequence of definitions and statements of [20] in the particular case of our setting. Recalling that $\{X_t, t \geq 0\}$ denotes the *a priori* continuous time stochastic process with infinitesimal generator $L = \mathscr{L}_A - I$ and initial probability μ_A. We denote by $\mathbb{P}_{\mu_A}$ the probability on the Skorohod space $\mathscr{D}$ associated to such stationary process. We will begin with the occupational time for $\{X_t, t \geq 0\}$. Define, for all $t \geq 0$, $\omega \in \mathscr{D}$ and for any Borel subset Γ of the $\{1,\ldots,d\}^{\mathbb{N}}$,

$$L_t^{\omega}(\Gamma) = \frac{1}{t}\int_0^t \mathbf{1}_\Gamma(X_s(\omega))\,\mathrm{d}s\,.$$

Observe that for t and ω fixed we have L_t^{ω} is a measure on $\{1,\ldots,d\}^{\mathbb{N}}$ and it is called *empirical measure*. Moreover, if we consider the canonical version of the process $\{X_t, t \geq 0\}$, we can rewrite the expression above above as

$$\int_{\{1,\ldots,d\}^{\mathbb{N}}} \mathbf{1}_\Gamma(y)\,L_t^{\omega}(dy) = \frac{1}{t}\int_0^t \mathbf{1}_\Gamma(\omega_s)\,\mathrm{d}s\,.$$

Fixing $t \geq 0$ and $\omega \in \mathscr{D}$, using the fact that L_t^{ω} is a measure on $\{1,\ldots,d\}^{\mathbb{N}}$, moreover, using the expression above and usual arguments for approximating bounded (or positive) functions, we have

$$\int_{\{1,\ldots,d\}^{\mathbb{N}}} f(y)\,L_t^{\omega}(dy) \;=\; \frac{1}{t}\int_0^t f(\omega_s)\,\mathrm{d}s\,, \tag{25}$$

for all $f : \{1,\ldots,d\}^{\mathbb{N}} \to \mathbb{R}$ bounded (or positive) mensurable function.

Finally, by the Ergodic Theorem, for any $f \in \mathscr{C}^+$, we have

$$\lim_{t\to\infty}\int_{\{1,\ldots,d\}^{\mathbb{N}}} f(y)\,L_t^{\omega}(dy) \;=\; \int_{\{1,\ldots,d\}^{\mathbb{N}}} f(y)\,\mu_A(dy)\,, \quad \mathbb{P}_{\mu_A} - \text{almost surely in } \omega\,,$$

in other words, $\lim_{t\to\infty} L_t^{\omega} = \mu_A$, $\mathbb{P}_{\mu_A}$-almost surely in ω, in the sense of weak convergence of measures. Since the measure L_t^{ω} is random, there is some deviation to this convergence. We will study now the rate of convergence. In order to do it, we will prove the large deviation principle at level two for the a priori *process* $\{X_t, t \geq 0\}$. We say there exists a large deviation principle at level two, if there exists a lower semicontinuous functional I, defined on $\mathscr{M}(\{1,\ldots,d\}^{\mathbb{N}})$, such that:

i) *for any closed set $K \subset \mathscr{M}(\{1,\ldots,d\}^{\mathbb{N}})$*

$$\limsup_{t\to\infty} \frac{1}{t}\,\log \mathbb{P}_{\mu_A}\big[L_t \in K\big] \;\leq\; -\inf_{\nu\in K} I(\nu)\,,$$

ii) *for any open set $G \subset \mathscr{M}(\{1,\ldots,d\}^{\mathbb{N}})$*

$$\liminf_{t\to\infty} \frac{1}{t}\,\log \mathbb{P}_{\mu_A}\big[L_t \in G\big] \;\geq\; -\inf_{\nu\in G} I(\nu)\,.$$

We call I the deviation function, or the rate function.

In order to prove the result above, we observe that by the equality (25), we get

$$e^{\,t\int_{\{1,\ldots,d\}^{\mathbb{N}}} f(y)\,L_t^{\omega}(dy)} \;=\; e^{\int_0^t f(\omega_s)\,\mathrm{d}s}\,,$$

for all $t \geq 0$, $\omega \in \mathscr{D}$ and $f : \{1,\ldots,d\}^{\mathbb{N}} \to \mathbb{R}$ bounded (or positive) mensurable function. Then, we integrate both sides of the equality above concerning $\mathbb{P}_x$, and we obtain

$$\int_{\mathscr{D}} e^{\,t\int_{\{1,\ldots,d\}^{\mathbb{N}}} f(y)\,L_t^{\omega}(dy)}\,\mathrm{d}\mathbb{P}_x(\omega) \;=\; \int_{\mathscr{D}} e^{\int_0^t f(\omega_s)\,\mathrm{d}s}\,\mathrm{d}\mathbb{P}_x(\omega)\,,$$

for all $t \geq 0$ and $f : \{1,\dots,d\}^{\mathbb{N}} \to \mathbb{R}$ bounded (or positive) mensurable function. We recall that $\mathbb{P}_x$ is a probability on $\mathscr{D}$ induced by the initial measure δ_x and the Markov process $\{X_t; t \geq 0\}$.

Using Lemma 22, and (24) in the previous section, and the last fact, we have

$$Q(V) = \lambda_V = \lim_{T\to\infty} \frac{1}{T} \log \int_{\mathscr{D}} e^{\int_0^T V(\omega_r)\,dr}\, \mathrm{d}\mathbb{P}_{\mu_A}(\omega) = \lim_{T\to\infty} \frac{1}{T} \log \int_{\mathscr{D}} e^{\,T\int_{\{1,\dots,d\}^{\mathbb{N}}} V(y) L_T^{\omega}(dy)}\, \mathrm{d}\mathbb{P}_{\mu_A}(\omega).$$

This shows that $Q(V)$ is the same one given in (1.3) of [20], then this will allow us to find the functional rate I. From the general setting of [20] (there is no mention of eigenvalue in the below expression), we get

$$0 \leq I(\nu) = \sup_{V\in\mathscr{C}} \Big(\int V\,\mathrm{d}\nu - Q(V) \Big) = \sup_{\substack{V\in\mathscr{C} \\ \text{and } V \text{ is Lipschitz}}} \Big(\int V\,\mathrm{d}\nu - Q(V) \Big),$$

for any ν on $\mathscr{P}(\{1,\dots,d\}^{\mathbb{N}})$ and $I(\mu) = \infty$ for all other $\mu \in \mathscr{M}(\{1,\dots,d\}^{\mathbb{N}})$. We point out that the above expression for I is in agreement with the one in Lemma 17 by Lemma 22. Since the dual space of $\mathscr{C}$ is the space $\mathscr{M}(\{1,\dots,d\}^{\mathbb{N}})$, we have

$$Q(V) = \sup_{\mu\in\mathscr{P}(\{1,\dots,d\}^{\mathbb{N}})} \Big(\int V\,\mathrm{d}\mu - I(\mu) \Big).$$

Following [20] we say that $\mu_V \in \mathscr{P}(\{1,\dots,d\}^{\mathbb{N}})$ is an *equilibrium state for* V, if

$$Q(V) = \int V\,\mathrm{d}\mu_V - I(\mu_V).$$

A major result in the theory is Theorem 2.1 in [20]. We will state a particular version of this result in Theorem 23.

Theorem 23. *If for each Lipschitz function $V : \{1,\dots,d\}^{\mathbb{N}} \to \mathbb{R}$ the equilibrium state μ_V is unique, then, the large deviation principle at level two is true with the deviation function*

$$I(\nu) = \sup_{V\in\mathscr{C}} \Big(\int V\,d\nu - Q(V) \Big).$$

From (24) we get the upper bound estimate for I and from Theorem 20 (uniqueness) we get the lower bound estimate. Then, we can state one of our main results (Theorem A in the Introduction):

Theorem 24. *Let $\{X_t, t \geq 0\}$ be the* a priori *process, then the large deviation principle at level two is true for our setting with the deviation function I given by*

$$I(\nu) = \sup_{V\in\mathscr{C}} \Big(\int V\,d\nu - Q(V) \Big),$$

for any probability ν on $\{1,\dots,d\}^{\mathbb{N}}$ and $I(\nu) = \infty$ in any other case.

We point out that Lemma 17 characterizes the equilibrium state in our setting. We can state a major result due to Y. Kifer which follows by the reasoning of Section 4 in [20]. This was adapted from the original claim.

Theorem 25. *If for each Lipschitz function $V : \{1,\dots,d\}^{\mathbb{N}} \to \mathbb{R}$ there exists a positive eigenfunction for the associated continuous time Ruelle operator, then, the deviation function I is also given by*

$$I(\nu) = -\inf_{u\in\mathscr{C}^+} \int \frac{L(u)}{u}\,d\nu.$$

It follows from last subsection (see Lemma 17) that the above expression is true in our setting. In this way our description of the Large Deviation Principle at level two is completed. We refer the reader to Lemma 17 for explicit expressions related to the above result.

We point out that the above Theorem 25 in [20] (see also [21]) is presented in a different setting: the state space is a Riemannian manifold and it is considered a certain class of differential operators as infinitesimal generators. We do not consider such differentiable structure. However, from last section we were able to adapt such reasoning to our setting.

APPENDIX A. THE SPECTRUM OF $\mathscr{L}_A - I + V$ ON $\mathbb{L}^2(\mu_A)$ AND DIRICHLET FORM.

For any $f \in \mathbb{L}^2(\mu_A)$ the Dirichlet form of f is

$$\mathscr{E}_A(f,f) := \langle (I - \mathscr{L}_A)(f), f \rangle_{\mu_A}.$$

Notice that

$$\mathscr{E}_A(f,f) = \frac{1}{2} \int \sum_{\sigma(y)=x} e^{A(y)} [f(x) - f(y)]^2 \, \mathrm{d}\mu_A(x) \geq 0. \tag{26}$$

Indeed,

$$\langle (I - \mathscr{L}_A)(f), f \rangle_{\mu_A} = \int \sum_{\sigma(y)=x} e^{A(y)} [f(x) - f(y)] \, f(x) \, \mathrm{d}\mu_A(x).$$

By the other hand,

$$\begin{aligned}\langle (I - \mathscr{L}_A)(f), f \rangle_{\mu_A} &= \langle f, f \rangle_{\mu_A} - \langle \mathscr{L}_A(f), f \rangle_{\mu_A} = \int [\mathscr{L}_A(f^2) - \mathscr{L}_A(f) f] \, \mathrm{d}\mu_A \\ &= \int \Big\{ \sum_{\sigma(y)=x} e^{A(y)} [f(y) - f(x)] \, f(y) \Big\} \mathrm{d}\mu_A(x).\end{aligned}$$

These two equalities imply that

$$\langle (I - \mathscr{L}_A)(f), f \rangle_{\mu_A} = \frac{1}{2} \int \sum_{\sigma(y)=x} e^{A(y)} [f(x) - f(y)]^2 \, \mathrm{d}\mu_A(x).$$

From expression (26) we have that $\mathscr{E}_A(f,f) = 0$ implies $f = 0$.

We point out that we will consider bellow eigenvalues in $\mathbb{L}^2(\mu_A)$ which are not necessarily Lipschtiz.

Dirichlet forms are quite important (see [22]), among other reasons, because they are particulary useful when there is an spectral gap. However, this will not be the case here.

Proposition 26. *Let a Lipschitz function $V : \{1, \ldots, d\} \to \mathbb{R}$ such that $\sup V - \inf V < 2$. There are eigenvalues c for $\mathscr{L}_A - I + V$ in $\mathbb{L}^2(\mu_A)$ such that $\big[(\sup V - 2) \vee 0\big] < c < \inf V$. Each eigenvalue has infinite multiplicity. Therefore, in this case, there is no spectral gap.*

Proof. The existence of positive eigenvalues c for the operator $\mathscr{L}_A - I + V$ satisfying $\big[(\sup V - 2) \vee 0\big] < c < \inf V$ will obtained from solving the twisted cohomological equation. In order to simplify the reasoning we will present the proof for the case $E = \{0,1\}^{\mathbb{N}}$. From section 2.2 in [5], we know that given functions $z : E \to \mathbb{R}$ and $C : E \to \mathbb{R}$ one can solve in α the twisted cohomological equation

$$\frac{z(y)}{C(y)} = \frac{1}{C(y)} \alpha(y) - \alpha(\sigma(y)), \tag{27}$$

in the case that $|C| < 1$. Indeed, just take

$$\alpha(y) = \sum_{j=0}^{\infty} \frac{\frac{z(\sigma^j(y))}{C(\sigma^j(y))}}{(C(y) C(\sigma(y)) \ldots C(\sigma^j(y)))^{-1}}.$$

Note that this function α is measurable and bounded but not Lipschitz.

Take $z(y) = (-1)^{y_0} e^{-A(y)}$, when $y = (y_0, y_1, y_2, \ldots)$. Now, for $c \in \big([(\sup V - 2) \vee 0], \inf V\big)$ fixed, consider $C(y) = 1 - V(\sigma(y)) + c$. Notice that $|C| < 1$. Then, the equation (27) becomes

$$(-1)^{y_0} = e^{A(y)} \Big\{ \alpha(y) - \alpha(\sigma(y)) \big(1 - V(\sigma(y)) + c\big) \Big\}.$$

Let $x \in \{1, \ldots, d\}^{\mathbb{N}}$. Adding the equations above when $y = 0x$ and when $y = 1x$, we get

$$(\mathscr{L}_A - I + V)(\alpha)(x) = c\alpha(x),$$

because $\sigma(0x) = x = \sigma(1x)$, and the potential A is normalized.

Is is also easy to show that changing a little bit the argument one can get an infinite dimensional set of possible α associated to the same eigenvalue.

□

APPENDIX B. BASIC TOOLS FOR CONTINUOUS TIME MARKOV CHAINS

In this section we present the proofs of the Lemma 3 and Lemma 5. In order to do that, we will present another way to analyze the properties of a continuous time Markov chain.

Suppose the process $\{X_t, t \geq 0\}$ is a continuous time Markov chain. In an alternative way we can described it by considering its skeleton chain (see [27] [34]). Let $\{\xi_n\}_{n\in\mathbb{N}}$ be a discrete time Markov chain with transition probability given by $p(x,y) = \mathbf{1}_{[\sigma(y)=x]}e^{A(y)}$. Consider a sequence of random variables $\{\tau_n\}_{n\in\mathbb{N}}$, which are independent and identically distributed according to an exponential law of parameter 1. For $n \geq 0$, define

$$T_0 = 0\,, \qquad T_{n+1} = T_n + \tau_n = \tau_0 + \tau_1 + \cdots + \tau_n\,.$$

Thus, X_t can be rewritten as $\sum_{n=0}^{+\infty} \xi_n \mathbf{1}_{[T_n \leq t < T_{n+1}]}$, for all $t \geq 0$.

Proof of Lemma 3. Using the above, we are able to describe expression (1) in a different way:

$$\begin{aligned}
P_T^V(f)(x) &= \mathbb{E}_x\big[e^{\int_0^T V(X_r)\,dr} f(X_T)\big] = \sum_{n=0}^{+\infty} \mathbb{E}_x\big[e^{\int_0^T V(X_r)\,dr} f(X_T)\mathbf{1}_{[T_n\leq T<T_{n+1}]}\big]\\
&= \sum_{n=0}^{+\infty} \mathbb{E}_x\big[e^{T_1V(\xi_0)+(T_2-T_1)V(\xi_1)+\cdots+(T_n-T_{n-1})V(\xi_{n-1})+(T-T_n)V(\xi_n)} f(\xi_n)\mathbf{1}_{[T_n\leq T<T_{n+1}]}\big]\\
&= \sum_{n=0}^{+\infty} \mathbb{E}_x\big[e^{\tau_0V(\xi_0)+\tau_1V(\xi_1)+\cdots+\tau_{n-1}V(\xi_{n-1})+(T-\sum_{i=0}^{n-1}\tau_i)V(\xi_n)} f(\xi_n)\mathbf{1}_{[\sum_{i=0}^{n-1}\tau_i\leq T<\sum_{i=0}^{n}\tau_i]}\big]\\
&= \mathbb{E}_x\big[e^{TV(\xi_0)} f(\xi_0)\mathbf{1}_{[T<\tau_0]}\big] +\\
&\sum_{n=1}^{+\infty}\sum_{a_1=1}^{d}\cdots\sum_{a_n=1}^{d} \mathbb{E}_x\big[e^{\tau_0V(\xi_0)+\cdots+(T-\sum_{i=0}^{n-1}\tau_i)V(\xi_n)} f(\xi_n)\mathbf{1}_{[\sum_{i=0}^{n-1}\tau_i\leq T<\sum_{i=0}^{n}\tau_i]}\mathbf{1}_{[\xi_1=a_1x,\ldots,\xi_n=a_n\ldots a_1x]}\big]\,,
\end{aligned}$$

where $\sigma^n(a_n\ldots a_1x) = x$. The first term above is equal to $e^{TV(x)}f(x)e^{-T}$. The summand in the second one is equal to

$$\begin{aligned}
&\mathbb{E}_x\Big[e^{\tau_0V(\xi_0)+\cdots+(T-\sum_{i=0}^{n-1}\tau_i)V(\xi_n)} f(\xi_n)\mathbf{1}_{[\sum_{i=0}^{n-1}\tau_i\leq T<\sum_{i=0}^{n}\tau_i]}\Big|\,\xi_1 = a_1x,\ldots,\xi_n = a_n\ldots a_1x\Big]\cdot\\
&\quad\cdot\,\mathbb{P}_x\big[\xi_1 = a_1x,\ldots,\xi_n = a_n\ldots a_1x\big]\,.
\end{aligned}$$

Using the transition probability of the Markov chain $\{\xi_n\}_n$, we get

$$\mathbb{P}_x\big[\xi_1 = a_1x,\ldots,\xi_n = a_n\ldots a_1x\big] = e^{A(a_1x)}\ldots e^{A(a_n\ldots a_1x)}\,.$$

Recalling that the random variables $\{\tau_i\}$ are independent and identically distributed according to an exponential law of parameter 1, we have

$$\begin{aligned}
&\mathbb{E}_x\Big[e^{\tau_0V(\xi_0)+\cdots+(T-\sum_{i=0}^{n-1}\tau_i)V(\xi_n)} f(\xi_n)\mathbf{1}_{[\sum_{i=0}^{n-1}\tau_i\leq T<\sum_{i=0}^{n}\tau_i]}\Big|\,\xi_1 = a_1x,\ldots,\xi_n = a_n\ldots a_1x\Big]\\
&= \mathbb{E}_x\Big[e^{\tau_0V(x)+\cdots+(T-\sum_{i=0}^{n-1}\tau_i)V(a_n\ldots a_1x)} f(a_n\ldots a_1x)\mathbf{1}_{[\sum_{i=0}^{n-1}\tau_i\leq T<\sum_{i=0}^{n}\tau_i]}\Big]\\
&= f(a_n\ldots a_1x)\int_0^\infty dt_n\ldots\int_0^\infty dt_0\, e^{t_0V(x)+\cdots+(T-\sum_{i=0}^{n-1}t_i)V(a_n\ldots a_1x)}\mathbf{1}_{[\sum_{i=0}^{n-1}t_i\leq T<\sum_{i=0}^{n}t_i]}e^{-t_0}\ldots e^{-t_n}\,.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&P_T^V(f)(x) = \mathbb{E}_x\big[e^{\int_0^T V(X_r)\,dr} f(X_T)\big] = e^{TV(x)}f(x)e^{-T} +\\
&\sum_{n=1}^{+\infty}\sum_{a_1=1}^{d}\cdots\sum_{a_n=1}^{d} e^{A(a_1x)}\ldots e^{A(a_n\ldots a_1x)} f(a_n\ldots a_1x)\cdot\\
&\int_0^\infty dt_n\ldots\int_0^\infty dt_0\, e^{t_0V(x)+\cdots+(T-\sum_{i=0}^{n-1}t_i)V(a_n\ldots a_1x)}\mathbf{1}_{[\sum_{i=0}^{n-1}t_i\leq T<\sum_{i=0}^{n}t_i]}e^{-t_0}\ldots e^{-t_n}\,.
\end{aligned}$$

□

Proof of Lemma 5. We begin analyzing

$$\begin{aligned}
&\mathscr{I}_V^T(a_n\dots a_1x)\\
&=\int_0^\infty dt_n\dots\int_0^\infty dt_0\; e^{t_0V(x)+\dots+(T-\sum_{i=0}^{n-1}t_i)V(a_n\dots a_1x)}\mathbf{1}_{[\sum_{i=0}^{n-1}t_i\le T<\sum_{i=0}^{n}t_i]}e^{-t_0}\dots e^{-t_n}\\
&\le e^{TC_Vd(x,y)+TC_Vd(a_1x,a_1y)+\dots+TC_Vd(a_n\dots a_1x,a_n\dots a_1y)}\cdot\\
&\cdot\int_0^\infty dt_n\dots\int_0^\infty dt_0\; e^{t_0V(y)+\dots+(T-\sum_{i=0}^{n-1}t_i)V(a_n\dots a_1y)}\mathbf{1}_{[\sum_{i=0}^{n-1}t_i\le T<\sum_{i=0}^{n}t_i]}e^{-t_0}\dots e^{-t_n}\\
&\le e^{TC_V(1+\theta+\dots+\theta^n)d(x,y)}\cdot\\
&\cdot\int_0^\infty dt_n\dots\int_0^\infty dt_0\; e^{t_0V(y)+\dots+(T-\sum_{i=0}^{n-1}t_i)V(a_n\dots a_1y)}\mathbf{1}_{[\sum_{i=0}^{n-1}t_i\le T<\sum_{i=0}^{n}t_i]}e^{-t_0}\dots e^{-t_n}\\
&\le e^{TC_V(1-\theta)^{-1}d(x,y)}\mathscr{I}_V^T(a_n\dots a_1y)
\end{aligned}\tag{28}$$

and $e^{TV(x)}e^{-T}\le e^{TC_Vd(x,y)}e^{TV(y)}e^{-T}$. Since the potential A is also Lipschitz, we get

$$\begin{aligned}
e^{A(a_1x)}\dots e^{A(a_n\dots a_1x)}&\le e^{C_A(\theta+\dots+\theta^n)d(x,y)}e^{A(a_1y)}\dots e^{A(a_n\dots a_1y)}\\
&\le e^{C_A\theta(1-\theta)^{-1}d(x,y)}e^{A(a_1y)}\dots e^{A(a_n\dots a_1y)}.
\end{aligned}\tag{29}$$

By the hypothesis we assume for f, we get

$$f(a_n\dots a_1x)\le e^{C_f\theta^n d(x,y)}f(a_n\dots a_1y)\le e^{C_f\theta d(x,y)}f(a_n\dots a_1y).$$

Thus,

$$\begin{aligned}
P_T^V(f)(x)&=e^{TV(x)}e^{-T}+\sum_{n=1}^{+\infty}\sum_{a_1=1}^{d}\dots\sum_{a_n=1}^{d}e^{A(a_1x)}\dots e^{A(a_n\dots a_1x)}f(a_n\dots a_1x)\mathscr{I}_V^T(a_n\dots a_1x)\\
&\le e^{TC_Vd(x,y)}e^{TV(y)}e^{-T}\\
&+e^{[(C_A\theta+TC_V)(1-\theta)^{-1}+C_f\theta]d(x,y)}\sum_{n=1}^{+\infty}\sum_{a_1=1}^{d}\dots\sum_{a_n=1}^{d}e^{A(a_1y)}\dots e^{A(a_n\dots a_1y)}f(a_n\dots a_1y)\mathscr{I}_V^T(a_n\dots a_1y)\\
&\le e^{[(C_A\theta+TC_V)(1-\theta)^{-1}+C_f\theta]d(x,y)}\Big[e^{TV(y)}e^{-T}\\
&+\sum_{n=1}^{+\infty}\sum_{a_1=1}^{d}\dots\sum_{a_n=1}^{d}e^{A(a_1y)}\dots e^{A(a_n\dots a_1y)}f(a_n\dots a_1y)\mathscr{I}_V^T(a_n\dots a_1y)\Big]\\
&\le e^{[(C_A\theta+TC_V)(1-\theta)^{-1}+C_f\theta]d(x,y)}P_T^V(f)(y).
\end{aligned}$$

□

APPENDIX C. RADON-NIKODYM DERIVATIVE

Let $\{\mathscr{F}_T, T\ge 0\}$ be the natural filtration.

Proposition 27. *The Radon-Nikodim derivative of the measure $\mathbb{P}_\mu$ (associated to the* a priori *process) concerning the admissible measure $\tilde{\mathbb{P}}_\mu$ (see Definition 10) restricted to $\mathscr{F}_T$ is*

$$\frac{d\mathbb{P}_\mu}{d\tilde{\mathbb{P}}_\mu}\Big|_{\mathscr{F}_T}=\exp\left\{\int_0^T(\tilde{\gamma}(X_s)-1)\,ds+\sum_{s\le T}\mathbf{1}_{[\sigma(X_s)=X_{s^-}]}\Big(A(X_s)-\tilde{A}(X_s)-\log\big(\tilde{\gamma}(\sigma(X_s))\big)\Big)\right\}.$$

Proof. The probabilities $\tilde{\mathbb{P}}_\mu$ and $\mathbb{P}_\mu$ on $\mathscr{D}$ are equivalent, because the initial measure and the allowed jumps are the same. Thus, the expectation under $\mathbb{E}_\mu$ of all bounded function $\psi:\mathscr{D}\to\mathbb{R}$, $\mathscr{F}_T$-measurable, is

$$\tilde{\mathbb{E}}_\mu\left[\psi\frac{\mathrm{d}\mathbb{P}_\mu}{\mathrm{d}\tilde{\mathbb{P}}_\mu}\Big|_{\mathscr{F}_T}\right].$$

The goal here is to obtain a formula for the Radon-Nikodim derivative $\frac{\mathrm{d}\mathbb{P}_\mu}{\mathrm{d}\tilde{\mathbb{P}}_\mu}$. Since every bounded $\mathscr{F}_T$-measurable function can be approximated by functions depending only on a finite number of coordinates, then, it is enough to work with these functions. For $k\ge 1$, consider a sequence of times $0\le t_1<\dots<t_k\le T$

and a bounded function $F:\left(\{1,\dots,d\}^{\mathbb{N}}\right)^k \to \mathbb{R}$. Using the skeleton chain, presented in the proof of Lemma 3, we get

$$\mathbb{E}_\mu[F(X_{t_1},\dots,X_{t_k})] = \sum_{n\geq 0} \mathbb{E}_\mu\left[F(X_{t_1},\dots,X_{t_k})\mathbf{1}_{[T_n\leq T<T_{n+1}]}\right].$$

Since $F(X_{t_1},\dots,X_{t_k})$ restricted to the set $[T_n \leq T < T_{n+1}]$ depends only on $\xi_1, T_1, \dots, \xi_n, T_n$, there exist functions $\bar{F}_n$ such that

$$\mathbb{E}_\mu[F(X_{t_1},\dots,X_{t_k})] = \sum_{n\geq 0} \mathbb{E}_\mu\left[\bar{F}_n(\xi_1,T_1,\dots,\xi_n,T_n)\mathbf{1}_{[T_n\leq T<T_{n+1}]}\right].$$

Through some calculations that are similar to the one used on the Corollary 2.2 in Appendix 1 of the [22], the last probability is equal to

$$\sum_{n\geq 0} \mathbb{E}_\mu\left[\bar{F}_n(\xi_1,T_1,\dots,\xi_n,T_n)\,\mathbf{1}_{[T_n\leq T]}\; e^{-\lambda(\xi_n)(T-T_n)}\right]. \tag{30}$$

Then, we need to estimate for each $n \in \mathbb{N}$ and, moreover, for all bounded measurable function $G:\left(\{1,\dots,d\}^{\mathbb{N}}\times(0,\infty)\right)^n \to \mathbb{R}$ the expectation

$$\mathbb{E}_\mu\big[G(\xi_1,T_1,\dots,\xi_n,T_n)\big] = \int_{\{1,\dots,d\}^{\mathbb{N}}} \mathbb{E}_x\big[G(\xi_1,T_1,\dots,\xi_n,T_n)\big]\,\mathrm{d}\mu(x).$$

Notice that, for all $x \in \{1,\dots,d\}^{\mathbb{N}}$,

$$\begin{aligned}
\mathbb{E}_x\big[G(\xi_1,T_1,\dots,\xi_n,T_n)\big] &= \sum_{a_1=1}^{d}\cdots\sum_{a_n=1}^{d} e^{A(a_1x)}\dots e^{A(a_n\dots a_1x)} \\
&\quad\cdot\left\{\int_0^\infty dt_{n-1}\dots\int_0^\infty dt_0\; e^{-t_0}\dots e^{-t_{n-1}}\,G(a_1x,t_0,\cdots,a_n\dots a_1x,t_{n-1}+\cdots+t_0)\right\} \\
= \sum_{a_1=1}^{d}\cdots\sum_{a_n=1}^{d} & e^{\tilde{A}(a_1x)}\dots e^{\tilde{A}(a_n\dots a_1x)}\left\{\int_0^\infty dt_{n-1}\dots\int_0^\infty dt_0\;\tilde{\gamma}(x)e^{-\tilde{\gamma}(x)t_0}\dots\tilde{\gamma}(a_{n-1}\dots a_1x)e^{-\tilde{\gamma}(a_{n-1}\dots a_1x)t_{n-1}}\right. \\
&\quad\cdot\; e^{A(a_1x)-\tilde{A}(a_1x)}\dots e^{A(a_n\dots a_1x)-\tilde{A}(a_n\dots a_1x)}\;\frac{e^{(\tilde{\gamma}(x)-1)t_0}}{\tilde{\gamma}(x)}\cdots\frac{e^{(\tilde{\gamma}(a_{n-1}\dots a_1x)-1)t_{n-1}}}{\tilde{\gamma}(a_{n-1}\dots a_1x)} \\
&\quad\left.\cdot\; G(a_1x,t_0,\cdots,a_n\dots a_1x,t_{n-1}+\cdots+t_0)\right\} \\
= \tilde{\mathbb{E}}_x\Bigg[G(\xi_1,T_1,\dots,\xi_n,T_n)\,&\exp\left\{\sum_{i=0}^{n-1}(\tilde{\gamma}(\xi_i)-1)\tau_i\right\}\prod_{i=0}^{n-1} e^{A(\xi_{i+1})-\tilde{A}(\xi_{i+1})}\,\frac{1}{\tilde{\gamma}(\xi_i)}\Bigg].
\end{aligned}$$

We can write $\sum_{i=0}^{n-1}(\tilde{\gamma}(\xi_i)-1)\tau_i$ as

$$\sum_{i=0}^{n-1}(\tilde{\gamma}(\xi_i)-1)\int_0^{T_n}\mathbf{1}_{[T_i\leq s<T_{i+1}]}\,\mathrm{d}s = \int_0^{T_n}\sum_{i=0}^{\infty}(\tilde{\gamma}(\xi_i)-1)\,\mathbf{1}_{[T_i\leq s<T_{i+1}]}\,\mathrm{d}s = \int_0^{T_n}(\tilde{\gamma}(X_s)-1)\,\mathrm{d}s,$$

and, we can write $e^{A(\xi_{i+1})-\tilde{A}(\xi_{i+1})}\,\frac{1}{\tilde{\gamma}(\xi_i)}$ as

$$\begin{aligned}
&\exp\left\{\sum_{i=0}^{n-1}(A(\xi_{i+1})-\tilde{A}(\xi_{i+1})-\log\tilde{\gamma}(\xi_i))\right\} \\
=&\exp\left\{\sum_{i=0}^{n-1}\mathbf{1}_{[\sigma(\xi_{i+1})=\xi_i]}\big(A(\xi_{i+1})-\tilde{A}(\xi_{i+1})-\log\tilde{\gamma}(\sigma(\xi_{i+1}))\big)\right\} \\
=&\exp\left\{\sum_{s\leq T_n}\mathbf{1}_{[\sigma(X_s)=X_{s^-}]}\Big(A(X_s)-\tilde{A}(X_s)-\log\big(\tilde{\gamma}(\sigma(X_s))\big)\Big)\right\}.
\end{aligned}$$

The expectation under $\mathbb{P}_x$ of $G(\xi_1,T_1,\dots,\xi_n,T_n)$ becomes

$$\tilde{\mathbb{E}}_x\left[G(\xi_1,T_1,\dots,\xi_n,T_n)\,\exp\left\{\int_0^{T_n}(\tilde{\gamma}(X_s)-1)\,\mathrm{d}s+\sum_{s\leq T_n}\mathbf{1}_{[\sigma(X_s)=X_{s^-}]}\Big(A(X_s)-\tilde{A}(X_s)-\log\big(\tilde{\gamma}(\sigma(X_s))\big)\Big)\right\}\right].$$

Using the formula above in the equation (30), the expectation under $\mathbb{E}_\mu$ of $F(X_{t_1},\dots,X_{t_k})$ is equal to

$$\sum_{n\geq 0}\tilde{\mathbb{E}}_\mu\Bigg[\bar{F}_n(\xi_1,T_1,\dots,\xi_n,T_n)\,\mathbf{1}_{[T_n\leq T]}\,e^{-\lambda(\xi_n)(T-T_n)}$$
$$\cdot\,\exp\Big\{\int_0^{T_n}(\tilde{\gamma}(X_s)-1)\,\mathrm{d}s+\sum_{s\leq T_n}\mathbf{1}_{[\sigma(X_s)=X_{s^-}]}\Big(A(X_s)-\tilde{A}(X_s)-\log\big(\tilde{\gamma}(\sigma(X_s))\big)\Big)\Big\}\Bigg].$$

Once again, we use some calculations similarly to the Corollary 2.2 in Appendix 1 of the [22] and we rewrite the expression above as

$$\sum_{n\geq 0}\tilde{\mathbb{E}}_\mu\Bigg[\bar{F}_n(\xi_1,T_1,\dots,\xi_n,T_n)\,\mathbf{1}_{[T_n\leq T<T_{n+1}]}$$
$$\cdot\,\exp\Big\{\int_0^{T}(\tilde{\gamma}(X_s)-1)\,\mathrm{d}s+\sum_{s\leq T}\mathbf{1}_{[\sigma(X_s)=X_{s^-}]}\Big(A(X_s)-\tilde{A}(X_s)-\log\big(\tilde{\gamma}(\sigma(X_s))\big)\Big)\Big\}\Bigg],$$

and, this sum is equal to

$$\tilde{\mathbb{E}}_\mu\Big[F(X_{t_1},\dots,X_{t_k})\exp\Big\{\int_0^{T}(\tilde{\gamma}(X_s)-1)\,\mathrm{d}s+\sum_{s\leq T}\mathbf{1}_{[\sigma(X_s)=X_{s^-}]}\Big(A(X_s)-\tilde{A}(X_s)-\log\big(\tilde{\gamma}(\sigma(X_s))\big)\Big)\Big\}\Big].$$

This finish the proof.

□

Appendix D. Proof of Lemma 11

Proof of Lemma 11. We claim that

$$M_T^G(\omega)=\sum_{s\leq T}\mathbf{1}_{\{\sigma(\omega_s)=\omega_{s^-}\}}G(\omega_s)-\int_0^T\tilde{\gamma}(\omega_s)G(\omega_s)\,\mathrm{d}s$$

is a $\tilde{\mathbb{P}}_\mu$ - martingale. Then, this lemma will follow from $\tilde{\mathbb{E}}_\mu\big[M_T^G\big]=\tilde{\mathbb{E}}_\mu\big[M_0^G\big]=0$. In order to prove this claim it is enough to prove that

$$M_T(\omega)=\sum_{s\leq T}\mathbf{1}_{\{\sigma(\omega_s)=\omega_{s^-}\}}-\int_0^T\tilde{\gamma}(\omega_s)\,\mathrm{d}s \tag{31}$$

is a $\tilde{\mathbb{P}}_\mu$ - martingale, because $M_T^G=\int G\,\mathrm{d}M_T$ will be a $\tilde{\mathbb{P}}_\mu$ - martingale (see [36]).

Now, we prove (31). Let $\{\mathscr{F}_T,T\geq 0\}$ be the natural filtration. For all $S<T$, we prove that $\tilde{\mathbb{E}}_\mu\big[M_T-M_S|\mathscr{F}_S\big]=0$. By Markov property, we only need to show that $\tilde{\mathbb{E}}_x\big[M_t\big]=0$.

Denote by $\mathscr{D}_x$ the space of all trajectories ω in $\mathscr{D}$ such that $\omega_0=x$. Observe that, for all ω in $\mathscr{D}_x$,

$$\int_0^t\tilde{\gamma}(\omega_s)\,\mathrm{d}s=\sum_{k\geq 1}\sum_{i_1=1}^{d}\cdots\sum_{i_k=1}^{d}\tilde{\gamma}(i_k\dots i_1x)\int_0^t\mathbf{1}_{[\omega_s=i_k\dots i_1x]}\,\mathrm{d}s. \tag{32}$$

For all $s\geq 0$ and $y\in\{1,\dots,d\}^{\mathbb{N}}$, $N_s(y)$ denotes the number of times that the exponential clock rang at site y. Thus, the first term on the right side of (31) can be rewritten as

$$\sum_{s\leq t}\mathbf{1}_{\{\sigma(\omega_s)=\omega_{s^-}\}}=\sum_{k\geq 1}\sum_{i_1=1}^{d}\cdots\sum_{i_k=1}^{d}N_t(i_k\dots i_1x), \tag{33}$$

for all ω in $\mathscr{D}_x$.

Since (32) and (33) are true, in order to conclude this prove, it is sufficient to show that

$$\tilde{\mathbb{E}}_x\big[N_t(y)-\tilde{\gamma}(y)\int_0^t\mathbf{1}_{[X_s=y]}\,\mathrm{d}s\big]=0, \tag{34}$$

for all $y\in\{1,\dots,d\}^{\mathbb{N}}$.

Let $0=t_0<t_1<\cdots<t_n=t$ be a partition of the interval $[0,t]$. The expression (34) can be rewritten as

$$\sum_{i=0}^{n-1}\tilde{\mathbb{E}}_x\Big[N_{t_{i+1}}(y)-N_{t_i}(y)+\tilde{\gamma}(y)\int_{t_i}^{t_{i+1}}\mathbf{1}_{[X_s=y]}\,\mathrm{d}s\Big].$$

Observe that

$$\begin{aligned}
&\tilde{\mathbb{E}}_x\Big[\int_{t_i}^{t_{i+1}} \mathbf{1}_{[X_s=y]}\,\mathrm{d}s\Big] = \tilde{\mathbb{E}}_y\Big[\int_0^{t_{i+1}-t_i} \mathbf{1}_{[X_s=y]}\,\mathrm{d}s\Big]\\
=&\tilde{\mathbb{E}}_y\Big[\int_0^{t_{i+1}-t_i} \mathbf{1}_{[X_s=y]}\,\mathrm{d}s\,\mathbf{1}_{[N_{t_{i+1}-t_i}(y)=0]}\Big] + \tilde{\mathbb{E}}_y\Big[\int_0^{t_{i+1}-t_i} \mathbf{1}_{[X_s=y]}\,\mathrm{d}s\,\mathbf{1}_{[N_{t_{i+1}-t_i}(y)>0]}\Big]\\
=&(t_{i+1}-t_i) + O_{\tilde{\gamma}}\big((t_{i+1}-t_i)^2\big)\,,
\end{aligned}$$

where the function $O_{\tilde{\gamma}}$ satisfies $O_{\tilde{\gamma}}(h) \leq C_{\tilde{\gamma}}h$. Then, we only need to prove that

$$\tilde{\mathbb{E}}_x\big[N_{t_{i+1}}(y) - N_{t_i}(y)\big] = \tilde{\gamma}(y)(t_{i+1}-t_i)\,.$$

By the Markov Property, it is enough to see that $\tilde{\mathbb{E}}_x[N_h(y)] = \tilde{\gamma}(y)h$. This is a consequence of the $\tilde{\gamma}(y)$ being the parameter of the exponential clock at the site y. □

Appendix E. Basic properties of $Q(V)$

Lemma 28. $|Q(V)-Q(U)| \leq \|V-U\|_\infty$.

Proof. Since

$$P_T^V(1)(x) = \mathbb{E}_x\Big[e^{\int_0^T V(X_r)\,dr}\Big] \leq \mathbb{E}_x\Big[e^{T\|V-U\|_\infty} e^{\int_0^T U(X_r)\,dr}\Big] = e^{T\|V-U\|_\infty} P_T^U(1)(x)\,,$$

then,

$$\begin{aligned}
&|Q(V)-Q(U)| = \lim_{T\to\infty}\frac{1}{T}\log\frac{\int P_T^V(1)(x)\,\mathrm{d}\mu_A(x)}{\int P_T^U(1)(x)\,\mathrm{d}\mu_A(x)}\\
&\leq \lim_{T\to\infty}\frac{1}{T}\log\frac{\int e^{T\|V-U\|_\infty}(P_T^U 1)(x)\,\mathrm{d}\mu_A(x)}{\int P_T^U(1)(x)\,\mathrm{d}\mu_A(x)}\\
&= \|V-U\|_\infty\,.
\end{aligned}$$

□

Lemma 29. *The functional $V \to Q(V)$ is convex, i.e., for all $\alpha \in (0,1)$, we have*

$$Q(\alpha V + (1-\alpha)U) \leq \alpha Q(V) + (1-\alpha)Q(U)\,.$$

Proof. Using the Holder's inequality, we have

$$\begin{aligned}
&\int P_T^{\alpha V+(1-\alpha)U}(1)(x)\,\mathrm{d}\mu_A(x) = \mathbb{E}_{\mu_A}\Big[e^{\int_0^T \alpha V(X_r)\,dr} e^{\int_0^T (1-\alpha)U(X_r)\,dr}\Big]\\
&\leq \Big(\mathbb{E}_{\mu_A}\Big[e^{\int_0^T V(X_r)\,dr}\Big]\Big)^{\alpha}\Big(\mathbb{E}_{\mu_A}\Big[e^{\int_0^T U(X_r)\,dr}\Big]\Big)^{(1-\alpha)}\,.
\end{aligned}$$

Thus,

$$\begin{aligned}
Q(\alpha V+(1-\alpha)U) &= \lim_{T\to\infty}\frac{1}{T}\log\int P_T^{\alpha V+(1-\alpha)U}(1)(x)\,\mathrm{d}\mu_A(x)\\
&\leq \lim_{T\to\infty}\frac{1}{T}\log\Big(\int \mathbb{E}_{\mu_A}\Big[e^{\int_0^T V(X_r)\,dr}\Big]\Big)^{\alpha}\\
&\quad\times\Big(\int \mathbb{E}_{\mu_A}\Big[e^{\int_0^T U(X_r)\,dr}\Big]\Big)^{(1-\alpha)}\\
&=\alpha\lim_{T\to\infty}\frac{1}{T}\log\int \mathbb{E}_x\Big[e^{\int_0^T V(X_r)\,dr}\Big]\,\mathrm{d}\mu_A(x)\\
&\quad+(1-\alpha)\lim_{T\to\infty}\frac{1}{T}\log\int \mathbb{E}_x\Big[e^{\int_0^T U(X_r)\,dr}\Big]\,\mathrm{d}\mu_A(x)\,.
\end{aligned}$$

□

APPENDIX F. THE ASSOCIATED SYMMETRIC PROCESS AND THE METROPOLIS ALGORITHM

We can consider in our setting an extra parameter $\beta \in \mathbb{R}$ which plays the role of the inverse of temperature. For a given fixed potential V we can consider the new potential βV, $\beta \in \mathbb{R}$, and applying what we did before, we get continuous time equilibrium states described by $\gamma_\beta := \gamma_{\beta V}$ and $B_\beta := B_{\beta V}$, in the previous notation. In other words, we consider the infinitesimal generator $(\mathscr{L}_A - I) + \beta V$, $\beta > 0$, and the associated main eigenvalue $\lambda_\beta := \lambda_{\beta V}$. We denote by $L^{V,\beta}$ the infinitesimal generator of the process that is the continuous time Gibbs state for the potential βV, then $L^{V,\beta}$ acts on functions f as $L^{V,\beta}(f)(x) = \gamma_\beta(x)\sum_{\sigma(y)=x} e^{B_\beta(y)}\big[f(y) - f(x)\big]$. We are interested in the stationary probability $\mu_\beta := \mu_{B_{\beta V},\gamma_{\beta V}}$ for the semigroup $\{e^{tL^{V,\beta}}, t \geq 0\}$, and its weak limit as $\beta \to \infty$. This limit would correspond to the continuous time Gibbs state for temperature zero (see [7], [31] and [28] for related results).

The dual of $L^{V,\beta}$ on the Hilbert space $\mathbb{L}^2(\mu_\beta)$ is $L^{V,\beta^*} = \gamma_\beta\,(\mathscr{K} - I)$, where $\mathscr{K}$ is the Koopman operator. Notice that the probability μ_β is also stationary for the continuous time process with symmetric infinitesimal generator $L^{V,\beta}_{sym} := \frac{1}{2}(L^{V,\beta} + L^{V,\beta^*})$. In this new process the particle at x can jump to a $\sigma-$preimage y with probability $\frac{1}{2}e^{B_\beta(y)}$, or with probability $\frac{1}{2}$, to the forward image $\sigma(x)$, but, in both ways, according to a exponential time of parameter $\gamma_\beta(x)$.

The eigenfunction of the continuous time Markov chain with infinitesimal generator $L^{V,\beta}_{sym}$ can be different from the one with generator $L^{V,\beta}$. Given V and β, we denote $\lambda(\beta)_{sym}$ the main eigenvalue that we obtained from $\beta\,V$ and the generator $L^{V,\beta}_{sym}$. The eigenvalues of $L^{V,\beta}$ and L^{V,β^*} are the same as before. Now, we will look briefly at how to obtain $\lambda(\beta)_{sym}$. From the symmetric assumption and [12], we get, for a fixed β,

$$\begin{aligned}
\lambda(\beta)_{sym} &= \sup_{\substack{\phi\in\mathbb{L}^2(\mu_\beta),\\ \|\phi\|_2=1}} \int \phi^{1/2}\Big[\frac{\gamma_\beta}{2}\big([\mathscr{L}_\beta + \mathscr{K}] - 2I\big) + \beta V\Big](\phi^{1/2})\,\mathrm{d}\mu_\beta \\
&= \sup_{\substack{\phi\in\mathbb{L}^2(\mu_\beta),\\ \|\phi\|_2=1}} \int \phi^{1/2}\Big[\frac{1}{2}\big([\mathscr{L}_\beta + \mathscr{K}] - 2I\big) + \frac{1}{\gamma_\beta}\beta V\Big](\phi^{1/2})\,\frac{\mathrm{d}\mu_{B_\beta}}{\int \frac{1}{\gamma_\beta}\,\mathrm{d}\mu_{B_\beta}} \\
&= \sup_{\substack{\phi\in\mathbb{L}^2(\mu_\beta),\\ \|\phi\|_2=1}} \int \Big\{\phi^{1/2}\mathscr{L}_\beta(\phi^{1/2}) - 1 + \frac{1}{\gamma_\beta}\beta V|\phi|\Big\}\frac{\mathrm{d}\mu_{B_\beta}}{\int \frac{1}{\gamma_\beta}\,\mathrm{d}\mu_{B_\beta}}.
\end{aligned}$$

The second equality is due to the Definition (12), and the last one is by the dual, $\mathscr{L}^*_\beta$, on $\mathbb{L}^2(\mu_\beta)$ is $\mathscr{K}$.

Suppose one changes β in such way that β increases converging to ∞, then one can ask about the asymptotic behavior of the stationary Gibbs probability μ_β. One should analyze first what that happens with the optimal ϕ (or almost optimal) in the maximization problem above. In order to answer this last question, we use, in $\mathbb{L}^2(\mu_\beta)$, the Schwartz inequality, and we obtain

$$|\langle\phi^{1/2},\, \mathscr{L}_\beta(\phi^{1/2})\rangle_{\mu_\beta}| \leq \|\phi\|_2\, \|\mathscr{L}_\beta(\phi^{1/2})\|_2 \leq d\|\phi\|_2 = d.$$

Note that, for a fixed large β, the positive value $\gamma_\beta(x) = 1 - \beta V(x) + \lambda_{\beta V}$ became smaller close by the supremum of V. Which means that $\frac{1}{\gamma_\beta(x)}$ became large close by the supremum of V. Moreover, for fixed β, the part $\int \beta V|\phi|\,\frac{1}{\gamma_\beta}\,\frac{\mathrm{d}\mu_{B_\beta}}{\int \frac{1}{\gamma_\beta}\mathrm{d}\mu_{B_\beta}}$ of the above expression increase if we consider $|\phi|$ such that the big part of its mass is more and more close by to the supremum of βV. Note that, for fixed β, the part $\int\{\phi^{1/2}\mathscr{L}_\beta(\phi^{1/2}) - 1\}\frac{\mathrm{d}\mu_{B_\beta}}{\int \frac{1}{\gamma_\beta}\mathrm{d}\mu_{B_\beta}}$ of the above expression is bounded and just depends on ϕ. The supremum of $\int \beta V|\phi|\,\frac{1}{\gamma_\beta}\,\frac{\mathrm{d}\mu_{B_\beta}}{\int \frac{1}{\gamma_\beta}\mathrm{d}\mu_{B_\beta}}$ grows with β at least of order β.

Therefore, for large β, the maximization above should be obtained by taking $\phi = \phi_\beta$ in $\mathbb{L}^2(\mu_\beta)$ such that is more and more concentrated close by the supremum of βV. In this way, when $\beta \to \infty$ the "almost" optimal ϕ has a tendency to localize the points where the supremum of V is attained. If there is a unique point z_0 where V is optimal, then $\lambda_\beta \sim \beta V(z_0)$. The probability μ_β will converge to the delta Dirac on the point z_0. This procedure is quite similar with the process of determining ground states for a given potential via an approximation by Gibbs states which have a very small value of temperature (see for instance [1]).

The Metropolis algorithm has several distinct applications. In one of them, it can be used to maximize a function on a quite large space (see [15] and [23]). Suppose V has a unique point of maximal value. The basic idea is to produce a random algorithm that can explore the state space and localize the point of maximum, this problem may happen with a deterministic algorithm. The use of continuous time paths resulted in some advantages in the method. The randomness assures that the algorithm does note stuck on a point of local maximum of some function V. The setting we consider here has several similarities with the usual procedure. When we take β large, then the probability μ_β will be very close to the delta Dirac on the point of maximum for V as we just saw. This is so because the parameter $\frac{1}{\gamma_\beta(x)}$ of the exponential distribution became large close by the supremum of V. In the classical Metropolis algorithm there is link on β and t which is necessary for the convergence (cooling schedule in [38]). In a forthcoming paper, using our large deviation results, we will investigate the question: given small ε and δ, with probability bigger than $1-\delta$, the empirical path on the one-dimensional spin lattice will stay, up to a distance smaller the ε of the maximal value, a proportion $1-\delta$ of the time t, if t and β are chosen in a certain way (to be understood). In order to do that we have to use the large deviation results we get before.

APPENDIX G. ERGODICITY OF THE SHIFT $\Theta_t : \mathscr{D} \to \mathscr{D}$ RELATIVE TO $\mathbb{P}_{\mu_A}$

The probability $\mathbb{P}_{\mu_A}$ was obtained from $\{X_t = X_t^{\mu_A},\ t \geq 0\}$.

This section is devoted to show the ergodicity for the continuous time shift $\Theta_t : \mathscr{D} \to \mathscr{D}$, when we have that the limit below exists:

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t P_s(F)(x)\,ds = \int F\,d\mu_A.$$

The ideas presented here are based in [25] and [26].

Consider f, g functions of n variables. For all $0 \leq t_1 < \cdots < t_n$ define the functions F and G in one variable by

$$F(x) = \mathbb{E}_x\big[f(X_0, X_{t_2-t_1}\ldots, X_{t_n-t_1})\big]$$
$$G(X_{t_n}) = \mathbb{E}_{\mu_A}\big[g(X_{t_1},\ldots,X_{t_n})|X_{t_n}\big]$$

Using the Markov property, for $s > t_n - t_1$, we can write

$$\begin{aligned}\mathbb{E}_{\mu_A}\big[g(X_{t_1},\ldots,X_{t_n})f(X_{t_1+s},\ldots,X_{t_n+s})\big] &= \mathbb{E}_{\mu_A}\Big[g(X_{t_1},\ldots,X_{t_n})\mathbb{E}_{\mu_A}\big[f(X_{t_1+s},\ldots,X_{t_n+s})|\mathscr{F}_{t_n}\big]\Big]\\ &= \mathbb{E}_{\mu_A}\Big[\mathbb{E}_{\mu_A}\big[g(X_{t_1},\ldots,X_{t_n})|X_{t_n}\big]\mathbb{E}_{\mu_A}\big[f(X_{t_1+s},\ldots,X_{t_n+s})|\mathscr{F}_{t_n}\big]\Big]\\ &= \mathbb{E}_{\mu_A}\Big[G(X_{t_n})\mathbb{E}_{X_{t_n}}\big[f(X_{t_1+s-t_n},\ldots,X_s)\big]\Big].\end{aligned}$$

Since $\{X_t, t\geq 0\}$ is stationary, we obtain

$$\begin{aligned}\mathbb{E}_{\mu_A}\big[f(X_{t_1+s},\ldots,X_{t_n+s})\big] &= \mathbb{E}_{\mu_A}\Big[G(X_0)\mathbb{E}_{X_0}\big[f(X_{t_1+s-t_n},\ldots,X_s)\big]\Big]\\ &= \mathbb{E}_{\mu_A}\Big[G(X_0)\mathbb{E}_{X_0}\big[\mathbb{E}_{\mu_A}[f(X_{t_1+s-t_n},\ldots,X_s)|\mathscr{F}_{t_1+s-t_n}]\big]\Big].\end{aligned}$$

Applying again the Markov property, we get

$$\begin{aligned}\mathbb{E}_{\mu_A}\Big[G(X_0)\mathbb{E}_{X_0}\big[\mathbb{E}_{X_{t_1+s-t_n}}[f(X_0,\ldots,X_{X_{t_n-t_1}})]\big]\Big] &= \mathbb{E}_{\mu_A}\Big[G(X_0)\mathbb{E}_{X_{t_1+s-t_n}}\big[f(X_0,\ldots,X_{X_{t_n-t_1}})\big]\Big]\\ &= \mathbb{E}_{\mu_A}\big[G(X_0)F(X_{t_1+s-t_n})\big]\\ &= \int \mathbb{E}_x\big[G(X_0)F(X_{t_1+s-t_n})\big]d\mu_A(x)\\ &= \int G(x)P_{t_1+s-t_n}(F)(x)d\mu_A(x).\end{aligned}$$

Then, we have that

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t \mathbb{E}_{\mu_A}\big[g(X_{t_1},\ldots,X_{t_n})f(X_{t_1+s},\ldots,X_{t_n+s})\big]\,ds = \lim_{t\to\infty}\int G(x)\frac{1}{t}\int_0^t P_{t_1+s-t_n}(F)(x)\,ds\,d\mu_A(x)$$

The limit

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t P_{t_1+s-t_n}(F)(x)\,ds$$

exists and it is equal to $\int F\,d\mu_A$ (see o beginning of the Section 2). Thus,

$$\begin{aligned}\lim_{t\to\infty}\frac{1}{t}\int_0^t \mathbb{E}_{\mu_A}\left[g(X_{t_1},\dots,X_{t_n})f(X_{t_1+s},\dots,X_{t_n+s})\right]ds &= \int G(x)\,d\mu_A(x)\int F(x)\,d\mu_A(x)\\ &= \mathbb{E}_{\mu_A}\left[g(X_{t_1},\dots,X_{t_n})\right]\mathbb{E}_{\mu_A}\left[f(X_{t_1},\dots,X_{t_n})\right].\end{aligned}$$

Now, consider f such that $f(\Theta_s\circ(X_{t_1}(w),\dots,X_{t_n})(w)) = f(X_{t_1}(w),\dots,X_{t_n}(w))$, for all s and w, then

$$\begin{aligned}E_{\mu_A}\left[g(X_{t_1},\dots,X_{t_n})f(X_{t_1},\dots,X_{t_n})\right] &= \lim_{t\to\infty}\frac{1}{t}\int_0^t \mathbb{E}_{\mu_A}\left[g(X_{t_1},\dots,X_{t_n})f(X_{t_1},\dots,X_{t_n})\right]ds\\ &= \lim_{t\to\infty}\frac{1}{t}\int_0^t \mathbb{E}_{\mu_A}\left[g(X_{t_1},\dots,X_{t_n})f(\Theta_s\circ(X_{t_1},\dots,X_{t_n}))\right]ds\\ &= \lim_{t\to\infty}\frac{1}{t}\int_0^t \mathbb{E}_{\mu_A}\left[g(X_{t_1},\dots,X_{t_n})f(X_{t_1+s},\dots,X_{t_n+s})\right]ds\\ &= \mathbb{E}_{\mu_A}\left[g(X_{t_1},\dots,X_{t_n})\right]\mathbb{E}_{\mu_A}\left[f(X_{t_1},\dots,X_{t_n})\right].\end{aligned}$$

Take g equal to f, then

$$E_{\mu_A}\left[f(X_{t_1},\dots,X_{t_n})^2\right] = \mathbb{E}_{\mu_A}\left[f(X_{t_1},\dots,X_{t_n})\right]^2.$$

The last equality implies that f is constant (almost surely).

Considering that $\{X_t, t\geq 0\}$ is the canonical process, i.e., $X_t(w) = w_t$, for all $t\geq 0$, and $w\in\mathscr{D}$, we can rewritten our result as: given a function

$$w\in\mathscr{D}\mapsto f(w_{t_1},\dots,w_{t_n}),$$

which is invariant for the continuous time shift $\Theta_s:\mathscr{D}\to\mathscr{D}$, we get that this function is constant.

Note that all mensurable function $H:\mathscr{D}\to\mathbb{R}$ depends on a countable set of coordinates. Then, without loss of generality, suppose that $H(w) = h(w_{s_1},\dots,w_{s_n},\dots)$, where $h:\left(\{1,\dots,d\}^{\mathbb{N}}\right)^{\mathbb{N}}\to\mathbb{R}$. Therefore, using approximation arguments one also get that H is constant.

References

[1] A. Baraviera, R. Leplaideur and A. O. Lopes, Selection of ground states in the zero temperature limit for a one-parameter family of potentials. SIAM J. Appl. Dyn. Syst. 11 (2012), no. 1, 243-260.

[2] A. Baraviera, R. Exel and A. Lopes, A Ruelle Operator for continuous time Markov chains, São Paulo Journal of Mathematical Sciences. vol 4 n. 1, pp 1-16 (2010)

[3] A. T. Baraviera, L. M. Cioletti, A. Lopes, J. Mohr and R. R. Souza, On the general one dimensional *XY* Model: positive and zero temperature, selection and non-selection, Reviews in Math. Physics. Vol. 23, N. 10, pp 1063-1113 (2011).

[4] V. Baladi, Positive transfer operators and decay of correlations. World Scientific, Singapore, 2000,

[5] V. Baladi and D. Smania, Linear response formula for piecewise expanding unimodal maps, Nonlinearity 21 (2008) n.4, 677–711.

[6] N. Berger, C. Kenyon, E. Mossel and Y. Peres, Glauber Dynamics on Trees and Hyperbolic Graphs, Probability Theory and Related Fields, 131, no. 3 311-340 (2005).

[7] G. Contreras, A. O. Lopes e Ph. Thieullen, Lyapunov minimizing measures for expanding maps of the circle, *Ergodic Theory and Dynamical Systems* **21** (2001), 1379-1409.

[8] M. Craizer, Teoria Ergódica das Transformações expansoras, Master Dissertation, IMPA, Rio de Janeiro (1985)

[9] A. Dembo and O. Zeitouni, Large Deviations Techniques and Applications, Springer Verlag, 1998.

[10] M. Donsker and S. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time I. Comm. Pure Appl. Math. 28, 1–47; (1975)

[11] Jean-Dominique Deutschel and D. Stroock, Large Deviations, AMS (1989).

[12] N. Dunford and J. Schwartz, Linear Operators, Part I, Interscience (1957)

[13] F. den Hollander, Large Deviations, AMS (2000)

[14] P. Diaconis I and L. Saloff-Coste, Nash Inequalities for Finite Markov Chains, Journal of Theoretical Probability, Vol. 9, No. 2, (1996)

[15] P. Diaconis I and L. Saloff-Coste, What do we know about the Metropolis algorithm ? 27 th Annual ACM Symposium on the Theory of Computing (STOC 95) (Las Vegas, NV). J. Comput. System Sci. 57 (1998), no. 1, 2-36.

[16] Paul Dupuis and Yufei Liu, On the large deviation rate function for the empirical measures of reversible jump Markov processes, Arxiv (2013)

[17] R. Ellis, Entropy, Large Deviations, and Statistical Mechanics, Springer Verlag

[18] S. N. Ethier and T. G. Kurtz, Markov processes - Characterization and Convergence, Wiley (2005)
[19] J. Feng and T. Kurtz, Large Deviations for Stochastic Processes, AMS (2006)
[20] Y. Kifer, Large Deviations in Dynamical Systems and Stochastic processes, TAMS, Vol 321, N.2, 505–524 (1990)
[21] Y. Kifer, Principal eigenvalues, topological pressure, and stochastic stabilty of equilibrium states, Israel Journal of Mathematics, Vol. 70, No. I, pp 1-47 (1990)
[22] C. Landim and C. Kipnis, Scaling limits of interacting particle systems. Grundlehren der Mathematischen Wissenschaften, 320. (1999) Springer-Verlag, Berlin.
[23] G. Lebeau, Introduction a l'analyse de l'algorithme de Metropolis, Preprint, http://math.unice.fr/ sdescomb/MOAD/CoursLebeau.pdf
[24] C. Leonard, Large deviations for Poisson random measures and processes with independent increments, Stochastic Processes and their Applications, 85 (2000), 93-121
[25] T. Liggett, Interacting Particle Systems, Springer, (1985).
[26] T. Liggett, Stocastic Interacting Systems: Contact, Voter and Exclusion Process, Springer, (1991).
[27] T. Liggett, Continuous Time Markov Processes, AMS, (2010).
[28] A. Lopes, J. Mengue, J. Mohr and R. R. Souza, Entropy and Variational Principle for one-dimensional Lattice Systems with a general a-priori probability: positive and zero temperature, Arxiv (2012)
[29] A. O. Lopes, An analogy of the charge on distribution on Julia sets with the Brownian motion, Journal Math. Physics. Vol. 30, N. 9, pp 2120-2124 (1989).
[30] A. O. Lopes, Entropy and Large Deviations, NonLinearity, Vol. 3, N 2, pp. 527-546, (1990).
[31] A. O. Lopes, J. Mohr, R. Souza and Ph. Thieullen, Negative entropy, zero temperature and stationary Markov chains on the interval, *Bulletin of the Brazilian Mathematical Society* **40**, 1-52, 2009.
[32] A. O. Lopes. Thermodynamic Formalism, Maximizing Probabilities and Large Deviations. Work in progress (2012). Lecture Notes - Dynamique en Cornouaille
[33] W. Parry and M. Pollicott, Zeta functions and the periodic orbit structure of hyperbolic dynamics, *Astérisque* **187-188** (1990).
[34] P. Protter, Stochastic Integration and Differential Equations, Springer Verlag, (1990).
[35] Dana Randall and P. Tetali, Analyzing Glauber dynamics by comparison of Markov chains, Journal of Mathematical Physics Vol 41, Number 3, march, 1598-1615 (2000)
[36] D. Revuz and M. Yor, Continuous Martingales and Brownian Motion, 3rd ed., Springer Verlang, (1999).
[37] D. Ruelle, Thermodynamic formalism, Addison-Wesley, Reading (Mass.),1978.
[38] D. Stroock, An Introduction to Markov Processes, Springer Verlag, (2000)
[39] D. Stroock, An Introduction to the Theory of Large Deviations, Springer Verlag
[40] D. Stroock and B. Zegarlinski, On the ergodic properties of Glauber dynamics, Journal of Statistical Physics, December 1995, Volume 81, Issue 5-6, pp 1007-1019
[41] K. Yoshida, Functional Analysis, Springer Verlag, (1978).

UFRGS, INSTITUTO DE MATEMÁTICA, AV. BENTO GONÇALVES, 9500. CEP 91509-900, PORTO ALEGRE, BRASIL
E-mail address: arturoscar.lopes@gmail.com

UFRGS, INSTITUTO DE MATEMÁTICA, AV. BENTO GONÇALVES, 9500. CEP 91509-900, PORTO ALEGRE, BRASIL
E-mail address: aneumann@impa.br

INSTITUT DE MATHÉMATIQUES, UNIVERSITÉ BORDEAUX 1, BOURDEAUX, FRANCE
E-mail address: philippe.thieullen@math.u-bordeaux1.fr