TUNNELING AND METASTABILITY OF CONTINUOUS TIME MARKOV CHAINS II, THE NONREVERSIBLE CASE

J. BELTRÁN, C. LANDIM

Abstract. We proposed in [1] a new approach to prove the metastable behavior of reversible dynamics based on potential theory and local ergodicity. In this article we extend this theory to nonreversible dynamics based on the Dirichlet principle proved in [11].

1. Introduction

Metastability is a relevant dynamical phenomenon in non-equilibrium statistical mechanics which occurs in the vicinities of first order phase transitions [17], [10]. In the sequel of [15, 7, 5], we proposed in [1] a new definition of metastability for continuous time Markov chains which visit points, and we presented in the reversible case three simple conditions which ensure the metastable behavior of a chain. All three conditions are formulated in terms of the stationary measure and of the capacity between the metastable sets.

This method has been applied successfully in [2] to prove the metastable behavior of the condensate in reversible zero range processes evolving on finite sets, in [3, 4] to examine the metastability of reversible Markov chains evolving on fixed finite sets, and in [12, 13] to investigate the scaling limits of trap models on random graphs.

The proof of the metastable behavior of a reversible Markov chain presented in [1] relies essentially on two ingredients. On the one hand, the potential theory of reversible Markov chains [6, 10] permits to express the capacity between two sets in terms of a variational problem, the so-called Dirichlet principle, and provides simple formulas for the expectations of time integrals, carried over time intervals ending at the hitting time of a set, in terms of the equilibrium potentials. On the other hand, the local ergodicity allows to replace a time integral of a function by the time integral of the average of the function over the metastable sets.

We generalized recently in [11] the potential theory of reversible Markov chains to the non-reversible context by defining a capacity between two sets and by proving a Dirichlet principle which involves, in contrast with the reversible case, a double variational formula.

In this article, we extend the theory elaborated in [1] to nonreversible dynamics aiming to examine the metastable behavior of the condensate in asymmetric zero range processes evolving on a one-dimensional finite torus [14]. The main result, stated in Theorem 2.1 below, asserts that the metastable behavior of a sequence of continuous time Markov chains follows from the same three conditions formulated in the reversible case.

One of the conditions, assumption (**H0**) below, requires the average jump rates of the trace process to converge. While in the reversible case the average rates can be expressed in terms of capacities, [1, Remark 2.9], in the nonreversible case such formula is not available. We may, however, formulate a double variational problem whose optimal solution is equal to the average rate. This is the content of Proposition 2.2.

It may seem that the characterization of the average jump rate provided by Proposition 2.2 is deprived of interest and impossible to use in concrete examples. This is not the case as shown in [14] where we apply this result to obtain the asymptotic value of the average jump rates. This is done in two steps. We first compute the limit of (2.5) as $N \uparrow \infty$ and derive the optimal value for f on the set $\mathcal{E}_N^x$. We then show that if the supremum in (2.5) is carried over functions h which are not equal to the optimal value for f on $\mathcal{E}_N^x$, the limit of (2.5) is strictly smaller than the one obtained before.

The last main result of the article, Proposition 2.3, asserts that the adjoint of a Markov chain which exhibits a metastable behavior is also metastable and its asymptotic dynamics is the adjoint of the asymptotic dynamics of the original process.

2. NOTATION AND RESULTS

Fix a sequence $(E_N : N \geq 1)$ of countable state spaces. The elements of E_N are denoted by the Greek letters η, ξ. A sequence of states $\boldsymbol{\eta} = (\eta^N \in E_N : N \geq 1)$ is said to be a point in a sequence $\mathcal{A}$ of subsets of E_N, $\mathcal{A} = (A_N \subseteq E_N : N \geq 1)$, if η^N belongs to A_N for every $N \geq 1$.

For each $N \geq 1$ consider a matrix $R_N : E_N \times E_N \to \mathbb{R}$ such that $R_N(\eta, \xi) \geq 0$ for $\eta \neq \xi$, $-\infty < R_N(\eta, \eta) \leq 0$ and $\sum_{\xi \in E_N} R_N(\eta, \xi) = 0$ for all $\eta \in E_N$. Denote by L_N the generator which acts on bounded functions $f : E_N \to \mathbb{R}$ as

$$(L_N f)(\eta) \;=\; \sum_{\xi \in E_N} R_N(\eta, \xi) \left\{ f(\xi) - f(\eta) \right\} . \tag{2.1}$$

Let $\{\eta_t^N : t \geq 0\}$ be the *minimal* right-continuous Markov process associated to the generator L_N. It is well known that $\{\eta_t^N : t \geq 0\}$ is a strong Markov process with respect to the filtration $\{\mathcal{F}_t^N : t \geq 0\}$ given by $\mathcal{F}_t^N = \sigma(\eta_s^N : s \leq t)$ [9, 16]. We assume throughout this section that for each $N \geq 1$ the process $\{\eta_t^N : t \geq 0\}$ is *irreducible* and *positive recurrent*.

Denote by μ^N the unique stationary probability measure and let $\lambda^N(\eta) = -R(\eta, \eta)$, $\eta \in E_N$, be the holding rates of the process η_t^N. We assume that the holding rates are integrable:

$$\sum_{\eta \in E_N} \lambda_N(\eta) \mu_N(\eta) \;<\; \infty \,, \quad \forall N \geq 1 \,. \tag{2.2}$$

Denote by $D(\mathbb{R}_+, E_N)$ the space of right-continuous trajectories with left limits endowed with the Skorohod topology. Let $\mathbf{P}_\eta^N$, $\eta \in E_N$, be the probability measure on $D(\mathbb{R}_+, E_N)$ induced by the Markov process $\{\eta_t^N : t \geq 0\}$ starting from η. Expectation with respect to $\mathbf{P}_\eta^N$ is denoted by $\mathbf{E}_\eta^N$ and we frequently omit the index N in $\mathbf{P}_\eta^N$, $\mathbf{E}_\eta^N$.

Denote by H_A (resp. H_A^+), $A \subseteq E_N$, the hitting time of (resp. return time to) the set A:

$$H_A := \inf\{t > 0 : \eta_t^N \in A\},$$

$$H_A^+ := \inf\{t > 0 : \eta_t^N \in A, \eta_s^N \neq \eta_0^N \text{ for some } 0 < s < t\}.$$

Denote by $L^2(\mu_N)$ the space of square integrable functions $f : E_N \to \mathbb{R}$, $\sum_{\eta \in E_N} f(\eta)^2 \mu_N(\eta) < \infty$, endowed with the usual scalar product. Let L^* be the adjoint of L in $L^2(\mu_N)$. The operator L^* corresponds to the generator of a Markov process, denoted by $\{\eta_t^{*,N} : t \geq 0\}$, whose rates are represented by R_N^*. Denote by $\mathbf{P}^* = \mathbf{P}_\eta^{*,N}$, $\eta \in E_N$, the probability measure on $D(\mathbb{R}_+, E_N)$ induced by the adjoint Markov process starting from η. Expectation with respect to $\mathbf{P}_\eta^{*,N}$ is represented by $\mathbf{E}_\eta^{*,N}$

Fix a finite number of disjoint subsets $\mathcal{E}_N^1, \ldots, \mathcal{E}_N^\kappa$, $\kappa \geq 2$, of E_N: $\mathcal{E}_N^x \cap \mathcal{E}_N^y = \varnothing$, $x \neq y$. Let $S = \{1, \ldots, \kappa\}$, $\mathcal{E}_N = \cup_{x \in S} \mathcal{E}_N^x$ and let $\Delta_N = E_N \setminus \mathcal{E}_N$ so that

$$E_N = \mathcal{E}_N^1 \cup \cdots \cup \mathcal{E}_N^\kappa \cup \Delta_N = \mathcal{E}_N \cup \Delta_N. \tag{2.3}$$

Let $\breve{\mathcal{E}}_N^x := \mathcal{E}_N \setminus \mathcal{E}_N^x$,

$$\mathcal{E} = (\mathcal{E}_N : N \geq 1), \quad \mathcal{E}^x = (\mathcal{E}_N^x : N \geq 1) \quad \text{and} \quad \breve{\mathcal{E}}^x = (\breve{\mathcal{E}}_N^x : N \geq 1).$$

Denote by $\{\eta_t^{\mathcal{E}_N} : t \geq 0\}$ the trace of $\{\eta_t^N : t \geq 0\}$ on $\mathcal{E}_N$ [1]. Let $R_N^{\mathcal{E}} : \mathcal{E}_N \times \mathcal{E}_N \to \mathbb{R}_+$ be the transition rates of the trace process $\{\eta_t^{\mathcal{E}_N} : t \geq 0\}$, and denote by $r_N(\mathcal{E}_N^x, \mathcal{E}_N^y)$ the mean rate at which the trace process jumps from $\mathcal{E}_N^x$ to $\mathcal{E}_N^y$:

$$r_N(\mathcal{E}_N^x, \mathcal{E}_N^y) := \frac{1}{\mu_N(\mathcal{E}_N^x)} \sum_{\eta \in \mathcal{E}_N^x} \mu_N(\eta) \sum_{\xi \in \mathcal{E}_N^y} R_N^{\mathcal{E}}(\eta, \xi), \tag{2.4}$$

and let

$$r_N(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x) = \sum_{y \neq x} r_N(\mathcal{E}_N^x, \mathcal{E}_N^y).$$

For two disjoint subsets A, B of E_N, denote by $\mathrm{cap}_N(A, B)$ the capacity between A and B, defined in [11] by (6.7) for irreducible positive recurrent Markov processes. For a point $\boldsymbol{\xi}_x = (\xi_x^N : N \geq 1)$ in $\mathcal{E}^x$, let

$$\mathrm{cap}_N(\boldsymbol{\xi}_x) = \inf_{\eta \in \mathcal{E}_N^x} \mathrm{cap}_N(\{\eta\}, \{\xi_x^N\}).$$

Recall the notion of tunneling behavior and recall properties (**M1**), (**M2**), (**M3**) introduced in Definition 2.2 of [1].

Theorem 2.1. *Suppose that there exists a sequence $\boldsymbol{\theta} = (\theta_N : N \geq 1)$ of positive numbers such that, for every pair $x, y \in S$, $x \neq y$, the following limit exists*

$$r(x, y) := \lim_{N \to \infty} \theta_N r_N(\mathcal{E}_N^x, \mathcal{E}_N^y). \tag{H0}$$

Suppose that for each $x \in S$, there exists a point $\boldsymbol{\xi}_x = (\xi_x^N : N \geq 1)$ in $\mathcal{E}^x$ such that

$$\lim_{N \to \infty} \frac{\mathrm{cap}_N(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x)}{\mathrm{cap}_N(\boldsymbol{\xi}_x)} = 0. \tag{H1}$$

*Then, for any points $\{\boldsymbol{\zeta}_x \in \mathcal{E}^x : x \in S\}$, properties (**M1**) and (**M2**) of tunneling hold on the time-scale $\boldsymbol{\theta}$, with metastates $\{\mathcal{E}^x : x \in S\}$, metapoints $\{\boldsymbol{\zeta}_x : x \in S\}$ and asymptotic Markov dynamics characterized by the rates $r(x, y)$, $x, y \in S$.*

If in addition we have that (**M3**) *holds for each* $x \in S$ *which is an absorbing state of the Markov dynamics on* S *determined by the rates* r *and that*

$$\lim_{N\to\infty} \frac{\mu_N(\Delta_N)}{\mu_N(\mathcal{E}_N^x)} = 0 \qquad\qquad\text{(H2)}$$

for each non-absorbing state $x \in S$, *then property* (**M3**) *holds for every* $x \in S$.

In contrast with the reversible case, we are not able to express the average rates $r_N(\mathcal{E}_N^x, \mathcal{E}_N^y)$, $x, y \in S$, in terms of capacities. We may, however, formulate a variational problem whose optimal solution is equal to the average rate.

Proposition 2.2. *Fix* $x \neq y$ *in* S *and consider the variational problem*

$$\inf_f \sup_h \left\{ 2\langle f, L_N h\rangle_{\mu_N} - \langle h, (-L_N)h\rangle_{\mu_N} \right\}, \qquad (2.5)$$

where the infimum is carried over all functions $f : E_N \to \mathbb{R}$ *which are equal to* 1 *on* $\mathcal{E}_N^y$, 0 *on* $\mathcal{E}_N \setminus (\mathcal{E}_N^x \cup \mathcal{E}_N^y)$, *and which are constant on* $\mathcal{E}_N^x$, *while the supremum is carried over all functions* $h : E_N \to \mathbb{R}$ *which are equal to* 0 *on* $\mathcal{E}_N \setminus (\mathcal{E}_N^x \cup \mathcal{E}_N^y)$ *and which are constant over* $\mathcal{E}_N^y$, $\mathcal{E}_N^x$, *with possibly different values at each set. Then, the optimal function* h *associated to the optimal function* f *is such that*

$$h(\eta) = \frac{r_N(\mathcal{E}_N^x, \mathcal{E}_N^y)}{r_N(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x)} \quad \text{for all} \quad \eta \in \mathcal{E}_N^x \,.$$

This result is a particular case of Proposition 4.2 below. We conclude this section with two remarks on metastability. The first one concerns the metastable behavior of the adjoint of processes which exhibit a metastable behavior, while the second one points out that condition (**H0**) alone guarantees the metastable behavior of the collapsation of the trace process.

Proposition 2.3. *Suppose that a sequence of Markov process* $\{\eta_t^N : t \geq 0\}$ *satisfies the hypotheses of Theorem 2.1, that* $\mu_N(\mathcal{E}_N^x)$ *converges to a strictly positive limit* $m(x)$ *for each* $x \in S$ *and that Markov dynamics on* S *determined by the rates* r *has no absorbing states. Then,* m *is a stationary probability measure for this asymptotic dynamics, the adjoint process* $\{\eta_t^{*,N} : t \geq 0\}$ *also satisfies the assumptions of Theorem 2.1, and the limiting jump rates introduced in condition* (**H0**) *for the adjoint process, denoted by* $r_*(x, y)$, *are the adjoint rates of the asymptotic dynamics with respect to the measure* m. *In particular, the adjoint process also exhibits a tunneling behavior.*

Proof. Denote by $r_N^*(\mathcal{E}_N^x, \mathcal{E}_N^y)$, $x \neq y \in S$, the mean jump rates of the trace on $\mathcal{E}_N$ of the adjoint process. By definition of m and r, $\sum_{y\neq x} m(y)r(y,x)$ is equal to

$$\lim_{N\to\infty} \sum_{y\neq x} \mu_N(\mathcal{E}_N^y)\, \theta_N\, r_N(\mathcal{E}_N^y, \mathcal{E}_N^x) = \lim_{N\to\infty} \sum_{y\neq x} \mu_N(\mathcal{E}_N^x)\, \theta_N\, r_N^*(\mathcal{E}_N^x, \mathcal{E}_N^y)$$

where the last identity follows from the explicit expressions for the jump rates of the trace processes presented in [1, Proposition 6.1]. By (6.8) and (6.10),

$$\mu_N(\mathcal{E}_N^x) \sum_{y\neq x} r_N^*(\mathcal{E}_N^x, \mathcal{E}_N^y) = \mu_N(\mathcal{E}_N^x) r_N^*(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x) = \mathrm{cap}_N^*(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x) \,,$$

where cap_N^* represents the capacity with respect to the adjoint process. By [11, Lemma 2.3] and by formula (2.4) in that article, $\mathrm{cap}_N^*(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x) = \mathrm{cap}_N(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x)$

so that

$$\lim_{N\to\infty} \sum_{y\neq x} \mu_N(\mathcal{E}_N^y)\,\theta_N\,r_N(\mathcal{E}_N^y,\mathcal{E}_N^x) \;=\; \lim_{N\to\infty} \mu_N(\mathcal{E}_N^x)\sum_{y\neq x}\theta_N\,r_N(\mathcal{E}_N^x,\mathcal{E}_N^y)\,,$$

that is, $\sum_{y\neq x} m(y)r(y,x) \;=\; m(x)\sum_{y\neq x} r(x,y)$, proving that m is a stationary measure.

To prove the second assertion of the proposition, note that condition (**H1**) is satisfied because the measure and the capacities of the adjoint process coincides with the measure and the capacities of the original process.

On the other hand, as we have seen above,

$$\mu_N(\mathcal{E}_N^x)\,r_N^*(\mathcal{E}_N^x,\mathcal{E}_N^y) \;=\; \mu_N(\mathcal{E}_N^y)\,r_N(\mathcal{E}_N^y,\mathcal{E}_N^x)$$

for all $x \neq y$ in S. Hence, by condition (**H0**) for the original process and by assumption,

$$\lim_{N\to\infty}\theta_N\,r_N^*(\mathcal{E}_N^x,\mathcal{E}_N^y) \;=\; \frac{m(y)}{m(x)}r(y,x) \;=:\; r_*(x,y)\,.$$

This proves condition (**H0**) for the adjoint process and shows that the dynamics on S induced by the rates $r_*(x,y)$ has no absorbing point since

$$\sum_{y\neq x} r_*(x,y) \;=\; \sum_{y\neq x} r(x,y) \;>\; 0$$

by assumption. In particular, condition (**H2**) for the adjoint process is in force. □

Recall that we denote by $\{\eta_t^{\mathcal{E}_N} : t \geq 0\}$ the trace of the process η_t^N on $\mathcal{E}_N$. Denote furthermore by $\{\eta_t^S : t \geq 0\}$ the Markov process on $S = \{1,\ldots,\kappa\}$ obtained from $\eta_t^{\mathcal{E}_N}$ by collapsing each set $\mathcal{E}_N^x$ to the point x [11]. The jump rates of the Markov process η_t^S, denoted by $r_N^S(x,y)$, are given by

$$r_N^S(x,y) \;=\; r_N(\mathcal{E}_N^x,\mathcal{E}_N^y)\,,\quad x\neq y \in S\,.$$

In particular, under assumption (**H0**) the speeded up Markov process $\eta_{t\theta_N}^S$ converges to the Markov process on S characterized by the rates $r(x,y)$.

3. Proof of Theorem 2.1

We prove in this section the main result of the article. Given two disjoint subsets $A, B \subset E_N$, let

$$f_{AB}(\eta) \;:=\; \mathbf{P}_\eta\big[\,H_A < H_B\,\big]\,,\quad f_{AB}^*(\eta) \;:=\; \mathbf{P}_\eta^*\big[\,H_A < H_B\,\big]\,.$$

In addition, for point $\boldsymbol{\eta} = (\eta^N : N \geq 1)$, $\boldsymbol{\xi} = (\xi^N : N \geq 1)$ in $\mathcal{E}$, $\eta^N \neq \xi^N$, and a set $\mathcal{A} = (A_N : N \geq 1)$, set

$$f_N^*(\boldsymbol{\eta},\boldsymbol{\xi}) \;=\; f_{\{\eta^N\}\{\xi^N\}}^*\,,\quad f_N^*(\boldsymbol{\eta},\mathcal{A}) \;=\; f_{\{\eta^N\}A_N}^*\,,\quad \mathrm{cap}_N(\boldsymbol{\eta},\boldsymbol{\xi}) = \mathrm{cap}_N(\{\eta^N\},\{\xi^N\})\,,$$

and a similar notation without the upper index $*$.

For two subsets A, B of E_N, denote by $H_A(B)$ the time the process η_t^N spent on the set B before hitting the set A:

$$H_A(B) \;=\; \int_0^{H_A} \mathbf{1}\{\eta_s^N \in B\}\,ds\,.$$

Fix $x \in S$. By the definition of $H_{\xi_x^N}(\mathcal{E}_N^x)$ given just above, by (6.13) below, and by identity (6.8) for the capacity,

$$\mathbf{E}_{\eta^N}\Big[\int_0^{H_{\xi_x}} R_N(\eta_s^N, \breve{\mathcal{E}}_N^x)\,\mathbf{1}\{\eta_s^N \in \mathcal{E}_N^x\}\,ds\Big] \;=\; \frac{\langle\, R_N(\,\cdot\,, \breve{\mathcal{E}}_N^x)\,\mathbf{1}\{\mathcal{E}_N^x\}\,,\, f_N^*(\boldsymbol{\eta},\boldsymbol{\xi})\,\rangle_{\mu_N}}{\mathrm{cap}_N(\boldsymbol{\eta},\boldsymbol{\xi})}\;,$$

$$r_N(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x)\,\mathbf{E}_{\eta^N}[H_{\xi_x^N}(\mathcal{E}_N^x)] \;=\; \frac{\mathrm{cap}_N(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x)}{\mu_N(\mathcal{E}_N^x)}\,\frac{\langle\,\mathbf{1}\{\mathcal{E}_N^x\}\,,\, f_N^*(\boldsymbol{\eta},\boldsymbol{\xi})\,\rangle_{\mu_N}}{\mathrm{cap}_N(\boldsymbol{\eta},\boldsymbol{\xi})}\;.$$

Since $0 \le f_N^*(\boldsymbol{\eta},\boldsymbol{\xi}) \le 1$, and since $\langle\, R_N(\,\cdot\,, \breve{\mathcal{E}}_N^x)\,\mathbf{1}\{\mathcal{E}_N^x\}\,\rangle_{\mu_N} = \mu_N(\mathcal{E}_N^x)r_N(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x)$, by identity (6.8) again, we obtain the following estimates which will be used repeatedly below

$$\begin{aligned}
\mathbf{E}_{\eta^N}\Big[\int_0^{H_{\xi_x}} R_N(\eta_s^N, \breve{\mathcal{E}}_N^x)\,\mathbf{1}\{\eta_s^N \in \mathcal{E}_N^x\}\,ds\Big] &\le \frac{\mathrm{cap}_N(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x)}{\mathrm{cap}_N(\boldsymbol{\xi}_x)}\;,\\
r_N(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x)\,\mathbf{E}_{\eta^N}[H_{\xi_x^N}(\mathcal{E}_N^x)] &\le \frac{\mathrm{cap}_N(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x)}{\mathrm{cap}_N(\boldsymbol{\xi}_x)}\;.
\end{aligned} \tag{3.1}$$

Lemma 3.1 below is Proposition 5.10 of [1]. The proof needs some modifications to take into account the lack of reversibility. Let $R_N^{*,\mathcal{E}_N} : \mathcal{E}_N \times \mathcal{E}_N \to \mathbb{R}_+$ be the transition rates of the trace on $\mathcal{E}_N$ of the adjoint process $\eta_t^{*,N}$. Denote by $r_N^*(\mathcal{E}_N^x, \mathcal{E}_N^y)$ the mean rate at which this trace process jumps from $\mathcal{E}_N^x$ to $\mathcal{E}_N^y$:

$$r_N^*(\mathcal{E}_N^x, \mathcal{E}_N^y) \;:=\; \frac{1}{\mu_N(\mathcal{E}_N^x)}\sum_{\eta \in \mathcal{E}_N^x} \mu_N(\eta)\sum_{\xi \in \mathcal{E}_N^y} R_N^{*,\mathcal{E}_N}(\eta,\xi)\;,$$

and let $R_N^{*,\mathcal{E}_N}(\eta, \mathcal{E}_N^y) = \sum_{\xi \in \mathcal{E}_N^y} R_N^{*,\mathcal{E}_N}(\eta,\xi)$, $y \ne x$,

$$r_N^*(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x) \;=\; \sum_{y \ne x} r_N^*(\mathcal{E}_N^x, \mathcal{E}_N^y)\;.$$

Lemma 3.1. *Fix $x \in S$. Assume condition (**H1**) for some point $\boldsymbol{\xi}_x = (\xi_x^N : N \ge 1)$ in $\mathcal{E}^x$. Then, for every point $\boldsymbol{\zeta} = (\zeta^N : N \ge 1)$ in $\mathcal{E}^x$,*

$$\lim_{N \to \infty} \frac{\mathrm{cap}_N(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x)}{\mathrm{cap}_N(\zeta^N, \breve{\mathcal{E}}_N^x)} \;=\; 1\;, \tag{3.2}$$

and

$$\lim_{N \to \infty} \inf_{\eta \in \mathcal{E}^x} \mathbf{P}_\eta^*[\,H_{\zeta^N} < H_{\breve{\mathcal{E}}_N^x}\,] \;=\; 1\;, \quad \lim_{N \to \infty} \inf_{\eta \in \mathcal{E}^x} \mathbf{P}_\eta[\,H_{\zeta^N} < H_{\breve{\mathcal{E}}_N^x}\,] \;=\; 1\;. \tag{3.3}$$

Proof. Fix $x \in S$ and some point $\boldsymbol{\xi}_x = (\xi_x^N : N \ge 1)$ in $\mathcal{E}^x$ for which (**H1**) holds. In view of (6.9), by (3.1) for the adjoint process and by assumption (**H1**), for every point $\boldsymbol{\eta} = (\eta^N : N \ge 1)$ in $\mathcal{E}^x$,

$$\lim_{N \to \infty} \mathbf{E}_{\eta^N}^*\Big[\int_0^{H_{\xi_x^N}} R_N^{*,\mathcal{E}_N}(\eta_s^{*,N}, \breve{\mathcal{E}}_N^x)\,\mathbf{1}\{\eta_s^{*,N} \in \mathcal{E}_N^x\}\,ds\Big] \;=\; 0\;,$$

$$\lim_{N \to \infty} r_N^*(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x)\,\mathbf{E}_{\eta^N}^*[H_{\xi_x^N}(\mathcal{E}_N^x)] \;=\; 0\;.$$

The assumptions of [1, Proposition 5.1] are thus fulfilled for the adjoint process and with $\mathcal{W} = \mathcal{E}^x$, $\mathcal{B}^c = \breve{\mathcal{E}}^x$ and the point $\boldsymbol{\xi}_x = (\xi_x^N : N \ge 1)$.

Fix an arbitrary point $\boldsymbol{\zeta} = (\zeta^N : N \ge 1)$ in $\mathcal{E}^x$. We show that conditions (i) and (ii) of [1, Proposition 4.2] are in force for the adjoint process and with $\mathcal{W} = \mathcal{E}^x$, $\mathcal{B}^c = \breve{\mathcal{E}}^x$, $\boldsymbol{\xi} = \boldsymbol{\zeta}$.

Fix a point $\boldsymbol{\eta} = (\eta^N : N \geq 1)$ in $\mathcal{E}^x$. By (6.13) applied to η^N, $g = \mathbf{1}\{\mathcal{E}_N^x\}$, $B = \{\xi^N\}$, and to ξ^N, $g = \mathbf{1}\{\mathcal{E}_N^x\}$, $B = \{\zeta^N\}$,

$$
\begin{aligned}
\mathbf{E}_{\eta^N}^*[H_{\zeta^N}(\mathcal{E}_N^x)] &\leq \mathbf{E}_{\eta^N}^*[H_{\xi^N}(\mathcal{E}_N^x)] + \mathbf{E}_{\xi^N}^*[H_{\zeta^N}(\mathcal{E}_N^x)] \\
&\leq \frac{\mu_N(\mathcal{E}_N^x)}{\mathrm{cap}_N(\eta_N, \xi_N)}\mathbf{1}\{\eta^N \neq \xi^N\} + \frac{\mu_N(\mathcal{E}_N^x)}{\mathrm{cap}_N(\zeta_N, \xi_N)}\mathbf{1}\{\zeta^N \neq \xi^N\} \\
&\leq \frac{2\,\mu_N(\mathcal{E}_N^x)}{\mathrm{cap}_N(\xi_N)}\,.
\end{aligned}
$$

It follows from this estimate, identities (6.8), (6.9) and assumption (**H1**), that

$$
\lim_{N \to \infty} r_N^*(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x)\, \mathbf{E}_{\eta^N}^*[H_{\zeta^N}(\mathcal{E}_N^x)] = 0\,.
$$

Therefore, applying the assertion (5.1) of [1, Proposition 5.1], to $\mathcal{W} = \mathcal{E}^x$, $\mathcal{B}^c = \breve{\mathcal{E}}^x$,

$$
\lim_{N \to \infty} \frac{\mathbf{E}_{\eta^N}^*[H_{\zeta^N}(\mathcal{E}_N^x)]}{\mathbf{E}_{\eta^N}^*[H_{\breve{\mathcal{E}}_N^x}(\mathcal{E}_N^x)]} = \lim_{N \to \infty} r_N^*(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x)\, \mathbf{E}_{\eta^N}^*[H_{\zeta^N}(\mathcal{E}_N^x)] = 0\,.
$$

This result shows that condition (i) of [1, Proposition 4.2] holds.

Fix a point $\boldsymbol{\eta} = (\eta^N : N \geq 1)$ in $\mathcal{E}^x$. By [1, Proposition 5.1], under $\mathbf{P}_{\eta^N}^*$, $H_{\breve{\mathcal{E}}_N^x}(\mathcal{E}_N^x)/\mathbf{E}_{\eta^N}^*[H_{\breve{\mathcal{E}}_N^x}(\mathcal{E}_N^x)]$ converges to a mean one exponential distribution, which is condition (ii) of [1, Proposition 4.2].

By items (i) and (ii) of Proposition 4.2 we conclude that $(\mathcal{E}^x, \mathcal{E}^x, \boldsymbol{\zeta})$ is a valley for the trace on $\mathcal{E}_N$ of the adjoint process. Hence, for any point $\boldsymbol{\eta} = (\eta^N : N \geq 1)$ in $\mathcal{E}^x$

$$
\lim_{N \to \infty} \mathbf{P}_{\eta^N}^*[H_{\zeta^N}(\mathcal{E}_N) < H_{\breve{\mathcal{E}}_N^x}(\mathcal{E}_N)] = 1\,,
$$

which implies condition (**V1**) for the triple $(\mathcal{E}^x, \mathcal{E}, \boldsymbol{\zeta})$ because clearly $\{H_{\zeta^N}(\mathcal{E}_N) < H_{\breve{\mathcal{E}}_N^x}(\mathcal{E}_N)\} \subseteq \{H_{\zeta^N} < H_{\breve{\mathcal{E}}_N^x}\}$ $\mathbf{P}_{\eta^N}^*$-a.s., proving the first assertion in (3.3). The proof of the second one is exactly the same replacing the adjoint process by the original one.

The arguments presented in the first part of the proof, applied to the original process instead of the adjoint process, show that the assumptions of [1, Proposition 5.1] are in force for the original process and with $\mathcal{W} = \mathcal{E}^x$, $\mathcal{B}^c = \breve{\mathcal{E}}^x$ and the point $\boldsymbol{\xi}_x = (\xi_x^N : N \geq 1)$. Fix a point $\boldsymbol{\zeta} = (\zeta^N : N \geq 1)$ in $\mathcal{E}^x$. By (6.13) with $A = \{\zeta^N\}$, $B = \breve{\mathcal{E}}_N^x$ and $g = \mathbf{1}\{\mathcal{E}_N^x\}$, and by identity (6.8), the limit (5.1) in [1, Proposition 5.1] can be re-written as

$$
\lim_{N \to \infty} \frac{\langle \mathbf{1}\{\mathcal{E}_N^x\}, f_N^*(\boldsymbol{\zeta}, \breve{\mathcal{E}}^x)\rangle_{\mu_N}\, \mathrm{cap}_N(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x)}{\mu_N(\mathcal{E}_N^x)\, \mathrm{cap}_N(\zeta^N, \breve{\mathcal{E}}_N^x)} = 1\,.
$$

By (3.3), the infimum of $f_N^*(\boldsymbol{\zeta}, \breve{\mathcal{E}}^x)$ over $\mathcal{E}_N^x$ converges to 1 as $N \uparrow \infty$. Therefore, (3.2) follows from this observation and the previous identity. $\qquad\square$

Proof of Theorem 2.1. The proof of this result is similar to the one of [1, Theorem 2.7]. We first need to check that all assumptions of [1, Theorem 2.4] are in force.

By (3.1), conditions (**C1**) and (**C2**) in [1, Theorem 2.4] follow from the stronger condition (**H1**). Therefore, by [1, Theorem 2.4], properties (**M1**) and (**M2**) of tunneling hold on the time-scale $\boldsymbol{\theta}$, with metastates $\{\mathcal{E}^x : x \in S\}$, metapoints $\{\boldsymbol{\xi}_x : x \in S\}$ and asymptotic Markov dynamics characterized by the rates $r(x, y)$, $x, y \in S$. The estimate (3.3) shows that we can replace the metapoint $\boldsymbol{\xi}_x$ by any point $\boldsymbol{\zeta}_x$ in $\mathcal{E}^x$.

We turn to condition (**M3**). Fix $x \in S$ and $t > 0$. Decomposing the time interval $[0, t\theta_N]$ according to the successive visits of the trace process $\eta_t^{\mathcal{E}_N}$ to the metastates $\mathcal{E}_N^z$, $z \in S$, we see that to prove condition (**M3**) for all $x \in S$, it is enough to show that this condition is in force for each absorbing state x for the Markov dynamics on S determined by the rates r and that

$$\limsup_{N \to \infty} \sup_{\eta \in \mathcal{E}_N^x} \frac{1}{\theta_N} \mathbf{E}_\eta \big[H_{\breve{\mathcal{E}}_N^x}(\Delta_N) \big] \ = \ 0$$

holds for each non-absorbing state x.

Fix a non-absorbing state $x \in S$. By (6.13) and since $f^*_{\eta \breve{\mathcal{E}}_N^x}$ is bounded by 1, the expectation is less than or equal to $\mu_N(\Delta_N)/\mathrm{cap}_N(\eta, \breve{\mathcal{E}}_N^x)$. By Lemma 3.1, we may replace asymptotically η by $\mathcal{E}_N^x$ in the previous ratio. Since $\mathrm{cap}(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x) = \mu_N(\mathcal{E}_N^x) r_N(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x)$, we have shown that

$$\limsup_{N \to \infty} \sup_{\eta \in \mathcal{E}_N^x} \frac{1}{\theta_N} \mathbf{E}_\eta \big[H_{\breve{\mathcal{E}}_N^x}(\Delta_N) \big] \ \le \ \limsup_{N \to \infty} \frac{1}{\theta_N \, r_N(\mathcal{E}_N^x, \breve{\mathcal{E}}_N^x)} \frac{\mu_N(\Delta_N)}{\mu_N(\mathcal{E}_N^x)} \ .$$

As x is a non-absorbing point, by assumptions (**H0**), (**H2**), the right hand side vanishes, which concludes the proof of the theorem. $\qquad\square$

4. A FORMULA FOR THE AVERAGE JUMP RATE

Consider an irreducible chain $\{\eta(t) : t \ge 0\}$ on a countable state space E. Let $B \subset F$ be two non-empty subsets of E. Denote by $\{\eta^C(t) : t \ge 0\}$ the Markov process on $(E \setminus B) \cup \{\mathfrak{d}\}$, where $\mathfrak{d} \notin E$ is an extra pointed added to represent the collapsed set B, obtained from $\eta(t)$ by collapsing the set B to a point $\mathfrak{d}$ [11], and denote by $\{\eta^T(t) : t \ge 0\}$ the trace of the process $\eta(t)$ on the set F [1]. Denote furthermore by $\{\eta^{TC}(t) : t \ge 0\}$ the trace of the Markov process $\eta^C(t)$ on the set $G = (F \setminus B) \cup \{\mathfrak{d}\}$, and by $\{\eta^{CT}(t) : t \ge 0\}$ the process $\eta^T(t)$ where the set B has been collapsed to a point $\mathfrak{d}$. Note that the state space of both processes is $G = (F \setminus B) \cup \{\mathfrak{d}\}$

Lemma 4.1. *The processes $\{\eta^{TC}(t) : t \ge 0\}$ and $\{\eta^{CT}(t) : t \ge 0\}$ have the same law.*

Proof. Both processes are Markov processes on $G = (F \setminus B) \cup \{\mathfrak{d}\}$. To prove the assertion it is therefore enough to check that the jump rates are equal.

Denote by R_C, R_T, R_{CT}, R_{TC} the jump rates of the Markov processes $\eta^C(t)$, $\eta^T(t)$, $\eta^{CT}(t)$, $\eta^{TC}(t)$, respectively. Let $\eta \in F \setminus B$. We show that $R_{TC}(\mathfrak{d}, \eta) = R_{CT}(\mathfrak{d}, \eta)$ and leave the other identities to the reader.

By [1, Proposition 6.1], $R_{TC}(\mathfrak{d}, \eta) = \lambda_C(\mathfrak{d}) \mathbf{P}_{\mathfrak{d}}^C[H_G^+ = H_\eta^+]$, where $\mathbf{P}_{\mathfrak{d}}^C$ stands for the probability measure on the path space $D(\mathbb{R}_+, (E \setminus B) \cup \{\mathfrak{d}\})$ induced by the Markov process $\eta^C(t)$ starting from $\mathfrak{d}$. The probability $\mathbf{P}_{\mathfrak{d}}^C[H_G^+ = H_\eta^+]$ can be written as the sum of the probabilities of all paths connecting $\mathfrak{d}$ to η through points which do not belong to G:

$$\mathbf{P}_{\mathfrak{d}}^C \big[H_G^+ = H_\eta^+ \big] \ = \ \sum_\gamma p_C(\mathfrak{d}, \xi_1) \cdots p_C(\xi_{n-1}, \eta) \ ,$$

where the sum is carried over all paths $\gamma = (\xi_0 = \mathfrak{d}, \xi_1, \ldots, \xi_{n-1}, \xi_n = \eta)$ such that $\xi_j \notin G$, $1 \le j \le n-1$ or, equivalently, such that $\xi_j \in E \setminus F$. In this formula, p_C

stands for the jump probabilities of the collapsed chain. Hence,

$$R_{TC}(\mathfrak{d}, \eta) \;=\; \sum_{\gamma} R_C(\mathfrak{d}, \xi_1) p_C(\xi_1, \xi_2) \cdots p_C(\xi_{n-1}, \eta) \,.$$

By equation (3.3) in [11], $p_C(\xi, \zeta) = p(\xi, \zeta)$ if ξ, $\zeta \in E \setminus B$ and $R_C(\mathfrak{d}, \xi_1) = \mu(B)^{-1} \sum_{\zeta \in B} \mu(\zeta) R(\zeta, \xi_1)$. The previous expression is thus equal to

$$R_{TC}(\mathfrak{d}, \eta) \;=\; \frac{1}{\mu(B)} \sum_{\zeta \in B} \mu(\zeta) \sum_{\gamma} R(\zeta, \xi_1) p(\xi_1, \xi_2) \cdots p(\xi_{n-1}, \eta)$$

$$=\; \frac{1}{\mu(B)} \sum_{\zeta \in B} \mu(\zeta) \lambda(\zeta) \mathbf{P}_\zeta \big[H_F^+ = H_\eta^+ \big] \,.$$

On the other hand, by equation (3.3) in [11] and Proposition 6.1 in [1], for $\eta \in F \setminus B$,

$$R_{CT}(\mathfrak{d}, \eta) \;=\; \frac{1}{\mu(B)} \sum_{\zeta \in B} \mu(\zeta) R_T(\zeta, \eta) \;=\; \frac{1}{\mu(B)} \sum_{\zeta \in B} \mu(\zeta) \lambda(\zeta) \mathbf{P}_\zeta \big[H_F^+ = H_\eta^+ \big] \,.$$

This proves the lemma. $\qquad \square$

In the next result we consider an irreducible Markov process $\{\eta(t) : t \geq 0\}$ on a finite state space E. Let A_0, A_1, B be three disjoint subsets of E and let R_T denote the jump rates of $\eta^T(t)$, the trace of the process $\eta(t)$ on the set $F = A_0 \cup A_1 \cup B$. We denote by $r_T(B, A_i)$, $i = 0, 1$, the average rate at which the trace process jumps from B to A_i:

$$r_T(B, A_i) \;=\; \frac{1}{\mu(B)} \sum_{\eta \in B} \mu(\eta) \sum_{\xi \in A_i} R_T(\eta, \xi) \;=\; \frac{1}{\mu(B)} \sum_{\eta \in B} \mu(\eta) R_T(\eta, A_i) \,.$$

Proposition 4.2. *Consider the variational problem*

$$\inf_{f} \sup_{h} \big\{ 2\langle f, Lh \rangle_\mu - \langle h, (-L)h \rangle_\mu \big\}, \tag{4.1}$$

where the infimum is carried over all function $f : E \to \mathbb{R}$ which are constant at each set A_0, A_1, B, equal to 1 at A_1 and equal to 0 at A_0, while the supremum is carried over all functions h which are equal to 0 at A_0 and which are constant over A_1, B, with possibly different values at each set. Then, the optimal function h associated to the optimal function f is such that

$$h(B) \;=\; \frac{r_T(B, A_1)}{r_T(B, A_0) + r_T(B, A_1)} \,.$$

Proof. Denote by L_C the generator of the Markov process $\eta^C(t)$ on $E_B = (E \setminus B) \cup \{\mathfrak{d}\}$, where the set B has been collapsed to a point $\mathfrak{d} \notin E$. For a function $g : E \to \mathbb{R}$ constant over B, denote by $\overline{g} : E_B \to \mathbb{R}$ the function which is equal to g at $E \setminus B$, and such that $\overline{g}(\mathfrak{d}) = g(B)$.

Denote by $\overline{\mu}$ the probability measure on E_B equal to μ on $E \setminus B$ and such that $\overline{\mu}(\mathfrak{d}) = \mu(B)$. A calculation, performed below equation (3.7) in [11], shows that for any pair of functions f, $h : E \to \mathbb{R}$ constant over B,

$$\langle f, Lh \rangle_\mu \;=\; \langle \overline{f}, L_C \overline{h} \rangle_{\overline{\mu}} \,. \tag{4.2}$$

Conversely, given $\overline{f}$, $\overline{h} : E_B \to \mathbb{R}$, if we define f, $h : E \to \mathbb{R}$ to be equal to f, h on $E \setminus B$ and such that $f(\eta) = \overline{f}(\mathfrak{d})$, $h(\eta) = \overline{h}(\mathfrak{d})$, $\eta \in B$, (4.2) holds.

It follows from (4.2) that the variational problem (4.1) is equivalent to the variational problem
$$\inf_{\overline{f}} \sup_{\overline{h}} \left\{ 2\langle \overline{f}, L_C \overline{h} \rangle_{\overline{\mu}} - \langle \overline{h}, (-L_C) \overline{h} \rangle_{\overline{\mu}} \right\},$$
where the infimum is carried over all functions $\overline{f} : E_B \to \mathbb{R}$ which are equal to 1 at A_1 and equal to 0 at A_0, while the supremum is carried over all functions h which are constant over A_0, A_1. This variational problem corresponds to the variational problem for the capacity between A_1 and A_0 for the Markov process $\eta^C(t)$. By [11, Theorem 2.4], the optimal function $\overline{h} : E_B \to \mathbb{R}$ associated to the optimal function $\overline{f}$ is the harmonic function $\overline{h}(\eta) = \mathbf{P}_\eta^C[H_{A_1} < H_{A_0}]$. In particular, $\overline{h}(\mathfrak{d}) = \mathbf{P}_{\mathfrak{d}}^C[H_{A_1} < H_{A_0}]$ and the optimal function $h : E \to \mathbb{R}$ associated to the optimal function f for the variational problem (4.1) is such that
$$h(B) \;=\; \mathbf{P}_{\mathfrak{d}}^C\left[H_{A_1} < H_{A_0}\right]. \tag{4.3}$$

Decompose the event $\{H_{A_1} < H_{A_0}\}$ according to whether $H_{\mathfrak{d}}^+ < H_{A_1}$ or $H_{\mathfrak{d}}^+ > H_{A_1}$ and use the strong Markov property to get that
$$\mathbf{P}_{\mathfrak{d}}^C\left[H_{A_1} < H_{A_0}\right] \;=\; \frac{\mathbf{P}_{\mathfrak{d}}^C\left[H_{A_1} < \min\{H_{\mathfrak{d}}^+, H_{A_0}\}\right]}{1 - \mathbf{P}_{\mathfrak{d}}^C\left[H_{\mathfrak{d}}^+ < \min\{H_{A_1}, H_{A_0}\}\right]}.$$

The denominator can be written as $\mathbf{P}_{\mathfrak{d}}^C[H_{A_1} < \min\{H_{\mathfrak{d}}^+, H_{A_0}\}] + \mathbf{P}_{\mathfrak{d}}^C[H_{A_0} < \min\{H_{\mathfrak{d}}^+, H_{A_1}\}]$. Taking the trace of the process $\eta^C(t)$ on the set $A_0 \cup A_1 \cup \{\mathfrak{d}\}$, we get that
$$\mathbf{P}_{\mathfrak{d}}^C\left[H_{A_1} < H_{A_0}\right] \;=\; \frac{\mathbf{P}_{\mathfrak{d}}^{TC}\left[H_{A_1} < \min\{H_{\mathfrak{d}}^+, H_{A_0}\}\right]}{\mathbf{P}_{\mathfrak{d}}^{TC}\left[H_{A_0} < \min\{H_{\mathfrak{d}}^+, H_{A_1}\}\right] + \mathbf{P}_{\mathfrak{d}}^{TC}\left[H_{A_1} < \min\{H_{\mathfrak{d}}^+, H_{A_0}\}\right]}.$$

By Lemma 4.1, we may replace the probability $\mathbf{P}_{\mathfrak{d}}^{TC}$ by the probability $\mathbf{P}_{\mathfrak{d}}^{CT}$ in this formula. Since $\mathbf{P}_{\mathfrak{d}}^{TC}[H_{A_0} < \min\{H_{\mathfrak{d}}^+, H_{A_1}\}]$ is equal to $p_{CT}(\mathfrak{d}, A_0)$, where p_{CT} stands for the jump probabilities of the process $\eta^{CT}(t)$, multiplying the denominator and the numerator by $\lambda_{CT}(\mathfrak{d})$, the previous ratio becomes
$$\frac{R_{CT}(\mathfrak{d}, A_1)}{R_{CT}(\mathfrak{d}, A_0) + R_{CT}(\mathfrak{d}, A_1)}.$$

By [1, Proposition 6.1] and since the stationary measure for the trace process is the conditional measure, this expression is equal to
$$\frac{\mu(B)^{-1} \sum_{\eta \in B} \mu(\eta) R_T(\eta, A_1)}{\mu(B)^{-1} \sum_{\eta \in B} \mu(\eta) R_T(\eta, A_0) + \mu(B)^{-1} \sum_{\eta \in B} \mu(\eta) R_T(\eta, A_1)}.$$

By definition of $r_T(B, A_i)$ we conclude that
$$\mathbf{P}_{\mathfrak{d}}^C\left[H_{A_1} < H_{A_0}\right] \;=\; \frac{r_T(B, A_1)}{r_T(B, A_0) + r_T(B, A_1)}.$$

This identity concludes the proof of the proposition in view of (4.3). $\square$

5. Birth and death process

We prove in this subsection the metastable behavior of a class of birth and death processes which includes the zero range processes examined in [2] evolving on two sites. Proposition 5.1 presented below was used in [2] to prove the metastability of reversible zero range processes.

Fix $a < b$ in $\mathbb{R}$ and consider a nonnegative smooth function $H : [a, b] \to \mathbb{R}_+$. Assume that H vanishes only at a finite number of points denoted by $a_1 < a_2 < \cdots < a_m$:

$$H(x) = 0 \quad \text{if and only if} \quad x \in \{a_1, \ldots, a_m\} .$$

We do not exclude the possibility that H vanishes at the boundary points a, b.

For each $i = 1, \ldots, m$, assume that there exist a neighborhood V_{a_i} of a_i and $\alpha_i > 0$ such that

$$H(x) = |x - a_i|^{\alpha_i} \quad \text{for all } x \in V_{a_i} ,$$

and that $V_{a_i} \cap V_{a_j} = \varnothing$ for $i \neq j$. Let $\alpha = \max\{\alpha_i : 1 \leq i \leq m\}$. Assume that $\alpha > 1$ and that there are at least two exponents α_i equal to α:

$$\kappa := \big|\{i : \alpha_i = \alpha\}\big| \geq 2 ,$$

where $|A|$ indicates the cardinality of a finite set A. Denote by $b_1 < b_2 < \cdots < b_\kappa$ the elements of $\{a_1, \ldots, a_m\}$ whose associated exponents are α.

The definition of the state space of the birth and death process requires some notation. Fix $N \geq 1$, $1 \leq i \leq m - 1$, and let

$$k_0^N = \min\left\{k \geq 0 : a > a_1 - (k + 1)/N\right\} ,$$

$$k_i^N = \min\left\{k \geq 0 : a_i + (k + 1)/N > a_{i+1} - (k + 1)/N\right\} ,$$

$$k_m^N = \min\left\{k \geq 0 : a_m + (k + 1)/N > b\right\} .$$

For $1 \leq i \leq m - 1$, set

$$G_{N,0} = \left\{a_1 - k_0^N/N, \ldots, a_1 - 1/N, a_1\right\} ,$$

$$G_{N,i} = \left\{a_i + 1/N, a_i + 2/N, \ldots, a_i + k_i^N/N, a_{i+1} - k_i^N/N, \ldots, a_{i+1}\right\} ,$$

$$G_{N,m} = \left\{a_m + 1/N, a_m + 2/N, \ldots, a_m + k_m^N/N\right\}$$

and let the state space E_N be $\bigcup_{i=0}^m G_{N,i}$. Note that the exact definition of E_N is not important for the meta-stability behavior discussed in this section. The elements of E_N are denoted by the letters x, y, z. Two points $x < y$ are said to be neighbors in E_N if there is no z in E_N such that $x < z < y$.

Let ν_N be the probability measure on E_N defined by

$$\nu_N(x) = \begin{cases} \dfrac{1}{Z_N} \dfrac{1}{H(x)} & \text{if } x \notin \{a_1, \ldots, a_m\} , \\ \dfrac{1}{Z_N} N^{\alpha_i} & \text{if } x = a_i \text{ for some } 1 \leq i \leq m . \end{cases}$$

In this formula Z_N is a normalizing constant. An elementary computation shows that

$$\lim_{N \to \infty} \frac{Z_N}{N^\alpha} = \sum_{i=1}^\kappa \left\{1 + \sigma_i \sum_{k \geq 1} \frac{1}{k^\alpha}\right\} . \tag{5.1}$$

where $\sigma_i = 1$ if $b_i \in \{a, b\}$ and $\sigma_i = 2$ otherwise. In particular, if we denote

$$m(b_i) = 1 + \sigma_i \sum_{k \geq 1} \frac{1}{k^\alpha} , \quad \text{for } i = 1, \ldots, \kappa$$

then

$$\lim_{N \to \infty} \nu_N(b_i) = \left\{\sum_{j=1}^\kappa m(b_j)\right\}^{-1} > 0$$

for every $1 \le i \le \kappa$.

Let $\{\ell_N : N \ge 1\}$ be a sequence of positive integers such that $1 \ll \ell_N \ll N$,

$$\lim_{N \to \infty} \frac{\ell_N}{N} = 0 \quad \text{and} \quad \lim_{N \to \infty} \ell_N = \infty , \tag{5.2}$$

and, for each b_i, $1 \le i \le \kappa$, define

$$\mathcal{E}_N^i := E_N \cap \left[b_i - \frac{\ell_N}{N}, b_i + \frac{\ell_N}{N} \right].$$

Since $N^{-1}\ell_N \to 0$, for N large enough $\mathcal{E}_N^i \subseteq V_{b_i}$ for every $1 \le i \le \kappa$. In particular, for N large enough, $\mathcal{E}_N^i \cap \mathcal{E}_N^j = \varnothing$ for all $i \ne j$. Moreover, since $\ell_N \to \infty$,

$$\lim_{N \to \infty} \nu_N(\Delta_N) = 0 \quad \text{and} \quad \lim_{N \to \infty} \nu_N(\mathcal{E}_N^i) = \frac{m(b_i)}{\sum_{j=1}^{\kappa} m(b_j)} \tag{5.3}$$

for all $1 \le i \le \kappa$, where $\Delta_N := E_N \setminus (\cup_{i=1}^{\kappa} \mathcal{E}_N^i)$.

Fix a positive function $\Phi : [a, b] \to \mathbb{R}_+$ bounded above and below by a strictly positive constant:

$$0 < \delta \le \Phi(x) \le \delta^{-1} .$$

This assumption is not needed but we do not seek optimal assumptions. Consider a birth and death process $\{Z_t^N : t \ge 0\}$ on E_N with rates given by

$$R_N(x, y) = \begin{cases} \Phi(x) & \text{if } x > y, \\ \Phi(y) \dfrac{\nu_N(y)}{\nu_N(x)} & \text{if } x < y , \end{cases}$$

provided x and y are neighbors in E_N. The process is of course reversible with respect to the measure ν_N.

Proposition 5.1. *The Markov process $\{Z_t^N : N \ge 1\}$ exhibits a tunneling behavior in the sense of [1, Definition 2.2] on the time scale $N^{1+\alpha}$ with metastates $\mathcal{E}_N^i$, metapoints b_i, $1 \le i \le \kappa$, and asymptotic Markovian dynamics on $\{1, \dots, \kappa\}$ characterized by the rates*

$$r(i, i+1) = \frac{1}{m(b_i)} \frac{1}{\int_{b_i}^{b_{i+1}} \{H(u)/\Phi(u)\} \, du} ,$$

$$r(i+1, i) = \frac{1}{m(b_{i+1})} \frac{1}{\int_{b_i}^{b_{i+1}} \{H(u)/\Phi(u)\} \, du} , \quad 1 \le i < \kappa .$$

Proof. We show that the hypotheses of [1, Theorem 2.7] are in force. Let $\xi_N^i = b_i$ for $1 \le i \le \kappa$. The asymptotic dynamics has no absorbing point and condition **(H2)** of [1, Theorem 2.7] follows from (5.3).

To check conditions **(H0)**, **(H1)** we take advantage from the one-dimensional setting to get explicit expressions for capacities. For two disjoint subsets A, B of E_N, denote by $\mathrm{cap}_N(A, B)$ the capacity between A and B. When $A = \{a\}$ we represent $\mathrm{cap}_N(A, B)$ by $\mathrm{cap}_N(a, B)$ with the same convention for B. Let $x < y$ be points in E_N. Recall that $\mathrm{cap}_N(x, y) = D_N(f_{x,y})$ where $f_{x,y} : E_N \mapsto \mathbb{R}$ solves the equation $L_N f_{x,y}(z) = 0$ for $z \notin \{x, y\}$ with boundary conditions $f_{x,y}(x) = 1$ and $f_{x,y}(y) = 0$. An elementary computation gives that $f(z) = 1$ for $z \le x$, $f(z) = 0$ for $z \ge y$ and

$$f(z + 1/N) - f(z) = \frac{\{\nu_N(z)R_N(z, z+1/N)\}^{-1}}{\sum_{z=x}^{y-1/N} \{\nu_N(z)R_N(z, z+1/N)\}^{-1}}$$

for $z \in E_N \cap [x, y)$. Hence,

$$\mathrm{cap}_N(x, y) = \frac{1}{\sum_{z=x}^{y-1/N} \left\{ \nu_N(z) R_N(z, z+1/N) \right\}^{-1}} . \qquad (5.4)$$

In last two formulae, there is a slight abuse of notation since E_N is not the set $\{z/N : z \in \mathbb{Z} \cap [aN, bN]\}$, but the meaning is clear. In particular, if $\{x^N : N \geq 1\}$, $\{y^N : N \geq 1\}$ are two sequences in E_N such that $x^N \to a'$ and $y^N \to b'$ for some $a \leq a' < b' \leq b$, by (5.4),

$$\lim_{N \to \infty} N^{1+\alpha} \mathrm{cap}_N(x^N, y^N) = \frac{1}{\sum_{i=1}^{\kappa} m(b_i)} \left\{ \int_{a'}^{b'} \{H(u)/\Phi(u)\} \, du \right\}^{-1} > 0 . \qquad (5.5)$$

Denote by $r_N(\mathcal{E}_N^i, \mathcal{E}_N^j)$ the average jump rate from $\mathcal{E}_N^i$ to $\mathcal{E}_N^j$ of the trace of the process Z_t^N on $\mathcal{E}_N = \cup_{1 \leq m \leq \kappa} \mathcal{E}_N^m$, defined by (2.4) in a general context. Clearly, $r_N(\mathcal{E}_N^i, \mathcal{E}_N^j) = 0$ for $|i - j| > 1$. Fix an arbitrary $1 \leq i < \kappa$ and let $G_1 = \cup_{j \leq i} \mathcal{E}_N^j$, and $G_2 = \cup_{j > i} \mathcal{E}_N^j$ so that

$$\nu_N(\mathcal{E}_N^i) \, r_N(\mathcal{E}_N^i, \mathcal{E}_N^{i+1}) = \nu_N(b_i + \frac{\ell_N}{N}) \, R_N^{\mathcal{E}}(b_i + \frac{\ell_N}{N}, b_{i+1} - \frac{\ell_N}{N})$$
$$= \nu_N(G_1) \, r_N(G_1, G_2) ,$$

where $R_N^{\mathcal{E}}(x, y)$, $x \neq y \in \mathcal{E}_N$, represents the jumps rates of the trace of Z_t^N on $\mathcal{E}_N$. Therefore, by (6.8),

$$r_N(\mathcal{E}_N^i, \mathcal{E}_N^{i+1}) = \frac{\nu_N(G_1) \, r_N(G_1, G_2)}{\nu_N(\mathcal{E}_N^i)} = \frac{\mathrm{cap}_N(G_1, G_2)}{\nu_N(\mathcal{E}_N^i)}$$
$$= \frac{\mathrm{cap}_N(b_i + \ell_N/N, b_{i+1} - \ell_N/N)}{\nu_N(\mathcal{E}_N^i)} .$$

Analogously, we obtain that

$$r_N(\mathcal{E}_N^i, \mathcal{E}_N^{i-1}) = \frac{\mathrm{cap}_N(b_i - \ell_N/N, b_{i-1} + \ell_N/N)}{\nu_N(\mathcal{E}_N^i)}$$

for any $1 < i \leq \kappa$. Therefore, by (5.5) and (5.3),

$$\lim_{N \to \infty} N^{1+\alpha} r_N(\mathcal{E}_N^i, \mathcal{E}_N^{i+1}) = \frac{1}{m(b_i)} \left\{ \int_{b_i}^{b_{i+1}} \{H(u)/\Phi(u)\} \, du \right\}^{-1}$$

and

$$\lim_{N \to \infty} N^{1+\alpha} r_N(\mathcal{E}_N^i, \mathcal{E}_N^{i-1}) = \frac{1}{m(b_{i+1})} \left\{ \int_{b_i}^{b_{i+1}} \{H(u)/\Phi(u)\} \, du \right\}^{-1}$$

for any $1 \leq i < \kappa$, which concludes the proof of assumption (**H0**).

The same arguments show that

$$\lim_{N \to \infty} N^{1+\alpha} \, \mathrm{cap}_N(\mathcal{E}_N^i, \breve{\mathcal{E}}_N^i)$$
$$= \frac{1}{\sum_{1 \leq i \leq \kappa} m(b_i)} \left\{ \frac{1}{\int_{b_i}^{b_{i+1}} \{H(u)/\Phi(u)\} \, du} + \frac{1}{\int_{b_{i-1}}^{b_i} \{H(u)/\Phi(u)\} \, du} \right\} , \qquad (5.6)$$

provided $b_i \neq a$, b, with similar identities if $b_i = a$ or if $b_i = b$.

It remains to check condition **(H1)**. For any $1 \le i \le \kappa$ and N large enough, $H(x) = |x - b_i|^\alpha$ for all $x \in \mathcal{E}_N^i$. In consequence, by (5.1) and (5.4), there exists a positive constant C_0, independent of N, such that

$$\text{cap}_N(x, b_i) \ge \frac{C_0}{\ell_N^{\alpha+1}} \,,$$

for any $1 \le i \le \kappa$ and $x \in \mathcal{E}_N^i$. Therefore, since $\ell_N \ll N$, by (5.6),

$$\lim_{N\to\infty} \sup_{x \in \mathcal{E}_N^i} \frac{\text{cap}_N(\mathcal{E}_N^i, \breve{\mathcal{E}}_N^i)}{\text{cap}_N(x, b_i)} \;=\; 0$$

for all $1 \le i \le \kappa$, which concludes the proof of the proposition. $\qquad\square$

6. Potential theory for positive recurrent processes

We state in this section several properties of continuous time Markov chains used throughout the article. Consider a countable set E and a matrix $R : E \times E \to \mathbb{R}$ such that $R(\eta, \xi) \ge 0$, $\eta \ne \xi$, $-\infty < R(\eta, \eta) < 0$, $\sum_\xi R(\eta, \xi) = 0$, $\eta \in E$. Let $\lambda(\eta) = -R(\eta, \eta)$. Since $\lambda(\eta)$ is finite and strictly positive, we may define the transition probabilities $\{p(\eta, \xi) : \eta, \xi \in E\}$ as

$$p(\eta, \xi) \;=\; \frac{1}{\lambda(\eta)} \, R(\eta, \xi) \quad \text{for } \eta \ne \xi \,, \tag{6.1}$$

and $p(\eta, \eta) = 0$ for $\eta \in E$. We assume throughout this section that $\{p(\eta, \xi) : \eta, \xi \in E\}$ are the transition probabilities of an irreducible and recurrent discrete time Markov chain, denoted by $Y = \{Y_n : n \ge 0\}$.

Let $\{\eta(t) : t \ge 0\}$ be the unique strong Markov process associated to the rates $R(\eta, \xi)$. We shall refer to $R(\cdot, \cdot)$, $\lambda(\cdot)$ and $p(\cdot, \cdot)$ as the transition rates, holding rates and jump probabilities of $\{\eta(t) : t \ge 0\}$, respectively. Since the jump chain Y is irreducible and recurrent, so is the corresponding Markov process $\{\eta(t) : t \ge 0\}$. We shall assume throughout this section that $\eta(t)$ is *positive recurrent*. In consequence, $\eta(t)$ has a unique invariant probability measure μ. Moreover,

$$M(\eta) \;:=\; \lambda(\eta)\mu(\eta) \,, \quad \eta \in E \,, \tag{6.2}$$

is an invariant measure for the jump chain Y, unique up to scalar multiples. The proofs of these assertions can be found in Sections 3.4 and 3.5 of [16]. We assume furthermore that the holding rates are summable with respect to μ:

$$\sum_{\eta \in E} \lambda(\eta) \, \mu(\eta) \;<\; \infty \,, \tag{6.3}$$

so that M is a finite measure. Assumption (6.3) reduces the potential theory of continuous time Markov chains to the potential theory of discrete time Markov chains.

Let $L^2(\mu)$, $L^2(M)$ be the space of square integrable functions $f : E \to \mathbb{R}$ endowed with the usual scalar product $\langle f, g \rangle_m = \sum_{\eta \in E} f(\eta)g(\eta)m(\eta)$, with $m = \mu$, M, respectively. Denote by P the bounded operator in $L^2(M)$ defined by

$$(Pf)(\eta) \;=\; \sum_{\xi \in E} p(\eta, \xi) \{f(\xi) - f(\eta)\} \tag{6.4}$$

for $f \in L^2(M)$, and by L the generator of the Markov process $\{\eta(t) : t \geq 0\}$. Thus, for every finitely supported function $f : E \to \mathbb{R}$,

$$(Lf)(\eta) \;=\; \sum_{\xi \in E} R(\eta, \xi) \left\{ f(\xi) - f(\eta) \right\} .$$

Let L^* be the adjoint of the operator L on $L^2(\mu)$. L^* is the generator of a Markov process $\{\eta^*(t) : t \geq 0\}$ with holding rates λ^* and jump rates R^* given by $\lambda^*(\eta) = \lambda(\eta)$, $R^*(\eta, \xi) = \mu(\xi)R(\xi, \eta)/\mu(\eta)$, so that $\mu(\eta)R^*(\eta, \xi) = \mu(\xi)R(\xi, \eta)$. Let P^* be the adjoint of P in $L^2(M)$. Clearly, the bounded operator P^* is given by (6.4) with p^* in place of p. Denote by $\{Y_n^* : n \geq 0\}$ the discrete time Markov chain associated to P^*.

Similarly, let S be the symmetric part of the operator L on $L^2(\mu)$. S is the generator of a Markov process $\{\eta^s(t) : t \geq 0\}$ with holding rates λ^s and jump rates R^s given by $\lambda^s(\eta) = \lambda(\eta)$, $R^s(\eta, \xi) = (1/2)\{R(\eta, \xi) + R^*(\eta, \xi)\}$, so that $\mu(\eta)R^s(\eta, \xi) = \mu(\xi)R^s(\xi, \eta)$.

Let $\mathbf{P}_\eta$, $\mathbf{P}_\eta^*$, $\mathbf{P}_\eta^s$, $\eta \in E$, be the probability measure on the path space $D(\mathbb{R}_+, E)$ of right continuous trajectories with left limits induced by the Markov process $\{\eta(t) : t \geq 0\}$, $\{\eta^*(t) : t \geq 0\}$, $\{\eta^s(t) : t \geq 0\}$ starting from η, respectively. Expectation with respect to $\mathbf{P}_\eta$, $\mathbf{P}_\eta^*$, $\mathbf{P}_\eta^s$ are denoted by $\mathbf{E}_\eta$, $\mathbf{E}_\eta^*$, $\mathbf{E}_\eta^s$, respectively.

Denote by H_A (resp. H_A^+), $A \subseteq E$, the hitting time of (resp. return time to) the set A:

$$H_A \;=\; \inf\{t > 0 : \eta(t) \in A\} ,$$
$$H_A^+ \;=\; \inf\{t > 0 : \eta(t) \in A, \eta(s) \neq \eta(0) \text{ for some } 0 < s < t\} .$$

Let F be a proper subset of E. Denote by $\{\mathcal{T}_t : t \geq 0\}$ the time spent on the set F by the process $\eta(s)$ in the time interval $[0, t]$:

$$\mathcal{T}_t \;:=\; \int_0^t \mathbf{1}\{\eta(s) \in F\} \, ds .$$

Notice that $\mathcal{T}_t \in \mathbb{R}_+$, $\mathbf{P}_\eta$-a.s. for every $\eta \in E$ and $t \geq 0$. Denote by $\{\mathcal{S}_t : t \geq 0\}$ the generalized inverse of $\mathcal{T}_t$:

$$\mathcal{S}_t \;:=\; \sup\{s \geq 0 : \mathcal{T}_s \leq t\} .$$

Since $\{\eta(t) : t \geq 0\}$ is irreducible and recurrent, $\lim_{t \to \infty} \mathcal{T}_t = \infty$, $\mathbf{P}_\eta$-a.s. for every $\eta \in E$. Therefore, the random path $\{\eta^F(t) : t \geq 0\}$, given by $\eta^F(t) = \eta(\mathcal{S}_t)$, is $\mathbf{P}_\eta$-a.s. well defined for all $\eta \in E$ and takes value in the set F. We call the process $\{\eta^F(t) : t \geq 0\}$ the trace of $\{\eta(t) : t \geq 0\}$ on the set F.

Denote by $R^F(\eta, \xi)$ the jump rates of the trace process $\{\eta^F(t) : t \geq 0\}$. By Propositions 6.1 and 6.3 in [1], $\{\eta^F(t) : t \geq 0\}$ is an irreducible, recurrent strong Markov process whose invariant measure μ^F is given by

$$\mu^F(\xi) \;=\; \frac{1}{\mu(F)} \mu(\xi) , \quad \xi \in F . \tag{6.5}$$

For each pair A, B of disjoint subsets of F, denote by $r_F(A, B)$ the average rate at which the trace process jumps from A to B:

$$r_F(A, B) \;:=\; \frac{1}{\mu(A)} \sum_{\eta \in A} \mu(\eta) \sum_{\xi \in B} R^F(\eta, \xi) .$$

By [1, Proposition 6.1],

$$r_F(A, B) \;=\; \frac{1}{\mu(A)} \sum_{\eta \in A} M(\eta) \, \mathbf{P}_\eta \big[\, H_F^+ = H_B^+ \,\big] \,. \tag{6.6}$$

We shall refer to $r_F(\cdot, \cdot)$ as the mean set rates associated to the trace process.

Recall from [11] that the capacity between two disjoint subsets A, B of E, denoted by $\mathrm{cap}(A, B)$, is defined as

$$\mathrm{cap}(A, B) \;:=\; \sum_{\eta \in A} M(\eta) \, \mathbf{P}_\eta \big[\, H_B^+ < H_A^+ \,\big] \,. \tag{6.7}$$

Hence, by (6.6) for any two disjoint subsets A, B of E,

$$\mathrm{cap}(A, B) \;=\; \mu(A) \, r_{A \cup B}(A, B) \,. \tag{6.8}$$

Let $\mathrm{cap}^*(A, B)$, $\mathrm{cap}^s(A, B)$ be the capacity between two disjoint subsets A, B of E for the adjoint, symmetric process, respectively. By equation (2.4) and Lemmata 2.3, 2.5 in [11],

$$\mathrm{cap}^*(A, B) \;=\; \mathrm{cap}(B, A) \;=\; \mathrm{cap}(A, B) \,, \quad \mathrm{cap}^s(A, B) \;\leq\; \mathrm{cap}(A, B) \,. \tag{6.9}$$

Denote by cap_F the capacity with respect to the trace process $\eta^F(t)$. By the proof of [1, Lemma 6.9], for every subsets A, B of F, $A \cap B = \varnothing$,

$$\mu(F) \, \mathrm{cap}_F(A, B) \;=\; \mathrm{cap}(A, B) \,. \tag{6.10}$$

Next result presents an identity between the capacities $\mathrm{cap}^s(A, B)$ and $\mathrm{cap}(A, B)$, somehow surprising in view of inequality (6.9).

Lemma 6.1. *Let A, B, C be three disjoint subsets of E. Then,*

$$\mathrm{cap}(A, B \cup C) \;+\; \mathrm{cap}(B, A \cup C) \;-\; \mathrm{cap}(C, A \cup B)$$
$$=\; \mathrm{cap}^s(A, B \cup C) \;+\; \mathrm{cap}^s(B, A \cup C) \;-\; \mathrm{cap}^s(C, A \cup B) \,.$$

Proof. Taking $F = A \cup B \cup C$ in (6.10), we may assume that A, B, C forms a partition of E. In this case, since $\mathrm{cap}(C, A \cup B) = \mathrm{cap}(A \cup B, C)$ and since $E = A \cup B \cup C$, by (6.8) the left hand side of the previous equation is equal to

$$\sum_{\eta \in A} \mu(\eta) R(\eta, B \cup C) \;+\; \sum_{\eta \in B} \mu(\eta) R(\eta, A \cup C) \;-\; \sum_{\eta \in A \cup B} \mu(\eta) R(\eta, C)$$
$$=\; \sum_{\eta \in A} \mu(\eta) R(\eta, B) \;+\; \sum_{\eta \in B} \mu(\eta) R(\eta, A) \;=\; 2 \sum_{\eta \in A} \mu(\eta) R^s(\eta, B) \,,$$

where $R(\xi, D) = \sum_{\zeta \in D} R(\xi, \zeta)$. Performing the computations backward with R^s in place of R we conclude the proof. $\qquad\square$

We conclude this section proving a relation between expectations of time integrals of functions and capacities. Fix two disjoint subsets A, B of E. Denote by f_{AB}, $f_{AB}^* : E \to \mathbb{R}$ the harmonic functions defined as

$$f_{AB}(\eta) \;:=\; \mathbf{P}_\eta \big[\, H_A < H_B \,\big] \,, \quad f_{AB}^*(\eta) \;:=\; \mathbf{P}_\eta^* \big[\, H_A < H_B \,\big] \,.$$

An elementary computation shows that f_{AB} solves the equation

$$\begin{cases} (Lf)(\eta) = 0 & \eta \in E \setminus (A \cup B) \,, \\ f(\eta) = 1 & \eta \in A \,, \\ f(\eta) = 0 & \eta \in B \,. \end{cases} \tag{6.11}$$

and that f_{AB}^* solves the same equation with the adjoint L^* replacing L. Clearly, we may replace the generator L by the operator $I - P$ in the above equation, and (6.11) has a unique solution in $L^2(M)$ given by f_{AB}.

Define the harmonic measure ν_{AB}, ν_{AB}^* on A as

$$\nu_{AB}(\eta) \;=\; \frac{M(\eta)\,\mathbf{P}_\eta\big[\,H_B^+ < H_A^+\,\big]}{\mathrm{cap}(A,B)}\,, \quad \nu_{AB}^*(\eta) \;=\; \frac{M(\eta)\,\mathbf{P}_\eta^*\big[\,H_B^+ < H_A^+\,\big]}{\mathrm{cap}^*(A,B)} \quad \eta \in A\,.$$

Denote by $\mathbf{E}_{\nu_{AB}}$ the expectation associated to the Markov process $\{\eta(t) : t \geq 0\}$ with initial distribution ν_{AB}.

Proposition 6.2. *Fix two disjoint subsets A, B of E. Let $g : E \to \mathbb{R}$ be a μ-integrable function. Then,*

$$\mathbf{E}_{\nu_{AB}^*}\Big[\int_0^{H_B} g(\eta(t))\,dt \Big] \;=\; \frac{\langle g\,, f_{AB}^*\rangle_\mu}{\mathrm{cap}(A,B)}\,, \tag{6.12}$$

where $\langle \cdot, \cdot \rangle_\mu$ represents the scalar product in $L^2(\mu)$.

Proof. We first claim that the proposition holds for indicator functions of states. Fix an arbitrary state $\xi \in E$. If ξ belongs to B the right hand the left hand side of (6.12) vanish. We may therefore assume that ξ does not belong to B. In this case we may write the expectation appearing in the statement of the lemma as

$$\mathbf{E}_{\nu_{AB}^*}\Big[\sum_{n=0}^{\mathbb{H}_B - 1} \frac{e_n}{\lambda(\xi)}\,\mathbf{1}\{Y_n = \xi\} \Big]\,,$$

where $\{Y_n : n \geq 0\}$ is the discrete time embedded Markov chain, $\{e_n : n \geq 0\}$ is a sequence of i.i.d. mean one exponential random variables independent of the jump chain $\{Y_n : n \geq 0\}$, and $\mathbb{H}_B$ the hitting time of the set B for the discrete time Markov chain Y_n. By the Markov property and by definition of the harmonic measure ν_{AB}^*, this expression is equal to

$$\frac{1}{\lambda(\xi)} \sum_{\eta \in A} \sum_{n \geq 0} \nu_{AB}^*(\eta)\,\mathbf{P}_\eta\big[Y_n = \xi\,, n < \mathbb{H}_B\big]$$

$$= \; \frac{1}{\lambda(\xi)\,\mathrm{cap}^*(A,B)} \sum_{n \geq 0} \sum_{\eta \in A} M(\eta)\,\mathbf{P}_\eta^*\big[\,H_B < H_A^+\,\big]\mathbf{P}_\eta\big[Y_n = \xi\,, n < \mathbb{H}_B\big]\,.$$

We may replace the hitting time and the return time H_B, H_A^+ by the respective times $\mathbb{H}_B$, $\mathbb{H}_A^+$ for the discrete chain. On the other hand, since η and ξ do not belong to B, the event $\{Y_0 = \eta\,, Y_n = \xi\,, n < \mathbb{H}_B\}$ represents all paths that started from η, reached ξ at time n without passing through B. In particular, by the detailed balanced relations between the process and its adjoint, $M(\eta)\,\mathbf{P}_\eta[Y_n = \xi\,, n < \mathbb{H}_B] = M(\xi)\,\mathbf{P}_\xi^*[Y_n = \eta\,, n < \mathbb{H}_B]$ and the last sum becomes

$$\frac{M(\xi)}{\lambda(\xi)\,\mathrm{cap}^*(A,B)} \sum_{n \geq 0} \sum_{\eta \in A} \mathbf{P}_\eta^*\big[\,\mathbb{H}_B < \mathbb{H}_A^+\,\big]\mathbf{P}_\xi^*\big[Y_n = \eta\,, n < \mathbb{H}_B\big]$$

$$= \; \frac{M(\xi)}{\lambda(\xi)\,\mathrm{cap}^*(A,B)} \sum_{n \geq 0} \sum_{\eta \in A} \mathbf{P}_\xi^*\big[Y_n = \eta\,, n < \mathbb{H}_B\,, \mathbb{H}_B \circ \theta_n < \mathbb{H}_A^+ \circ \theta_n\big]\,,$$

where we used the Markov property in the last step. In this formula $\{\theta_k : k \geq 1\}$ stands for the group of discrete time shift. Summing over η the sum can be written

as

$$\frac{M(\xi)}{\lambda(\xi)\,\mathrm{cap}^*(A,B)}\sum_{n\geq 0}\mathbf{P}^*_\xi\big[Y_n\in A,\, n<\mathbb{H}_B,\, \mathbb{H}_B\circ\theta_n<\mathbb{H}^+_A\circ\theta_n\big]$$

The set inside the probability represents the event that the process Y_k visits A before visiting B and that its last visit to A before reaching B occurs at time n. Hence, since $M(\xi)=\lambda(\xi)\mu(\xi)$, since by (6.9) $\mathrm{cap}^*(A,B)=\mathrm{cap}(A,B)$ and since g is the indicator of the state ξ, summing over n we get that the previous expression is equal to

$$\frac{1}{\mathrm{cap}^*(A,B)}\,\mu(\xi)\,\mathbf{P}^*_\xi\big[\mathbb{H}_A<\mathbb{H}_B\big]\;=\;\frac{\langle g\,,\,f^*_{AB}\rangle_\mu}{\mathrm{cap}(A,B)}\;.$$

By linearity and the monotone convergence theorem we get the desired result for positive and then μ-integrable functions. $\square$

In the particular case where $A=\{\eta\}$ for $\eta\notin B$ we have that

$$\mathbf{E}_\eta\Big[\int_0^{H_B}g(\eta(s))\,ds\Big]\;=\;\frac{\langle g\,,\,f^*_{\{\eta\}B}\rangle_\mu}{\mathrm{cap}(\{\eta\},B)}\tag{6.13}$$

for any μ-integrable function g. This formula provides an estimation for the thermalization.

Let S be a finite set, let $\pi=\{A^x:x\in S\}$ be a partition of E, and let ξ_x be a state in A^x for each $x\in S$. For each μ-integrable function g denote by $\langle g|\pi\rangle_\mu:E\to\mathbb{R}$ the conditional expectation of g, under μ, given the σ-algebra generated by π:

$$\langle g|\pi\rangle_\mu\;=\;\sum_{x\in S}\frac{\langle g\,\mathbf{1}\{A^x\}\rangle_\mu}{\mu(A^x)}\,\mathbf{1}\{A^x\}\,.$$

For each $x\in S$, let

$$\mathrm{cap}(\xi_x)\;:=\;\inf_{\eta\in A^x\setminus\{\xi_x\}}\mathrm{cap}(\{\eta\},\{\xi_x\})\,.$$

The next result shows that if the process thermalizes quickly in each set of the partition, we may replace time averages of a bounded function by time averages of the conditional expectation. This statement plays a key role in the investigation of metastability. It assumes, however, the existence of an attractor.

Corollary 6.3. *Let $g:E\to\mathbb{R}$ be a μ-integrable function. Then, for every $t>0$,*

$$\sup_{\eta\in E}\Big|\mathbf{E}_\eta\Big[\int_0^t\{g-\langle g|\pi\rangle_\mu\}(\eta(s))\,ds\Big]\Big|\;\leq\;4\sum_{x\in S}\frac{\langle |g|\,\mathbf{1}\{A^x\}\rangle_\mu}{\mathrm{cap}(\xi_x)}\,,$$

where $|g|(\eta)=|g(\eta)|$ for all η in E.

The proof of this result follows from [1, Corollary 6.5], formula (6.13) and the fact that f^*_{AB} is bounded by one.

REFERENCES

[1] J. Beltrán, C. Landim: Tunneling and metastability of continuous time Markov chains. J. Stat. Phys. **140** 1065–1114 (2010).

[2] J. Beltrán, C. Landim: Metastability of reversible condensed zero range processes on a finite set. Probab. Th. Rel. Fields **152** 781–807 (2012)

[3] J. Beltrán, C. Landim; Metastability of reversible finite state Markov processes. Stoch. Proc. Appl. **121** 1633–1677 (2011).

[4] J. Beltrán, C. Landim; Tunneling of the Kawasaki dynamics at low temperatures in two dimensions. arXiv:1109.2776 (2011).

[5] A. Bovier, M. Eckhoff, V. Gayrard, M. Klein. Metastability in stochastic dynamics of disordered mean field models. Probab. Theory Relat. Fields **119**, 99-161 (2001)

[6] A. Bovier, M. Eckhoff, V. Gayrard, M. Klein. Metastability and low lying spectra in reversible Markov chains. Commun. Math. Phys. **228**, 219–255 (2002).

[7] M. Cassandro, A. Galves, E. Olivieri, M. E. Vares. Metastable behavior of stochastic dynamics: A pathwise approach. J. Stat. Phys. **35**, 603–634 (1984).

[8] K. L. Chung, *Markov chains with stationary transition probabilities*. Second edition. Die Grundlehren der mathematischen Wissenschaften, Band 104 Springer-Verlag New York, Inc., New York 1967.

[9] David Freedman. *Markov chains*. Holden-Day, San Francisco (1971).

[10] A. Gaudillière. Condenser physics applied to Markov chains: A brief introduction to potential theory. Online available at http://arxiv.org/abs/0901.3053.

[11] A. Gaudillière, C. Landim: A Dirichlet principle for non reversible Markov chains and some recurrence theorems. arXiv:1111.2445 (2011).

[12] M. Jara, C. Landim, A. Teixeira; Quenched scaling limits of trap models. Ann. Probab. **39**, 176–223 (2011).

[13] M. Jara, C. Landim, A. Teixeira; Quenched scaling limits of trap models in random graphs. preprint (2012).

[14] C. Landim: Metastability for a non-reversible dynamics: the evolution of the condensate in totally asymmetric zero range processes. arXiv:1204.5987

[15] J. L. Lebowitz, O. Penrose: Rigorous treatment of the van der Waals–Maxwell theory of the liquid–vapor transition. J. Math. Phys. **7**, 98–113 (1966).

[16] J. R. Norris. *Markov chains*. Cambridge University Press, Cambridge (1997).

[17] E. Olivieri and M. E. Vares. *Large deviations and metastability*. Encyclopedia of Mathematics and its Applications, vol. 100. Cambridge University Press, Cambridge, 2005.

IMCA, Calle los Biólogos 245, Urb. San César Primera Etapa, Lima 12, Perú and PUCP, Av. Universitaria cdra. 18, San Miguel, Ap. 1761, Lima 100, Perú.
E-mail: johel@impa.br

IMPA, Estrada Dona Castorina 110, CEP 22460 Rio de Janeiro, Brasil and CNRS UMR 6085, Université de Rouen, Avenue de l'Université, BP.12, Technopôle du Madrillet, F76801 Saint-Étienne-du-Rouvray, France.
E-mail: landim@impa.br