Self-Similar Markov Processes on Cantor Set

Yuri Bakhtin*

October 17, 2008

Abstract

We define analogues of Brownian motion on the triadic Cantor set by introducing a few natural requirements on the Markov semigroup. We give a detailed description of these symmetric self-similar processes and study their properties such as mixing and moment asymptotics.

1 Introduction

In a recent paper [PB08], the authors use noncommutative geometry to describe an analogue of the Riemannian structure for ultrametric Cantor sets. This eventually leads them to a definition of a certain family of operators that play the role of the Laplace–Beltrami operators for Cantor sets. It is natural to treat these operators as infinitesimal generators of Markov semigroups on Cantor sets and call the associated Cantor-set-valued stochastic processes to be the analogues of the Brownian motion.

The goal of this note is to provide an alternative definition of Brownian motion on the classical triadic Cantor set. We use the axiomatic method and describe several natural requirements, most important of which are isometry invariance and scale invariance, that should hold for a reasonable analogue of Brownian motion on the Cantor set. Then we give a complete description of Markov processes satisfying these requirements. We call these processes symmetric self-similar (SSS).

The parametrization of SSS processes involves two degrees of freedom. One of these is responsible just for uniform time changes, so that effectively this family of SSS processes is parametrized by one parameter of scaling, or self-similarity. Since we are basing our approach on scaling properties, the SSS processes are, in fact, analogous to symmetric stable Lévy processes in $\mathbb{R}$. All processes on the tridaic Cantor set described in [PB08] are SSS, but the converse is not true.

Our approach is somewhat similar to Schramm's celebrated characterization of SLE via conformal invariance and Markov property, see e.g. [Law05, Chapter 6].

*School of Mathematics, Georgia Tech, Atlanta GA, 30332-0160; email:bakhtin@math.gatech.edu, 404-894-9235 (office phone), 404-894-4409(fax)

It is worth mentioning that random walks on self-similar fractals have been studied in the literature, see e.g. [BN98]. However, to the best of our knowledge, our results for a *disconnected ultrametric* Cantor set are new. We also stress that our approach can be easily implemented for other Cantor sets with rich self-semilarity properties. However, it is not clear how it can be used for a general Cantor set.

The paper is organized as follows. In Section 2 we introduce the setting by a description of the geometry of the Cantor set. In Section 3, we give our main result, a complete characterization of SSS Markov processes on the Cantor set, and discuss the relationship with the processes studied in [PB08]. In Section 4 we describe an explicit construction of SSS processes and study their jump statistics. The role of the Laplacian on the Cantor set is played by the generators of SSS processes, and in Section 5 we describe some of their spectral properties. In Section 6, we prove that the Bernoulli measure on C is a unique invariant measure for any SSS process. We also prove that it is exponentially attracting. In Section 7, we study the asymptotics of displacement moments of SSS processes for small transition times.

Acknowledgements. The author is grateful to Jean Bellissard and John Pearson for introducing him to their work on the geometry of Cantor sets and for several stimulating discussions that led to writing this paper. He is also grateful to NSF for partial support of this research via CAREER award DMS-0742424.

2 The ultrametric structure on the triadic Cantor set

A Cantor set is a topological space that is non-empty, compact, perfect, totally disconnected and metrizable. In this paper, we study a classical example, the triadic Cantor set

$$C = \left\{ x : \; x = \sum_{k=1}^{\infty} \frac{2x_k}{3^k}, \; x_k = 0, 1, \text{ for all } k \in \mathbb{N} \right\}.$$

In many situations it is natural to identify $x \in C$ with the sequence $(x_k)_{k \in \mathbb{N}}$ which in turn may be identified with an infinite simple path on the infinite rooted binary tree.

For every $x, y \in C$ we define

$$d(x, y) = 3^{-c(x,y)},$$

where

$$c(x, y) = \min\{k \in \mathbb{N} : x_k \neq y_k\}.$$

(We agree that $c(x, x) = \infty$ and $d(x, x) = 0$ for any $x \in C$.) It is easy to see that d is equivalent to the Euclidean metric:

$$\frac{1}{3}|x - y| \leq d(x, y) \leq |x - y|, \quad x, y \in C.$$

It is also easy to see that d is an ultrametric, i.e. it satisfies the strong triangle inequality:

$$d(x,y) \le \max\{d(x,z), d(z,y)\}, \quad x,y,z \in C.$$

The metric space (C,d) has rich structure that involves self-similarity properties and a rich group of isometries. Let us denote

$$\pi_n x = (x_1,\dots,x_n), \quad x \in C, n \in \mathbb{N}.$$

Then, for every $n \in \mathbb{N}$, the set C may be decomposed into a disjoint family of 2^n sets:

$$[v] = \{x \in C:\ \pi_n = v\}, \quad v \in L_n,$$

where $L_n = \{0,1\}^n$.

Each of these sets is similar to C with scaling coefficient 3^n. One of the similarity maps is given by

$$(v_1, v_2, \dots, v_n, x_{n+1}, x_{n+2}, \dots) \mapsto (x_{n+1}, x_{n+2}, \dots).$$

In particular, these sets are isometric to each other and have diameter 3^{-n-1}. This leads to the following complete description of all isometries. Let g be an isometry. Then for each $n \in \mathbb{N}$, g permutes sets $[v]$, $v \in L_n$, i.e. for every $v \in L_n$, there is $v' \in L_n$ such that

$$g([v]) = [v'],$$

which we shall abbreviate as

$$g_n(v) = v'.$$

Obviously,

$$g_n(v) = v' \Rightarrow g_{n-1}(\pi_{n-1}v) = \pi_{n-1}v', \tag{1}$$

where we used the notation

$$\pi_m(v_1,\dots,v_n) = (v_1,\dots,v_m), \quad m \le n,\ v \in L_n,$$

and one can easily show that a sequence of bijections $g_n : L_n \to L_n$ satisfying the consistency condition (1) generates an isometry.

3 Markov processes, the main characterization theorem

Our aim is to obtain an analogue of the Brownian Motion (or symmetric stable Levy process) on (C,d). Since C is completely disconnected, there is no hope that a nontrivial Markov process will have continuous trajectories. So we shall relax the continuity requirement, and consider stochastically continuous Markov processes with càdlàg (right-continuous and with left limits) trajectories. Let us recall that a homogeneous Markov processes with transition probability function

$$P(t,x,B), \quad t \in [0,\infty),\ x \in C,\ B \in \mathcal{B}(C,d),$$

where $\mathcal{B}(C,d)$ is the Borel σ-algebra on (C,d), is called stochastically continuous if for any open set U and any point $x \in U$,

$$\lim_{t\downarrow 0} P(t,x,U) = 1.$$

Transition probability function $P(\cdot,\cdot,\cdot)$ is called Feller if the space $\mathcal{C}(C)$ of continuous real-valued functions on C is invariant under the semigroup $(S^t)_{t\geq 0}$ generated by $P(\cdot,\cdot,\cdot)$ and defined via:

$$(S^t f)(x) = \int_C f(y) P(t,x,dy). \tag{2}$$

(We refer to Chapter I of [Lig85] for a concise exposition of the necessary background on Markov semigroups.)

Let $(P_x)_{x\in C}$ be a Markov family compatible with $P(\cdot,\cdot,\cdot)$, i.e. for each x, P_x is a measure on the space of càdlàg C-valued trajectories on $[0,\infty)$ such that for the canonical process X, $P_x\{X(0)=x\}=1$ and, under P_x, X is a Markov process with transition probabilitiy $P(\cdot,\cdot,\cdot)$.

We are going to impose restrictions on $P(\cdot,\cdot,\cdot)$ one by one and show that these restrictions together lead to a concise analytic description of the generator of $(S^t)_{t\geq 0}$ allowing for a parametrization of the set of allowed processes by two parameters. Then, for any choice of these two parameters, we construct a unique process with required properties.

We say that $P(\cdot,\cdot,\cdot)$ is invariant under isometries if for any isometry g,

$$P(t,g(x),g(B)) = P(t,x,B), \quad t\in[0,\infty),\ x\in C,\ B\in\mathcal{B}(C,d).$$

Lemma 1 *Let $(X_t)_{t\geq 0}$ be a homogeneous Markov process with transition function $P(\cdot,\cdot,\cdot)$ invariant under isometries. Then $(\pi_n X_t)_{t\geq 0}$ is also a homogeneous Markov process.*

Proof: We notice that the isometry invariance straightforwardly implies

$$\begin{aligned}
&\mathsf{P}\{\pi_n X_0 = v_0, \pi_n X_{t_1} = v_1, \ldots, \pi_n X_{t_m} = v_m\}\\
&= \int_{[v_0]} \nu(dx_0) \int_{[v_1]} P(t_1 - t_0, x_0, dx_1) \int_{[v_2]} P(t_2 - t_1, x_1, dx_2) \ldots \int_{[v_m]} P(t_m - t_{m-1}, x_{m-1}, dx_m)\\
&\qquad = \nu([v_0]) P_n(t_1 - t_0, v_0, v_1) \ldots P_n(t_m - t_{m-1}, v_{m-1}, v_m),
\end{aligned}$$

where ν is the distribution of X_0, and the quantities

$$P_n(s,u,v) = P(s,x,[v]), \quad \pi_n x = u,$$

are well-defined due to the isometry invariance.

Since $\pi_n(X_t)$ takes finitely many values, we conclude that $\pi_n X$ is a Markov chain with transition matrix $P_n(s,v,u)$. $\square$

The process $\pi_n X$ inherits the stochastic continuity from X. Since it takes finitely many values, the transition rates

$$q(u,v) = \lim_{t\downarrow 0} \frac{P_n(t,u,v) - \delta_{uv}}{t},$$

are well-defined, where δ_{uv} is the Kronecker symbol, see e.g. [Chu67][Theorem 5, Section II.2]. If $v, u \in L_n$ for some n, then for any $x \in u$, we set $q(x,v) = q(u,v)$.

Let us introduce

$$d(u,v) = d([u],[v]) = \inf\{d(x,y) : x \in [u], y \in [v]\}.$$

Due to the isometry invariance, $q(u,v_1) = q(u,v_2)$ if $d(u,v_1) = d(u,v_2)$, so that $q_n(u,v)$ is actually a function of $d(u,v)$, and we can write

$$q(u,v) = q(d(u,v)).$$

If $n \le m$ and $v \in L_n$ then

$$q(x,v) = \sum_{\substack{w \in L_m \\ \pi_n w = v}} q(x,w), \quad x \in C.$$

If, moreover, $x \notin [v]$, then considering the family of isometries that leave x fixed and permute points w in the above summation, we can conclude that all the terms $q(x,w)$ coincide. Therefore, in this case,

$$q(x,v) = 2^{m-n} q(x,w), \quad w \in L_m, \pi_n w = v. \tag{3}$$

For any $x \in C$ and $n \in \mathbb{N}$, we define $q_n = q(x, v_n(x))$, where

$$v_n(x) = (x_1, \ldots, x_{n-1}, 1 - x_n). \tag{4}$$

For any n, q_n does not depend on x due to the isometry invariance. Using (3), we see that all rates $q(x,v)$ can be expressed in terms of q_n. Namely, for $v \in L_n$ and $x \notin [v]$,

$$q(x,v) = 2^{-(n-c(x,v))} q_{c(x,v)}, \tag{5}$$

where $c(x,v) = \min\{k : x_k \ne v_k\}$. We conclude that the distribution of X under P_x is completely determined by the family of jump rates $(q_n)_{n\in\mathbb{N}}$.

Let us recall that the infinitesimal generator for the semigroup $(S^t)_{t\ge 0}$ defined in (2) is given by

$$Af = \lim_{t\downarrow 0} \frac{S^t f - f}{t} \tag{6}$$

for $f \in \mathcal{D}(A)$, where $\mathcal{D}(A)$ is the domain of A, i.e., set of all functions f such that the r.h.s. of (6) is well-defined as a uniform limit. If f is a cylindric function, i.e., $f(x) = h(\pi_n x)$ for some function $h : L_n \to \mathbb{R}$, then $f \in \mathcal{D}(A)$ and

$$Af(x) = \sum_{v \in L_n} q(x,v)(h(v) - h(\pi_n x)). \tag{7}$$

The next property we would like to require is self-similarity, or scale invariance. Let us recall that for any n and any $v \in L_n$, $[v]$ is similar to C. So we shall require that the distribution of the Markov process X confined to $[v]$ coincides with the appropriately scaled distribution of the unrestricted process on the entire C.

To make this precise, we need to introduce confinements of Markov processes. For any $T > 0$, $n \in \mathbb{N}$, $v \in L_n$, $x \in [v]$ and consider the conditional measure

$$P_x^{v,T}\left\{X_{[0,T]} \in \cdot\ \right\} = P_x\left\{X_{[0,T]} \in \cdot \mid X_t \in [v], t \le T\right\},$$

where $X_{[0,T]}$ denotes the restriction of the canonical process X onto the time interval $[0,T]$.

Lemma 2 *For any $T > 0$, $n \in \mathbb{N}$, $v \in L_n$, $x \in [v]$, the canonical process X is Markov under measure $P_x^{v,T}$ on $[0,T]$. Moreover, these distributions are consistent for different values of T.*

Proof: Choose any $n' > n$, $l \ge 1$, and $u_1, \ldots, u_l \in L_{n'}$ such that $[u_k] \subset [v]$ for all k, and $0 < t_1 < \ldots < t_l = T$. Then

$$P_x\big\{X_{t_1} \in [u_1], \ldots, X_{t_l} \in [u_l] \mid X_t \in [v], t \in [0,T]\big\} = \frac{b}{a}. \tag{8}$$

Here

$$\begin{aligned} a = P_x\{X_t \in [v], t \le T\} &= \lim_{m\to\infty} P_x\left\{X_t \in [v], t = \frac{T}{m}, \frac{2T}{m}, \ldots, \frac{mT}{m}\right\} \\ &= \lim_{m\to\infty} P\left(\frac{T}{m}, v, v\right)^m \\ &= \exp\{q(v,v)T\}, \end{aligned}$$

(we used the right-continuity of trajectories for the first identity, the isometry invariance plus for the second, and the definition of q_n for the third one), and

$$\begin{aligned} b &= P_x\{X_{t_1} \in [u_1], \ldots, X_{t_l} \in [u_l];\ X_t \in [v], t \in [0,T]\} \\ &= \lim_{m\to\infty} \sum P\left(\frac{t_1}{m}, [\pi_n x], [w_1^1]\right) P\left(\frac{t_1}{m}, [w_1^1], [w_2^1]\right) \ldots P\left(\frac{t_1}{m}, [w_{m-1}^1], [u_1]\right) \\ &\quad \times P\left(\frac{t_2 - t_1}{m}, [u_1], [w_1^2]\right) P\left(\frac{t_2 - t_1}{m}, [w_1^2], [w_2^2]\right) \ldots P\left(\frac{t_2 - t_1}{m}, [w_{m-1}^2], [u_2]\right) \\ &\quad \ldots \\ &\quad \times P\left(\frac{t_l - t_{l-1}}{m}, [u_{l-1}], [w_1^l]\right) P\left(\frac{t_l - t_{l-1}}{m}, [w_1^l], [w_2^l]\right) \ldots P\left(\frac{t_l - t_{l-1}}{m}, [w_{m-1}^l], [u_l]\right) \\ &= \exp\{t_1 Q\}_{\pi_n x, u_1} \exp\{(t_2 - t_1)Q\}_{u_1,u_2} \ldots \exp\{(t_l - t_{l-1})Q\}_{u_{l-1},u_l}, \end{aligned}$$

where $Q = (Q_{z_1,z_2})_{z_1,z_2 \in L_n}$ is the matrix given by

$$Q_{z_1,z_2} = q(z_1,z_2)\mathbf{1}_{[z_1],[z_2] \subset [v]}.$$

Therefore, the l.h.s. of (8) equals

$$P^v(t_1, \pi_n x, u_1) P^v(t_2 - t_1, u_1, u_2) \dots P^v(t_l - t_{l-1}, u_{l-1}, u_l),$$

where

$$P^v(s, z_1, z_2) = \exp(q(v, v)s) \exp\{sQ\}_{z_1, z_2}, \quad z_1, z_2 \in L_n, [z_1], [z_2] \subset [v]. \tag{9}$$

Clearly, $P^v(s, z_1, z_2)$ is a Markov transition matrix, and it does not depend on T which completes the proof of the lemma. □

The lemma above means that our Markov process conditioned on the fact that it stays within $[v]$ up to time T is also a Markov process with transition probabilities that do not depend on T. Therefore, we can consistently define this process up to infinite time. We denote the resulting measure on infinite paths in $[v]$ by P^v_x. The collection of these measures for all $x \in C$ is a Markov family.

Let us now give a precise notion of self-similarity. We say that the Markov family $(P_x)_{x \in C}$ is self-similar, if for any n there is a number α_n such that for every $v \in L_n$ and every $x \in [v]$, and any map h realizing the similarity between $[v]$ and C,

$$P^v_x\{X_{t_1} \in B_1, \dots, X_{t_l} \in B_l\} = P_{h(x)}\{X_{\alpha_n t_1} \in h(B_1), \dots, X_{\alpha_n t_l} \in h(B_l)\}. \tag{10}$$

We shall say that a Markov family (as well as the associated Markov process, transition function, and semigroup) on C is SSS (symmetric and self-similar) if it is stochastically continuous, Feller, isometry invariant, self-similar, and the trajectories of the associated Markov process are a.s.-càdlàg.

Suppose that the Markov family P_x is SSS. Let h be a similarity map between $[v]$ and C. Then under P^v_x, $h(X)$ is a Markov process that inherits the stochastic continuity, Feller property and isometry invariance from the original Markov family. Therefore, all the above reasoning for the Markov family (P_x) applies to $h(X)$. In particular the distribution of $h(X)$ under (P^v_x) is completely determined by rates q^v_n that are defined for $h(X)$ under (P^v_x) in the same way as the rates q_n are defined for X under (P_x) (these rates do not depend on T as well).

Due to (9), the jump rates for the confined process X under (P^v_x) are

$$q^v(z_1, z_2) = \lim_{t \downarrow 0} \frac{P^v(t, z_1, z_2) - \delta_{z_1, z_2}}{t} = q(z_1, z_2) + \delta_{z_1, z_2} q(v, v). \tag{11}$$

This means that the process confined to $[v]$ can be viewed as the original process except the jumps out of $[v]$ are prohibited or ignored. Taking $v = (0)$ and the left 1-shift on sequences $(0, x_2, x_3 \dots)$ for h, we get for the jump rates of $h(X)$ under (P^v_x):

$$q^{(0)}_n = q_{n+1}, \quad n \in \mathbb{N}. \tag{12}$$

By the self-similarity hypothesis we must have

$$q^{(0)}_n = \alpha_1 q_n, \quad n \in \mathbb{N}. \tag{13}$$

Comparing (12) and (13), we see that

$$q_{n+1} = \alpha_1 q_n, , \quad n \in \mathbb{N},$$

so that

$$q_n = \alpha_1^{n-1} q_1, \quad n \in \mathbb{N}. \tag{14}$$

Theorem 1 *Suppose a Markov family on (C, d) is SSS. Then there are numbers $\theta, \gamma \geq 0$ such that for every cylindric function $f = h \circ \pi_n$, the generator Af is given by*

$$Af(x) = \gamma \sum_{k=1}^{n} \theta^k (\langle h \rangle_{n,k,x} - h(\pi_n x)), \tag{15}$$

where

$$\langle h \rangle_{n,k,x} = 2^{-(n-k)} \sum_{v \in L_n : c(x,v) = k} h(v).$$

Proof: Set $\theta = \alpha_1$, $\gamma = q_1/\alpha_1$ ($\gamma = 0$ if $\alpha_1 = 0$), and use (7), (5), and (14). □

Let us recall that a linear operator A defined on a vector subspace $\mathcal{D}$ of $\mathcal{C}(C)$ is called a Markov pregenerator (see [Lig85, Chapter 1, Definition 2.1]), if

1. $1 \in \mathcal{D}$ and $A1 = 0$.
2. $\mathcal{D}$ is dense in $\mathcal{C}(C)$.
3. If $f \in \mathcal{D}$, $\mu \geq 0$ and $f - \mu Af = g$, then

$$\min_{x \in C} f(x) \geq \min_{x \in C} g(x).$$

Lemma 3 *The operator A defined on cylindric functions via* (15) *is a Markov pregenerator.*

Proof: First two properties are obvious, and the third one follows from [Lig85, Chapter 1, Proposition 2.2] since A satisfies the following easily verifiable condition: if $f \in \mathcal{D}$ and $f(x^*) = \min_{x \in C} f(x)$, then $Af(x^*) \geq 0$. □

Lemma 3 implies that the closure of A denoted by $\bar{A}$ is also a well-defined closed Markov pregenerator, see [Lig85, Chapter 1, Proposition 2.5].

Lemma 4 *The operator $\bar{A}$ is a Markov generator, i.e., it is a closed Markov pregenerator satisfying*

$$\mathcal{R}(I - \mu A) = \mathcal{C}(C), \quad \text{for all } \mu > 0. \tag{16}$$

Proof: We need only to show (16). It is easy to see that for any cylindric function h, there is a cylindric function f such that $f - \mu \bar{A} f = h$ (this can be derived from the fact that $\pi_n x$ is a Markov process for every n, and the one-to-one correspondence between Markov semigroups and Markov generators given by the Hille–Iosida theorem). The lemma follows since the set of cylindric functions is dense, and $\mathcal{R}(I - \mu\bar{A})$ is closed (see [Lig85, Chapter 1, Proposition 2.6]). $\square$

We can now summarize the above.

Theorem 2 *If a Markov family $(P_x)_{x\in C}$ is SSS, then the generator of the associated Markov semigroup coincides with the closure of the operator A defined on cylindric functions via* (15) *for some $\gamma, \theta \geq 0$.*

Proof: The result follows now from the Hille-Iosida theorem (see e.g. [Lig85, Chapter 1, Theorem 2.9]) which establishes a one-to-one correspondence between Markov generators and Markov semigroups. $\square$

Theorem 2 gives a necessary condition for a Markov semigroup to be SSS. The next result shows that this condition is, in fact, sufficent.

Theorem 3 *For any $\gamma, \theta \geq 0$ there is a unique Markov family with Markov generator coinciding with A defined on cylindric functions via* (15)*. That Markov family is SSS (with scaling parameter given by $\alpha_n = \theta^n$, $n \in \mathbb{N}$).*

Proof: The existence-uniqueness and the Feller property follows from the Hille–Iosida theorem. The isometry invariance follows from that of A. The stochastic continuity follows from

$$P(t, x, [v]) = 1 - t\gamma \sum_{k=1}^{n} \theta^k + o(t), \quad t \to 0,$$

for any $n \in \mathbb{N}$, any $v \in L_n$, and any $x \in [v]$. In particular, a càdlàg version of the canonical process exists.

For the self-similarity, we must take any $n \in \mathbb{N}$, $v \in L_n$, $x \in [v]$, and consider the process X emitted from x under the condition that it stays within $[v]$. Due to Lemma 2, the conditioned process is Markov, and so is $h(X)$ under P^v, where h is a similarity map between $[v]$ and C. Computing the transition probabilities and jump rates for this process:

$$q_k^v = q_{n+k} = \gamma\theta^{n+k} = \theta^n\gamma\theta^k = \theta^n q_k,$$

so that $h(X)$ under P has the same distribution as X under P^v and time change $t \to \theta^n t$. The proof is complete. $\square$

Theorems 2 and 3 give a complete characterization of SSS processes, the analogues of the Wiener process on (C, d) via their Markov generators. SSS processes are naturally parametrized by γ and θ. We shall write $SSS(\gamma, \theta)$ to denote the SSS process with parameters γ, θ. It is clear that γ is responsible for uniform time changes, and it is often sufficient to study the case $\gamma = 1$, since by a simple time rescaling one can obtain the process with any given γ. However, the self-similarity parameter θ is essential, and there are qualitative differences between processes with different values of θ.

Our scaling factor θ is equal to 3^{s_0+2-s} in the notation of [PB08], where only values of $s \geq s_0 = \frac{\ln 2}{\ln 3}$ (the box dimension of C) were considered. Therefore, our approach extends the class of symmetric self-similar processes on C by removing the constraint $\theta \leq 9$.

Since SSS processes on the Cantor set play the role of symmetric diffusion, the generators A play the role of the Laplacian. We shall study spectral properties of A in Section 5.

4 Explicit construction and jump statistics

We begin with an explicit construction of $SSS(\gamma, \theta)$. Let $(\xi_{kj})_{k,j=1}^{\infty}$ be a family of independent random variables such that for each $k \in \mathbb{N}$, $(\xi_{k1}, \xi_{k2}, \ldots)$ are exponentially distributed with parameter $\gamma\theta^k$. Let $S_{kn} = \xi_{k1} + \ldots + \xi_{kn}$. Clearly, $N_k(t) = \max\{n : S_{kn} \leq t\}$, $t \geq 0$, is a càdlàg Poisson process with intensity $\gamma\theta^{k-1}$. We shall say that there is a jump at level k at time τ if $N_k(\tau) = N_k(\tau-) + 1$. Let $\tilde{N}_k(t) = N_1(t) + \ldots + N_{k-1}$ be the Poisson process that counts the jumps of Poisson processes at all levels below k, i.e. at levels $1, 2, \ldots, k-1$.

To define our Markov process X we shall also need a family of i.i.d. $\frac{1}{2}$-Bernoulli random variables $(\varkappa_{kj})_{k,j=1}^{\infty}$ independent of the Poisson processes described above.

We set $X(0) = x = (x_1, x_2, \ldots)$ and let the evolution of the k-th coordinate X_k to be defined by the following rules:

1. X_k stays constant while the processes N_k and $\tilde{N}_k$ are constant.
2. If at time τ the process N_k makes a jump, X_k also makes a jump so that $X_k(\tau) = 1 - X_k(\tau-)$.
3. If at time τ the process $\tilde{N}_k$ makes a jump, then $X_k(\tau) = \varkappa_{k\tilde{N}_k(\tau)}$ no matter what the value of $X_k(\tau-)$ was.

In other words, when a jump occurs at a level k, $X_1, \ldots, X_{k-1}$ do not change, X_k gets flipped, and $(X_{k+1}, X_{k+2}, \ldots)$ are re-initialized according to the $\frac{1}{2}$-Bernoulli product measure. We exclude the event of probability 0 on which two jumps happen at the same time. The process $(X_1, \ldots, X_n)$ makes finitely many jumps in finite time for any finite n.

It is easy to see that this procedure uniquely defines a Markov process with pregenerator described in the last section.

We see that the value $\theta = 1$ is critical. If $\theta < 1$ then with probability 1, X makes finitely many jumps in a finite time. However, if $\theta \geq 1$, then with probability 1, X makes infinitely many jumps in a finite time.

The value $\theta = 1$ is also special due to the following: if $\theta = 1$ then for any $n \in \mathbb{N}$ and any $v \in L_n$, the distribution of the restricted process $h(X)$ under P^v coincides precisely with that of X (no time change is needed). Notice also that in terms of this model our conditioning means that the process $\tilde{N}_n + N_n$ makes no jumps.

5 Spectral structure of the generator

In this section we fix the values of $\theta > 0$ and $\gamma = 1$ and study eigenvalues and eigenvectors of the infinitesimal operator A. The eigenvectors will be given by Haar function that we proceed to introduce. For any n and any $v \in L_n$, we define

$$\psi_v = 2^{n/2}(\chi_{v0} - \chi_{v1}),$$

where $\chi_u = \chi_{[u]}$ denotes the characteristic function (or indicator) of $[u]$ for any u. Notice that ψ with no indices denotes $\chi_0 - \chi_1$.

Theorem 4 *1. Eigenvalues of A are given by $\lambda_0 = 0$, and*

$$\lambda_n = -\sum_{k=1}^{n-1} \theta^k - 2\theta^n, \quad n \in \mathbb{N}$$
$$\left(= -\frac{2\theta^{n+1} - \theta^n - \theta}{\theta - 1} \text{ if } \theta \neq 1 \right).$$

2. *The unique (up to a multiplicative constant) eigenfunction associated to λ_0 is 1. For $n \in \mathbb{N}$, λ_n has multiplicity 2^{n-1}, and its eigenspace is spanned by $M_n = \{\psi_v : \ v \in L_{n-1}\}$.*

3. *The eigenfunctions described above form a complete basis in $\mathcal{C}(C)$ and $L^2(C)$. This basis is orthonormal in $L^2(C, b)$, where b is the $1/2$-Bernoulli measure on C (also known as the Cantor measure).*

4. *If $\theta < 1$, A is bounded with spectral radius equal to*

$$r(A) = \sum_{k=1}^{\infty} \theta^k = \frac{\theta}{1-\theta}.$$

Moreover, $A + r(A)I$ is compact where I is the identity operator, so that A is a compact perturbation of a multiple of the identity.

5. *If $\theta \geq 1$ then A has compact resolvent.*

This theorem is a result of straightforward computations and arguments. We omit the proof.

6 Unique ergodicity, exponential mixing

Theorem 5 *1. For any $\gamma, \theta > 0$, the $\frac{1}{2}$-Bernoulli product measure $b = \mu^{\mathbb{N}}$ on C is a unique invariant measure for $SSS(\gamma, \mu)$.*

2. *There are constants $K, \nu > 0$, such that for any point $x \in C$,*

$$|P(t, x, \cdot) - b(\cdot)|_{TV} \leq K e^{-\nu t},$$

where $|\cdot|_{TV}$ denotes the total variation norm.

Proof: The Bernoulli measure b is clearly invariant for SSS since it is invariant under isometries. The uniqueness of the invariant measure follows from the second statement of the theorem. The latter can be proved by exploiting the spectral gap provided by Theorem 4, but we choose another simple method instead.

Let us work with the explicit model introduced in Section 4. It is easy to see that for every $N \geq 1$, the distribution of $X(t)$ conditioned on $\{N_1(t) = N\}$ is given by $\delta_{(N+x_1)(\text{mod }2)} \times b$. This implies that for any set B,

$$\begin{aligned}
P(t,x,B) = &\sum_{N=1}^{\infty} \mathsf{P}\{N_1(t) = N\}(\delta_{(N+x_1)(\text{mod }2)} \times b)(B) + \beta(t,B) \\
=&(\delta_{(x_1+1)(\text{mod }2)} \times b)(B) e^{-\gamma\theta t} \sum_{m=1}^{\infty} \frac{(\gamma\theta t)^{2m-1}}{(2m-1)!} \\
&+ (\delta_{x_1(\text{mod }2)} \times b)(B) e^{-\gamma\theta t} \sum_{m=1}^{\infty} \frac{(\gamma\theta t)^{2m}}{(2m)!} + \beta(t,B) \\
=&(\delta_{(x_1+1)(\text{mod }2)} \times b)(B) e^{-\gamma\theta t} \sinh(\gamma\theta t) \\
&+ (\delta_{x_1(\text{mod }2)} \times b)(B) e^{-\gamma\theta t}(\cosh(\gamma\theta t) - 1) + \beta(t,B) \\
=&\frac{1}{2}(\delta_{(x_1+1)(\text{mod }2)} \times b)(B)(1 - e^{-2\gamma\theta t}) \\
&+ \frac{1}{2}(\delta_{x_1(\text{mod }2)} \times b)(B)(1 + e^{-2\gamma\theta t} - 2e^{-\gamma\theta t}) + \beta(t,B), \\
=&b(B) - \frac{1}{2}(\delta_{(x_1+1)(\text{mod }2)} \times b)(B) e^{-2\gamma\theta t} \\
&+ \frac{1}{2}(\delta_{x_1(\text{mod }2)} \times b)(B)(e^{-2\gamma\theta t} - 2e^{-\gamma\theta t}) + \beta(t,B),
\end{aligned}$$

where

$$0 \leq \beta(t,B) \leq \mathsf{P}\{N_1(t) = 0\} = e^{-\gamma\theta t}.$$

So the second statement of the theorem follows with $K = \frac{3}{2}$, $\nu = \gamma\theta$. □

7 Displacement moments

In this section we study the behavior of the moments of displacement of the process X given by

$$M_r(t) = \mathsf{E}_x d^r(x, X(t)) = \sum_{k=1}^{\infty} 3^{-rk} P_x(c(x, X(t)) = k),$$

as $t \to 0$. Due to the isometry invariance it is sufficient to consider $x = \bar{0} = (0, 0, \ldots)$.

Theorem 6 *Let $r > 0$.*

1. *If $3^r > \theta$, then*

$$M_r(t) = \sum_{k=1}^{\infty} 3^{-rk}\theta^k (t + o(t)), \quad t \downarrow 0.$$

2. *If $3^r < \theta$, then*

$$0 < \liminf_{t\downarrow 0} \frac{M_r(t)}{t^{\frac{\ln 3}{\ln \theta}r}} \leq \limsup_{t\downarrow 0} \frac{M_r(t)}{t^{\frac{\ln 3}{\ln \theta}r}} < \infty.$$

Remark 1 *It is shown in [PB08] that if $3^r = \theta$, then $M_r(t)$ behaves as $t\ln(1/t)$ as $t \to 0$.*

Proof of Theorem 6: For any $k \in \mathbb{N}$,

$$\begin{aligned} P(x,t,[v_k(x)]) &= \int_{y\neq x} \chi_{[v_k(x)]}(y)P(x,t,dy) \\ &= S^t \chi_{[v_k(x)]}(x) \\ &= \langle \chi_{[v_k(x)]}, 1\rangle \chi_{[v_k(x)]}(x) + \sum_{i=0}^{k-1} e^{\lambda_{i+1}t} \sum_{u\in L_i} \langle \chi_{[v_k(x)]}, \psi_u\rangle \psi_u(x), \end{aligned}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in the Hilbert space $L^2(C, b)$. The first term in the r.h.s. is 0, and for each i in the sum above, the only nonzero contribution in the sum comes from

$$\langle \chi_{[v_k(x)]}, \psi_{\pi_i x}\rangle = \begin{cases} 2^{\frac{i}{2}-k}, & i < k-1 \\ -2^{-\frac{k+1}{2}} & i = k-1, \end{cases}$$

so that

$$P(x,t,[v_k(x)]) = \sum_{i=0}^{k-2} e^{\lambda_{i+1}t} 2^{\frac{i}{2}-k} \bar{\psi}_i(\bar{0}) - e^{\lambda_k t} 2^{-\frac{k+1}{2}} \bar{\psi}_{k-1}(\bar{0}),$$

where

$$\bar{\psi}_i = \psi_{\underbrace{00\ldots0}_{i}}.$$

Since $\psi_i(\bar{0}) = 2^{i/2}$, we obtain

$$P(x,t,[v_k(x)]) = \sum_{i=0}^{k-2} 2^{i-k} e^{\lambda_{i+1}t} - 2^{-1} e^{\lambda_k t},$$

and

$$M_r(t) = \sum_{k=1}^{\infty} 3^{-kr} \left(\sum_{i=0}^{k-2} 2^{i-k} e^{\lambda_{i+1}t} - 2^{-1} e^{\lambda_k t} \right)$$

If $3^r > \theta$, then the series

$$\sum_{k=1}^{\infty} 3^{-kr} \left(\sum_{i=0}^{k-2} \lambda_{i+1} 2^{i-k} e^{\lambda_{i+1} t} - 2^{-1} \lambda_k e^{\lambda_k t} \right)$$

converges uniformly in $t \geq 0$. Therefore, it represents $\partial_t M_r(t)$. Evaluating this series at 0 produces

$$\frac{d^+}{dt} M_r(t) = \sum_{k=1}^{\infty} 3^{-kr} \left(\sum_{i=0}^{k-2} \lambda_{i+1} 2^{i-k} - 2^{-1} \lambda_k \right).$$

One can show by induction that

$$\sum_{i=0}^{k-2} \lambda_{i+1} 2^{i-k} - 2^{-1} \lambda_k = \theta^k, \quad k \in \mathbb{N},$$

so that in this case

$$\frac{d^+}{dt} M_r(t) = \sum_{k=1}^{\infty} 3^{-rk} \theta^k,$$

and the first statement of the theorem follows.

Let now $\theta > 3^r > 1$. Consider a sequence $t_n = R_n \theta^{-n}$, $n \in \mathbb{N}$ with $R_n \in [1, \theta]$.

We have

$$M_r(t_n) = \sum_{k=1}^{\infty} a_{k,n} = \sum_{m=-n+1}^{\infty} a_{n+m,n}, \tag{17}$$

where

$$a_{k,n} = 3^{-kr} \left(\sum_{i=0}^{k-2} 2^{i-k} e^{\lambda_{i+1} t_n} - 2^{-1} e^{\lambda_k t_n} \right) > 0.$$

Taking $k = m + n$, we obtain

$$\begin{aligned} \frac{a_{m+n,n}}{t^{\frac{\ln 3}{\ln \theta} r}} &= \frac{3^{-(n+m)r}}{R_n^{\frac{\ln 3}{\ln \theta} r} 3^{-nr}} \left(\sum_{i=0}^{n+m-2} 2^{i-(n+m)} e^{\lambda_{i+1} R_n \theta^{-n}} - 2^{-1} e^{\lambda_{n+m} R_n \theta^{-n}} \right) \\ &= R_n^{-\frac{\ln 3}{\ln \theta} r} 3^{-mr} \left(\sum_{l=1}^{n+m-1} 2^{-1-l} e^{\lambda_{n+m-l} R_n \theta^{-n}} - 2^{-1} e^{\lambda_{n+m} R_n \theta^{-n}} \right), \end{aligned}$$

where we used a change of variables $l = n + m - 1 - i$. Therefore, for $m \geq 0$,

$$\frac{a_{m+n,n}}{t^{\frac{\ln 3}{\ln \theta} r}} \leq 3^{-mr} \left(\sum_{l=1}^{n+m-1} 2^{-1-l} + 2^{-1} \right) \leq 3^{-mr}. \tag{18}$$

If $m \leq 0$, then we use another estimate:

$$\frac{a_{m+n,n}}{t^{\frac{\ln 3}{\ln \theta} r}} \leq 3^{-mr} \left(1 - e^{\lambda_{n+m} R_n \theta^{-n}} \right). \tag{19}$$

Since

$$|\lambda_{n+m}R_n\theta^{-n}| \leq \frac{2\theta^{n+m+1} - \theta^{n+m} - \theta}{\theta - 1}\theta^{-n+1} \leq K_1\theta^m,$$

for some $K_1 > 0$, all $n > 0$, $-n+1 \geq m \leq 0$, inequality (19) can be continued as

$$\frac{a_{m+n,n}}{t^{\frac{\ln 3}{\ln\theta}r}} \leq K_2 3^{-mr}\theta^m, \tag{20}$$

for some $K_2 > 0$ and all $m \leq 0$. The upper estimate in the theorem follows now from (18) and (20). To prove the lower bound, take $m = 0$:

$$\frac{M_r(t)}{t^{\frac{\ln 3}{\ln\theta}r}} \geq \frac{a_{n,n}}{t^{\frac{\ln 3}{\ln\theta}r}} \geq R_n^{-\frac{\ln 3}{\ln\theta}r}\left(\sum_{l=1}^{n-1} 2^{-1-l}e^{\lambda_{n-l}R_n\theta^{-n}} - 2^{-1}e^{\lambda_n R_n\theta^{-n}}\right),$$

Notice that

$$\lambda_{n-l}\theta^{-n} = -\frac{2\theta^{-l+1} - \theta^{-l} - \theta^{1-n}}{\theta - 1} \;\to\; -\nu\theta^{-l}, \quad n \to \infty, \tag{21}$$

where $\nu = \frac{2\theta-1}{\theta-1}$. So,

$$\begin{aligned}\frac{M_r(t)}{t^{\frac{\ln 3}{\ln\theta}r}} &\geq \theta^{-\frac{\ln 3}{\ln\theta}r}\left(\sum_{l=1}^{n-1} 2^{-1-l}e^{-R_n\nu\theta^{-l}} - 2^{-1}e^{-R_n\nu}\right)\\ &+ \theta^{-\frac{\ln 3}{\ln\theta}r}\left(\sum_{l=1}^{n-1} 2^{-1-l}(e^{\lambda_{n-l}R_n\theta^{-n}} - e^{-R_n\nu\theta^{-l}}) - 2^{-1}(e^{\lambda_n R_n\theta^{-n}} - e^{-R_n\nu})\right).\end{aligned}$$

Due to (21), the second sum converges to zero, and we have

$$\liminf_{n\to\infty}\frac{M_r(t)}{t^{\frac{\ln 3}{\ln\theta}r}} \geq \theta^{-\frac{\ln 3}{\ln\theta}r}\inf_{R\in[1,\theta]}\left(\sum_{l=1}^{\infty} 2^{-1-l}e^{-R\nu\theta^{-l}} - 2^{-1}e^{-R\nu}\right) > 0,$$

and the proof is completed. □

References

[BN98] M. T. Barlow and D. Nualart. *Lectures on probability theory and statistics*, volume 1690 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1998. Lectures from the 25th Saint-Flour Summer School held July 10–26, 1995, Edited by P. Bernard.

[Chu67] Kai Lai Chung. *Markov chains with stationary transition probabilities.* Second edition. Die Grundlehren der mathematischen Wissenschaften, Band 104. Springer-Verlag New York, Inc., New York, 1967.

[Law05] Gregory F. Lawler. *Conformally invariant processes in the plane*, volume 114 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2005.

[Lig85] Thomas M. Liggett. *Interacting particle systems*, volume 276 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1985.

[PB08] John Pearson and Jean Bellissard. Noncommutative Riemannian geometry and diffusion on ultrametric Cantor sets. *Submitted to Journal of Noncommutative Geometry. Also available at http://arxiv.org/abs/0802.1336v1*, 2008.