Séminaire de probabilités (Strasbourg)

Martin Baxter

Markov processes on the boundary of the binary tree

Séminaire de probabilités (Strasbourg), tome 26 (1992), p. 210-224
<http://www.numdam.org/item?id=SPS_1992__26__210_0>

Markov processes on the boundary of the binary tree

MARTIN BAXTER

Statistical Laboratory, University of Cambridge, Cambridge CB2 1SB

1. Introduction and Summary

We consider a Markov chain on the nodes of the binary tree, I:

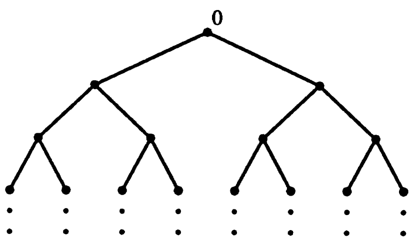

Figure 1. Graph of Binary Tree

By choice of jump rates and the relative up-down weightings, we can ensure that the chain is reversible, positive recurrent, and able to hit infinity and return in finite time. Our basic structural assumption, on which we shall lean heavily, is of lateral symmetry — that is that jump rates depend on the state only through its level and that the process is equally likely to go left as right on any down jump.

Rogers and Williams [4] allows us the existence of the chain with reflection at its boundary. (David Williams asks me to report that Ivor McGillivray has explained to him that the reflection in the example on page 156 is off the Kuramochi rather than the Martin boundary. In our example, the two agree.) The Ray-Knight compactification can be thought of as:

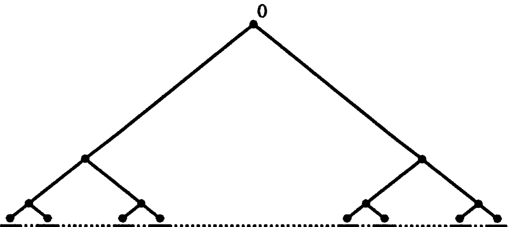

Figure 2. Ray-Knight Compactification

Let I be the points of the nodes of the graph, C be the Cantor set of limit points of I, and F be $I \cup C$, the Ray-Knight compactification of I. There are no branch points, and no 'irrelevant' points in the sense of III.81 of Williams [6].

Construction. We use Rogers and Williams [4], and results in Williams [6] to construct the Markov chain on the graph with reflection at the boundary. Unidentified references in this paragraph refer to Williams [6]. The time-truncation arguments of

Rogers and Williams [4] let us take the limit of finite chains on the tree which reflect at level n to give us, their Theorem 9.13, a π-symmetric transition matrix $P(t)$ and Feller resolvent R_λ on I. From III.57, we extend R_λ to $C(F)$, and there are no branch points, as we can condition the chain to hit any particular boundary point (III.49). By symmetry either all the boundary points are relevant ($\lambda r_\lambda(\xi, I) = 1$) or they are all not, but as we must visit some of them, it is the former which holds (III.81). From III.60, we see that $P(t)$ is Feller-Dynkin (FD), and we can thus, III.11, construct a strong Markov F-valued honest R-process X with law $P(t)$.

Given the F-valued process X, the projection process $n(X_t)$, the level of X_t, is also a Markov chain (by the symmetry) and is a birth-death process on the non-negative integers with the one-point compactification at infinity. This level process reflects from infinity, as the projection of the time-truncation of X is the same process as the time-truncation of the reflecting birth-death process.

WARNING: Throughout, we shall switch from the graph process to the level process and back with impunity, using i, j and so on for states in I, and n to denote states of the level process. For example $\pi_i = 2^{-n(i)}\pi_{n(i)}$, where $n(i)$ is the level of state i.

The chain can be fully quantified by the jump rates from level n.

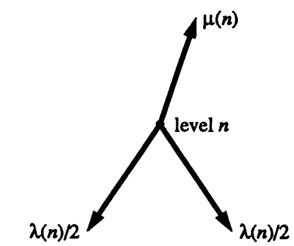

Figure 3. Graph process

Figure 4. Projected BD process

The up-jump rate is μ_n and the left and right down-jump rates are both $\frac{1}{2}\lambda_n$, and we shall define q_n to be $\lambda_n + \mu_n$. We put $\pi_n := (\lambda_0 \dots \lambda_{n-1})/(\mu_1 \dots \mu_n)\pi_0$, the invariant measure for the BD process, choosing π_0 to make π a distribution if it is a finite measure.

We can think in terms of the level process as being the time-change of a Brownian motion reflecting at each end of a bounded interval, $[0, a]$. This is made rigorous later (section 4). The set of times at which this Brownian process is at 0 is a random Cantor set, which in particular is uncountable and perfect. Our first question is inspired by noticing that when the graph process first hits the boundary, it must almost immediately return uncountably often to the boundary. But there are uncountably many other boundary points nearby. Does it return to the same point of the boundary ? That is, is the boundary point regular ? One can think of a bolt of lightning which bounces back off the ground and back down again repeatedly. From all the possible bits of ground, how can it ever find again the exact spot where it first hit ?

Assuming (wisely) that we can arrange it so that all boundary points are regular (and thus have individual local times), our second question is: can we find a jointly-continuous version of the local time on F ?

The precise answers to all these questions are contained in the following theorem:

THEOREM 1. *Let X be the graph process as constructed above. Then*
(1) X is positive recurrent $\iff \sum_n \pi_n < \infty$, and then
(2) X reaches the boundary in finite time $\iff \sum_n \frac{1}{\lambda_n \pi_n} < \infty$, and then
(3) any (and hence each) boundary point is regular $\iff \sum_n b_n < \infty$, and then
(4) there exists a jointly-continuous local time on $F \iff \sum_n \sqrt{\frac{1}{n} c_n} < \infty$, and then
(5) X has visited all the states of F by a finite time,
where $b_n := \frac{2^n}{\lambda_n \pi_n}$, and $c_n := \sum_{r=n}^{\infty} b_r$.

COROLLARY 2. *If (1)–(3) hold then*

$$\sum_n n^{1+\epsilon} b_n < \infty \quad \Rightarrow \quad \text{there exists a jointly-continuous local time on } F,$$
$$\text{and} \quad \sum_n n b_n = \infty \quad \Rightarrow \quad \text{there does not.}$$

We thus have enough to construct a boundary (Cantor set) valued process, which we examine in the final section.

2. Proof of Theorem

This section contains the mathematics of the proof, the next contains the arithmetic.

Parts (1) and (2) of the theorem are basic Markov chain theory. See, for example, 4-3 of Wolff [7].

Part (3). Rogers and Williams [4] have given us a standard honest π-symmetric transition matrix function $P(t) = (p_{ij}(t))_{i,j \in I}$, such that $\mathsf{P}_i(X_t = j) = p_{ij}(t)$. We define its Laplace transform $R(\lambda)$ and a π-normalised symmetric resolvent u_λ by

$$r_{ij}(\lambda) := \int_0^\infty e^{-\lambda t} p_{ij}(t)\, dt,$$
$$\text{and} \quad u_\lambda(i,j) = u_\lambda(j,i) := r_{ij}(\lambda)/\pi_j$$

respectively. It is known that a boundary point ξ is regular if and only if (for any and hence all λ) u_λ has a continuous extension to (ξ, ξ) and $u_\lambda(\xi, \xi) < \infty$. By the definition of the Ray-Knight compactification, in for example III.57 of Williams [6], we see that u_λ has a finite continuous extension to $F \times I$, and hence by symmetry to $F \times F \backslash \{(\xi, \xi) : \xi \in F \backslash I\}$.

Let us put a partial order on F by saying $x < y$ for x, y in F, if x is one of the points between y and the root of the tree 0. We say that x is *before* y, and that y is *beyond* x. For any pair x and y, we let $x \wedge y$ be the $<$-greatest point which is before both of them. Pick an i before ξ and let $I_i := \{j \in I : i \not< j\}$ be the set of all points not beyond i. For any k beyond i, and for any j in I_i, all paths from j to k must pass through i, so that the strong Markov property implies that

$$u_\lambda(j,k) = \mathsf{E}_j\left(e^{-\lambda H(i)}\right) u_\lambda(i,k),$$

where $H(i)$ is the time to first hit state i. Letting k tend to $\xi \in C$, we find that

$$u_\lambda(j,\xi) = \mathsf{E}_j\left(e^{-\lambda H(i)}\right) u_\lambda(i,\xi).$$

Then

$$\sum_{j\in I_i} \pi_j u_\lambda(j,\xi) = \sum_{j\in I_i} \pi_j \mathsf{E}_j\left(e^{-\lambda H(i)}\right) u_\lambda(i,\xi),$$

$$\text{and so} \qquad u_\lambda(i,\xi) = \frac{\lambda \sum_{j\in I_i} \pi_j u_\lambda(j,\xi)}{\lambda \sum_{j\in I_i} \pi_j \mathsf{E}_j\left(e^{-\lambda H(i)}\right)}.$$

As i goes to ξ, $I_i \uparrow I$, and the numerator tends upwards to $\lambda r_\lambda(\xi, I) = 1$ (Williams [6], III.81). Therefore

$$u_\lambda(\xi,\xi) = \frac{1}{\lambda \sum_I \pi_j \mathsf{E}_j\left(e^{-\lambda H(\xi)}\right)},$$

and $u_\lambda(\xi,\xi) < \infty$ if and only if $H(\xi) < \infty$ (a.s.).

We let the "up-jump time", V_n, be a random variable distributed as the time to hit level $(n-1)$ starting at level n, and let the "left-down-jump time", T_n, be a random variable distributed as the time to hit the point below and to the left of a start point on level n. Then we can control the means and variances of these in the following theorem.

THEOREM 3. *If (1) and (2) hold then*

$$\mathsf{E}V_n = \frac{\pi[n]}{\lambda_{n-1}\pi_{n-1}}, \qquad \mathrm{Var}(V_n) = \frac{1}{\lambda_{n-1}\pi_{n-1}} \sum_{r=n}^{\infty} \left(\frac{\pi[r]^2}{\lambda_{r-1}\pi_{r-1}} + \frac{\pi[r+1]^2}{\lambda_r \pi_r} \right),$$

$$\mathsf{E}T_n = \frac{2^{n+1} - \pi[n+1]}{\lambda_n \pi_n}, \qquad \mathrm{Var}(T_n) \leqslant K \frac{2^n}{\lambda_n \pi_n} \sum_{r=0}^{n} \frac{2^r}{\lambda_r \pi_r}, \quad \textit{for some } K,$$

where $\pi[r] := \pi(\{r, r+1, \ldots\})$.

We defer this proof till section 3, but the method of calculation in each case is just to find the minimal non-negative solution to a system of equations induced by conditioning on the first jump. We find that means are enough for upper bounds and sufficiency, but we need control away from 0, that is variance information, for lower bounds and necessity.

Sufficiency of (3). If $\sum_n b_n < \infty$ (where $b_n = 2^n/\lambda_n \pi_n$) then Theorem 3 shows that $\mathsf{E}_0(H(\xi)) = \sum_n \mathsf{E}T_n < \infty$ and so ξ is a regular boundary point.

Necessity of (3). Conversely, if $\sum_n b_n = \infty$ we use the following lemma:

LEMMA 4. (LOWER-BOUND LEMMA) *If* $X : \Omega \to [0,\infty]$ *is a random variable such that* $\mathsf{E}(X^2) \leqslant K\mathsf{E}(X) < \infty$ *for some* K, *then*

$$\mathsf{E}\left(1 - e^{-X}\right) \geqslant \alpha \mathsf{E}(X)$$

$$\textit{where} \quad \alpha = \alpha(K) = \frac{1 - e^{-4K}}{8K}$$

The proof (in section 3) uses concavity coupled with the variance control.

By Theorem 3 (noticing that $(\mathsf{E}T_n)^2$ is of no greater order than $\mathrm{Var}(T_n)$) we have that

$$\mathsf{E}(T_n^2) \leqslant K d_n \mathsf{E}(T_n),$$

(for some new K) where $d_n = \sum_{r=0}^{n} b_r$. And so the Lower Bound Lemma 4 tells us that

$$\begin{aligned}\mathsf{E}\left(1-e^{-T_n}\right) &\geqslant \frac{1-e^{-4Kd_n}}{8Kd_n}\mathsf{E}(T_n)\\ &\geqslant \alpha\frac{b_n}{d_n} \quad \text{for some } \alpha > 0,\end{aligned}$$

as $d_n \to \infty$ as $n \to \infty$. Kronecker's Lemma tells us that $\sum_n \frac{b_n}{d_n} = \infty$, and we deduce that $\sum_n \mathsf{E}(1-e^{-T_n}) = \infty$, and hence $\mathsf{E}_0(e^{-H(\xi)}) = \prod_n \mathsf{E}(e^{-T_n}) = 0$, giving us the necessity of condition (3). Here we have used, and will use again, the useful analysis lemma that for a sequence (x_n) in $(0,1)$, $\sum_n(1-x_n)$ is finite if and only if $\prod_n x_n$ is positive.

Part (4). Given (1)–(3), we can assume an individual local time $L(x,t)$ for each point x of F. For (4) we use an excellent paper of Marcus and Rosen [2], which uses an Isomorphism theorem of Dynkin between the Markov chain on the graph, and zero-mean Gaussian process on the graph with covariance equal to the 1-potential density $u_1(\cdot,\cdot)$. Their Theorems II and 9.1 together state that

THEOREM 5 (MARCUS AND ROSEN). *Let X be a strongly symmetric standard Markov Process with continuous 1-potential density u_1. Let $L = \{L(x,t) : x \in F,\ t \in \mathsf{R}^+\}$ be the joint local time of X, then L is continuous a.s. if and only if there exists a probability measure m on F such that*

$$\sup_{x\in F}\int_0^{\delta}\left[\log\frac{1}{m(B_d(x,r))}\right]^{1/2} dr \longrightarrow 0 \qquad \text{as } \delta \to 0, \qquad (*)$$

where $B_d(x,r)$ is the radius-r closed ball centered on x under the metric d, where

$$d(x,y) = \left[u_1(x,x) + u_1(y,y) - 2u_1(x,y)\right]^{1/2}.$$

X is *strongly symmetric* if u_1 exists as a symmetric π-density for the Laplace transform of $P(t)$, which here is true. Our first step is to show (in section 3)

LEMMA 6. *Given (1)–(3), for x, y in F, with $n(x) < n(y)$, then*

$$\alpha\mathsf{E}_{x\wedge y}(H(y)) \leqslant d^2(x,y) \leqslant A\mathsf{E}_{x\wedge y}(H(y)),$$

for some universal constants α and A.

As already hinted, the upper bound follows from simple inequalities and knowledge of the means, whilst the lower is derived from variance control and the more subtle analysis of the Lower Bound Lemma 4. The lemma tells us that d is "equivalent" to the sequence (c_n), and we essentially translate condition $(*)$ of Theorem 5 into a statement about (c_n), preserving both necessity and sufficiency.

Sufficiency of (4). We elect m to be $\frac{1}{2}p + \frac{1}{2}c$, where p is the probability on I giving mass $2^{-(2n+1)}$ to each point on level n, and c is the Cantor distribution (Hausdorff measure) on C. Our condition holds, that is

$$\sum_n \sqrt{\frac{1}{n}c_n} < \infty \qquad \text{where} \quad c_n := \sum_{r=n}^{\infty} b_r, \quad \text{and} \quad b_n := \frac{2^n}{\lambda_n \pi_n}.$$

Putting $x = \xi \in C$, we can deduce from Theorem 3 and Lemma 6 that (for new α and A)

$$\alpha c_{n(y\wedge\xi)} \leqslant d^2(\xi, y) \leqslant A c_{n(y\wedge\xi)}.$$

Thus for r chosen to lie in $\sqrt{Ac_n} \leqslant r < \sqrt{Ac_{n-1}}$, then $I_n := \{y \geqslant \xi_n\} \subset B_d(\xi, r)$, where ξ_n is the point on level n before ξ. So

$$m(B_d(\xi, r)) \geqslant m(I_n) \geqslant \tfrac{1}{2}c(I_n) = 2^{-(n+1)},$$

and

$$\begin{aligned}
\int_0^{\sqrt{Ac_n}} \left[\log \frac{1}{m(B_d(\xi,r))}\right]^{1/2} dr &\leqslant \sum_{r=n}^{\infty} \sqrt{A \log 2}\left(\sqrt{c_r} - \sqrt{c_{r+1}}\right)\sqrt{r+2} \\
&\leqslant K\left(\sqrt{nc_n} + \sum_{r=n}^{\infty}\left(\sqrt{r+1} - \sqrt{r}\right)\sqrt{c_r}\right) \\
&\leqslant K\left(\sqrt{nc_n} + \sum_{r=n}^{\infty}\sqrt{\frac{1}{r}c_r}\right).
\end{aligned}$$

By the monotonicity of (c_n), $\sqrt{nc_n} < 2\sum_{r=\lceil n/2\rceil}^{\infty} \sqrt{\frac{1}{r}c_r}$ which goes to 0 as n goes to infinity, so that the whole right-hand side goes to 0 as we wish.

We also have to get a similar result when $x = i \in I$. Let N be $n(i)$, the level of i, and let i_n be the point on level n before i, for $n \leqslant N$. As before, for r such that $\sqrt{Ac_n} \leqslant r < \sqrt{Ac_{n-1}}$, then $\{y \geqslant i_n\} \subset B_d(i, r)$. In addition for $0 < r < \sqrt{Ac_N}$, then $\{i\} \subset B_d(i, r)$, so $m(B_d(i, r)) \geqslant m\{i\} = 2^{-(2N+1)}$. And so

$$\int_0^{\sqrt{Ac_M}} \left[\log \frac{1}{m(B_d(i,r))}\right]^{1/2} dr \leqslant K' \begin{cases} \sqrt{Mc_M} + \sqrt{Nc_N} + \sum_{r=M}^{N-1}\sqrt{\frac{1}{r}c_r} & M < N, \\ \sqrt{Nc_M} & M \geqslant N, \end{cases}$$

which goes to 0 uniformly in N as $M \to \infty$. Thus the sufficiency is proved.

Necessity of (4). For this we use the lower bound for d. If m is any probability measure on F, set ξ_0 to be 0, and recursively define ξ_{n+1} to be the point immediately beyond ξ_n which has no more m-mass in its subtree than the other point immediately beyond ξ_n, so that $m\{y \geqslant \xi_n\} \leqslant 2^{-n}$. We set $\xi := \lim_n \xi_n$ be a boundary point. If r is such that $\sqrt{\alpha c_n} \leqslant r < \sqrt{\alpha c_{n-1}}$, then $B_d(\xi, r) \subset \{y \geqslant \xi_n\}$, so

$$\int_0^{\sqrt{\alpha c_N}} \left[\log \frac{1}{m(B_d(\xi, r))}\right]^{1/2} dr \geqslant \sum_{n=N}^{\infty} \sqrt{\alpha \log 2}\,(\sqrt{c_n} - \sqrt{c_{n+1}})\,\sqrt{n+1}$$
$$\geqslant k \sum_{n=N+1}^{\infty} \sqrt{\frac{1}{n} c_n}, \quad \text{for some } k > 0.$$

Part (5). Fix $\xi \in C$ and let $A_n := \{\zeta \in C : \zeta > \xi_n\}$, where ξ_n is the point on level n before ξ. Let $\tau_\xi(t) := \inf\{s : L^X(\xi, s) > t\}$, and set

$$p(n,t) := \mathbf{P}_\xi(L^X(\cdot, \tau_\xi(t)) > 0 \text{ on } A_n).$$

The function p is monotone in each co-ordinate, and as L^X is jointly-continuous $\lim_n p(n,t) = 1$, $(t > 0)$. Thus p is positive for some (and hence all) n, and the strong Markov property gives us that

$$1 - p(0, Nt) \leqslant (1 - p(0,t))^N,$$

whence we deduce that $\lim_t p(0,t) = 1$. If we now set $C_t := \{\zeta \in C : L^X(\zeta, t) > 0\}$, which is open as L^X is continuous, we have proved that $C_t \uparrow C$ as $t \to \infty$, and by the compactness of C, we deduce that $C_T = C$ for some finite T. □

3. Various Proofs

Proof of Corollary 2. Firstly, if $\sum_n n^{(1+\epsilon)} b_n < \infty$, then by Hölder's inequality

$$\sum_n \sqrt{\frac{1}{n} c_n} \leqslant \left(\sum_n \frac{1}{n^{1+\epsilon}}\right)^{1/2} \left(\sum_n n^\epsilon c_n\right)^{1/2} < \infty,$$

because $$\sum_n n^\epsilon c_n = \sum_n b_n \left(\sum_{r=1}^n r^\epsilon\right) \leqslant \sum_n n^{(1+\epsilon)} b_n < \infty.$$

Secondly, if $\sum_n n b_n = \infty$, then as

$$\sum_{n=1}^N n b_n = \sum_{n=1}^N \sum_{r=n}^N b_r \leqslant \sum_{n=1}^N c_n,$$

we see that $\sum_n c_n = \infty$. We consider the sequence (d_n), defined by $d_n := \sqrt{c_n/n}$, and look at the set

$$A := \{n : d_n \geqslant 1/n\} = \{n : c_n \geqslant d_n\}.$$

If A is finite, then (d_n) is eventually more than (c_n) so its sum diverges. If A is infinite, there exists an increasing sequence (n_i) in A, so that by the monotonicity of (d_n)

$$\sum_n d_n \geqslant \sum_i \frac{n_i - n_{i-1}}{n_i} = \sum_i \left(1 - \frac{n_{i-1}}{n_i}\right) = \infty, \qquad \text{as } \prod_i \frac{n_{i-1}}{n_i} = 0. \qquad □$$

Proof of Theorem 3. Let k_n be set to $\mathsf{E}(V_n)$, the expected time to jump-up one level from n. We can expand V_n conditionally on the first jump as

$$V_n = \mathcal{E}(q_n) + \begin{cases} 0 & \text{with prob. } \mu_n/q_n \\ V_{n+1} + \tilde{V}_n & \text{with prob. } \lambda_n/q_n, \end{cases}$$

where $\tilde{V}_n$ has the same distribution as V_n, $\mathcal{E}(\alpha)$ is exponentially distributed with rate α, and all variables on the right-hand side are independent. We know the (k_n) are the minimal non-negative solutions to

$$k_n = \frac{1}{q_n} + \frac{\lambda_n}{q_n}(k_n + k_{n+1}), \qquad (n \geqslant 1),$$
$$\text{or} \qquad (\lambda_{n-1}\pi_{n-1}k_n) = \pi_n + (\lambda_n \pi_n k_{n+1})$$

(using $\mu_n \pi_n = \lambda_{n-1}\pi_{n-1}$). This has the required solution

$$k_n = \frac{\pi[n]}{\lambda_{n-1}\pi_{n-1}}.$$

Similarly the variance sequence $(\mathrm{Var}(V_n))$ will satisfy

$$\mathrm{Var}(V_n) = \frac{1}{q_n^2} + \frac{\lambda_n}{q_n}\left[\mathrm{Var}(V_{n+1}) + \mathrm{Var}(V_n)\right] + \frac{\lambda_n \mu_n}{q_n^2}\left[\mathsf{E}(V_{n+1} + V_n)\right]^2,$$
$$\text{or} \qquad \left[\lambda_{n-1}\pi_{n-1}\,\mathrm{Var}(V_n)\right] = \left[\lambda_n \pi_n\,\mathrm{Var}(V_{n+1})\right] + \frac{\pi[n]^2}{\lambda_{n-1}\pi_{n-1}} + \frac{\pi[n+1]^2}{\lambda_n \pi_n},$$

which has the desired solution.

Now we let h_n be equal to $\mathsf{E}(T_n)$, the expected time to down-jump one level from n to a particular point. We can decompose T_n as

$$T_0 = \mathcal{E}(\lambda_0) + \begin{cases} 0 & \text{with prob. } \frac{1}{2} \\ V_1 + \tilde{T}_0 & \text{with prob. } \frac{1}{2}, \end{cases}$$

$$T_n = \mathcal{E}(q_n) + \begin{cases} 0 & \text{with prob. } \frac{1}{2}\lambda_n/q_n \\ V_{n+1} + \tilde{T}_n & \text{with prob. } \frac{1}{2}\lambda_n/q_n \\ T_{n-1} + \tilde{T}_n & \text{with prob. } \mu_n/q_n, \end{cases} \qquad (n \geqslant 1)$$

where $\tilde{T}_n$ has the same distribution as T_n. So (h_n) is the minimal solution to

$$h_0 = \frac{1}{\lambda_0} + \tfrac{1}{2}h_0 + \tfrac{1}{2}\frac{\pi[1]}{\lambda_0 \pi_0}$$
$$h_n = \frac{1}{q_n} + \frac{\mu_n + \frac{1}{2}\lambda_n}{q_n}h_n + \frac{\mu_n}{q_n}h_{n-1} + \frac{\frac{1}{2}\lambda_n}{q_n}\frac{\pi[n+1]}{\lambda_n \pi_n} \qquad (n \geqslant 1),$$

or

$$\lambda_0 \pi_0 h_0 = 2 - \pi[1]$$
$$(\lambda_n \pi_n h_n) = 2(\lambda_{n-1}\pi_{n-1}h_{n-1}) + \pi_n + \pi[n] \qquad (n \geqslant 1).$$

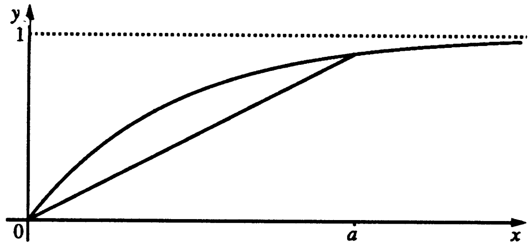

Figure 5. The concave function $y = 1 - e^{-x}$ and a subchord

We can use induction to show that

$$(\lambda_n \pi_n h_n) = 2^{n+1} - \pi[n+1], \qquad (n \geqslant 0)$$

whence the result.

Finally $\mathrm{Var}(T_n)$ will be the minimal non-negative solution to the following equations:

$$\mathrm{Var}(T_0) = \frac{1}{\lambda_0^2} + \tfrac{1}{2}\left[\mathrm{Var}(V_1) + \mathrm{Var}(T_0)\right] + \tfrac{1}{4}\left[\mathsf{E}(V_1 + T_0)\right]^2$$

$$\begin{aligned}\mathrm{Var}(T_n) = {} & \frac{1}{q_n^2} + \frac{\frac{1}{2}\lambda_n}{q_n}\left[\mathrm{Var}(T_n) + \mathrm{Var}(V_{n+1})\right] + \frac{\mu_n}{q_n}\left[\mathrm{Var}(T_n) + \mathrm{Var}(T_{n-1})\right] \\ & + \frac{\frac{1}{2}\lambda_n(\mu_n + \frac{1}{2}\lambda_n)}{q_n^2}\left[\mathsf{E}(T_n + V_{n+1})\right]^2 + \frac{\lambda_n \mu_n}{q_n^2}\left[\mathsf{E}(T_n + T_{n-1})\right]^2 \\ & - \frac{\lambda_n \mu_n}{q_n^2}\mathsf{E}(T_n + V_{n+1})\mathsf{E}(T_n + T_{n-1}), \qquad (n \geqslant 1),\end{aligned}$$

which, on setting u_n to equal $2^{-n}\lambda_n \pi_n \mathrm{Var}(T_n)$, can be rearranged to give

$$u_0 = \frac{(1+\pi_0)^2}{\lambda_0 \pi_0} + 2\sum_{r=0}^{\infty} \frac{\pi[r+1]^2}{\lambda_r \pi_r} < \infty$$

$$\begin{aligned}u_n = {} & u_{n-1} + 2^{-n}\sum_{r=n}^{\infty}(1 + I_{(r>n)})\frac{\pi[r+1]^2}{\lambda_r \pi_r} + 4\left(\frac{2^n - \pi[n+1]}{\lambda_n \pi_n} + \frac{2^{n-1} - \pi[n]}{\lambda_{n-1}\pi_{n-1}}\right) \\ & + \frac{2^{1-n}\pi_n}{q_n} + \frac{\mu_n}{q_n}\frac{\pi[n+1]^2}{\lambda_n \pi_n} + \frac{\lambda_n}{q_n}\frac{\pi[n]^2}{\lambda_{n-1}\pi_{n-1}} + \frac{2\pi[n]\pi[n+1]}{\lambda_n \pi_n + \lambda_{n-1}\pi_{n-1}} \qquad (n \geqslant 1).\end{aligned}$$

Hence

$$\begin{aligned}u_n \leqslant {} & u_{n-1} + 2^{1-n}\left(\sum_{r=n}^{\infty} \frac{\pi[r+1]^2}{\lambda_r \pi_r}\right) + 4\left(\frac{2^n}{\lambda_n \pi_n} + \frac{2^{n-1}}{\lambda_{n-1}\pi_{n-1}}\right) \\ & + \frac{2^{1-n}\pi_n}{q_n} + \left(\frac{\pi[n+1]^2}{\lambda_n \pi_n}\right) + 3\left(\frac{\pi[n]^2}{\lambda_{n-1}\pi_{n-1}}\right)\end{aligned}$$

So, remembering that (1) and (2) hold

$$u_n \leqslant A + B\sum_{r=0}^{n} \frac{2^r}{\lambda_r \pi_r} \leqslant K\sum_{r=0}^{n} \frac{2^r}{\lambda_r \pi_r} \qquad (n \geqslant 0),$$

for some constants A, B, and K. This delivers the required result. □

Proof of Lower Bound Lemma 4. The function $f(x) = 1 - e^{-x}$ is concave on the bounded interval $[0, a]$ (Fig. 5), so

$$f(x) \geqslant x\left(\frac{1-e^{-a}}{a}\right), \qquad \text{for } x \in [0, a].$$

Now

$$\mathsf{E}(f(X)) \geqslant \mathsf{E}(f(X); X \leqslant a) \geqslant \left(\frac{1-e^{-a}}{a}\right)\mathsf{E}(X; x \leqslant a),$$

and by Hölder's inequality

$$\mathsf{E}(X; X > a) = \|X\, I_{(X>a)}\|_1 \leqslant \|X\|_2 \mathsf{P}(X > a)^{1/2} \leqslant \left(K\mathsf{E}(X)\mathsf{P}(X > a)\right)^{1/2}.$$

Further $a\mathsf{P}(X > a) \leqslant \mathsf{E}(X)$, so we deduce that

$$\mathsf{E}(X; X > a) \leqslant (K/a)^{1/2}\mathsf{E}(X).$$

Choosing a to be $4K$, then $\mathsf{E}(X; X \leqslant a) \geqslant \frac{1}{2}\mathsf{E}(X)$ and

$$\mathsf{E}(f(X)) \geqslant \left(\frac{1-e^{-4K}}{8K}\right)\mathsf{E}(X).$$

□

Proof of Lemma 6. We can write $d^2(x, y)$ as

$$d^2(x,y) = u_1(x,x) + u_1(y,y) - \mathsf{E}_x(e^{-H(y)})u_1(y,y) - \mathsf{E}_y(e^{-H(x)})u_1(x,x).$$

For the upper bound, we use the fact that $(1 - e^{-x}) \leqslant x$ to show that

$$d^2(x,y) \leqslant \left[\sup_{z\in F} u_1(z,z)\right]\left(\mathsf{E}_x(H(y)) + \mathsf{E}_y(H(x))\right).$$

We can split the expected hitting times into four summands, two being sums of (V_n)'s from x and y to $x \wedge y$, and two of (T_n)'s from $x \wedge y$ to x and y. Theorem 3 tells us that the largest will be the down time from $x \wedge y$ to y, so

$$d^2(x,y) \leqslant 4\left[\sup_{z\in F} u_1(z,z)\right]\mathsf{E}_{x\wedge y}(H(y)).$$

For the upper bound we throw away some terms to reveal that

$$\begin{aligned} d^2(x,y) &\geqslant \left[\inf_{z\in F} u_1(z,z)\right]\left(\mathsf{E}_x(1 - e^{-H(y)})\right) \\ &\geqslant \left[\inf_{z\in F} u_1(z,z)\right]\left(\mathsf{E}_{x\wedge y}(1 - e^{-H(y)})\right). \end{aligned}$$

The function $u_1(z, z)$ is a continuous positive function on the compact space F, so the sup and the inf are finite and positive. Remembering that $\sum \frac{2^n}{\lambda_n \pi_n}$ is finite, Theorem 3 and the Lower Bound Lemma 4 together give us that

$$\mathsf{E}_{x\wedge y}(1 - e^{-H(y)}) \geqslant \alpha\mathsf{E}_{x\wedge y}(H(y)),$$

for some positive α. □

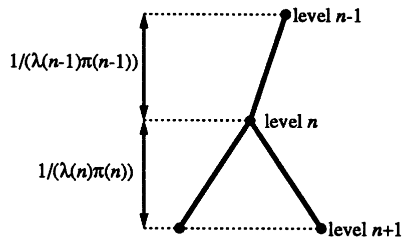

Figure 6. Construction of diffusion

4. Time Substitution

We aim to tie the Markov chain on the nodes, F, of the graph together with a Brownian diffusion on the graph, G, comprising of F and the edges. We can construct the diffusion by building up excursions from a point on level n as follows.

(We are only going to define the height and current edge of the process, the horizontal position being thus determined.) Given a reflecting Brownian motion, we take its excursions from 0 in order and make them excursions from our start point. The edge that each excursion follows is randomly selected according to the law assigning probability $\frac{1}{2}$ to going up, and probability $\frac{1}{4}$ to each of the down edges. Run this process until it hits levels $(n-1)$ or $(n+1)$, then repeat starting from the new node. We identify level n with the height $x_n := \sum_{r \geqslant n} \frac{1}{\lambda_r \pi_r}$ (Fig. 6).

The height process then becomes a reflecting Brownian motion on the interval $[0, \sum_n \frac{1}{\lambda_n \pi_n}]$. We notate the G-valued process as $(\tilde{X}_t)$, and the height process as $(\tilde{Y}_t)$. Then Trotter's Theorem allows us a jointly-continuous local time $\tilde{L}^Y$ for the height process. For a good treatment of local times see V.3 of Blumenthal and Getoor [1]. We can then time change $\tilde{Y}$ via

$$A_t := \sum_n \pi_n \tilde{L}^Y(x_n, t)$$
$$\tau_t := \inf\{s \geqslant 0 : A_s > t\}.$$

We note that A is continuous and (weakly) increasing; τ is right-continuous and strictly increasing; $A(\tau_t) = t$; and $\tau(A_t) \geqslant t$ with equality if and only if t is a point of right increase of A. We time change the diffusion by setting Y_t to be $\tilde{Y}(\tau_t)$, which by III.37 of Williams [6] is a strong Markov process on the support of A ($\{0\} \cup \{x_n : n \in \mathbb{N}\}$). The local time of $\tilde{Y}$ at a level before it hits an adjacent level (the holding time of the Y-process) is exponentially distributed, by the strong Markov property, and with the right normalisation of $\tilde{L}^Y$, our choice of (x_n) has ensured that the jump rates of Y agree with those of the BD-chain. In fact they are the same process. We can then define the local time of Y, L^Y, on $\mathbb{R}^+ \times \mathbb{N}$ by

$$L^Y(n,t) := \frac{1}{\pi_n} \int_0^t I_n(Y_s)\, ds,$$

and notice that by change of variable

$$\tilde{L}^Y(x_n, t) = \frac{1}{\pi_n} \int_0^t I_{x_n}(\tilde{Y}_s)\, dA_s = \frac{1}{\pi_n} \int_{J_+ \cap [0,t]} I_{x_n}(\tilde{Y}_s)\, dA_s = L^Y(n, A_t),$$

where J_+ is the set of the points of right-increase of A, which is all but countably many points of the set of points of increase of A.

We can also time change the G-diffusion by τ to produce $X_t := \tilde{X}(\tau_t)$, which is similarly a Markov chain on F with the same jump-rates as the process we studied in previous sections. Consideration of the time-truncation arguments of Rogers and Williams [6] should convince that the processes are the same. Given that conditions (1)–(4) of Theorem 1 hold, we can construct a jointly-continuous local time, L^X for X on F. It then follows that $\tilde{L}^X := \{L^X(x, A_t)\}_{x\in F}$ is a local time for $\tilde{X}$ at the points F in G. We can extend $\tilde{L}^X$ by interpolation on the edges to be continuous on G.

It is now possible to construct processes on the boundary, C, via time changes of X and $\tilde{X}$ induced by

$$A_t^\partial := \int_C L^X(\xi,t)\, c(d\xi) = L^Y(\infty,t),$$

$$\text{and} \qquad \tilde{A}_t^\partial := \int_C \tilde{L}^X(\xi,t)\, c(d\xi) = \tilde{L}^Y(0,t),$$

with τ^∂ and $\tilde{\tau}^\partial$ respectively the right-continuous inverses. This gives us the strong Markov R-processes $Z_t := X(\tau_t^\partial)$ and $\tilde{Z}_t := \tilde{X}(\tilde{\tau}_t^\partial)$. By the continuity of the local times,

$$A^\partial(A_t) = \lim_{n\to\infty} 2^{-n} \sum_{x\in\text{level } n} L^X(x, A_t) = \lim_{n\to\infty} 2^{-n} \sum_{x\in\text{level } n} \tilde{L}^X(x,t) = \tilde{A}_t^\partial.$$

For any (A,τ)-type pair, $\tau_t < s \iff t < A_s$, hence

$$\tilde{\tau}_t^\partial < s \iff t < \tilde{A}_s^\partial \iff t < A^\partial(A_s) \iff \tau_t^\partial < A_s \iff \tau(\tau_t^\partial) < s,$$

whence we can deduce that $\tilde{\tau}_t^\partial = \tau(\tau_t^\partial)$ and that $Z_t = \tilde{Z}_t$. The process Z also has a jointly-continuous local time L^Z, given by

$$L^Z(\xi,t) = L^X(\xi,\tau_t^\partial) = \tilde{L}^X(\xi,\tilde{\tau}_t^\partial).$$

In summary we can say that Figure 7 commutes. We have thus produced the same process by taking local time on the boundary of both the chain and the diffusion, which allows us to work with whichever is more appropriate for the current problem.

$$\begin{array}{ccccc} \tilde{Y} & \xleftarrow{n} & \tilde{X} & \xrightarrow{\tilde{\tau}^\partial} & \tilde{Z} \\ \big\downarrow \tau & & \big\downarrow \tau & & \big\| \\ Y & \xleftarrow{n} & X & \xrightarrow{\tau^\partial} & Z \end{array}$$

Figure 7. A commutative diagram of processes

5. The Boundary Process

We now finally turn our attention to the boundary process Z. We know that the graph processes (both chain and diffusion) spend no intervals of time on the boundary, but rather the set of times at which they visit the boundary is a Cantor set obtained by removing the open excursion intervals from the time axis. The process will (almost surely) not be back in its original position at the right-hand endpoints of these intervals — even though it will return to its original position uncountably often almost immediately. As the boundary Cantor set is totally disconnected, we see that Z must be a very discontinuous process. In fact Z is discontinuous at a dense, though countable, set of times.

LEMMA 7. *Z is FD.*

Proof. By adapting the argument at the end of III.38 of Williams [6], we can show that Z is FD if $\mathsf{E}_\xi(1 - e^{-\tilde{H}(\eta)})$ goes to 0 as $\eta \to \xi$ in C, where $\tilde{H}(\eta) := \inf\{t \geqslant 0 : Z_t = \eta\}$. Now $\tilde{H}(\eta)$ is almost surely a point of right-increase of A^∂, so $\tilde{H}(\eta) = A^\partial(H(\eta))$. We can write $H(\eta)$ as

$$H(\eta) \stackrel{\mathcal{D}}{=} U_n := \sum_{r=n}^{\infty}(T_r + V_{r+1}), \qquad \text{where } n = n(\xi \wedge \eta),$$

$$\text{and} \qquad A^\partial(H(\eta)) \stackrel{\mathcal{D}}{=} A^*(U_n) := \sum_{r=n}^{\infty} A_r^*,$$

with A_r^* the local time on the boundary notched up while a version of the process did an up-down $T_r + V_{r+1}$. Then $U_n \downarrow 0$, $A^*(U_n) \downarrow 0$, and thus $(1 - e^{-A^*(U_n)}) \downarrow 0$ as $\eta \to \xi$, giving the result. □

By VI.28 of Rogers and Williams [5], there exists a *Lévy system* (N, H) for Z. In our case $H_t = t$, and N as usual is a *kernel*, that is a function

$$N : (C, \mathcal{B}(C)) \longrightarrow [0, \infty],$$
$$\text{such that } N(\cdot, \Gamma) \text{ is } \mathcal{B}(C)\text{-measurable, for all } \Gamma \text{ in } \mathcal{B}(C),$$
$$\text{and } N(\xi, \cdot) \text{ is a } \sigma\text{-finite measure on } \mathcal{B}(C), \text{ for all } \xi \text{ in } C.$$

In addition $N(\xi, \{\xi\}) = 0$ for all ξ in C, and N has the Lévy property, in the sense that for any non-negative borel-measurable function f on $C \times C$ with $f(\xi, \xi) = 0$ for all ξ in C, then

$$M_t^f := \sum_{s \leqslant t} f(Z_{s-}, Z_s) - \int_{(0,t]} ds \int_C N(Z_{s-}, d\xi) f(Z_{s-}, \xi)$$

is a martingale, if the expectation of either term is finite. We can think of $N(\xi, d\eta)$ as the rate at which jumps from ξ to $d\eta$ of Z occur.

We can calculate this directly using excursion theory, and we will not need any more than is in Rogers [3]. By thinking of the diffusion height process, Proposition 2 of Rogers [3] tells us that the rate of excursions from ξ in C to level n or before is

$$\frac{1}{x_n} = \frac{1}{\sum_{r=n}^{\infty} a_r}, \qquad \text{where} \quad a_r := \frac{1}{\lambda_r \pi_r}.$$

(The factor $\frac{1}{2}$ is lost because we have reflection at the boundary so all our excursions go up.) Therefore the rate of excursions from ξ which have their furthest extent on level n is simply the difference

$$\frac{1}{x_n} - \frac{1}{x_{n-1}} = \begin{cases} \frac{a_{n-1}}{x_n x_{n-1}} & n \geqslant 1 \\ \frac{1}{x_0} & n = 0. \end{cases}$$

The chance that such an excursion ends up in $d\eta \subset C$ is then (by symmetry) exactly $2^n c(d\eta)$ if $\xi \wedge \eta$ is on or beyond level n, and 0 if not. We deduce that the rate of excursions from ξ to $d\eta$ is given by the following:

LEMMA 8.

$$N(\xi, d\eta) = c(d\eta)\left(\frac{1}{x_0} + \sum_{n=1}^{n(\xi\wedge\eta)} 2\frac{b_{n-1}}{x_n x_{n-1}}\right).$$

Proof. For ξ in C, let ξ_n be the point on level n before ξ, and let A_n be $\{\zeta \in C : \zeta > \xi_n\}$. For any subset B of $C \backslash A_n$, we notice that $N(\cdot, B)$ is constant on A_n by symmetry. We can set $f(x,y) := I_{A_n}(x) I_B(y)$, and $T := \inf\{t : Z_t \notin A_n\}$. Then $\mathsf{E}(M_T^f) = 0$ implies that

$$N(\xi, B) = \frac{\mathsf{P}(Z_T \in B)}{\mathsf{E}(T)},$$

and the result is proved. □

Example. In the geometric case, with

$$\lambda_n = \alpha^n, \qquad \mu_n = \alpha^n/\gamma, \qquad \text{and} \quad \pi_n = \left(\frac{\alpha - \gamma}{\alpha}\right)(\gamma/\alpha)^n,$$

$$\text{then} \quad \frac{1}{\lambda_n \pi_n} = \frac{\alpha\gamma^{-n}}{\alpha - \gamma}, \quad x_n = \frac{\alpha\gamma^{-(n-1)}}{(\alpha-\gamma)(\gamma-1)}, \quad b_n = \frac{\alpha(2/\gamma)^n}{\alpha - \gamma},$$

$$\text{and} \quad N(\xi, d\eta) = c(d\eta)\big(A(2\gamma)^n + B\big) \qquad \text{where } n = n(\xi \wedge \eta);\ A,\ B > 0.$$

The conditions of Theorem 1 translate as (1) $\alpha > \gamma$; (2) $\gamma > 1$; (3)&(4) $\gamma > 2$, and we assume that all these hold. Then H_n, the first time to leave A_n, will be exponentially distributed with rate

$$N(\xi, A_n^c) = \frac{A(\gamma^n - 1)}{2(\gamma - 1)} + B(1 - 2^{-n}).$$

We can form an analogue of the Hausdorff dimension of a diffusion as

$$\lim_{n\to\infty} \frac{\log \mathsf{E}(H_n)}{-n} = \log \gamma.$$

This can be seen as a measure of the asymptotic neighbourhood escape rate of the process. As γ gets larger it takes longer to escape as the downward pressure inhibits larger excursions. The normal scaling logarithm in the denominator is missing as there is no obviously natural metric on C. □

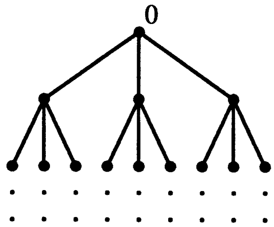

Figure 8. Ternary Tree

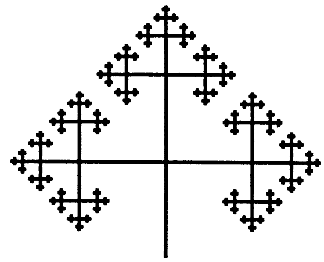

Figure 9. Ray-Knight compactification

There is (at least) one easy generalisation of the chain, keeping the same basic structure, by taking the M-ary tree with M down edges, all equally likely, from each node (Fig. 8). In the case of $M = 3$, the Ray-Knight compactification can be thought of as a tree-like graph (Fig. 9).

Everything thus described still holds, with the alteration of the down-jump time line of Theorem 3 to

$$\mathsf{E}T_n = \frac{M^{n+1} - \pi[n+1]}{\lambda_n \pi_n}, \quad \mathrm{Var}(T_n) \leqslant K \frac{M^n}{\lambda_n \pi_n} \sum_{r=0}^{n} \frac{M^r}{\lambda_r \pi_r}, \quad \textit{for some } K.$$

And Theorem 1 holds with $b_n := M^n / \lambda_n \pi_n$. The Lévy kernel N is as stated above for these new values of (b_n), and the number 2 replaced by M.

Acknowledgement. This paper would not have been possible without the inspiration of David Williams, who saw from afar that this chain would do exciting things, and whose great enthusiasm sparked my own. I am also indebted to Ben Hambly for his advice at crucial points throughout, and to Martin Barlow, David Hobson and David Dean for many helpful comments.

References

[1] R. M. Blumenthal and R. K. Getoor. *Markov Processes and Potential Theory* (Academic Press 1968).

[2] M. B. Marcus and J. Rosen. Sample Path Properties of the Local Times of Strongly Symmetric Markov Processes via Gaussian Processes. *Ann. Prob.* (to appear)

[3] L. C. G. Rogers. A Guided Tour through Excursions. *Bull. London Math. Soc.*, **21** (1989), 305–341.

[4] L. C. G. Rogers and D. Williams. Construction and Approximation of Transition Matrix Functions. *Analytic and Geometric Stochastics*, suppl. to *Adv. App. Prob.*, **18** (1986), 133–160.

[5] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales, vol. 2: Itô Calculus* (Wiley 1987).

[6] D. Williams. *Diffusions, Markov Processes, and Martingales, vol. 1: Foundations* (Wiley, 1979).

[7] R. W. Wolff. *Stochastic Modelling and the Theory of Queues* (Prentice-Hall 1989).