MARTIN CAPACITY FOR MARKOV CHAINS

Itai Benjamini[1], Robin Pemantle [2] and Yuval Peres [3]

Cornell University, University of Wisconsin and University of California

Abstract

The probability that a transient Markov chain, or a Brownian path, will ever visit a given set Λ, is classically estimated using the capacity of Λ with respect to the Green kernel $G(x,y)$. We show that replacing the Green kernel by the Martin kernel $G(x,y)/G(0,y)$ yields improved estimates, which are exact up to a factor of 2. These estimates are applied to random walks on lattices, and also to explain a connection found by R. Lyons between capacity and percolation on trees.

Keywords : capacity, Markov chain, hitting probability, Brownian motion, tree, percolation.

Subject classification : Primary: 60J45, 60J10; Secondary: 60J65, 60J15, 60K35.

[1]Research partially supported by the U. S. Army Research Office through the Mathematical Sciences Institute of Cornell University.

[2]Research supported in part by National Science Foundation grant # DMS 9300191, by a Sloan Foundation Fellowship, and by a Presidential Faculty Fellowship.

[3]Research partially supported by NSF grant # DMS-9404391 and a Junior Faculty Fellowship from the Regents of the University of California .

1 Introduction

Kakutani (1944) discovered that a compact set $\Lambda \subseteq \mathbb{R}^d$ is hit with positive probability by a d-dimensional Brownian motion ($d \geq 3$) if and only if Λ has positive Newtonian capacity. A more quantitative relation holds between this probability and capacity. Under additional assumptions on the set Λ it is well known that the hitting probability is estimated by capacity up to a constant factor (often unspecified). Our main result, Theorem 2.2, estimates this probability (for general Markov chains) by a different capacity up to a factor of 2. We first state this estimate for Brownian motion.

Proposition 1.1 *Let $\{B_d(t)\}$ denote standard d-dimensional Brownian motion with $B_d(0) = \mathbf{0}$ and $d \geq 3$. Let Λ be any closed set in $\mathbb{R}^d$. Then*

$$\frac{1}{2}\mathrm{Cap}_K(\Lambda) \leq \mathbf{P}[\exists t > 0 : B_d(t) \in \Lambda] \leq \mathrm{Cap}_K(\Lambda) \tag{1}$$

where

$$K(x,y) = \frac{||y||^{d-2}}{||x-y||^{d-2}}$$

for $x \neq y$ in $\mathbb{R}^d$, and $K(x,x) = \infty$. Here $||x-y||$ is the Euclidean distance and

$$\mathrm{Cap}_K(\Lambda) = \left[\inf_{\mu(\Lambda)=1} \int_\Lambda \int_\Lambda K(x,y)\, d\mu(x)\, d\mu(y)\right]^{-1}.$$

Remarks:
1. More detailed definitions will be given in the next section.
2. The constants 1/2 and 1 in (1) are sharp (see Section 4).

Note that while the Green kernel $G(x,y) = ||x-y||^{2-d}$, and hence the corresponding capacity, are translation invariant, the hitting probability of a set Λ by standard d-dimensional Brownian motion is not translation invariant, but is invariant under scaling. This scale-invariance is shared by the Martin kernel $K(x,y) = G(x,y)/G(0,y)$.

The rest of this paper is organized as follows. Section 2 states and proves the connection between the probability of a Markov chain hitting a set and the Martin capacity of the set. Section 3 gives several examples, including a relation between simple random walk in three dimensions and the time-space chain arising from simple random walk in the plane. The ratio of 2 between the two sides in the estimate (1) may remind the reader of a theorem of Lyons (1992) that gives a precise relation between capacity and independent percolation on trees. In SectionT 4 we show how recognizing a "hidden" Markov chain in the percolation setting leads to a very short proof of this theorem. In Section 5 we give the proof of Proposition 1.1 concerning Brownian motion. Section 6 discusses motivations and extensions.

2 Main result

First we recall some potential theory notions.

Definition 2.1 *Let Λ be a set and $\mathcal{B}$ a σ-field of subsets of Λ. Given a measurable function $F : \Lambda \times \Lambda \to [0, \infty]$ and a finite measure μ on $(\Lambda, \mathcal{B})$, the F-energy of μ is*

$$I_F(\mu) = \int_\Lambda \int_\Lambda F(x, y)\, d\mu(x)\, d\mu(y).$$

The capacity of Λ in the kernel F is

$$\mathrm{Cap}_F(\Lambda) = \left[\inf_\mu I_F(\mu)\right]^{-1}$$

where the infimum is over probability measures μ on $(\Lambda, \mathcal{B})$ and by convention, $\infty^{-1} = 0$.

If Λ is contained in Euclidean space, we always take $\mathcal{B}$ to be the Borel σ-field; if Λ is countable, we take $\mathcal{B}$ to be the σ-field of all subsets. When Λ is countable we also define the *asymptotic capacity* of Λ in the kernel F:

$$\mathrm{Cap}_F^{(\infty)}(\Lambda) = \inf_{\{\Lambda_0 \text{ finite}\}} \mathrm{Cap}_F(\Lambda \setminus \Lambda_0). \tag{2}$$

Let $\{p(x,y) : x, y \in Y\}$ be transition probabilities on the countable set Y, i.e. $\sum_y p(x,y) = 1$ for every $x \in Y$. Let $\rho \in Y$ be a distinguished starting state and let $\{X_n : n \geq 0\}$ be a Markov chain with $\mathbf{P}[X_{n+1} = y \,|\, X_n = x] = p(x,y)$.

Define the Green function

$$G(x,y) = \sum_{n=0}^{\infty} p^{(n)}(x,y) = \sum_{n=0}^{\infty} \mathbf{P}_x[X_n = y]$$

where $p^{(n)}(x,y)$ are the n-step transition probabilities and $\mathbf{P}_x$ is the law of the chain $\{X_n : n \geq 0\}$ when $X_0 = x$. We want to estimate the probability that a sample path $\{X_n\}$ visits a set $\Lambda \subseteq Y$. We assume that the Markov chain $\{X_n\}$ is transient; in fact, it suffices to assume that $G(x,y) < \infty$ for all $x, y \in \Lambda$.

Theorem 2.2 *Let $\{X_n\}$ be a transient Markov chain on the countable state space Y with initial state ρ and transition probabilities $p(x,y)$. For any subset Λ of Y we have*

$$\frac{1}{2}\mathrm{Cap}_K(\Lambda) \leq \mathbf{P}_\rho[\exists n \geq 0 : X_n \in \Lambda] \leq \mathrm{Cap}_K(\Lambda) \tag{3}$$

and

$$\frac{1}{2}\mathrm{Cap}_K^{(\infty)}(\Lambda) \leq \mathbf{P}_\rho[X_n \in \Lambda \textit{ infinitely often }] \leq \mathrm{Cap}_K^{(\infty)}(\Lambda) \tag{4}$$

where K is the Martin kernel

$$K(x,y) = \frac{G(x,y)}{G(\rho,y)} \tag{5}$$

defined using the initial state ρ.

Remarks:

1. The Martin kernel $K(x,y)$ can obviously be replaced by the symmetric kernel $\frac{1}{2}(K(x,y) + K(y,x))$ without affecting the energy of measures or the capacity of sets.
2. If the Markov chain starts according to an initial measure π on the state space, rather than from a fixed initial state, the theorem may be applied by adding an abstract initial state ρ with transition probabilities $p(\rho, y) = \pi(y)$ for $y \in Y$.

PROOF: (i) The right hand inequality in (3) follows from an entrance time decomposition. Let τ be the first hitting time of Λ and let ν be the (possibly defective) hitting measure $\nu(x) = \mathbf{P}_\rho[X_\tau = x]$ for $x \in \Lambda$. Then

$$\nu(\Lambda) = \mathbf{P}[\exists n \geq 0 : X_n \in \Lambda]\,. \tag{6}$$

Now for all $y \in \Lambda$:

$$\int G(x,y)\, d\nu(x) = \sum_{x \in \Lambda} \mathbf{P}_\rho[X_\tau = x] G(x,y) = G(\rho, y).$$

Thus $\int K(x,y)\, d\nu(x) = 1$ for every $y \in \Lambda$. Consequently

$$I_F\left(\frac{\nu}{\nu(\Lambda)}\right) = \nu(\Lambda)^{-2} I_F(\nu) = \nu(\Lambda)^{-1},$$

so that $\mathrm{Cap}_K(\Lambda) \geq \nu(\Lambda)$. By (6), this proves half of (3).

To establish the left hand inequality in (3) we use the second moment method. Given a probability measure μ on Λ, consider the random variable

$$Z = \int_\Lambda G(\rho, y)^{-1} \sum_{n=0}^{\infty} \mathbf{1}_{\{X_n = y\}}\, d\mu(y).$$

By Tonelli and the definition of G,

$$\mathbf{E}_\rho Z = 1. \tag{7}$$

Now we bound the second moment:

$$\begin{aligned}
\mathbf{E}_\rho Z^2 &= \mathbf{E}_\rho \int_\Lambda \int_\Lambda G(\rho,y)^{-1} G(\rho,x)^{-1} \sum_{m,n=0}^{\infty} \mathbf{1}_{\{X_m = x, X_n = y\}}\, d\mu(x)\, d\mu(y) \\
&\leq 2\mathbf{E}_\rho \int_\Lambda \int_\Lambda G(\rho,y)^{-1} G(\rho,x)^{-1} \sum_{0 \leq m \leq n < \infty} \mathbf{1}_{\{X_m = x, X_n = y\}}\, d\mu(x)\, d\mu(y).
\end{aligned}$$

For each m we have

$$\mathbf{E}_\rho \sum_{n=m}^{\infty} \mathbf{1}_{\{X_m = x, X_n = y\}} = \mathbf{P}_\rho[X_m = x] G(x,y).$$

Summing this over all $m \geq 0$ yields $G(\rho, x)G(x, y)$, and therefore

$$\mathbf{E}_\rho Z^2 \leq 2 \int_\Lambda \int_\Lambda G(\rho, y)^{-1} G(x, y)\, d\mu(x)\, d\mu(y) = 2 I_K(\mu).$$

By Cauchy-Schwarz and (7),

$$\mathbf{P}_\rho[\exists n \geq 0 : X_n \in \Lambda] \geq \mathbf{P}_\rho[Z > 0] \geq \frac{(\mathbf{E}_\rho Z)^2}{\mathbf{E}_\rho Z^2} \geq \frac{1}{2 I_K(\mu)}.$$

Since the left hand side does not depend on μ, we conclude that

$$\mathbf{P}_\rho[\exists n \geq 0 : X_n \in \Lambda] \geq \frac{1}{2} \mathrm{Cap}_K(\Lambda)$$

as claimed.

To infer (4) from (3) observe that since $\{X_n\}$ is a transient chain, almost surely every state is visited only finitely often and therefore

$$\{X_n \in \Lambda \ \text{ infinitely often }\} = \bigcap_{\Lambda_0 \text{ finite}} \{\exists n \geq 0 : X_n \in \Lambda \setminus \Lambda_0\} \quad \text{a.s.}$$

Applying (3) and the definition (2) of asymptotic capacity yields (4). $\square$

3 Corollaries and examples

This section is devoted to deriving some consequences of Theorem 2.2. The first involves a widely applicable equivalence relation between distributions of random sets.

Definition: Say that two random subsets W_1 and W_2 of a countable space are *intersection-equivalent* (or more precisely, that their *laws* are intersection-equivalent) if there exist positive finite constants C_1 and C_2, such that for every subset A of the space,

$$C_1 \leq \frac{\mathbf{P}[W_1 \cap A \neq \emptyset]}{\mathbf{P}[W_2 \cap A \neq \emptyset]} \leq C_2\,.$$

It is easy to see that if W_1 and W_2 are intersection-equivalent then $C_1 \leq \mathbf{P}[|W_1 \cap A| = \infty]/\mathbf{P}[|W_2 \cap A| = \infty] \leq C_2$ for all sets A, with the same constants C_1 and C_2. An immediate corollary of Theorem 2.2 is the following, one instance of which is given in Corollary 3.4.

Corollary 3.1 *Suppose the Green's functions for two Markov chains on the same state space (with the same initial state) are bounded by constant multiples of each other. (It suffices that this bounded ratio property holds for the corresponding Martin kernels $K(x,y)$ or for their symmetrizations $K(x,y)+K(y,x)$.) Then the ranges of the two chains are intersection-equivalent.*

Lamperti (1963) gave an alternative criterion for $\{X_n\}$ to visit the set Λ infinitely often. Fix $b>1$. With the notations of Theorem 2.2, denote $Y(n)=\{x\in Y : b^{-n-1}<G(\rho,x)\le b^{-n}\}$.

Corollary 3.2 (Lamperti's Wiener Test) *Assume that the set $\{x\in Y : G(\rho,x)>1\}$ is finite. Also, assume that there exists a constant C such that for all sufficiently large m and n we have*

$$G(x,y)<Cb^{-(m+n)} \tag{8}$$

for all $x\in Y(m)$ and $y\in Y(m+n)$. Then

$$\mathbf{P}[X_n\in\Lambda \textit{ infinitely often}]>0 \iff \sum_{n=1}^{\infty} b^{-n}\mathrm{Cap}_G(\Lambda\cap Y(n))=\infty\,. \tag{9}$$

Sketch of proof: Clearly $\sum_{n=1}^{\infty} b^{-n}\mathrm{Cap}_G(\Lambda\cap Y(n))=\infty$ if and only if $\sum_n \mathrm{Cap}_K(\Lambda\cap Y(n))=\infty$. The equivalence (9) then follows from a version of the Borel-Cantelli lemma proved in Lamperti's paper (a better proof is in Kochen and Stone (1964)).

Lamperti's Wiener test is useful in many cases; however the condition (8) excludes some natural transient chains such as simple random walk on a binary tree. Next, we deduce from Theorem 2.2 a criterion for a recurrent Markov chain to visit its initial state infinitely often within a prescribed time set.

Corollary 3.3 *Let $\{X_n\}$ be a recurrent Markov chain on the countable state space Y, with initial state $X_0=\rho$ and transition probabilities $p(x,y)$. For nonnegative integers $m\le n$ denote*

$$\tilde{G}(m,n)=\mathbf{P}[X_n=\rho\,|\,X_m=\rho]=p^{(n-m)}(\rho,\rho)$$

and

$$\tilde{K}(m,n) = \frac{\tilde{G}(m,n)}{\tilde{G}(0,n)}.$$

Then for any set of times $A \subseteq \mathbf{Z}^+$:

$$\frac{1}{2}\mathrm{Cap}_{\tilde{K}}(A) \leq \mathbf{P}[\exists n \in A : X_n = \rho] \leq \mathrm{Cap}_{\tilde{K}}(A) \tag{10}$$

and

$$\frac{1}{2}\mathrm{Cap}^{(\infty)}_{\tilde{K}}(A) \leq \mathbf{P}[\sum_{n \in A} \mathbf{1}_{\{X_n=\rho\}} = \infty] \leq \mathrm{Cap}^{(\infty)}_{\tilde{K}}(A). \tag{11}$$

PROOF: Consider the space-time chain $\{(X_n, n) : n \geq 0\}$ on the state space $Y \times \mathbf{Z}^+$. This chain is obviously transient; let G denote its Green function. Since $G((\rho,m),(\rho,n)) = \tilde{G}(m,n)$ for $m \leq n$, applying Theorem 2.2 with $\Lambda = \{\rho\} \times A$ shows that (10) and (11) follow respectively from (3) and (4). $\square$

Example 1: *Random walk on* $\mathbf{Z}$. The moral of this example will be that Borel-Cantelli does not always correctly settle questions about return times of random walks; similar examples may be found in Ruzsa and Székely (1982) and Lawler (1991).

Let S_n be the partial sums of mean-zero, finite variance, i.i.d. integer random variables. By the local central limit theorem (cf. Spitzer 1964),

$$\tilde{G}(0,n) = \mathbf{P}[S_n = 0] \approx cn^{-1/2}$$

provided that the summands $S_n - S_{n-1}$ are aperiodic. Therefore

$$\mathbf{P}[\sum_{n \in A} \mathbf{1}_{\{S_n=0\}} = \infty] > 0 \Leftrightarrow \mathrm{Cap}^{(\infty)}_F(A) > 0, \tag{12}$$

with $F(m,n) = (n^{1/2}/(n-m+1)^{1/2})\mathbf{1}_{\{m \leq n\}}$. By the Hewitt-Savage zero-one law, the event in (12) must have probability zero or one. Consider the special case in which A consists of separated blocks of integers:

$$A = \bigcup_{n=1}^{\infty} [2^n, 2^n + L_n]. \tag{13}$$

A standard calculation (e.g., with the Wiener test applied to the time-space chain) shows that in this case $S_n = 0$ for infinitely many $n \in A$ with probability one, if and only if $\sum_n L_n^{1/2} 2^{-n/2} = \infty$. On the other hand, the expected number of returns $\sum_{n \in A} \mathbf{P}[S_n = 0]$ is infinite if and only if $\sum_n L_n 2^{-n/2} = \infty$. Thus an infinite expected number of returns in a time set does not suffice for almost sure return in the time set. When the walk is periodic, i.e.

$$r = \gcd\{n : \mathbf{P}[S_n = 0] > 0\} > 1,$$

the same criterion holds as long as A is contained in $r\mathbf{Z}^+$.

In some cases, the criterion of Corollary 3.3 can be turned around and used to estimate asymptotic capacity. For instance, if $\{S'_n\}$ is an independent random walk with the same distribution as $\{S_n\}$ and A is the random set $A = \{n : S'_n = 0\}$, then the positivity of $\mathrm{Cap}_F^{(\infty)}(A)$ follows from the recurrence of the planar random walk $\{(S_n, S'_n)\}$. This implies that the "discrete Hausdorff dimension" of A (in the sense of Barlow and Taylor (1992)) is almost surely $1/2$; detailed estimates of the discrete Hausdorff measure of A were obtained by Khoshnevisan (1993).

Example 2: *Random walk on* $\mathbf{Z}^2$. Now we assume that S_n are partial sums of aperiodic, mean-zero, finite variance i.i.d. random variables in $\mathbf{Z}^2$. Let $A \subseteq \mathbf{Z}$. Again, $\mathbf{P}[S_n = 0$ for infinitely many $n \in A]$ is zero or one and it is one if and only if $\mathrm{Cap}_F^{(\infty)}(A) > 0$ where $F(m,n) = (n/(1+n-m))\mathbf{1}_{\{m \le n\}}$. This follows from the local central limit theorem (cf. Spitzer 1964) which ensures that

$$\tilde{G}(0,n) = \mathbf{P}[S_n = 0] \approx cn^{-1} \text{ as } n \to \infty.$$

For instance, if A consists of disjoint blocks

$$A = \bigcup_n [2^n, 2^n + L_n]$$

then $\mathrm{Cap}_F^{(\infty)}(A) > 0$ if and only if $\sum_n 2^{-n} L_n / \log L_n = \infty$. The expected number of returns to zero is infinite if and only if $\sum 2^{-n} L_n = \infty$.

Comparing the kernel F with the Martin kernel for simple random walk on $\mathbf{Z}^3$ leads to the next corollary.

Corollary 3.4 *For $d = 2,3$, let $\{S_n^{(d)}\}$ be a truly d-dimensional random walk on the d-dimensional lattice, with increments of mean zero and finite variance. Assume that the walks are aperiodic, i.e., the set of positive integers n for which $\mathbf{P}[S_n^{(d)} = 0] > 0$ has g.c.d. $= 1$. Then there exist positive finite constants C_1 and C_2 such that for any set of positive integers A,*

$$C_1 \leq \frac{\mathbf{P}[S_n^{(2)} = 0 \textit{ for some } n \in A]}{\mathbf{P}[S_n^{(3)} \in \{0\} \times \{0\} \times A \textit{ for some } n]} \leq C_2\,, \tag{14}$$

where $\{0\} \times \{0\} \times A = \{(0,0,k) : k \in A\}$. Consequently,

$$\mathbf{P}[S_n^{(2)} = 0 \textit{ for infinitely many } n \in A] = \mathbf{P}[S_n^{(3)} \in \{0\} \times \{0\} \times A \textit{ infinitely often}]. \tag{15}$$

Note that both sides of (15) take only the values 0 or 1. Corollary 3.4 follows from Corollary 3.1, in conjunction with Example 2 above and the asymptotics $G(0,x) \sim c/|x|$ as $|x| \to \infty$ for the random walk $S_n^{(3)}$ (cf. Spitzer (1964)). The Wiener test implies the equality (15) but not the estimate (14). Erdös (1961) and McKean (1961) showed that for $A = \{\text{primes}\}$, the left-hand side of (15) is 1. The corresponding result for the right-hand side is in Kochen and Stone (1964). To see why Corollary 3.4 is surprising, observe that the space-time chain $\{(S_n^{(2)}, n)\}$ travels to infinity faster than $S_n^{(3)}$, yet by Corollary 3.4, the same subsets of lattice points on the positive z-axis are hit infinitely often by the two processes.

Example 3: *Riesz-type kernels.* The analogues of the Riesz kernels in the discrete setting are the kernels

$$F_\alpha(x,y) = \frac{||y||^\alpha}{1 + ||x-y||^\alpha}$$

on $\mathbf{Z}^d$, where $||\cdot||$ is any norm. We write $\text{Cap}_\alpha^{(\infty)}$ for $\text{Cap}_{F_\alpha}^{(\infty)}$. By Theorem 2.2, the asymptotics for the Green function, and the Hewitt-Savage law, simple random walk on $\mathbf{Z}^d$ visits a set $\Lambda \subseteq \mathbf{Z}^d$ i.o. a.s. if and only if $\text{Cap}_{d-2}^{(\infty)}(\Lambda) > 0$. More generally, if a random walk $\{S_n\}$ on the d-dimensional lattice has a Green function satisfying $G(0,x) \sim c|x|^{\alpha-d}$ as $|x| \to \infty$, then Theorem 2.2 implies that $S_n \in \Lambda$ for infinitely many n a.s. iff $\text{Cap}_{d-\alpha}^{(\infty)}(\Lambda) > 0$. These asymptotics for the Green function are known to hold for many increment distributions in the domain of attraction of an α-stable distribution. (cf. Williamson (1968) for some sufficient conditions.)

Given a set of digits $D \subseteq \{0, 1, \ldots, b-1\}$ containing zero, consider "the integer Cantor set"

$$\Lambda(D, b) = \{\sum_{n=0}^{N} a_n b^n : a_n \in D \,\text{for all}\, n\,, \text{and}\, N \geq 0\}.$$

It may be shown that $\text{Cap}_\alpha^{(\infty)}(\Lambda(D, b)) > 0$ if and only if $|D| \geq b^\alpha$. This, together with Example 3, motivates defining the (discrete) dimension of $\Lambda \subseteq \mathbf{Z}^d$ by

$$\dim(\Lambda) = \inf\{\alpha : \text{Cap}_\alpha^{(\infty)}(\Lambda) = 0\}. \tag{16}$$

Corollary 8.4 in Barlow and Taylor (1992) shows that this definition is equivalent to the definition of discrete Hausdorff dimension in that paper.

When applying Theorem 2.2, it is often useful to know whether for the Markov chain under consideration, the probability of visiting a set infinitely often must be either 0 or 1. As remarked before, random walks on $\mathbf{Z}^d$ (or any abelian group) have this property by the Hewitt-Savage zero-one law. Easy examples show that this fails for random walk on a free group. More generally, the following "folklore" criterion holds.

Proposition 3.5 *Let μ be a probability measure whose support generates a countable group Y, and let $\{S_n\}$ be the random walk with step distribution $S_n S_{n-1}^{-1} \sim \mu$. Then the probability $\mathbf{P}[S_n \in \Lambda$ infinitely often$]$ takes only the values 0 and 1 as Λ ranges over subsets of Y, if and only if every bounded μ-harmonic function is constant. (Recall that $h : Y \to I\!R$ is μ-harmonic if $h(x) = \int_Y h(yx)\, d\mu(y)$ for all $x \in Y$.)*

Remark: When all bounded harmonic functions are constant, one says that the *Poisson boundary* of (Y, μ) is trivial; see Kaimanovich and Vershik (1982) for background.

PROOF: Given a set $\Lambda \subseteq Y$, the function $h(x) = \mathbf{P}[S_n x \in \Lambda$ infinitely often$]$ is bounded and μ-harmonic. The Markov property and the martingale convergence theorem imply that

$$h(S_m) = \mathbf{P}\left[\{S_k : k \geq 0\} \text{ visits } \Lambda \text{ i.o. } \mid S_1, S_2, \ldots, S_m\right] \to \mathbf{1}_{\{S_k \text{ visits } \Lambda \text{ i.o.}\}}$$

as $m \to \infty$. Thus if all bounded harmonic functions are constant, the zero-one law holds. Conversely, assume the zero-one law holds and let h be a bounded μ-harmonic function. For $\alpha \in \mathbb{R}$, let $\Lambda_\alpha = \{y \in Y : h(y) < \alpha\}$. If $\mathbf{P}[S_n x \text{ visits } \Lambda_\alpha \text{ i.o.}] = 0$ then we consider the stopping time $\tau = \min\{n : h(S_n x) \geq \alpha\}$ and obtain $h(x) = h(S_0 x) = \mathbf{E}h(S_\tau x) \geq \alpha$. Similarly, if $\mathbf{P}[S_n x \text{ visits } \Lambda_\alpha \text{ i.o.}] = 1$ then $h(x) \leq \alpha$. Since the support of μ generates Y,

$$\mathbf{P}[S_n \text{ visits } \Lambda_\alpha \text{ i.o.}] = 0 \Leftrightarrow \mathbf{P}[S_n x \text{ visits } \Lambda_\alpha \text{ i.o.}] = 0$$

and it follows that $h(x) = h(e)$ for all $x \in Y$. $\square$

4 Independent percolation on trees

Theorem 2.2 yields a short proof of a fundamental result of R. Lyons concerning percolation on trees. This theorem and its variants have been used in the analysis of a variety of probabilistic processes on trees, including random walks in a random environment, first-passage percolation and the Ising model. (See Lyons (1989, 1990, 1992); Lyons and Pemantle (1992); Benjamini and Peres (1994); Pemantle and Peres (1994).) Recently, this theorem has also been applied to study intersections of sample paths in Euclidean space (cf. Peres (1994)).

Notation: Let T be a finite, rooted tree. Vertices of degree one in T (apart from the root ρ) are called *leaves*, and the set of leaves is the *boundary* ∂T of T. The set of edges on the path connecting the root to a leaf x is denoted $\texttt{Path}(x)$.

Independent percolation on T is defined as follows. To each edge e of T, a parameter p_e in $[0, 1]$ is attached, and e is removed with probability $1 - p_e$, retained with probability p_e, with mutual independence among edges. Say that a leaf x *survives the percolation* if all of $\texttt{Path}(x)$ is retained, and say that the tree boundary ∂T survives if some leaf of T survives.

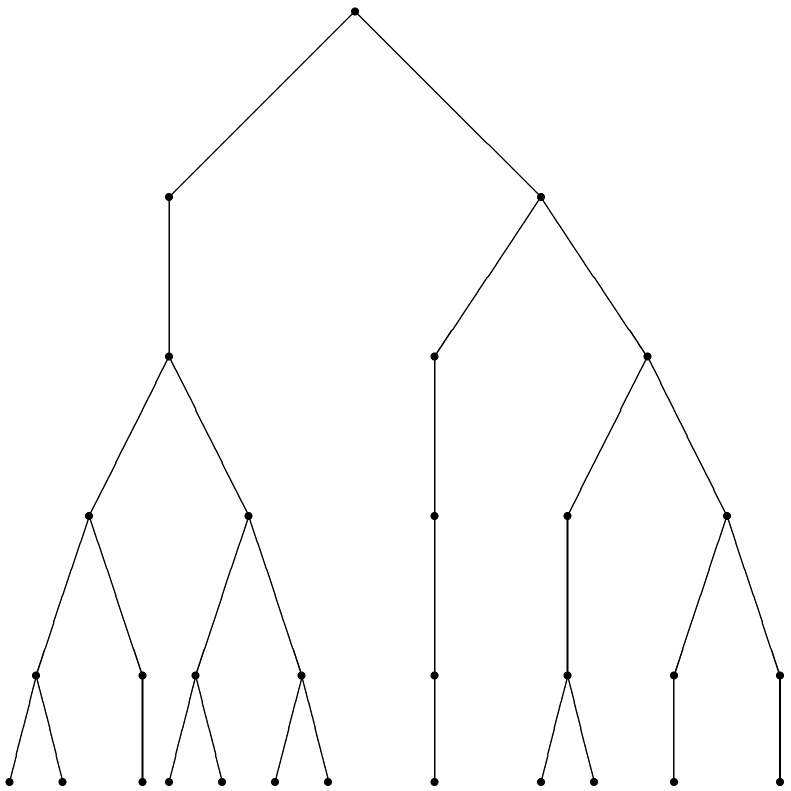

Figure 1: a tree

Theorem 4.1 (Lyons (1992)) *With the notation above, define a kernel F on ∂T by $F(x,y) = \prod\{p_e^{-1} \,:\, e \in \mathtt{Path}(x) \bigcap \mathtt{Path}(y)\}$ for $x \neq y$ and $F(x,x) = 2\prod\{p_e^{-1} \,:\, e \in \mathtt{Path}(x)\}$. Then*

$$\mathrm{Cap}_F(\partial T) \leq \mathbf{P}[\partial T \textit{ survives the percolation}] \leq 2\mathrm{Cap}_F(\partial T)$$

(The kernel F differs from the kernel used in Lyons (1992) on the diagonal, but this difference is unimportant in all applications).

Proof: Embed T in the lower half-plane, with the root at the origin. The random set of $r \geq 0$

leaves that survive the percolation may be enumerated from left to right as $V_1, V_2, \ldots, V_r$. The key observation is that The random sequence $\rho, V_1, V_2, \ldots, V_r, \Delta, \Delta, \ldots$ is a Markov chain on the state space $\partial T \bigcup \{\rho, \Delta\}$ (where ρ is the root and Δ is a formal absorbing cemetery).

Indeed, *given* that $V_k = x$, all the edges on $\mathtt{Path}(x)$ are retained, so that survival of leafs to the right of x is determined by the edges strictly to the right of $\mathtt{Path}(x)$, and is thus conditionally independent of $V_1, \ldots, V_{k-1}$. This verifies the Markov property, so Theorem 2.2 may be applied.

The transition probabilities for the Markov chain above are complicated, but it is easy to write down the Green kernel. Clearly, $G(\rho, y) = \mathbf{P}[y \text{ survives the percolation}] = \prod_{e \in \mathtt{Path}(y)} p_e$. Also, if x is to the left of y, then $G(x, y)$ is equal to the probability that the range of the Markov chain contains y given that it contains x, which is just the probability of y surviving given that x survives. Therefore

$$G(x, y) = \prod_{e \in \mathtt{Path}(y) \setminus \mathtt{Path}(x)} p_e$$

and hence

$$K(x, y) = \frac{G(x, y)}{G(\rho, y)} = \prod_{e \in \mathtt{Path}(x) \cap \mathtt{Path}(y)} p_e^{-1}.$$

Thus $K(x, y) + K(y, x) = F(x, y)$ for all $x, y \in \partial T$, and Lyons' Theorem follows from Theorem 2.2. $\square$

Remark: The same method of recognizing a "hidden" Markov chain may be used to prove more general results on random labeling of trees due to Evans (1992) and Lyons (1992).

5 Martin capacity and Brownian motion

Proof of Proposition 1.1: To bound from above the probability of ever hitting Λ, consider the stopping time $\tau = \min\{t > 0 : B_d(t) \in \Lambda\}$. The distribution of $B_d(\tau)$ on the event $\tau < \infty$ is a possibly defective distribution ν satisfying

$$\nu(\Lambda) = \mathbf{P}[\tau < \infty] = \mathbf{P}[\exists t > 0 : B_d(t) \in \Lambda]. \tag{17}$$

Now recall the standard formula, valid when $0 < \epsilon < ||y||$:

$$\mathbf{P}[\exists t > 0 : ||B_d(t) - y|| < \epsilon] = \frac{\epsilon^{d-2}}{||y||^{d-2}} . \tag{18}$$

By a first entrance decomposition, the probability in (18) is at least

$$\mathbf{P}[||B_d(\tau) - y|| > \epsilon \text{ and } \exists t > \tau : ||B_d(t) - y|| < \epsilon] \ = \ \int_{x:||x-y||>\epsilon} \frac{\epsilon^{d-2}}{||x-y||^{d-2}} \, d\nu(x) .$$

Dividing by ϵ^{d-2} throughout and letting $\epsilon \to 0$, we obtain

$$\int_\Lambda \frac{d\nu(x)}{||x-y||^{d-2}} \leq \frac{1}{||y||^{d-2}} ,$$

i.e. $\int_\Lambda K(x,y)\, d\nu(x) \leq 1$ for all $y \in \Lambda$. Therefore $I_K(\nu) \leq \nu(\Lambda)$ and thus

$$\mathrm{Cap}_K(\Lambda) \geq [I_K(\nu/\nu(\Lambda))]^{-1} \geq \nu(\Lambda) ,$$

which by (17) yields the upper bound on the probability of hitting Λ.

To obtain a lower bound for this probability, a second moment estimate is used. It is easily seen that the Martin capacity of Λ is the supremum of the capacities of its compact subsets, so we may assume that Λ itself is compact. For $\epsilon > 0$ and $y \in \mathbb{R}^d$ let $D(y,\epsilon)$ denote the Euclidean ball of radius ϵ about y and let $h_\epsilon(||y||)$ denote the probability that a standard Brownian path will hit this ball:

$$h_\epsilon(r) = \begin{cases} (\epsilon/r)^{d-2} & \text{if } r > \epsilon \\ 1 & \text{otherwise} \end{cases} . \tag{19}$$

Given a probability measure μ on Λ, and $\epsilon > 0$, consider the random variable

$$Z_\epsilon = \int_\Lambda \mathbf{1}_{\{\exists t>0: B_d(t) \in D(y,\epsilon)\}} h_\epsilon(||y||)^{-1} \, d\mu(y) .$$

Clearly $\mathbf{E} Z_\epsilon = 1$. We compute the second moment of Z_ϵ in order to apply Cauchy-Schwarz as in the proof of Theorem 2.2.

By symmetry,

$$\begin{aligned}
\mathbf{E}Z_\epsilon^2 &= 2\mathbf{E}\int_\Lambda\int_\Lambda \mathbf{1}_{\{\exists t>0: B_d(t)\in D(x,\epsilon) \text{ and } \exists s>t: B_d(s)\in D(y,\epsilon)\}} \frac{d\mu(x)d\mu(y)}{h_\epsilon(||x||)h_\epsilon(||y||)} \\
&\le 2\mathbf{E}\int_\Lambda\int_\Lambda \mathbf{1}_{\{\exists t>0: B_d(t)\in D(x,\epsilon)\}} \frac{h_\epsilon(||y-x||-\epsilon)}{h_\epsilon(||x||)h_\epsilon(||y||)}\, d\mu(x)\, d\mu(y) \\
&= 2\int_\Lambda\int_\Lambda \frac{h_\epsilon(||y-x||-\epsilon)}{h_\epsilon(||y||)}\, d\mu(x)\, d\mu(y)\,. \qquad (20)
\end{aligned}$$

The last integrand is bounded by 1 if $||y|| \le \epsilon$. On the other hand, if $||y|| > \epsilon$ and $||y-x|| \le 2\epsilon$ then $h_\epsilon(||y-x||-\epsilon) = 1 \le 2^{d-2}h_\epsilon(||y-x||)$, so that the integrand on the right-hand side of (20) is at most $2^{d-2}K(x,y)$. Thus

$$\begin{aligned}
&\mathbf{E}Z_\epsilon^2 \le 2\mu(D(0,\epsilon)) + 2^{d-1}\int\int \mathbf{1}_{\{||y-x||\le 2\epsilon\}} K(x,y)\, d\mu(x)\, d\mu(y) \qquad (21) \\
&+2\int\int \mathbf{1}_{\{||y-x||>2\epsilon\}} \left(\frac{||y||}{||y-x||-\epsilon}\right)^{d-2} d\mu(x)\, d\mu(y)\,.
\end{aligned}$$

Since the kernel is infinite on the diagonal, any measure with finite energy must have no atoms. Restricting attention to such measures μ, we see that the first two summands in (21) drop out as $\epsilon \to 0$ (by dominated convergence) . This leaves

$$\lim_{\epsilon\downarrow 0} \mathbf{E}Z_\epsilon^2 \le 2I_K(\mu)\,. \qquad (22)$$

Clearly the hitting probability $\mathbf{P}[\exists t > 0, y \in \Lambda : B_d(t) \in D(y,\epsilon)]$ is at least

$$\mathbf{P}[Z_\epsilon > 0] \ge \frac{(\mathbf{E}Z_\epsilon)^2}{\mathbf{E}Z_\epsilon^2} = (\mathbf{E}Z_\epsilon^2)^{-1}\,.$$

Transience of Brownian motion implies that if the Brownian path visits every ϵ-neighborhood of the compact set Λ then it almost surely intersects Λ itself. Therefore, by (22):

$$\mathbf{P}[\exists t > 0 : B_d(t) \in \Lambda] \ge \lim_{\epsilon\downarrow 0}(\mathbf{E}Z_\epsilon^2)^{-1} \ge \frac{1}{2I_K(\mu)}.$$

Since this is true for all probability measures μ on Λ, we get the desired conclusion:

$$\mathbf{P}[\exists t > 0 : B_d(t) \in \Lambda] \ge \frac{1}{2}\mathrm{Cap}_K(\Lambda). \qquad (23)$$

$\square$

Remark: The right–hand inequality in (1) is sometimes an equality– a sphere centered at the origin has hitting probability and Martin capacity both equal to 1. To see that the constant 1/2 in (23) cannot be increased, consider the spherical shell

$$\Lambda_R = \{x \in \mathbb{R}^d : 1 \leq ||x|| \leq R\}\,.$$

We claim that $\lim_{R\to\infty} \mathrm{Cap}_K(\Lambda_R) = 2$. Indeed by Proposition 1.1, the Martin capacity of *any* compact set is at most 2, while lower bounds tending to 2 for the capacity of Λ_R are established by computing the energy of the probability measure supported on Λ_R, with density a constant multiple of $||x||^{1-d}$ there.

Next, we pass from the local to the global behavior of Brownian paths. Barlow and Taylor (1992) noted that for $d \geq 2$ the set of nearest-neighbour lattice points to a Brownian path in $\mathbb{R}^d$ is a subset of $\mathbf{Z}^d$ with dimension 2, using their definition of dimension which is equivalent to (16). This is a property of the path near infinity; another such property is given by

Proposition 5.1 *Let $B_d(t)$ denote d-dimensional Brownian motion. Let $\Lambda \subseteq \mathbb{R}^d$ with $d \geq 3$ and let Λ_1 be the cubical fattening of Λ defined by*

$$\Lambda_1 = \{x \in \mathbb{R}^d : \exists y \in \Lambda \;\; s.t. \;\; ||y-x||_\infty \leq 1\}.$$

Then a necessary and sufficient condition for the almost sure existence of times $t_j \uparrow \infty$ at which $B_d(t_j) \in \Lambda_1$ is that $\mathrm{Cap}_{d-2}^{(\infty)}(\Lambda_1 \cap \mathbf{Z}^d) > 0$.

The proof is very similar to the proof of Theorem 2.2 and is omitted.

6 Concluding remarks

1. With the exception of Section 5, this paper is concerned with *discrete* Markov chains. Of course the proof of Proposition 1.1 given in that section extends without difficulty to some other Markov processes in continuous time, but a classification of the processes for which this extension is possible is beyond the scope of this paper. Nevertheless, we do mention explicitly the range of a stable subordinator of index $1/2$, since this range can be viewed as the zero set of a one-dimensional Brownian motion, and is therefore of wider interest.

Corollary 6.1 *Let $\{B(t)\}$ denote standard one-dimensional Brownian motion, and let A be any closed set in $(0,\infty)$. Then*

$$\frac{1}{2}\mathrm{Cap}_K(A) \le \mathbf{P}[\exists t \in A \,:\, B(t)=0] \le \mathrm{Cap}_K(A) \quad \textit{where}$$

$$K(s,t) = \begin{cases} (t/t-s)^{1/2} & \textit{if } s<t \\ \infty & \textit{if } s=t \\ 0 & \text{otherwise}\ . \end{cases}$$

SKETCH OF PROOF: Use the obvious estimate $\mathbf{P}(|B(t)|<\epsilon) \sim 2\epsilon/\sqrt{2\pi t}$ as $\epsilon \downarrow 0$, and mimic the proof of Proposition 1.1.

2. A probabilist might wonder what is gained by capacity estimates such as Proposition 1.1 and Theorem 2.2, since the quantity of interest, the hitting probability, is estimated by a quantity which appears more complicated. Indeed only in special situations can the capacity of a set be calculated exactly. Capacity estimates are useful because of their robustness (see corollaries 2.3 and 2.6, as well as the proof of the stability of the Nash-Williams recurrence criterion in Lyons (1992)) and the ease with which they yield lower bounds for hitting probabilities. Finally, in the continuous setting, such estimates allow one to exploit the information amassed on capacity by analysts studying singularities of solutions to PDE's.

3. The restriction to dimension $d \ge 3$ in Proposition 1.1 is natural since planar Brownian motion will hit any measurable set with probability 0 or 1. However, by killing the motion at a finite time one may obtain a planar version of the proposition.

4. The Martin kernel is most often encountered in constructions of the Martin boundary, where only its asymptotics matter. In Lyons, MacGibbon and Taylor (1984) the Kernel $G(x,y)/G(0,y)$

is used to compare the probability of Brownian motion hitting a set, to the probability of hitting its projection on a hyperplane. However, the denominator plays a different role there, as the Brownian motion is not started at 0, and is stopped when it leaves the upper half-space.

5. The methods of this paper do not seem to yield upper estimates for the probability that a set will be hit by the *intersection* of the ranges of two Markov chains. Such estimates were obtained, in a very general setting, in a remarkable paper by Fitzsimmons and Salisbury (1989). However, the estimates in that paper required that the initial distribution for each chain be an equilibrium measure, so that for fixed initial states only qualitative (though important) information was obtained. After we showed Tom Salisbury the statement of Proposition 1.1, he observed that the methods of his paper with P. Fitzsimmons may be used to estimate the hitting probability of a set by the intersection of two chains (with no restrictions imposed on the initial distributions), in terms of the product of the corresponding Martin kernels. See Salisbury (1994) for a very readable exposition.

References

[1] Barlow, M. and Taylor, S.J. (1992). Defining fractal subsets of $\mathbf{Z}^d$. *Proc. Lon. Math. Soc.* **64** 125 - 152.

[2] Benjamini, I. and Peres, Y. (1994). Tree-indexed random walks on groups and first-passage percolation. *Probab. Th. Rel. Fields* **98** 91 - 112.

[3] Erdös, P. (1961). A problem about prime numbers and the random walk II. *Ill. J. Math.* **5**, 352 - 353.

[4] Evans, S. (1992). Polar and nonpolar sets for a tree-indexed process. *Ann. Prob.* **20**, 579-590.

[5] Fitzsimmons, P.J. and Salisbury, T. (1989). Capacity and energy for multiparameter Markov processes. *Ann. Inst. Henri Poincarè, Probab.* **25** 325-350.

[6] Kahane, J. P. (1985). Some random series of functions. Second edition, Cambridge University Press: Cambridge.

[7] Kaimanovich, V. and Vershik, A. (1983). Random walks on discrete groups: boundary and entropy. *Ann. Prob.* **11**, 457 - 490.

[8] Kakutani, S. (1944). Two dimensional Brownian motion and harmonic functions. *Proc. Imp. Acad. Tokyo* **20**, 648-652.

[9] Khoshnevisan, D. (1994). A discrete fractal in $\mathbf{Z}_+^1$. *Proc. Amer. Math. Soc.* **120**, 577-584.

[10] Kochen, S. and Stone, C. (1964). A note on the Borel-Cantelli Lemma. *Illinois J. of Math.* **8** 248-251.

[11] Lamperti, J. (1963). Wiener's test and Markov chains. *J. Math. Anal. Appl.* **6** 58 - 66.

[12] Lawler, G. (1991). Intersections of random walks. Birkhäuser: New York.

[13] Lyons, R. (1989). The Ising model and percolation on trees and tree-like graphs. *Comm. Math. Phys.* **125** 337 - 353.

[14] Lyons, R. (1990). Random walks and percolation on trees. *Ann. Probab.* **18** 931-958.

[15] Lyons, R. (1992). Random walks, capacity, and percolation on trees. *Ann. Probab.* **20** 2043 - 2088.

[16] Lyons, R. and Pemantle, R. (1992). Random walk in a random environment and first-passage percolation on trees. *Ann. Probab.* **20**, 125 - 136.

[17] Lyons, T.J., MacGibbon, K.B. and Taylor, J.C. (1984). Projection theorems for hitting probabilities and a theorem of Littlewood. *J. Functional analysis* **59**, No. 3, 470-489.

[18] McKean, H. (1961). A problem about prime numbers and the random walk I. *Ill. J. Math.* **5**, 351.

[19] Pemantle, R. and Peres, Y. (1994). Domination between trees and application to an explosion problem. *Ann. Probab.* **22**, 180 - 194.

[20] Peres, Y. (1994). Intersection-equivalence of Brownian paths and certain branching processes. *Preprint.*

[21] Ruzsa, I. Z. and Székely, G. J. (1982). Intersections of traces of random walks with fixed sets. *Annals Probab.* **10**, 132-136.

[22] Salisbury, T. (1994). Energy, and intersections of Markov chains. To appear in *Proceedings of the IMA workshop on Random Discrete Structures*, D. Aldous and R. Pemantle, editors.

[23] Spitzer, F. (1964). Principles of random walk. Van Nostrand: New York.

[24] Williamson, J. A. (1968). Random walks and Riesz kernels. *Pacific J. Math.* **25**, 393-415.

Itai Benjamini, Mathematical Sciences Institute, 409 College Ave., Ithaca, NY, 14853

Robin Pemantle, Department of Mathematics, University of Wisconsin-Madison, Van Vleck Hall, 480 Lincoln Drive, Madison, WI 53706 .

Yuval Peres, Department of Statistics, 367 Evans Hall University of California, Berkeley, CA 94720