

Markov Chain Intersections and the Loop-Erased Walk

BY RUSSELL LYONS, YUVAL PERES AND ODED SCHRAMM

Abstract. Let X and Y be independent transient Markov chains on the same state space that have the same transition probabilities. Let L denote the “loop-erased path” obtained from the path of X by erasing cycles when they are created. We prove that if the paths of X and Y have infinitely many intersections a.s., then L and Y also have infinitely many intersections a.s.

Résumé. Soit X et Y deux chaînes de Markov indépendantes et transientes à même espace d'états et aux mêmes probabilités de transition. Soit L le “chemin aux boucles effacées” obtenu du chemin de X en effaçant les cycles aux moments de création. Nous démontrons que si les chemins de X et Y ont une infinité d'intersections p.s., alors L et Y ont aussi une infinité d'intersections p.s.

§1. Introduction.

Erdős and Taylor [3] proved that two independent simple random walk paths in \mathbb{Z}^d intersect infinitely often if $d \leq 4$, but not if $d > 4$. Lawler [9] proved that for $d = 3, 4$, there are still infinitely many intersections even if one of the paths is replaced by its loop-erasure, which is obtained by erasing cycles in the order in which they are created. Lawler's proof relied on estimates that are only available in Euclidean space, so it remained unclear how general this phenomenon is. Our main result, Theorem 1.1 below, extends Lawler's result to general transient Markov chains. Our primary motivation for studying intersections of Markov chain paths and loop-erased paths is the connection with uniform spanning forests, which will be recalled in the next section. The precise definition of loop-erasure is as follows. We denote sequences by angle brackets $\langle \dots \rangle$ and sets by braces $\{\dots\}$. Cardinality is denoted by bars $|\dots|$.

DEFINITION. Let \mathcal{S} be a countable set, and consider a (finite or infinite) sequence $\gamma = \langle v_0, v_1, v_2, \dots \rangle$ in \mathcal{S} , such that each $v \in \mathcal{S}$ occurs only finitely often in γ . The **loop-erasure**

2000 Mathematics Subject Classification. Primary 60J10. Secondary 60G17.

Key words and phrases. Random walk, uniform spanning forests.

Research partially supported by the Institute for Advanced Studies, Jerusalem (Lyons), NSF grant DMS-0104073 (Peres), and the Sam and Ayala Zacks Professorial Chair (Schramm).

of γ , denoted $\text{LE}(\gamma) = \langle u_0, u_1, u_2, \dots \rangle$, is constructed inductively as follows. Set $u_0 := v_0$. If u_j has been chosen, then let k be the last index such that $v_k = u_j$. If v_k is the last element of γ , then let $\text{LE}(\gamma) := (u_1, \dots, u_j)$; otherwise, set $u_{j+1} := v_{k+1}$ and continue.

We shall let $\text{LE}(\gamma)$ stand for the sequence as well as the set. In this notation, our main result can be stated as follows:

THEOREM 1.1. *Let $\langle X_m \rangle$ and $\langle Y_n \rangle$ be independent transient Markov chains on the same state space \mathcal{S} that have the same transition probabilities, but possibly different initial states. Then on the event that $|\{X_m\} \cap \{Y_n\}| = \infty$, almost surely $|\text{LE}\langle X_m \rangle \cap \{Y_n\}| = \infty$.*

The key to proving this theorem is the following quantitative lemma.

LEMMA 1.2. *Let $\langle X_m \rangle$ and $\langle Y_n \rangle$ be independent transient Markov chains on the same state space \mathcal{S} that have the same transition probabilities, but possibly different initial states. Then*

$$\mathbf{P}[\text{LE}\langle X_m \rangle \cap \{Y_n\} \neq \emptyset] \geq 2^{-8} \mathbf{P}[\{X_m\} \cap \{Y_n\} \neq \emptyset].$$

REMARK 1.3. Lemma 1.2 also applies to Markov chains that are killed when they exit a certain set, or killed at an exponential time; in this form, it contains new information even when the underlying chain is simple random walk in \mathbb{Z}^d for $d \geq 3$.

Theorem 1.1 reduces the question of the intersection of a Markov chain and the loop-erasure of an independent copy to the simpler problem of the intersection of two independent copies of the chain. However, for a general Markov chain, it may still be difficult to find the probability that two independent sample paths of the chain have infinitely many intersections. Consider the Green function

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

Because of the Borel-Cantelli lemma, if the independent sample paths X and Y satisfy

$$\begin{aligned} \sum_{m,n=0}^{\infty} \mathbf{P}_{x,x}[X_m = Y_n] &= \sum_{z \in \mathcal{S}} \sum_{m,n=0}^{\infty} \mathbf{P}_x[X_m = z] \mathbf{P}_x[Y_n = z] \\ &= \sum_{z \in \mathcal{S}} G(x, z)^2 < \infty, \end{aligned} \tag{1.1}$$

then the number of intersections of X and Y is a.s. finite. In general, the converse need not hold; see Example 6.1. Nevertheless, if the transition probabilities are invariant under a transitive group of permutations of the state space, then the converse to (1.1) is valid.

THEOREM 1.4. *Let $p(\cdot, \cdot)$ be a transition kernel on a countable state space \mathcal{S} . Suppose that Π is a group of permutations of \mathcal{S} that acts transitively (i.e., with a single orbit) and satisfies $p(\pi x, \pi y) = p(x, y)$ for all $\pi \in \Pi$ and $x, y \in \mathcal{S}$. Suppose that*

$$\sum_{z \in \mathcal{S}} G(o, z)^2 = \infty, \quad (1.2)$$

where o is a fixed element of \mathcal{S} . Then two independent chains X and Y with transition probabilities $p(\cdot, \cdot)$ and initial state $o \in \mathcal{S}$ have infinitely many intersections a.s. Moreover, Y_n is in $\text{LE}(\langle X_m \rangle_{m \geq 0})$ for infinitely many n a.s.

For simple random walk on a vertex-transitive graph, (1.2) holds if and only if the graph has polynomial volume growth of degree at most 4; see Corollary 5.3.

Next, we consider triple intersections. It is well known that three independent simple random walks in \mathbb{Z}^3 have infinitely many mutual intersections a.s. (see Lawler [9], Section 4.5). To illustrate the versatility of Theorem 1.1, we offer the following refinement.

COROLLARY 1.5. *Let $X = \langle X_m \rangle$, $Y = \langle Y_n \rangle$ and $Z = \langle Z_k \rangle$ be independent simple random walks in the lattice \mathbb{Z}^3 . Denote by $L_Z(X)$ the “partially loop-erased” path, obtained from X by erasing any cycle that starts (and ends) at a node in Z , where the erasure is made when the cycle is created. Then the (setwise) triple intersection $L_Z(X) \cap Y \cap Z$ is a.s. infinite.*

See Corollary 6.2 in Section 6 for an extension and the (very short) proof. We note that $L_Z(X)$ cannot be replaced by $\text{LE}(X)$ in this corollary; this follows from Lawler [9], Section 7.5.

In the next section, we recall the connection to spanning forests and sketch a heuristic argument for Theorem 1.1. In Section 3, we discuss the reverse second moment method for two Markov chains, following Salisbury [13]. Lemma 1.2 and Theorem 1.1 are proved in Section 4 by combining ideas from the two preceding sections. Section 5 contains a proof of Theorem 1.4 on transitive chains. Concluding remarks and questions are in Section 6.

§2. Spanning Forests and Heuristics.

Loop-erased random walks and uniform spanning trees are intimately related. Let G be a finite graph with x, y two vertices of G . Let L be the loop-erasure of the random walk path started at x and stopped when it reaches y . On the other hand, let T be a spanning tree of G chosen uniformly and let L_T be the shortest path in T that connects x and y . Pemantle [12] showed that L_T has the same distribution as L . Given that $L_T = \ell$ for some simple path ℓ , the remainder of T has the uniform distribution among spanning trees of the graph obtained from G by contracting ℓ . Therefore, it follows immediately from $L_T \stackrel{D}{=} L$ that a uniform spanning tree can also be chosen as follows. Pick any vertex x_0 of G . Let L_0 be loop-erased random walk from any vertex x_1 to x_0 . Pick any vertex x_2 and let L_1 be loop-erased random walk from x_2 to L_0 . Pick any vertex x_3 and let L_1 be loop-erased random walk from x_3 to $L_0 \cup L_1$. Continue until a spanning tree $T := L_0 \cup L_1 \cup \dots$ is created. Then T has the uniform distribution.

This is known as Wilson's algorithm for generating a uniform spanning tree. Wilson [14] showed that an analogous algorithm exists corresponding to any Markov chain, not merely to reversible Markov chains.

Next, we discuss the analogous object on infinite graphs. The **wired uniform spanning forest** (WUSF) in an infinite graph G may be defined as a weak limit of uniform random spanning trees in an exhaustion of G by finite subgraphs G_n , with the boundary of G_n identified to a single point ("wired"). The resulting measure on spanning forests does not depend on the exhaustion. The WUSF was implicit in Pemantle [12] and was made explicit by Häggström [5]; see Benjamini, Lyons, Peres, and Schramm [1] for details. The connection of the WUSF to loop-erased walks was discovered by Pemantle [12]:

PROPOSITION 2.1. *Let G be a locally-finite connected graph. The wired uniform spanning forest (WUSF) is a single tree a.s. iff from every (or some) vertex, simple random walk and an independent loop-erased random walk intersect infinitely often a.s. Moreover, the probability that u and v belong to the same WUSF component equals the probability that a simple random walk path from u intersects an independent loop-erased walk from v .*

Just as the relation between spanning trees and loop-erased walks in finite graphs was clarified by the algorithm of Wilson [14] for generating uniform spanning trees, this algorithm was extended to infinite graphs in [1] to generate the WUSF. With this extended algorithm, Proposition 2.1 becomes obvious. This proposition illustrates why Theorem 1.1 is useful in the study of the WUSF.

We now sketch a heuristic argument for Theorem 1.1. On the event that $X_m = Y_n$, the continuation paths $X' := \langle X_j \rangle_{j \geq m}$ and $Y' := \langle Y_k \rangle_{k \geq n}$ have the same distribution, whence

the chance is at least $1/2$ that Y' intersects $L := \text{LE}\langle X_0, \dots, X_m \rangle$ at an earlier (in the clock of L) point than X' . On this event, the earliest intersection point of Y' and L will remain in $\text{LE}\langle X_j \rangle_{j \geq 0} \cap \langle Y_k \rangle_{k \geq 0}$. The difficulty in making this heuristic precise lies in selecting a pair (m, n) such that $X_m = Y_n$, given that such pairs exist. The natural rules for selecting such a pair (*e.g.*, lexicographic ordering) affect the law of at least one of the continuation paths, and invalidate the argument above; R. Pemantle (private communication, 1996) showed that this holds for *all* selection rules. Our solution to this difficulty is based on applying a second moment argument to a weighted count of intersections.

§3. The Second Moment Method and a Converse.

THEOREM 3.1. *Let X and Y be two independent Markov chains on the same countable state space \mathcal{S} , with initial states x_0 and y_0 , respectively. Let*

$$A \subset \mathbb{N} \times \mathcal{S} \times \mathbb{N} \times \mathcal{S},$$

and denote by $\text{hit}(A)$ the event that $(m, X_m, n, Y_n) \in A$ for some $m, n \in \mathbb{N}$. Given any weight function $w : \mathcal{S} \rightarrow [0, \infty)$ that vanishes outside of A , consider the random variable

$$S_w := \sum_{m,n=0}^{\infty} w(m, X_m, n, Y_n).$$

If $\mathbf{P}[\text{hit}(A)] > 0$, then there exists such a w satisfying $0 < \mathbf{E}[S_w^2] < \infty$ and

$$\mathbf{P}[\text{hit}(A)] \leq 64 \frac{(\mathbf{E}S_w)^2}{\mathbf{E}[S_w^2]}. \quad (3.1)$$

Note that this provides a converse estimate to that provided by the Cauchy-Schwarz inequality (often referred to as “the second-moment method”): If $0 < \mathbf{E}[S_w^2] < \infty$, then

$$\mathbf{P}[\text{hit}(A)] \geq \mathbf{P}[S_w > 0] \geq \frac{(\mathbf{E}S_w)^2}{\mathbf{E}[S_w^2]}. \quad (3.2)$$

Theorem 3.1 is essentially contained in Theorem 2 of Salisbury [13], which is, in turn, based on the ideas in the path-breaking paper of Fitzsimmons and Salisbury [4]. We include the proof of this theorem, since our focus on time-space chains allows us to avoid the subtle time-reversal argument in Salisbury’s paper. The ratio 64 between the upper and lower bounds in (3.1) and (3.2), respectively, improves on the ratio 1024 obtained in Salisbury [13], but we suspect that it is still not optimal. We remark that the lower bound for the hitting probability which is stated but not proved in Corollary 1 of Salisbury [13] (*i.e.*, the left-hand inequality in the last line of the statement) is incorrect as stated, but we shall not need that inequality.

We start with two known lemmas.

LEMMA 3.2. Let $(\Omega, \mathcal{B}, \mathbf{P})$ be a probability space. Suppose that $\mathcal{F}, \mathcal{G}, \mathcal{H}$ are sub- σ -fields of \mathcal{B} such that $\mathcal{G} \subset \mathcal{F} \cap \mathcal{H}$ and \mathcal{F}, \mathcal{H} are conditionally independent given \mathcal{G} . Then

(i) for any $f \in L^1(\mathcal{F}) := L^1(\Omega, \mathcal{F}, \mathbf{P})$,

$$\mathbf{E}[f \mid \mathcal{H}] = \mathbf{E}[f \mid \mathcal{G}]; \quad (3.3)$$

(ii) for any $\varphi \in L^1(\mathcal{B}) := L^1(\Omega, \mathcal{B}, \mathbf{P})$,

$$\mathbf{E}[\mathbf{E}[\varphi \mid \mathcal{F}] \mid \mathcal{H}] = \mathbf{E}[\varphi \mid \mathcal{G}]. \quad (3.4)$$

Proof.

(i) For any $h \in L^\infty(\mathcal{H})$ and $g \in L^1(\mathcal{G})$, we have $\int hg d\mathbf{P} = \int \mathbf{E}[h \mid \mathcal{G}]g d\mathbf{P}$ by definition of conditional expectation. In particular,

$$\int h\mathbf{E}[f \mid \mathcal{G}] d\mathbf{P} = \int \mathbf{E}[h \mid \mathcal{G}]\mathbf{E}[f \mid \mathcal{G}] d\mathbf{P} = \int \mathbf{E}[hf \mid \mathcal{G}] d\mathbf{P} = \int hf d\mathbf{P}, \quad (3.5)$$

where the second equality follows from the conditional independence assumption. Since $\mathbf{E}[f \mid \mathcal{G}]$ is \mathcal{H} -measurable, the identity (3.5) for all $h \in L^\infty(\mathcal{H})$ implies that (3.3) holds.

(ii) Write $f := \mathbf{E}[\varphi \mid \mathcal{F}]$. Since $\mathbf{E}[f \mid \mathcal{G}] = \mathbf{E}[\varphi \mid \mathcal{G}]$, (3.4) follows from (3.3). \blacksquare

We shall use the following inequality from Burkholder, Davis and Gundy [2].

LEMMA 3.3. Suppose that $\langle \mathcal{F}_m \rangle$ is an increasing or decreasing sequence of σ -fields and that $\langle \varphi(m) \rangle$ is a sequence of nonnegative random variables. Then

$$\mathbf{E}\left[\left(\sum_m \mathbf{E}[\varphi(m) \mid \mathcal{F}_m]\right)^2\right] \leq 4\mathbf{E}\left[\left(\sum_m \varphi(m)\right)^2\right].$$

Proof of Theorem 3.1. Consider the σ -fields

$$\begin{aligned} \mathcal{F}_m &:= \sigma(X_1, \dots, X_m), & \widehat{\mathcal{F}}_m &:= \sigma(X_m, X_{m+1}, \dots), \\ \mathcal{G}_n &:= \sigma(Y_1, \dots, Y_n), & \widehat{\mathcal{G}}_n &:= \sigma(Y_n, Y_{n+1}, \dots). \end{aligned} \quad (3.6)$$

Abbreviate $\mathcal{F} := \widehat{\mathcal{F}}_0$ and $\mathcal{G} = \widehat{\mathcal{G}}_0$.

We begin with the lexicographically minimal stopping time (τ, λ) defined as follows:

if there exist $m, n \geq 0$ such that $(m, X_m, n, Y_n) \in A$, then let

$$\begin{aligned} \tau &:= \min\{m ; \exists n \ (m, X_m, n, Y_n) \in A\}, \\ \lambda &:= \min\{n ; (\tau, X_\tau, n, Y_n) \in A\}; \end{aligned}$$

otherwise, set $\tau := \lambda := \infty$.

Consider

$$\psi(m, n) := \mathbf{1}_{\{\tau=m, \lambda=n\}}.$$

Since $\psi(m, n)$ is $\mathcal{F}_m \vee \mathcal{G}$ -measurable and $\mathcal{F}_m \vee \mathcal{G}$ is conditionally independent of $\widehat{\mathcal{F}}_m \vee \mathcal{G}$ given $\sigma(X_m) \vee \mathcal{G}$, Lemma 3.2(i) implies that

$$\psi_1(m, n) := \mathbf{E}[\psi(m, n) \mid \widehat{\mathcal{F}}_m \vee \mathcal{G}] = \mathbf{E}[\psi(m, n) \mid \sigma(X_m) \vee \mathcal{G}].$$

Let $\psi_2(m, n) := \mathbf{E}[\psi_1(m, n) \mid \mathcal{F} \vee \mathcal{G}_n]$. Two applications of Lemma 3.2(ii) yield that

$$\psi_3(m, n) := \mathbf{E}[\psi_2(m, n) \mid \mathcal{F} \vee \widehat{\mathcal{G}}_n] = \mathbf{E}[\psi_1(m, n) \mid \mathcal{F} \vee \sigma(Y_n)] = \mathbf{E}[\psi(m, n) \mid \sigma(X_m, Y_n)].$$

Thus

$$\psi_3(m, n) = \mathbf{P}[\tau = n, \lambda = m \mid X_m, Y_n] = \sum_{x, y \in \mathcal{S}} w(m, x, n, y) \mathbf{1}_{\{X_m=x, Y_n=y\}}, \quad (3.7)$$

where

$$w(m, x, n, y) := \mathbf{P}[\tau = m, \lambda = n \mid X_m = x, Y_n = y].$$

Applying Lemma 3.3 to the random variables $\varphi(m) := \sum_{n \geq 0} \psi(m, n)$ and the σ -fields $\widehat{\mathcal{F}}_m \vee \mathcal{G}$, we obtain

$$\mathbf{E}\left[\left(\sum_{m, n} \psi_1(m, n)\right)^2\right] \leq 4 \mathbf{E}\left[\left(\sum_{m, n} \psi(m, n)\right)^2\right] = 4 \mathbf{P}[\text{hit}(A)],$$

since $\sum_{m, n} \psi(m, n) = \mathbf{1}_{\text{hit}(A)}$. By (3.7), we have

$$S_w := \sum_{m, n} w(m, X_m, n, Y_n) = \sum_{m, n} \psi_3(m, n).$$

Two applications of Lemma 3.3, first with the variables $\varphi_2(n) := \sum_m \psi_2(m, n)$ and the σ -fields $\mathcal{F} \vee \widehat{\mathcal{G}}_n$, then with the variables $\varphi_1(n) := \sum_m \psi_1(m, n)$ and the σ -fields $\mathcal{F} \vee \mathcal{G}_n$, yield

$$\mathbf{E}[S_w^2] \leq 4 \mathbf{E}\left[\left(\sum_{m, n} \psi_2(m, n)\right)^2\right] \leq 16 \mathbf{E}\left[\left(\sum_{m, n} \psi_1(m, n)\right)^2\right] \leq 64 \mathbf{P}[\text{hit}(A)].$$

Since $\mathbf{E}[S_w] = \mathbf{P}[\text{hit}(A)]$, the previous inequality is equivalent to (3.1). ■

§4. Intersecting the Loop-Erasure.

We shall prove the following extension of Lemma 1.2.

LEMMA 4.1. *Let $\langle X_m \rangle_{m \geq 0}$ and $\langle Y_n \rangle_{n \geq 0}$ be independent transient Markov chains on \mathcal{S} that have the same transition probabilities, but possibly different initial states x_0 and y_0 . Given $k \geq 0$, fix $\langle x_j \rangle_{j=-k}^{-1}$ in \mathcal{S} and set $X_j := x_j$ for $-k \leq j \leq -1$. Then the probability that the loop-erasure of $\langle X_m \rangle_{m \geq -k}$ intersects $\{Y_n\}_{n \geq 0}$ is at least $2^{-8} \mathbf{P}[\exists m \geq 0 \ \exists n \geq 0 \ X_m = Y_n]$.*

Proof. For $A := \{(m, x, n, x) ; m, n \geq 0, x \in \mathcal{S}\}$, choose a weight function $w : A \rightarrow [0, \infty)$ as in Theorem 3.1, which defines the sum S_w . Denote

$$L^m := \langle L_j^m \rangle_{j=0}^{J(m)} := \mathsf{LE} \langle X_{-k}, X_{1-k}, \dots, X_m \rangle.$$

On the event $\{X_m = Y_n\}$, define

$$j(m, n) := \min \{j \geq 0 ; L_j^m \in \{X_m, X_{m+1}, X_{m+2}, \dots\}\}, \quad (4.1)$$

$$i(m, n) := \min \{i \geq 0 ; L_i^m \in \{Y_n, Y_{n+1}, Y_{n+2}, \dots\}\}. \quad (4.2)$$

Note that the sets on the right-hand sides of (4.1) and (4.2) both contain $J(m)$ if $X_m = Y_n$. Define $j(m, n) := i(m, n) := 0$ on the event $\{X_m \neq Y_n\}$. Let $\chi(m, n) := 1$ if $i(m, n) \leq j(m, n)$, and $\chi(m, n) := 0$ otherwise. Given $\{X_m = Y_n = x\}$, the continuations $\langle X_m, X_{m+1}, X_{m+2}, \dots \rangle$ and $\langle Y_n, Y_{n+1}, Y_{n+2}, \dots \rangle$ are exchangeable with each other, so for every $x \in \mathcal{S}$,

$$\mathbf{E}[\chi(m, n) | X_m = Y_n = x] = \mathbf{P}[i(m, n) \leq j(m, n) | X_m = Y_n = x] \geq \frac{1}{2}. \quad (4.3)$$

Observe that if $X_m = Y_n$ and $i(m, n) \leq j(m, n)$, then $L_{i(m, n)}^m$ is in $\mathsf{LE} \langle X_r \rangle_{r=-k}^\infty \cap \{Y_\ell\}_{\ell=0}^\infty$.

Consider the random variable

$$\Upsilon_w := \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} w(m, X_m, n, Y_n) \chi(m, n).$$

Obviously $\Upsilon_w \leq S_w$ everywhere. On the other hand, by conditioning on X_m, Y_n and applying (4.3), we see that

$$\mathbf{E}[\Upsilon_w] = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \mathbf{E} \left[w(m, X_m, n, Y_n) \mathbf{E}[\chi(m, n) | X_m, Y_n] \right] \geq \frac{1}{2} \mathbf{E}[S_w].$$

By our choice of w and Theorem 3.1, we conclude that

$$P[\Upsilon_w > 0] \geq \frac{(\mathbf{E} \Upsilon_w)^2}{\mathbf{E}[\Upsilon_w^2]} \geq \frac{(\mathbf{E} S_w)^2}{4 \mathbf{E}[S_w^2]} \geq \frac{1}{256} \mathbf{P}[\text{hit}(A)].$$

The observation following (4.3) and the definition of A conclude the proof. ■

The next corollary follows immediately from Lemma 4.1 and the Markov property of $\langle X_m \rangle$ at a fixed time k .

COROLLARY 4.2. *Let $\langle X_m \rangle_{m \geq 0}$ and $\langle Y_n \rangle_{n \geq 0}$ be independent transient Markov chains on \mathcal{S} that have the same transition probabilities, but possibly different initial states x_0 and y_0 . Suppose that the event $B := \{X_1 = x_1, \dots, X_k = x_k\}$ has $\mathbf{P}[B] > 0$. Then*

$$\mathbf{P}_{x_0, y_0} \left[\text{LE}(\langle X_m \rangle_{m \geq 0} \cap \{Y_n\}) \neq \emptyset \mid B \right] \geq \frac{1}{256} \mathbf{P}_{x_0, y_0} \left[\exists m \geq k \ \exists n \geq 0 \ X_m = Y_n \mid B \right]. \quad (4.4)$$

Proof of Theorem 1.1. Denote by Λ the event that Y_n is in $\{X_m\}_{m \geq 0}$ for infinitely many n . Let Γ_ℓ denote the event that $Y_n \notin \text{LE}(\langle X_m \rangle_{m \geq 0})$ for all $n \geq \ell$, and define $\Gamma := \cup_{\ell \geq 0} \Gamma_\ell$. We must show that $\mathbf{P}_{x_0, y_0}[\Lambda \cap \Gamma] = 0$ for any choice of initial states x_0, y_0 . Suppose, on the contrary, that

$$\exists x_0, y_0 \quad \mathbf{P}_{x_0, y_0}[\Lambda \cap \Gamma] > 0. \quad (4.5)$$

Then for ℓ sufficiently large, $\mathbf{P}_{x_0, y_0}[\Lambda \cap \Gamma_\ell] > 0$. By Lévy's zero-one law, for any $\epsilon > 0$, if k, r are large enough, then there exist $x_1, \dots, x_k, y_1, \dots, y_r$ such that the events

$$B := \{X_1 = x_1, \dots, X_k = x_k\}, \quad B' := \{Y_1 = y_1, \dots, Y_r = y_r\}$$

satisfy $\mathbf{P}_{x_0, y_0}[B \cap B'] > 0$ and $\mathbf{P}_{x_0, y_0}[\Lambda \cap \Gamma_\ell \mid B \cap B'] > 1 - \epsilon$. We fix such events B, B' with $r > \ell$. Starting the chain Y at y_r instead of y_0 and using the Markov property, we infer that

$$\mathbf{P}_{x_0, y_r}[\Lambda \cap \Gamma_0 \mid B] > 1 - \epsilon. \quad (4.6)$$

However, Corollary 4.2 implies that

$$\mathbf{P}_{x_0, y_r}[\Gamma_0^c \mid B] \geq \frac{1}{256} \mathbf{P}_{x_0, y_r}[\Lambda \mid B] > \frac{1 - \epsilon}{256}.$$

Adding the preceding two inequalities, we obtain

$$1 \geq \mathbf{P}_{x_0, y_r}[\Lambda \cap \Gamma_0 \mid B] + \mathbf{P}_{x_0, y_r}[\Gamma_0^c \mid B] > \frac{257(1 - \epsilon)}{256}.$$

Taking $\epsilon < 1/257$ yields a contradiction to the assumption (4.5) and completes the proof. ■

§5. Transitive Markov Chains.

One ingredient of the proof of Theorem 1.4 will be the following lemma, which does not require transitivity.

LEMMA 5.1. *Let X and Y be independent transient Markov chains on the same state space that have the same transition probabilities. Denote by Λ the event that the paths of X and Y intersect infinitely often, and let $u(x, y) := \mathbf{P}_{x,y}[\Lambda]$, where the subscripts indicate the initial states of X and Y respectively. Then $u(x, x) \geq 2u(x, y) - 1$ for all x, y .*

Proof. Since $u(x, \cdot)$ is harmonic, the sequence $\langle u(x, Y_n) \rangle_{n \geq 0}$ is a bounded martingale. Therefore

$$\begin{aligned} u(x, y) - u(x, x) &= \lim_{n \rightarrow \infty} \mathbf{E}_y[u(x, Y_n)] - \lim_{m \rightarrow \infty} \mathbf{E}_x[u(x, X_m)] \\ &= \mathbf{E}_{x,y} \left[\lim_{n \rightarrow \infty} u(x, Y_n) - \lim_{m \rightarrow \infty} u(x, X_m) \right]. \end{aligned} \quad (5.1)$$

On the event Λ , the two limits in (5.1) coincide; therefore,

$$u(x, y) - u(x, x) \leq \mathbf{P}_{x,y}[\Lambda^c] = 1 - u(x, y).$$

This is equivalent to the assertion of the lemma. ■

Proof of Theorem 1.4. Here both Markov chains X, Y are started at o , so we write \mathbf{P} rather than \mathbf{P}_o , etc. Denote $G_n(o, x) := \sum_{k=0}^n \mathbf{P}[X_k = x]$. By transitivity,

$$\sum_{w \in \mathcal{S}} G_n(z, w)^2 = \sum_{w \in \mathcal{S}} G_n(o, w)^2 \quad (5.2)$$

for all $z \in \mathcal{S}$. Let $I_n := \sum_{k=0}^n \sum_{m=0}^n \mathbf{1}_{\{X_k = Y_m\}}$ be the number of intersections of X and Y by time n . Then

$$\begin{aligned} \mathbf{E}[I_n] &= \sum_{z \in \mathcal{S}} \sum_{k=0}^n \sum_{m=0}^n \mathbf{P}[X_k = z = Y_m] \\ &= \sum_{z \in \mathcal{S}} \sum_{k=0}^n \mathbf{P}[X_k = z] \cdot \sum_{m=0}^n \mathbf{P}[Y_m = z] \\ &= \sum_{z \in \mathcal{S}} G_n(o, z)^2. \end{aligned} \quad (5.3)$$

To estimate the second moment of I_n , observe that

$$\begin{aligned} \sum_{k,i=0}^n \mathbf{P}[X_k = z, X_i = w] &= \sum_{k=0}^n \sum_{i=k}^n \mathbf{P}[X_k = z] \mathbf{P}[X_i = w \mid X_k = z] \\ &\quad + \sum_{i=0}^n \sum_{k=i+1}^n \mathbf{P}[X_i = w] \mathbf{P}[X_k = z \mid X_i = w] \\ &\leq G_n(o, z) G_n(z, w) + G_n(o, w) G_n(w, z). \end{aligned}$$

Therefore

$$\begin{aligned}
\mathbf{E}[I_n^2] &= \sum_{z,w \in \mathcal{S}} \sum_{k,m=0}^n \sum_{i,j=0}^n \mathbf{P}[X_k = z = Y_m, X_i = w = Y_j] \\
&= \sum_{z,w \in \mathcal{S}} \sum_{k,i=0}^n \mathbf{P}[X_k = z, X_i = w] \cdot \sum_{m,j=0}^n \mathbf{P}[Y_m = z, Y_j = w] \\
&\leq \sum_{z,w \in \mathcal{S}} [G_n(o,z) G_n(z,w) + G_n(o,w) G_n(w,z)]^2 \\
&\leq \sum_{z,w \in \mathcal{S}} 2[G_n(o,z)^2 G_n(z,w)^2 + G_n(o,w)^2 G_n(w,z)^2] \\
&= 4 \sum_{z,w \in \mathcal{S}} G_n(o,z)^2 G_n(z,w)^2. \tag{5.4}
\end{aligned}$$

Summing first over w and using (5.2), then (5.3), we deduce that

$$\mathbf{E}[I_n^2] \leq 4 \left(\sum_{z \in \mathcal{S}} G_n(o,z)^2 \right)^2 = 4 \mathbf{E}[I_n]^2. \tag{5.5}$$

By a consequence of the Cauchy-Schwarz inequality (see, e.g., Kahane [8], p. 8),

$$\mathbf{P}[I_n \geq \epsilon \mathbf{E}[I_n]] \geq (1 - \epsilon)^2 \frac{\mathbf{E}[I_n]^2}{\mathbf{E}[I_n^2]} \geq \frac{(1 - \epsilon)^2}{4}. \tag{5.6}$$

As in Lemma 5.1, denote by Λ the event that the path-sets X and Y have infinitely many intersections, and let $u(x,y) := \mathbf{P}_{x,y}[\Lambda]$. Define $\mathcal{F}_n, \mathcal{G}_m$ as in (3.6). Apply the transience of X and Y and the Markov property to obtain

$$\mathbf{P}[\Lambda \mid \mathcal{F}_n \vee \mathcal{G}_n] = \mathbf{P}[\Lambda \mid X_n, Y_n] = u(X_n, Y_n).$$

Therefore, by Lévy's zero-one law, $\lim_{n \rightarrow \infty} u(X_n, Y_n) = \mathbf{1}_\Lambda$ a.s.

By (5.6) and the hypothesis (1.2), $\mathbf{P}[\Lambda] = \mathbf{P}[\lim I_n = \infty] \geq 1/4$. On the event Λ , we have by Lemma 5.1 that

$$\lim_{n \rightarrow \infty} u(X_n, X_n) \geq 2 \lim_{n \rightarrow \infty} u(X_n, Y_n) - 1 = 1,$$

whence $u(o,o) = 1$ by transitivity. The assertion concerning loop-erased walk now follows from Theorem 1.1. ■

REMARK 5.2. The calculation leading to (5.5) follows Le Gall and Rosen [11], Lemma 3.1. More generally, their argument gives $\mathbf{E}[I_n^k] \leq (k!)^2 (\mathbf{E}I_n)^k$ for every $k \geq 1$.

COROLLARY 5.3. *Let Δ be an infinite, locally finite, vertex-transitive graph. Denote by V_n the number of vertices in Δ at distance at most n from a fixed vertex o .*

- (i) If $\sup_n V_n/n^4 = \infty$, then two independent sample paths of simple random walk in Δ have finitely many intersections a.s.
- (ii) Conversely, if $\sup_n V_n/n^4 < \infty$, then two independent sample paths of simple random walk in Δ intersect infinitely often a.s.

Proof. For independent simple random walks, reversibility and regularity of Δ imply that

$$\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \mathbf{P}_{x,x}[X_m = Y_n] = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \mathbf{P}_x[X_{m+n} = x] = \sum_{n=0}^{\infty} (n+1) \mathbf{P}_x[X_n = x]. \quad (5.7)$$

- (i) The assumption that $\sup_n V_n/n^4 = \infty$ implies that $V_n \geq cn^5$ for some $c > 0$ and all n : see Theorem 5.11 in Woess [15]. Corollary 14.5 in the same reference yields $\mathbf{P}_x[X_n = x] \leq Cn^{-5/2}$. Thus the sum in (5.7) converges.
- (ii) Combining the results (14.5), (14.12) and (14.19) in Woess [15], we infer that the assumption $V_n = O(n^4)$ implies that $P_x[X_{2n} = x] \geq cn^{-2}$ for some $c > 0$ and all $n \geq 1$. Thus the series (5.7) diverges, so the assertion follows from (1.1) and Theorem 1.4. ■

§6. Concluding Remarks.

The following example shows that the invariance assumption in Theorem 1.4 cannot be omitted.

EXAMPLE 6.1. Consider the graph H , defined as the union of copies of \mathbb{Z}^5 and \mathbb{Z} , joined at a single common vertex o . The Green function for simple random walk in H satisfies

$$G(o, z) = \frac{\deg z}{\deg o} G(z, o) = \frac{2}{12} G(o, o)$$

provided $z \neq o$ is in the copy of \mathbb{Z} . In particular, $\sum_z G(o, z)^2 = \infty$. However, two independent simple random walks on H will have finitely many intersections a.s.

We continue by proving an extension of Corollary 1.5.

COROLLARY 6.2. *Let $X = \langle X_m \rangle$ and $Y = \langle Y_n \rangle$ be independent transient Markov chains on a state space \mathcal{S} and that have the same transition probabilities. Let Z be a subset of \mathcal{S} such that Z is a.s. hit infinitely often by X (and so by Y). Denote by $L_Z(X)$ the sequence obtained from X by erasing any cycle that starts (and ends) at a state in Z , where the*

erasure is made when the cycle is created. Then on the event that $X \cap Y \cap Z$ is infinite, almost surely $L_Z(X) \cap Y \cap Z$ is also infinite.

Proof. Let $m(0) = 0$ and $m(j+1) := \min\{k > m(j) : X_k \in Z\}$ for all $j \geq 0$. Then $X^Z = \langle X_j^Z \rangle := \langle X_{m(j)} \rangle$ is a Markov chain (“the chain X induced on Z ”). Similarly, let $Y^Z = \langle Y_i^Z \rangle = \langle Y_{n(i)} \rangle$ denote the chain Y induced on Z . Since $\text{LE}(X^Z) = L_Z(X) \cap Z$, and $Y^Z = Y \cap Z$ as sets of vertices, the assertion follows by applying Theorem 1.1 to the chains X^Z and Y^Z on the state space Z . \blacksquare

A natural question suggested by Theorem 1.1 is the following.

QUESTION 6.3. Let $X = \langle X_m \rangle$ and $Y = \langle Y_n \rangle$ be independent transient Markov chains on a state space \mathcal{S} and that have the same transition probabilities. Suppose that $|\{X_m\} \cap \{Y_n\}| = \infty$ a.s. Must $|\text{LE}\langle X_m \rangle \cap \text{LE}\langle Y_n \rangle| = \infty$ a.s?

This question is open even for simple random walk in \mathbb{Z}^3 . For simple random walk in \mathbb{Z}^4 , an affirmative answer was given by Lawler [10].

Our final question arose from an attempt to compare the stationary σ -fields defined by a Markov chain and by its loop-erasure.

QUESTION 6.4. Let $\langle X_m \rangle$ be a transient Markov chain. Consider $\langle L_j \rangle = \text{LE}\langle X_m \rangle_{m \geq 0}$ and $\langle L_j^* \rangle = \text{LE}\langle X_m \rangle_{m \geq 1}$. Does there a.s. exist some integer k such that $L_j^* = L_{j+k}$ for all large j ?

The answer is certainly positive if $\langle X_m \rangle$ has infinitely many “cutpoints” a.s.; this is the case for transient random walks in \mathbb{Z}^d : see James and Peres [7]. However, there exist transient chains without cutpoints (see James [6]).

Acknowledgement. We are indebted to Itai Benjamini, François Ledrappier, Robin Pemantle, Tom Salisbury, and Jeff Steif for useful discussions. We thank the referee for helpful comments and corrections.

REFERENCES

- [1] BENJAMINI, I., LYONS, R., PERES, Y., and SCHRAMM, O. (2001). Uniform spanning forests, *Ann. Probab.* **29**, 1–65.
- [2] BURKHOLDER, D.L., DAVIS, B.J. and GUNDY, R.F. (1972). Integral inequalities for convex functions of operators on martingales, *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley 1970/1971), Vol. II: Probability theory*, 223–240, Univ. California Press, Berkeley.
- [3] ERDŐS, P. and TAYLOR, S.J. (1960). Some intersection properties of random walk paths, *Acta Math. Acad. Sci. Hungar.* **11**, 231–248.
- [4] FITZSIMMONS, P.J. and SALISBURY, T. (1989). Capacity and energy for multiparameter Markov processes. *Ann. Inst. H. Poincaré Probab. Statist.* **25**, 325–350.
- [5] HÄGGSTRÖM, O. (1995). Random-cluster measures and uniform spanning trees, *Stoch. Proc. Appl.* **59**, 267–275.
- [6] JAMES, N. (1996). *Ph.D. Thesis*. University of California, Berkeley.
- [7] JAMES, N. and PERES, Y. (1996). Cutpoints and exchangeable events for random walks. *Teor. Veroyatnost. i Primenen.* **41**, 854–868. Reproduced in *Theory Probab. Appl.* **41** (1997), 666–677.
- [8] KAHANE, J.-P. (1985). *Some Random Series of Functions*. Second edition, Cambridge University Press, Cambridge.
- [9] LAWLER, G. (1991). *Intersections of Random Walks*. Birkhäuser, Boston.
- [10] LAWLER, G. (1998). Loop-erased walks intersect infinitely often in four dimensions, *Electronic Communications in Probability* **3**, 35–42.
- [11] LE GALL, J.-F. and ROSEN, J. (1991). The range of stable random walks, *Ann. Probab.* **19**, 650–705.
- [12] PEMANTLE, R. (1991). Choosing a spanning tree for the integer lattice uniformly, *Ann. Probab.* **19**, 1559–1574.
- [13] SALISBURY, T. (1996). Energy, and intersections of Markov chains. *Random Discrete Structures*, IMA Volume 76, Aldous, D. and Pemantle, R. (Editors), Springer-Verlag, New York, 213–225.
- [14] WILSON, D.B. (1996). Generating random spanning trees more quickly than the cover time, *Proceedings of the Twenty-eighth Annual ACM Symposium on the Theory of Computing (Philadelphia, PA, 1996)*, 296–303, ACM, New York.
- [15] WOESS, W. (2000). *Random Walks on Infinite Graphs and Groups*. Cambridge Tracts in Mathematics, **138**, Cambridge University Press, Cambridge.

DEPARTMENT OF MATHEMATICS, INDIANA UNIVERSITY, BLOOMINGTON, IN 47405-5701, USA
`rdlyons@indiana.edu`
`http://php.indiana.edu/~rdlyons/`

AND

SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA, GA 30332-0160
`rdlyons@math.gatech.edu`

DEPARTMENT OF STATISTICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CA 94720-3860, USA
`peres@stat.berkeley.edu`
`http://www.stat.berkeley.edu/~peres/`

MICROSOFT RESEARCH, ONE MICROSOFT WAY, REDMOND, WA 98052, USA
`schramm@microsoft.com`
`http://research.microsoft.com/~schramm/`