The Annals of Probability
1991, Vol. 19, No. 4, 1559–1574

CHOOSING A SPANNING TREE FOR THE INTEGER LATTICE UNIFORMLY[1]

BY ROBIN PEMANTLE

Cornell University

Consider the nearest neighbor graph for the integer lattice $\mathbf{Z}^d$ in d dimensions. For a large finite piece of it, consider choosing a spanning tree for that piece uniformly among all possible subgraphs that are spanning trees. As the piece gets larger, this approaches a limiting measure on the set of spanning graphs for $\mathbf{Z}^d$. This is shown to be a tree if and only if $d \le 4$. In this case, the tree has only one topological end, that is, there are no doubly infinite paths. When $d \ge 5$ the spanning forest has infinitely many components almost surely, with each component having one or two topological ends.

1. Introduction. Let $\mathbf{Z}^d$ be the nearest neighbor graph on the d-dimensional integer lattice, so there is an edge between $(v_1, \dots, v_d)$ and $(w_1, \dots, w_d)$ if and only if $\sum_i |v_i - w_i| = 1$. The term *subgraph* will be used to denote any subcollection of these edges. A subgraph of $\mathbf{Z}^d$ *spans* $\mathbf{Z}^d$ if it contains at least one edge incident to each vertex. A graph is a *forest* if it has no loops and a *tree* if it is a connected forest. A spanning tree on $\mathbf{Z}^d$ is thus a connected, loopless subgraph of $\mathbf{Z}^d$ that spans $\mathbf{Z}^d$.

For measure theoretic purposes, subgraphs are viewed as maps from the set of edges of $\mathbf{Z}^d$ to $\{0, 1\}$. Topologize the space of all subgraphs by the product topology, generated by the *cylinder sets*, namely, those sets depending on only finitely many edges. There is a Borel σ-field for this topology and it is also generated by the *elementary* cylinder sets, $C(A)$, where A is a finite set of edges and $C(A)$ is the set of subgraphs containing all the edges in A. For measures on the Borel σ-field, $\nu_n \to \nu$ weakly iff $\nu_n(C) \to \nu(C)$ for every cylinder set C; it suffices to check this for elementary cylinder events $C(A)$.

This paper is concerned with the following method of picking a spanning tree on $\mathbf{Z}^d$ at random. Let B_n be the box of diameter $2n$ centered at the origin, so it has $(2n + 1)^d$ vertices and all the nearest neighbor edges between these vertices. Let $|v - w|$ denote the metric $\max\{v_i - w_i\}$; this is convenient for counting and for making B_n a sphere, although any equivalent metric could be substituted throughout with no change to the theorems. There are finitely many spanning trees on B_n so there is a uniform measure $\mu_1(B_n)$ on spanning trees of B_n. Any spanning tree on B_n is a subgraph of $\mathbf{Z}^d$, so one

Received March 1990; revised September 1990.
[1]Research supported by a National Science Foundation postdoctoral fellowship and by a Mathematical Sciences Institute postdoctoral fellowship.
AMS 1980 *subject classifications*. 60C05, 60K35.
Key words and phrases. Spanning tree, spanning forest, loop-erased random walk.

may view the measure $\mu_1(B_n)$ as a measure on subgraphs of $\mathbf{Z}^d$. It turns out that these measures converge weakly as $n \to \infty$ to a measure μ on spanning forests of $\mathbf{Z}^d$. For notational convenience, abbreviate $\mu\{\mathbf{T}: \cdots\}$ to $\mu(\cdots)$.

The main tool for proving this basic result is the equivalence (for finite graphs) between uniform spanning trees and random walks. Together with the further equivalence between random walks and electrical networks, this provides a basis for proving that the measures $\mu_1(B_n)$ converge as well as proving some ergodic properties of the limiting measure μ that will be important later. This groundwork is laid in Section 2.

The rest of the paper is concerned with the geometry of the typical sample from the measure μ. It is easy to see that μ concentrates on spanning forests of $\mathbf{Z}^d$. The first result is that in dimensions $d \le 4$ the measure concentrates on spanning trees, while in dimensions $d \ge 5$ the spanning forest will almost surely have infinitely many components. The shape can be further described by the number of topological ends. For a tree, the number of topological ends is just the number of infinite, self-avoiding paths from any fixed vertex. It turns out that when $d \le 4$ the measure concentrates on spanning trees with only one end. When $d \ge 5$ the measure concentrates on spanning forests in which each of whose components has one or two topological ends.

The machinery used to prove these shape results is Lawler's theory of loop-erased random walks (LERW). These are defined in Section 3 and the required basic results about LERW are referenced or proved. The shape results are then proved in Section 4.

2. Uniform spanning trees, random walks and electrical networks. For any connected finite graph G, let $\mu_1(G)$ be the uniform measure on spanning trees of G, as in Section 1. Let v be any vertex of G. The following defines a measure $\mu_2(G, v)$ which will turn out to be the same as $\mu_1(G)$, independently of v: Let $\gamma = \gamma(0), \gamma(1), \dots$ be a simple random walk (SRW) on G starting from $v = \gamma(0)$. Let $\mathbf{T}(\gamma)$ be the subgraph of G containing precisely those edges $\overline{\gamma(i)\gamma(i+1)}$ for which there is no $j < i$ with $\gamma(j) = \gamma(i+1)$. Another way to describe $\mathbf{T}(\gamma)$ is "walk along gamma and draw in each edge as you go except when drawing in an edge would close a loop." The graph $\mathbf{T}(\gamma)$ depends only on $\gamma(0), \dots, \gamma(\tau)$, where τ is the first time γ has visited every vertex. The SRW measure on paths γ projects to a measure $\mu_2(G, v)$ on subgraphs of G. By viewing these edges as oriented from $\gamma(i)$ to $\gamma(i+1)$, it is easy to see that the resulting subgraph is a spanning tree on G oriented away from v.

Lemma 2.1. *For any vertex v of a finite graph G, $\mu_1(G) = \mu_2(G, v)$.*

Proof. This result is due to Diaconis and Doyle; a more complete account can be found in Aldous (1988) or Broder (1988). Let $\{v_i: i \in \mathbf{Z}\}$ be the stationary Markov chain corresponding to a SRW on G. Let $\mathbf{T}_i$ be the rooted tree whose oriented edges are just those edges $\overline{v_j v_{j+1}}$ for which v_{j+1} is distinct from every v_k, for $i \le k < j$. It is easy to check that $\mathbf{T}_i$ is indeed loopless and almost surely connected and that all edges are oriented away from v_i, which is taken to be the root. Furthermore, it may be verified that $\{\mathbf{T}_i\}$ is a stationary Markov chain on the space of rooted spanning trees of G and that a unique stationary measure for it is given by letting the measure of each rooted tree be proportional to the number of neighbors of the root. This means that conditioning on the root of the tree (which is just v_0) leaves a uniform unrooted spanning tree. Now the SRW measure from v is just the stationary Markov measure on $\{v_i: i \geq 0\}$ conditioned on $v_0 = v$. Thus $\mu_2(G, v)$ is distributed as $\mathbf{T}(v_0, v_1, \dots)$, where $\{v_i\}$ is a stationary Markov chain conditioned on $v_0 = v$. This has just been shown to be uniform, and the proof is done. □

For any edge $e = \overline{vw}$ in a finite graph G, define the *contraction* of G by e to be the graph G/e obtained by removing e and identifying v and w. This may result in parallel edges, which must still be regarded as distinct, or in loops (edges whose endpoints are not distinct), which may for the purposes of what follows be thrown away. The *deletion* of e from G is just the graph $G - e$ consisting of all edges of G except e. Contraction commutes and associates with deletion, so it makes sense to speak of the graph G with $e_1, \dots, e_r$ contracted and $e'_1, \dots, e'_s$ deleted. Note that there are natural identifications $\phi^{(-e)}$ and $\phi^{(/e)}$ between edges of G other than e and edges of either G/e or $G - e$.

Now another measure will be defined on subgraphs of a given graph G that turns out to be the same as $\mu_1(G)$. Let $\mathscr{E} = e_1, e_2, \dots$ be any enumeration of the edges of a finite graph G. Define $\mu_3(G, \mathscr{E})$ recursively as follows: The start of the recursion is that if G is a single vertex then $\mu_3(G)$ is the point mass at G. To continue the recursion, assume that $\mu_3(G)$ is defined for all contractions and deletions of G and all enumerations. To define $\mu_3(G, \mathscr{E})$ begin by throwing out all loops and putting a 1-ohm resistor along each edge. Put the terminals of a battery at the two ends of e_1. Look at the total current that flows through the battery and see what fraction of it flows through the resistor at e_1. Call this fraction p. There is a random walk interpretation for p: It is the probability that a simple random walk started at one end of the edge e reaches the other end for the first time by moving along e. Let the $\mu_3(G)$ measure give probability p to the event $e_1 \in \mathbf{T}$ and $1 - p$ to the complementary event. The specification of μ_3 is completed by stating the conditional distributions of μ_3 given $e_1 \notin \mathbf{T}$ and $e_1 \in \mathbf{T}$. To do this write $\mathscr{E}' = e_2, e_3, \dots,$ where $e_2, e_3, \dots$ are viewed as edges in $G - e$ or G/e via the natural identifications $\phi^{(-e)}$ and $\phi^{(/e)}$. Then the distribution of $\mu_3(G, \mathscr{E})$ given $e_1 \notin \mathbf{T}$ is just $\mu_3(G - e_1, \mathscr{E}')$, which is a measure on subgraphs of $G - e_1$, hence on subgraphs of G via $\phi^{(-e)}$. Let the distribution of $\mu_3(G, \mathscr{E})$ given $e_1 \in \mathbf{T}$ be given by adding the edge e_1 to a subgraph of G chosen by picking a subgraph of G/e_1 from $\mu(G/e_1, \mathscr{E}')$ and viewing it as a subgraph of G by the natural identification $\phi^{(/e)}$.

LEMMA 2.2. *For any enumeration $\mathscr{E} = e_1, e_2, \dots$ of the edges of a finite connected graph G, measure $\mu_3(G, \mathscr{E})$ is equal to $\mu_1(G)$.*

PROOF. The idea of the proof is that μ_1 satisfies the same recursion as μ_3. Begin by observing that the spanning trees of G that do not contain an edge e are in one-to-one correspondence with the spanning trees of $G - e$. Second, observe that the spanning trees of G that do contain e are in one-to-one correspondence with the spanning trees of G/e, where the correspondence is given by subtracting the edge e. This is because the identification of the endpoints of e in G/e makes a set of edges of G/e contain a loop if and only if the set together with e contains a loop in G. It is clear that single edge loops of G/e may be thrown out.

These observations imply that $\mu_1(G)$ conditioned on $e \in \mathbf{T}$ is just $\mu_1(G/e)$ and $\mu_1(G)$ conditioned on $e \notin \mathbf{T}$ is just $\mu_1(G - e)$. The next thing is to see that the event $B = \{e_1 \in \mathbf{T}\}$ has the same probability under μ_1 as it does under $\mu_3(G, \mathscr{E})$ for any enumeration $\mathscr{E}$ beginning with $e_1 = \overline{vw}$. By Lemma 2.1, $\mu_1(B)$ is the probability that a SRW on G from v has just traveled across e when it hits w for the first time. By the well-known correspondence between random walks and electrical networks [see Doyle and Snell (1984), Section 3.4], this is precisely the fraction p of the current that flows across e_1 in the electrical scenario used to define μ_3.

Now it follows that if $\mu_1(G/e_1) = \mu_3(G/e_1, \mathscr{E}')$ and if either $G - e$ is disconnected or $\mu_1(G - e_1) = \mu_3(G - e_1, \mathscr{E}')$, then $\mu_1(G) = \mu_3(G, \mathscr{E})$. The initial conditions are certainly the same: If G is a single vertex then $\mu_1(G)$ is the point mass at G. By induction on the number of edges, it follows that $\mu_1(G) = \mu_3(G, \mathscr{E})$ for all finite connected graphs and enumerations. □

THEOREM 2.3. *Let $\{B_n\}$ be any sequence of finite sets of edges of $\mathbf{Z}^d$, $d \geq 2$, converging to $\mathbf{Z}^d$ in the sense that any edge is in all but finitely many sets B_n. Then the measures $\mu_1(B_n)$ converge weakly to a limiting measure μ in the sense that $\mu_1(B_n)(C) \to \mu(C)$ for any cylinder event C. The measure μ is concentrated on spanning forests of $\mathbf{Z}^d$ and is translation invariant.*

PROOF. For weak convergence it suffices to show that $\mu_1(B_n)(C)$ converges for the special case where C is the event $C(A)$ that all edges in a finite set A are in the random subgraph. This is because the probabilities of $C(A)$ determine the probabilities of all cylinder events by inclusion–exclusion and because, if all cylinder probabilities converge, the limits of these must define a measure.

Proceed by fixing a set $A = e_1, \dots, e_k$. When n is sufficiently large, so $A \subseteq B_n$, let $\mathscr{E}_n$ be an enumeration of the edges of B_n that begins with $e_1, \dots, e_k$. Then, by Lemma 2.2, $\mu_1(B_n)(C(A)) = \mu_3(B_n, \mathscr{E})(C(A)) = \prod_{j=1}^k p_j^{(n)}$, where $p_j^{(n)}$ is the $\mu_3(B_n/e_1/\cdots/e_{j-1}, \mathscr{E}^{(j-1)})$ probability of $\{e_j \in \mathbf{T}\}$. This is just the fraction of current that flows through e_j when a battery is placed across e_j in the resistor network $B_n/e_1/\cdots/e_{j-1}$.

Consider for a moment the special case where B_n is a box of diameter $2n$ centered at the origin. Then, for $r < 0$, B_n is just B_{n+r} with a lot of edges removed. Since contraction and deletion commute, $B_n/e_1/\cdots/e_{j-1}$ is just a deletion of $B_{n+r}/e_1/\cdots e_{j-1}$. It follows from Raleigh's monotonicity law [Doyle and Snell (1984), Chapter 4] that more current flows in $B_{n+r}/e_1/\cdots/e_{j-1}$ than in $B_n/e_1/\cdots/e_{j-1}$. Since the same current flows directly across the edge e_j, it follows that $p_j^{(n)} \geq p_j^{(n+r)}$ and, by taking the product, that $\mu_1(B_n)(C(A)) \geq \mu_1(B_{n+r})(C(A))$. The sequence of probabilities is therefore decreasing in n and must converge for each A.

For general B_n, note that the B_n eventually contain any finite box and are each contained in some finite box. The monotonicity proof worked for any graphs, one containing the other. Then the probabilities $\mu_1(B_n)(C(A))$ interlace the sequence of probabilities of $C(A)$ for boxes of diameter $2n$ and hence converge to the same limit.

The rest is immediate. There are no loops in the final measure μ, because any loop $e_1, e_2, \ldots, e_k$ is a finite cylinder event and has probability zero under each $\mu_1(B_n)$. Also, the event that vertices $v_1, \ldots, v_k$ are a component not connected to the rest of the graph is a cylinder event on any box B_n big enough to contain all edges incident to any v_i. The $\mu_1(B_n)$ probability of this event is zero, since $\mu_1(B_n)$ concentrates on connected graphs, so the limit is zero. For stationarity, note that $\mu_1(B_n)(C(\pi A)) = \mu_1(\pi^{-1}B_n)(C(A))$ for any translation π. The interlacing argument shows that using the sequence $\pi^{-1}B_n$ in place of B_n does not affect the limit, so $\mu(C(\pi A)) = \mu(A)$ for any event $C(A)$. These events determine the measure, hence μ is translation invariant. □

For any set A of edges, let $\sigma(A)$ denote as usual the σ-field generated by the events $C(A')$ for A' a finite subset of A. Let $\mathscr{T}$ denote the tail σ-field, which is just the intersection of $\sigma(A)$ over all cofinite sets A.

THEOREM 2.4. *Let μ be the measure defined previously on spanning forests of $\mathbf{Z}^d$. Then the tail field is trivial, that is, $\mu(C) = 0$ or 1 for every $c \in \mathscr{T}$.*

PROOF. First the electrical viewpoint will be used to reduce the statement to a more specialized proposition and then the random walk construction will be used to prove the proposition.

Begin with the device used to prove Kolmogorov's zero–one law: An event is trivial if it is independent from every event in a sufficiently large set. Letting C be any tail event, it suffices to show that if $\mu(C) > 0$, then the conditional probabilities $\mu(\cdot|C)$ agree with μ on elementary cylinder sets. For $n > 0$, let B_n be boxes of diameter $2n$ centered at the origin and let C_n be cylinder sets in $\sigma(\mathbf{Z}^d \setminus B_n)$ such that $\mu(C_n \vartriangle C) \to 0$. In particular, the sequence $\{\mu(C_n)\}$ has a positive lim inf and it will suffice to show that for each finite set of edges A, $\mu(C(A)|C_n) \to \mu(C(A))$ for n such that $\mu(C_n) \neq 0$. By Lemma 2.2, it suffices to show that for any sequence of boxes B'_n big enough so that $C_n \in \sigma(B'_n)$, $\mu_3(B'_n)(C(A)|C_n) \to \mu_3(B'_n)(A)$ at least for those n such that $\mu_3(B'_n)(C_n) \neq 0$.

To do this, consider the electrical networks G_1 and G_2 where G_1 is just B_n and G_2 is gotten by contracting all edges outside of B_n, which is electrically the same as short-circuiting the boundary ∂B_n of the box B_n. I claim that $\mu_3(B'_n)(C(A)|D)$ is bounded below by $\mu_3(G_2)(C(A))$ and above by $\mu_3(G_1)(C(A))$

for any event $D \in \sigma(B'_n \setminus B_n)$. To see this, let $\mathscr{E}$ be an enumeration of the edges in B'_n beginning with those not in B_n. The event D is a union of cylinder events that specify precisely which edges in $B'_n \setminus B_n$ are present. Conditioning on such an event is, by the construction of μ_3, the same as doing the electrical computations on a contraction–deletion of B'_n. Thus $\mu_3(B'_n, \mathscr{E})(\cdot \mid D)$ is a mixture of $\mu_3(G, \mathscr{E}')(\cdot)$ as G ranges over contraction–deletions of B'_n (where $\mathscr{E}'$ is what is left of the enumeration when you get to B_n). The claim is then just Raleigh's monotonicity; $\mu_3(G, \mathscr{E}')(C(A))$ is a product of conditional probabilities p_j as in the proof of Theorem 2.3; any contraction–deletion of B'_n can be contracted to G_2 or deleted to G_1; monotonicity says that contracting increases total current and deleting decreases it, so each p_j increases with deletion and decreases with contraction, and the claim is shown.

It remains to show that $\mu_3(G_1)(C(A)) - \mu_3(G_2)(C(A)) \to 0$ as $n \to \infty$ for each A. For this, use the random walk scenario. Let B_M be a box containing A. Let $\varepsilon > 0$ be arbitrary and L be large enough so that the union of L independent SRW's started anywhere on ∂B_M will cover all the edges of A with probability at least $1 - \varepsilon$. The following fact can be found in or deduced from Lawler (1991): The hitting measure of ∂B_M for SRW on B_n started from the vertex v converges as v goes to infinity and n varies arbitrarily with $v \in B_n$. This implies that, for sufficiently large n, the total variation distance between the hitting measures on ∂B_M from any two vertices on ∂B_n can be made less than ε/L.

Now view G_1 and G_2 as graphs and couple SRW's γ_i from the origins on G_i as follows: They are the same until they hit the boundary (which has been collapsed to a single point in G_2). Then they are coupled so that their next hits of ∂B_M occur in the same place (though not necessarily at the same time) with probability as close to 1 as possible; this probability is at least $1 - \varepsilon/L$. Then they make the same moves until they hit ∂G_i, become recoupled as often as possible when they hit ∂B_M again, and so on. The probability is at least $1 - \varepsilon$ that γ_1 and γ_2 are coupled whenever they are inside B_M up to the first L hits of ∂B_M. At this point, the probability is at least $1 - \varepsilon$ that all edges in B_M have been traversed, in which case the subgraph $\mathbf{T}(\gamma_1)$ is in the event $C(A)$ if and only if $\mathbf{T}(\gamma_2)$ is. Thus $|\mu_3(G_1)(C(A)) - \mu_3(G_2(C(A))| < 2\varepsilon$. Since ε was arbitrary, that sandwiches $\mu(C(A) \mid C_n)$ between sequences with the same limit and proves the theorem. □

3. Loop-erased random walk. This section contains lemmas about loop-erased random walks. The reason that loop-erased random walks are relevant to this paper will be clear later but, briefly, it is the following: When $\mu_2(G, v)$ is used to construct a random spanning tree on G, the unique path connecting a vertex w to v is given by a loop-erased random walk from w to v. The section is self-contained, but not formal. For a more complete development, see Lawler (1991) or Lawler (1980, 1983 and 1986).

Let G be any graph and let γ be a path on G. The following notational conventions will be used throughout. The ith vertex visited by γ is denoted

$\gamma(i)$, beginning at $\gamma(0)$. If γ is finite then $l(\gamma)$ denotes the length of γ and γ' denotes the time reversal of γ, so $\gamma'(0) = \gamma(l(\gamma))$. If in addition there is a path β with $\beta(0) = \gamma'(0)$, then $\gamma * \beta$ denotes γ followed by β. The paths β and γ are said to intersect whenever $\beta(i) = \gamma(j)$ for some i and j not necessarily equal but not both zero. Finally, $\gamma \wedge n$ denotes the initial segment $\{\gamma(i): i \le n\}$ of γ and $\gamma \vee n$ denotes γ from step n onwards, so $\gamma = (\gamma \wedge n)*(\gamma \vee n)$.

For finite paths γ the loop-erasure operator LE is defined intuitively as follows: If γ is a self-avoiding path [meaning that the vertices $\gamma(i)$ are distinct], then $LE(\gamma) = \gamma$. Otherwise, the first time γ visits a vertex v twice, erase the loop at v. In other words, if $\gamma(i) = \gamma(j)$, $i < j$, and j is minimal for this, delete from the sequence $\{\gamma(k)\}$ all the vertices with $i < j \le k$. If the result is still not self-avoiding, then repeat this step until it is. The map LE preserves the initial and final points of a path. For a given initial and final point LE maps onto the set of self-avoiding paths with the given endpoints but it is not one-to-one. Let α be a self-avoiding path and m a positive integer and [following Lawler (1983), using slightly different notation] define $\Gamma^m(\alpha)$ to be $LE^{-1}(\alpha) \cap \{\text{paths of length } m\}$.

If γ is an infinite path that hits every vertex finitely often then the paths $LE(\gamma \wedge n)$ converge to an infinite path which will be called $LE(\gamma)$. When $G = \mathbf{Z}^d$, $d \ge 3$, and γ is a SRW from some vertex v, then γ hits each vertex finitely often almost surely. Consequently $LE(\gamma)$ is almost surely well-defined. The law of $LE(\gamma)$ is called the loop-erased random walk measure on $\mathbf{Z}^d$ from v, or simply LERW. (LERW can be defined on Z^2 as well but will not be needed here.)

Commonly, an alternative construction for LERW is used. Let $\gamma(0)$ be given and let the measure of the event $\gamma(1) = v$ be given by the probability that $\beta(1) = v$, where β is a SRW conditioned never to return to $\gamma(0)$. In general, let the measure of $\gamma(i+1) = v$ conditional on $\{\gamma(j): j \le i\}$ be given by the probability that $\beta(1) = v$, where β is a SRW from $\gamma(i)$ conditioned never to return to $\{\gamma(j), j \le i\}$. A similar construction gives the law of $LE(\gamma)$ when γ is a SRW from v on a finite graph G, stopped upon hitting some vertex w. In this case the conditional probability of $\gamma(i+1) = v$ given $\gamma(j)$ for $j \le i$ is given by the next step of a random walk conditioned to hit w before returning to $\{\gamma(j), j \le i\}$. These characterizations are easy to prove and will be assumed freely when convenient.

LEMMA 3.1. *Let v and w be vertices in $\mathbf{Z}^d$, $d \ge 3$. Let β and γ be independent LERW from v and SRW from w, respectively. Then if $d = 3$ or 4, β and γ intersect infinitely often almost surely. On the other hand, if $d \ge 5$, β and γ intersect finitely often almost surely and the probability that they intersect at all (other than at v if $v = w$) is bounded between $c_1(d)|v-w|^{4-d}$ and $c_2(d)|v-w|^{4-d}$, for some constants $0 < c_1(d) < c_2(d) < \infty$.*

PROOF. The statement for $d = 3$ is proved in Lawler [(1988), equation 3.1] and for $d = 4$ is proved in Lawler [(1986), Theorem 5.1]. For $d \ge 5$, the fact that β and γ intersect finitely often almost surely can be deduced from the corresponding facts for two SRW's and the fact that LERW is a subsequence of SRW. To prove the quantitative bounds for $d \geq 5$, proceed as follows.

Let X be the random number of intersection points of a LERW from v and an independent SRW from w, counted with multiplicity k if the point is hit k times by the SRW. The upper bound is a consequence of the following upper bound on $\mathbf{E}X^2$, which can be found in Lawler [(1991), Chapter 3]:

$$\mathbf{E}X^2 \leq c|v - w|^{4-d}. \tag{1}$$

Since X is an integer-valued random variable, this immediately establishes that $\mathbf{P}(X > 0) \leq c|v - w|^{4-d}$, which is the desired upper bound on the probability that LERW from v intersects an independent SRW from w. The lower bound will be proved by showing

$$\mathbf{E}X \geq c|v - w|^{4-d}. \tag{2}$$

To see that (1) and (2) actually imply $\mathbf{P}(X > 0) \geq c|v - w|^{4-d}$, write

$$\begin{aligned}\mathbf{E}X^2 &= \mathbf{P}(X > 0)\mathbf{E}(X^2|X > 0)\\ &\geq \mathbf{P}(X > 0)(\mathbf{E}(X|X > 0))^2\\ &= \mathbf{P}(X > 0)\left[\frac{\mathbf{E}X}{\mathbf{P}(X > 0)}\right]^2\\ &= (\mathbf{E}X)^2\mathbf{P}(X > 0)^{-1},\end{aligned}$$

hence $\mathbf{P}(X > 0) \geq (\mathbf{E}X)^2/\mathbf{E}X^2$.

To show (2), let β be a SRW from v and $\gamma = LE(\beta)$ be the corresponding LERW from v. Write $G(x, y)$ for the Green's function, that is, the expected number of visits to y of SRW starting at x. It is known [e.g., Lawler (1991)] that $G(x, y)$ is bounded between constant multiples of $|x - y|^{2-d}$ in each dimension greater than or equal to 3; in this regard, let 0^{-n} denote the constant $G(x, x)$ to avoid making explicit exceptions for zero in the summations. Then (2) is implied by

$$\mathbf{P}(x \in \gamma) \geq c|v - x|^{2-d} \tag{3}$$

since this implies

$$\begin{aligned}\mathbf{E}X &= \sum_x \mathbf{P}(x \in \gamma)G(w, x)\\ &\geq c\sum_s |\{x\colon |v - x| = s\}| s^{2-d}(s + |v - w|)^{2-d}\\ &\geq c\sum_s s^{2-d}s^{d-1}(s + |v - w|)^{2-d}\\ &\geq c\sum_{s \geq |v-w|} s^{2-d}s^{d-1}(2s)^{2-d},\end{aligned}$$

which is just $c|v - w|^{4-d}$.

Finally, to show (3) let τ be the first time (possibly infinity) that β hits x, and write

$$\mathbf{P}(x \in \gamma) \geq \mathbf{P}(\tau < \infty)\mathbf{P}(\beta \wedge \tau \text{ is disjoint from } \beta \vee \tau | \tau < \infty).$$

The first factor is at least $c|v - x|^{2-d}$ so it remains to bound the second factor away from zero. Since $\beta \vee \tau$ is independent of $\beta \wedge \tau$ given $\tau < \infty$, the second factor is the probability that $\beta \wedge \tau$ is disjoint from an independent SRW β_1 from x, where β is a SRW from v conditioned to hit x. Write $\beta_2 = (\beta \wedge \tau)'$, so β_2 is a SWR from x conditioned to hit v and stopped when it hits v. Since two independent SRW's from x are disjoint with positive probability for $d \geq 5$, it remains to show that conditioning one of the walks to reach v does not alter this. We may assume that $|v - x|$ is greater than some fixed constant r_0, since (3) is immediate for $|v - x| \leq r_0$ just from transience of the SRW.

Let γ_1 and γ_2 be independent SRW's from x. Fix any positive ε. Since independent SRW's from x intersect finitely often with probability 1, an M can be chosen large enough so that $\mathbf{P}(\gamma_1 \vee M \text{ intersects } \gamma_2) < \varepsilon$. By transience of SRW, an $M' > M$ can be chosen so that $\mathbf{P}(\gamma_2 \vee M' \text{ intersects } B(x, M)) < \varepsilon$, where $B(y, k)$ is the cube of radius k centered at y. It is known, via triviality of the Martin boundary for SRW [e.g., Lawler (1991), Chapter 2], that SRW from x conditioned to hit y converges weakly to unconditioned SRW from x as $|x - y| \to \infty$, so r_0 may be chosen such that $|x - y| \geq r_0/4$ implies that the total variation difference $\gamma_1 \wedge M$ and $\beta_2 \wedge M$ is less than ε. Similarly, let $r = |x - v|$ and let α be a SRW from x conditioned to avoid $B(v, 3r/4)$; then the same argument about the Martin boundary shows that the distribution of α converges weakly to that of γ_2 as $r \to \infty$, so r_0 can be chosen large enough so that $r \geq r_0$ implies that the total variation distance between $\gamma_2 \wedge M'$ and $\alpha \wedge M'$ is at most ε.

Now let $p_1 = \mathbf{P}(\gamma_1 \wedge M \text{ is disjoint from } \gamma_2)$. Let $p_2 = \mathbf{P}(\gamma_2 \text{ is disjoint from } B(v, 3r/4))$ and let $p_3 = \min_{y \in B(v, r/2)} \mathbf{P}(\text{SRW from } y \text{ conditioned to hit } v \text{ does so before leaving } B(v, 3r/4))$. Note that p_1 is bounded away from zero by the standard result, while p_2 and p_3 are easily seen by scaling to be bounded away from zero in any fixed dimension. Let σ be the first time β_2 hits $B(v, r/2)$ and write

$$\begin{aligned}
&\mathbf{P}(\beta_2 \text{ is disjoint from } \beta_1) \\
&\quad\geq p_2 \mathbf{P}(\beta_2 \text{ is disjoint from } \alpha) \\
&\quad\geq p_2 \mathbf{P}(\beta_2 \wedge \sigma \text{ is disjoint from } \alpha) \mathbf{P}(\beta_2 \vee \sigma \text{ is disjoint from } \alpha | \alpha) \\
&\quad\geq p_2 \big[\mathbf{P}(\beta_2 \wedge M \text{ is disjoint from } \alpha \wedge M') \\
&\qquad\quad - \mathbf{P}(\beta_2 \wedge M \text{ intersects } \alpha \vee M') \\
&\qquad\quad - \mathbf{P}((\beta_2 \wedge \sigma) \vee M \text{ intersects } \alpha)\big] \mathbf{P}(\beta_2 \vee \sigma \text{ is disjoint from } \alpha | \alpha) \\
&\quad\geq p_2 \big[\mathbf{P}(\gamma_1 \wedge M \text{ is disjoint from } \gamma_2 \wedge M') \\
&\qquad\quad - 2\varepsilon - \mathbf{P}(\alpha \vee M' \text{ is disjoint from } B(x, M)) \\
&\qquad\quad - \mathbf{P}((\beta_2 \wedge \sigma) \vee M \text{ intersects } \alpha)\big] p_3 \\
&\quad\geq p_2 p_3 \big[p_1 - 2\varepsilon - 2\varepsilon - \mathbf{P}((\beta_2 \wedge \sigma) \vee M \text{ intersects } \alpha)\big]
\end{aligned}$$

by choice of M and M'. Since p_i are all bounded away from zero, it remains to show that $\mathbf{P}((\beta_2 \wedge \sigma) \vee M$ intersects $\alpha)$ is small. But the distribution of $\beta_2 \wedge \sigma$ is given by a SRW conditioned to hit $B(v, r/2)$ at some random point y, stopped when it does so, reweighted by $\mathbf{P}$(SRW from y hits v) and normalized. Scaling shows that $\mathbf{P}$(SRW from x hits $B(v,(r/2))$) is bounded below and, as y varies over the boundary of $B(v, r/2)$ in a fixed dimension, the ratios of these reweights are bounded. Thus the Radon–Nikodym derivative $d(\beta_2 \wedge \sigma)/d(\text{SRW} \wedge \sigma)$ is bounded above, and hence $\mathbf{P}((\beta_2 \wedge \sigma) \vee M$ intersects $\alpha)$ is bounded by a constant times $\mathbf{P}(\gamma_1 \vee M$ intersects $\alpha)$ and the latter is at most $p_2^{-1}\varepsilon$. This completes the proof that $\mathbf{P}(\beta_2$ is disjoint from $\beta_1)$ is bounded away from zero, thus proving (3) and (2). □

LEMMA 3.2. *Let G be any graph and α a finite path in G. Let $\Phi^m(\alpha) = \{\beta: \beta' \in \Gamma^m(\alpha')\}$ be the set of paths of length m whose "backward loop-erasure" is α. Then for each m there is a bijection $T^{m,\alpha}$ between $\Gamma^m(\alpha)$ and $\Phi^m(\alpha)$ such that the multiset of sites visited by γ is the same as the multiset of sites visited by $T^{m,\alpha}(\gamma)$.*

PROOF. Lawler [(1983), Proposition 2.1] states this for $G = \mathbf{Z}^d$ and for sets instead of multisets. The proof actually shows that multisets are preserved. Clearly, if the proposition is true for $\mathbf{Z}^d$ it is true for subgraphs of $\mathbf{Z}^d$, which is all that is used below. It is easy, however, to see that Lawler's proof is valid for any graph. □

LEMMA 3.3. *Let w be any vertex in $\mathbf{Z}^d$, $d \geq 3$. For any positive integer L, let x be a vertex in B_n at distance at least L from w, where B_n is large enough to contain w. Let γ be a SRW from x on B_n conditioned to hit w before returning to x and let $\alpha = LE(\gamma')$. Then the distribution of the first M steps of α converges as $n, L \to \infty$ to the distribution of the first M steps of LERW on $\mathbf{Z}^d$ from w, the convergence being uniform over choices of x.*

PROOF. First note that by time reversal, γ' is distributed as SRW from w conditioned to hit x before returning to w. It suffices to show that, for each self-avoiding path β of length $j < M$ from w and for each neighbor v of $\beta(j)$, the conditional probability that $\alpha \wedge j + 1(j+1) = v$ given $\alpha \wedge j = \beta$ approaches $\mathbf{P}(\text{LERW} \wedge j + 1(j+1) = v | \text{LERW} \wedge j = \beta)$. By the alternative construction for LERW, the latter probabilities for fixed β are proportional to the quantities $p(v)$ defined by $p(v) = \mathbf{P}$(SRW from v never hits β) and are thus given by $p(v)$ normalized to sum to 1. Similarly, the former probabilities are proportional to $q(v) = \mathbf{P}$(SRW on B_n from v hits x before hitting β).

Let K be the box such that $x \in \partial B_K$, so that, for fixed w, $K \to \infty$ as $L \to \infty$. Let $Q(\cdot)$ be the hitting measure on the boundary of B_K for SRW from w conditioned to avoid β. It is known [e.g., Lawler (1991), Theorem 2.1.2] that $Q(y)$ is bounded between $1 - \varepsilon(K, \beta)$ and $1 + \varepsilon(K, \beta)$ times the hitting measure for SRW starting from the origin, where $\varepsilon(K, \beta) \to 0$ as $K \to \infty$. To make use of this, write

$$q(v) = \mathbf{P}(\text{SRW on } B_n \text{ from } v \text{ hits the boundary of } B_K \text{ before hitting } \beta)$$

$$\times \sum_{y \in \partial B_K} Q(y)\mathbf{P}(\text{SRW on } B_n \text{ from } y \text{ hits } x \text{ before hitting } \beta(0), \dots, \beta(j)). \tag{4}$$

The first factor on the right-hand side of (4) converges to $p(v)$ as $K \to \infty$. The second factor, according to the preceding observation about Q, may only vary with v by a factor of at most $1 \pm \varepsilon(K, \beta)$. Thus, for fixed β, $q(v)$ normalized converges to $p(v)$ normalized as n, $K \to \infty$ uniformly in $x \in B_K$, hence as n, $L \to \infty$ uniformly in x at distance at least L from y, and the proof is done. □

LEMMA 3.4. *Remove the conditioning in Lemma* 3.3 *so that* γ *may return any number of times to* x *before hitting* w. *Then*:

(i) *the conclusion that* $\alpha \wedge M$ *converges to* $LERW|M$ *uniformly in* x *still holds*;
(ii) $(LE(\gamma'))'$ *has the same distribution as* $LE(\gamma)$.

PROOF. For finite paths β from x in B_n, let $W(\beta) = W(B_n, \beta)$ denote $\mathbf{P}(\gamma \wedge l(\beta) = \beta)$, which can be written $\prod_i(\text{number of neighbors of } \beta(i))^{-1}$. To prove (ii), write

$$\mathbf{P}(LE(\gamma') = \alpha') = \sum_m \sum_{\beta \in \Gamma^m(\alpha') \cap S} W(\beta'), \tag{5}$$

where S is the set of paths that never return to w. Since the bijections $T^{m,\alpha}$ of Lemma 3.2 preserve the multiset of sites visited, they preserve W and can be used to rewrite (5) as

$$\sum_m \sum_{\beta \in \Phi^m(\alpha') \cap S} W(\beta'),$$

which is, by definition of Φ^m, just

$$\sum_m \sum_{\beta' \in \Gamma^m(\alpha) \cap S} W(\beta'),$$

which is $\mathbf{P}(LE(\gamma) = \alpha)$.

To prove (i), note that the distribution of $LE(\gamma)$ is independent of the number of times γ returns to x. Then, by (ii), the distribution of $LE(\gamma')$ is independent of the number of times γ returns to x. In particular, it is unaffected by conditioning on this number being zero; thus Lemma 3.3 holds even after conditioning. □

4. Number and shape of the components.

The following easy lemma connects the loop-erased random walk to the random walk method of generating a random spanning tree of a finite graph. Recall the definition of $\mathbf{T}(\gamma)$ at the beginning of Section 2.

LEMMA 4.1. *Let v and w be distinct vertices of a finite graph G and let γ be any path from v to w, not necessarily self-avoiding. Then the unique path connecting w to v in $\mathbf{T}(\gamma)$ is given by $LE(\gamma')$.*

PROOF. Let α be $LE(\gamma')$ and β be the path connecting w to v in $\mathbf{T}(\gamma)$. Clearly $\alpha(0) = \beta(0) = w$. Now assume for induction that $\alpha(i) = \beta(i)$. Then $\beta(i+1)$ is the unique x for which $\mathbf{T}(\gamma)$ has an oriented edge $\overline{x\beta(i)}$. This is just $\gamma(j-1)$, where j is minimal such that $\gamma(j) = \beta(i)$. This is also equal to $\gamma'(j+1)$, where j is maximal for $\gamma'(j) = \beta(i) = \alpha(i)$. Then, when applying loop-erasure to γ', the edge from $\alpha(i)$ to x is never erased; hence $\alpha(i+1) = x$. By induction, $\alpha = \beta$. $\square$

The main theorem on connectedness can now be proved.

THEOREM 4.2. *Let μ be the limiting measure on subgraphs of $\mathbf{Z}^d$, $d \geq 3$, constructed in Section 2. Then, for $d = 3$ or 4, μ concentrates on connected graphs. For $d \geq 5$, μ concentrates on graphs with infinitely many components. In this case, $|v-w|^{d-4}\mathbf{P}(v$ and w are connected) is bounded between $c_1(d)$ and $c_2(d)$ for $v \neq w$ and some constants $0 < c_1(d) < c_2(d) < \infty$.*

PROOF. Fix d for the moment. If $d = 2$, μ can be defined via μ_2 without a limiting procedure, since SRW in Z^2 hits every point, and connectedness follows immediately. So assume without loss of generality that $d > 2$. Let v and w be distinct vertices. The main project will be determining whether v and w are almost surely connected. If so, then by countable additivity the whole graph is almost surely connected. If not, then another few sentences will show that there are almost surely infinitely many components.

Fix the vertices v and w. The argument will use the random walk scenario, writing μ as the limit of $\mu^n = \mu_2(B_n, v)$ as $n \to \infty$. Let C be the event $\{v$ is connected to $w\}$. Since the convergence is weak and the indicator function $\mathbf{1}_C$ is not continuous, $\mu^n(C)$, which is always 1, does not necessarily converge to $\mu(C)$. To get information about μ we must work instead with the continuous events $C_M = \{v$ is connected to w by a path of length less than or equal to $M\}$. Specifically, weak convergence implies $\mu^n \to \mu$ on each C_M, hence

$$\mu(C) = \lim_{M\to\infty} \lim_{n\to\infty} \mu^n(C_M). \tag{6}$$

Another way to say this is to let L_n be the length of the path connecting v and w under μ^n. Then v and w are μ-almost surely connected if and only if the L_n's are tight. Equation (6) will be used to show that $\mu(C)$ is equal to the probability that LERW from w intersects an independent SRW from v [see (9)].

To analyze μ^n, run a SRW β and v on B_n. Let τ be the first time β hits w and let $\gamma = \beta \wedge \tau$. The path connecting v and w in $\mathbf{T}(\beta)$ is determined by γ. There are two possibilities: either β hits ∂B_n before hitting w or vice versa. If it hits w first, it is easy to check that the conditional distribution of the length of γ is tight as $n \to \infty$.

To examine the other possibility, condition (hereafter) on β hitting ∂B_n before w and let x be the first point where β hits ∂B_n. Write $\gamma = \gamma_1 * \gamma_2$, where γ_1 is the initial segment of γ up to the first hit of x and γ_2 is all the rest. Then γ_1 is distributed as SRW from v stopped upon hitting the boundary and conditioned to do this before it hits w. Then as $n \to \infty$ the first M steps of γ_1 converge for each M to the first M steps of an infinite SRW from v conditioned never to hit w.

Recall from Lemma 4.1 that the path connecting w to v is given by $LE(\gamma') = LE(\gamma_2' * \gamma_1')$. Fix any M. Observe that $LE(\gamma_2') \wedge M = LE(\gamma') \wedge M$ whenever $LE(\gamma_2') \wedge M$ is disjoint from γ_1. This is because $LE(\gamma') = LE(LE(\gamma_2') * \gamma_1'))$ and the addition of γ_1' cannot alter any initial segment of $LE(\gamma_2')$ that it does not intersect. It should now be clear where Lemma 3.1 comes in; the rest of the work will be in identifying the distributions of γ_1 and $LE(\gamma_2')$ and taking limits correctly.

Let α be a LERW from w independent from β. Recall from Lemma 3.4 that

$$LE(\gamma_2') \wedge M \to_{\mathscr{D}} \alpha \wedge M \tag{7}$$

as $n \to \infty$, even when conditioned on x. (Here the dependence of γ_2 on n is supressed in the notation.) Since γ_1 and γ_2 are conditionally independent given x, it follows that for any M, the pair $(LE(\gamma_2') \wedge M, \gamma_1 \wedge M)$ converges to $(\alpha \wedge M, \beta \wedge M)$ as $n \to \infty$.

Let D be the event that α and β intersect. Let D_M be the event that $\alpha \wedge M$ and $\beta \wedge M$ intersect, and let D_M' be the event that $\alpha \wedge M$ and β intersect. Then $D_M, D_M' \uparrow D$, so $\mathbf{P}(D_M), \mathbf{P}(D_M') \uparrow \mathbf{P}(D)$.

Recall that C_M is the event that the path connecting v to w in $\mathbf{T}$ has length at most M. Then

$$\gamma_2' \wedge M \cap \gamma_1 \wedge M \neq \varnothing \quad \Rightarrow \quad C_{2M} \quad \Rightarrow \quad \gamma_2' \wedge 2M \cap \gamma_1 \neq \varnothing. \tag{8}$$

It follows from (7) that

$$\lim_{n \to \infty} \mu^n(LE(\gamma_2') \wedge M \cap \gamma_1 \wedge M \neq \varnothing) = \mathbf{P}(D_M).$$

Let $u(M)$ be large enough so that

$$\mathbf{P}(\alpha \wedge 2M \cap \beta \wedge u(M) \neq \varnothing | \alpha \wedge 2M \cap \beta \neq \varnothing) > 1 - 1/M.$$

Then it also follows from (7) that

$$\begin{aligned}\lim_{n \to \infty} \mu^n(LE(\gamma_2') \wedge 2M \cap \gamma_1 \neq \varnothing) &\leq \frac{M}{M-1}\mathbf{P}(\alpha \wedge 2M \cap \beta \wedge u(M) \neq \varnothing) \\ &\leq \frac{M}{M-1}\mathbf{P}(D_{2M}').\end{aligned}$$

Now, taking limits as $n \to \infty$ of (8) gives

$$\mathbf{P}(D_M) \leq \lim_{n \to \infty} \mu^n(C_{2M}) \leq \frac{M}{M-1}\mathbf{P}(D_{2M}').$$

Taking the limit in M and using (6) gives

$$\mathbf{P}(D) = \mu(C). \tag{9}$$

Now, if $d = 3$ or 4, Lemma 3.1 says that the probability of α intersecting an independent SRW from v is 1; since β is distributed as an independent SRW from v conditioned on an event of positive probability, this means $\mathbf{P}(D) = 1$, from which the statement of the theorem follows immediately.

On the other hand, consider the case $d \geq 5$. By Lemma 3.1, the probability that α intersects an independent SRW from v is bounded between constants times $|v - w|^{4-d}$. Since the event that SRW from v actually hits w is of order $|v - w|^{2-d}$, β is distributed as a SRW conditioned on an event of probability $1 - c|v - w|^{2-d}$, and it follows from $\mathbf{P}(A)/\mathbf{P}(B) \geq \mathbf{P}(A|B) \geq (\mathbf{P}(A) - \mathbf{P}(B^c))/\mathbf{P}(B)$ that $\mathbf{P}(D)$ is bounded between constant multiples of $|v - w|^{4-d}$, hence $\mathbf{P}(C)$ is also, which was to be shown. It follows immediately that the measure μ does not concentrate on connected graphs.

To see that the measure concentrates on graphs with infinitely many components, recall from Theorem 2.3 that μ is stationary and from Theorem 2.4 that the tail field is trivial. Then μ is ergodic, so the number of components is some constant K almost surely. To bound K, write $I(x, y)$ for the indicator function of the event that x is connected to y and calculate

$$\mathbf{E} \sum_{x, y \in B_n} I(x, y) = \sum_{x, y \in B_n} \mathbf{E} I(x, y) = \sum_{x, y \in B_n} O\big(|x - y|^{4-d}\big) = O(n^{d+4}).$$

On the other hand, if B_n is partitioned into at most K connected components, $K < \infty$, then

$$\sum_{x, y \in B_n} I(x, y) \geq n^{2d}/K.$$

When $d \geq 5$ this is greater than $O(n^{d+4})$ for any finite K, so K must be infinite almost surely. □

The last theorem is about the shape of the tree when $d \leq 4$.

THEOREM 4.3. *If $d \leq 4$, then the measure μ concentrates on trees having only one topological end, that is, trees for which removal of any vertex divides the tree into components precisely one of which is infinite.*

PROOF. Call a vertex x in a subgraph of $\mathbf{Z}^d$ a separator if removal of x leaves more than one infinite component. Call x a branch point if its removal leaves more than two infinite components. Burton and Keane [(1989), Theorem 2] show that the set of branch points for a subgraph of the integer lattice may not be a set of vertices of positive density. By stationarity and ergodicity it follows that there are no branch points at all almost surely. Then the tree has at most two topological ends.

The number of topological ends is translation invariant, hence almost surely constant. Assume for contradiction that there are almost surely two. Then the spanning tree $\mathbf{T}$ looks like a doubly infinite line to which has been attached at each vertex a finite (possibly empty) tree. The vertices on the infinite line are precisely those vertices that are separators and by ergodicity and tail triviality this set has a density $D_{\text{sep}} > 0$ that is almost surely constant.

For any vertices v_1, v_2 and v_3, say that v_2 separates v_1 and v_3 if the unique path in $\mathbf{T}$ from v_1 to v_3 passes through v_2. Observe that if v_1, v_2 and v_3 are all on the infinite line in $\mathbf{T}$, then one of them separates the other two. Thus, for any v_1, v_2, v_3,

$$\sum_i \mathbf{P}(v_i \text{ separates the other two}) \geq \mathbf{P}(v_1, v_2, v_3 \text{ are all separators}).$$

Now triviality of the tail implies that μ is mixing of all orders, and, in particular, 2-mixing implies

$$\mathbf{P}(v_1, v_2, v_3 \text{ are all separators}) \to D_{\text{sep}}^3$$

as the pairwise distances $|v_i - v_j|$ all go to infinity. To get a contradiction, then, it suffices to show that

$$\mathbf{P}(x \text{ separates } v \text{ and } w) \to 0 \tag{10}$$

as the pairwise distances between v, w and x all go to infinity.

Assume then that the pairwise distances between the vertices v, w and x are at least L for some $L > 0$. Use the random walk scenario with $\mu = \lim \mu^n$, where μ^n is constructed as $\mu_2(B_n, v)$ for B_n large enough to contain v, w and x. Fix n for the moment and let γ be the initial segment of the random walk from v up to the first hitting of w. Here is how γ determines whether x separates v and w. If γ does not hit x, then x does not separate v from w. If γ does hit x, then let γ_1 be γ up to the first hitting of x and γ_2 be the rest of γ. The path connecting v and w in $\mathbf{T}$ is the path connecting them in $\mathbf{T}(\gamma)$, which is given by

$$LE(\gamma') = LE(\gamma_2' * \gamma_1') = LE(LE(\gamma_2') * \gamma_1').$$

Now x appears only once in $LE(\gamma_2') * \gamma_1'$, namely, at the point where they join. Thus x separates v and w if and only if it does not get erased when LE is applied to $LE(\gamma_2') * \gamma_1'$. If γ_1' is disjoint from $LE(\gamma_2)$ except at x, it is clear that the loop-erasure on $LE(\gamma_2') * \gamma_1'$ acts only on the γ_1' part and x never gets erased. Conversely, the first time that γ_1' intersects $LE(\gamma_2')$, the vertex x will be erased. Therefore, x is erased if and only if γ_1' and $LE(\gamma_2')$ are disjoint except at x. It remains to show that the probability of these paths being disjoint goes to zero as $n \to \infty$ and then $L \to \infty$.

For each M, the probability that $LE(\gamma_2') \wedge M$ and $\gamma_1' \wedge M$ are disjoint is an upper bound for the probability that $LE(\gamma_2')'$ and γ_1' are disjoint. Then to show (10) it suffices to show

$$\inf_M \lim_{L\to\infty} \lim_{n\to\infty} \mathbf{P}\big(LE(\gamma_2')' \wedge M \cap \gamma_1' \wedge M \neq \{x\}\big) = 0. \tag{11}$$

Now, by Lemma 3.4(ii), $LE(\gamma_2')'$ has the same distribution as $LE(\gamma_2)$. Combine the fact that γ_1 and γ_2 are independent with the fact from Lemma 3.3 that

$LE(\gamma_1) \wedge M$ converges to LERW $\wedge\, M$ and the fact that $\gamma_2 \wedge M$ converges to an independent SRW $\wedge\, M$ to rewrite (11) as

$$\inf_M \mathbf{P}(\text{LERW} \wedge M \cap \text{SRW} \wedge M \neq \{x\}) = 0,$$

where LERW and SRW are independent starting from x. This is a direct consequence of Lemma 3.1. Thus (11) and (10) are shown and the theorem is proved. □

Acknowledgment. All of the questions studied in this paper were asked by Russ Lyons.

REFERENCES

ALDOUS, D. (1988). The random walk construction of uniform spanning trees and uniform labelled trees. Technical Report, Dept. Statistics, Univ. California, Berkeley.

BRODER, A. (1988). Generating random spanning trees. Unpublished manuscript.

BURTON, R. and KEANE, M. (1989). Density and uniqueness in percolation. *Comm. Math. Phys.* **121** 501–505.

DOYLE, P. and SNELL, J. L. (1984). *Random Walks and Electric Networks. Carus Math. Monographs* **22**. Math. Assoc. Amer., Washington, D.C.

LAWLER, G. (1980). A self-avoiding random walk. *Duke Math J.* **47** 655–694.

LAWLER, G. (1983). A connective constant for loop-erased self-avoiding random walk. *J. Appl. Probab.* **20** 264–276.

LAWLER, G. (1986). Gaussian behavior of loop-erased self-avoiding random walk in four dimensions. *Duke Math J.* **53** 249–269.

LAWLER, G. (1988). Loop-erased self-avoiding walk in two and three dimensions. *J. Statist. Phys.* **50** 91–108.

LAWLER, G. (1991). *Intersections of Random Walks*. Birkhäuser, Boston. To appear.

DEPARTMENT OF MATHEMATICS
368 KIDDER HALL
OREGON STATE UNIVERSITY
CORVALLIS, OREGON 97331-4605