ANNALES DE L'I. H. P., SECTION B

HARRY KESTEN

Subdiffusive behavior of random walk on a random cluster

Annales de l'I. H. P., section B, tome 22, nº 4 (1986), p. 425-487
<http://www.numdam.org/item?id=AIHPB_1986__22_4_425_0>

Ann. Inst. Henri Poincaré,
Vol. 22, n° 4, 1986, p. 425-487.

Probabilités et Statistiques

Subdiffusive Behavior of Random Walk on a Random Cluster

by

Harry KESTEN [1]
Department of Mathematics, Cornell University,
Ithaca, N. Y. 14853

ABSTRACT. — We consider two cases of a random walk $\{X_n\}_{n\geq 0}$ on a random graph $\mathscr{G}$. *i*) In this case $\mathscr{G}$ is the family tree of a critical branching process, conditioned on no extinction. If $h(x)$ denotes the distance on $\mathscr{G}$ from x to the root of the tree, then the distribution of $n^{-1/3}h(X_n)$ converges. *ii*) In this case $\mathscr{G}$ is the « incipient infinite cluster » of two-dimensional bond percolation at criticality. Now we can merely prove that for some $\varepsilon > 0$ the family $\{ n^{-\frac{1}{2}+\varepsilon} X_n \}$ is tight.

Key-words and phrases: Anomalous diffusion, incipient infinite cluster, random walk, random walk on a random graph, subdiffusive behavior, Alexander-Orbach conjecture, backbone.

RÉSUMÉ. — Nous considérons deux cas de marche aléatoire $\{X_n\}_{n\geqq 0}$ sur un graphe aléatoire $\mathscr{G}$.

i) Dans le cas où $\mathscr{G}$ est l'arbre d'un processus de branchement critique, conditionné par la non-extinction, si $h(x)$ dénote la distance sur $\mathscr{G}$ de x à la racine de l'arbre, alors la distribution de $n^{-1/3}h(X_n)$ converge.

ii) Dans le cas où $\mathscr{G}$ est « l'amas infini d'un processus de percolation » de dimension 2 au seuil critique, nous prouvons seulement que pour un $\varepsilon > 0$ la famille $\{ n^{-\frac{1}{2}+\varepsilon} X_n \}$ est tendue.

[1] Research supported by the NSF through a grant to Cornell University.

1. INTRODUCTION

Let $\mathcal{G}$ be a graph in which all vertices have finite degree. Denote the degree of x by $d(x)$. By the *random walk on* $\mathcal{G}$ we mean the Markov chain with state space the vertex set of $\mathcal{G}$ and transition probabilities

$$(1.1) \qquad P(x, y) = \begin{cases} \dfrac{1}{d(x)} \text{ if } y \text{ is adjacent to } x \text{ on } \mathcal{G}, \\ 0 \quad \text{otherwise.} \end{cases}$$

We denote this random walk by $\{X_n^{\mathcal{G}}\}$. In the cases which we consider $\mathcal{G}$ will always be a subgraph of some fixed graph $\mathcal{H}$ and will contain a fixed vertex $\underline{0}$ of $\mathcal{H}$. We usually take $X_0^{\mathcal{G}} = \underline{0}$, and can then view $\{X_n^{\mathcal{G}}\}$ also as a Markov chain with state space $\mathcal{V}$, the vertex set of $\mathcal{H}$. In the sequel we shall choose $\mathcal{G}$ as a random subgraph of $\mathcal{H}$, according to some probability measure ν. It will always be the case that $\nu\{\underline{0} \in \mathcal{G}\} = 1$. We denote by $\{X_n\}_{n \geq 0}$ the random walk in random environment, for which $\mathcal{G}$ is first chosen according to ν, and once $\mathcal{G}$ is fixed, X_n starts at $X_0 = \underline{0}$ and moves on $\mathcal{G}$ as $\{X_n^{\mathcal{G}}\}$. Thus $\{X_n\}_{n \geq 0}$ is a (non-Markovian) process on

$$\Omega := \prod_0^\infty \mathcal{V}.$$

The probability measure governing $(\mathcal{G}, \{X_n\})$ will be denoted by $\mathbb{P}$, and expectation with respect to this measure by $\mathbb{E}$. $\mathbb{P}$ is given by

$$(1.2) \qquad \mathbb{P}\{\mathcal{G} \in G, \{X_n\} \in B\} = \int_G \nu(d\mathcal{G}) P\{\{X_n^{\mathcal{G}}\} \in B \mid X_n^{\mathcal{G}} = \underline{0}\},$$

for G and B measurable subsets of the collection of subgraphs of $\mathcal{H}$ and of Ω, respectively.

De Gennes [*4*] proposed the study of the asymptotic behavior of the random walk $\{X_n\}$ when $\mathcal{G}$ is a random cluster in a percolation system near criticality. He suggested that this should give insight into the conductivity properties of such clusters when viewed as random electrical networks. In spite of a seizable literature on this subject (see for instance [*1*] [*11*] [*18*] [*20*] [*22*] and some of their references) we do not know of a rigorously established relationship between the behavior of $X_n^{\mathcal{G}}$ for large n and the conductivity properties of $\mathcal{G}$ [2]. In the above named references (see also [*24*]) the asymptotic behavior of $\{ X_n \}$ has been studied by means of one or more of the following quantities:

(1.3) $$\mathbb{P}\{ X_n = \underline{0} \} \quad \text{for large } n.$$

(1.4) The distribution or moments of X_n itself, or of $h(X_n)$ for a suitable function $h(.)$ on $\mathcal{V}$. The former is appropriate when $\mathcal{G}$ is a subgraph of $\mathbb{Z}^d$ (with $\underline{0}$ = the origin), while we use the latter in case *i*) when $\mathcal{G}$ is a random tree and $\underline{0}$ its root. In the latter case we take $h(x)$ as the generation number of x; this is also called the « chemical distance » from x to the root. *Warning*: It is useful to think of the binary tree as being imbedded in infinite dimensional Euclidean space, with each edge between the k-th and $(k+1)$-th generation lying along the positive or negative k-th coordinate axis. This explains why many results for trees should also be true for percolation in sufficiently high dimension. Because of this, many physicists use the Euclidean distance $\| x \|$ on the tree. Then $\| x \| = \sqrt{k}$ for a k-th generation vertex x, for which $h(x) = k$. As a consequence many physics papers have $\bar{d} = 4$ and state a limit theorem for $n^{-1/6} \| X_n \|$ instead of our normalization in case *i*).

(1.5) The distribution of the exit time $\mathcal{T}$ be of the ball of radius k. More precisely, assume that we have some distance function $\rho(.,.)$ defined on $\mathcal{V}$. Let

$$\mathcal{B}(k) = \{ x \in \mathcal{V} : \rho(\underline{0}, x) \leq k \}.$$

Then

(1.6) $$\mathcal{T}(k) = \inf \{ n : X_n \notin \mathcal{B}(k) \}.$$

These aspects of the asymptotic behavior of X_n are of course closely related. For instance if $m(n)$ is such that $P\{ \rho(\underline{0}, X_n) \leq m(n) \}$ is close to 1/2, and the distribution of X_n is sufficiently smooth, then we expect [3]

(1.7) $$\mathbb{P}\{ X_n = \underline{0} \} = \int \nu(d\mathcal{G}) P\{ X_n^{\mathcal{G}} = \underline{0} \mid X_n^{\mathcal{G}} = \underline{0} \} \sim C_1 \{ \mathcal{N}(m(n)) \}^{-1},$$

---

(2) To be sure, there is the relationship pointed out by Straley [*27*] between the Green function of $X_n^{\mathcal{G}}$ and the resistance between $\underline{0}$ and ground in the *enhanced* network in which each edge of $\mathcal{G}$ incident to a vertex x has resistance $d(x)/t$ and each vertex x is connected to ground via a resistance of $(1-t)^{-1}$, for some $0 < t < 1$. The resistance between $\underline{0}$ and ground in this network equals $\Sigma_0^\infty t^n P\{ X_n^{\mathcal{G}} = \underline{0} \mid X_0^{\mathcal{G}} = \underline{0} \}$.

(3) C_i denotes a strictly positive and finite constant, whose specific value is of no significance for our considerations. Its value may differ at different appearances.

where

$\mathcal{N}(k) := \mathbb{E}\ \#\ (\mathscr{B}(k) \cap \mathscr{G})$ = expectation of the number of vertices in $\mathscr{B}(k) \cap \mathscr{G}$.

Of course we also have the obvious relation

$$\{ \mathscr{T}(k) > n \} \subset \{ \rho(\underline{0}, X_n) \leq k \} \tag{1.8}$$

with the corresponding relation between the probabilities of these events.

Alexander and Orbach [1] conjectured another relationship between the behavior of X_n and the structure of $\mathscr{G}$ for a not precisely described class of (random) graphs $\mathscr{G}$. They consider graphs for which in some (average) sense

$$\mathcal{N}(k) \sim C_2 k^{\bar{d}}, \qquad k \to \infty, \tag{1.9}$$

for some constant $\bar{d}$. $\bar{d}$ is sometimes called the fractal dimension; it is widely believed that the incipient infinite cluster of percolation at criticality satisfies (1.9). It is further assumed that for some $\theta > 0$

$$\mathbb{E}\rho^2(\underline{0}, X_n) \sim C_3 n^{2/(2+\theta)} \tag{1.10}$$

or (using (1.7), (1.9))

$$\rho\{ X_n = \underline{0} \} \sim C_4 n^{-\bar{d}/(2+\theta)}. \tag{1.11}$$

The Alexander-Orbach conjecture states that

$$\frac{\bar{d}}{2+\theta} = \frac{2}{3}. \tag{1.12}$$

It is unclear for what kind of graphs $\mathscr{G}$ this should hold. Heuristic arguments have been given in [18] and [24] for (1.12) when $\mathscr{G}$ is the family tree of a critical branching process conditioned on nonextinction. When the branching process satisfies (1.13) and (1.17) below, then $\bar{d}=2$ (see (1.22)) and our results (1.16) and (1.19) confirm (1.11) and (1.12) in this case (see Remark 1.22). It should be noted though that the *Alexander-Orbach conjecture is false even on some trees*, namely the family trees of branching processes with infinite variance considered in Theorem 1.21 (again see Remark 1.22). For the second case mentioned in the abstract, when $\mathscr{G}$ is the incipient infinite cluster of two-dimensional bond percolation we only obtain in Theorem 1.27 that $\{ n^{\frac{1}{2}-\varepsilon} X_n \}$ is tight. This merely shows that X_n has subdiffusive behavior. If (1.10) holds, then θ will indeed be strictly positive.

In both cases, the result is obtained by analyzing the imbedded random walk on a subgraph of $\mathcal{G}$ which is called the *backbone*. Attached to the backbone are « dangling ends » which are of no help for X_n to get far out in the graph $\mathcal{G}$. If X enters a dangling end it has to return to the backbone in order to go to infinity. The time spent in the dangling ends is responsible for $|X_n|$ growing slower than $n^{1/2}$ in the second case. In the first case we can analyze the time spent in the dangling ends more precisely to obtain the result that X_n behaves like a time changed version of the imbedded random walk on the backbone. The imbedded random walk itself is very much like a simple one-dimensional random walk about which we can obtain enough information to derive our results (cf. proof of Theorem 1.19).

To state our results precisely we now describe the measure ν governing the choice of $\mathcal{G}$. For the branching process in case *i*) we consider a critical Bienaymé-Galton-Watson (4) branching process $\{Z_n\}$, starting with one individual in the zeroth generation ($Z_0 = 1$) and non-degenerate offspring distribution with

$$\mu = E\{Z_1 \mid Z_0 = 1\} = 1. \tag{1.13}$$

We denote by $T_{[n]}(T)$ the family tree of the generations $0, 1, \ldots, n$ (all generations) of this process. (See [*10*], Ch. VI.2 or [*13*], Ch. 1.2 for family trees.) We view $T_{[n]}$ and T as rooted labelled trees (5). The root corresponds to the particle in generation 0, the progenitor of the process, and it is denoted by $\langle 0 \rangle$. A generic particle of the k-th generation is indexed as $\langle 0, l_1, \ldots, l_k \rangle$, $l_r \geq 1$, $1 \leq r \leq k$. The particles $\langle 0, l_1, \ldots, l_{k-1}, j \rangle$, $j = 0, 1, \ldots$ denote the children of the particle $\langle 0, l_1, \ldots, l_{k-1} \rangle$. Of course, not for all j does $\langle 0, l_1, \ldots, l_{k-1}, j \rangle$ correspond to an actual vertex of T. If $\langle 0, l_1, \ldots, l_{k-1}, j \rangle$ is a vertex of T, let $N(0, l_1, \ldots, l_{k-1})$ be the number of children of $\langle 0, l_1, \ldots, l_{k-1} \rangle$ in the branching process. Then $\langle 0, l_1, \ldots, l_{k-1}, j \rangle$ is a vertex of T for $1 \leq j \leq N(0, l_1, \ldots, l_{k-1})$. $\langle 0, l_1, \ldots, l_{k-1}, j \rangle$ is called the j-th child of $\langle 0, l_1, \ldots, l_{k-1} \rangle$. Z_k is the total number of vertices in the k-th generation, i. e.

$$Z_k = \sum_{\langle 0, l_1, \ldots, l_{k-1} \rangle \in T} N(0, l_1, \ldots, l_{k-1}).$$

---

(4) The more traditional name for these processes is merely « Galton-Watson processes » (see [*2*], [*10*]).

(5) Viewing the family tree as labelled is a bit of a nuisance, as will become apparent in the sequel. Nevertheless in our view, it makes the proofs of the Lemmas 2.2, 2.10 and 2.14 clearer.

T is finite if and only if $Z_k = 0$ from some k on. The subtree having as vertices $\langle 0, l_1, \ldots, l_k \rangle$ and all its descendants will be denoted by $T(0, l_1, \ldots, l_k)$. This too is a rooted labelled tree, with root $\langle 0, l_1, \ldots, l_k \rangle$. If t is a rooted labelled tree and t' is obtained by permuting the subtrees

$$t(0, l_1, \ldots, l_{k-1}, j), \qquad j=1, \ldots, N(0, l_1, \ldots, l_{k-1}),$$

then t and t' are isomorphic as abstract graphs. E. g. the two trees in Figure 1 are isomorphic.

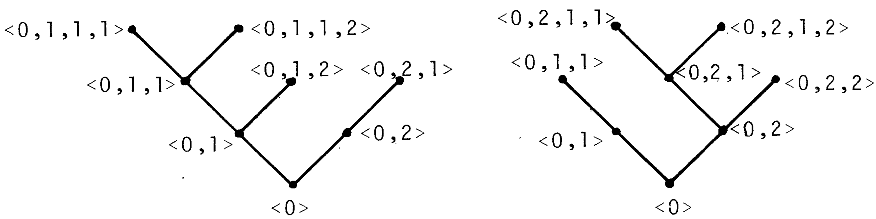

FIG. 1. — Two isomorphic trees which we view as distinct labelled trees.

However t and t' are *not* the same as rooted labelled trees. Two such trees are identified only if the offspring numbers $N(0, l_1, \ldots, l_k)$ and $N'(0, l_1, \ldots, l_k)$ for these two trees are the same for all k and l_j.

If t is a rooted labelled tree we write t_k for the collection of vertices in its k-th generation, and $\# t_k$ for the number of such vertices. (Thus $Z_k = \# T_k$.) We define the *height* $h(x)$ of a vertex x of t by its generation number, i. e. $h(\langle 0, l_1, \ldots, l_k \rangle) = k$. We say that t is a *tree of k generations* if $\# t_k \neq 0$ but $\# t_{k+1} = 0$. The following lemma is immediate from $P\{Z_n \neq 0\} \to 0$ and Kolmogorov's consistency criterion ([*19*], Part I, Section 4.3) (see also the proof of 2.6).

(1.14) LEMMA. — *If (1.13) holds, then for any rooted labelled tree t of k generations*

$$\lim_{n\to\infty} P\{T_{[k]} = t \mid Z_n \neq 0\} = (\# t_k).P\{T_{[k]} = t\}.$$

If we set

$$(1.15) \qquad \nu\{T_{[k]} = t\} = (\# t_k).P\{T_{[k]} = t\}$$

then ν *has a unique extension to a probability measure on the rooted labelled infinite trees.*

We interpret the measure ν of this lemma as the distribution of the family tree of the branching process $\{Z_n\}$ conditioned on no extinction.

Note that we need such a roundabout definition of this concept because extinction occurs w. p. 1 in a non-degenerate critical branching process ([10], Theorem I.6.1). We can think of ν as a measure on the subtrees of the tree $\mathscr{H}$, which has the vertex set $\langle 0, l_1, \ldots, l_k \rangle$, $k = 0, 1, \ldots, l_j = 1, \ldots$ and edges only between pairs $\langle 0, l_1, \ldots, l_k \rangle$ and $\langle 0, l_1, \ldots, l_{k+1} \rangle$. The special vertex $\underline{0}$ is taken as $\langle 0 \rangle$ and $\rho(\langle x \rangle, \langle y \rangle)$ as the number of edges in the shortest path on $\mathscr{H}$ from $\langle x \rangle$ to $\langle y \rangle$. In particular $\rho(\underline{0}, \langle x \rangle) = h(x)$. Our result for case *i*) of the abstract is as follows.

(1.16) THEOREM. — *Assume that (1.13) holds as well as*

$$0 < \sigma^2 := \sigma^2 \{ Z_1 \mid Z_0 = 1 \} < \infty . \tag{1.17}$$

Let $\mathscr{G}$ be a rooted labelled tree chosen with the distribution ν of Lemma (1.14), and let X_n be the random walk on $\mathscr{G}$ governed by $\mathbb{P}$ as defined in (1.2). Then under $\mathbb{P}$ the process $\{ Y_n(t) \}_{t \geq 0}$ defined by

$$Y_n(t) = \frac{1}{n^{1/3}} \{ (nt - k)h(X_{k+1}) + (k + 1 - nt)h(X_k) \} \quad \text{for} \quad \frac{k}{n} \leq t \leq \frac{k+1}{n}$$

converges weakly in $C[0, \infty)$ to a process $\{ Y(t) \}_{t \geq 0}$ which is not the zero process.

(1.18) REMARK. — *Y_n is the scaled linear interpolation of $\{ X_k \}_{k \geq 0}$.* In particular the theorem states that $n^{-1/3}h(X_n)$ converges in distribution, and the limit distribution is not concentrated at 0. In fact one can show that it has no mass at 0 at all. One can also show that $\{ Y(t) \}$ visits 0 infinitely often. Note that we only state convergence of the $\mathbb{P}$-probabilities. We do not make any statement about the distribution of $\{ X_n^{\mathscr{G}} \}$ for any fixed realization G of the random graph $\mathscr{G}$.

At this time we only have a monstrously long proof of (1.16) and we therefore restrict ourselves to the following weaker result, which sufficiently illustrates most of the interesting aspects of the problem. In particular it still confirms the Alexander-Orbach conjecture under (1.13) and (1.17). Copies of the proof of (1.16) are available from the author.

(1.19) THEOREM. — *Assume (1.13) and (1.17) and let $\mathscr{G}$ and X_n be as in Theorem 1.16. Let*

$$\mathscr{T}(m) = \inf \{ n : h(X_n) = m \} .$$

Then for all $\varepsilon > 0$ there exist constants $0 < x_1(\varepsilon), x_2(\varepsilon) < \infty$ such that

$$\mathbb{P} \{ x_1 \leq m^{-3}\mathscr{T}(m) \leq x_2 \} \geq 1 - \varepsilon, \qquad m \geq 1 .$$

The normalization for $\mathcal{T}(m)$ changes when one drops the variance assumption (1.17). Specifically, the following result can be proven along the lines of Theorem 1.19. We shall not give the proof, but only point out that most of the branching process estimates needed to replace Lemma 2.41 below can be found in [26].

(1.20) THEOREM. — *Let $\mathcal{G}$ and X_n be as in Theorem 1.16 and $\mathcal{T}$ as in Theorem 1.19. If (1.13) holds and for some* $0 < \alpha < 1$

$$P\{ Z_1 \geq k \mid Z_0 = 1 \} \sim C_1 k^{-1-\alpha}, \qquad k \to \infty, \tag{1.21}$$

then for all $\varepsilon > 0$ *there exist constants* $0 < x_1(\varepsilon)$, $x_2(\varepsilon) < \infty$ *such that*

$$\mathbb{P}\{ x_1 < m^{-2-\frac{1}{\alpha}} \mathcal{T}(m) \leq x_2 \} \geq 1 - \varepsilon, \qquad m \geq 1.$$

(1.22) REMARK. — [23] Theorem 5, shows that under (1.13) and (1.17)

$$m^{-2} \# \mathcal{G}_{[m]} = m^{-2} \sum_{k=0}^{m} Z_k$$

has a limit distribution (under ν) when $m \to \infty$. Thus $\bar{d} = 2$ in this case, while Theorems 1.16 and 1.19 can be interpreted as saying that in (1.10) $\theta = 1$, which agrees with (1.12). However, if (1.17) is replaced by (1.21), then one can deduce from the Corollary and Lemma 2 in [26], that

$$m^{-1-\frac{1}{\alpha}} \# \mathcal{G}_{[m]} = m^{-1-\frac{1}{\alpha}} \sum_{k=0}^{m} Z_k$$

has a limit distribution (under ν), so that $\bar{d} = 1 + \alpha^{-1}$. On the other hand Theorem 1.20 says that $\theta = 1/\alpha$, in the sense that at time $\mathcal{T}(m)$—which is of order $m^{2+1/\alpha}$ — $h(X_{\mathcal{T}(m)}) = m$ or $h(X_t)$ is of order $t^{\frac{1}{2+1/\alpha}}$ for $t = \mathcal{T}(m)$. Of course Theorem 1.20 also implies

$$\mathbb{P}\{ h(X_{x_1 m^{2+1/\alpha}}) \leq m \} \geq 1 - \varepsilon,$$

as explained at (1.8). Thus (1.12) does *not* hold in the situation of Theorem 1.20.

(1.23) REMARK. — We have chosen $\mathcal{G}$ according to the measure ν of Lemma 1.14. We can interpret this as conditioning on no extinction of the branching process, ever. For Theorems (1.19) and (1.20) one could equally well have conditioned on $\{ Z_m > 0 \}$. This is enough to guarantee that the family tree has m generations so that $\mathcal{T}(m)$ is finite. Such a change in conditioning would not have influenced the results. ///

We turn to the percolation case *ii*). The analogue of Lemma 1.14 for this case is much harder, but was derived in [*16*]. We quote the result for bond percolation on $\mathbb{Z}^2$, using the following notation: $\mathcal{E}$ = edge set of $\mathbb{Z}^2$, $\Omega^* = \{0, 1\}^{\mathcal{E}}$ and P_p = probability measure on Ω^* which assigns to each edge, independently of all other edges, a 1 (respectively 0) with probability p respectively $q := 1 - p$). An edge e with a 1 (0) assigned to it is called *open* (respectively *closed*). A path on $\mathbb{Z}^2$ is called open if all its edges are open. An *open cluster* is a maximal connected set of open edges. The critical probability for this model equals 1/2, i. e. infinite open clusters exist if and only if $p > 1/2$ (cf. [*15*], Application 3.4 *ii*)).

(1.24) LEMMA. — *For each event* E *which depends on finitely many edges only, the following two limits exist and are equal*:

$$(1.25)\quad \lim_{n\to\infty} P_{1/2}\{\, E \mid \text{there exists an open path from } \underline{0} \text{ to } \mathbb{R}^2/[-n, n]^2 \,\},$$

$$(1.26)\quad \lim_{p\downarrow 1/2} P_p\{\, E \mid \text{there exists an infinite open path starting at } \underline{0} \,\}.$$

If we denote the common value of (1.25) and (1.26) by $\nu(E)$, *then* $\nu(.)$ *extends uniquely to a probability measure on* Ω^* *(which we again denote by* ν*).* A. e. $[\nu]$ $\underline{0}$ *belongs to an infinite open cluster and there is no other infinite open cluster.*

We now consider a random configuration of open and closed edges, chosen according to the measure ν, and denote by $\mathcal{G}$ the random subgraph of $\mathbb{Z}^2$ consisting of the open edges and their endpoints only. The component of $\underline{0}$ in $\mathcal{G}$ is denoted by $\tilde{W}$. By the lemma $\tilde{W}$ is a. e. $[\nu]$ the unique infinite open cluster. We proposed in [*16*] that $\tilde{W}$ be considered the « incipient infinite cluster ». We further consider the random walks $X_n^{\mathcal{G}}$ and X_n as in (1.2). Note that since X_n starts at $\underline{0}$, it can never leave $\tilde{W}$, so that X_n is a random walk on the incipient infinite cluster $\tilde{W}$.

(1.27) THEOREM. — *There exists an* $\varepsilon > 0$ *such that for* X_n, *the random walk on the incipient infinite cluster of two-dimensional bond percolation defined above, the family* $\{ n^{-\frac{1}{2}+\varepsilon} X_n \}_{n\geq 1}$ *is tight.*

ACKNOWLEDGEMENT

The author wishes to thank R. Durrett for suggesting that Theorem (1.19) should be true.

2. PROOF OF THEOREM 1.19

Throughout this section (1.13) and (1.17) will be assumed.

« Tree » will always mean « rooted labelled tree ». T will denote a random tree with the distribution of the family tree of the original Bienaymé-Galton-Watson process. The offspring distribution for this process is determined by

(2.1) $$p_j := \mathrm{P}\{ Z_1 = j \mid Z_0 = 1 \}, \qquad j \geq 0 .$$

$\tilde{\mathrm{T}}$ will be used to denote a random tree with the distribution ν of Lemma 1.14. If $\langle 0, l_1, \ldots, l_k \rangle$ is a k-th generation vertex of a tree t, then we denote by $\zeta(0, l_1, \ldots, l_k; t)$ its *lifetime*, i. e., the maximal $n \geq 0$ for which there exists a vertex $\langle 0, l_1, \ldots, l_k, l_{k+1}, \ldots, l_{k+n} \rangle$ in $t(0, l_1, \ldots, l_k)$. $\zeta(0; t)$ will be abbreviated to $\zeta(t)$, and if no confusion is likely the t may be dropped from these symbols. The *m-backbone* of a tree t is the subtree of t containing exactly those vertices $\langle x \rangle$ of t with $\zeta(x; t) \geq m$. This backbone is denoted by $b(m; t)$; for simplicity we write $\tilde{\mathrm{B}}(m)$ for the backbone $b(m; \tilde{\mathrm{T}})$ of $\tilde{\mathrm{T}}$.

We shall first derive some facts about the distribution of $\tilde{\mathrm{T}}$, in particular its relations with the distribution of T. Next we describe the distribution of $\tilde{\mathrm{B}}$, and the conditional distribution of $\tilde{\mathrm{T}}$, given $\tilde{\mathrm{B}}$. We can then start on the main step which is a comparison of $\{ X_n^{\tilde{\mathrm{T}}} \}$ with its embedded random walk $\{ U_k \} = \{ U_k(m; \tilde{\mathrm{T}}) \}$ on the backbone $\tilde{\mathrm{B}}(m)$. This will require some branching process estimates in order to estimate the time spent in $\tilde{\mathrm{T}} \backslash \tilde{\mathrm{B}}(m)$. After that we obtain our principal result by estimating how much actual time it takes for a certain (random) number of steps of the imbedded walk.

(2.2) LEMMA. — A. e. $[\nu]$ $\tilde{\mathrm{T}}$ *has a single infinite line of descent. i. e., there exist unique integers* $i_1, i_2, \ldots,$ *such that* $\langle 0, i_1, \ldots, i_n \rangle \in \tilde{\mathrm{T}}$ *for all* n, *and hence* $\zeta(0, i_1, \ldots, i_n; \tilde{\mathrm{T}}) = \infty$, *while* $\zeta(0, j_1, \ldots, j_n; \tilde{\mathrm{T}}) < \infty$ *for all* $(j_1, \ldots, j_n) \neq (i_1, \ldots, i_n)$, $n \geq 0$.

The joint distribution of the i_n *and* $\mathrm{N}(0, i_1, \ldots, i_n)$, $n \geq 0$ *is given by*

(2.3) $$\begin{aligned} &\nu\{ \mathrm{N}(0) = l_0 + 1, \quad i_1 = \alpha_1, \\ &\mathrm{N}(0, \alpha_1) = l_1 + 1, \ i_2 = \alpha_2, \ldots, \mathrm{N}(0, \alpha_1, \ldots, \alpha_{k-1}) = l_{k-1} + 1, \ i_k = \alpha_k \} \\ &= \prod_{j=0}^{k-1} p_{l_j + 1}, \qquad l_j \geq 0, \qquad 1 \leq \alpha_j \leq l_{j-1} + 1 . \end{aligned}$$

Finally, conditionally on $i_1, i_2, \ldots,$ *and* $N(0, i, \ldots, i_n)$, $n \geq 0$, *the subtrees* $\tilde{T}(0, i_1, \ldots, i_l, j)$ *with* $l \geq 0$, $1 \leq j \leq N(0, i_1, \ldots, i_l)$, $j \neq i_{l+1}$ *are independent copies of* T.

Proof. — Let t be a tree of k generations (i. e., $\zeta(t) = k$). Then by the definition of v, for any $\langle 0, i_1, \ldots, i_k \rangle \in t$ we have

$$v\{\tilde{T}_{[k]} = t \text{ and } \zeta(0, i_1, \ldots, i_k; \tilde{T}) = \infty\} = \lim_{r \to \infty} v\{\tilde{T}_{[k]} = t, \zeta(0, i_1, \ldots, i_k; \tilde{T}) \geq r\}$$

and

$$(2.4) \quad v\{\tilde{T}_{[k]} = t, \zeta(0, i_0, \ldots, i_k; \tilde{T}) \geq r\} = \lim_{s \to \infty} \frac{1}{P\{\zeta(T) \geq s\}} P\{T_{[k]} = t, \zeta(0, i_0, \ldots, i_k; T) \geq r \text{ and } \zeta(0, j_0, \ldots, j_k; T) \geq s - k \text{ for some } \langle 0, j_0, \ldots, j_k \rangle \in T_{[k]}\}.$$

Finally, by the branching property we may for $s - k \geq r$ write the last probability as the sum of

$$P\{T_{[k]} = t, \zeta(0, i_1, \ldots, i_k; T) \geq s - k\} = P\{T_{[k]} = t\}.P\{\zeta(T) \geq s - k\}$$

and

$$P\{T_{[k]} = t, r \leq \zeta(0, i_1, \ldots, i_k; T) < s - k$$

but for some $\langle 0, j_1, \ldots, j_k \rangle \in T_{[k]}$ $\zeta(0, j_1, \ldots, j_k; T) \geq s - k\}$

$$\leq P\{T_{[k]} = t\}.P\{\zeta(T) \geq r\}.(\# t_k)P\{\zeta(T) \geq s - k\}.$$

Since by Theorem I.9.1 of [2]

$$(2.5) \qquad P\{\zeta(T) \geq s\} \sim \frac{2}{s\sigma^2},$$

it is easy to see from this that the right hand side of (2.4) equals

$$\lim_{r \to \infty} P\{T_{[k]} = t\}\left(1 + 0\left(\frac{\# t_k}{r\sigma^2}\right)\right) = P\{T_{[k]} = t\},$$

so that

$$(2.6) \quad v\{\tilde{T}_{[k]} = t \quad \text{and} \quad \zeta(0, i_1, \ldots, i_k; \tilde{T}) = \infty\} = P\{T_{[k]} = t\}.$$

Similar estimates show that

$$(2.7) \quad v\{\tilde{T}_{[k]} = t \text{ and } \zeta(0, i_1, \ldots, i_k; \tilde{T}) = \zeta(0, j_1, \ldots, j_k; \tilde{T}) = \infty \text{ for two distinct } \langle 0, i_1, \ldots, i_k \rangle \text{ and } \langle 0, j_1, \ldots, j_k \rangle \text{ in } \tilde{T}_k\} = 0.$$

(2.7) shows that in each generation of $\tilde{T}$ there is at most one vertex with infinite lifetime. On the other hand, a special case of (2.6) says $\nu\{\zeta(0;\tilde{T})=\infty\}=1$, so that each generation of $\tilde{T}$ must contain at least one vertex with infinite lifetime. The first part of the lemma is immediate from this.

(2.3) is obtained from (2.6) by summing over all t which contain $\langle 0, i_1, \dots, i_k \rangle$ and satisfy $N(0, i_1, \dots, i_j) = l_j + 1$, $0 \le j \le k$.

The last statement of the lemma, giving the fact that the $\tilde{T}(0, i_1, \dots, i_l, j)$ are conditionally independent with the distribution of T, also follows from (2.6) and the branching property. For example, if $t'_1, \dots, t'_{l_0+1}$ are fixed trees with $\zeta(t'_i) \le k-1$ then

$$(2.8) \quad \nu\{N(0, i_1, \dots, i_j) = l_j + 1, i_{j+1} = \alpha_j, 0 \le j \le k-1, \\ \tilde{T}(0, r) = t'_r, 1 \le r \le l_0 + 1, r \ne \alpha_1\} = \Sigma' P\{T_{[k]} = t\},$$

where Σ' is the sum over all t of k generations with $N(0) = l_0 + 1$ and $T(0, r)=t'_r$, $1 \le r \le l_0+1$, $r \ne \alpha_1$, $N(0, \alpha_1, \dots, \alpha_j)=l_j+1$, $1 \le j \le k-1$. By the branching property (2.8) equals

$$p_{l_0+1} \prod_{\substack{r=1 \\ r\ne\alpha_1}}^{l_0+1} P\{T = t'_r\} \prod_{j=1}^{k-1} p_{l_k+1}. \qquad \square$$

(2.9) COROLLARY. — *The numbers of children in $\tilde{T}$ of the vertices on the infinite line of descent are i. i. d with distribution*

$$\nu\{N(x) = l + 1\} = (l + 1)p_{l+1}, \qquad l \ge 0,$$

for any $\langle x \rangle$ on the infinite line of descent.

Proof. — Sum (2.3) over α_j, $1 \le \alpha_j \le l_{j-1} + 1$. $\square$

We turn to the description of $\tilde{B}(m) = b(m;\tilde{T})$, the *m-backbone* of $\tilde{T}$. Clearly $\tilde{B}(m)$ is a subtree of $\tilde{T}$ which contains the unique infinite line of descent of $\tilde{T}$ as described in Lemma 2.2. We again view $\tilde{B}(m)$ as a rooted labelled tree. The root is $\langle 0 \rangle$, the same as the root of $\tilde{T}$, There is, however, a complication due to the fact that the l-th child in $\tilde{B}$ of a certain vertex may be the l'-th child (with $l' > l$) of that same vertex in $\tilde{T}$. This is because $\tilde{T}$ has more vertices than $\tilde{B}$. In some arguments it is necessary to distinguish the labeling of the vertices in $\tilde{B}$ and of those same vertices in $\tilde{T}$. Whenever it seems necessary to explicitly indicate that we are using the labels in $\tilde{B}(m)$ we shall use the notation $\langle 0, l_1, \dots, l_k; m \rangle$ for the label in $\tilde{B}$. This label is attached to the l_k-th child in $\tilde{B}$ of the vertex $\langle 0, l_1, \ldots, l_{k-1}; m \rangle \in \tilde{B}$. Of course $\langle 0, l_1, \ldots, l_k; m \rangle$ has to equal some vertex $\langle 0, \alpha_1, \ldots, \alpha_k \rangle$ of the k-th generation of $\tilde{T}$. We shall order the children in $\tilde{B}$ of a given vertex in the same way as they are ordered in $\tilde{T}$. That is, if $\langle 0, l_1, \ldots, l_k; m \rangle = \langle 0, \alpha_1, \ldots, \alpha_k \rangle$ and $\langle 0, \alpha_1, \ldots, \alpha_k, r \rangle$ belongs to $\tilde{B}$ exactly for the indices $r_1 < r_2 < \ldots < r_s$, then $\langle 0, \alpha_1, \ldots, \alpha_k, r_t \rangle$ will be the t-th child of $\langle 0, l_1, \ldots, l_k; m \rangle$ in $\tilde{B}$ and hence have the label $\langle 0, l_1, \ldots, l_k, t; m \rangle$ in $\tilde{B}$.

We also need the following notation

$$\tilde{N}(0, l_1, \ldots, l_k; m) = \text{number of children in } \tilde{B}(m) \text{ of } \langle 0, l_1, \ldots, l_k; m \rangle,$$

$$\tilde{B}(0, l_1, \ldots, l_k; m) = \text{subtree of } \tilde{B}(m) \text{ consisting of } \langle 0, l_1, \ldots, l_k; m \rangle \text{ and its descendants in } \tilde{B}(m),$$

$$\tilde{B}_{[k]}(m) = \text{subtree of first } k \text{ generations of } \tilde{B}(m),$$

$$\tilde{\zeta}(0, l_1, \ldots, l_k; m) = \text{lifetime in } \tilde{B}(m) \quad \text{of} \quad \langle 0, l_1, \ldots, l_k; m \rangle.$$

(2.10) LEMMA. — A. e. $[\nu]$ $\tilde{B}(m)$ *contains a single infinite line of descent, i. e., there exist unique integers* $j_1, j_2, \ldots,$ *such that* $\tilde{\zeta}(0, j_1, \ldots, j; m) = \infty$ *for all* k. *The joint distribution of the* j_n *and* $\tilde{N}(0, j_1, \ldots, j_n; m)$, $n > 0$, *is given by*

$$(2.11) \quad \nu\{ \tilde{N}(0; m) = l_0 + 1, \ j_1 = \beta_1, \ \tilde{N}(0, \beta_1; m) = l_1 + 1, \\ j_2 = \beta_2, \ldots, \tilde{N}(0, \beta_1, \ldots, \beta_k; m) = l_k + 1, \ j_{k+1} = \beta_{k+1} \} = \prod_{j=0}^{k} q_{l_j+1}(m)$$

for $l_j \geq 0$, $1 \leq \beta_j \leq l_{j-1} + 1$, *where*

$$(2.12) \quad q_r(m) = \sum_{s \geq r} p_s \binom{s}{r} [P\{ \zeta(T) \geq m \}]^{r-1} . [P\{ \zeta(T) < m \}]^{s-r}.$$

Conditionally on $j_1, j_2, \ldots,$ *and* $\tilde{N}(0, j_1, \ldots, j_n; m)$, $n \geq 0$, *the subtrees* $\tilde{B}(0, j_1, \ldots, j_k, j; m)$ *with* $k \geq 0$, $1 \leq j \leq \tilde{N}(0, j_1, \ldots, j_k; m)$, $j \neq j_{k+1}$, *are i. i. d., each with distribution of* $b(m; T)$ *given* $\zeta(t) \geq m$.

Proof. — We do not give a detailed proof of this lemma. It is essentially a consequence of (2.2) or (2.6) if we take into account that a vertex $\langle x \rangle$

of $\tilde{T}$ belongs to $\tilde{B}(m)$ if and only if $\zeta(x) \geq m$. Thus, for instance, if $b_1, \ldots, b_{l_0+1}$ are trees with lifetime $\geq m$, then

$$\nu\{\tilde{N}(0;m)=l_0+1,\ j_1=\beta_1 \quad \text{and} \quad \tilde{B}(0,j;m)=(b_j, 1\leq j\leq l_0+1, j\neq j_1\}$$

$$= \sum_{\substack{s\geq l_0+1 \\ 1\leq i_1<\ldots<i_{l_0+1}\leq s}} \nu\{N(0)=s,\ \zeta(0,i)\geq m \text{ exactly when } i\in\{i_1,\ldots,i_{l_0+1}\},$$

$$\zeta(0,i_{j_1})=\infty,\ B(0,i_r;m,\tilde{T})=b_r,\ 1\leq r\leq l_0+1,\ r\neq j_1\},$$

where $B(x;m,\tilde{T})$ is the m-backbone of $\tilde{T}(x)$. It is easy to see from (2.6) that the right hand side of this equation equals

$$\sum_{\substack{s\geq l_0+1 \\ 1\leq i_1<\ldots<i_{l_0+1}\leq s}} p_s[P\{\zeta(T)\geq m\}]^{l_0}$$

$$.\ [P\{\zeta(T)<m\}]^{s-l_0-1}\prod_{\substack{j=1 \\ j\neq j_1}}^{l_0+1} P\{b(m;T)=b_j\mid\zeta(T)\geq m\}$$

$$= q_{l_0+1}\prod_{\substack{j=1 \\ j\neq j_1}}^{l_0+1} P\{b(m;T)=b_j\mid\zeta(T)\geq m\}.$$

(2.13) COROLLARY. — *The numbers of children in $\tilde{B}(m)$ of the vertices on the infinite line of descent are i. i. d. with distribution*

$$\nu\{\tilde{N}(x;m)=l+1\}=(l+1)q_{l+1}$$

for any $\langle x;m\rangle$ on the infinite line of descent.

We need some additional concepts to describe the conditional distribution of $\tilde{T}$, given $\tilde{B}(m)$. We call a vertex $\langle x\rangle$ of a tree t a *branchpoint* (of order $p\geq 2$) if $\langle x\rangle$ has at least two children (exactly p children) in t. Note that between branchpoints a tree looks like a segment of $\mathbb{Z}$, since each vertex which is not a branchpoint has at most one child. We call $\langle x\rangle$ an *endpoint* of t if $\langle x\rangle$ has no children in t (such vertices are often called « leaves »). The degree of a branchpoint of order p (other than the root) equals $p+1$, while the degree of an endpoint equals 1.

Some more notation:

$M(y;m)$ = number of children in $\tilde{T}\backslash\tilde{B}(m)$ of a vertex $\langle y\rangle$ of $\tilde{B}(m)$.

(2.14) LEMMA. — *Given* $\tilde{B}(m)$, *the family trees of the offspring in* $\tilde{T}\setminus\tilde{B}(m)$ *of all the vertices* $\langle y \rangle$ *in* $\tilde{B}(m)$ *are conditionally independent with the following distributions*

(2.15) *If* $\langle y \rangle \in \tilde{B}(m)$ *is not an endpoint of* $\tilde{B}(m)$ *then*

$$v\{ M(y;m) = k \} = p_{k+\tilde{N}(y;m)}\binom{k+\tilde{N}(y;m)}{k}[P\{\zeta(T) < m-1\}]^k \cdot \left[\sum_{l\geq 0} p_{l+\tilde{N}(y;m)}\binom{l+\tilde{N}(y;m)}{l}[P\{\zeta(T) < m-1\}]^l\right]^{-1}.$$

Given $M(y,m)$ *the subtrees* $\tilde{T}(y,j)$ *for each of the* $M(y;m)$ *children of* $\langle y \rangle$ *in* $\tilde{T}\setminus\tilde{B}(m)$ *are conditionally independent, each with the conditional distribution of* T, *given* $\zeta(T) < m-1$.

(2.16) *If* $\langle y \rangle$ *is an endpoint of* $\tilde{B}(m)$, *then* $\tilde{T}(y)$ *has the conditional distribution of* T, *given* $\zeta(T) = m$.

Proof. — Again we do not give a detailed proof of this lemma; it is a direct consequence of Lemma 2.2. We merely point out that, for an endpoint $\langle y \rangle$ of $\tilde{B}(m)$ we must have

$$\zeta(y;\tilde{T}) = m. \tag{2.17}$$

Indeed $\zeta(y;\tilde{T}) \geq m$ because $\langle y \rangle \in \tilde{B}$, while $\zeta(y;\tilde{T}) > m$ is excluded by the fact that $\langle y \rangle$ has no children in $\tilde{B}(m)$. (2.16) is immediate from (2.17). The situation is slightly different if $\langle y \rangle$ is not an endpoint at $\tilde{B}(m)$ and hence $\tilde{N}(y;m) \geq 1$. Then the condition $\langle y \rangle \in \tilde{B}(m)$ merely says that all the children of $\langle y \rangle$ in $\tilde{T}\setminus B(\tilde{m})$ have lifetimes $< m-1$. The condition $\zeta(y) \geq m$—which has to be satisfied because $\langle y \rangle \in \tilde{B}(m)$—is taken care of through the $\tilde{N}(y;m)$ children of $\langle y \rangle$ in $\tilde{B}(m)$. This leads to (2.15). □

We now prepare our first result for the imbedded random walk of $\{X_n^T\}$ on $\tilde{B}(m)$. First its definition. Let t be a fixed rooted labelled tree, and let $\{X_n^t\}$ be the random walk on t, as defined in (1.1). Let $b(m;t)$ be the m-backbone of t as before. Define (6)

$$\tau_0 = \tau_0(m;t) = \inf\{ n \geq 0 : X_n^t \in b(m;t) \}, \tag{2.18}$$

$$\tau_{i+1} = \tau_{i+1}(m;t) = \inf\{ n > \tau_i : X_n^t \in b(m;t) \quad \text{but} \quad X_n^t \neq X^t(\tau_i) \}, \qquad i \geq 0,$$

$$U_k = U_k(m;t) = X^t(\tau_k(m;t)). \tag{2.19}$$

---

(6) In order to avoid double subscripts we shall often replace subscripts by arguments. For instance we write $X(r_i)$ for X_{r_i} in the following lines.

$\{ U_k \}$ is the *imbedded random walk* on $b(m; t)$. Its values are the values of X^t at its successive visits to $b(m; t)$, not counting an immediate return to the last visited point of $b(m; t)$. (In not counting such returns we deviate from the usual definition of an imbedded Markov chain). $\{ U_k \}$ is a Markov chain with state space $b(m; t)$ and it is not hard to write down its transition probabilities. Indeed, if $U_k = X(\tau_k) = \langle y \rangle \in b(m; t)$ and $X(\tau_k + 1) = \langle z \rangle$, and $\langle z \rangle$ also lies in $b(m; t)$, then $\tau_{k+1} = \tau_k + 1$ and $U_{k+1} = \langle z \rangle$. If $\langle z \rangle \in t \backslash b(m; t)$, then the fact that $b(m; t)$ is a subtree of the tree t, implies that X^t_n can return from $\langle z \rangle$ to $b(m; t)$ only by visiting $\langle y \rangle$ first (see Figure 2). This next visit to $\langle y \rangle$ does not count as a

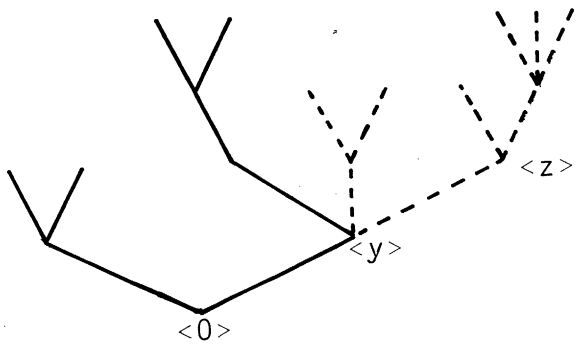

FIG. 2. — The solid edges belong to the backbone $b(m; t)$; the dotted edges to $t \setminus b(m; t)$.

step of U., because $U_k = \langle y \rangle$. Moreover, once at y we start afresh. Either the next step takes $X^t_\cdot$ to a $\langle z \rangle \in b(m, t)$ or $X^t_\cdot$ moves again into $t \backslash b(m; t)$ and returns to $\langle y \rangle$. From this one easily sees (compare (1.1)) that for two adjacent points $\langle y \rangle$ and $\langle z \rangle$ of $b(m; t)$

$$(2.20) \quad P\{ U_{k+1} = \langle z \rangle \mid U_k = \langle y \rangle \} = P\{ X_{n+1} = \langle z \rangle \mid X_n = \langle y \rangle, X_{n+1} \in b(m; t) \} = \frac{1}{\Delta(y)},$$

where

$$(2.21) \quad \Delta(y) = \Delta(y; b(m; t)) = \# \text{ of neighbors of } x \text{ in } b(m; t).$$

The above argument yields more than just (2.20). For $\langle y \rangle \in b(m; t)$ define

$$(2.22) \qquad \Xi(y) = \{ \text{vertices in } t(y) \backslash b(m; t) \} \cup \{ \langle y \rangle \}.$$

$\Xi(y)$ contains the descendants of y which do not lie on the backbone (but in the dashed part of Figure 2). By definition of height and backbone we have $0 \le h(v) - h(y) \le m$ for each $v \in \Xi(y)$. Therefore

(2.23) $$0 \le h(X_n^t) - h(U_k) \le m \qquad \text{for} \qquad \tau_k \le n < \tau_{k+1}.$$

This will allow us later on to translate a limit theorem for $\{ U_k \}$ into one for $\{ X_n \}$.

We should also note that the transitions of $\{ U_n \}$ do not depend on $t \backslash b(m; t)$. In the argument above, $\Xi(y)$ has no influence on U_{k+1} when we already know $U_k = \langle y \rangle$. This results in the following independence property.

(2.24) Conditionally on $\tilde{B}(m)$, $\tilde{T} \backslash \tilde{B}(m)$ and $\{ U_n(m; \tilde{T}) \}_{n \ge 0}$ are independent.

Also on $\{ U_k = \langle y \rangle \}$, $\tau_{k+1} - \tau_k$ is just the time spent by $\{ X. \}$ in $\Xi(y)$, and this depends only on $\Xi(y)$ and $\Delta(y)$, and hence on U_k, but not on the U_n with $n \ne k$. Thus.

(2.25) The conditional distribution of $\tau_{k+1} - \tau_k$ given $\tilde{T}$ and $\{ U_n \}_{n \ge 0}$ is the same as the conditional distribution of $\tau_{k+1} - \tau_k$ given $\tilde{T}$ and U_k.

(2.26) Conditionally on $\tilde{T}$ and $\{ U_n \}$, the $\{ \tau_{k+1} - \tau_k \}_{k \ge 0}$ are independent.

We now derive some properties of the conditional distribution of $\tau_{k+1} - \tau_k$, given $\tilde{T}$ and $\{ U_n \}$. We define for a tree t and $\langle y \rangle$ a vertex of t

(2.27) $$d(y; t) = \text{degree of } \langle y \rangle \text{ in } t.$$

(2.28) LEMMA. — *If t is a tree with a single infinite line of descent and $\langle y \rangle$ a vertex of $b(m; t)$, then*

(2.29) $$E \{ \tau_{k+1}(m; t) - \tau_k(m; t) \mid U_k(m; t) = \langle y \rangle \} = \alpha(y, m; t) := \frac{1}{\Delta(y; b(m; t))} \sum_{\langle v \rangle \in \Xi(y)} d(v; t).$$

Also

(2.30) $$\beta(y) = \beta(y; m; t) := E \{ [\tau_{k+1} - \tau_k]^2 \mid U_k(m; t) = \langle y \rangle \} \le 4 [\alpha(y)]^2 + \frac{8m}{\Delta(y)} \Sigma_j \Big[\sum_{\langle v \rangle \in t(y, j)} d(v; t) \Big]^2,$$

where Σ_j runs over all children $\langle y, j \rangle$ of $\langle y \rangle$ in $t \backslash b(m; t)$.

Proof. — First we note that $\{ X_n^t \}$ must be recurrent. This is so because we assumed that t has a single infinite line of descent. This means that all side trees of this line are finite, and $X_n^t \to \infty$ is possible only if the imbedded random walk on the infinite line of descent moves out to infinity. Moreover, just as in (2.20), the latter is a simple random walk on a halfline (reflected at the origin), hence recurrent. Thus also $\{ X^t_n \}$ is recurrent. Consequently, w. p. 1. $\tau_k < \infty$ for all k. Moreover, for any two vertices $\langle u \rangle$, $\langle v \rangle$ of t, when $X^t_n = \langle u \rangle$ for some n, then $X^t_.$ will visit $\langle v \rangle$ some time after n.

We next show that the hitting probability

$$(2.31) \qquad P\{ X^t_. \text{ hits } \langle v \rangle \text{ before } \langle w \rangle \mid X^t_0 = \langle u \rangle \} = \frac{r}{q+r},$$

when

$\langle v \rangle = \langle 0, l_1, \ldots, l_p \rangle$, $\langle u \rangle = \langle 0, l_1, \ldots, l_{p+q} \rangle$ and $\langle w \rangle = \langle 0, l_1, \ldots, l_{p+q+r} \rangle$

(i. e., $\langle u \rangle$ and $\langle w \rangle$ are descendants of $\langle v \rangle$ and $\langle u \rangle$, respectively). To see (2.31) observe that the left hand side is again a hitting probability for the imbedded one-dimensional random walk on the line of descent from $\langle v \rangle$ to $\langle w \rangle$, $\{ \langle 0, l_1, \ldots, l_s \rangle : p \le s \le p+q+r \}$. Excursions by $X^t_.$ into subtrees of t with root on this segment are of no consequence, since X^t has to return to the root in order to reach $\langle v \rangle$ or $\langle w \rangle$. Thus (2.31) is just the one-dimensional gambler's ruin formula ([8], Ex. XIV.9.1).

Now let $\langle y \rangle = \langle 0, l_1, \ldots, l_p \rangle \in b(m; t)$ and $\langle v \rangle = \langle 0, l_1, \ldots, l_{p+q} \rangle$ a descendant of $\langle y \rangle$ in $\Xi(y)$. Let

$$(2.32) \qquad \sigma = \inf \{ n > 0 : X^t_n \in b(m; t) \}.$$

Note that on $\{ X^t_0 = \langle y \rangle \}$, σ is the minimum of τ_1 and the hitting time of $\langle y \rangle$. We claim that

$$(2.33) \qquad \begin{aligned} &E\{ \# \text{ of visits by } X^t_. \text{ to } \langle v \rangle \text{ in the time interval } [0, \sigma) \mid X^t_0 = \langle y \rangle \} \\ &= P\{ X_. \text{ hits } \langle v \rangle \text{ before } \sigma \mid X^t_0 = \langle v \rangle \} \\ &\quad . \ [1 - P\{ X_. \text{ returns to } \langle v \rangle \text{ before } \sigma \mid X^t_0 = \langle v \rangle \}]^{-1} = \frac{d(v)}{d(y)}. \end{aligned}$$

Only the second equality needs proof. The argument for (2.20) shows that, starting at $\langle y \rangle$, $X^t_.$ can hit $\langle v \rangle$ before σ only if $X^t_1 = \langle 0, l_1, \ldots, l_{p+1} \rangle$, and if then $X^t_.$ hits $\langle v \rangle$ before returning to $\langle y \rangle$. Thus, by (1.1) and (2.31)

$$(2.34) \qquad P\{ X_. \text{ hits } \langle v \rangle \text{ before } \sigma \mid X^t_0 = \langle y \rangle \} = \frac{1}{d(y)} \cdot \frac{1}{q}.$$

Similarly (again using the fact that t is a tree)

$$\begin{aligned} &1 - P\{ X_. \text{ returns to } \langle v \rangle \text{ before } \sigma \mid X^t_0 = \langle v \rangle \} \\ &\quad = P\{ X_. \text{ hits } \langle y \rangle \text{ before returning to } \langle v \rangle \mid X^t_0 = \langle v \rangle \} = \frac{1}{d(v)} \cdot \frac{1}{q}. \end{aligned}$$

Thus (2.33) holds. (2.33) remains valid for $v = y$.

It is now easy to prove (2.29). By the argument for (2.20), when $U_k = \langle y \rangle$ then $\tau_{k+1} - \tau_k$ equals the number of visits to $\Xi(y)$ in the time interval $[\tau_k, \tau_{k+1})$. If we denote by $\sigma_0 = \tau_k < \sigma_1 < \ldots < \sigma_\lambda$ all successive times X^t visits $\langle y \rangle$ in the interval $[\tau_k, \tau_{k+1})$, and by $v_{k,j}(v)$ the number of visits to $\langle v \rangle$ in the interval (7) $[\sigma_j, \sigma_{j+1} \wedge \tau_{k+1})$, then on $\{ U_k = \langle y \rangle \}$

$$(2.35) \qquad \tau_{k+1} - \tau_k = \sum_{\langle v \rangle \in \Xi(y)} \sum_{j=0}^{\lambda} v_{k,j}(v) .$$

To simplify notation we shall denote for the remainder of this lemma conditional expectation given $b(m; t)$ and $U_k = \langle y \rangle$ by E^y. We now have by Wald's identity ([3], Theorem 5.3.1) and (2.33)

$$E^y \{ \tau_{k+1} - \tau_k \} = \sum_{\langle v \rangle \in \Xi(y)} E^y \{ \lambda + 1 \} E^y \{ v_{k,0}(v) \}$$
$$= E^y \{ \lambda + 1 \} \sum_{\langle v \rangle \in \Xi(y)} \frac{d(v)}{d(y)} .$$

(2.29) follows since λ has the geometric distribution

$$(2.36) \quad P \{ \lambda = l \mid U_k = \langle y \rangle \} = \left[\frac{d(y) - \Delta(y)}{d(y)} \right]^l \frac{\Delta(y)}{d(y)}, \qquad l \geq 0 .$$

For (2.30) we begin with the estimate

$$(2.37) \qquad \beta(y) \leq 2E^y \left\{ \left[\sum_{j=0}^{\lambda} \sum_{\langle v \rangle \in \Xi(y)} \left\{ v_{k,j}(v) - \frac{d(v)}{d(y)} \right\} \right]^2 \right\}$$
$$+ 2E^y \left\{ \left[\sum_{j=0}^{\lambda} \sum_{\langle v \rangle \in \Xi(y)} \frac{d(v)}{d(y)} \right]^2 \right\},$$

which is immediate from (2.35). By (2.33) and the second moment analogue of Wald's identity ([3], Theorem 5.3.3) the right hand side of (2.37) equals

$$2E^y \{ \lambda + 1 \} E^y \left\{ \left(\sum_{\langle v \rangle \in \Xi(y)} \left[v_{k,0}(v) - \frac{d(v)}{d(y)} \right) \right]^2 \right\}$$
$$+ 2E^y \{ (\lambda + 1)^2 \} \left[\sum_{\langle v \rangle \in \Xi(y)} \frac{d(v)}{d(y)} \right]^2 .$$

(7) $a \wedge b = \min(a, b)$.

Since, by (2.36)

$$E^y\{\lambda+1\}=\frac{d(y)}{\Delta(y)},\qquad E^y\{(\lambda+1)^2\}\le 2\left[\frac{d(y)}{\Delta(y)}\right]^2,$$

it suffices to estimate

$$(2.38)\qquad \sum_{\langle v\rangle\in\Xi(y)}\sum_{\langle w\rangle\in\Xi(y)} E^y\left\{\left(\nu_{k,0}(v)-\frac{d(v)}{\Delta(y)}\right)\cdot\left(\nu_{k,0}(w)-\frac{d(w)}{d(y)}\right)\right\}.$$

Both $\langle v\rangle$ and $\langle w\rangle$ in this sum have to be descendants of $\langle y\rangle$ or equal $\langle y\rangle$ itself. Let $\langle y\rangle=\langle 0,l_1,\dots,l_p\rangle$ and $\langle v\rangle=\langle 0,l_1,\dots,l_{p+q}\rangle$ as above. Assume that the label of $\langle w\rangle$ agrees with that of $\langle v\rangle$ exactly until the $(p+r)$-th generation. Then $\langle w\rangle=\langle 0,l_1,\dots,l_{p+r},j_{p+r+1},\dots,j_{p+s}\rangle$ for some j's with $j_{p+r+1}\neq l_{p+r+1}$ or $\langle w\rangle$ is a descendant of $\langle v\rangle$ or *vice versa*. In order for $X^t_\cdot$ to reach $\langle v\rangle$ or $\langle w\rangle$ it must first reach

$$\langle u\rangle=\langle 0,l_1,\dots,l_{p+r}\rangle.$$

Consequently with σ as in (2.32)

$$(2.39)\quad E^y\left\{\left(\nu_{k,0}(v)-\frac{d(v)}{d(y)}\right)\left(\nu_{k,0}(w)-\frac{d(w)}{d(y)}\right)\right\}$$

$$=E^y\{\nu_{k,0}(v)\nu_{k,0}(w)\}-\frac{d(v)}{d(y)}\frac{d(w)}{d(y)}\le E^y\{\nu_{k,0}(v)\nu_{k,0}(w)\}$$

$$=P\{X^t_\cdot\text{ reaches }\langle u\rangle\text{ before }\sigma\mid X^t_0=y\}$$

$$.\,E\{(\#\text{ of visits to }\langle v\rangle\text{ before }\sigma)\,(\#\text{ of visits to }\langle w\rangle\text{ before }\sigma)\mid X^t_0=\langle u\rangle\}.$$

We must now distinguish two cases. First $r=0$. Then $\langle v\rangle$ and $\langle w\rangle$ are descendants of different children of $\langle y\rangle$, or at least one of $\langle v\rangle$, $\langle w\rangle$ equals $\langle y\rangle$. In the former case $\nu_{k,0}(v)\nu_{k,0}(w)=0$, since X^t_n cannot enter both $t(y,j_1)$ and $t(y,j_2)$ for $j_1\neq j_2$ between two successive visits to y. In the latter case, if $\langle v\rangle=y$, then $\nu_{k,0}(v)=1=d(v)/d(y)$. Thus, the left hand side of (2.39) is ≤ 0 whenever $r=0$. The second case is $r\ge 1$. In this case the probability in the last member of (2.39) equals $1/(d(y)r)$ by (2.34). It seems that we must imitate the proof of (2.29) to estimate the remaining expectation. Let $0=\rho_0<\rho_1<\dots<\rho_\kappa<\sigma$ be the successive times during $[0,\sigma)$ when $X^t_\cdot$ visits $\langle u\rangle$, and set

$$\xi_j(v)=\#\text{ of visits by }X^t_\cdot\text{ to }\langle v\rangle\text{ in the time interval }[\rho_j,\rho_{j+1}\wedge\sigma),$$

and similarly with $\langle v \rangle$ replaced by $\langle w \rangle$. Then the last expectation in (2.39) equals

$$(2.40)\quad \mathrm{E}\left\{\sum_{a=0}^{\kappa}\sum_{b=0}^{\kappa}\xi_a(v)\xi_b(w)\,|\,\mathrm{X}_0^t = \langle u \rangle\right\}$$

$$= \sum_{a=0}^{\infty} \mathrm{P}\{\kappa > a-1\,|\,\mathrm{X}_0^t = \langle u \rangle\}\,\mathrm{E}\{\xi_0(v)\xi_0(w)\,|\,\mathrm{X}_0^t = \langle u \rangle\}$$

$$+ \sum_{\substack{a,b=0\\ a\neq b}}^{\infty} \mathrm{E}\{\xi_a(v)\xi_b(w)\mathrm{I}[\kappa > \max(a,b)-1]\,|\mathrm{X}_0^t = \langle u \rangle\}.$$

If both $\langle v \rangle \neq \langle u \rangle$ and $\langle u \rangle \neq \langle w \rangle$, then any excursion of $\mathrm{X}^t_\cdot$ starting at $\langle u \rangle$ cannot reach both $\langle v \rangle$ and $\langle w \rangle$ before returning to $\langle u \rangle$ (by the tree structure of t and the choice of u). Thus $\mathrm{E}\{\xi_0(v)\xi_0(w)\,|\,\mathrm{X}_0^t = \langle u \rangle\} = 0$ unless $\langle v \rangle = \langle u \rangle$ or $\langle w \rangle = \langle u \rangle$. If $\langle v \rangle = \langle u \rangle$, then $\xi_0(v) = 1$ and (see (2.33))

$$\mathrm{E}\{\xi_0(v)\xi_0(w)\,|\,\mathrm{X}_0^t = \langle u \rangle\} = \mathrm{E}\{\xi_0(w)\,|\,\mathrm{X}_0^t = \langle u \rangle\} = \frac{d(w)}{d(u)}.$$

(This remains valid even if $\langle v \rangle = \langle w \rangle = \langle u \rangle$). As for the second sum in (2.40), if $a < b$, then the Markov property of $\mathrm{X}^t_\cdot$ gives

$$\begin{aligned}
&\mathrm{E}\{\xi_a(v)\xi_b(w)\mathrm{I}[\kappa > \max(a,b)-1]\,|\,\mathrm{X}_0^t = \langle u \rangle\}\\
&\quad = \mathrm{E}\{\xi_0(w)\,|\,\mathrm{X}_0^t = \langle u \rangle\}\,\mathrm{E}\{\xi_a(v)\mathrm{I}[\kappa > b-1]\,|\,\mathrm{X}_0^t = \langle u \rangle\}\\
&\quad = \mathrm{E}\{\xi_0(w)\,|\,\mathrm{X}_0^t = \langle u \rangle\}\,\mathrm{P}\{\kappa > a-1\,|\,\mathrm{X}_0^t = \langle u \rangle\}\,.
\end{aligned}$$

$$\begin{aligned}
&\mathrm{E}\{\xi_0(v)\mathrm{I}[\kappa \geq 1]\,|\,\mathrm{X}_0^t = \langle u \rangle\}\,.\,\mathrm{P}\{\kappa > b-a-2\,|\,\mathrm{X}_0^t = \langle u \rangle\}\\
&\quad \leq \mathrm{E}\{\xi_0(w)\,|\,\mathrm{X}_0^t = \langle u \rangle\}\,\mathrm{E}\{\xi_0(v)\,|\,\mathrm{X}_0^t = \langle u \rangle\}\,\mathrm{P}\{\kappa > a-1\,|\,\mathrm{X}_0^t = \langle u \rangle\}\\
&\qquad\qquad .\ \mathrm{P}\{\kappa > b-a-2\,|\,\mathrm{X}_0^t = \langle u \rangle\}\\
&\quad = \frac{d(w)}{d(u)}\frac{d(v)}{d(u)}\mathrm{P}\{\kappa > a-1\,|\,\mathrm{X}_0^t = \langle u \rangle\}\,\mathrm{P}\{\kappa > b-a-2\,|\,\mathrm{X}_0^t = \langle u \rangle\}\\
&\qquad\qquad \text{(by (2.33))}.
\end{aligned}$$

Substitution of these results and obvious symmetry properties show that (2.40) is in all cases at most

$$2\frac{d(v)d(w)}{d^2(u)}\left[\mathrm{E}\{(\kappa+1)\,|\,\mathrm{X}_0^t = \langle u \rangle\} + [\mathrm{E}\{(\kappa+1)\,|\,\mathrm{X}_0^t = \langle u \rangle\}]^2\right]$$

$$\leq 4\frac{d(v)d(w)}{d^2(u)}(d(u)r)^2 = 4r^2 d(v)d(w)\,.$$

(For the inequality we use that κ has a geometric distribution with parameter $1-(d(u)r)^{-1}$, by (2.34) with $\langle y \rangle$ and $\langle v \rangle$ replaced by $\langle u \rangle$ and $\langle y \rangle$, respectively. Finally, if we take into account that $r \leq m$ (as in (2.23) since $\langle v \rangle, \langle w \rangle \in \Xi(y)$), we obtain from the above observations that

$$\beta(y) \leq 4\left[\sum_{v\in\Xi(y)} \frac{d(v)}{\Delta(y)}\right]^2 + \frac{8m}{\Delta(y)}\Sigma^* d(v)d(w),$$

where Σ^* runs over all pairs $\langle v \rangle, \langle u \rangle \in \Xi(y)$ for which the corresponding $r \geq 1$. $r \geq 1$ can happen only if $\langle v \rangle$ and $\langle w \rangle$ are both children of one $\langle y, j \rangle$. This proves (2.30). □

To prove Theorem 1.19 we need some information on the conditional distribution of $\alpha(y, m; \tilde{T})$ and $\beta(y, m; \tilde{T})$, given $\tilde{B}(m)$ ($\tilde{T}$ is random now).

(2.41) LEMMA. — *Given* $\tilde{B}(m)$, *the* $\{\alpha(y), \beta(y) : \langle y \rangle \in \tilde{B}(m)\}$ *are conditionally independent. They have the following properties for* $\langle y \rangle \in \tilde{B}(m)$, *uniformly in* $\tilde{B}(m)$ *and* $\langle y \rangle$ [8]:

(2.42) *If* $\tilde{N}(y; m) = 0$, *i. e.*, $\langle y \rangle$ *is an endpoint of* $\tilde{B}(m)$, *then as* $m \to \infty$

$$\nu\{\alpha(y) \mid \tilde{B}(m)\} \sim C_1 m^2,$$
$$\nu\{\beta(y) \mid \tilde{B}(m)\} \leq C_2 m^5,$$
$$\nu\{m^{-2}\alpha(y) \geq \lambda \mid \tilde{B}(m)\} \to \varphi_1(\lambda) \quad \text{for some } 0 \leq \varphi_1(\lambda) \leq 1 \text{ when } \lambda > 0.$$

(2.43) *If* $\tilde{N}(y; m) = 1$ *and* $\langle y \rangle \neq \langle 0 \rangle$, *then as* $m \to \infty$

$$\nu\{\alpha(y) \mid \tilde{B}(m)\} \sim C_3 m,$$
$$\nu\{\beta(y) \mid \tilde{B}(m)\} \leq C_4 m^4,$$
$$m\nu\{m^{-2}\alpha(y) \geq \lambda \mid \tilde{B}(m)\} \to \varphi_2(\lambda) \quad \text{for some } 0 < \varphi_2(\lambda) < \infty \text{ when } \lambda > 0;$$
$$\text{as } \lambda \downarrow 0 \qquad \varphi_2(\lambda) \sim C_5 \lambda^{-1/2}.$$

(2.44) *If* $\tilde{N}(y; m) = 2$ *or* $\langle y \rangle = \langle 0 \rangle$ *and* $0 < \tilde{N}(0; m) \leq 2$, *then as* $m \to \infty$

$$\frac{1}{m^2}\nu\{\alpha(y) \mid \tilde{B}(m)\} \to 0.$$

Proof. — Note that $\alpha(y)$ and $\beta(y)$ depend on $\tilde{B}(m)$ only through $\Delta(y)$ and $d(y)$, by their definition and (2.35). Also, their conditional distributions, given $\tilde{B}(m)$, depend only on $\Delta(y)$, by Lemma 2.14. Therefore, all the statements in this lemma are automatically uniform in $\tilde{B}$ and $\langle y \rangle \in \tilde{B}$.

---

[8] Here and in the sequel we write $\nu\{X\}$ for the expectation of a random variable X with respect to ν.

Before starting on the moment estimates observe that for $\langle v \rangle \neq \langle 0 \rangle$ $d(v;t) = (1 +$ the number of children of $\langle v \rangle$ in $t)$. Therefore, for $k \geq 1$,

$$\sum_{\langle v \rangle \in t_k} d(v;t) = \# t_k + \# t_{k+1},$$

where, as before, t_k denotes the k-th generation of t. Thus. For any finite tree t

$$\sum_{v \in t} d(v;t) = 2(\# \text{ of vertices in } t) - 1. \tag{2.45}$$

Also note that for $\langle y \rangle \neq \langle 0 \rangle, \Delta(y) = \tilde{N}(y;m) + 1$ while for $\langle y \rangle = \langle 0 \rangle$, $\Delta(y) = \tilde{N}(0;m)$. This is the only reason why $\langle 0 \rangle$ plays a special role in (2.43), (2.44).

We now prove (2.43). For $\langle y \rangle \neq 0$ and $\tilde{N}(y;m) = 1$ a decomposition according to $M(y;m)$, (2.15), (1.13) and (1.17) show (9).

$$\begin{aligned}
v\{\alpha(y) \mid \tilde{B}(m)\} = & \left[\sum_{l \geq 0} p_{l+1}(l+1)[P\{\xi(T) < m-1\}]^l\right]^{-1} \\
& \cdot \sum_{k \geq 0} p_{k+1}(k+1)[P\{\xi(T) < m-1\}]^k \\
& \cdot \frac{1}{2}\left[k+1+kE\left\{\sum_{\langle v \rangle \in T} d(v;T) \mid \xi(T) < m-1\right\}\right] \\
& \sim \frac{1}{2}\left(\sigma^2+1+\sigma^2 E\left\{\left[\sum_{\langle v \rangle \in T} d(v;T)\right]; \zeta(T) < m-1\right\}\right) (\text{as } m \to \infty).
\end{aligned}$$

Also, by (2.45)

$$\begin{aligned}
E\left\{\left[\sum_{\langle v \rangle \in T} d(v;T)\right]; \zeta(T) < m-1\right\} \\
= 2\sum_{k=0}^{m-1} E\{Z_k; \zeta(T) < m-1\} - P\{\zeta(T) < m-1\}.
\end{aligned}$$

(9) $\mathbb{E}\{X;E\}$ denotes the integral of X over the set E (with respect to $\mathbb{P}$), that is $\mathbb{E}\{X;E\} = \mathbb{E}\{XI_E\}$.

Further, by Theorem I.9.2 of [2] and (2.5)

$$E\{Z_k; \zeta(T) < m-1\} = \sum_{r=0}^{\infty} P\{Z_k = r\} r [P\{\zeta(T) < m-1-k\}]^r$$

$$\sim \frac{2}{k\sigma^2}\frac{2}{\sigma^2}\int_0^\infty kx \exp\left(-\frac{2x}{\sigma^2}\left(1+\frac{k}{m-1-k}\right)\right)dx = \left(\frac{m-1-k}{m-1}\right)^2.$$

For the approximate equality we need the simple domination (see [10], I.(5.3))

$$\sum_{r \geq Ak} rP\{Z_k = r\} \leq \frac{1}{Ak} E\{Z_k^2\} = \frac{1}{Ak}(k\sigma^2 + 1).$$

The first formula in (2.43) now follows easily.

For the second formula in (2.43) we estimate the last member of (2.30). For $\Delta(y) = 2$

$$4[\alpha(y, m; \tilde{T})]^2 \leq \left[d(y; \tilde{T}) + \Sigma_j \sum_{\langle v \rangle \in T(y,j)} d(v; \tilde{T})\right]^2.$$

Thus, by (2.30) and $d(y; \tilde{T}) = \Delta(y; \tilde{B}(m)) + M(y; m)$

$$(2.46) \quad \beta(y) \leq C_1 \left\{ 1 + M^2(y; m) + m\Sigma_j \left[\sum_{\langle v \rangle \in T(y,j)} d(v; \tilde{T})\right]^2 + \sum_{j_1 \neq j_2} \left[\sum_{\langle v \rangle \in T(y,j_1)} d(v; \tilde{T})\right]\left[\sum_{\langle v \rangle \in T(y,j_2)} d(v; \tilde{T})\right] \right\}.$$

By (2.15) and (2.45),

$$\nu\left\{\left[\sum_{\langle v \rangle \in T(y,j)} d(v; \tilde{T})\right]^2 \mid \tilde{B}(m), \langle y, j \rangle \in \tilde{T}\backslash\tilde{B}(m)\right\}$$

$$\leq 4E\left\{\left[\sum_{k \geq 0} Z_k\right]^2 \mid \zeta(T) < m-1\right\} \leq C_2 E\left\{\sum_{k=0}^{m-1}\sum_{l=k}^{m-1} Z_k Z_l\right\}.$$

Further, by [10], Theorem I.5.1, for $l \geq k$

$$E\{Z_k Z_l\} = E\{Z_k E\{Z_l \mid Z_k\}\} = EZ_k^2 = 1 + k\sigma^2.$$

Thus (again use (2.15))

$$\nu\left\{\Sigma_j\left[\sum_{\langle v \rangle \in \tilde{T}(y,j)} d(v; \tilde{T})\right]^2 \mid \tilde{B}(m)\right\} \leq C_3 m^3 \nu\{M(y; m) \mid \tilde{B}(m)\} \leq C_4 m^3.$$

Similarly, using (2.15) for $\tilde{N}(y; m) = 1$, we have

$$v\left\{ \sum_{j_1 \neq j_2} \left[\sum_{\langle v \rangle \in \tilde{T}(y, j_1)} d(v; \tilde{T}) \right] \left[\sum_{\langle v \rangle \in \tilde{T}(y, j_2)} d(v; \tilde{T}) \right] \mid \tilde{B}(m) \right\}$$

$$\leq C_5 v \{ M^2(y; m) \mid \tilde{B}(m) \} \left[E\left\{ \sum_{k=0}^{m-1} Z_k \right\} \right]^2$$

$$\leq C_5 m^2 \sum_{k \geq 0} (k+1) k^2 p_{k+1} [P\{ \zeta(T) < m - 1 \}]^k .$$

This expression is at most $C_6 m^3$ by (1.17) and the following estimate based on (2.5), valid for large m:

$$kP\{ \zeta(T) < m - 1 \} \leq k\left(1 - \frac{1}{\sigma^2 m}\right)^k \leq k e^{-k/(m\sigma^2)} \leq C_6 m . \tag{2.47}$$

Together with (2.46) these estimates prove the second formula in (2.43).

For the third parts of (2.43) as well as (2.42) we need the fact that

$$\lim_{l \to \infty} P\left\{ \sum_0^l Z_k \leq \lambda l^2 \mid \zeta(T) = l \right\} = H(\lambda) \tag{2.48}$$

for some honest distribution function $H(\lambda)$ which does not put all mass at zero. (2.48) was essentially proved by Durrett [*7*] and Pakes [*23*]. In particular, [*7*], Theorem 4.5 shows that for fixed $\varepsilon > 0$ the process $\{ l^{-1} Z_{l(1-\varepsilon)t} \}_{0 \leq t \leq 1}$, conditioned on $Z_{l(1-\varepsilon)} > 0$, converges weakly in $D[0, 1]$ to some Markov process. Consequently the conditional joint distribution of

$$l^{-2} \sum_0^{l(1-\varepsilon)} Z_k \quad \text{and} \quad l^{-1} Z_{l(1-\varepsilon)},$$

given $Z_{l(1-\varepsilon)} > 0$, converges. We write $H_\varepsilon(y, z)$ for the limit distribution function and use the asymptotic relation ([*2*], Cor. I.9.1 with $s = 0$)

$$P\{ \zeta(T) = l \} \sim \frac{2}{\sigma^2 l^2}, \qquad l \to \infty . \tag{2.49}$$

We then obtain

$$(2.50)\quad P\left\{ \sum_{0}^{l(1-\varepsilon)} Z_k \le \lambda l^2 \mid \zeta(T) = l \right\}$$

$$\sim \frac{\sigma^2 l^2}{2} \sum_{r=1}^{\infty} P\left\{ \sum_{k=0}^{l(1-\varepsilon)} Z_k \le \lambda l^2, Z_{l(1-\varepsilon)} = r \mid Z_{l(1-\varepsilon)} > 0 \right\}$$

$$. P\{ \zeta(T) = l \mid Z_{l(1-\varepsilon)} = r \} . P\{ Z_{l(1-\varepsilon)} > 0 \}.$$

By (2.5) and (2.49)

$$P\{ \zeta(T) = l \mid Z_{l(1-\varepsilon)} = r \} = [P\{ \zeta(T) \le \varepsilon l \}]^r - [P\{ \zeta(T) \le \varepsilon l - 1 \}]^r$$

$$\sim rP\{ \zeta(T) = \varepsilon l \} [P\{ \zeta(T) \le \varepsilon l \}]^{r-1} \sim \frac{2r}{\sigma^2 \varepsilon^2 l^2} \exp\left(- \frac{2r}{\sigma^2 \varepsilon l} \right), \qquad l \to \infty .$$

From this it follows that (2.50) converges to

$$\frac{2}{\varepsilon^2 \sigma^2 (1 - \varepsilon)} \int H_\varepsilon(\lambda, dy) y \exp\left(- \frac{2y}{\sigma^2 \varepsilon} \right).$$

This is almost (2.48). To complete the proof of (2.48) it suffices to show that for all large l

$$P\left\{ \sum_{l(1-\varepsilon)}^{l-1} Z_k > \sqrt{\varepsilon} l^2 \mid \zeta(T) = l \right\} \le C_{11} \sqrt{\varepsilon},$$

which in turn follows from

$$\sum_{l(1-\varepsilon)}^{l-1} E\{ Z_k \mid \zeta(T) = l \}$$

$$\le \sum_{k=l(1-\varepsilon)}^{l-1} \sum_{r=1}^{\infty} rP\{ Z_k = r \} P\{ \zeta(T) = l \mid Z_k = r \} \sigma^2 l^2 \qquad \text{(by (2.49))}$$

$$\le C_7 \sum_{k=l(1-\varepsilon)}^{l-1} \sum_{r=1}^{\infty} P\{ Z_k = r \} r^2 \frac{l^2}{(l-k)^2} \exp - C_8 \frac{r}{l-k}$$

$$\le C_9 \sum_{k=l(1-\varepsilon)}^{l-1} l^2 P\{ Z_k \ne 0 \} \le C_{10} l^2 \sum_{k=l(1-\varepsilon)}^{l-1} \frac{1}{k} \qquad \text{(by (2.5))}$$

$$\le C_{11} \varepsilon l^2.$$

This proves (2.48).

The third relation of (2.43) is now relatively easy. According to (2.45) and (2.15), for $\tilde{N}(y; m) = 1$

$$(2.51)\quad \nu\{\alpha(y) \geq \lambda m^2 \mid \tilde{B}(m)\} = \sum_{k=1}^{\infty} \nu\{M(y;m) = k \mid \tilde{B}(m)\}$$

$$. \, P\left\{\frac{1}{2} + \sum_{j=1}^{k} \# \, T(j) \geq \lambda m^2 \mid \zeta(T(j)) < m - 1 \quad \text{for} \quad 1 \leq j \leq k\right\},$$

where T(1), T(2), ... are independent copies of T, and # T(j) is the total number of vertices of T(j). By (2.48) and (2.49), as $m \to \infty$

$$mP\{\# \, T \geq \lambda m^2 \mid \zeta(T) < m - 1\} = [P\{\zeta(T) < m - 1\}]^{-1}$$

$$m \sum_{l=1}^{m-2} P\{\zeta(T) = l\} P\{l^{-2} \# \, T \geq \lambda l^{-2} m^2 \mid \zeta(T) = l\}$$

$$\sim m \sum_{l=1}^{m-2} \frac{2}{\sigma^2 l^2} [1 - H(\lambda l^{-2} m^2)] \to \frac{2}{\sigma^2} \int_0^1 \frac{1}{x^2}\left[1 - H\left(\frac{\lambda}{x^2}\right)\right] dx$$

$$= \frac{1}{\sigma^2 \sqrt{\lambda}} \int_\lambda^\infty \frac{1}{\sqrt{y}} [1 - H(y)] dy.$$

The asymptotics here are justified by the simple domination

$$m \sum_{l < \varepsilon m} P\{\zeta(T) = l, \# \, T \geq \lambda m^2\} = mP\{\# \, T \geq \lambda m^2, \zeta(T) < \varepsilon m\}$$

$$\leq mP\left\{\sum_{0}^{\varepsilon m - 1} Z_k \geq \lambda m^2\right\} \leq \frac{1}{\lambda m} E\left\{\sum_{0}^{\varepsilon m - 1} Z_k\right\} = \frac{\varepsilon}{\lambda}.$$

Finally, one easily sees (see also [*12*])

$$P\left\{\frac{1}{2} + \sum_{j=1}^{k} \# \, T(j) \geq \lambda m^2 \mid \zeta(T(j)) < m - 1 \quad \text{for} \quad 1 \leq j \leq k\right\}$$

$$\sim kP\{\# \, T \geq \lambda m^2 \mid \zeta(T) < m-1\} \sim \frac{k}{m\sigma^2\sqrt{\lambda}} \int_\lambda^\infty \frac{1}{\sqrt{y}} [1 - H(y)] dy$$

for each fixed k and $\lambda > 0$, as $m \to \infty$. Moreover, the left hand side is bounded by

$$\frac{1}{\lambda m^2 - \frac{1}{2}} \mathrm{E}\left\{ \sum_{j=1}^{k} \# \mathrm{T}_j \mid \zeta(\mathrm{T}(j)) < m - 1 \text{ for } 1 \le j \le k \right\} \le \frac{2k}{\lambda m}.$$

Substitution of these estimates in (2.51) yields the third relation of (2.43).

The proof of (2.42) is very similar to the proof of (2.43), if one takes into account that for $\tilde{\mathrm{N}}(y; m) = 0$ $\tilde{\mathrm{T}}(y)$ has the distribution of T, conditioned on $\zeta(\mathrm{T}) = m$ (by (2.16)). Moreover $\mathrm{P}\{\zeta(\mathrm{T}) = m\}$ behaves as indicated in (2.49). Actually, the proofs of (2.42) are easier because no decompositions with respect to the value of $\mathrm{M}(y; m)$ are necessary.

The proof of (2.44) is even easier. It is very similar to that of the first formula of (2.43), if in estimating $\nu\{\mathrm{M}(y; m)\}$ one takes into account that for all $\delta > 0$

$$\begin{aligned} &\sum_{k \ge 0} p_{k+2}(k+2)(k+1)k[\mathrm{P}\{\zeta(\mathrm{T}) < m-1\}]^k \\ &\qquad \le \delta m \sum_{k \le \delta m} \mathrm{P}_{k+2}(k+2)(k+1) + \mathrm{C}_6 m \sum_{k > \delta m} p_{k+2}(k+2)(k+1) \\ &\qquad \le \delta m \sigma^2 + o(m) \qquad \text{(by (2.47) and (1.17))}. \end{aligned}$$

Thus

$$\sum_{k \ge 0} p_{k+2}(k+2)(k+1)k[\mathrm{P}\{\zeta(\mathrm{T}) < m-1\}]^k = o(m). \tag{2.52}$$

We leave further details to the reader. □

Proof of Theorem 1.19. — Let $\tau_k = \tau_k(m; \tilde{\mathrm{T}})$ and $\mathrm{U}_k = \mathrm{U}_k(m; \tilde{\mathrm{T}})$ be as in (2.18),(2.19) but now with $\tilde{\mathrm{T}}$ random and distributed according to the measure ν of Lemma 1.14. Unless otherwise stated $\mathrm{U}_0 = \langle 0 \rangle$. Also define

$$i(n) = \max\{k : \tau_k(m; \tilde{\mathrm{T}}) \le n\}$$

and

$$\kappa(m) = \inf\{k : h(\mathrm{U}_k(m, \tilde{\mathrm{T}})) = m\}.$$

Then

$$h(\mathrm{X}(\tau_{\kappa(m)})) = h(\mathrm{U}_{\kappa(m)}) = m$$

so that

$$\mathcal{T}(m) \le \tau(\kappa(m)).$$

On the other hand, by virtue of (2.23), for $n \le \tau(\kappa(m))$ we have $i(n) \le \kappa(m)$

and

$$h(X_n) \leq h(U_{i(n)}) + m \leq 2m,$$

whence

$$\mathcal{T}(2m+1) > \tau(\kappa(m)), \qquad \mathcal{T}(m) > \tau\left(\kappa\left(\frac{1}{2}(m-1)\right)\right).$$

It therefore suffices to prove that for suitable $0 < x_3(\varepsilon), x_4(\varepsilon) < \infty$

$$\mathbb{P}\{ x_3(\varepsilon) \leq m^{-3}\tau(\kappa(m)) \leq x_4(\varepsilon) \} \geq 1 - \varepsilon, \quad m \geq 1.$$

To prove the full Theorem 1.16 we do indeed rely completely on the imbedded random walk $\{ U_k \}$. However, for (1.19) it is simpler to prove the estimate

$$\mathbb{P}\{ m^{-3}\mathcal{T}(m) > x_2 \} \leq \varepsilon \tag{2.53}$$

for some large x_2, directly. For this purpose, define for any vertex $\langle v \rangle$ of $\tilde{T}_{[m]}$ the local time

$$l(\langle v \rangle) = l(\langle v \rangle), \mathcal{T}(m) > = \{ \text{number of } k \leq \mathcal{T}(m) \text{ with } X_k = \langle v \rangle \}.$$

Then

$$\mathcal{T}(m) = \sum_{v \in \tilde{T}_{[m]}} l(\langle v \rangle).$$

We next prove the estimate

$$\mathbb{E}\{ l(\langle v \rangle) \mid \tilde{T}_{[m]} \} \leq (2m+1)d(v; \tilde{T}_{[m]}), \qquad \langle v \rangle \in T_{[m-1]}. \tag{2.54}$$

Let $\langle v \rangle = \langle 0, k_1, \ldots, k_p \rangle$ with $p \leq m-1$ and let $\langle x_0 \rangle = 0, \langle x_1 \rangle, \ldots, \langle x_n \rangle = \langle 0, l_1, \ldots, l_n \rangle$ be the vertices on the infinite line of descent of $\tilde{T}$. Assume that $\langle 0, k_1, \ldots, k_q \rangle = \langle x_q \rangle$ for $q \leq r$, but not for $q > r$. Then define

$$J = \{ \langle 0, k_1, \ldots, k_q \rangle : r+1 \leq q \leq p \} \cup \{ \langle x_s \rangle : r \leq s \leq m \}.$$

As a graph J is isomorphic to a segment of $\mathbb{Z}$ of length $p-r+m-r+1 \leq 2m+1$. Indeed $\langle 0, k_1, \ldots, k_{q+1} \rangle$ is adjacent to $\langle 0, k_1, \ldots, k_q \rangle$ and $\langle 0, k_1, \ldots, k_{r+1} \rangle$ is adjacent to $\langle x_r \rangle$, and $\langle x_{s+1} \rangle$ is adjacent to $\langle x_s \rangle$; no other pairs of vertices in J are adjacent. Next we introduce the subgraph K of $T_{[m]}$ whose vertex set consists of J plus all descendants of $\langle v \rangle$ in $T_{[m]}$. $\{ W_n \}$ will be the imbedded random walk on K; it is defined analogously to $\{ U_n \}$ (cf. (2.18), (2.19)) with $b(m; t)$ replaced by K. It is clear that any visit by X. to $\langle v \rangle$ also counts as a visit by W. to $\langle v \rangle$, so that

$$l(\langle v \rangle) = \text{number of visits by W. to } \langle v \rangle \text{ before } \mathcal{T}(m).$$

But it is easy to see that on $J\backslash\{\langle v\rangle\}$ W behaves exactly like a simple symmetric random walk. Specifically, if $w\in J\backslash\{\langle v\rangle\}$, then

$$P\{W_{n+1}=w'\mid W_n=w\}=\frac{1}{2}$$

for w' one of the two neighbors of w in J (compare (2.20)). Thus it follows from the gambler's ruin formula (just as in (2.31)) that

$$P\{\text{W reaches }\langle x_m\rangle\text{ before returning to }v\mid W_0=\langle 0,k_1,\ldots,k_{p-1}\rangle\}$$
$$\geq(\text{cardinality of J})^{-1}\geq\frac{1}{2m+1}.$$

Also, as in (2.20),

$$P\{W_{n+1}=\langle 0,k_1,\ldots,k_{p-1}\rangle\mid W_n=v\}=\frac{1}{d(v;\tilde{T}_{[m]})}.$$

These properties together show that

$$E\{\text{number of returns to }\langle v\rangle\text{ before reaching }\langle x_m\rangle\mid W_0=\langle v\rangle\}$$
$$\leq d(v;\tilde{T}_{[m]})(2m+1).$$

This proves (2.54). Of course if $\langle v\rangle\in\tilde{T}_m$, then $\mathcal{T}(m)\leq$ first hitting time of $\langle v\rangle$ and $l(\langle v\rangle)\leq 1$.

As a consequence of the above we have

$$\begin{aligned}\mathbb{E}\{\mathcal{T}(m)\}&=\mathbb{E}\left\{\mathbb{E}\left\{\sum_{v\in\tilde{T}_{[m]}}l(\langle v\rangle)\mid\tilde{T}_{[m]}\right\}\right\}\\&\leq(2m+1)\mathbb{E}\left\{\sum_{v\in\tilde{T}_{[m-1]}}[d(v;\tilde{T}_{[m]})]+\#\tilde{T}_m\right\}\\&\leq 6m\,\mathbb{E}\{\#\tilde{T}_{[m]}\}.\end{aligned}$$

For the last inequality we applied (2.45) to $\tilde{T}_{[m]}$. Finally

$$\begin{aligned}\mathbb{E}\{\#\tilde{T}_{[m]}\}&=\nu\{\#\tilde{T}_{[m]}\}\\&=\sum_{k=0}^{m}\nu\{\#\tilde{T}_k\}=E\{(\#T_k)^2\}\quad\text{(by (1.15))}\\&=\sum_{k=0}^{m}(\sigma^2k+1)\quad\text{(cf. [}10\text{], I (5.3))}\leq x_5m^2.\end{aligned}$$

Thus $\mathbb{E}\{\mathcal{T}(m)\}\leq 6x_5m^3$ and (2.53) follows.

We next turn to the lower bound for $\mathcal{T}(m)$, and start with proving

(2.55) $$\mathbb{P}\{m^{-2}\kappa(m)<x_6\}\leq\varepsilon,\qquad m>1$$

for sufficiently small $x_6(\varepsilon)$. For this and later parts of the proof we need the following estimates, valid for any $1 \leq c < \infty$:

(2.56) $\nu\{\tilde{B}_{[cm]}(m)$ contains any branchpoints of order $\geq 3\} \to 0 (m \to \infty)$,

(2.57) $\nu\{\tilde{B}_{[cm]}(m)$ contains more than $x_7 c^2$ branchpoints of order $2\} \leq \varepsilon$, uniformly in $c \geq 1$ and m, for large x_7.

(2.58) $\nu\{\tilde{B}_{[m]}(m)\backslash\tilde{B}_{[x_8 m]}(m)$ contains any branchpoint$\} \leq \varepsilon$, uniformly in m for x_8 close to $1, 0 < x_8 < 1$.

We first deduce (2.55) from (2.56)-(2.58). Let

$E_1 = \{\tilde{B}_{[m]}(m)$ contains no branchpoints of order ≥ 3 and at most x_7 branchpoints of order 2, while $\tilde{B}_{[m]}\backslash\tilde{B}_{[x_8 m]}$ contains no branchpoints at all).

Choose x_7 and x_8 such that $\nu(E_1^c) \leq \varepsilon$ and define

$$\kappa^*(x_8, m) = \max\{k \leq \kappa(m) : h(U_k) \leq x_8 m\}.$$

Then, by definition,

$$U_k \in \tilde{B}_{[m]}\backslash\tilde{B}_{[x_8 m]} \quad \text{for} \quad \kappa^*(x_8, m) < k \leq \kappa(m).$$

Thus, if E_1 occurs, then U_k does not visit any branchpoint of $\tilde{B}$ in the time interval $(\kappa^*(x_8 m), \kappa(m)]$. Consequently during this interval U_k has to lie on a single line of descent

$$\mathscr{L} := \{\langle 0, l_1, \ldots, l_p \rangle : x_8 m < p \leq m\}$$

without branchpoints. By (2.20), under the condition that $\{U_n\}$ has to stay in $\mathscr{L}$, it behaves like a simple symmetric random walk on $(x_8 m, m]$ reflected at $x_8 m + 1$ and m. Therefore, with S_n a simple symmetric random walk on $\mathbb{Z}$

$$\begin{aligned}\mathbb{P}\{\kappa(m) - \kappa^*(x_8, m) \leq x_6 m^2 \mid \tilde{B}(m), U_k \in \mathscr{L} \quad \text{for} \quad \kappa^* < k \leq \kappa\} \\ \leq P\{S_n \text{ reaches } (1 - x_8)m - 1 \text{ before time } x_6 m^2 \mid \\ S_0 = (1 - x_8)m/2, S_n \geq 0 \text{ until } S_n \text{ reaches } (1 - x_8)m - 1\}.\end{aligned}$$

The last probability is the same for all $\mathscr{L}$ and can be made $\leq \varepsilon$ uniformly in m by choosing x_6 small enough (by [*8*], Theorem III.7.3; the condition $S_n \geq 0$ has little influence). Therefore, for small enough $x_6(\varepsilon) > 0$

$$\begin{aligned}(2.59) \quad \mathbb{P}\{\kappa(m) \leq x_6 m^2\} \leq \nu\{E_1^c\} \\ + \mathbb{E}\{\mathbb{P}\{\kappa(m) - \kappa^*(x_8 m) \leq x_6 m^2 \mid \tilde{B}(m), U_k \in \mathscr{L} \text{ for } \kappa^* < k \leq \kappa\}; E_1\} \leq 2\varepsilon.\end{aligned}$$

Thus (2.54) will follow from (2.56)-(2.58).

We next give the proof of (2.56)-(2.58). Any branchpoint of $\tilde{B}_{[cm]}$ of order p is a point of $\tilde{T}_{[cm]}$ with exactly p children of lifetime $\geq m$. Thus, by (1.15) (with $k = (c+1)m$)

$$(2.60)\quad \nu\{\tilde{B}_{[cm]}(m) \text{ has at least } d \text{ branchpoints of order } p\}$$
$$= \nu\{\tilde{T}_{[cm]} \text{ has at least } d \text{ points with } p \text{ children of lifetime } \geq m\}$$
$$= E\{(\# T_{(c+1)m}).I[T_{[cm]} \text{ has at least } d \text{ points with } p \text{ children of lifetime } \geq m\}$$
$$\leq \frac{1}{d} E\{(\# T_{(c+1)m} \cdot [\# \text{ of points in } T_{[cm]} \text{ with } p \text{ children of lifetime } \geq m]\}$$
$$= \frac{1}{d} \sum_{k \leq cm} E\{Z_{(c+1)m} \sum_{\langle v \rangle \in T_k} I[\langle v \rangle \text{ has } p \text{ children of lifetime } \geq m].$$

Now for each $\langle v \rangle \in T_k$, $Z_{(c+1)m}$ counts the offspring of $\langle v \rangle$ in the $(c+1)m$-th generation plus the offspring in the $(c+1)m$-th generation of the $(Z_k - 1)$ individuals other than $\langle v \rangle$ in T_k. Given Z_k and the family tree $T(v)$ of the offspring of $\langle v \rangle$, the latter has conditional expectation $Z_k - 1$. Therefore, The expectation in the last member of (2.60) equals

$$E\left\{\sum_{\langle v \rangle \in T_k} E\{(\# T_{(c+1)m-k}(v) + Z_k - 1)I[\langle v \rangle \text{ has } p \text{ children of lifetime } \geq m]\}\right.$$
$$= E\{Z_k(Z_k - 1)\} P\{\langle 0 \rangle \text{ has } p \text{ children of lifetime } \geq m\}$$
$$+ E\{Z_k\} E\{Z_{(c+1)m-k} I[\langle 0 \rangle \text{ has } p \text{ children of lifetime } \geq m]\}.$$

Now, by (1.13), (1.17) and [*10*], I (5.3) and (2.5)

$$E\{Z_k\} = 1, \qquad E\{Z_k(Z_k - 1)\} = k\sigma^2,$$

$$(2.61)\quad P\{\langle 0 \rangle \text{ has 2 children of lifetime } \geq m\}$$
$$= \sum_{s \geq 2} p_s \binom{s}{2} [P\{\zeta \geq m\}]^2 [P\{\zeta < m\}]^{s-2}$$
$$\sim \Sigma \frac{1}{2} p_s s(s-1)\left(\frac{2}{m\sigma^2}\right)^2 = \frac{2}{\sigma^2}\frac{1}{m^2},$$

$$(2.62)\quad P\{\langle 0 \rangle \text{ has at least 3 children of lifetime } \geq m\}$$
$$\leq \sum_{3 \leq s \leq \delta m} p_s \binom{s}{3} [P\{\zeta \geq m\}]^3 + \sum_{s > \delta m} p_s \leq C_1 \frac{\delta m}{m^3} + \frac{1}{\delta^2 m^2} \sum_{s > \delta m} p_s s^2$$
$$= o\left(\frac{1}{m^2}\right) (\text{since } \delta > 0 \text{ is arbitrary; compare } (2.52)).$$

Also, for any $r \geq m$,

(2.63) $\mathrm{E}\{ Z_r I[\langle 0 \rangle \text{ has } p \text{ children of lifetime } \geq m]\}$

$$= \sum_{s \geq p} p_s \binom{s}{p} [\mathrm{P}\{\zeta \geq m\}]^p [\mathrm{P}\{\zeta < m\}]^{s-p} . p\mathrm{E}\{ Z_r \mid \zeta \geq m \}$$

$$= \sum_{s \geq p} p_s \binom{s}{p} [\mathrm{P}\{\zeta \geq m\}]^{p-1} [\mathrm{P}\{\zeta < m\}]^{s-p} p . \mathrm{E}\{ Z_r ; Z_m > 0 \}$$

$$= \sum_{s \geq p} s p_s \binom{s-1}{p-1} [\mathrm{P}\{\zeta \geq m\}]^{p-1} [\mathrm{P}\{\zeta < m\}]^{s-p},$$

since $\mathrm{E}\{ Z_r ; Z_m > 0 \} = \mathrm{E}\{ Z_r \} = 1$ for $r \geq m$. We can estimate the last member of (2.63) in the same way as (2.61) and (2.62). Substitution of these estimates with $d = 1$ and $p \geq 3$ in (2.60) yields (2.56). Substitution with $d = x_7 c^2$ and $p = 2$ proves (2.57). Finally, for (2.58) we should take $d = 1$, $p \geq 2$, but the sum in the last member of (2.60) should run only over $(1 - x_8)m < k \leq m$.

The preceding proves (2.56)-(2.58) and hence (2.55). This tells us that the imbedded random walk on $\tilde{\mathrm{B}}(m)$ has to take on the order of m^2 steps to reach height m. We next investigate how much time X. spends off the backbone between the successive steps of the imbedded random walk. This will tell us the order of the total time needed to reach height m. To this end we write for any M

$$\tau(\mathrm{M}) = \sum_{0}^{\mathrm{M}-1} (\tau_{k+1} - \tau_k),$$

and remind the reader that by (2.25) the conditional distribution of $(\tau_{k+1} - \tau_k)$, given $\tilde{\mathrm{T}}$ and $\{ \mathrm{U}_n \}$, depends on U_k only. Thus, if we introduce

(2.64) $\mathrm{L}(\langle v \rangle, \mathrm{M}) := \text{number of } k \in [0, \mathrm{M}] \text{ with } \mathrm{U}_k = \langle v \rangle,$

then

$$\tau(\mathrm{M}) = \sum_{\langle v \rangle \in \tilde{\mathrm{B}}(m)} \sum_{0}^{\mathrm{M}-1} (\tau_{k+1} - \tau_k) \mathrm{I}[\mathrm{U}_k = \langle v \rangle],$$

and, by virtue of (2.29),

(2.65) $$\mathbb{E}\{ \tau(\mathrm{M}) \mid \tilde{\mathrm{T}}, \{ \mathrm{U}_n \}_{n \geq 0} \} = \sum_{\langle v \rangle \in \mathrm{B}(m)} \mathrm{L}(\langle v \rangle, \mathrm{M} - 1) \alpha(v, m; \tilde{\mathrm{T}}).$$

As observed before, to obtain a lower bound for $\mathscr{T}(m)$, and hence to complete the proof of Theorem 1.19, we must merely prove

(2.66) $$\mathbb{P}\{\tau(\kappa(m)) < x_3 m^3, E_1\} \le \varepsilon.$$

As in the argument leading to (2.59) it suffices to show

(2.67) $$\mathbb{P}\{\kappa(m)) - \tau(\kappa^*(x_8, m)) \le x_3 m^3 \mid \tilde{B}(m), U_k \in \mathscr{L} \text{ for } \kappa^* < k \le \kappa\} \le \varepsilon,$$

uniformly for $\tilde{B}(m)$ such that E_1 occurs, and any choice of $\mathscr{L}$. To this end we first use the independence property (2.24) and (2.43). If

$$\mathscr{L} = \{\langle y_q \rangle : x_8 m < q \le m\} \quad \text{with} \quad \langle y_q \rangle = \langle 0, l_1, \ldots, l_q \rangle,$$

then on E_1 all points $\langle y_q \rangle$, $x_8 m < q < m$ have $\tilde{N}(y_q; m) = 1$, and given $\tilde{B}(m)$, $\{U_n\}$, the $\alpha(y_q)$, $x_8 m < q < m$ are i. i. d. (by (2.24) and (2.15)). Now set (for the C_5 of (2.43))

$$Q = Q(x_9) = \left\{ q : \left(\frac{1}{4} + \frac{3}{4} x_8\right) m \le q \le \left(\frac{3}{4} + \frac{1}{4} x_8\right) m,\ x_9 m^2 \le \alpha(y_q) \le 2 x_9 m^2 \right\},$$

$\# Q$ = cardinality of Q,

$$E_2 = \left\{ \# Q \ge \frac{1}{4}(1 - x_8) C_5 \left(1 - \frac{1}{\sqrt{2}}\right) x_9^{-1/2} \right\}.$$

By the above $\# Q$ has a binomial distribution corresponding to

$$\left| \left(\left(\frac{1}{4} + \frac{3}{4} x_8\right) m, \quad \left(\frac{3}{4} + \frac{1}{4} x_8\right) m \right) \right| \sim \frac{1}{2}(1 - x_8) m$$

trials, each with success probability $\sim C_5 m^{-1}[x_9^{-1/2} - (2x_9)^{-1/2}]$ (by (2.43)). We therefore have for all small x_9

$$\nu\{E_2^c\} \le \frac{\varepsilon}{4} \text{ for all large } m.$$

Consider a $\tilde{T}$ for which E_2 occurs and assume $U_k \in \mathscr{L}$ for $\kappa^* < k \le \kappa$. Then

$$\tau(\kappa(m)) - \tau(\kappa^*(x_8, m)) \ge \Sigma_Q(\tau_{k+1} - \tau_k),$$

where Σ_Q denotes the sum over those $k \in [\kappa^*, \kappa)$ for which $U_k \in Q$. We further define, again for the C_5 of (2.43),

$$E_3 = \left\{ \sum_{q \in Q} [L(\langle y_q \rangle, \kappa(m) - 1) - L(\langle y_q \rangle, \kappa^*(x_8, m) - 1)] \ge \frac{1}{40} C_5 x_{10} x_9^{-1/2}(1 - x_8) m \right\}.$$

Take $x_{10} = \frac{1}{16}\varepsilon(1-x_8)$. We claim that then

$$(2.68) \qquad \mathbb{P}\{ E_2 \cap E_3^c \mid \tilde{B}(m), U_k \in \mathscr{L} \quad \text{for} \quad \kappa^* < k \leq \kappa \} \leq \frac{\varepsilon}{4}.$$

This is due to the fact that $h(U_k)$ behaves like a simple symmetric random walk $\{ S_n \}$ on $(x_8 m, m]$ reflected at $x_8 m + 1$ until it hits m (see argument before (2.59)). Therefore, for any $\left(\frac{1}{4} + \frac{3}{4}x_8\right)m \leq q \leq \left(\frac{3}{4} + \frac{1}{4}x_8\right)m$ we have with

$$\Lambda(q) := L(\langle y_q \rangle, \kappa(m) - 1) - L(\langle y_q \rangle, \kappa^*(x_8, m) - 1),$$

$$\begin{aligned} & P\{ \Lambda(q) \leq x_{10}m \mid \tilde{T}, U_k \in \mathscr{L} \quad \text{for} \quad \kappa^* < k \leq \kappa \} \\ & \leq P\{ S_n \text{ visits } q \text{ at most } x_{10}m \text{ times before it reaches } m \mid S_0 = x_8 m, \\ & \qquad\qquad S_n > x_8 m \quad \text{for} \quad n > 0 \} \\ & \leq x_{10} m P\{ S_n \text{ reaches } m \text{ before returning to } q \mid S_0 = q, S_n > x_8 m, n > 0 \} \\ & = x_{10} m \frac{1}{2}\frac{1}{m-q} \text{ (cf. (2.34))} \\ & \leq 2(1 - x_8)^{-1} x_{10}. \end{aligned}$$

Therefore

$$\begin{aligned} & \mathbb{P}\left\{ \Lambda(q) \leq x_{10}m \text{ for more than } \frac{1}{2} \# Q \text{ values of } q \in Q \mid \tilde{T}, U_k \in \mathscr{L} \right. \\ & \qquad\qquad \left. \text{for } \kappa^* < k \leq \kappa \right\} \\ & \leq \frac{2}{\# Q} \mathbb{E}\{ \text{number of } q \in Q \text{ with } \Lambda(q) \leq x_{10}m \mid \tilde{T}, U_k \in \mathscr{L} \text{ for } \kappa^* < k \leq \kappa \} \\ & \leq \frac{4}{\# Q}(1-x_8)^{-1}x_{10} . \# Q . = \frac{\varepsilon}{4}. \end{aligned}$$

Since $\# Q \geq \frac{1}{20} C_5 (1-x_8)x_9^{-1/2}$ on E_2 we obtain (2.68)

Finally we use that the probability (2.67) is bounded by

$$\begin{aligned} (2.69) \quad & \mathbb{P}\{ E_2^c \cup E_3^c \mid \tilde{B}(m), U_k \in \mathscr{L} \quad \text{for} \quad \kappa^* < k \leq \kappa \} \\ & = \mathbb{E}\{ I[E_2 \cap E_3] \mathbb{P}\{ \Sigma_Q(\tau_{k+1} - \tau_k) \leq x_3 m^3 \mid \tilde{T}, \{ U_n \}, U_k \in \mathscr{L} \\ & \qquad \text{for} \quad \kappa^* < k \leq \kappa \} \mid \tilde{B}(m), \quad U_k \in \mathscr{L} \quad \text{for} \quad \kappa^* < k \leq \kappa \}. \end{aligned}$$

The first probability here is at most $\varepsilon/2$ for all small x_9 (by our choice of x_{10}), and we now estimate the inner probability in the second term on

$E_2 \cap E_3$. By the independence properties (2.25) (2.26) and by (2.29) (2.30) the summands in Σ_Q are conditionally independent, given $\tilde{T}$ and $\{U_n\}$, with expectation for $\tau_{k+1} - \tau_k$ $\alpha(y_q, m; \tilde{T}) \geq x_9 m^2$ and variance at most $\beta(y_q, m; \tilde{T}) \leq C_2 m [\alpha(y_q)]^2 \leq 4C_2 x_9^2 m^5$ whenever $U_k = \langle y_q \rangle$. (Note that the last member of (2.30) $\leq C_2 m [\alpha(y)]^2 \leq 4C_2 x_9^2 m^5$ for $q \in Q$). Therefore, on $E_2 \cap E_3$, just as in (2.65),

$$E\{\Sigma_Q(\tau_{k+1} - \tau_k) \mid \tilde{T}, \{U_n\}\} = \sum_{q \in Q} \Lambda(q)\alpha(y_q) \geq x_9 m^2 \sum_{q \in Q} \Lambda(q)$$
$$\geq \frac{1}{40} C_5 x_{10}(1 - x_8) x_9^{1/2} m^3,$$
$$\sigma^2(\Sigma_Q(\tau_{k+1} - \tau_k) \mid \tilde{T}, \{U_n\}) \leq \sum_{q \in Q} \Lambda(q)\beta(y_q) \leq 4C_2 x_9^2 m^5 \sum_{q \in Q} \Lambda(q).$$

Therefore, if we choose $x_3 = \frac{1}{80} C_5 x_{10}(1 - x_8) x_9^{1/2}$, the inner probability in (2.69) is by Chebyshev's inequality at most

$$16C_2 m \left[\sum_{q \in Q} \Lambda(q)\right]^{-1} \leq 640 C_2 (C_5 x_{10}(1 - x_8))^{-1} x_9^{1/2}$$
$$\leq C_6 \varepsilon^{-1} x_9^{1/2} (1 - x_8)^{-2}.$$

Lastly we choose $x_9 \leq (4C_6)^{-2} \varepsilon^4 (1 - x_8)^4$ so that this expression will also be $\leq \varepsilon/4$. Then (2.69) and (2.67) will be $\leq \varepsilon$ for all large m and (2.66) and Theorem 1.19 follow.

3. PROOF OF THEOREM 1.27

In outline the proof is very similar to that of Theorem 1.19. We define

$$S(m) = [-m, m]^2, \; S^c(m) = \mathbb{R}^2 \backslash S(m)$$
$$S(v, m) = [v(1) - m, v(1) + m] \times [v(2) - m, v(2) + m] \quad \text{for} \quad v = (v(1), v(2)),$$
$$\partial R = \text{boundary of } R, \text{ for any rectangle } R, \; \mathscr{T}(m) = \inf\{t : X_t \in \partial S(m)\},$$

where in this section X_t is the random walk on $\tilde{W}$, defined by (1.2) with $\mathscr{G}$ replaced by $\tilde{W}$, and $\tilde{W}$ is the incipient infinite percolation cluster of the origin $\underline{0}$, and governed by the measure ν of Lemma 1.24. To define the backbone $\tilde{B}(m)$ in the present context we set for any infinite connected subgraph w of $\mathbb{Z}^2$ containing $\underline{0}$

$b(m; w) = \{ v : v$ a vertex of $w \cap S(m)$ for which there exist two paths r_1 and r_2 on w which only have the vertex v in common, and which connect v to $\underline{0}$ and $\partial S(m)$, respectively $\}$.

$\underline{0}$ is taken as a point of $b(m; w)$ and so are all points of $W \cap \partial S(m)$. $\tilde{B}(m)$ is defined as $b(m; \tilde{W})$. We shall prove that for some $\varepsilon_1 > 0$

$$(3.1) \qquad \mathbb{P}\{ \mathscr{T}(m) \leq m^{2+\varepsilon_1} \} \to 0 .$$

This will imply Theorem 1.27, by virtue of (1.8) (with ρ the Euclidean distance on $\mathbb{R}^2$). To prove (3.1) we define the imbedded random walk $\{ U_n \} = \{ U_n(m; \tilde{W}) \}$ on $\tilde{B}(m)$, by a small variation on (2.18) (2.19). For an infinite connected subgraph w of $\mathbb{Z}^2$ containing $\underline{0}$, and with backbone $b(m; w)$, we first modify X_n^w to $\hat{X}_n^w := X_n^{w \cap S(m)}$, i. e., $\hat{X}_n$ has transition probabilities

$$P\{ \hat{X}_{n+1}^w = y \mid \hat{X}_n^w = x \} = \frac{1}{d(x; w \cap S(m))}$$

if x and y belong to $w \cap S(m)$ and are neighbors, and 0 otherwise. Of course

$$d(x; w \cap S(m)) = \text{number of neighbors of } x \text{ in } w \cap S(m) .$$

If $\hat{X}_0^w = \underline{0}$ this modification has the effect of restricting $\hat{X}$ to the component of w in $S(m)$ which contains the origin. Moreover, it does not effect $\mathscr{T}(m)$, in the sense that for $X_0^w = \hat{X}_0^w = \underline{0}$, X_0^w and $\hat{X}^w$ have the same value of $\mathscr{T}(m)$, and $\hat{X}_n^m = X_n^w$ for $0 \leq n \leq \mathscr{T}(m)$. As before we now set

$$\tau_0 = \tau_0(m; w) = \inf\{ n \geq 0 : \hat{X}_n^w \in b(m; w) \},$$
$$\tau_{i+1} = \tau_{i+1}(m; w) = \inf\{ n > \tau_i : \hat{X}_n^w \in b(m; w) \}, \quad i \geq 0,$$
$$U_k = U_k(m; w) = \hat{X}^w(\tau_k(m; w)) .$$

Note that this time we do not require $U_{k+1} \neq U_k$ in this section. This is the more convenient choice here, but not essential. As in the proof of 1.19 we set

$$\kappa(m) = \kappa(m; w) = \inf\{ k : U_k(m; w) \in \partial S(m) \} .$$

The first step is an analogue of (2.55), namely,

$$(3.2) \qquad P\{ \kappa(m; w) \leq C_2 m^2/\log m \mid U_0 = 0 \} \to 0 \text{ uniformly in } w,$$

for some universal constant C_2. We then complete the proof by estimating the ratio of the time spent in $\tilde{W}$ to the time spent in $\tilde{B}$, i. e., the ratio $\tau(\kappa(m))/\kappa(m)$. We shall show that for a typical realization of $\tilde{W}$, there are many points of $\tilde{W} \backslash \tilde{B}$ near most points of $\tilde{B}$. In addition we show that the number of visits to nearby points up till time $\mathscr{T}(m)$ have to be of the same order. The result of this is that $\mathscr{T}(m) = \tau(\kappa(m))$, the total time spent in $\tilde{W}$, is much larger than $\kappa(m)$, the total time spent in $\tilde{B}$. This, together with (3.2) will yield the theorem.

We begin with the proof of a quite general proposition for reversible Markov chains, of which 3.2 is only a very special case. It raises the interesting open problem 3.12 below.

(3.3) PROPOSITION. — *Let* $\{ U_n \}_{n \geq 0}$ *be a Markov chain with state space* w, *a subset of* $\mathbb{Z}^d$, *and transition probability matrix* $P = (P(x, y))$. *Assume that for some constants* K, K*

$$(3.4) \qquad P\{ | U_{n+1} - U_n | \leq K \} = 1$$

and

(3.5) P *is reversible with invariant measure* μ *on* w *which satisfies*

$$0 < \left| \frac{\mu(x)}{\mu(y)} \right| \leq K^*, \qquad x, y \in w.$$

Then for some universal constant C_1 *(independent of* w *and* $x_0 \in w$*)*

$$(3.6) \quad P\{ | U_n - x_0 | \geq \lambda\sqrt{n} \,|\, U_0 = x_0 \} \leq 2(K^* + 1)(2Kn + 1)^d \exp\left(- C_1\left[\frac{\lambda^2}{16K^2} \wedge \frac{\lambda\sqrt{n}}{4K}\right] + C_1\right).$$

Proof. — Define

$$\gamma(y) = E\{ U_{k+1} - U_k \,|\, U_k = y \} = \sum_z (z - y)P(y, z);$$

$$M_k = U_k - U_0 - \sum_0^{k-1} \gamma(U_j).$$

Then $\{ M_k \}$ is a martingale with increments bounded by 2K and $M_0 = \underline{0}$. Hence, by [*21*], Sect. 70 (with k replaced by n)

$$(3.7) \quad P\left\{ | M_n | \geq \frac{\lambda}{2}\sqrt{n} \right\} = P\left\{ \left| \frac{M_n}{2K} \right| \leq \frac{\lambda}{4K}\sqrt{n} \right\} \leq 2 \exp n[c - (c+1)\log(c+1)] \leq 2\exp - C_1 n[c \wedge c^2],$$

for

$$c = \frac{\lambda}{4\sqrt{n}K}.$$

Thus, the left hand side of (3.6) is at most

$$(3.8) \quad 2\exp - C_1 n[c \wedge c^2] + \sum_{|x - x_0| \geq \lambda\sqrt{n}} P\{ U_n = x, | M_n | < \frac{\lambda}{2}\sqrt{n} \,|\, U_0 = x_0 \},$$

where the sum over x may be restricted to the set

$$A := \{ x \in \mathbb{Z}^d : \lambda\sqrt{n} \leq | x - x_0 | \leq Kn \},$$

which contains at most $(2Kn + 1)^d$ vertices. We now use the reversibility to estimate the summands in (3.8). (The combined use of the martingale decomposition and the reversibility we learned from [5].) Set $x_n = x$. Then

$$(3.9)\quad P\left\{ U_n = x, | M_n | < \frac{\lambda}{2}\sqrt{n} \mid U_0 = x_0 \right\} = \sum_{x_1,\dots,x_{n-1}} \prod_{i=0}^{n-1} P(x_i, x_{i+1}),$$

where the sum rums over all $x_1, \dots, x_{n-1}$ for which

$$(3.10)\quad | M_n | = \left| x - x_0 - \sum_0^{n-1} \gamma(x_j) \right| < \frac{\lambda}{2}\sqrt{n}.$$

Since P is reversible

$$\mu(x_0) \prod_{i=0}^{n-1} P(x_i, x_{i+1}) = \mu(x) \prod_{i=0}^{n-1} P(x_{i+1}, x_i),$$

so that with $z_i = x_{n-i}$, the right hand side of (3.9) is

$$\sum_{z_1,\dots,z_{n-1}} \frac{\mu(x)}{\mu(x_0)} \prod_{i=0}^{n-1} P(z_i, z_{i+1}) \leq K^* \sum_{z_1,\dots,z_{n-1}} \prod_{i=0}^{n-1} P(z_i, z_{i+1}).$$

Since the x_i in (3.9) have to satisfy (3.10), the sum over the z_i is only over those z_i which satisfy

$$\begin{aligned}\left| z_n - z_0 - \sum_0^{n-1} \gamma(z_j) \right| &= \left| x_0 - x - \sum_1^{n} \gamma(x_j) \right| \\ &= \left| 2(x_0 - x) - \gamma(x) + \gamma(x_0) + x - x_0 - \sum_0^{n-1} \gamma(x_j) \right| \\ &\geq 2 | x_0 - x | - 2K - \frac{\lambda}{2}\sqrt{n} \geq \frac{3}{2}\lambda\sqrt{n} - 2K.\end{aligned}$$

Without loss of generality we may assume $\frac{3}{2}\lambda\sqrt{n} - 2K \geq \frac{1}{2}\lambda\sqrt{n}$, since (3.6) is trivial otherwise. Therefore (3.9) is bounded by

$$K^*P\left\{ U_n = x_0, | M_n | \geq \frac{\lambda}{2}\sqrt{n} \mid U_0 = x \right\} \leq 2K^* \exp - C_1 n [c \wedge c^2]$$

(as in (3.7)).

(3.6) follows by substitution of this estimate into (3.8) and taking into account that the sum can be restricted to $x \in A$.

(3.11) COROLLARY. — *Under the hypothesis of Prop. 3.3*

$$P\left\{ | U_n - x_0 | \geq 8K\left[\frac{d}{C_1} n \log n\right]^{1/2} \mid U_0 = x_0 \right\}$$
$$\leq 2(K^* + 1)(2K + 1)^{d+1} \frac{1}{n^{3d}}.$$

(3.12) *Open Problem.* — Can one improve the estimates to obtain under the hypotheses of (3.3)

$$P\{ | U_n - x_0 | \geq \lambda\sqrt{n} \mid U_0 = x_0 \} \to 0 \quad \text{as} \quad \lambda \to \infty,$$

uniformly in w, n and x_0?

Note that such an estimate is known for a nearest neighbor random walk U_n ($K = 1$) under the extra assumption

$$P(x, y) \geq a > 0 \quad \text{for} \quad | x - y | = 1, \qquad x, \qquad y \in \mathbb{Z}^d$$

(cf. [*17*], Theorem 2). Also in the continuous case for self-adjoint *uniformly elliptic* diffusions one has Nash's estimates for the fundamental solution, which are better than (3.6) and again give an affirmative answer to this problem under the extra assumption of uniform ellipticity. ///

We now return to our specific set up. w is an infinite connected subgraph of $\mathbb{Z}^2$ and $U_n = U_n(m; w)$ the imbedded random walk on $b(m; w)$. As in (2.21) we set

$$\Delta(y) = \Delta(y; b(m; w)) = \# \text{ of neighbors of } y \text{ in } b(m; w).$$

(3.13) LEMMA. — $\{ U_n(m; w) \}$ *has transition probabilities*

$$P(y, z) = \frac{1}{d(y; w \cap S(m))} \text{ if } z \text{ is adjacent to } y \text{ on } b(m; w),$$

$$P(y, z) = \frac{d(y; w \cap S(m)) - \Delta(y; b(m; w))}{d(x; w \cap S(m))} \text{ if } z = y, \qquad 0 \text{ otherwise}.$$

This P *is reversible with invariant measure* $\mu(x) = d(x; w \cap S(m))$, $x \in b(m; w)$. *Consequently,*

$$P\{\kappa(m; w) \le C_2 m^2/\log m \mid U_0 = \underline{0}\}$$
$$\le C_3\left(2C_2\frac{m^2}{\log m} + 1\right)^2\left[m\exp\left(-C_1\frac{m}{4}\right) + m^2\exp\left(-C_1\frac{\log m}{16C_2}\right)\right].$$

Proof. — In analogy with (2.22) we define for $y \in b(m; w)$

$\Xi(y) = \{y\} \cup \{z \in w \cap S(m)$: z not adjacent to y but there exists a path r on $w \cap S(m)$ from y to z which contains no point of $b(m; w)$ other than y or $z\} \cup \{z \in w \cap S(m)$: z adjacent to y but $z \notin b(m; w)\}$.

We claim that

$$(3.14) \qquad \Xi(y) \cap b(m; w) = \{y\},$$

so that if $\hat{X}^w_k = y$, $\hat{X}^w_{k+1} = z \in w \cap S(m) \backslash b(m; w)$, then the first return to $b(m; w)$ by $\hat{X}^w_.$ after time k has to be at y (10). To see (3.14) consider a path $z_1, z_2, \ldots, z_l$ on $w \cap S(m)$ with $z_1, \ldots, z_{l-1} \in w \cap S(m) \backslash b(m; w)$, z_1 adjacent to y and $z_l \in b(m; w) \cap S(m)$. To prove (3.14) it suffices to show that $z_l = y$ for any such path. Assume to the contrary that $z_l \neq y$. Since $y \in b(m; w)$ there exist two paths r_1, r_2 on $w \cap S(m)$ connecting y to $\underline{0}$ and $\partial S(m)$, respectively, such that $r_1 \cap r_2 = \{y\}$. Similarly there exist two paths r_3 and r_4 on $w \cap S(m)$ connecting z_l to $\underline{0}$ and $\partial S(m)$, respectively, with $r_3 \cap r_4 = \{z_l\}$. Now assume for the sake of argument that the path $r = (z_1, z_2, \ldots, z_l)$ does not intersect $r_1 \cup r_2$ and that $(r_3 \cup r_4) \cap (r_1 \cup r_2) = \underline{0}$ (see Figure 3*a*).

Then we can connect z_1 to $\partial S(m)$ by a path r'' made up of pieces of r, r_3 and r_4 which are disjoint from $r_1 \cup r_2$. Also we can connect z_1 to $\underline{0}$ by a path r' made up of the edge from z_1 to y and pieces of r_1 and r_2. r' and r'' would intersect only in z_1, contradicting the fact that $z_1 \notin b(m; w)$. Similar arguments work if $r_3 \cup r_4$ intersects $r_1 \cup r_2$ or if r intersects $r_1 \cup r_2$. In the latter case the first intersection of r with $r_1 \cup r_2$ necessarily is z_l itself and we have no need of paths z_3 and z_4 (see Figure 3*b*). Thus the assumption $z_l \neq y$ leads to a contradiction and this establishes (3.14). From here on the proof of the formula for $P(x, y)$ is the same as for (2.20).

(10) It is for this reason that $\Xi(y)$ is called a « dangling end ».

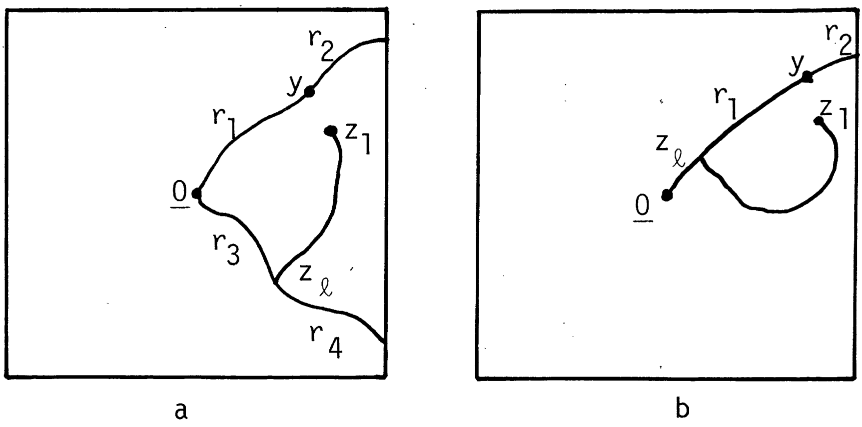

FIG. 3. — Two possible configurations for y, r and the r_i.

Clearly, the formula for P implies

$$d(x)P(x, y) = P(y, x)d(y), \qquad x, y \in b(m; w),$$

so that $\mu(x)$ is indeed an invariant measure and P is reversible with respect to this measure. Finally, the estimate for $\kappa(m; w)$ follows now from (3.6) with $K = 1$, $K^* = 2d$, and w replaced by $b(m; w)$, because for $U_0 = \underline{0}$

$$\{\kappa(m, w) \leq C_2 m^2/\log m\} \subset \bigcup_{n \leq C_2 m^2/\log m} \{|U_n - U_0| \geq m\}. \qquad \square$$

(3.2) is immediate from this lemma, and we now turn to the study of $\mathcal{T}(m)/\kappa(m)$. First we find a replacement for Lemma 2.28. (2.29) will not be used in exactly the same form, but would only require trivial changes. Moreover, since w is not necessarily a tree we have to use a different proof. Also, we do not have a direct analogue of (2.30) and have to replace (2.30) by estimates of a different type in Lemmas 3.18 and 3.20. For $y \in b(m; w)$ set

$$\sigma(n, y) = \inf\{k > n : \hat{X}_k^w = y\}.$$

(3.15) LEMMA. — *For* $y \in b(m; w)$, $x \in w \cap S(m)$, *we have on* $\{\hat{X}_0^w = \underline{0}, \hat{X}_n^w = y\}$.

(3.16) E { number of visits to x by $\hat{X}_\cdot^w$ during

$$[n, \sigma(n, y)) \mid \hat{X}_0^w, \ldots, \hat{X}_n^w\} = \frac{d(x; w \cap S(m))}{d(y; w \cap S(m))}.$$

If there exists a path r on $w \cap S(m)$ *from* y *to* x *of* l *edges, then on* $\{ \hat{X}_0^w = \underline{0}, \hat{X}_n^w = y \}$

(3.17) $P\{ \hat{X}_\cdot^w \text{ visits } x \text{ during } [n, \sigma(n, y)) \,|\, \hat{X}_0^w, \ldots, \hat{X}_n^w \} \geq [ld(y; w \cap S(m))]^{-1}$.

Proof. — Since $\hat{X}^w$ lives on the *finite* component of $w \cap S(m)$ which contains $\underline{0}$ it is recurrent. Therefore, by Theorem XV.11.1 of [*8*] the left hand side of (3.16) is given by the ratio of the invariant measures of x and y. Thus (3.16) is immediate from Lemma 3.13.

(3.17) is best proved by using electrical terminology. It is known (cf. [*6*], Sect. 2.6-2.8, [*9*], Theorem 2.1) that the left hand side of (3.17) equals $[d(y; w \cap S(m))R(x, y)]^{-1}$, where $R(x, y)$ is the resistance between x and y in the network $w \cap S(m)$, when each edge of this network has resistance 1 ohm. Moreover, this resistance can only increase by removing edges from the network. Therefore, we obtain a lower bound for the left hand side of (3.17) by removing all edges from $w \cap S(m)$ which are neither incident to y nor belong to r. Moreover, after this removal the network consists of r plus $(d - 1)$ edges incident to y and the hitting probability of x before returning to y is now given by the right hand side of (3.17) (see [*8*], XIV.2.5). □

Next we define the « local times » for the random walk $\hat{X}^w$ (rather than U as in the last section) by

$$L(x, n) = \text{number of } k \leq n \quad \text{with} \quad \hat{X}_k^w = x\,.$$

We shall show that L is « smooth in the space variable ».

Specifically, we put

$\| x - y \|_{m,w}$ = number of edges in shortest path on $w \cap S(m)$ from x to y. ($= \infty$ if no such path exists).

(3.18) LEMMA. — *For some constants* $0 < C_4, C_5 < \infty$ *and all* $x, y \in w \cap S(m)$ *and* $\lambda \geq 1$

$$P\left\{ \exists n \quad \text{with} \quad L(y, n) \geq \lambda \| x - y \|_{m,w} \quad \text{but} \quad L(x, n) \leq \frac{1}{2}\frac{d(x)}{d(y)} L(y, n) \right\} \leq C_4 \| x - y \|_{m,w} \exp(-\lambda C_5)\,.$$

Proof. — Many similar estimates are known. For the sake of completeness we give a proof of (3.18). Fix y and let $0 \leq \xi_1 < \xi_2 < \ldots$ be the successive times at which $\hat{X}^w_\cdot$ visits y. Thus $\hat{X}^w(\xi_k) = y$, and $L(y, n) = k$ for $\xi_k \le n < \xi_{k+1}$. The random variables

$$M(x, k) := L(x, \xi_{k+1}) - L(x, \xi_k),\ k \ge 1,$$

are i. i. d. If we define

$$\pi(x, z) = P\{\ \hat{X}^w_\cdot \text{ visits } z \text{ before it returns to } x \mid \hat{X}^w_0 = x\ \},$$

then the distribution of $M(x, k)$ can be written as follows

$$\begin{aligned} &P\{\ M(x, k) = 0\ \} = 1 - \pi(y, x),\\ &P\{\ M(x, k) = s\ \} = \pi(y, x)(1 - \pi(x, y))^{s-1}\pi(x, y), \qquad s \ge 1. \end{aligned}$$

Consequently

$$E\{\ M(x, k)\ \} = \frac{\pi(y, x)}{\pi(x, y)}.$$

One easily sees that this expectation is just the same as the one evaluated in (3.16) so that

$$\frac{\pi(y, x)}{\pi(x, y)} = \frac{d(x)}{d(y)}.$$

Also

$$\begin{aligned} Ee^{-\theta M(x,k)} &= 1 - \frac{\pi(y, x)(1 - e^{-\theta})}{\pi(x, y)[1 + (1 - \pi(x, y))(\pi(x, y))^{-1}(1 - e^{-\theta})]}\\ &\le \exp\left(-\frac{d(x)}{d(y)}\frac{1 - e^{-\theta}}{[1 + (\pi(x, y))^{-1}(1 - e^{-\theta})]}\right). \end{aligned}$$

Consequently, for any $\theta \ge 0$

$$\begin{aligned} (3.19)\quad & P\left\{ L(x, \xi_{k+1}) \le \frac{1}{2}\frac{d(x)}{d(y)}(k + 1)\right\}\\ &\le P\left\{\exp\left(-\theta\sum_{j=1}^{k} M(x, j)\right) \ge \exp\left(-\frac{\theta}{2}\frac{d(x)}{d(y)}(k + 1)\right)\right\}.\\ &\le \exp\left(\frac{\theta}{2}\frac{d(x)}{d(y)}\right)\exp\left(k\frac{d(x)}{d(y)}\left[\frac{\theta}{2} - \frac{1 - e^{-\theta}}{1 + (\pi(x, y))^{-1}(1 - e^{-\theta})}\right]\right). \end{aligned}$$

Finally, (3.17) (with x and y interchanged) says

$$\pi(x, y) \ge [d(x)\,\|\,x - y\,\|_{m,w}]^{-1} \ge [4\,\|\,x - y\,\|_{m,w}]^{-1}.$$

We can therefore find a C_6 such that for $\theta = C_6 [\, \| x - y \|_{m,w}]^{-1}$ the expression in square brackets in the last member of (3.19) is at most

$$- \theta/4 = - C_6 [4 \, \| x - y \| \,]^{-1} .$$

We thus obtain

$$P\left\{ L(x, \xi_{k+1}) \le \frac{1}{2} \frac{d(x)}{d(y)} (k+1) \right\} \le C_7 \exp(- kC_6 [16 \, \| x - y \| \,]^{-1}) .$$

The lemma follows by summing this inequality over $k \ge \lambda \, \| x - y \| - 1$. □

So far we have not used the structure of $\tilde{W}$ or $\tilde{B}$. The next lemma contains the only properties of $\tilde{W}$ and $\tilde{B}$ which are needed. Most of the ideas for its proof are already in [*16*]. We therefore show first how to derive Theorem 1.27 from Lemma 3.20, and only then prove the lemma. The following notation and terminology will be used. P_p is defined before (1.24) and ν is the measure of Lemma 1.24. A $\to$ B (A $\to$ B in S) means that there exists an open path on $\mathbb{Z}^2$ (respectively on $\mathbb{Z}^2$ and inside the set S) from a vertex in A to a vertex in B,

$$\pi(n) = P_{\frac{1}{2}} \{ \underline{0} \to [n, \infty) \times \mathbb{R} \},$$

$\rho(n) = \pi_{\frac{1}{2}} \{ \underline{0}$ is connected to $\partial S(n)$ by two open paths on $\mathbb{Z}^2$ which intersect only in $\underline{0} \}$

$\# A =$ number of vertices of $\mathbb{Z}^2$ in A,

$\mathring{R} =$ interior of R, $\overline{R} =$ closure of R.

If $R_1 \subset R_2$ are rectangles whose corners have integer coordinates, then an *open crossing of the annulus* $R_2 \backslash R_1$ is an open path on $\mathbb{Z}^2$ in $\overline{R}_2 \backslash \mathring{R}_1$ which connects ∂R_1 with ∂R_2.

(3.20) LEMMA. — *There exist constants* $0 < C_i, \eta_i < \infty$ *such that*

$$C_1 n^{-1/2} \le \pi(n) \le C_2 n^{-\eta_1}, \tag{3.21}$$

$$(2n)^{-1} \le \rho(n) \le 16 \pi^2(n) \le C_3 \pi(n) n^{-\eta_1} . \tag{3.22}$$

Let

$$D(\underline{j}) = D(\underline{j}, q) = [j_1 q, (j_1 + 1)q) \times [j_2 q, (j_2 + 1)q),$$
$$F(\underline{j}) = F(\underline{j}, q) = [(j_1 - 1)q, (j_1 + 2)q] \times [(j_2 - 1)q, (j_2 + 2)q] .$$

Then for every $t \geq 1$ *there exists a constant* C_t *such that for all* $\underline{j}=(j_1, j_2) \in \mathbb{Z}^2$, $1 \leq q \leq m$, $\varepsilon > 0$,

$$\nu \{ \# (\tilde{\mathrm{B}}(m) \cap \mathrm{F}(\underline{j})) \geq q^{2+\varepsilon}\rho(q) \} \leq \frac{C_t}{q^{t\varepsilon}} \tag{3.23}$$

and

(3.24) $\nu \{ \exists$ open crossing r of $\mathrm{F}(\underline{j}) \backslash \mathrm{D}(\underline{j})$ in $\mathrm{S}(m)$, but the number of vertices $x \in \mathrm{F}(\underline{j})$ such that $x \to r$ in $\mathrm{S}(m) \cap \mathrm{F}(\underline{j}) \backslash \mathrm{D}(\underline{j})$ is less than $q^{2-\varepsilon}\pi(q) \}$ $\leq C_1 \exp(-C_2 q^{\varepsilon/8})$.

Proof of Theorem 1.27. — Let $X(0) = \underline{0}$, $X(1), \ldots, X(\mathcal{T}(m))$ be the path of the random walk on $\tilde{W}$ from $\underline{0}$ to $\partial S(m)$. For a given q, much smaller than m, we now choose two finite sequences $\underline{j}_0, \ldots, \underline{j}_\lambda$ and $l_0, \ldots, l_\lambda$. These are chosen such that $0 = l_0 < l_1 < \ldots < l_\lambda \leq \mathcal{T}(m)$ and such that $\underline{j}_i$ is the unique point of $\mathbb{Z}^2$ for which

$$X(l_i) \in D(\underline{j}_i, q) . \tag{3.25}$$

As indicated above, we start with $l_0 = 0$. (3.25) then gives us $\underline{j}_0 = \underline{0}$, and once $l_0, \ldots, l_i, \underline{j}_0, \ldots, \underline{j}_i$ have been found, we take

$$l_{i+1} = \min \{ l > l_i : X(l) \in \partial F(\underline{j}_i, q) \} , \tag{3.26}$$

provided such an l exists. If no such l exists then we stop. Let l_λ be the last l which we choose. Then by definition

$$X(l) \in F(\underline{j}_\lambda, q), \qquad l \geq l_\lambda . \tag{3.27}$$

Also

$$X(l) \in F(\underline{j}_i, q), \qquad l_i \leq l \leq l_{i+1} . \tag{3.28}$$

Now define for each $i \in [0, \lambda]$

$$\mathscr{C}_i = \text{component of } \tilde{W} \cap F(\underline{j}_i, q) \cap S(m) \text{ which contains } X(l_i),$$

where x, y belong to the same component of a set S if there exists a path on $\mathbb{Z}^2$ from x to y with all its vertices in S. By (3.27) and (3.28) each $X(l)$ belongs to some $\mathscr{C}_i$ with $0 \leq i \leq \lambda$. Of course $\mathscr{C}_{i_1}$ may equal $\mathscr{C}_{i_2}$ for some $i_1 \neq i_2$. In order not to count such components more than once we define for $0 \leq i \leq \lambda$

$$f(i) = \text{smallest index } \imath \text{ for which } \underline{j}_i = \underline{j}_\imath \text{ and } \mathscr{C}_i = \mathscr{C}_\imath .$$

We also define

$$\Lambda(\imath) = \sum_{x \in \mathscr{C}_i} L(x, \mathcal{T}(m)) \qquad \text{for any } i \text{ with } f(i) = \imath .$$

and

$$\Theta(\iota) = \sum_{x \in \tilde{B}(m) \cap \mathscr{C}_i} L(x, \mathscr{T}(m)) \qquad \text{for any } i \text{ with } f(i) = \iota .$$

We claim that

$$1 + \mathscr{T}(m) = \sum_x L(x, \mathscr{T}(m)) \geq \frac{1}{16} \sum_\iota \Lambda(\iota) , \tag{3.29}$$

where the sum over ι runs over those ι which equal $f(i)$ for some i. Indeed, any x for which $L(x, \mathscr{T}(m)) > 0$ must be visited by $X_.$ during $[0, \mathscr{T}(m)]$ and hence during some interval $[l_i, l_{i+1})$ $(l_{\lambda+1} = \infty)$. For such an i $x \in \mathscr{C}_i$ since $X(l_i), \ldots, X(l_{i+1})$ automatically belongs to the component $\mathscr{C}_i$ of $\tilde{W} \cap F(\underline{j}_i, q)$ (by (3.27), (3.28)). Thus, each $L(x, \mathscr{T}(m))$ is counted at least once in $\Sigma\, \Lambda(\iota)$. Since any x belongs to $F(\underline{j}, q)$ for at most 16 distinct $\underline{j}$'s, no x can be counted more than 16 times in $\Sigma_\iota \Lambda(\iota)$. This implies (3.29). Note that

$$X(\mathscr{T}(m)) \in \partial S(m) \cap \tilde{W} \subset \tilde{B}(m) .$$

Therefore $\mathscr{T}(m)$ is one of the $\tau_k(m; \tilde{W})$ and in fact must equal $\tau_{\kappa(m)}$. Since $1 + \kappa(m)$ counts all visits to $\tilde{B}(m)$ until $X.$ hits $\partial S(m)$ we have

$$1 + \kappa(m) = \sum_{x \in \tilde{B}(m)} L(x, \mathscr{T}(m)) \leq \sum_\iota \Theta(\iota) .$$

Thus,

$$\frac{1 + \mathscr{T}(m)}{1 + \kappa(m)} \geq \frac{1}{16} \frac{\Sigma_\iota \Lambda(\iota)}{\Sigma_\iota \Theta(\iota)} . \tag{3.30}$$

In order to find a lower bound for the right hand side of (3.30) we shall restrict ourselves to the intersection of a number of events, each of which has high probability. To be specific, take

$$E_1 = \{ \kappa(m) > C_2 m^2 / \log m \}$$

with $C_2 = C_1/128$ and C_1 as in (3.6),

$$E_2 = \{ \# \tilde{B}(m) \leq m^{2+\varepsilon} \rho(m) \} ,$$

$E_3 = \left\{ \frac{1}{8} \leq \frac{L(x, \mathscr{T} m))}{L(X(l_\iota), \mathscr{T}(m)} \leq 8 \right.$ for any ι which is an $f(i)$ and has $L(X(l_\iota),$ $\mathscr{T}(m)) \geq 400 C_5^{-1} q^2 \log m$, and for all $x \in \mathscr{C}_\iota \cup \mathscr{C}_{\iota-1} \}$.

$E_4 = \{ L(x, \mathscr{T}(m) \leq 3\,200\, C_5^{-1} q^2 \log m$ for all $x \in \mathscr{C}_\iota$ with $L(X(l_\iota),$ $\mathscr{T}(m)) \leq 400 C_5^{-1} q^2 \log m$ and ι equal to some $f(i) \}$.

Finally we take

$$C_6 = (C_2C_5/6\,400)^{1/2}, \qquad \varepsilon \leq \frac{1}{8}\eta_1 \quad \text{and} \quad q = \left[\frac{C_6}{\log m} \{ m^{\varepsilon} \rho(m) \}^{-1/2} \right],$$

$E_5 = \underset{\smile}{v} \{ \# (\tilde{B}(m) \cap F(\underline{j}, q)) < q^{2+\varepsilon}\rho(q)$ for all $\underline{j}$ $\}$,

$E_6 = \{$ For any $\underline{j}$ with $F(\underline{j}, q) \cap S(m) \neq \emptyset$ and any open crossing r of $F(\underline{j}, q)\backslash D(\underline{j}, q)$ by $X_0, \ldots, X_{\mathcal{T}(m)}$, the number of vertices x such that $x \to r$ in $F(\underline{j}, q) \cap S(m)$ is $\geq q^{2-\varepsilon}\pi(q)$ $\}$.

Our first task is to show that $\mathbb{P}\{ E_k^c \} \to 0$ as $m \to \infty$ for $1 \leq k \leq 6$. $\mathbb{P}\{ E_1^c \} \to 0$ by Lemma 3.13 (condition on $\tilde{W}$), and $\mathbb{P}\{ E_2^c \} \to 0$ by (3.23) with $\underline{j} = 0$ and $q = m$. Furthermore by (3.21), (3.22) and $\varepsilon \leq \eta_1$ we have

$$C_7 \frac{m^{\frac{1}{2}\eta_1}}{\log m} \leq q \leq o(m^{1/2}).$$

Therefore $\mathbb{P}\{ E_5^c \} \to 0$ and $\mathbb{P}\{ E_6^c \} \to 0$ by (3.23) and (3.24) respectively, if we take into account that $\tilde{B}(m) \cap F(\underline{j}, q) \subset S(m)$, and that $F(\underline{j}, q)$ intersects $S(m)$ for at most $(2m/q) + 4)^2$ values of $\underline{j}$. We prove further that $\mathbb{P}\{ E_3^c \} \to 0$ and leave $\mathbb{P}\{ E_4^c \}$ to the reader. Note that if $x \in \mathcal{C}_\iota \cup \mathcal{C}_{\iota-1}$ then x can be connected to $X(l_\iota)$ by a path on $\tilde{W} \cap (F(\underline{j}_{\iota-1}, q) \cup F(\underline{j}_\iota, q))$. For any $x \in \mathcal{C}_\iota$ this is clear from the definition of $\mathcal{C}_\iota$. But also, if $x \in \mathcal{C}_{\iota-1}$, then x can be connected to $X(l_{\iota-1})$ by a path on $\tilde{W} \cap F(j_{\iota-1}, q)$, and $X(l_{\iota-1})$ is connected to $X(l_\iota)$ by the path $X(l_{\iota-1}), X(l_{\iota-1}+1), \ldots, X(l_\iota) \subset \tilde{W} \cap F(\underline{j}_{\iota-1}, q)$. In view of these facts for any $x \in \mathcal{C}_\iota \cup \mathcal{C}_{\iota-1}$ we have

$$\| x - X(l_\iota) \|_{m,w} \leq \# (F(\underline{j}_{\iota-1}, q) \cup F(\underline{j}_\iota, q)) \leq (6q+1)^2.$$

Assume now that E_3^c occurs. There then exists a y of the form $X(l_\iota)$ and an x with $\| x - y \|_{m,\tilde{W}} \leq (6q+1)^2$, $x, y \in S(m)$ and

$$\text{(3.31)} \qquad L(x, \mathcal{T}(m)) \leq \frac{1}{2}\frac{d(x)}{d(y)} L(y, \mathcal{T}(m)), \quad \text{and}$$

$$L(y, \mathcal{T}(m)) \geq 400 C_5^{-1} q^2 \log m \geq 8 C_5^{-1} (\log m) \| x - y \|_{m,\tilde{W}}$$

or

$$\text{(3.32)} \qquad L(y, \mathcal{T}(m)) \leq \frac{1}{2}\frac{d(y)}{d(x)} L(x, \mathcal{T}(m)), \quad \text{and}$$

$$L(x, \mathcal{T}(m)) \geq 8 L(y, \mathcal{T}(m)) \geq 8 C_5^{-1} (\log m) \| x - y \|_{m,\tilde{W}}.$$

$\left(\text{Here we used that } 1 \le d(z) \le 4 \text{ for all } z \in \tilde{W} \text{ so that } \frac{1}{8} \le d(y)(2d(x))^{-1} \text{ and } \frac{1}{8} \le d(x)(2d(y))^{-1}.\right)$ By (3.18), and the obvious estimate $\| x-y \| \le \# S(m)$, the conditional probability, given $\tilde{W} \cap S(m)$, that (3.31) or (3.32) occurs for any pair $x, y \in S(m)$ is at most

$$2(\# S(m))^3 C_4 \exp(-8 \log m) = 0(m^{-2}).$$

Thus, indeed $\mathbb{P}\{E_3^c\} \to 0$, and similarly for $\mathbb{P}\{E_4^c\}$.

Now consider a sample point in $\bigcap_1^6 E_k$. Let Σ^* denote the sum over those ι which equal some $f(i)$ and with

$$L(X(l_\iota), \mathcal{T}(m)) \ge 400 C_5^{-1} q^2 \log m.$$

Then, since we are on $E_1 \cap E_2 \cap E_4$, and by the choice of q,

$$\left(\sum_\iota - \sum^*\right)\Theta(\iota) \le (3\,200 C_5^{-1} q^2 \log m) \,\#\, B(m)$$

$$\le 3\,200\, C_5^{-1} q^2 (\log m) m^{2+\varepsilon} \rho(m) \le \frac{1}{2}\kappa(m).$$

Thus, by (3.29)

$$\frac{1+\mathcal{T}(m)}{1+\kappa(m)} \ge \frac{1}{32}\frac{\Sigma^*\Lambda(\iota)}{\Sigma^*\Theta(\iota)}.$$

Consider an ι which occurs in Σ^* and satisfies $\iota < \lambda$. Since we are on E_3

$$\Lambda(\iota) \ge \frac{1}{8} L((l_\iota), \mathcal{T}(m)). \,\#\, \mathscr{C}_\iota$$

and

$$\Theta(\iota) \le 8 L(X(l_\iota), \mathcal{T}(m)). \,\#\, (\tilde{B}(m) \cap \mathscr{C}_\iota).$$

In addition, $\mathscr{C}_\iota$ contains the points $X(l_\iota), \ldots, X(l_{\iota+1})$, which, in turn, contain an open crossing of $F(\underline{j}_\iota, q) \backslash D(\underline{j}_\iota, q)$. Since we are on E_6 we therefore have $\# \mathscr{C}_\iota \ge q^{2-\varepsilon}\pi(q)$. Also, on E_5,

$$\#(\tilde{B}(m) \cap \mathscr{C}_\iota) \le \#(\tilde{B}(m) \cap F(\underline{j}_\iota, q)) \le q^{2+\varepsilon}\rho(q).$$

Thus, for any $\iota < \lambda$ which occurs in Σ^* we have

$$\frac{\Lambda(\iota)}{\Theta(\iota)} \ge \frac{1}{64}\frac{\#\mathscr{C}_\iota}{\#(\tilde{B}(m) \cap \mathscr{C}_\iota)} \ge \frac{1}{64}\frac{\pi(q)}{q^{2\varepsilon}\rho(q)} \ge [2^{10} q^{2\varepsilon}\pi(q)]^{-1} \quad \text{(see (3.22))}.$$

A small change is needed in the estimate for $\Lambda(\imath)$ if $\imath = \lambda$ is an $f(i)$ which occurs in Σ^*. Indeed, we are not sure that $X(l_\lambda), \ldots, X(\mathscr{T}(m))$ contains an open crossing of $F(j_\lambda, q)\backslash D(j_\lambda, q)$. In this case we therefore use the fact that

$$1 + \mathscr{T}(m) \geq \sum_{x \in \mathscr{C}_{\lambda-1}} L(x, \mathscr{T}(m)) \geq \frac{1}{8} L(X(l_\lambda), \mathscr{T}(m)) . \# \mathscr{C}_{\lambda-1}$$
$$\geq \frac{1}{8} L(X(l_\lambda), \mathscr{T}(m)) q^{2-\varepsilon} \pi(q) \geq \frac{1}{64} q^{-2\varepsilon} \frac{\pi(q)}{\rho(q)} \Theta(\lambda) .$$

It follows that

$$\frac{1 + \kappa(m)}{1 + \mathscr{T}(m)} \leq 2^{11} q^{2\varepsilon} \frac{\rho(q)}{\pi(q)} + 32 \frac{\sum^*_{\imath < \lambda} \Theta(\imath)}{\sum^*_{\imath < \lambda} \Lambda(\imath)} \leq 2^{12} q^{2\varepsilon} \frac{\rho(q)}{\pi(q)} \leq C_7 q^{2\varepsilon} \pi(q) .$$

On E_1 this finally implies

$$1 + \mathscr{T}(m) \geq C_2 C_7^{-1} q^{-2\varepsilon} m^2 [\pi(q) \log m]^{-1}$$
$$\geq C_8 q^{-2\varepsilon + \eta_1} m^2 [\log m]^{-1} \quad \text{(by (3.21))}$$
$$\geq C_9 m^{2 + \frac{1}{2}\eta^1 - \varepsilon\eta_1} [\log m]^{-2} \quad \text{(by choice of } q) .$$

Since we already proved

$$\mathbb{P}\left\{ \bigcup_1^6 E_k^c \right\} \to 0$$

this implies (3.1) with $\varepsilon_1 = \frac{1}{4} \eta_1^2$, and hence Theorem 1.27.

Proof of Lemma 3.20. — Much of this is implicit in [*16*], and for part of the proof we shall merely direct the reader to the relevant parts of [*16*]. Note that the fundamental hypothesis (1) of [*16*] holds with $\Lambda = 0$ for bond percolation on $\mathbb{Z}^2$, as explained in [*16*]. It was also explained in [*16*] (see (5) and (13)) that (3.21) and (3.22) follow from [*15*], Lemma 8.5, and [*28*], Cor. 3.15. Only the left hand inequality in (3.22) was not mentioned in [*16*]. However, this follows from (See [*25*], p. 234)

$$\frac{1}{2} = P_{\frac{1}{2}} \{ \exists \text{ open path in } [0, 2n] \times [0, 2n-1] \text{ from the left to the right edge} \}$$
$$\leq \sum_{k=0}^{n-1} P_{\frac{1}{2}} \{ (n, k) \text{ has two open connections to } \partial S((n, k), n) \text{ which intersect only in } (n, k) \} = n\rho(n) .$$

Next we prove (3.23). Since $\tilde{B}(m) \subset S(m)$, by definition, we have $\tilde{B}(m) \cap F(\underline{j}) = \tilde{B}(m) \cap F(\underline{j}) \cap S(m)$. Since (3.23) only becomes stronger if $F(\underline{j})$ is replaced by a larger square and since $q^2\pi(q) \geq C_1 m^2 \pi(m/3)$ for some $C_1 > 0$ if $q \geq m/3$ (by (7) in [*16*]), we may replace $F(\underline{j})$ by some $(3q \wedge m) \times (3q \wedge m)$ square F contained in $S(m)$. First consider the case with $\underline{0} \notin F \subset S(m)$. Then any $x \in \tilde{B}(m) \cap F$ has to have two disjoint open connections to $\underline{0}$ and $\partial S(m)$ respectively, and pieces of these connections will form twó connections to ∂F. Thus, if $\underline{0} \notin F$

$$\tilde{B}(m) \cap F \subset K_F := \{ x \in F : \exists \text{ two open paths in F from } x \text{ to } \partial F \text{ which only have the point } x \text{ in common} \}.$$

If $\underline{0} \in F$, then a similar argument shows

$$\tilde{B}(m) \cap F \subset K_F \cup \tilde{B}(3q \wedge m).$$

It was stated in [*16*], that the proof of Theorem 8 also works for Theorem 14, so that (recall that (6) and (7) in [*16*] also hold with π replaced by ρ)

$$\nu \{ [\# B(3q \wedge m)]^t \} \leq C'_t [(3q \wedge m)^2 \rho(3q \wedge m)]^t \leq C''_t [q^2 \rho(q)]^t.$$

Moreover, if one replaces the definition of G_i in the argument following (43) in [*16*] by

$G_i(\underline{v}) = \{ v_i$ is connected to ∂F by two open paths in F which intersect only in $v_i \}$,

then the argument from (43) to (53) gives also

$$\nu \{ [\# K_F]^t \} \leq C''_t [q^2 \rho(q)]^t.$$

Thus

$$\nu \{ [\# \tilde{B}(m) \cap F(\underline{j})]^t \} \leq C_t [q^2 \rho(\rho)]^t$$

and

$$\nu \{ \# \tilde{B}(m) \cap F(\underline{j}) \geq q^{2+\varepsilon} \rho(q) \} \leq [q^{2+\varepsilon} \rho(q)]^{-t} \nu \{ \# \tilde{B}(m) \cap F(\underline{j})]^t \} \leq \frac{C_t}{q^{t\varepsilon}}.$$

This leaves us with (3.24) to prove. First we distil the relevant facts from [*16*]. For a vertex v of $\mathbb{Z}^2$ and path r from $\partial S(v, n)$ to $\partial S(v, 27n)$ inside $S(v, 27n) \backslash \mathring{S}(v, n)$ let

$$A(v, n) = \text{annulus } S(v, 3n) \backslash S(v, n),$$
$$Z(v, n, r) = \# \{ x \in A(v, 3n) : x \to r \text{ in } S(v, 27n) \backslash \mathring{S}(v, n) \},$$
$$Y(v, n) = \# \{ x \in A(v, 3n) : x \to \partial S(v, n) \cup \partial S(v, 27n) \text{ in } S(v, 27n) \backslash \mathring{S}(v, n) \}.$$

(see Figure 4). The argument after (57) in [*16*] can be used with out signifi-

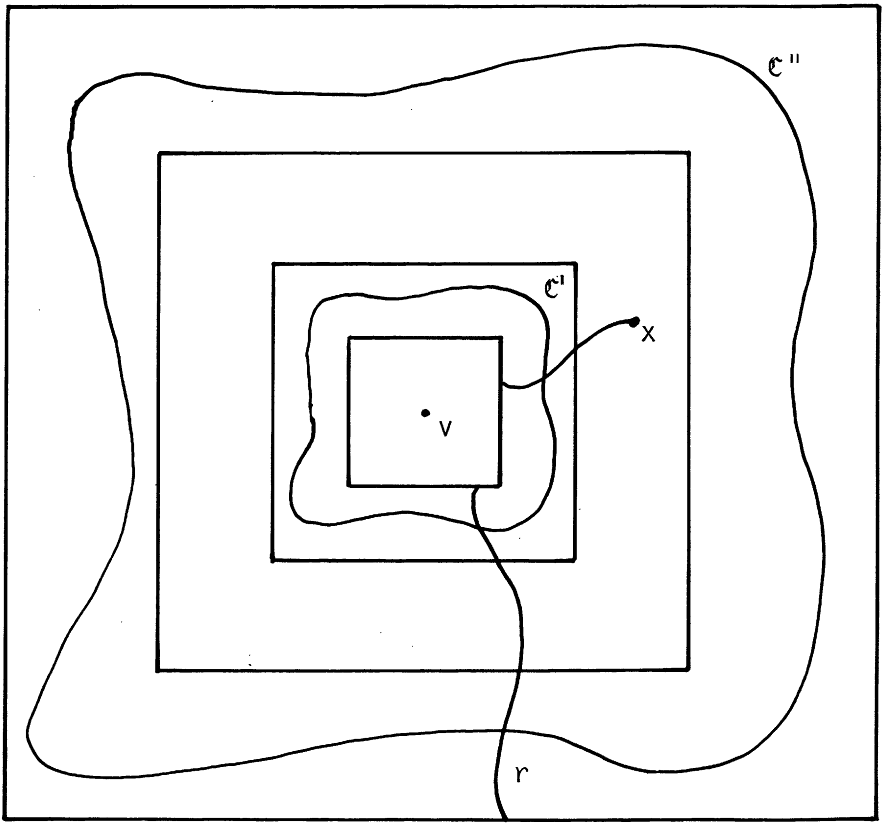

FIG. 4. — The squares (from the inside to the outside) are $S(v, n)$, $S(v, 3n)$, $S(v, 9n)$ and $S(v, 27n)$.

cant changes to show that there exist constants $C_1, C_2 > 0$ such that uniformly in r

$$
\begin{aligned}
& P_{\frac{1}{2}}\{ Z(v, n, r) \geq C_1 n^2 \pi(n) \} \\
& \quad \geq P_{\frac{1}{2}}\{ \exists \text{ open circuits } \mathscr{C}' \text{ in } A(v, n) \text{ and } \mathscr{C}'' \text{ in } A(v, 9n) \text{ and } Y(v, n) \\
& \qquad \geq C_1 n^2 \pi(n) \} \geq C_2 .
\end{aligned}
$$

As a consequence for any path r from $\partial S(v, n^{(1-\delta)})$ to $S(v, n)$ inside $S(v, n) \backslash \overset{\circ}{S}(v, n^{(1-\delta)})$

$$
\begin{aligned}
(3.33) \quad & P_{\frac{1}{2}}\{ \# \{ x \in S(v, n) \backslash S(v, n^{(1-\delta)}) : x \to r \text{ in } S(v, n) \backslash \overset{\circ}{S}(v, n^{(1-\delta)}) \} \\
& \geq C_1 n^{2(1-\delta)} \pi(n) \}
\end{aligned}
$$

$$\geq P_{\frac{1}{2}}\{\exists \text{ open circuits in } A(v, 3^{3j}) \text{ and } A(v, 3^{3j+2}) \text{ and } Y(v, 3^{3j}) \geq C_1 3^{6j}\pi(3^{3j}) \text{ for some } n^{1-\delta} \leq 3^{3j} \leq n/27\} \geq 1-(1-C_2)^{C_3\delta\log n}.$$

This argument from [*16*] will be used below. The remaining part of the proof, which is new, will be broken up into 3 steps. First we prove an analogue of (3.33) but now not for a fixed path r, but for the random collection of all open crossings of a rectangle. In step (*ii*) this estimate is improved to give (3.24) but with ν replaced by $P_{\frac{1}{2}}$. The last step shows that taking the ν-measure and not the $P_{\frac{1}{2}}$-measure in (3.24) makes little difference.

STEP *i*). — Let $\bar{J} := [0, t] \times [0, 3t]$ and consider the family $\mathscr{R}$ of all open paths r which connect the left and right edge of $\bar{J}$ and lie in $\mathring{J}$ (= interior of $\bar{J}$), except for their initial and endpoints. For any such r we define $J^+(r)(J^-(r))$ as the component of $\mathring{J}\backslash r$ above (below) r, i. e., the component with $[0, t] \times \{3t\}([0, t] \times \{0\})$ as part of its boundary (see Figure 5).

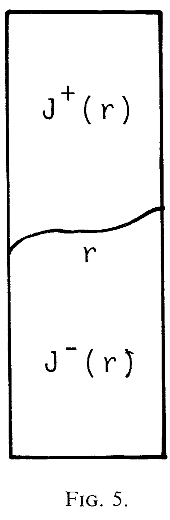

FIG. 5.

If $\mathscr{R} \neq \phi$, then there exists a lowest crossing R_1 in $\mathscr{R}$, i. e., a crossing R_1 for which $J^-(R_1)$ is minimal. (See [*14*], Lemma 1 or [*15*], Prop. 2.3). We can successively define the k-th lowest crossing R_k, $k > 1$, as the lowest crossing in $\bar{J}\backslash(J^-(R_{k-1}) \cup R_{k-1})$, i. e. the crossing r with minimal $J^-(r)$ among all $r \in \mathscr{R}$, $r \in \bar{J}\backslash(J^-(R_{k-1}) \cup R_{k-1})$. Note that by definition R_k has to be disjoint from R_{k-1}, and in fact with the exception of its endpoints $R_k \subset J^+(R_{k-1})$, i. e., R_k lies above R_{k-1}. It is not hard to see that

$$J^-(R_k) \supset J^-(R_{k-1}) \supset \dots \supset J^-(R_1).$$

Thus all the R_i have to be disjoint. Therefore, if K is the highest number k for which we can find R_k, then by [28], Theorem 3.3,

$$(3.34)\quad P_{\frac{1}{2}}\{K \ge k\} \le P_{\frac{1}{2}}\{\exists k \text{ disjoint open crossings from left to right in } \bar{J}\}$$
$$\le [P_{\frac{1}{2}}\{\exists \text{ at least one open crossing from left to right in } \bar{J}\}]^k \le (1-C_4)^k$$

for some $C_4 > 0$. Note that $C_4 > 0$ because the probability that there exists an open left right crossing of $\bar{J}$ is at most

$$1 - P_{\frac{1}{2}}\{\exists \text{ a closed crossing on the dual of } \mathbb{Z}^2 \text{ from the bottom to the top of } \left[-\frac{1}{2}, t+\frac{1}{2}\right] \times \left[\frac{1}{2}, 3t-\frac{1}{2}\right]\} < 1$$

uniformly in t (compare [25], Theorem 4.1). Next observe that any $r \in \mathcal{R}$ must intersect one of the R_i, $1 \le i \le K$. Therefore, if

$$Z^*(t, r) := \{x \in [-t, t] \times [-t, 4t] : x \to r \quad \text{in} \quad [-t, t] \times [-t, 4t]\},$$

then

$$(3.35)\quad P_{\frac{1}{2}}\{\exists r \in \mathcal{R} \quad \text{with} \quad Z^*(t, r) \le C_1 t^{2(1-\delta)} \pi(t)\}$$
$$\le \sum_{k=1}^{\infty} P_{\frac{1}{2}}\{K \ge k \quad \text{and } Z^*(t, R_k) \le C_1 t^{2(1-\delta)} \pi(t)\}.$$

To estimate the k-th summand in the right hand side we write it as

$$(3.36)\quad \sum_r P_{\frac{1}{2}}\{K \ge k \quad \text{and} \quad R_k = r \quad \text{and} \quad Z^*(t, r) \le C_1 t^{2(1-\delta)} \pi(t)\}.$$

Next we observe that for a fixed left-right crossing r of $\bar{J}$, the event $\{K \ge k, R_k = r\}$ depends only on the edges in the closure of $J^-(r)$. This is explained in [14], p. 45, 46 or [15], Prop. 2.3 for $k=1$. Essentially the same argument works for any k. Therefore the general term in (3.36) may also be written as

$$(3.37)\quad E_{\frac{1}{2}}\{P_{\frac{1}{2}}\{Z^*(t, r) \le C_1 t^{2(1-\delta)} \pi(t) \mid J^-(r)\}; K \ge k, R_k = r\},$$

where $E_{\frac{1}{2}}\{\,.\,|\,J^-(r)\,\}$ denotes conditional expectation with respect to the σ-field generated by the edges in the closure of $J^-(r)$. We claim that the method for (3.33) also gives

$$(3.38)\quad P_{\frac{1}{2}}\{\,Z^*(t,r) < C_1 t^{2(1-\delta)}\pi(t)\,|\,J^-(r)\,\} \le (1 - C_5)^{C_3\delta\log t} \quad \text{on}\quad \{\,K \ge k,\ R_k = r\,\}.$$

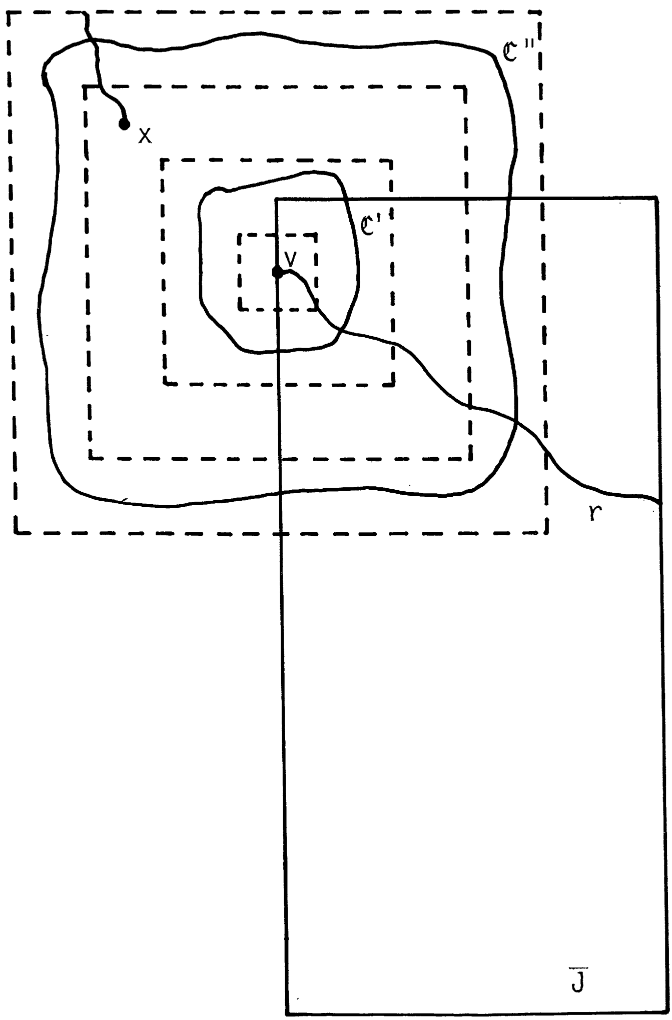

FIG. 6. — $\mathscr{C}'_j$ is a circuit in the annulus $A(v, 3^{3j})$; $\mathscr{C}''_j$ is a circuit in $A(v, 3^{3j+2})$. X lies in $A(v, 3^{3j+1}) \cap ([-t, -3^{3j+1}] \times \mathbb{R}$.

To see this, let v be the endpoint of r on the left edge $\{0\} \times [0, 3t]$ and draw the annuli $A(v, 3^k)$, $t^{1-\delta} \leq 3^k \leq t/27$. Assume there exist circuits $\mathscr{C}'_j$ in $A(v, 3^{3j})$ and $\mathscr{C}''_j$ in $A(v, 3^{3j+2})$ such that all edges of $\mathscr{C}'_j$ and $\mathscr{C}''_j$ which do not lie in the closure of $J^-(r)$ are open. Then for $3^{3j} \leq t/27$ any vertex $x \in A(v, 3^{3j+1}) \cap ([-t, 0) \times \mathbb{R})$ which is connected to $\partial S(v, 3^{3j}) \cup \partial S(v, 3^{3j+1})$ by an open path in $(S(v, 3^{3j+3}) \backslash \mathring{S}(v, 3^{3j})) \cap ([-t, 0) \times \mathbb{R})$ contributes to $Z^*(t, r)$ (see Figure 6).

Now the conditional probability of the existence of $\mathscr{C}'_j$ and $\mathscr{C}''_j$ as above, given the configuration in the closure of $J^-(r)$ is at least as much as

$$P_{\frac{1}{2}}\{\exists \text{ open circuit } \mathscr{C}'_j \text{ in } A(v, 3^{3j}) \text{ and an open circuit } \mathscr{C}''_j \text{ in } A(v, 3^{3j+2})\} \geq C_6 > 0.$$

(This can be seen by simply making a new and independent choice of the configuration in the closure of $J^-(r)$ and then requiring that $\mathscr{C}'_j$ and $\mathscr{C}''_j$ have all their edges open when we use the new configuration in the closure of $J^-(r)$; for the last inequality use (28) of [16].) Also, it is not hard to see that for any $x \in A(v, 3^{3j+1}) \cap ([-t, -3^{3j+1}] \times \mathbb{R})$

$$P_{\frac{1}{2}}\{x \mapsto \partial S(v, 3^j) \cap \partial S(v, 3^{3j+3}) \text{ in } (S(v, 3^{3j+3}) \backslash S(v, 3^{3j})) \cap [-t, 0) \times \mathbb{R} \mid J^-(r),$$
$$\exists \mathscr{C}'_j \text{ in } A(v, 3^{3j})) \text{ and } \mathscr{C}'' \text{ in } A(v, 3^{3j+2}) \text{ with only open edges outside the closure of } J^-(r)\} \geq C_7 \pi(3^{3j}),$$

because the configuration in $[-t, 0) \times \mathbb{R}$ is independent of the configuration in $\bar{J}$, and by the FKG inequality. We can now repeat the argument for (3.33) (including the argument following (54) in [*16*]) to obtain that the left hand side of (3.38) is at most

$$P_{\frac{1}{2}}\{\text{for each } j \text{ with } t^{1-\delta} \leq 3^{3j} \leq t/27 \text{ either there is no } \mathscr{C}'_j \text{ or no } \mathscr{C}''_j \text{ as above,}$$
$$\text{or the number of } x \in A(v, 3^{j+1}) \cap ([-t, 0) \times \mathbb{R}) \text{ with } x \to \partial S(v, 3^j) \cup \partial S(v, 3^{j+3})$$
$$\text{is } < C_1 3^{6j} \pi(3^j) \mid J^-(r)\} \leq (1 - C_5)^{C_3 \delta \log t}.$$

This proves (3.38). Finally, substitution of this estimate in (3.34)-(3.37) gives

$$(3.39) \quad P_{\frac{1}{2}}\{\exists r \in \mathscr{R} \text{ with } Z^*(t, r) \leq C_1 t^{2(1-\delta)} \pi(t)\}$$
$$\leq \sum_k \sum_r P_{\frac{1}{2}}\{K \geq k, R_k = r\} (1 - C_5)^{C_3 \delta \log t}$$
$$\leq (1 - C_5)^{C_3 \delta \log t} \sum_k P_{\frac{1}{2}}\{K \geq k\} \leq C_4^{-1} t^{-C_8 \delta}.$$

This concludes the first step.

STEP *ii*). — In order to prove (3.24) we need an estimate which goes to zero faster than the small power of t in the right hand side of (3.39). To obtain this we apply the block-technique of the beginning of the proof of Theorem 1.27 and a Peierls argument. Assume that r is an open crossing of $F(\underline{j}, q)\backslash D(\underline{j}, q)$ for some fixed $\underline{j}$ and that the center of $D(\underline{j}, q)$ is

$$v = \left(\left(j(1) + \frac{1}{2}\right)q, \left(j(2) + \frac{1}{2}\right)q\right)$$

(where $\underline{j} = (j(1), j(2))$. Then r connects $\partial D(\underline{j}, q) = S(v, q/2)$ with

$$\partial F(\underline{j}, q) = S(v, 3q/2).$$

We choose

$$\delta = \frac{\varepsilon}{4}, \qquad t = q^{\frac{2-\varepsilon}{2-3\delta}} < q$$

r contains a crossing $\bar{r}$ of $S(v, 3q/2 - 5t)\backslash S(v, q/2 + 5t)$. Let $x(0), \ldots, x(\xi)$ be the vertices of $\bar{r}$ with $x(0) \in \partial S(v, q/2+5t) =$ boundary of $[j(1)q - 5t, (j(1)+1)q+5t] \times [j(2)q-5t, (j(2)+1)q+5t]$, $x(\xi) \in \partial S(v, 3q/2-5t) =$ boundary of $[(j(1)-1)q+5t, (j(1)+2)q-5t] \times [(j(2)-1)q+5t, (j(2)+2)q-5t]$ and $x(i) \in \mathring{S}(v, 3q/2-5t)\backslash S(v, q/2+5t)$ for $1 \le i < \xi$. We then pick $\underline{j}_0, \ldots, \underline{j}_\lambda$, $l_0, \ldots, l_\lambda$ as in (3.25)-(3.28) with q replaced by t and X by x. Thus we have $l_0 = 0$,

(3.40) $\quad x(l_i) \in D(\underline{j}_i, t), \ 0 \le i \le \lambda, \ l_{i+1} = \min\{ l > l_i : x(l) \in \partial F(\underline{j}_i, t)\},$

(3.41) $\quad x(l) \in F(\underline{j}_\lambda, t), \ l \ge l_\lambda, \qquad x(l) \in F(\underline{j}_i, t), \ l_i \le l \le l_{i+1}, \ 0 \le i < \lambda.$

From these definitions we have

$$x(l_{i+1}) \in \partial F(\underline{j}_i, t) \cap D(\underline{j}_{i+1}, t), \tag{3.42}$$

which combined with (3.40) shows

$$| x(l_{i+1}) - x(l_i) | \le 2t. \tag{3.43}$$

On the other hand, since $x(0) \in \partial S(v, q/2 + 5t)$, $x(\xi) \in \partial S(v, 3q/2 - 5t)$ we have from (3.43) and (3.41)

$$q - 10t \le | x(\xi) - x(0) | \le \sum_{i=0}^{\lambda-1} | x(l_{i+1}) - x(l_i) | + | x(\xi) - x(l_\lambda) | \le 2\lambda t + 2t.$$

Consequently, for large q

$$\lambda \ge \frac{q}{2t} - 6 \ge \frac{1}{4} q^{\delta/2}. \tag{3.44}$$

Another consequence of (3.40) and (3.42) is that if $\underline{j}_i = (j_i(1), j_i(2))$ then

$$|j_{i+1}(k) - j_i(k)| \leq 2, \qquad k = 1, 2 .$$

Therefore, for given $\underline{j}_i$ there are at most 25 choices for $\underline{j}_{i+1}$. For $\underline{j}_0$ there are at most $4(q/t + 11)$ choices, since $x(0) \in \partial S(v, q/2 + 5t)$. Thus for given λ, there are at most

$$4\left(\frac{q}{t} + 11\right)(25)^{\lambda} \tag{3.45}$$

possibilities for $\underline{j}_0, \ldots, \underline{j}_\lambda$. Moreover, if r is open, then the piece of r from $x(\underline{j}_i)$ to $x(\underline{j}_{i+1})$ contains an open crossing, call it r_i, between the long sides of a $t \times 3t$ or $3t \times t$ rectangle. E. g. if $x(l_i) \in D(\underline{j}_i, t)$ and $x(l_{i+1})$ belongs to the right edge of $\partial F(\underline{j}_i, t)$ (which is $\{ (j_i(1)+2)t \} \times [(j_i(2)-1)t, (j_i(2)+2)t])$, then r_i crosses from the left edge to the right edge of

$$[(j_i + 1)t, (j_i + 2)t] \times [(j_i(2) - 1)t, (j_i(2) + 2)t] .$$

Denote this $t \times 3t$ or $3t \times t$ rectangle crossed by r_i by R_i. If $R_i = [a_i, a_i + t] \times [b_i, b_i + 3t]$ write R_i^* for the $2t \times 5t$ rectangle $[a_i - t, a_i + t] \times [b_i - t, b_i + 4t]$. Similarly, if $R_i = [a_i, a_i + 3t] \times [b_i, b_i + t]$, let $R_i^* = [a_i - t, a_i + 4t] \times [b_i - t, b_i + t]$. By construction $r_i \subset S(v, 3q/2 - 5t) \setminus \mathring{S}(v, q/2 + 5t)$ and hence $R_i^* \subset \mathring{S}(v, 3q/2) \setminus S(v, q/2)$. Thus, if we write

$$\hat{Z}(v, q/2, r) = \{ x \in S(v, 3q/2) : x \to r \quad \text{in} \quad S(v, 3q/2) \setminus S(v, q/2) \}$$

and

$$Z_i^*(t, r_i) = \{ x \in R_i^* : x \to r_i \quad \text{in} \quad R_i^* \},$$

then

$$\hat{Z}(v, q/2, r) \geq \max_{0 \leq i \leq \lambda} Z_i^*(t, r_i) .$$

Therefore,

$$P_{\frac{1}{2}} \{ \exists \text{ open crossing } r \text{ of } F(\underline{j}, q) \setminus D(\underline{j}, q) \text{ with } \hat{Z}(v, q/2, r) \leq q^{2-\varepsilon}\pi(q) \} \tag{3.46}$$

$$\leq \sum_{R^*, \ldots, R_\lambda^*} P_{\frac{1}{2}} \{ \text{For all } 0 \leqslant i \leqslant \lambda \, \exists \text{ open crossing } r_i \text{ in } R_i \text{ between its long sides with } Z_i^*(t, r_i) \leq q^{2-\varepsilon}\pi(q) \}.$$

As we saw in (3.44), $\lambda \geq q^{\delta/2}/4$. Also, given λ (3.45) gives an upper bound for the number of choices of the $\underline{j}_i$. For fixed $\underline{j}_i$, there are 4 choices for R_i and hence for R_i^*. Thus, in the right hand side of (3.46) we can let λ run from $q^{\delta/2}/4$ to ∞, and for each fixed λ there are most $16(q/t + 11)10^{2\lambda}$ summands. Finally, from any collection $R_0^*, \ldots, R_\lambda^*$ we can select a subcollection of $C_9\lambda$ disjoint ones (note that the corners of the R_i^* have coordinates divisible by t; each R_i^* can intersect at most a fixed number of the other ones.) For disjoint R_i^* the events

(3.47) $\{\exists$ open crossing r_i in R_i between its long sides with $Z_i^*(t, r_i) \le q^{2-\varepsilon}\pi(q)\}$

are independent. Moreover, by (3.39) and the choice of t the probability of (3.47) is at most $C_4^{-1}t^{-C_8\delta}$, since for large q

$$q^{2-\varepsilon}\pi(q) \le C_1 t^{2(1-\delta)}\pi(t).$$

Thus the right hand side of (3.46) is at most

$$\sum_{\lambda \ge \frac{1}{4}q^{\delta/2}} 16q10^{2\lambda}(C_4^{-1}t^{-C_8\delta})^{C_9\lambda}.$$

If $F(\underline{j}, q) \subset S(m)$, then the left hand side of (3.46) is the left hand side of (3.24) with v replaced by $P_{\frac{1}{2}}$. With a little more care we can handle a general $F(\underline{j}, q)$ which intersects $S(m)$ in the same way. We merely have to make sure that all connecting paths in steps *i*) and *ii*) stay in $S(m)$. As soon as t is so large that

$$10^2(C_4^{-1}t^{-C_8\delta})^{C_9} \le \frac{1}{2}$$

we therefore have (3.24) with v replaced by $P_{\frac{1}{2}}$.

STEP *iii*). — It remains to show that once we have (3.24) with the $P_{\frac{1}{2}}$-measure, then it also holds in the v-measure. For this we repeat the argument for (41) in [*16*]. We remind the reader that an event E is called *increasing* if its indicator function can only increase if a number of closed edges is changed into open edges. The Harris-FKG inequality says that for any two increasing events E_1, E_2 which depend on finitely many edges only, one has

$$P_p\{E_1 \cap E_2\} \ge P_p\{E_1\}P_p\{E_2\}$$

(see [*15*], Ch. 4.1). Unfortunately the event in braces in (3.24) is not increasing, so that we cannot immediately apply (41) of [*16*]. Denote the event in braces in (3.24) by E_0, and recall that v is the center of $D(\underline{j}, q)$, $F(\underline{j}, q) = S(v, 3q/2)$. For the time being assume that

(3.48) $$\underline{0} \notin S(v, 3q).$$

By definition

$$(3.49)\quad \nu\{E_0\} = \lim_{l\to\infty} [P_{\frac{1}{2}}\{\underline{0} \to \partial S(l)\}]^{-1} . P_{\frac{1}{2}}\{\underline{0} \to \partial S(l) \quad \text{and} \quad E_0\}.$$

Let

$E_1 = \{\underline{0} \to \partial S(l)\}$ and $E_2 = \{\exists$ open circuit in $S(v, 3q)\backslash S(v, 3q/2)$ which surrounds $S(v, 3q/2)\}$.

Now E_0 depends only on the edges is $S(v, 3q/2)$, and conditionally on the configuration in $S(v, 3q/2)$, the events E_1 and E_2 are increasing in the edges outside $S(v, 3q/2)$. We may therefore apply the Harris-FKG inequality to the outside of $S(v, 3q/2)$, conditionally on the configuration inside $S(v, 3q/2)$. This yields

$$P_{\frac{1}{2}}\{E_1 \cap E_2 \mid E_0\} \geq P_{\frac{1}{2}}\{E_1 \mid E_0\} P_{\frac{1}{2}}\{E_2 \mid E_0\}.$$

But E_2 and E_0 depend on disjoint collections of edges, hence are independent. Also, by (28) of [*16*], $P_{\frac{1}{2}}\{E_2\} \geq C_3 > 0$. Thus

$$P_{\frac{1}{2}}\{E_0 \cap E_1 \cap E_2\} \geq C_3 P_{\frac{1}{2}}\{E_0 \cap E_1\}.$$

Thus

$$(3.50)\quad \nu\{E_0\} \leq C_3^{-1} \lim_{l\to\infty} [P_{\frac{1}{2}}\{E_1\}]^{-1} P_{\frac{1}{2}}\{E_0 \cap E_1 \cap E_2\}.$$

Finally, the lines following (42) in [*16*] show that for l so large that $S(v, 3q) \subset S(l)$ we have under (3.48)

$$E_1 \cap E_2 \subset \{\underline{0} \to \partial S(l) \quad \text{in} \quad S^c(v, 3q/2)\},$$

so that

$$P_{\frac{1}{2}}\{E_0 \cap E_1 \cap E_2\} \leq P_{\frac{1}{2}}\{E_0 \text{ and } \underline{0} \to \partial S(l) \text{ in } S^c(v, 3q/2)\}$$
$$\leq P_{\frac{1}{2}}\{E_0\} P_{\frac{1}{2}}\{\underline{0} \to \partial S(l)\}.$$

Substitution into (3.50) yields

$$\nu\{E_0\} \leq C_3^{-1} P_{\frac{1}{2}}\{E_0\}.$$

Thus, if (3.48) holds we also have (3.24) in its original form, at the cost of replacing C_1 by $C_1 C_3^{-1}$.

If

$$(3.51)\quad \underline{0} \in S(v, 3q)$$

we must pay a slightly higher price to go from (3.24) with $P_{\frac{1}{2}}$ to (3.24) with ν.

Now set

$E_3 = \{ \partial S(v, 3q)$ is connected to $\partial S(l)$ by an open path which lies outside $S(v, 3q)$ with the exception of its initial point on $\partial S(v, 3q) \}$.

Then for all large l

$$E_1 = \{ \underline{0} \to \partial S(l) \} \subset E_3$$

so that

$$(3.52) \qquad \nu\{E_0\} = \lim_{l\to\infty} [P_{\frac{1}{2}}\{E_1\}]^{-1} P_{\frac{1}{2}}\{E_0 \cap E_1\}$$
$$\leq \limsup_{l\to\infty} [P_{\frac{1}{2}}\{E_1\}]^{-1} P_{\frac{1}{2}}\{E_0 \cap E_3\}$$
$$= P_{\frac{1}{2}}\{E_0\} \limsup_{l\to\infty} [P_{\frac{1}{2}}\{E_1\}]^{-1} P_{\frac{1}{2}}\{E_3\}.$$

Finally we apply (29) of [*16*] with

$$G = E_3 \cap \{ \underline{0} \to \partial S(v, 3q) \},$$
$$A_1^* = S(v, 3q) \backslash S(v, 3q/2),$$
$$A_2^* = S(v, 6q) \backslash S(v, 3q).$$

Note that if G occurs and there exist open circuits $\mathscr{C}_i$ in A_i^*, $i = 1, 2$, surrounding $S(v, 3q/2)$ and $S(v, 3q)$, respectively, and $\mathscr{C}_1 \to \mathscr{C}_2$, then $\underline{0} \to \partial S(l)$ by an open path which runs from $\underline{0}$ to $\mathscr{C}_1$, then along $\mathscr{C}_1$ and the open connection from $\mathscr{C}_1$ to $\mathscr{C}_2$, then along $\mathscr{C}_2$ and the connection from $\mathscr{C}_2$ to $\partial S(l)$. Therefore, by (29) in [*16*]

$$P_{\frac{1}{2}}\{ \underline{0} \to \partial S(l) \} \geq P_{\frac{1}{2}}\{ G, \exists \text{ open circuits } \mathscr{C}_i \text{ in } A_i^*, i = 1, 2, \text{ surrounding } S(v, 3q/2) \text{ and } S(v, 3q), \text{ respectively, and } \mathscr{C}_1 \to \mathscr{C}_2 \}$$
$$\geq C_4 P_{\frac{1}{2}}\{G\} = C_4 P_{\frac{1}{2}}\{ \underline{0} \to \partial S(v, 3q) \} P_{\frac{1}{2}}\{E_3\}.$$

This, together with (3.51) and (3.21), shows

$$[P_{\frac{1}{2}}\{E_1\}]^{-1} P_{\frac{1}{2}}\{E_3\} \leq [C_4 P_{\frac{1}{2}}\{ \underline{0} \to \partial S(v, 3q) \}]^{-1}$$
$$\leq [C_4 \pi(3q)]^{-1} \leq C_5 q^{1/2}.$$

Substitution into (3.52) shows that under (3.51)

$$\nu\{E_0\} \leq C_5 q^{1/2} P_{\frac{1}{2}}\{E_0\}.$$

Again (3.24) follows for all large q for the ν-measure. We merely have to decrease C_2 somewhat. Of course, once (3.24) holds for all large q we get it for all q by raising C_1.

REFERENCES

[1] S. Alexander and R. Orbach, Density of states on fractals: « fractons » ; *J. Physique Lett.*, t. **43**, 1982, L625-631.

[2] K. B. Athreya and P. E. Ney, *Branching Processes*, 1972, Springer-Verlag.

[3] Y. S. Chow and H. Teicher, *Probability Theory*, 1978, Springer-Verlag.

[4] P. G. de Gennes, La percolation : un concept unificateur, *La Récherche*, t. **7**, 1976, p. 919-927.

[5] A. de Masi, P. A. Ferrari, S. Goldstein and W. D. Wick, *An invariance principle for reversible Markov processes with application to random motions in random environments*, (1985, preprint).

[6] P. A. Doyle and J. L. Snell, *Random Walk and Electric Networks*, Carus Math. Monograph, No. 22, 1984, Math. Assoc. of America.

[7] R. Durrett, Conditioned limit theorems for some null recurrent Markov processes, *Ann. Probab.*, t. **6**, 1978, p. 798-828.

[8] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. I, 3rd ed., John Wiley and Sons, 1968.

[9] D. Griffeath and T. M. Liggett, Critical phenomena for Spitzer's reversible nearest particle systems, *Ann. Probab.*, t. **10**, 1982, p. 881-895.

[10] T. E. Harris, *The Theory of Branching Processes*, Springer-Verlag and Prentice Hall, 1963.

[11] S. Havlin, D. Movshovitz, B. Trus and G. H. Weiss, Probability densities for the displacement of random walks on percolation clusters, *J. Phys. A. Math. Gen.*, t. **18**, 1985, L719-722.

[12] C. C. Heyde, On large deviation probabilities in the case of attraction to a non-normal stable law, *Sankhya, Ser. A*, t. **30**, 1968, p. 253-258.

[13] P. Jagers, *Branching Processes with Biological Applications*, John Wiley and Sons, 1950.

[14] H. Kesten, The critical probability of bond percolation on the square lattice equals 1/2, *Comm. Math. Phys.*, t. **74**, 1980, p. 41-59.

[15] H. Kesten, *Percolation Theory for Mathematicians*, Birkhauser-Boston, 1982.

[16] H. Kesten, The incipient infinite cluster in two-dimensional percolation, to appear in *Theor. Prob. Rel. Fields*, 1986.

[17] R. Künnemann, The diffusion limit for reversible jump processes on Z^d with periodic random bond conductivities, *Comm. Math. Phys.*, t. **90**, 1983, p. 27-68.

[18] F. Leyvraz and H. E. Stanley, To what class of fractals does the Alexander-Orbach conjecture apply? *Phys. Rev. Lett.*, t. **51**, 1983, p. 2048-2051.

[19] M. Loeve, *Probability Theory*, 4th ed., Springer Verlag, 1977.

[20] B. B. Mandelbrot and S. Kirkpatrick, Solvable fractal family, and its possible relation to the backbone at percolation, *Phys. Rev. Lett.*, t. **47**, 1981, p. 1771-1774.

[21] P. A. Meyer, Martingales and Stochastic Integrals I, *Lecture Notes in Math*, t. **284**, 1972, Springer-Verlag.

[22] C. D. Mitescu and J. Roussenq, Diffusion on percolation structures, in Percolation Structures and Processes, *Ann. Israel. Phys. Soc.*, t. **5**, 1983, Eds. G. Deutscher, R. Zallen and J. Adler.

[23] A. G. Pakes, Some limit theorems for the total progeny of a branching process, *Adv. Appl. Prob.*, t. **3**, 1971, p. 176-192.

[24] R. Rammal and G. Toulouse, Random walk on fractal structures and percolation clusters, *J. Physique-Lett.*, t. **44**, 1983, L13-22.

[25] P. D. SEYMOUR and D. J. A. WELSH, Percolation probabilities on the square lattice, *Ann. Discrete Math.*, t. **3**, 1978, p. 227-245.
[26] R. S. SLACK, A branching process with mean one and possibly infinite variance, *Z. Wahrsch. verw. Geb.*, t. **9**, 1968, p. 139-145.
[27] J. STRALEY, The ant in the labyrinth: diffusion in random networks near the percolation threshold, *J. Phys., Solid State Phys.*, t. **C13**, 1980, p. 2991-3002.
[28] J. VAN DEN BERG and H. KESTEN, Inequalities with applications to percolation and reliability, *J. Appl. Prob.*, t. **22**, 1985, p. 556-569.

(Manuscrit reçu le 9 janvier 1986)