The Annals of Probability
1981, Vol. 9, No. 5. 753–768

GENERALIZED RANDOM WALK IN A RANDOM ENVIRONMENT

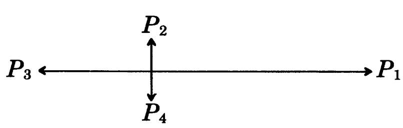
BY STEVEN A. KALIKOW

Stanford University

Random Walk in a Random Environment, a process where a randomly chosen transition probability is assigned to each point, has been exhaustively studied in the one dimensional nearest neighbor case but is much more difficult in any more general case. This paper develops crude techniques for handling transience results and zero-one laws in a generalized set up.

1. Introduction. In Section 1, all notation will be defined only in special cases and only heuristically. It will be reintroduced rigorously, in a general setting, in Section 2.

Consider a nearest neighbor random walk on $\mathbb{Z}$ where at each time the process moves to the left with probability ⅓ and to the right with probability ⅔. We abbreviate this probability law by writing $⅓ \leftrightarrow ⅔$. A probability law such as $⅓ \leftrightarrow ⅔$ is called an environment. If we say that 5 has environment $⅓ \leftrightarrow ⅔$, we mean that for any n, if the process is at 5 at time n, then at time $n + 1$, it will be at 4 with probability ⅓ and at 6 with probability ⅔. Random walk can be defined to be a process in which each element of $\mathbb{Z}$ has the same environment. Random Walk in a Random Environment, abbreviated R.W.I.R.E., is a process in which the environment of each point of $\mathbb{Z}$ is chosen randomly, in an i.i.d. (independent, identically distributed) manner. Once the environments are chosen, they are held fixed through time, and the process starts at 0 and takes place on these environments. This definition can be easily extended to a nearest neighbor process on $\mathbb{Z}^2$ where an environment is



with $P_1 + P_2 + P_3 + P_4 = 1$, or to a process which is not nearest neighbor.

The one dimensional nearest neighbor case has been exhaustively studied in [3], [5], [6] and [7] but in the more general case virtually nothing was known. In particular, example 2 (below) was previously unknown.

EXAMPLE 1. Consider the one dimensional case where the environment at each element of $\mathbb{Z}$ is taken to be $.001 \leftrightarrow .999$ with probability .999, and $.501 \leftrightarrow .499$ with probability .001. We denote this model in the following fashion:

$$g_1 = .001 \leftrightarrow .999, \qquad g_2 = .501 \leftrightarrow .499,$$

$$\mu(\{g_1\}) = .999, \qquad \mu(\{g_2\}) = .001.$$

Then it is easily proved that the process is transient; that it moves to the right with probability 1.

EXAMPLE 2. Now consider the two dimensional nearest neighbor case, where

Received March 31, 1978; revised August 11, 1980.
AMS subject classification. 60J10, 60J15
Key words and phrases. Random Walk in a Random Environment, Markov chain, transience, recurrence.

$$g_1 = \frac{.001}{2} \longleftarrow \overset{1/4}{\underset{1/4}{\updownarrow}} \longrightarrow \frac{.999}{2}, \qquad g_2 = \frac{.501}{2} \longleftarrow \overset{1/4}{\underset{1/4}{\updownarrow}} \longrightarrow \frac{.499}{2}$$

$$\mu(\{g_1\}) = .999, \qquad \mu(\{g_2\}) = .001.$$

It seems still obvious, but is difficult to prove, that this process is transient and moves to the right.

Theorem 1 establishes sufficient conditions for an R.W.I.R.E. $\{X_n\}_{n\geq 0}$ on a countable abelian group, for $\ell(X_n)$ to approach $+\infty$; here ℓ denotes a real-valued linear functional (i.e., $\ell(x + y) = \ell(x) + \ell(y)$ for x, y in the group). In particular, this theorem handles example 2. The generalization allows for unbounded step sizes. It should be noted that this provides new information in the one dimensional case because even in that case R.W.I.R.E. is unexplored when step sizes exceed 1.

REMARK 1. Actually, none of the structure of an abelian group, neither associativity, commutativity, or the existence of an inverse, is actually used in the proofs. All proofs go through word for word if we eliminate the word "abelian" and, indeed, with minor changes we could prove all results for a countable set endowed with an arbitrary binary relation. However, it is notationally simpler to do everything on an abelian group. □

Generalizing example 2, consider the two dimensional case where

$$g_1 = c_1 \longleftarrow \overset{b_1}{\underset{d_1}{\updownarrow}} \longrightarrow a_1, \qquad g_2 = c_2 \longleftarrow \overset{b_2}{\underset{d_2}{\updownarrow}} \longrightarrow a_2$$

$$\mu(\{g_1\}) = p, \qquad \mu(\{g_2\}) = 1 - p$$

where $0 < p < 1$, $a_1 + b_1 + c_1 + d_1 = 1 = a_2 + b_2 + c_2 + d_2$,

$$a_1, b_1, c_1, d_1, a_2, b_2, c_2, d_2, > 0, \ a_1 > c_1, \ a_2 < c_2.$$

We will show the following result for this example:

COROLLARY TO THEOREM 1. If

$$\frac{p(a_1 - c_1)}{(1-p)(c_2 - a_2)} > \max\left(\frac{a_1}{a_2}, \frac{b_1}{b_2}, \frac{c_1}{c_2}, \frac{d_1}{d_2}\right)$$

then the process moves to the right (i.e., the first coordinate approaches ∞) with probability 1.

Although the conclusion of example 2 seems obvious, none of the previous techniques seem to handle it. Harmonic functions for X_n, which can be easily computed in dimension 1, cannot be computed here. The problem is to show that the process does not somehow 'seek out' the bad environments. We do this by considering a fixed vector $v \in \mathbb{Z}^2$, fixing the environment off that vector, and bounding the maximum effect that a change of environment on v can have on how often the process visits v. Using this technique, we construct a certain average Markov chain which has probabiity laws similar to the R.W.I.R.E. itself, and which has a drift to the right at each point.

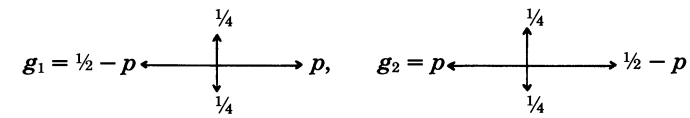
In Section 5, under very restrictive conditions, a powerful zero-one law is proved. However, if those conditions are relaxed, most zero-one conjectures are still unsolved. Here are some problems which remain open.

Problem 1. When the reader reads my proof of Theorem 1, it will become clear that the techniques used in the proof are so crude that the result is probably far from optimal. Indeed, in the one dimensional nearest neighbor case, the problem is completely settled and the conditions of Theorem 1 for transience, at least in that case, are known to be much worse than necessary. In cases other than the one dimensional case, nothing is known except Theorem 1. One such case is the case of the corollary to Theorem 1 which has already been stated. Here, for example, the reader can easily construct examples where the conditions of the corollary are not met and where it nonetheless seems obvious that the process is transient and moves to the right. Can Theorem 1, or at least the corollary to Theorem 1, be improved?

Problem 2. In the two dimensional case, consider the event that the x-coordinate approaches ∞. Need this event have probability either zero or one?

Problem 3. Is every 3-dimensional R.W.I.R.E. transient?

Problem 4. Let $0 < p < 1/2$,



$$\mu(\{g_1\}) = \mu(\{g_2\}) = 1/2.$$

Is this R.W.I.R.E. recurrent? For that matter, can any reasonable criterion for recurrence be established?

2. Statement of Theorem 1.

DEFINITION 1. *Random Walk in a Random Environment* (R.W.I.R.E.).

N are the nonnegative integers.

C is a countable abelian group (e.g., $\mathbb{Z}^2$).

$\vec{0}$ is the zero element of C.

$\Omega = C^N$. Elements of Ω are written $\{X_n\}_{n\in N}$ where each $X_n \in C$. $\mathscr{A}$ is the standard σ-algebra on Ω (generated by cylinders). G is the collection of all probability measures on C. Hence, for $g \in G$, $\sum_{v\in C} g(v) = 1$. $S = G^C$. Elements s of S are called environments and are written $s = \{s_v\}_{v\in C}$ where $s_v \in G$ for all v. Hence $\Sigma_{v_1\in C}\, s_v(v_1) = 1$. To each environment s is associated a measure P_s on Ω which makes $\{X_n\}$ into a Markov chain with stationary transition probabilities;

$$P_s(X_0 = \vec{0}) = 1,\ P_s(X_{n+1} = v + v_1 \mid X_n = v) = s_v(v_1),\ v,\ v_1 \in C.$$

E_s is the expectation operator corresponding to P_s. μ is a probability measure on G (here we presume the standard σ-algebra on G, the borel sets when G is given the topology of weak convergence). e is an S-valued random variable such that $\{e_v\}_{v\in C}$ are i.i.d. G-valued with common probability distribution μ. Hence P_e is a random measure on $(\Omega, \mathscr{A})$. e is called a random environment and, for $v \in C$, e_v is called the random environment at v. For each $A \in \mathscr{A}$, let $P(A) = E(P_e(A))$.

The process $\{X_n\}_{n\in N}$ under probability law P is called a *Random Walk in a Random Environment* (R.W.I.R.E.). R.W.I.R.E. will also be used to refer to P itself.

E will be used ambiguously to refer to either the expectation operator corresponding to the probability law of e or to the expectation operator corresponding to P. Hence, for any $\mathscr{A}$ measurable random variable X, $E(X) = E(E_e(X))$. For $s \in S$, $P_s(v_1, v_2)$ will denote $s_{v_1}(v_2 - v_1)$ so that $P_s(v_1, v_2) = P_s(X_{n+1} = v_2 \mid X_n = v_1)$ and $P_e(v_1, v_2) = e_{v_1}(v_2 - v_1)$.

THEOREM 1. *Let ℓ be a real valued linear functional on C. Suppose that*

a) *There exists a positive real function H such that*
   i) $\int_0^\infty H(\lambda)\, d\lambda < \infty$.
   ii) $\sum_{\{v \in C: |\ell(v)| > \lambda\}} g(v) \leq H(\lambda)$ *for all* $\lambda > 0$

with μ probability 1. (*Recall that μ is a probability measure picking g from G.*)

b) *For any v_1 and v_2 in C where $v_1 \neq v_2$, there exists $\varepsilon_0 > 0$ such that $P(p_e$(after time 0, $\{X_n\}$ hits v_1 before $v_2) > \varepsilon_0) = 1$*

(*i.e., for almost every choice of e, the p_e Markov chain has more than ε_0 probability that the first time after time 0 that X_n hits v_1 precedes the first time after time 0 that X_n hits v_2.*

The following notation will be used in the statement of condition c. Suppose $f_1, f_2 \in \mathbb{R}^C$. Define $f_i \cdot f_2$ by

$$f_1 \cdot f_2 = \sum_{v \in C} f_1(v) f_2(v).$$

c) *Let $F = \{f \in [0, 1]^C : f(v) \neq 0$ for all $v \neq \vec{0}\}$.*

$$\inf_{f \in F} \left(\int \frac{1}{f \cdot g}\, d\mu(g) \right)^{-1} \int \frac{\ell \cdot g}{f \cdot g}\, d\mu(g) > 0$$

Under these conditions

$$\lim_{n \to \infty} \ell(X_n) = \infty \text{ with probability } 1.$$

Note that, by translation invariance, condition b implies that for every v_1, v_2, v_3, n, where $v_2 \neq v_3$, there exists $\varepsilon_0 > 0$ such that, for almost every e, if the process is at v_1 at time n, there is P_e probability of more than ε_0 that, after time n, it will hit v_2 before v_3.

It should be noticed that condition b is not satisfied in the one dimensional nearest neighbor case. It would be possible to weaken condition b slightly in order to handle this case but the resulting statement would be complicated to state and the proof of the theorem would become more messy. Furthermore, the one dimensional nearest neighbor case has already been completely settled and the purpose of this paper is to deal with the other cases.

Condition c says that certain weighted averages of $l(v)$ are positive. The reader who wishes a better understanding of this formula should skip ahead and read the proof of the corollary to Theorem 1, which works out the meaning of this condition in the special case where the conditions of that corollary hold.

3. A Markov Chain. Section 3 develops a process which is used in the proof of Theorem 1. Note that R.W.I.R.E. is not itself a Markov chain. Definition 2 defines a measure $\hat{P}$ which does make $\{X_n\}$ into a Markov chain and is closely related to the R.W.I.R.E. measure P. We assume the notation of Definition 1. Recall that $P_s(v_1, v_2) = (s_{v_1} - v_1)$. In this section we presume $\{X_n\}$, P is an R.W.I.R.E. obeying condition b of Theorem 1.

DEFINITION 2. *Markov chain corresponding to P, θ.*

Let $\theta \subset C - \{\vec{0}\}$, $\theta \neq \phi$.
τ = least n such that $X_n \in \theta$, ∞ if none exists.
For $v \in C$, $R(v) = \#\{m, 0 \leq m \leq \tau \text{ and } X_m = v\}$.
$D = \{v \in C: P(R(v) > 0) > 0\}$.
$\Pi_s(v) = E_s(R(v)), \hat{\Pi}(v) = E(\Pi_e(v)), v \in C$.

Note that, for $v \in C$, $v \in D$ iff $\hat{\Pi}(v) > 0$. Note also that by condition b of Theorem 1, there exists $\varepsilon_0 > 0$, such that for almost every choice of e, whenever the P_e Markov chain hits v it has more the ε_0 P_e probability of hitting θ by the next time it hits v. Therefore, under P_e, conditioned on the event $R(v) \geq 1$, $R(v) - 1$ is geometrically distributed with success probability less than $1 - \varepsilon_0$ and hence has expectation less than $1/\varepsilon_0$. It follows that $\hat{\Pi}(v) = E(P_e R(v)) < 1 + 1/\varepsilon_0 < \infty$.

$$\hat{P}(v_1, v_2) = E(\Pi_e(v_1) P_e(v_1, v_2))/\hat{\Pi}(v_1),\ v_1 \in D - \theta,\ v_2 \in C.$$

$$\hat{P}(v_1, v_2) = \delta_{v_1,v_2}\ v_1 \in D \cap \theta.$$

Note that, if $v_1 \in D$, $\hat{P}(v_1, v_2) > 0$, then $v_2 \in D$ because $\Pi_e(v_1) P_e(v_1, v_2) \leq \Pi_e(v_2)$. It is immediate that

$$\textstyle\sum_{v \in D} \hat{P}(v_1, v) = 1,\ 0 \leq \hat{P}(v_1, v_2) \leq 1,\ v_1, v_2 \in D.$$

Let $\hat{P}$ be the probability measure on $D^N \subset \Omega$ which makes $\{X_n\}$ into a Markov chain with $\hat{P}(X_0 = \vec{0}) = 1$ and stationary transition probabilities

$$\hat{P}(X_{n+1} = v_2 \mid X_n = v_1) = \hat{P}(v_1, v_2),\ v_1, v_2 \in D,\ n \in N.$$

$\hat{E}$ is the expectation operator corresponding to $\hat{P}$.

$\hat{P}$ is called the Markov chain corresponding to P, θ.

PROPOSITION 1. *If*

$$\hat{P}(\tau < \infty) = 1, \tag{1}$$

then, for every $v \in \theta$, $\hat{P}(X_\tau = v) = P(X_\tau = v)$ *and hence* $P(\tau < \infty) = 1$.

PROOF. For $v \in C$,

$$\hat{\Pi}(v) = E\Pi_e(v) = E(E_e R(v)) = ER(v). \tag{2}$$

Easy Markov chain arguments give

$$\begin{aligned} \Pi_e(v_1) &= \textstyle\sum_{n=0}^{\infty} P_e\{X_n = v_1,\ \tau \geq n\} = \delta_{v_1,\vec{0}} + \sum_{n=1}^{\infty} \sum_{v \in C-\theta} P_e\{X_{n-1} \\ &= v,\ X_n = v_1\ \tau \geq n - 1\} = \delta_{v_1,\vec{0}} + \textstyle\sum_{v \in C-\theta} \Pi_e(v) P_e(v, v_1). \end{aligned} \tag{3}$$

Thus

$$\begin{aligned} \textstyle\sum_{v \in D-\theta} \hat{\Pi}(v) \hat{P}(v, v_1) + \delta_{v_1,\vec{0}} &= \textstyle\sum_{v \in D-\theta} E(\Pi_e(v) P_e(v, v_1) + \delta_{v_1,\vec{0}}) \\ &= \textstyle\sum_{v \in C-\theta} E(\Pi_e(v) P_e(v, v_1) + \delta_{v_1,\vec{0}}) \\ &= E\Pi_e(v_1) = \hat{\Pi}(v_1),\ v_1 \in D. \end{aligned} \tag{4}$$

The first equality follows from the definition of $\hat{P}$.

For all $v_1 \in D$, define $\hat{\pi}_n(v)$, $n = 0, 1, \cdots$ by induction as follows:

$$\hat{\pi}_0(v_1) = \delta_{v_1,\vec{0}}$$

$$\hat{\pi}_{n+1}(v_1) = \textstyle\sum_{v \in D-\theta} \hat{\pi}_n(v) \hat{P}(v, v_1) + \delta_{v_1,\vec{0}}.$$

Then

$$\hat{\pi}_n(v) = \hat{E}(\#\{m: 0 \leq m \leq \min(n, \tau) \text{ and } X_m = v\}.$$

Consequently,

$$\lim_{n\to\infty} \hat{\pi}_n(v) = \hat{E}(R(v)), \qquad v \in D.$$

Furthermore, by (4) and induction on n,

$$\hat{\pi}_n(v) \leq \hat{\Pi}(v) \qquad \text{for all } n.$$

Thus

$$\hat{E}(R(v)) \leq \hat{\Pi}(v), \qquad v \in D. \tag{5}$$

Note that $R(v)$ is the indicator function of $X_\tau = v$ for $v \in \theta$. Thus $v \in \theta$ implies $P(X_\tau = v) = ER(v)$ and $\hat{P}(X_\tau = v) = \hat{E}(R(v))$. (5) and (2) give

$$\hat{P}(X_\tau = v) \leq P(X_\tau = v), \qquad v \in D \cap \theta. \tag{6}$$

By (1)

$$\textstyle\sum_{v \in D \cap \theta} P(X_\tau = v) \leq 1 = \sum_{v \in D \cap \theta} \hat{P}(X_\tau = v).$$

Hence by (6)

$$\textstyle\sum_{v \in D \cap \theta} P(X_\tau = v) = 1$$

$$P(X_\tau = v) = \hat{P}(X_\tau = v), \qquad v \in D \cap \theta$$

$$P(X_\tau = v) = 0 = \hat{P}(X_\tau = v), \qquad v \in \theta - D. \quad \square$$

4. Proof of Theorem 1 and Corollary. For any positive real M, let θ_M be the set of all $v \in C$ such that $|\ell(v)| > M$ and let Φ_M be the set of all $v \in C$ such that $\ell(v) > M$ or $\ell(v) < 0$. Let τ_M and $\check{\tau}_M$ be the hitting times of θ_M and Φ_M, respectively. Let

$$A_M = \{\omega \in \Omega: \tau_M < \infty \quad \text{and} \quad \ell(X_{\tau_M}) > M\}$$

$$B_M = \{\omega \in \Omega: \check{\tau}_M < \infty \quad \text{and} \quad \ell(X_{\check{\tau}_M}) > M\}.$$

We will show

$$\lim_{M \to \infty} P(A_M) = 1 \tag{7}$$

$$\lim_{M \to \infty} P(B_M) > 0.$$

This implies that $\limsup_{n\to\infty} \ell(X_n) = \infty$ with probability 1 and that $\ell(X_n) \geq 0$ for all n with positive probability. These two results in turn imply $\lim_{n\to\infty} \ell(X_n) = \infty$ by a proof analogous to the proof of Theorem 3. The idea is that since $\limsup_{n\to\infty} \ell(X_n) = \infty$, $\ell(X_n)$ becomes higher than it has ever been before infinitely often, and each time it does so it has a probability p of never dropping below its new height, where $p = P(\ell(X_n > 0 \text{ for all } n)$. Consequently, it infinitely often reaches new heights approaching ∞ and infinitely many of those times it never again drops below its new height. For details of this technique see the proof of Theorem 3.

Thus we need only establish (7). Fix M, and let $\hat{P}, \check{P}$ be the Markov chains corresponding to P, θ_M and P, Φ_M, respectively, let $\tau = \tau_M$ and $\check{\tau} = \check{\tau}_M$. By Proposition 1, if we can establish

$$\hat{P}(\tau < \infty) = 1, \qquad \check{P}(\check{\tau} < \infty) = 1, \tag{8}$$

then $P(A_M) = \hat{P}(A_M)$, $P(B_M) = \check{P}(B_M)$. Thus (7) will follow if we can establish (8) and

(9) For every $\delta > 0$ there exists M such that the $\hat{P}$ process defined in terms of M satisfies $\hat{P}(A_M) > 1 - \delta$; there exists $\delta > 0$ such that, for all M, the $\check{P}$ process defined in terms of M satisfies $\check{P}(B_M) > \delta$.

From this point onward we will speak only of the $\hat{P}$ process and not of the $\check{P}$ process. The point is that, for every statement to be made of the $\hat{P}$ process, the corresponding statement for the $\check{P}$ process (where all notation of Definition 2 defined in terms of P, θ_M is replaced by the corresponding notation defined in terms of P, Φ_M) can be made and proved with the exact same proof. Here we assume all notation of Definition 2 for the $\hat{P}$ process.

To conclude, we now establish (8) and (9) by demonstrating that $\ell(X_n)$ for the $\hat{P}$ process has a positive drift. From Condition a of Theorem 1 and the definition of $\hat{P}$ we have for $X_n \in D - \theta_M$ and for all $\lambda > 0$:

$$\hat{P}(|\ell(X_{n+1}) - \ell(X_n)| > \lambda \,|\, X_n) \leq H(\lambda) \tag{10}$$

where $\int_0^\infty H(\lambda)\,d\lambda < \infty$ and H does not depend on M. Let

$$\varepsilon = \inf_{f\in F}\left(\int \frac{1}{f\cdot g}\,d\mu(g)\right)^{-1} \int \frac{\ell\cdot g}{f\cdot g}\,d\mu(g).$$

Then $\varepsilon > 0$ by condition c and ε does not depend on M. We will show that for any $v_1 \in D - \theta_M$:

$$\hat{E}(\ell(X_{n+1})-\ell(X_n)\,|\,X_n = v_1) > \varepsilon. \tag{11}$$

The proof that (10) and (11) establish (8) and (9) follows from Lemma 1 of [4] and is left to the appendix.

REMARK 3. The heuristic idea of the proof of (11) is as follows. Note that $\hat{E}$ is a certain average of expectations with respect to P_s, and that

$$E_s(\ell(X_{n+1})-\ell(X_n)\,|\,X_n = v)$$

can be negative for certain environments s. However, it will turn out that if we fix the environment off v and average only over the value of the environment at v, then the conditional expectation of $\ell(X_{n+1})-\ell(X_n)$ becomes positive. The technique is to express $\Pi_s(v)$ in terms of quantities which depend only on the environment off v.

PROOF OF (11). For $s \in S$, $v_1 \in C$, let

$$f_s(v_1) = P_s(\{X_n\} \text{ hits } v_1 \text{ for some } n \text{ with } 0 \leq n < \tau).$$

For any $n_0 \in N$, v_1, $v_2 \in C$, let

$$f_s(v_1, v_2) = P_s(X_n \neq v_1 \text{ for all } n \in (n_0, \tau]\,|\,X_{n_0+1} = v_1 + v_2,\ n_0 < \tau).$$

Note that $f_s(v_1, \bar{0}) = 0$ for any $v_1 \in C$, and that by the Markov property, $f_s(v_1, v_2)$ does not depend on n_0. Note also that $f_s(v_1)$ and $f_s(v_1, v_2)$ do not depend on s_{v_1}, but only on $\{s_v : v \neq v_1\}$, because s_{v_1} only influences the probabilities of paths which at some time exit from v_1 and $f_s(v_1)$, $f_s(v_1, v_2)$ are both sums of probabilities of finite paths which never exit from v_1. For $s \in S$, v_1, $v \in C$, recall that $s_v(v_1) = P_s(v, v + v_1)$ (see Definition 1). If at a given time n_0 before τ, we know the process to be at v_1, the P_s probability that it will not return to v_1 before τ is

$$\textstyle\sum_{v_2\in C} s_{v_1}(v_2)\, f_s(v_1, v_2).$$

Under P_s, $R(v) \geq 1$ with probability $f_s(v)$ and for $v \notin \theta_M$, conditioned on the event $R(v) \geq 1$, $R(v) - 1$ is geometrically distributed with failure probability $\sum_{v_1\in C} s_v(v_1)\, f_s(v, v_1)$. Thus, for $v \notin \theta_M$,

$$\Pi_s(v) = f_s(v)/\textstyle\sum_{v_1\in C} s_v(v_1)\, f_s(v, v_1). \tag{12}$$

For $v \in D$ we get

$$\begin{aligned}
&\hat{E}(\ell(X_{n+1}) - \ell(X_n)\,|\,X_n = v) = \textstyle\sum_{v_1\in C}\hat{P}(v, v + v_1)\ell(v_1)\\
&= \textstyle\sum_{v_1\in C}E(\ell(v_1)\Pi_e(v)P_e(v, v + v_1))/E(\Pi_e(v)) \qquad (13)\\
&= (\textstyle\sum_{v_1\in C}E(\ell(v_1)\Pi_e(v)e_v(v_1)))/E(\Pi_e(v)).
\end{aligned}$$

By (12) we get

$$\begin{aligned}
&\textstyle\sum_{v_1\in C}E(\ell(v_1)\Pi_e(v)e_v(v_1)) \qquad (14)\\
&= \textstyle\sum_{v_1\in C}E(\ell(v_1)e_v(v_1)f_e(v)/\sum_{v_1\in C} e_v(v_1)f_e(v, v_1))
\end{aligned}$$

$$= E(E(f_e(v)\textstyle\sum_{v_1\in C}(e_v(v_1)\ell(v_1))/\sum_{v_1\in C}(e_v(v_1)f_e(v,v_1)))\,|\,\{e_{v_1}: v_1\neq v\})$$

$$= E(\textstyle\int(f_e(v)\sum_{v_1\in C}g(v_1)\ell(v_1)/\sum_{v_1\in C}g(v_1)f_e(v,v_1))\,d\mu(g))$$

The last equality holds because e_v is distributed with distribution μ and is independent of $\{e_{v_1}: v_1 \neq v\}$ while $f_e(v)$, $f_e(v, v_1)$ depend only on $\{e_{v_1}: v_1 \neq v\}$.

By similar reasoning

$$E(\Pi_e(v)) = E\int (f_e(v)/\textstyle\sum_{v_1\in C} g(v_1)f_e(v,v_1))\,d\mu(g). \tag{15}$$

With F as in condition c of Theorem 1, and $f(\cdot)$ defined by $f(v_1) = f_s(v, v_1)$ for any fixed v, s, we have $f \in F$ for almost all s by condition b of Theorem 1. Hence for almost all choices of e

$$\left(\int f_e(v)/\textstyle\sum_{v_1\in C} g(v_1)f_e(v,v_1)\,d\mu(g)\right)^{-1} f_e(v)\sum_{v_1\in C}g(v_1)\ell(v_1)/\sum_{v_1\in C}g(v_1)f_e(v,v_1))\,d\mu(g)$$

$$= \left(\int (1/(g\cdot f_e(v,\cdot)))\,d\mu(g)\right)^{-1}\left(\frac{g\cdot\ell}{g\cdot f_e(v,\cdot)}\right)d\mu(g) \geq \varepsilon.$$

It follows from 14 and 15 that the right hand side of (13) is at least ε. □

Proof of Corollary to Theorem 1. Recall the conditions on a_1, b_1, c_1, d_1, a_2, b_2, c_2, d_2 stated in the paragraph preceding the statement of the corollary. We assume the hypothesis of the corollary and verify conditions a, b and c of Theorem 1 with $\ell(v)$ = the first coordinate of v. Condition a holds with $H(\lambda) = 1$ for $0 \leq \lambda \leq 1$ and $H(\lambda) = 0$ for $\lambda > 1$. Condition b holds because for all admissible environments, every nearest neighbor path (i.e., path where $X_{n+1} - X_n \in \{(0, 1), (0, -1), (-1, 0), (1, 0)\}$ for all n) has positive, bounded away from zero probability. For $f \in F$, define A_f by

$$A_f = \left(\int\frac{1}{f\cdot g}\,d\mu(g)\right)^{-1}\left(\int\frac{\ell\cdot g}{f\cdot g}\,d\mu(g)\right)$$

$$= (p/\textstyle\sum_{v\in C}f(v)g_1(v)) + (1-p)/\sum_{v\in C}f(v)g_2(v))^{-1} \tag{16}$$

$$((p(a_1-c_1)/\textstyle\sum_{v\in C}f(v)g_1(v)) + (1-p)(a_2-c_2)/\sum_{v\in C}f(v)g_2(v)).$$

Also define

$$m = \max(a_1/a_2, b_1/b_2, c_1/c_2, d_1/d_2)/(p(a_1-c_1)/((1-p)(c_2-a_2))).$$

Then, dividing both numerator and denominator of the right hand side of 16 by $\sum_{v\in C} f(v)g_2(v)$ and then solving for $\sum_{v\in C} f(v)g_1(v)/\sum_{v\in C} f(v)g_2(v)$ yields

$$\begin{aligned} p(a_1-c_1-A_f)/(1-p)(c_2-a_2+A_f) &= \textstyle\sum_{v\in C}f(v)g_1(v)/\sum_{v\in C}f(v)g_2(v)\\ &= f(1,0)a_1 + f(0,1)b_1 + f(-1,0)c_1\\ &\quad + f(0,-1)\,d_1)/(f(1,0)a_2\\ &\quad + f(0,1)b_2 + f(-1,0)c_2 + f(0,-1)\,d_2)\\ &\leq \max(a_1/a_2, b_1/b_2, c_1/c_2, d_1/d_2)\\ &= mp(a_1-c_1)/(1-p)(c_2-a_2). \end{aligned}$$

Since we assumed $a_1 > c_1$, $a_2 < c_2$, the condition of the corollary states $m < 1$. Thus A_f is positive and bounded below by a function of a_1, c_1, a_2, c_2, m. □

5. Zero-One Laws. Most interesting questions about zero-one laws of R.W.I.R.E. are still open problems. In this section, for simplicity, we restrict ourselves to the case $C = \mathbb{Z}^2$.

LEMMA 1. *Suppose that the* R.W.I.R.E. *is irreducible with probability* 1 (*i.e., for almost all environments e and for all pairs of points* v_1, v_2 *in* $\mathbb{Z}^2$, *there is a path of positive probability connecting* v_1 *to* v_2). *Then for any finite set* $\theta \subset \mathbb{Z}^2$, $P(X_n$ *hits* θ *infinitely often*) *is* 0 *or* 1 *and is independent of* θ.

PROOF. For any fixed environment s, X_n under probability law P_s is a Markov chain. If P_s is irreducible, then this Markov chain is either recurrent or transient. Let

$$f(s) = \begin{cases} 1 & \text{if } P_s \text{ is recurrent} \\ 0 & \text{otherwise.} \end{cases}$$

Then $P_s(X_n$ hits θ infinitely often$) = f(s)$ if P_s is irreducible. Hence, since we are assuming that P_e is irreducible with probability 1, the lemma reduces to proving that f is constant with probability 1. We do this by demonstrating that f is measurable with respect to the completion of the tail field of the i.i.d. random variables $\{e_v\}_{v \in Z^2}$. In fact, we prove that $f(s_1) = f(s_2)$ when s_1 and s_2 are any two irreducible environments which differ on only finitely many points of $\mathbb{Z}^2$. Let Ξ be a finite set such that s_1 and s_2 agree off Ξ. Suppose that $f(s_1) = 1$. Let v be an element of $\mathbb{Z}^2 \backslash \Xi$. Then for any $m > 0$,

$$P_{s_2}(X_{n+m} \in \Xi \text{ for some } n > 0 \mid X_m = v) =$$

$$P_{s_1}(X_{n+m} \in \Xi \text{ for some } n > 0 \mid X_m = v) = 1.$$

The first equality holds because both expressions are the sums of probabilities of finite paths which never exit from points of Ξ; P_{s_1} and P_{s_2} agree on all such paths. Thus $f(s_1) = 1$ implies $f(s_2) = 1$ and the converse is proved similarly. □

DEFINITION. If for all finite θ, $P(X_n$ hits θ infinitely often$) = 1$ we say that the R.W.I.R.E. is recurrent.

Now suppose that θ is an infinite subset of $\mathbb{Z}^2$. Can we conclude $P(X_n$ hits θ infinitely often) equals zero or one? This is still an open question. However, if we impose the additional conditions that the vertical components of all environments are nonrandom and that there is a lower bound on the horizontal components, a much stronger result can be proved. For the convenience of the reader we introduce some standard notation. For any random variable X, let $\sigma(X)$ be the σ-algebra generated by X. For any collection of σ-algebras $\{\sigma_i\}_{i=1}^{\infty}$ let $\bigvee_{i=1}^{\infty} \sigma_i$ be the smallest σ-algebra containing all σ_i and $\bigwedge_{i=1}^{\infty} \sigma_i$ be the intersection of the $\{\sigma_i\}_{i=1}^{\infty}$. Then the tailfield of a sequence of random variables $\{X_i\}_{i=1}^{\infty}$ is defined to be $\bigwedge_{i=1}^{\infty}(\bigvee_{j=i}^{\infty} \sigma(X_j))$.

THEOREM 2. *Let* a, b, *and* ε *be reals* $0 < \varepsilon$, $0 \le a, b < 1$. *Suppose that with* μ-*probability* 1

i) $g((-1, 0)) + g((1, 0)) + g((0, -1)) + g((0, 1)) = 1$.

ii) $g((0, 1)) = a, g((0, -1)) = b$.

iii) $\min\{g((1, 0)), g((-1, 0))\} \ge \varepsilon$.

Then either the R.W.I.R.E. *is recurrent or the tailfield of* $\{X_n\}_{n=1}^{\infty}$ *is trivial.*

PROOF. If $a = b = 0$ this reduces to the one-dimensional case which can be easily handled with a proof similar to the one given below in the other cases. If $a = 0$, $b \neq 0$ or $b = 0$, $a \neq 0$ the process is clearly transient and if $a \neq 0$, $b \neq 0$ the process is irreducible so in either case the conclusion of Lemma 1 holds. Thus, throughout the proof we will feel free to use the conclusion of Lemma 1.

This proof uses the well known technique of coupling (see, for example, [2]) which we explain presently. Let $(\Omega_i, \mathscr{A}_i, P_i)_{i=1,2}$ be two probability spaces, $\mathscr{A}_1 \times \mathscr{A}_2$ the usual product

σ-algebra on $\Omega_1 \times \Omega_2$. Using the natural imbeddings $\mathscr{A}_i \to \mathscr{A}_1 \times \mathscr{A}_2$, $i = 1, 2$, we can regard $\mathscr{A}_i$, $i = 1, 2$, as sub-sigma algebras of $\mathscr{A}_1 \times \mathscr{A}_2$ (here $\mathscr{A}_1 \to \mathscr{A}_1 \times \mathscr{A}_2$ is defined by $A \to A \times \Omega$). A probability space $(\Omega_1 \times \Omega_2, \mathscr{A}_1 \times \mathscr{A}_2, P)$ is called a coupling of $(\Omega_1, \mathscr{A}_1, P_1)$ and $(\Omega_2, \mathscr{A}_2, P_2)$ if P restricted to $\mathscr{A}_1$ is P_1 and P restricted to $\mathscr{A}_2$ is P_2. Hence, regarding Ω_1 and Ω_2 as path spaces, when we couple two stochastic processes together, we get a process where two particles move jointly in such a way that if we look only at one of the particles and forget the other, we get back one of the original processes.

Let m be any positive integer. Note that since we are considering a nearest neighbor process, the sum of the coordinates of X_m has the same parity as m. With this in mind, let $p = (i_1, i_2)$ be an element of $\mathbb{Z}^2$ with $(-1)^{i_1+i_2} = (-1)^m$. Now consider a fixed choice of environment s such that for $v \in \mathbb{Z}^2$, s_v obeys conditions i, ii, and iii. Let process 1 start at the origin at time 0 and move as a Markov chain in accordance with environment s (i.e., process 1 is simply $\{X_n\}$ under probability law P_s). We now couple a process 2 to process 1. We construct process 2 so that if we consider only process 2 it also moves as a Markov chain in environment s, but it no longer starts at $\vec{0}$. In fact we start process 2 at time m at point p and leave it undefined for times less than m. I will now define how process 2 moves inductively. Suppose $n \geq m$ and we know where processes 1 and 2 are at time n. If the y-coordinates of the two processes are different, then we let process 2 take its next step in accordance with environment s independently of the movement of process 1. In the case where the y-coordinates are the same, but the x-coordinates differ, let v be the position of process 2 at time n and let $g = s_v$. In that case, if process 1 moves one step upward or downward, let process 2 also move one step up or down, respectively. If process 1 moves horizontally, in either direction, let process 2 move one step left with probability $g((-1, 0))/(g((-1, 0)) + g((1, 0)))$ and one step right with probability $g((1, 0))/(g((-1, 0)) + g((1, 0)))$. In the third case, where the two processes are together at time n, let them stay together, i.e., process 2 takes the same next step as process 1.

We have coupled two Markov chains, both governed by P_s, one of which starts at $\vec{0}$ at time 0 and the other starts at p at time m.

LEMMA 2. *For almost all environments e, the coupled processes defined above will eventually coincide with P_e probability* 1.

PROOF OF LEMMA 2. First assume we are in a situation where the y-coordinates are different. Because all vertical components of the environment s are the same, the difference between the y-coordinates behaves like a symmetric random walk with step sizes $-2, -1, 0, 1, 2$ until the two y-coordinates agree. Since a symmetric random walk with bounded step sizes is recurrent, the difference between the y-coordinates will at some time be zero. Once that happens, the vertical movement of process 1 is matched by the vertical movement of process 2 and the y-coordinates remain the same. Furthermore, by induction and our initial assumption regarding the parity of p, the difference between the x-coordinates will be an even integer. Because there are only countably many pairs v_1, v_2 of elements of $\mathbb{Z}^2$, it suffices to show:

(17) For almost all choices of s, and for any choice of pairs v_1, v_2 of elements of $\mathbb{Z}^2$ whose y-coordinates are the same and x-coordinates differ by an even integer, the coupled processes 1 and 2 will with probability 1 meet if there exists a time at which process 1 is at v_1 and process 2 is at v_2.

Since the y coordinates will remain the same, this amounts to showing that the x-coordinates will become the same. We introduce some notation. For fixed environment s and fixed $v \in \mathbb{Z}^2$, define the probability law P_s^v analogously to our definition of P_s except, instead of initial condition $X_0 = \vec{0}$, assume $X_0 = v$. Then define P_e^v, P^v, E^v analogously to our definitions of P_e, P and E (i.e., for $A \in \mathscr{A}$, $P^v(A) = E(P_e^v(A))$, and E^v is the expectation operator corresponding to P^v.) For each fixed environment s, couple processes $P_s^{v_1}$ and $P_s^{v_2}$ together in the same way that processes 1 and 2 above were coupled (i.e., since v_1 and v_2

initially have the same y-coordinates, let their y-coordinates remain the same and their x-coordinates move independently until they meet. If and when they meet, let them stay together). Let $\check{P}_s$ be the probability law of the coupled paths $\{(X_{1,n}, X_{2,n})\}_{n=1}^{\infty}.\{X_{1,n}\}_{n=1}^{\infty}$ moves according to $P_s^{v_1}$ and $\{X_{2,n}\}_{n=1}^{\infty}$ moves to $P_s^{v_2}$. Using the strong Markov property, (17) reduces to proving

(18) For almost all s, $\check{P}_s(X_{1,n} = X_{2,n}$ for sufficiently large $n) = 1$.

For any $v \in \mathbb{Z}^2$, let v' be the x-coordinate of v. The probability laws P^{v_1} and P^{v_2} are identical except for translation. Consequently, for any integer n, $E^{v_2}(X'_n) - E^{v_1}(X'_n) = v'_2 - v'_1$. Let $\check{E}_s$ be the expectation operator corresponding to $\check{P}_s$. It follows that

$$\begin{aligned} E(\check{E}_e(X'_{2,n} - X'_{1,n})) &= E(\check{E}_e(X'_{2,n}) - \check{E}_e(X'_{1,n})) \\ &= E(E_e^{v_2}(X'_n)) - E(E_e^{v_1}(X'_n)) = E^{v_2}(X'_n) - E^{v_1}(X'_n) = v'_2 - v'_1 \end{aligned} \tag{19}$$

for any $n \geq 0$.

Assume W.L.O.G. that $v'_2 > v'_1$. Since $v'_2 - v'_1$ is an even integer, $X'_{2,n} - X'_{1,n}$ will always be even and hence cannot become negative without first hitting zero. However, once $X'_{1,n}$ and $X'_{2,n}$ meet, they must stay together. Thus, for the coupled process,

$$X'_{2,n} - X'_{1,n} \geq 0 \qquad \text{for all } n. \tag{20}$$

By (19) it follows that for any positive real r_1 and any integer n,

$$P(\hat{E}_e(X'_{2,n} - X'_{1,n}) > r_1) \leq (v'_2 - v'_1)/r_1$$

For $n \geq 0$, let θ_{n,r_1} be the set of all s such that $\check{E}_s(X'_{2,n} - X'_{1,n}) > r_1$ or such that there is a site at which either i, ii, or iii fails. Then $P(e \in \theta_{n,r_1}) \leq (v'_2 - v'_1)/r_1$. Let σ_1, σ_2 be any positive real numbers and r_2 be any positive integer. Let σ be the first time the two processes meet. Clearly we can choose n_0 so large that

$$E(\check{P}_e(n_0 \leq \sigma \leq n_0 + r_2)) < \delta_1.$$

Let $\equiv_{r_2,\delta_2}$ denote the set of environments s such that

$$\check{P}_s(n_0 \leq \sigma \leq n_0 + r_2) > \delta_2.$$

It follows that

$$P(e \in \Xi_{r_2,\delta_2}) < \frac{\delta_1}{\delta_2}.$$

Now suppose $s \notin (\theta_{n_0,r_1} \cup \Xi_{r_2,\delta_2})$. Then $\check{E}_s(X'_{2,n_0} - X'_{1,n_0}) \leq r_1$. Therefore,

$$\check{P}_s(X'_{2,n_0} - X'_{1,n_0} > r_2) \leq r_1/r_2. \tag{21}$$

Now suppose $0 < X'_{2,n_0} - X'_{1,n_0} \leq r_2$. Since $X'_{2,n_0} - X'_{1,n_0} \neq 0$, the processes have not yet met by time n_0. It is possible that from time n_0 onward the processes will do nothing but move directly toward each other until they meet. By condition iii this will happen with probability at least ε^{r_2}. If this happens, the processes will first meet between times n_0 and $n_0 + r_2$ and this event has $\check{P}_s$ probability not exceeding δ_2. Hence

$$\check{P}_s(0 < X'_{2,n_0} - X'_{1,n_0} \leq r_2)\, \varepsilon^{r_2} \leq \delta_2.$$

Using this, (20), and (21) we get

$$\check{P}_s(X_{2,n_0} \neq X_{1,n_0}) \leq \frac{\delta_2}{\varepsilon^{r_2}} + \frac{r_1}{r_2}.$$

We also know

$$P(e \in \theta_{n_0,r_1} \cup \Xi_{r_2,\delta_2}) \leq P(e \in \theta_{n_0,r_1}) + P(e \in \Xi_{r_2,\delta_2}) \leq \frac{v'_2 - v'_1}{r_1} + \frac{\delta_1}{\delta_2}.$$

By successively choosing r_1 large, then r_2 large, then δ_2 small, then δ_1 small, then n_0 large we can insure successively that

$$\frac{v_2' - v_1'}{r_1}, \frac{r_1}{r_2}, \frac{\delta_2}{\varepsilon^{r_2}}, \frac{\delta_1}{\delta_2}$$

are arbitrarily small. (18) follows immediately. The lemma is proved.

PROOF OF THEOREM 2. Suppose that the R.W.I.R.E. is not recurrent. Let θ be the set of environments s which obey i, ii, and iii at every site, for which the P_s process is not recurrent, and for which any coupled processes such as those of Lemma 2 will eventually meet with probability 1. By Lemma 1, $e \in \theta$ with probability 1. Suppose $s \in \theta$. Consider the coupled processes of Lemma 2. Let A be any event in the tailfield of $\{X_n\}_{n=1}^{\infty}$. Since the two processes stay together after a finite amount of time, A occurs for one of the processes if and only if A occurs for the other. Thus, using coupling, we have proved for any p of proper parity, and any A in the tailfield,

$$P_s(A \mid X_m = p) = P_s(A)$$

or in other words

$$P_s(A \mid X_m) = P_s(A).$$

By the Markov property,

$$P_s(A \mid X_m) = P_s(A \mid X_1, X_2 \cdots X_m)$$

and thus A is independent of $V_{i=1}^{m} X_i$ under P_s. Since m is arbitrary, A is independent of $V_{i=1}^{\infty}(X_i)$ under P_s. $A \in V_{i=1}^{\infty}(X_i)$ so A is independent of itself. $P_s(A) \in \{0, 1\}$.

It remains to be shown that the function $e \to P_e(A)$ has the same value for almost all e, which will follow if we can show it measurable with respect to the tailfield of the i.i.d. random variables $\{e_v\}_{v \in Z^2}$. In fact, for $s_1, s_2 \in \theta$ where s_1 and s_2 differ only on a finite set, we will show $P_{s_1}(A) = P_{s_2}(A)$. Let $\mathbb{F}$ be a finite subset of $\mathbb{Z}^2$ such that s_1 and s_2 agree off $\mathbb{F}$. We couple processes P_{s_1} and P_{s_2} as follows. Let them run independently if either process is on $\mathbb{F}$. If neither process is on $\mathbb{F}$, couple them as in the coupling of Lemma 2. Since neither process is recurrent, eventually both processes will leave $\mathbb{F}$ forever. Once that happens, the two processes are both on s_1 (because $s_2 = s_1$ off $\mathbb{F}$) and by the proof of Lemma 2, they will eventually meet and stay together. Hence, in the coupled process, A occurs for process 1 if it occurs for process 2, i.e., $P_{s_1}(A) = P_{s_2}(A)$. □

Theorem 2 implies that if i, ii, and iii hold, events such as hitting a fixed set infinitely often, or the x-coordinate approaching ∞ must have probability 0 or 1. Note that example 2 obeys all these conditions.

It should be noted that for any R.W.I.R.E. which is not merely a random walk (i.e., where μ has a support of more than one environment), if the R.W.I.R.E. is recurrent, its tailfield cannot be trivial. If $p, q \in \mathbb{Z}^2$, the random variable

$$\limsup_{n\to\infty} (\#\{i: 0 < i < n, X_i = p, X_{i+1} = q\} / \#\{i: 0 < i < n, X_i = p\})$$

is measurable with respect to the tailfield if the process is recurrent. It is in general not constant, though, because it depends upon the environment at p; indeed it equals $e_p(q - p)$.

Theorem 2 can easily be generalized to the n-dimensional nearest neighbor case, where $n - 1$ directions are nonrandom and probabilities in the other direction are bounded below. In particular, the result holds for the one dimensional case, where probabilities of moving left or right are both bounded below.

Although very few zero-one conjectures can be proved when we remove the assumption of nonrandom vertical components, we can still show the following.

THEOREM 3. *Let $\varepsilon > 0$. Suppose that with μ-probability* 1,

$$\text{i)} \quad \min\{g((1, 0)), g((-1, 0))\} \geq \varepsilon.$$

Then the probability that the process hits the Y-axis infinitely often is either zero or one.

PROOF. For any $v \in \mathbb{Z}^2$, let v' be the first coordinate of v. Define events A, B, C and D by

$$A = \{X_n \text{ hits the } Y\text{-axis infinitely often}\}$$
$$B = \{\text{There exists an } N \text{ such that for all } n > N, X'_n > 0\}$$
$$C = \{\text{There exists an } N \text{ such that for all } n > N, X'_n < 0\}$$
$$D = \{X'_n > 0 \text{ for all } n > 0\}.$$

If $P(B) = 0$ and $P(C) = 0$, then $P(A) = 1$ and we are done. Thus we can assume that either $P(B) > 0$ or $P(C) > 0$. Assume W.L.O.G. $P(B) > 0$. We begin our proof by proving that $P(D) > 0$.

Suppose $P(D) = 0$. Then $P_s(D) = 0$ for almost all s. For $v \in Z^2$, define P^v, P_s^v as in the proof of Theorem 2. Let D^v be the event defined by

$$D^v = \{X'_n > v' \quad \text{for all} \quad n > 0\}.$$

Since the probability laws P, P^v are identical except for translation, $P^v(D^v) = P(D) = 0$. Therefore $P_s^v(D^v) = 0$ for almost all s. Let

$$\theta = \{s \in S: P_s^v(D^v) = 0 \quad \text{for all} \quad v \in \mathbb{Z}^2\}.$$

Then $P(e \in \theta) = 1$ because $\mathbb{Z}^2$ is countable. Suppose $s \in \theta$. By the Markov property

$$P_s(X'_n > v' \quad \text{for all} \quad n > N \mid X_N = v) = 0, N \geq 0.$$

Therefore,

$$P_s(X'_n \leq X'_N \quad \text{for some} \quad n > N) = 1, N \geq 0.$$

Thus, since $X'_0 = 0$,

$$P_s(X'_n \leq 0 \quad \text{i.o.}) = 1$$

and thus, since s is an arbitrary element of θ,

$$P(X'_n \leq 0 \quad \text{i.o.}) = 1$$

contradicting our assumption that $P(B) > 0$. We have proven $P(D) > 0$.

For $N > 0$, $v \in \mathbb{Z}^2$ define the event $F_{N,v}$ by

$$F_{N,v} = \{X'_n < v' \quad \text{for all} \quad n < N, X_N = v\}.$$

Let $Y = \{\hat{v} \in \mathbb{Z}^2: \hat{v}' \geq v'\}$. For any $\hat{v} \in Y$, the environment at $\hat{v}$, $e_{\hat{v}}$ is independent of any path $\{X_1, X_2, \cdots X_N\}$ in the set $F_{N,v}$ because $e_{\hat{v}}$, influences only the probabilities of paths exiting from $\hat{v}$ and no path $\{X_1, X_2, \cdots X_n\}$ in $F_{N,v}$ ever exist $\hat{v}$. On the other hand, the conditional probability of the event

$$G_{N,v} = \{X'_n > v' \quad \text{for all} \quad n > N, X_N = v\}$$

given that $X_N = v$ only depends on environments $e_{\hat{v}}$, where $\hat{v} \in Y$, because paths X_N, $X_{N+1}, \cdots$ in $G_{N,v}$ live entirely in Y. Thus

$$P(X'_n > v' \quad \text{for all} \quad n > N \mid X_1, X_2 \cdots X_N\} = P^v(D^v) = P(D)$$

on the set $F_{N,v}$. Define the event G by

$$G = \{\limsup X_n = \infty\}.$$

Equivalently,

$$G = \{\text{there are infinitely many } N \text{ such that there exists } v \in \mathbb{Z}^2 \text{ with } v' > 0 \text{ and } F_{N,v}\}.$$

Hence, on G, there are infinitely many N for which

$$P(X'_n > X'_N \quad \text{for all} \quad n > N \mid X_2 \cdots X_N) = P(D)$$

and $X'_N > 0$. Let A^c be the complement of A. Then it follows that on G,

$$\limsup_{N\to\infty} P(A^c \mid X_1 \cdots X_N) \geq P(D).$$

However, by the martingale convergence theorem,

$$\lim_{N\to\infty} P(A^c \mid X_1 \cdots X_N) = 0$$

on almost all of A. Therefore,

$$P(G \cap A) = 0.$$

Since $P(X'_{n+1} - X'_n = 1 \mid X_1, X_2, \cdots, X_n) \geq \varepsilon$, it follows that $P(X'_{n+R} - X'_n = R \mid X_1, X_2, \cdots, X_n) \geq \varepsilon^R$ for any R. Thus, for any R, $X'_n \geq R$ for infinitely many n on almost all of A (see [1], page 97, problem 9). Therefore,

$$P(G^c \cap A) = 0$$

and hence $P(A) = 0$. □

Theorem 3 can be generalized to arbitrary R.W.I.R.E. as follows. Let ℓ be a linear functional on C. Suppose there exists $\varepsilon > 0$ such that with μ probability 1,

$$\textstyle\sum_{\{v\in C:\ell(v)>\varepsilon\}} g(v) > \varepsilon \quad \text{and} \quad \sum_{\{v\in C:\ell(v)<-\varepsilon} g(v) > \varepsilon.$$

Then the event

$$\{\text{both sets } \{n : \ell(X_n) \geq 0\} \text{ and } \{n : \ell(X_n) \leq 0\} \text{ are infinite}\}$$

has probability either 0 or 1.

APPENDIX

Here we prove that in the proof of Theorem 1, (10), (11) and the corresponding statements for the $\check{P}$ process imply (8) and (9).

Since we are not concerned with the behavior of $\ell(X_n)$ after time τ it will not hurt to redefine the $\hat{P}$ process on $\theta_M \cup (C - D)$ so that (10) and (11) hold regardless of where X_n is. Consider the process to be so redefined.

Since

$$\int H(\lambda)\, d(\lambda) < \infty, \liminf_{\lambda\to\infty} \lambda H(\lambda) = 0.$$

Choose an integer I so that

$$IH(I) + \int_I^\infty H(\lambda)\, d(\lambda) < \varepsilon/3.$$

Let

$$Y_n = \begin{cases} \ell(X_{n+1}) - \ell(X_n) & \text{if } \ell(X_{n+1}) - \ell(X_n) > I \\ 0 & \text{otherwise.} \end{cases}$$

Then by (10),

$$\hat{E}(Y_n | X_n) < \varepsilon/3$$

and by (11),

$$\hat{E}(\ell(X_{n+1}) - \ell(X_n) - Y_n \mid X_n) > 2\varepsilon/3.$$

Let

$$p = \hat{P}(\ell(X_{n+1}) - \ell(X_n) - Y_n < \varepsilon/3 \mid X_n).$$

Since

$$\ell(X_{n+1}) - \ell(X_n) - Y_n \leq I,$$

it follows that

$$2\varepsilon/3 < (\varepsilon/3)\, p + I(1-p);\, p < (3I - 2\varepsilon)/(3I - \varepsilon).$$

Thus

$$\hat{P}(\ell(X_{n+1}) - \ell(X_n) \geq \varepsilon/3 \,|\, X_n) \geq \hat{P}(\ell(X_{n+1}) - \ell(X_n) - Y_n \geq \varepsilon/3 \,|\, X_n) > \varepsilon/3I - \varepsilon, \qquad n \geq 0. \tag{22}$$

Consequently, for any m

$$\hat{P}(\ell(X_{n+m}) - \ell(X_n) > m\,\varepsilon/3 \,|\, X_n) > (\varepsilon/(3I - \varepsilon))^m$$

from which it follows that for any m, the event $\ell(X_{n+m}) - \ell(X_n) > m\,\varepsilon/3$ will eventually occur with probability 1 and thus the process is unbounded (see problem 9, page 97 of [1]). This implies that (8) holds for $\hat{P}$ and an exactly analogous argument proves that (8) holds for $\check{P}$.

Let

$$Z_{n\ 1} = \ell(X_{n+1}) - \ell(X_n) - E(\ell(X_{n+1}) - \ell(X_n) \,|\, X_n).$$

By (10),

$$|E(\ell(X_{n+1}) - \ell(X_n) \,|\, X_n)| \leq \int_0^\infty H(s)\, ds$$

and hence for $\lambda > 2 \int_0^\infty H(s)\, ds$,

$$\hat{P}(|Z_{n+1}| > \lambda) \leq \hat{P}(|\ell(X_{n+1}) - \ell(X_n)| > \lambda/2) = \hat{E}(\hat{P}(|\ell(X_{n+1}) - \ell(X_n)| > \lambda/2 \,|\, X_n) \leq H(\lambda/2).$$

Let

$$G(\lambda) = \begin{cases} 1 \text{ if } \lambda < 2 \int_0^\infty H(s)\, ds \\ \min(1, H(\lambda/2)) \text{ otherwise.} \end{cases}$$

G is a non-negative real-valued non-increasing function on the non-negative reals with

$$G(0) = 1, \qquad \int_0^\infty G(\lambda)\, d\lambda < \infty$$

and

$$\hat{P}(|Z_{n+1}| > \lambda \,|\, X_n) \leq G(\lambda)$$

for all λ. Since $E(Z_{n+1} \,|\, Z_0, Z_1, \cdots, Z_n) = E(Z_{n+1} \,|\, X_n) = 0$ for all n, it follows from a result of H. Kesten (Lemma 1 of [4]) that there exists a function f dependent only on G and ε (hence only on H and ε) such that $\lim_{x\to\infty} f(x) = 0$ and for any real $R > 0$ and integer $N > 0$,

$$P(\exists\, n: \textstyle\sum_{1=N+1}^{N+n} (Z_i) < -R - (\varepsilon/2)n \,|\, X_N) < f(R).$$

By (11), setting $N = 0$, with probability at least $1 - f(R)$,

$$\ell(X_n) = \textstyle\sum_{i=1}^n (Z_i) + \sum_{i=1}^n \hat{E}(\ell(X_n) - \ell(X_{n-1}) \,|\, X_{n-1}) \geq -R - (\varepsilon/2)n + \varepsilon n = (\varepsilon/2)n - R$$

for all n. Thus (9) holds for $\hat{P}$.

Choose R so that $f(R) < 1$ and an integer $N > 3R/\varepsilon$. By (22),

$$\ell(X_i) - \ell(X_{i-1}) > \varepsilon/3 \qquad \text{for all} \quad 0 \leq i \leq N$$

with $\hat{P}$ probability at least $(\varepsilon/(3I-\varepsilon))^N$. If that happens, then $\ell(X_N) > R$, $\ell(X_i) > 0$ for $1 \le i \le N$ and if for all $n \ge 1$

$$\sum_{i=N+1}^{N+n} (Z_i) \ge -R - (\varepsilon/2)n,$$

then for any $n \ge 1$,

$$\begin{aligned}\ell(X_{N+n}) &= \ell(X_N) + \sum_{i=N+1}^{N+n} (\ell(X_i) - \ell(X_{i-1})) \\ &= \ell(X_N) + \sum_{i=N+1}^{N+n} (Z_i) + \sum_{i=N+1}^{N+n} \hat{E}(\ell(X_i) - \ell(X_{i-1}) \mid X_{i-1}) \\ &> R - R - (\varepsilon/2)n + \varepsilon n > 0.\end{aligned}$$

Thus for any M,

$$\hat{P}(B_M) \ge (\varepsilon/(3I-\varepsilon))^N(1-f(R)).$$

By an analogous argument,

$$\check{P}(B_M) \ge (\varepsilon/(3I-\varepsilon))^N(1-f(R))$$

and (9) holds for $\check{P}$. □

Acknowledgments. Since this is my thesis, I would like to thank my thesis advisor, Harry Kesten, for his substantial help guiding me through the topic, and assistance in writing the paper. I also wish to thank Kiyosi Ito and Frank Spitzer for their very personal help while teaching me probability theory. Finally, I wish to thank Barry Simon for pointing out an error in the original paper.

REFERENCES

[1] BREIMAN, L. (1968). *Probability*, Addison-Wesley, New York.
[2] GRIFFEATH, D. (1974–75). A maximal coupling for Markov chains. *Z. Wahrsch. verw. Gebiete* **31** 95–106.
[3] KESTEN, H. (1976). Recurrence criteria for multidimensional linear birth and death processes. *Adv. Appl. Probability* **8** 58–87.
[4] KESTEN, H. (1977). A renewal theorem for random walk in a random environment. *Proc. Symposia Pure Math.* **31** 67–77.
[5] KESTEN, H., KOSLOV, M., and SPITZER, F. (1975). A limit law for random walk in a random environment. *Compositio Math.* **30** 145–168.
[6] RITTER, G. A. (1976). Random walk in a random environment, critical case. Ph.D. thesis, Cornell University.
[7] SOLOMON, F. (1975). Random walks in a random environment. *Ann. Probability* **3** 1–31.

STANFORD UNIVERSITY
DEPARTMENT OF MATHEMATICS
STANFORD, CALIFORNIA 94305