Multidimensional Random Walks in Random Environments with Subclassical Limiting Behavior

Richard Durrett[*]

Department of Mathematics, University of California, Los Angeles, California[**], USA

Abstract. In this paper we will describe and analyze a class of multidimensional random walks in random environments which contain the one dimensional nearest neighbor situation as a special case and have the pleasant feature that quite a lot can be said about them. Our results make rigorous a heuristic argument of Marinari et al. (1983), and show that in any $d < \infty$ we can have (a) X_n is recurrent and (b) $X_n \sim (\log n)^2$.

1. Introduction

In this paper we will describe some new results concerning multidimensional random walks in random environments (hereafter abbreviated RWRE). To put our results in perspective we will start by describing the one dimensional nearest-neighbor model and stating the results that we will generalize. Let $p_x \in (0, 1)$, $x \in Z$ be a stationary sequence [i.e., the distribution of $(p_{x-j}, \ldots p_{x+j})$ is independent of x]. This sequence is the environment. We think of it as being generated at time 0 and then fixed for all time while a particle wanders around on Z moving as a discrete time Markov Chain with transition probabilities $p(x, x+1) = p_x$, $p(x, x-1) = 1 - p_x$, i.e., if it is at x it flips a coin with these probabilities to determine its next position. If we let X_n be the position of the particle at time n then X_n is our RWRE. When the environment is nonrandom (and by stationarity constant) X_n is a random walk, so with results about these systems in mind, the first question we would like to ask is "Will X_n return to 0 infinitely often?". If $p_x \equiv p$ then the answer is yes if and only if $p = 1/2$, so one might guess than the answer in the random case is $Ep_x = 1/2$. This is wrong, however, but the correct answer is not difficult to determine and is given by

Theorem A. *Suppose* p_x, *$x \in Z$ is ergodic. If $E \log(1 - p_x/p_x) \neq 0$ then for almost every environment $P(X_n = 0$ infinitely often$) = 0$. If $E \log(1 - p_x/p_x) = 0$ then for a.e. environment $P(X_n = 0$ i.o.$) = 1$.*

Since X_n is a birth and death chain and one has explicit formulas for the probability of hitting a before b the proof is almost trivial.

Solomon (1975, p. 4) gives the result in the i.i.d. case but it is trivial to extend his result to prove the first result above (all that was used is the strong law of large numbers which can be replaced by the ergodic theorem). The second result is more subtle. It is a consequence of a "well known" fact about random walks with stationary increments [see Durrett, 1984a, (2) on p. 1011]. A nice completely different self-contained proof can be found in Kotani (1985).

Solomon (1975) showed than when $E \log(1 - p_x/p_x) < 0$ (and hence $X_n \to \infty$ almost surely) we have $X_n/n \to$ a positive limit if and only if $E(1 - p_x/p_x) < 1$. This was the first indication that RWRE's can have much different limiting behavior from random walks (where $p_x \equiv p$), a fact which became clearer when Kesten et al. (1975) showed that in the sublinear case ($E(\log(1 - p_x/p_x) < 0$ and $E(1 - p_x/p_x) > 1$), if we let $k \in (0, 1)$ be the unique positive number so that $E(1 - p_x/p_x)^k = 1$, then as $h \to \infty X_n/n^k$ converges in distribution (hereafter denoted $\Rightarrow$) to a nondegenerate limit Z.

From the results above we see that when we are near the recurrent case X_t grows like a small power, so it is natural to guess that when $E \log(1 - p_x/p_x) = 0$, $|X_t|$ grows like a power of $\log t$. Ritter (1976) showed

$$\max_{0 \leq m \leq n} |X_m|/(\log n)^2 \Rightarrow Z,$$

where Z is a nondegenerate random variable. A complete solution of the problem had to wait until Sinai (1982) proved

Theorem B. *Suppose the p_x are iid with $\varepsilon \leq p_x \leq 1 - \varepsilon$, $E \log((1 - p_x)/p_x) = 0$, and $p_x \not\equiv \frac{1}{2}$, then as $n \to \infty$,*

$$X_n/(\log n)^2 \Rightarrow Z,$$

where Z is a nondegenerate random variable.

Our aim in this paper is to generalize the results of Solomon, Ritter, and Sinai to a class of multidimensional models. The new results are called Theorems 1–3, respectively. In Sect. 2 we will describe our model and state and prove Theorem 1. In Sect. 3 we will state Theorems 2 and 3 and describe their connection with the work of Marinari et al. (1983). In Sect. 4 we will apply Theorems 2 and 3 to two classes of examples: one dimensional nearest neighbor systems and a family of multidimensional processes constructed from random potentials with Gaussian distributions. In the latter examples X_n is recurrent (in any $d < \infty$), and depending upon the value of a parameter $\beta \in (0, 2)$ we have $X_n \sim (\log n)^{2/\beta}$.

The last conclusion casts doubt on (but does not contradict) speculations of Derrida and Luck (1983), Luck (1983, 1984), and Fisher (1984) that [in a model where the $p(x, \cdot) x \in Z^d$ are independent] "the critical dimension is 2. For $d > 2$ the mean square displacement $\langle x^2(t) \rangle$ is linear in time." We will discuss this briefly at the end of Sect. 4 and then turn to the proofs of Theorems 2 and 3 in Sect. 5.

2. Description of the Model, Theorem 1

Let $V: R^d \to R$ be a random function with stationary increments [i.e., the distribution of $(V(x+y_1) - V(x), \dots V(x+y_n) - V(x))$ does not depend on x]. We call V a random potential and think of the graph $(x, V(x)) \in R^{d+1}$ as a random hillside. For all $x, y \in Z^d$ with $|x - y| = 1$, we let $\alpha(x, y) = \exp(-V(x+y/2))$, the minus sign being put in so that our process will have a preference for stepping downhill, and we define

$$\alpha(x) = \sum_y \alpha(x, y), \quad p(x, y) = \alpha(x, y)/\alpha(x),$$

$\alpha(x, y) \geq 0$ so we have $p(x, y) \geq 0$ and $\sum_y p(x, y) = 1$.

In one dimension a little arithmetic shows

$$p(x, x+1) = \left(1 + \frac{\alpha(x, x-1)}{\alpha(x, x+1)}\right)^{-1} = (1 + \exp(V(x+\tfrac{1}{2}) - V(x-\tfrac{1}{2})))^{-1},$$

so we can generate any stationary sequence by picking V to have stationary increments, and in particular if we want the $p(x, x+1)$, $x \in Z$ to be independent then we want $V(x+\tfrac{1}{2})$, $x \in Z$ to be a random walk.

In dimensions $d > 1$ (or in $d = 1$ when nonnearest neighbor steps are allowed) the computations in the last paragraph break down and in fact not every RWRE can be generated in this way. Our definitions imply that

$$\alpha(x)p(x, y) = \alpha(x, y) = \alpha(y, x) = \alpha(y)p(y, x), \tag{1}$$

so

$$\frac{\alpha(x)}{\alpha(y)} = \frac{p(y, x)}{p(x, y)},$$

and if $x_0, x_1, \dots x_n$ is a sequence of points with $x_n = x_0$ and $p(x_{m-1}, x_m) > 0$ for $m = 1, \dots, n$

$$\prod_{m=1}^n \frac{p(x_m, x_{m-1})}{p(x_{m-1}, x_m)} = \prod_{m=1}^n \frac{\alpha(x_{m-1})}{\alpha(x_m)} = 1,$$

since $\alpha(x_m) = \alpha(x_0)$. From the last equation we see that if the vectors $\mathbf{p}(x) = (p(x, y) : |x - y| = 1)$ are independant and non-constant this condition will almost surely not be satisfied.

The observations in the last paragraph show that the "usual" case in which the $\mathbf{p}(x) x \in Z^d$ $d > 1$ are independent is never a special case of our model. This is a little disappointing, but having sacrificed some generality we have put ourselves in the "reversible" case and given ourselves a powerful tool: the "Dirichlet principle." To explain the words in quotation marks we observe that (1) implies

$$\sum_x \alpha(x)p(x, y) = \sum_x \alpha(y)p(y, x) = \alpha(y),$$

so α is a stationary measure for p, and furthermore satisfies the "detailed balance condition"

$$\alpha(x)p(x, y) = \alpha(y)p(y, x),$$

i.e., in equilibrium the amount of mass which goes from x to y in one step is equal to the amount which goes from y to x. When this happens we say $\alpha(x)$ is a reversible measure and we have

The Dirichlet Principle. Let $\Lambda \subset Z^d$ be a set with $0 \notin \Lambda$ and $Z^d - \Lambda$ finite. The function which minimizes the "energy" $\sum\limits_{xy} \alpha(x, y)(h(x) - h(y))^2$ among all the functions with $h(0) = 1$ and $h(x) = 0$ for $x \in \Lambda$ is $h(x) = P_x(T_0 < T_\Lambda)$, where $T_0 = \inf\{n \geq 0 : X_n = 0\}$ and $T_\Lambda = \inf\{n \geq 0 : X_n \in \Lambda\}$, and furthermore the minimum energy is $2\alpha(0)P_0(T_\Lambda < T_0^+)$, where $T_0^+ = \inf\{n \geq 1 : X_n = 0\}$.

Proof. See Griffeath and Liggett (1982, p. 885) or Liggett (1985, Sect. 2.6).

The first part of this is the analogue of the classical Dirichlet principle, i.e., the function which minimizes $\int\limits_G |\nabla u|^2 dx$ among all the functions (in a suitable class) which are equal to f on ∂G is the harmonic function with those boundary values.

The second part of the conclusion which gives the minimum energy is the more important for it implies that as the minimum energy $\downarrow 0$ as $\Lambda \downarrow \phi$ if and only if the process is recurrent [i.e., $P_0(T_\Lambda < T_0^+) \to 0$] and this gives us the following

Comparison Theorem. If $\tilde{\alpha}(x, y) \leq \alpha(x, y)$ for all x, y, then $\tilde{\alpha}$ is more recurrent that α, i.e.,

    (a) if α is recurrent then $\tilde{\alpha}$ is,

    (b) if α is transient then $\tilde{\alpha}$ is

[where we have used "α is transient" as shorthand for "the Markov Chain with transition probabilities $p(x, y) = \alpha(x, y)/\alpha(x)$ is transient"].

If we use the comparison theorem in $d = 2$ with $\alpha(x, y) = 1$ for $|x - y| = 1$ and $\tilde{\alpha}(x, y) = 1$ if $(x, y) \in$ "the infinite cluster in bond percolation" or any other subgraph of Z^2 we see that a random walk on such an object is recurrent.

If we let $\alpha(x, y) = \exp(-V(x + y)/2)$ and compare with

$$\tilde{\alpha}_i(x, y) = \begin{array}{ll} \alpha(x, y) & \text{if} \quad y = x \pm e_i \\ 0 & \text{otherwise} \end{array},$$

we get

Theorem 1. *Suppose the random potential V is such that the vector valued process $\mathbf{p}(x) = (p(x, y), |x - y| = 1)$ is stationary and ergodic. A necessary condition for X_n to be recurrent is*

$$E \log \left(\frac{p(x, x + e_i)}{p(x, x - e_i)} \right) = 0 \quad \text{for} \quad i = 1, \ldots d.$$

Proof. $\tilde{\alpha}_i$ is a family of one dimensional RWRE. If $E \log(p(x, x + e_i)/p(x, x - e_i)) \neq 0$, then it follows from Theorem A that $\tilde{\alpha}_i$ is transient and from the comparison theorem that α is.

I find it amazing that in the reversible case the recurrence condition in $d \geq 2$ is such a mindless generalization of the answer in $d = 1$. Theorem 1 should be thought of as the RWRE analogue of the following result for random walks.

Theorem 1'. *If S_n is a random walk (i.e., $\mathbf{p}(x)=\text{const}$) then a necessary condition for recurrence is*

$$E(S_n^i - S_{n-1}^i) = 0 \quad for \quad 1 = 1, \ldots d.$$

In words Theorem 1' says that "in recurrent random walks the increments must have mean zero, while Theorem 1 says that the same thing is true for RWRE if we define the mean of the increments to be the vector

$$E \log(p(x, x+e_i)/p(x, x-e_i)).$$

With such a simple answer in the reversible case it is tempting to conjecture that Theorem 1 gives the answer in general, but we will not do so because we believe that this is not correct and that in $d > 1$ the condition for recurrence will not be just in terms of the one-dimensional distributions.

3. Theorems 2 and 3

Theorem 1 generalizes half of Theorem A to higher dimensions. Theorems 2 and 3 give our extension of Theorem B. The first result makes precise a heuristic argument of Marinari et al. (1983) (which we discovered after we proved our theorem) so we will indicate their reasoning before we state our result.

On p. 2 of their paper they say (at various places up and down the page): "we assume that V belongs to an ensemble which is invariant (at least for the large distance behavior) under the transformation $V \to V^*$, where

$$V^*(\lambda x) - V^*(0) = \lambda^{\alpha}(V(x) - V(0)).$$

"We choose the diffusion equation to be

$$\frac{\partial c}{\partial t} = kV \cdot J, \quad J = Vc + cVV. \tag{$*$}$$

This is the continuous limit of a random walk on a lattice with nearest neighbor transition $i \to j$ proportional to $\exp\frac{1}{2}[v_i - v_j]$."

"The diffusion ($*$) is chosen such that it has an equilibrium distribution $c \sim \exp(-v)$ in a bounded box.... In the landscape created by the potential V in R^N one expects that $X(t)$ will occasionally go through a mountain pass and then rapidly relax to equilibrium in the intermediate valleys."

"Other things being similar, the flux through the mountain pass is proportional to the density $c = \exp(-v)$. When distances are multiplied by λ the height of the mountain pass is multiplied by λ^{α} and the flux through the pass changes from $g(1)\exp(-v)$ to $g(\lambda)\exp(-\lambda^{\alpha}v)$, where $g(\lambda)$ is a polynomial. The time scale is correspondingly multiplied by $g(1)/g(\lambda)\exp((\lambda^{\alpha}-1)V)$. Conversely multiplication of time by τ corresponds to multiplication of distances by a factor $\lambda(\tau) \sim (\log \tau)^{1/\alpha}$."

The argument in the last paragraph is not a proof and in fact their conclusion is not even believed by some. Fisher (1984) says, "We will argue that contrary to the logarithmic time dependence suggested in ref. 1 (Marinari et al.), in any dimension greater than two the mean square displacement of random walks in such random environments will be linear in time." Fisher's conclusions are based on renormaliz-

ation group calculations, a frequently used and usually reliable tool of the modern physicist, but this time it turns out that Marinari et al. are right.

Theorem 2. *Suppose* $V(0)=0$, V *is continuous, has stationary increments, and satisfies*

(i) *as* $a\to\infty$ $V(az)/a^\alpha$ *converges weakly to a limit* $W(z)$,

(ii) $z\to W(z)$ *is continuous,*

(iii) *with probability* 1, $G(\varepsilon)\equiv$ *the component of* $\{z:W(z)<\varepsilon\}$ *containing* 0 *is bounded for some* $\varepsilon>0$ *(which may depend on* ω*).*

Then for a.e. *environment*

(a) X_n *is recurrent,*

(b) $\sup_{m\leq n}|X_m|=0(\log^{1/\alpha}n)$ *that is for all* $\varepsilon>0$ *there is a* $K<\infty$ *so that*

$$\limsup_{n\to\infty} P\left(\sup_{m\leq n}|X_m|\geq K\log^{1/\alpha}n\right)\leq\varepsilon.$$

Assumption (i) is clearly in the spirit of the assumption Marinari et al. made concerning the long distance behavior of V, and I do not think they would object to (ii): continuity is natural if you think of V as a mountain range. (iii) is a technical assumption and may look a little mysterious at first. In words it says that "any path to ∞ must rise to a height $\geq\varepsilon(\omega)>0$ at some point." A simple argument using scaling and translation invariance (see Sect. 4) shows that if W has increments which are stationary, ergodic and satisfy (i) and (ii), then (iii) can only fail if the same paths have no valleys, i.e., there is no open set G which contains a point x_0 with $W(x)>W(x_0)$ for all $x\in\partial G$. This would not happen if by some miracle W was always a harmonic function, but in most other cases we expect this to be satisfied, and in any case the existence of valleys is necessary for the heuristic argument given above and our proof to be given below.

Theorem 2 gives bounds on the movement of X_n. The next result tries to describe where X_n goes and how long it stays there.

Theorem 3. *Let* $m(a)=\inf\{W(z):z\in G(a)\}$, *let* $a^*=\inf\{a:a-m(a)>1\}$, *and let* $T_n(a)=\inf\{m:V(X_m)/\log n>a\}$.

(c) *If* $\delta>0$ *then as* $n\to\infty$

$$P(T_n(a^*+\delta)\leq n)\to 0, \qquad P(T_n(a^*-\delta)\leq n)\to 1.$$

(d) $\log|\{m\leq n:X_m=0\}|/\log n\Rightarrow a^*$,

and (e) *the joint distributions of the occupation times*

$$\log|\{m\leq T_n(a):X_m=x\log^{1/\alpha}n\}|$$

converge to those of

$$\begin{cases}0 & x\in G(a)\\ a-W(x) & x\in G(a)\end{cases}.$$

(d) and (c) sharpen the conclusions (a) and (b) in Theorem 1. (c) is the multidimensional analogue of Ritter's limit theorem for $\max\{|X_m|:m<n\}$. (e) is our attempt to generalize Sinai's result. Its strength is that it tells us in great detail about the sites visited at times $m\leq T_n(a)$. For example, it allows us to conclude that
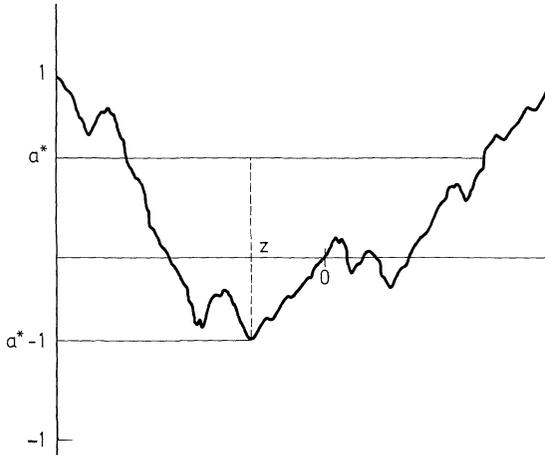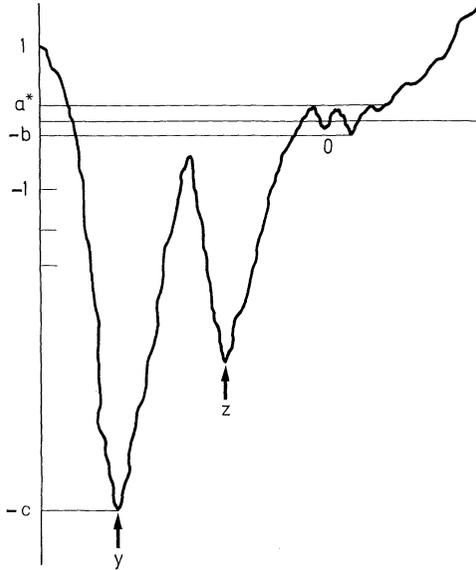
Fig. 1

Fig. 2

$T_n(a) \approx n^{a-m(a)}$, extending the result in (c), but the last observation also points out the weakness of our result: $a \rightarrow m(a)$ is nonincreasing and discontinuous, so there are large gaps between the times we can say something about.

We had hoped (and in fact claimed in some lectures we have given) to prove a result about the limiting behavior of the empirical distribution $\frac{1}{n} |\{m \leqq n : X_m \in \cdot\}|$. When $a^* - m(a^*) = 1$, it is a simple consequence of (c) and (e) in Theorem 3 that

$$\frac{1}{n} |\{m < n : X_m \in \cdot\}| \approx \delta_Z, \qquad (*)$$

where δ_Z denotes a point mass at Z the point indicated in Fig. 1.

Fig. 3

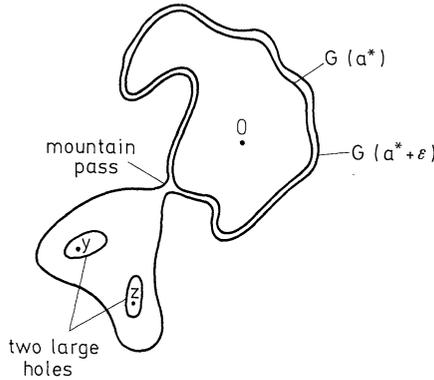The figure labels: G(a*), 0, mountain pass, G(a*+ε), y, z/ζ, two large holes

Unfortunately Fig. 1 is not the only possibility. There is also Fig. 2. From (e) we see that X_n will leave the little valleys near zero at a time $\leq n^{a^*+b+\varepsilon}$ with $a^*+b<1$ but not exit $G(a^*+\delta)$ until a time $\geq n^{a^*+c-\varepsilon}$ with $a^*+c>4$. In the outcome drawn it is not hard to show that X_m will reach Z in time $o(n)$ and then not escape from the corresponding valley until time n, so ($*$) is again true.

The last argument makes crucial use of the fact that in one dimension you have to go through the valley marked Z to get to the one marked Y. This is not true in two dimensions (see Fig. 3 this is a "top view" of something which from the "side view" would look like Fig. 2) and we have not been able to solve the problem of predicting which hole the process will be trapped in at time n.

4. Examples

In this section we will apply the abstract result in the last section to two examples. The first is the one dimensional case where we recover some earlier results. In the second we describe a family of examples which show that the behavior described in Theorem 2 can occur in any $d<\infty$.

Example 1. Suppose that the increments $V(x+\frac{1}{2})-V(x-\frac{1}{2})$, $x \in Z$ are such that the central limit theorem holds in the form $V(at)/a^{1/2} \Rightarrow \sigma B_t$, where B_t, $-\infty<t<\infty$ is a (two-sided) Brownian motion. This holds in particular if the $V(x+\frac{1}{2})-V(x-\frac{1}{2})$ are independent and have mean 0 and variance $0<\sigma^2 = E(V(x+\frac{1}{2})-V(x-\frac{1}{2}))^2 <\infty$, but the weaker statement in the last sentence is obviously all we need to conclude that (i) holds.

It is well known (see, e.g., Durrett, 1984b, Chap. 1) that Brownian motion has continuous paths and $\lim\sup_{t\to\infty} B_t = \infty$, so (ii) and (iii) hold and we can conclude from Theorem 2 that (a) the corresponding RWRE is recurrent (a fact which also follows from Theorem A) and (b) $\sup_{m<n} |X_m| = 0(\log^2 n)$. The last conclusion is almost Ritter's result. His theorem is one of the things we get from Theorem 3. Let $T_a^+ = \inf\{t>0 : B_t > a\}$, let $T_a^- = \sup\{t<0 : B_t > a\}$. It follows from (c) [with a little help from (e)] that

$$\frac{1}{\log n}\left(\min_{m\leq n} X_m, \max_{m\leq n}\right) \Rightarrow (T_{a^*}^-, T_{a^*}^+).$$

Part (d) is a new result. It shows that there are approximately n^{a^*} visits to 0 by time n. If you notice that for ordinary random walk the number of visits to 0 is $\sim n^{1/2}$ in $d=1$ and $\sim (\log n)$ in $d=2$, and you like mysterious phrases you could say that "the effective dimension is $2 - 2a^*$."

Finally part (e) gives the limiting behavior of the occupation times. With this result you can prove Sinai's (1982) theorem but there is no point. Schumacher (1984) has already proved better results.

While Theorem 2 asserts the possibility of subdiffusive behavior in any $d < \infty$, it is not complete until we show there is an example which satisfies the hypotheses.

Example 2. Let $V(x)$, $x \in R^d$ be a Gaussian process with $V(0) = 0$, $EV(x) = 0$, and $EV(x)V(y) = |x|^\beta + |y|^\beta - |x-y|^\beta$, where $0 < \beta < 2$.

A theorem of Schoenberg (1983) shows that the last expression is positive definite and hence a legitimate covariance function [see Gangolli (1965) or Pitt (1978) for facts we will use below]. It is easy to see that V has stationary increments,

$$E(V(x+z) - V(z))(V(y+z) - V(z)) = |x+z|^\beta + |y+z|^\beta - |x-y|^\beta$$

$$- |x+z|^\beta - |z|^\beta + |x|^\beta - |z|^\beta - |y+z|^\beta + |y|^\beta + 2|z|^\beta = |x|^\beta + |y|^\beta - |x-y|^\beta.$$

When $\beta = 1$, V is called Lévy's multiparameter Brownian motion because if $s, t > 0$,

$$EV(sx)V(tx) = (s \wedge t)x,$$

$$EV(-sx)V(tx) = 0,$$

so if $|x| = 1$, $V(tx)$ $t \in (-\infty, \infty)$ is a (two-sided) Brownian motion, and it follows from the observation above that V is a Brownian motion along any line.

The V's described above are the random potentials which give rise (by the rules described in Sect. 2) to the RWRE with $X_t \sim (\log t)^{2/\beta}$. We will try to argue below that the behavior in the independent case will be similar, so to prepare for that we would like to observe that although the V fluctuates wildly, the individual $p(x, y)$ have nonsingular distributions. Suppose without loss of generality $x = 0$, $y = e_1$,

$$p(0, e_1) = \exp(-V(e_1/2)) / \sum_y \exp(-V(y/2))$$

$$= \left[\sum_y \exp(V(e_1/2) - V(y/2)) \right]^{-1}$$

$$\geq \left[2d \max_y \exp(V(e_1/2) - V(y(2))) \right]^{-1}.$$

Since the increments have normal distributions with mean 0 and variance 1, it is easy to see that these are C, $\gamma \in (0, \infty)$ so that

$$P\left(\max_y (V(e_1/2) - V(y/2)) > \lambda \right) \leq Ce^{-\gamma\lambda^2},$$

and consequently

$$P(p(0, e_1) \leq \varepsilon) \leq C'e^{-\gamma'(\log \varepsilon)^2} = C'\varepsilon^{-\gamma' \log \varepsilon},$$

so $E(p(0, e_1)^{-k}) < \infty$ for all $k < \infty$.

The last computation shows that the $(\log t)^{2/\beta}$ behavior does not come from a few bad bonds. The cause of this behavior will be painfully obvious in the proofs in Sect. 5, so we turn now to the tack of checking that V satisfies the hypotheses of Theorem 2.

Brownian motion has the scaling property $B_{at} \overset{d}{=} a^{1/2} B_t$. A moments thought generalizes this to the examples above as $V(ax) \overset{d}{=} a^{\beta/2} V(x)$, so (i) holds with $\alpha = \beta/2$ and $W = V$. Levy (1948) has shown that (ii) holds [see also Gangolli (1965, p. 19) or Pitt (1978, p. 332], so that brings us to the task of checking (iii).

Let $\Omega_\varepsilon = \{G(\varepsilon)$ is bounded$\}$. We will first show that $P(\Omega_\varepsilon) > 0$. To do this we will use the "local nondeterministic" property of these processes first proved by Pitt (1978, see Lemma 7.1, p. 322).

(1) There is positive constant C so that for all $t \in R^N$ and any positive $r < |t|$,

$$\text{Var}(W(t) | W(s), |t - s| \geq r) = C r^\beta .$$

Since the conditional distribution of $W(t)$ given $W(s)$, $|t - s| \geq r$ is Gaussian it follows that with positive probability $W(t) < \inf \{W(s); |t - s| = r\}$, and it follows from the stationary increments property that $P(\Omega_\varepsilon) > 0$.

To improve the last conclusion to $P(\Omega_\varepsilon) = 1$ we observe that scaling implies $P(\Omega_\varepsilon)$ is independent of ε and $\varepsilon' < \varepsilon$ implies $\Omega_\varepsilon \subset \Omega_{\varepsilon'}$ so all the Ω_ε are the same (modulo null sets). The last situation is ridiculous if $P(\Omega_\varepsilon) < 1$, but to reach a contradiction easily we have to make matters worse. Let $\Omega_\varepsilon(x) = \{G(\varepsilon, x)$ is bounded$\}$, where $G(\varepsilon, x) = $ the component of $\{y : W(y) < \varepsilon\}$ containing x. If we let $M(x) = \sup \{W(\theta x) : \theta \in [0, 1]\}$, then for $\varepsilon > M(x)$ we have $G(\varepsilon, 0) = G(\varepsilon, x)$, and since the events $\Omega_\varepsilon(0)$ and $\Omega_\varepsilon(x)$ are independent of ε it follows that $\Omega_\varepsilon(0) = \Omega_\varepsilon(x)$ for all x, ε.

The last state of affairs is clearly impossible given (1). Let $\Omega_{\varepsilon, R}(x) = \{G(\varepsilon, x) \subset B(x, R)\}$. From (1) it follows that there is a $\delta > 0$ so that if $|x| > R + 1$,

$$P(\Omega^c_{\varepsilon, R}(0) \cap \Omega_{\varepsilon, 1}(x)) \geq \delta P(\Omega^c_{\varepsilon, R}(0)) .$$

If $\varrho \equiv P(\Omega^c_{\varepsilon, R}(0)) > 0$ and R is so large that $P(\Omega_{\varepsilon, R}(0)) \geq P(\Omega_\varepsilon(0)) - \delta \varrho / 2$, then we have a contradiction since

$$P(\Omega_\varepsilon(0)) = P(\Omega_\varepsilon(0) \cup \Omega_\varepsilon(x))$$

$$\geq P(\Omega_{\varepsilon, R}(0)) + P(\Omega^c_{\varepsilon, R}(0) \cap \Omega_{\varepsilon, 1}(x))$$

$$\geq P(\Omega_\varepsilon(0)) - \frac{\delta \varrho}{2} + \delta \varrho > P(\Omega_\varepsilon(0)) .$$

Having verified (iii) we can now apply Theorem 2 to conclude that the corresponding RWRE (a) is recurrent and (b) has

$$\sup_{m \leq n} |X_m| = 0(\log^{2/\beta} n) .$$

The last conclusion is in sharp contrast to (i) all previous regorous results in $d > 1$ and to (ii) physicists' speculations that [for models in which the $\mathbf{p}(x)$ are independent] "the critical dimension is 2," so we would like to take a moment and explain (i) what makes our system different from those previously studied and (ii)

why we think that the independent case will have behavior similar to that found above.

The answer to both questions can be found in the observation that the stationary measure for our process

$$\alpha(x) \approx 2d \exp(-V(x)) \approx \exp(\pm C|x|^\alpha)$$

fluctuates violently in contrast to all previous results in dimension $d > 1$ in which we have $\varepsilon \leq \alpha(x) \leq \varepsilon^{-1}$ (or in a few works moment conditions on α and α^{-1}) and not coincidentally the process behaves asymptotically like Brownian motion (where stationary measure is constant).

As for (ii) you can probably guess our answer given the last paragraph: we expect the stationary measure to fluctuate widely in the independent case. It is easy to see that this happens in 1 dimensional finite range systems, but deciding what happens in $d > 1$ seems like a difficult problem.

5. Proofs of Theorems 2 and 3

Having discussed our results at length, the time has come to prove them. The proof of Theorem 2 is quick and simple. Theorem 3 is proved by repeated application of the same ideas.

We begin by observing that we have assumed $V(a\theta)/a^\alpha \Rightarrow W(\theta)$ without really having explained what this means, so we will do that now: $\theta \to V(a\theta)/a^\alpha$ is a random element of $C(R^d, R) =$ continuous functions from R^d to R. This space has a natural topology (uniform convergence on compact sets) and an associated notion of weak convergence ($\mu_n \Rightarrow \mu$ if and only if $\int f d\mu_n \to \int f d\mu$ for all bounded continuous f).

Since the topology above can be generated by a complete metric on the space (exercise) it follows from a result of Skorokhod (1956, see Theorem 3.1.1, p. 281) or for a more recent treatment Dudley (1968) that we can construct $V_1, V_2, \ldots d\ V$ on the same probability space in such a way that with probability 1,

$$V_n(\theta \log^{1/\alpha} n)/\log n \to W(\theta),$$

uniformly on compact sets.

Having traded our distributional convergence in for almost sure uniform convergence we are ready to get to work and do the

Proof of Theorem 2. The key to the proof is the following

(1) *Upper Bound.* Let A be a bounded open set in R^d in which $W(\theta) \leq a$ and $W(\theta) = a$ for $\theta \in \partial A$.

Let $A_n = \{x \in Z^d : x/\log^{1/\alpha} n \in A\}$. If $\varepsilon > 0$ and n is sufficiently large, then for $x \in A_n$,

$$P_x(T_x^+ > T_{A_n^c}) \leq n^{-a + W(x/\log^{1/\alpha} n) + 3\varepsilon}.$$

Proof. We say that V_n is good on A if

(i) $$\sup_{\theta \in A} \left| \frac{V_n(\theta(\log n)^{1/\alpha})}{\log n} - W(\theta) \right| < \varepsilon,$$

and

(ii)
$$|V_n((x+y)/2)-V_n(x)|<\varepsilon\log n$$

when $x\in A_n$ and $|x-y|=1$.

If we let $\phi_n(x)=1$ on A_n, $=0$ on A_n^c, and let $\partial A_n=\{(x,y):x\in A_n,\,y\in A_n^c\}$, then ϕ_n has energy

$$D(\phi_n)=\tfrac{1}{2}\sum_{(x,y)\in\partial A_n}\alpha(x,y)\leq\tfrac{1}{2}n^{-a+\varepsilon}|\partial A_n|,$$

since it follows from (i) and (ii) that $V_n((x+y)/2)>(a-\varepsilon)\log n$ when $x\in A_n,\,y\in A_n^c$. If $A\subset\{x:|x_i|<R$ for $1\leq i\leq d\}$ the last quantity is

$$\leq\tfrac{1}{2}n^{-a+\varepsilon}2d(2R\log^{1/\alpha}n)^d,$$

so we have
(2) If V_n is good and n is sufficiently large,

$$D(\phi_n)\leq n^{-a+2\varepsilon}.$$

To translate this into an estimate on the hitting probability we observe that since $\phi_n(x)=1$ and $\phi_n(y)=0$ for $y\in A_n^c$, it follows from remarks in Sect. 2 that

$$\alpha(x)P_x(T_x^+>T_{A_n^c})\leq D(\phi_n)\leq n^{-a+2\varepsilon},$$

and we have

$$\alpha(x)=\sum_y\alpha(x,y)=\sum_y\exp\left(-V\left(\frac{x+y}{2}\right)\right)\geq 2d\exp(-W(x/\log^{1/\alpha}n)-\varepsilon\log n),$$

so replacing $2d$ by 1 and dividing by $\alpha(x)$ gives the desired result.

Let $G(1+4\varepsilon)=$ the component of $\{\theta:W(\theta)<1+4\varepsilon\}$ containing 0 and imitating the notation in (1), let

$$G_n(a)=\{x\in Z^d:x/\log^{1/\alpha}n\in G(a)\}.$$

By hypothesis this is bounded and we have $W(\theta)=1+4\varepsilon$ on $\partial G(1+4\varepsilon)$, so applying (1) with $x=0$ gives (for n sufficiently large)

$$P_0(T_0^+>T_n^c(1+4\varepsilon))\leq n^{-(1+\varepsilon)},$$

where we have used $T_n^c(1+4\varepsilon)$ to denote the hitting time of $G_n^c(1+4\varepsilon)$.

Letting $n\to\infty$ shows $P_0(T_0^+<\infty)=1$ so X_n is recurrent, and we have proved (a). To prove (b) we observe that when the last inequality holds the number of returns to 0 before we exit A_n is larger than N_n a random variable with $P(N_n>m)=(1-n^{-(1+\varepsilon)})^m$, $m=0,1,2,\ldots$.

From the last formula it is easy to see that $P(N_n>tn^{(1+\varepsilon)})\to e^{-t}$ as $n\to\infty$ so $P(N_n<n)\to 0$, and since each return to 0 requires at least 2 steps it follows that we have

(3)
$$P_0(T_n^c(1+4\varepsilon)\leq 2n)\to 0\quad\text{as}\quad n\to\infty.$$

Since $G_n(1+4\varepsilon)=\{x\in Z^d:x/\log^{1/\alpha}n\in G(1+4\varepsilon)\}$ and $G(1+4\varepsilon)$ is a bounded set we have proved (b).

Proof of Theorem 3. We begin by proving the first part of (c). To do this we begin by repeating the argument above with $G(a^* + 6\varepsilon)$ replacing $G(1 + 4\varepsilon)$ and 0 replaced by a point chosen near a minimum of W in $G(a^* + 6\varepsilon)$.

Since W is continuous it follows that if n is large we can pick a point $x_n^* \in G_n(a^* + 6\varepsilon) = \{x \in Z^d : x/\log^{1/\alpha} n \in G(a^* + 6\varepsilon)\}$ so that $|W(x_n^*/\log^{1/\alpha} n) - m(a^* + \varepsilon))| < \delta$. If V_n is good on $G(a^* + 6\varepsilon)$ (in the sense defined above) then it follows from (1) that we have

$$P_{x_n^*}(T_{x_n^*}^+ > T_n^c(a^* + 6\varepsilon)) \leqq n^{-(a^* + 6\varepsilon) + m(a^* + \varepsilon) + 4\varepsilon} \leqq n^{-(1+\varepsilon)},$$

since $(a^* + \varepsilon) - m(a^* + \varepsilon) > 1$, and repeating the last part of the proof of (b) shows

$$(4) \qquad\qquad P_{x_n^*}(T_n^c(a^* + 6\varepsilon) \leqq 2n) \to 0.$$

To improve the last conclusion to the one in the first part of (c) we need to show that $P_0(T_{x_n^*} < T_n^c(a^* + 6\varepsilon))$ is close to 1. To do this we will prove the following more general result.

(5) Let A be a bounded open set in R^d in which $W(\theta) < a$ and $W(\theta) = a$ for $\theta \in \partial A$. Let K be a connected compact subset of A. There is a constant γ so that if n is sufficiently large then for all $x, y \in K_n = \{z \in Z^d : z/\log^{1/\alpha} n \in K\}$, we have

$$P_x(T_y < T_{A_n^c}) \geqq 1 - n^{-\gamma}.$$

Proof. Now $h_n(y) = P_y(T_x < T_{A_n^c})$ is the function which minimizes the energy in the class of functions which are $= 1$ at x and $= 0$ on A_n^c, and the function ϕ_n introduced in the proof of (1) has these properties so it follows from (2) that

$$n^{-a + 2\varepsilon} \geqq D(\phi_n) \geqq D(h_n).$$

To get a lower bound on $D(h_n)$ we let z_k, $0 < k < m$, be a self-avoiding path from $z_0 = x$ to $z_m = y$ (i.e., $|z_{k-1} - z_k| = 1$, $1 \leqq k \leqq m$ and $z_i \neq z_j$ if $i \neq j$), and observe

$$D(h_n) \geqq \sum_{k=1}^{m} \alpha(z_{k-1}, z_k)(h_n(z_k) - h_n(z_{k-1}))^2$$

$$\geqq \left(\inf_k \alpha(z_{k-1}, z_k)\right) \sum_{k=1}^{m} (h_n(z_k) - h_n(z_{k-1}))^2$$

$$\geqq \left(\inf_k \alpha(z_{k-1}, z_k)\right) \frac{1}{m} (h_n(x) - h_n(y))^2,$$

by the Cauchy-Schwarz inequality, so recalling the definition of h_n gives

$$(1 - P_y(T_x < T_{A_n^c}))^2 \leqq n^{-a + 2\varepsilon} m \Big/ \inf_k \alpha(z_{k-1}, z_k).$$

If $A \subset \{x : |x_i| < R \text{ for } 1 \leqq i \leqq d\}$ the last quantity is

$$\leqq n^{-a + 3\varepsilon} \Big/ \left(\inf_k \alpha(z_{k-1}, z_k)\right),$$

for n sufficiently large.

Let γ be chosen so that $a - 3\gamma > \sup\{W(z) : z \in K\}$, and set $\varepsilon = \gamma/10$. It is easy to see that if n is sufficiently large, then for any $x, y \in K_n$ we can pick the path so that

$$\inf_k \alpha(z_{k-1}, z_k) \geqq n^{-a + 2\gamma + 5\varepsilon}.$$

Plugging this result into the inequality above and taking square roots proves the desired result.

To prove the second part of (c) we need to get an upper bound on $N_x^n(x) = $ the number of visits to x before X_m exists $G_n(a) = \{x \in Z^d : x/\log^{1/\alpha} n \in G(a)\}$. Now it is easy to see that

(6)
$$E_0 N_x^n(a) = P_0(T_x < T_n^c(a)) E_x N_x^n(a) \leqq E_x N_x^n(a)$$
$$= 1/P_x(T_x^+ > T_n^c(a)),$$

so we will get our upper bound on the number of visits by proving the following:

(7) *Lower Bound.* Let A be a bounded open set in R^d in which $W(\phi) < a$ and $W(\phi) = a$ for $\phi \in \partial A$. If $\varepsilon > 0$ and n is sufficiently large, then for all $x \in A_n = \{x \in Z^d : x/\log^{1/\alpha} n \in A\}$. We have

$$P_x(T_x^+ > T_{A_n^c}) \geqq n^{-a + W(x/\log^{1/\alpha} n) - 4\varepsilon}.$$

Proof. If V_n is good then $V_n((x+y)/2) < (a+\varepsilon) \log n$ for all $x \in A_n$, $y \in Z^d$, $|x - y| = 1$, so

$$D(h) \geqq n^{-(a+\varepsilon)} \sum_{(x,y) \in \partial A_n} (h(x) - h(y))^2.$$

To get a lower bound on the right-hand side we use an idea from the proof of (5). If z_k, $0 < k < m$ is a selfavoiding path from x to some point $y \in \partial D_n$, then

$$D(h) \geqq n^{-(a+\varepsilon)} \sum_{k=1}^{m} (h(z_k) - h(z_{k-1}))^2,$$

and the sum minimized by having all the increments equal so

$$D(h) \geqq n^{-(a+\varepsilon)} \frac{(h(x) - h(y))^2}{m} = n^{-(a+\varepsilon)}/m,$$

if $h(x) = 1$ and $h(y) = 0$. Now if $A \subset \{x : |x_i| < R$ for $1 \leqq i \leqq d\}$, then $m \leqq (2R \log^{1/\alpha} n)^d$, so we have

(8)
$$D(h) \geqq n^{-(a+\varepsilon)}(2R \log^{1/\alpha} n) \geqq n^{-(a+2\varepsilon)},$$

if n is sufficiently large. This is the lower bound analogous to

$$D(\phi_n) \leqq n^{-a+2\varepsilon}. \tag{2}$$

To translate it into an estimate on the hitting probabilities we observe that the minimum energy among functions $= 1$ at x and $= 0$ on A_n^c is $\alpha_n(x) P_x(T_x^+ > T_{A_n^c})$, which is $> n^{-(a+2\varepsilon)}$ by (8), so if V_n is good and n is sufficiently large then

$$P_x(T_x^+ > T_{A_n^c}) \geqq n^{-(a+2\varepsilon)}/\alpha_n(x) \geqq n^{-(a+2\varepsilon)}/2d \exp(W(x/\log^{1/\alpha} n) + \varepsilon \log n)$$
$$\geqq n^{-(a+4\varepsilon) + W(x/\log^{1/\alpha} n)},$$

proving (7).

To prove the second part of (c) now we apply (7) with $A = G(a^* - 5\varepsilon)$ and find that if n is sufficiently large then

$$P_x(T_x^+ > T_n^c(a^* - 5\varepsilon)) \geqq n^{\varepsilon - 1},$$

for $x \in G_n(a^* - 5\varepsilon)$, so it follows from (6) that

$$E_0 N_x^n(a^* - 5\varepsilon) \leq n^{1-\varepsilon},$$

and summing over $x \in G_n(a^* - 5\varepsilon)$ gives

$$E_0 T_n^c(a^* - 5\varepsilon) \leq n^{1-\varepsilon} |G_n(a^* - 5\varepsilon)|.$$

Now if $G(a^* - 5\varepsilon) \subset \{x : |x_i| < R \text{ for } 1 = 1, \dots d\}$, we have

$$E_0 T_n^c(a^* - 5\varepsilon) \leq n^{1-\varepsilon} (2R \log^{1/\alpha} n)^d,$$

and the desired result follows from Chebyshev's inequality

$$P_0(T_n^c(a^* - 5\varepsilon) \geq n) \leq n^{-\varepsilon}(2R \log^{1/\alpha} n)^d.$$

To prove (d) we observe that the upper and lower bounds show that

$$P_x(T_0^+ > T_{A_n^c}) = n^{-a \pm 4\varepsilon}.$$

Applying the observation above to $A = G(a^* + \varepsilon)$ we see that [to use the notation defined for (6)]

$$E_0 N_0^n(a^* + \varepsilon) \leq n^{a^* + 5\varepsilon},$$

so it follows from Chebyshev's inequality that

$$P_0(N_0^n(a^* + \varepsilon) \geq n^{a^* + 6\varepsilon}) \to 0,$$

or rearranging the inside inequality

(9) $$P_0(\log N_0^n(a^* + \varepsilon) \geq (a^* + 6\varepsilon) \log n) \to 0.$$

To get a lower bound on the number of visits to 0 we let $A = G(a^* - \varepsilon)$ and observe that the number of returns to 0 has a geometric distribution with success probability $1 - p_n$, where $p_n = P_0(T_0^+ > T_n^c(a^* - \varepsilon))$ that is

$$P_0(N_0^n(a^* - \varepsilon) \geq k) = (1 - p_n)^k, \qquad k = 0, 1, 2, \dots$$

(we are not counting the visit at time 0).

Now (7) implies $p_n \leq n^{-a^* + 5\varepsilon}$, so we have

$$P_0(N_0^n(a^* - \varepsilon) \geq n^{a^* - 6\varepsilon}) \to 1,$$

as $n \to \infty$, and rearranging the inside inequality gives

(10) $$P_0(\log N_0^n(a^* - \varepsilon) \geq (a^* - 6\varepsilon)(\log n)) \to 1.$$

To prove (e) we observe that the upper and lower bounds show that

$$P_x(T_x^+ > T_{A_n^c}) = n^{-a + W(x/\log^{1/\alpha} n) \pm 4\varepsilon},$$

so repeating the proof of (d) shows that

(9′) $$P_x(\log N_x^n(a + \varepsilon) \geq (a + 6\varepsilon) \log n) \to 0,$$

(10′) $$P_x(\log N_x^n(a - \varepsilon) \geq (a - 6\varepsilon) \log n) \to 1.$$

With (9′) and (10′) in hand all we have to do is show that if ε is small, then starting from 0 we hit x with high probability before we exit $G_n(a^* - \varepsilon)$, but this is a consequence of (5) so the proof is complete.

References

Derrida, B., Luck, J.M.: Diffusion on a random lattice: weak disorder expansion in arbitrary dimension. Phys. Rev. B 28, 7183–7190 (1983)

Durrett, R.: Oriented percolation in two dimensions. Ann. Probab. **12**, 999–1040 (1984a)

Durrett, R.: Brownian motion and martingales in analysis. Monterey, CA: Wadsworth 1984b

Durrett, R.: Reversible diffusion processes. In: Probability and harmonic analysis. Chao, J., Woyczynski, W. (eds.). New York: Dekker 1985a

Durrett, R.: Particle systems, random media, and large deviations. Conference proceedings. AMS Contemporary Math. Series, Vol. 41 (1985b)

Fisher, D.: Random walks in random environments. Bell Laboratories, preprint (1984)

Gangolli, R.: Abstract harmonic analysis and Lévy's Brownian motion of several parameters, 5th Berkeley Symp., Vol. II, pp. 13–30 (1965)

Griffeath, D., Liggett, T.: Critical phenomena for Spitzer's reversible nearest particle systems. Ann. Probab. **10**, 881–895 (1982)

Kesten, H., Kozlov, M., Spitzer, F.: A limit law for random walk in a random environment. Compos. Math. **30**, 145–168 (1975)

Kotani, S.: A lemma for stationary random sequences and its application to the recurrence of random walks in random media. Preprint (1985)

Lévy, P.: Processus stochastiques et mouvement Brownien. Paris: Gauthier-Villars (1948)

Liggett, T.: Interacting particle systems. Berlin, Heidelberg, New York: Springer 1985

Luck, J.M.: Diffusion in a random medium: a renormalization group approach. Nucl. Phys. B **225**, 169–184 (1983)

Luck, J.M.: A numerical study of diffusion and conduction in a 2D random medium. J. Phys. A **17**, 2069–2077 (1984)

Marinari, E., Parisi, G., Ruelle, D., Windey, P.: Random walks in a random environment and $1/f$ noise. Phys. Rev. Lett. **50**, 1223–1225 (1983a)

Marinari, E., Parisi, G., Ruelle, D., Windey, P.: On the interpretation of $1/f$ noise. Commun. Math. Phys. **89**, 1–12 (1983b)

Pitt, L.D.: Local times for Gaussian vector fields. Indiana Math. J. **27**, 309–330 (1978)

Ritter, G.A.: Random walk in a random environment, critical case. Ph. D. Thesis, Cornell (1976)

Schumacher, S. (1984): Diffusions with random coefficients. Ph. D. Thesis, UCLA. A summary appears in Durrett (1985b)

Sinai, Ya.: Limit behavior of one-dimensional random walks in random environments. Theor. Probab. Appl. **27**, 247–258 (1982)

Solomon, F.: Random walks in a random environment. Ann. Probab. **3**, 1–31 (1975)

Skorokhod, A.V.: Limit theorems for stochastic processes. Theor. Probab. Appl. **1**, 261–290 (1956)