The Annals of Probability
1989, Vol. 17, No. 2, 444–481

STATISTICAL MECHANICS OF CRABGRASS

BY M. BRAMSON,[1] R. DURRETT[1,2] AND G. SWINDLE[2]

University of Minnesota, Cornell University and University of California, Los Angeles

In this article we consider the asymptotic behavior of the contact process when the range M goes to ∞. We show that if λ is the total birth rate from an isolated particle, then the critical value $\lambda_c(M) \to 1$ as $M \to \infty$. The rate of convergence depends upon the dimension: $\lambda_c(M) - 1 \approx M^{-2/3}$ in $d = 1$, $\approx (\log M)/M^2$ in $d = 2$, and $\approx M^{-d}$ in $d \geq 3$.

1. Introduction. In this article we consider a "contact process" in which the state at time t is $\xi_t \subset \mathbb{Z}^d/M = \{z/M: z \in \mathbb{Z}^d\}$. ξ_t is the set of sites occupied by "particles." In our model (i) particles die at rate 1 and (ii) if x is occupied and y with $\|x - y\| \leq 1$ is vacant, births occur from x to y at rate $\lambda/v(M)$, where $v(M) = |\{x \in \mathbb{Z}^d/M: 0 < \|x\| \leq 1\}|$ is a constant chosen to make the total birth rate from an isolated particle equal to λ. v is for volume. Here and in what follows, $\|z\| = \sup|z_i|$, but other norms or birth rates of the form $\varphi(x - y)$ could be introduced without much difficulty. This type of model is appropriate for a lawn or meadow where the spacing between plants is small, hence the title of the article. More generally, the model can be used in any situation where offspring can be displaced a large distance (measured on the lattice) from their parents.

The point of this article is to show that if M is large, ξ_t behaves much like a branching random walk Z_t in which (i) particles die at rate 1, (ii) give birth at rate λ and (iii) an offspring of a particle at x goes to a y chosen at random from $\{y \in \mathbb{Z}^d/M: 0 < \|x - y\| \leq 1\}$. Comparison with the definition of ξ_t reveals that the contact process can be thought of as a coalescing branching random walk, that is, a system which follows rules (i)–(iii), and (iv) if a particle is sent to an occupied site, the two coalesce to one.

From the above comparison, it follows that if $Z_t^0(A)$ is the number of particles in A at time t in the branching random walk starting from a single particle at 0 at time 0, and $\xi_t^0(A) \equiv |\xi_t^0 \cap A|$ is the same thing for the contact process starting from $\xi_0^0 = \{0\}$, then the two processes can be constructed on the same probability space with $Z_t^0(A) \geq \xi_t^0(A)$. Here $|B|$ = the number of points in B. The reader should notice that Z_t^0 is a measure-valued process and ξ_t^0 is set-valued. To facilitate comparison, we will use $|Z_t^0|$ to denote $Z_t^0(\mathbb{R}^d)$.

From the last observation it follows that if we let

$$\Omega_\infty = \{\xi_t^0 \neq \varnothing \text{ for all } t\},$$
$$\lambda_c(M) = \inf\{\lambda: P(\Omega_\infty) > 0\},$$

Received January 1988.

[1]Partially supported by a grant from the National Science Foundation.

[2]Partially supported by the Army Research Office through the Mathematical Sciences Institute at Cornell University.

AMS 1980 *subject classifications*. Primary 60K35; secondary 60J80.

Key words and phrases. Contact process, renormalized bond construction, branching processes.

then $\lambda_c(M) \geq 1$ = the critical value for the branching random walk. It is not hard to show that if we keep T fixed and let $M \to \infty$, then the contact process ξ_t^0, $0 \leq t \leq T$, converges in distribution to a branching random walk. The next result shows that we can let $T \to \infty$ and then $M \to \infty$.

Theorem 1. *As $M \to \infty$, $\lambda_c(M) \to 1$. Furthermore*

$$\lambda_c(M) - 1 \approx \begin{cases} C/M^{2/3} & d = 1, \\ C(\log M)/M^2 & d = 2, \\ C/M^d & d \geq 3, \end{cases}$$

where $\approx$ means that if C is small (large) then the right-hand side is a lower (upper) bound for large M.

The conclusion that $\lambda_c(M) \to 1$ is analogous to results of Holley and Liggett (1981) and Griffeath (1983) who consider what happens in the basic contact process (in our notation, $M = 1$ and $\|x\| = |x_1| + \cdots + |x_d|$). They show that the critical value for the model in d dimensions converges to 1 as $d \to \infty$. The main point of Theorem 1 is to identify the rate of convergence. The content of Theorem 1 becomes a little more transparent if we write the conclusion in terms of volume $v(M)$:

$$\lambda_c(v) - 1 \approx \begin{cases} C/v^{2/3} & d = 1, \\ C(\log v)/v & d = 2, \\ C/v & d \geq 3. \end{cases}$$

Recast in these terms, the result suggests that there should be a power series for $\lambda_c(v)$ in $d \geq 3$, and resembles a (nonrigorous) result of Thouless (1969) for the Ising model: If q is the number of "interacting particles," then the shift of T_c from the mean field value is proportional to $q^{-1} \ln q$ in $d = 2$ and q^{-1} in $d \geq 3$. (The Ising model does not have a phase transition in $d = 1$.)

One final reason for writing things in terms of the volume can be seen in our bounds for $\lambda_c(M) - 1$ in $d \geq 3$. If U_n denotes the random walk which has steps uniformly distributed on $\{x \in \mathbb{R}^d\colon \|x\| \leq 1\}$, then Proposition 3.1 and 2.6 imply

$$\frac{1}{9v(M)} \leq \lambda_c(M) - 1 \leq \frac{2.01}{v(M)} \sum_{n=1}^{\infty} P(\|U_n\| \leq 1).$$

We think that in $d \geq 3$,

$$\lambda_c(M) - 1 \sim \frac{0.5}{v(M)} \sum_{n=1}^{\infty} P(\|U_n\| \leq 1),$$

where $a_M \sim b_M$ means $a_M/b_M \to 1$ as $M \to \infty$. Getting sharp asymptotics like this seems to be a difficult problem. Our bounds are ridiculously crude in $d = 1$ and $d = 2$.

Theorem 1 gives asymptotics for the critical value. Our last two results concern the survival probability $P(\Omega_\infty)$, where $\Omega_\infty = \{\xi_t^0 \neq \varnothing \text{ for all } t\}$. The first result is the analog of a result of Schonmann and Vares (1986) for the basic contact process with $d \to \infty$.

THEOREM 2. *If* $\lambda > 1$ *is fixed and* $M \to \infty$, *then* $P(\Omega_\infty) \to (\lambda - 1)/\lambda$.

The limit is the survival probability for the branching random walk, so Theorem 2 gives another sense in which the ξ_t^0 is like Z_t^0 when M is large. The next result refines the last conclusion.

THEOREM 3. *Let*

$$\lambda - 1 = \begin{cases} C/M^{2/3} & d = 1, \\ C(\log M)/M^2 & d = 2, \\ C/M^d & d \geq 3. \end{cases}$$

There is a constant $\delta(C)$, *which depends on the dimension and approaches* 0 *as* $C \to \infty$, *so that if* M *is large*

$$P(\Omega_\infty) \geq (1 - \delta(C))(\lambda - 1)/\lambda.$$

In the language of statistical physics, the last result identifies the scale on which the *crossover* occurs from contact process to branching process behavior. When C is large, the contact process survival probability is almost that of the branching process, but when C is small, the lower bounds on λ_c imply that the survival probability for the contact process is 0 for large M.

Having stated our results, the rest of the Introduction is devoted to an explanation of how they are proved. We begin with a "back-of-the-envelope" calculation which indicates why the answers in Theorem 1 are correct. Let $\lambda = 1 + \varepsilon$ and suppose

$$v(M) \sim \begin{cases} C\varepsilon^{-3/2} & d = 1, \\ C\varepsilon^{-1}\log(\varepsilon^{-1}) & d = 2, \\ C\varepsilon^{-1} & d \geq 3. \end{cases} \tag{1}$$

To prove Theorem 1, we need to show that the contact process dies out if C is small and survives if C is large. The key to this lies in two differential equations. If we use $x \sim y$ to denote $x, y \in \mathbb{Z}^d/M$ with $0 < \|x - y\| \leq 1$, then

$$\frac{d}{dt}E|\xi_t^0| = (\lambda - 1)E|\xi_t^0| - \frac{\lambda}{v(M)} \sum_{x,\, y \sim x} P(x, y \in \xi_t^0), \tag{2}$$

$$\frac{d}{dt}E|Z_t^0| = (\lambda - 1)E|Z_t^0|. \tag{3}$$

The solution to the second equation is trivial: $E|Z_t^0| = e^{(\lambda - 1)t}$. The right-hand side is approximately equal to 1 if $t \ll 1/\varepsilon$ and large if $t \gg 1/\varepsilon$. This suggests that $t = \theta/\varepsilon$ is the right time scale to look at, and on this scale Z_t^0 is much like a critical branching random walk. The second conclusion is a large leap, but will be substantiated by results given in Section 2. In percolation terminology, $1/\varepsilon$ is the *correlation length.*

The term which appears in (2) but not in (3) will be called the *interference term.* To estimate this term, we let

$$B_n^0 = [n_1, n_1 + 1) \times \cdots \times [n_d, n_d + 1)$$

and

$$B_n^1 = [n_1 - 1, n_1 + 1) \times \cdots \times [n_d - 1, n_d + 1)$$

for each $n = (n_1, \dots, n_d) \in \mathbb{Z}^d$. Now if $x, y \in B_n^0$ and $x \neq y$, then $x \sim y$, and $x \sim y$ implies $x, y \in B_n^1$, where $n_i = 1 + \min([x_i], [y_i])$ and $[x]$ is the greatest integer less than or equal to x. So

$$(4) \qquad \sum_n E\xi_t^0(B_n^0)^2 \le \sum_{x,\, y \sim x} P(x, y \in \xi_t^0) \le \sum_n E\xi_t^0(B_n^1)^2.$$

Careful readers will have noticed that the left-hand side should be $\Sigma E\xi_t^0(B_n^0)(\xi_t^0(B_n^0) - 1)$, but we will ignore this technicality. It is a minor mistake when compared to the next step.

If we replace ξ_t^0 by Z_t^0 and set $\lambda = 1$, we have a critical branching random walk. Well-known results for that process [see, e.g., Fleischman (1978)] imply

$$(5) \qquad \sum_n EZ_t^0(B_n^i)^2 \approx \begin{cases} t^{1/2} & d = 1, \\ \log t & d = 2, \\ 1 & d \ge 3. \end{cases}$$

Replacing ξ by Z in (4) and plugging the last asymptotics into (2), we see that for $t = \theta/\varepsilon$,

$$\text{the interference term} \approx \begin{cases} \varepsilon^{-1/2} & d = 1, \\ \log(1/\varepsilon) & d = 2, \\ 1 & d \ge 3. \end{cases}$$

The choice of $v(M)$ in (1) makes the interference term of order ε, that is, the same size as the first term.

The last few paragraphs explain the intuition behind the answer. To translate these heuristics into rigorous lower bounds on $\lambda_c(M)$, we observe that by comparing with a branching random walk in which a particle at x gives birth to a set of particles with distribution $x + \xi_t^0$, we get

$$(6) \qquad E|\xi_{kt}^0| \le \left(E|\xi_t^0|\right)^k.$$

So if $E|\xi_t^0| < 1$, then $P(\Omega_\infty) = 0$.

The techniques we use to show that $E|\xi_t^0|$ will fall below 1 when the C in (1) is small depending upon the dimension. In $d \ge 3$, a simple argument shows that the interference term is at least

$$\frac{1}{8v(M)} E|\xi_{t-\ln 2}^0|,$$

so $E|\xi_t^0| \to 0$ when $\lambda \le 1 + 1/9v(M)$ and M is large. In $d = 1$, we show that if C is small,

$$\frac{d}{dt} E|\xi_t^0| \le -\varepsilon \quad \text{when } t \in [1/10\varepsilon, 1/\varepsilon] \text{ and } E|\xi_t^0| \ge 0.9.$$

Since

$$E|\xi_{0.1/\varepsilon}^0| \le E|Z_{0.1/\varepsilon}^0| = \exp(0.1) \le 1.2,$$

it follows that $E|\xi_t^0|$ will be less than or equal to 0.9 at some time $t \le 1/\varepsilon$, and (6) implies the contact process dies out. The argument in $d = 2$ is similar to the one in $d = 1$, but we look at $t \in [a/2\varepsilon, a/\varepsilon]$ and suppose $E|\xi_t^0| \ge 1 - a$ where a is small.

To prove upper bounds on $\lambda_c(M)$, we use a block construction. Let $L = (K_2/\varepsilon)$, $l = L^{1/2}$ and $I_m = [(2m-1)l, (2m+1)l] \times [-l, l]^{d-1}$. We first show that if we start a branching random walk with at least K_1/ε particles in I_0 and do not allow births outside $[-2l, 2l] \times \mathbb{R}^{d-1}$, then with high probability there are at least $3K_1/\varepsilon$ particles in I_{-1} and in I_1 at time L. Let $\mathscr{L} = \{(m, n): m + n$ is even$\}$, and say that $(m, n) \in \mathscr{L}$ is occupied if there are at least K_1/ε particles in I_m in the contact process at time nL. To prove that the contact process survives when C in (1) is large, we show that the process of occupied sites dominates oriented site percolation with parameter p close to 1.

One technical problem with the approach above is that if all K_1/ε particles lie in $[0,1]^d$, then the contact process will die out with high probability. To avoid this difficulty, we start with an initial $A \subset I_0$ which has at most $K_3 h(\varepsilon)$ particles in any unit cube B_n^0, where

$$h(\varepsilon) = \begin{cases} \varepsilon^{-1/2} & d = 1, \\ \log \varepsilon^{-1} & d = 2, \\ 1 & d \ge 3. \end{cases}$$

Experienced readers will recognize these numbers as giving the size of a typical clump of particles in a critical branching random walk at time $1/\varepsilon$. We prove that if K_3 is large and if we kill all the particles in the branching process at time L which reside in crowded B_n^0, then with high probability we will still have at least $2K_1/\varepsilon$ particles in I_{-1} and I_1.

The last two paragraphs show that the branching process can do the things that we want the contact process to do. To complete the proof, we will show that if C in (1) is large, then the contact process is almost as large as the branching process for $t \le L$. The first step is to notice that

$$\sum_{x,\, y \sim x} P(x, y \in \xi_t^A) \le \sum_n E\left(Z_t^A(B_n^1)^2\right).$$

If we let $\delta_t^A = |Z_t^A| - |\xi_t^A|$ and subtract (2) from (3), then we get

$$\frac{d}{dt} E\delta_t^A \le \frac{\lambda}{v(M)} \sum_n E Z_t^A(B_n^1)^2.$$

Using the fact that A has at most $K_3 h(\varepsilon)$ particles in any B_n^0 and estimates for branching random walk, we show that if C in (1) is large, then $E|\xi_t^A| \geq (1-\delta)E|Z_t^A|$ for $t \leq L = K_2/\varepsilon$. Chebyshev's inequality then implies that $|Z_L^A| - |\xi_L^A| \leq K_1/\varepsilon$ with high probability. Since the branching random walk will have at least $2K_1/\varepsilon$ particles in I_1 and I_{-1} with high probability, this completes the proof of the upper bounds (modulo a few details).

The article is organized as follows: Preliminary results for branching processes and random walks are stated in Section 2. Section 3 contains proofs of the lower bounds in $d \neq 2$ and Section 4 treats the stubborn case $d = 2$. Readers who get tired of all the details in Section 4 will be happy to know that the rest of the article is independent of that section. The block construction which allows us to prove that the contact process survives is described in Section 5. In Section 6 we prove some estimates which imply that the block construction works for the branching random walk. In Section 7 we use the differential equations (2)–(4) to estimate the difference between the contact process and the branching process completing the proof of Theorem 1. In Section 8 we prove Theorem 2. In Section 9 we prove the sharper result in Theorem 3.

Formulas are numbered (1), (2), ... in each section. When formula (6) from Section 2 is referred to in a later section it is called (2.6). Most of our results are called lemmas, but those whose conclusions are part of Theorem 1 are called propositions. Two of the lemmas are interesting enough so that they could be called theorems. Lemma 4.6 shows that if S_n is the sum of n independent random variables with distribution F_n and F_n converges weakly to F, then (under suitable conditions) the local central limit theorem holds for S_n. Lemma 9.2 is an almost central limit theorem for almost critical branching random walks. Finally, the reader might find Lemma 4.5 useful.

2. Branching process preliminaries. In this section we will describe some results about branching processes and random walks which will be useful in what follows. The first result says that a branching process with mean $1 + \delta$ looks roughly like a critical branching process for times $t \leq K/\delta$. In reading the hypotheses the reader should keep in mind that we will apply this result to the discrete-time skeletons $Z_0, Z_1, Z_2, \ldots$ of continuous-time branching processes in which particles die at rate 1, and give birth to new ones at rate $1 + \varepsilon$ with $0 \leq \varepsilon \leq 1$. In this case, $1 + \delta = e^{\varepsilon}$.

LEMMA 2.1. *For any $\alpha > 0$, let $\mathscr{K}_\alpha$ be a class of Galton–Watson processes with reproduction variances not less than α, and such that for any $\eta > 0$ there is a $k(\eta)$ so that*

$$\sum_{k > k(\eta)} k^2 p_k < \eta$$

for all laws p in $\mathscr{K}_\alpha$. If p is a law in $\mathscr{K}_\alpha$ let $f(z) = \sum p_k z^k$, $\mu = f'(1)$, $\nu = f''(1)/2$ and

$$c_n = \begin{cases} \nu \dfrac{(1-\mu^{-n})}{(\mu - 1)} & \mu \neq 1, \\ \nu n & \mu = 1. \end{cases}$$

For any sequence of laws $p^n \in \mathscr{K}_\alpha$ with $\mu \to 1$ as $n \to \infty$,

(a) $$c_n P(Z_n > 0) \to 1,$$

(b) $$E(Z_n/\mu^n c_n | Z_n > 0) \to 1,$$

(c) $$P(Z_n/\mu^n c_n \le x | Z_n > 0) \to 1 - e^{-x}.$$

(Here μ, ν and c_n are the constants associated with p^n.)

REMARK. The lemma above is from Jagers (1975), page 63. There the result is stated incorrectly. The μ^n in (b) and (c) is missing. It is easy to see that if we accept (a), then (b) and (c) must have the form given above since

$$E(Z_n|Z_n > 0) = EZ_n/P(Z_n > 0) = \mu^n/P(Z_n > 0) \sim \mu^n c_n.$$

While the last result is useful for conceptualization, it plays a minor role in the proofs below, appearing only in Sections 3 and 9. Most of the work is carried out using moment equations. Let $Z_t^x(A)$ be the number of particles in A at time t in a branching random walk which starts with a single particle at x and in which (i) particles die at rate 1, and (ii) a particle at x gives birth to a new particle at $x + z$ at rate $\beta(dz)$, that is, β is a measure on $\mathbb{R}^d$ and the rate at which particles appear in $x + A = \{x + y\colon\ y \in A\}$ is $\beta(A)$.

Durrett (1979) gives the following integral equation for the moments.

LEMMA 2.2. *Let $m_k(t, x, A) = E(Z_t^x(A)^k)$ and omit the subscript when $k = 1$. For $k \ge 2$,*

$$\begin{aligned} m_k(t, x, A) &= m(t, x, A) \\ &\quad + \int \beta(dz) \sum_{j=1}^{k-1} \frac{1}{2} \binom{k}{j} \int_0^t ds \int m(t-s, x, dy) \\ &\quad \times m_j(s, y, A) m_{k-j}(s, y+z, A). \end{aligned}$$

Lemma 2.2 relates the kth moment to moments of order less than k, so all the moments can be calculated once we observe

(1) $$m(t, x, A) = e^{(\lambda - 1)t} P(X_t^x \in A).$$

Here $\lambda = \beta(\mathbb{R}^d)$ and X_t^x is a continuous-time random walk which starts at x, takes steps at rate λ and has jumps with distribution $F(dx) = \beta(dz)/\lambda$. To check (1), observe that

(2) $$\frac{\partial}{\partial t} m(t, x, A) = -m(t, x, A) + \int m(t, x, dy) \beta(A - y),$$

and that the right-hand side of (1) satisfies (2). In what follows we will use β_M and F_M to denote the quantities above when the range is M.

Lemma 2.2 has been used to analyze the spatial distribution of critical branching random walks. [See Sawyer (1976), Fleischman (1978) and Durrett (1979).] In view of Lemma 2.1, the reader should not be surprised to hear that for $\lambda = 1 + \varepsilon$ the spatial distribution is roughly the same as the critical case for

$t \le K/\varepsilon$. To translate this intuition into proofs, we will need estimates on $m(t, x, A)$ which are uniform in the random walk law. To do this, a concentration function inequality of Kesten (1969) will be useful. To state his result, we need some notation. If Y is a random variable let

$$Q(Y; L) = \sup_x P(x \le Y \le x + L)$$

be its concentration function.

LEMMA 2.3. *There is a constant C so that if S_n is a sum of n independent and identically distributed random variables, then*

$$Q(S_n; L) \le \frac{C}{n^{1/2}} \frac{L}{l} \frac{Q(S_1; L)}{[1 - Q(S_1; l)]^{1/2}}$$

for any l with $0 < l \le 2L$.

For the results that follow, we will need to generalize Kesten's result to continuous time and to $d > 1$. To do this, let X_t be a continuous-time random walk which jumps at rate λ and has jumps with distribution F_M, the uniform distribution on $\{x \in \mathbb{Z}^d/M\colon 0 < \|x\|_\infty \le 1\}$. X_t can be thought of as a random walk which jumps at rate $\lambda(1 + 1/V(M))$ and has jumps with distribution $\bar{F}_M$ = the uniform distribution on $\{x \in \mathbb{Z}^d/M\colon \|x\|_\infty \le 1\}$. If we do this and let $\bar{S}_n$ denote the position after the nth jump, then the d coordinates of $\bar{S}_n$ are independent. (Note that the last representation shows us that the coordinates of X_t are independent.)

Generalizing the concentration function by setting

$$Q(Y; L) = \sup_x P\big(x_i \le Y^i \le x_i + L \text{ for } 1 \le i \le d\big)$$

where $Y = (Y^1, \dots, Y^d)$, we have

$$Q(\bar{S}_n; L) \le \prod_{i=1}^{d} Q(\bar{S}_n^i; L), \tag{3}$$

since the components $\bar{S}_n^1, \dots, \bar{S}_n^d$ are independent. Now $\bar{S}_1^i$ is uniformly distributed on $\{-1, -1 + 1/M, \dots, 1\}$ so if we pick $L = 2$ and $l = 1$, then $Q(\bar{S}_1^i; 2) = 1$ and $Q(\bar{S}_1^i; 1) \le 2/3$ for $M \ge 1$. Using Lemma 2.3 and (3) now, it follows that

$$Q(\bar{S}_n; 2) \le C/n^{d/2}. \tag{4}$$

Here and in what follows C is the constant whose value is not important and which will change from line to line.

Our next step is to generalize (4) to continuous time.

LEMMA 2.4. *There is a constant C so that if $1 \le \lambda \le 2$ and $M \ge 1$, then*

$$Q(X_t; 2) \le C/t^{d/2}.$$

PROOF. If we let $\gamma = \lambda(1 + 1/v(M)) \geq 1$, then

$$Q(X_t; 2) \leq \sum_{n=0}^{\infty} e^{-\gamma t} \frac{(\gamma t)^n}{n!} Q(\bar{S}_n; 2).$$

A standard large deviations estimate implies

$$\sum_{n=0}^{[t/2]} e^{-\gamma t} \frac{(\gamma t)^n}{n!} \leq K e^{-\delta t},$$

where K, δ are independent of t and $\gamma \geq 1$. Using $Q(\bar{S}_n; 2) \leq 1$ for $n \leq [t/2]$, (4) for $n > t/2$ and summing gives the desired result. □

Our first application of Kesten's inequality is to give upper bounds on λ_c in $d \geq 3$. Griffeath (1983) has shown that if S_n is the random walk which takes steps with a uniform distribution on $\{x \in \mathbb{Z}^d/M\colon 0 < \|x\| \leq 1\}$, then we have

LEMMA 2.5. *Let $\gamma = P(S_n \neq 0$ for all $n \geq 1 | S_0 = 0)$. If $\gamma > 1/2$, then $\lambda_c \leq 1/(2\gamma - 1)$.*

To get an upper bound on λ_c, it suffices to bound γ. Noticing S_n has the same distribution as $\bar{S}(T_n)$ where $T_n = \inf\{m > T_{n-1}\colon \bar{S}_m \neq \bar{S}(T_{n-1})\}$ for $n \geq 1$ and $T_0 = 0$ gives

$$\begin{aligned} & P(S_n = 0 \text{ for some } n \geq 1) \\ & \leq \sum_{n=2}^{\infty} P(S_n = 0) \quad [\text{since } P(S_1 = 0) = 0] \\ & \leq \sum_{n=2}^{\infty} P(\bar{S}_n = 0) \\ & \leq \sum_{n=2}^{\infty} P(\|\bar{S}_{n-1}\| \leq 1) \frac{1}{v(M) + 1} \\ & \leq \sum_{n=2}^{\infty} \frac{C}{n^{d/2}} \frac{1}{v(M) + 1} \leq \frac{C'}{v(M)} \end{aligned} \tag{5}$$

by (4). The constant C' is presumably large, so the last result is not very good. It can be improved considerably by noting that if U_n is the random walk which has steps uniformly distributed over $\{x \in \mathbb{R}^d\colon \|x\| \leq 1\}$, then for each n,

$$P(\|S_n\| \leq 1) \to P(\|U_n\| \leq 1)$$

as $M \to \infty$. Dominated convergence and (5) allow us to conclude that the upper bound is asymptotically

$$\frac{1}{v(M)} \sum_{n=1}^{\infty} P(\|U_n\| \leq 1).$$

Plugging this bound for $1 - \gamma$ into Lemma 2.5 gives

PROPOSITION 2.6. *For M large*

$$\lambda_c \leq 1 + \frac{2.01}{v(M)} \sum_{n=1}^{\infty} P(\|U_n\| \leq 1).$$

Note: The idea of using Griffeath's result to bound λ_c is due to Ted Cox.

While the definition of S_n is still in the reader's mind, we would like to introduce one bit of notation which will come up again and again in what follows. Let σ_M^2 be the variance of one component of S_n. As $M \to \infty$,

$$(6) \qquad \sigma_M^2 \to 1/3,$$

1/3 being the variance of the uniform distribution on $[-1, 1]$.

Last and least, the following fact will be useful several times below to simplify computations.

LEMMA 2.7. *If* $0 \leq x \leq 1$, *then* $e^x \leq 1 + 2x$.

PROOF. The proof is trivial and is included as comic relief.

$$\frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots \leq \frac{x}{2}\left(1 + \frac{1}{3} + \frac{1}{3 \cdot 4} + \cdots\right)$$
$$\leq \frac{x}{2}\left(1 + \frac{1}{2} + \frac{1}{2 \cdot 2} + \cdots\right) = x. \qquad \square$$

3. Lower bounds for $d \neq 2$. In this section and the next we will prove the lower bounds on $\lambda_c(M)$ given in Theorem 1. The key to doing this is the differential equation given in the Introduction:

$$(1) \qquad \frac{d}{dt}E|\xi_t^0| = (\lambda - 1)E|\xi_t^0| - \frac{\lambda}{v(M)} \sum_x \sum_{y \sim x} P(x, y \in \xi_t^0).$$

We begin with a result which gives the desired bound in $d \geq 3$.

PROPOSITION 3.1. *In any dimension, if M is large*

$$\lambda_c(M) \geq 1 + 1/9v(M).$$

PROOF. To estimate the second term on the right-hand side of (1), we first show that if $\lambda \geq 1$,

$$(2) \qquad \sum_{y \sim x} P(x \in \xi_t^0, y \in \xi_t^0) \geq \tfrac{1}{8} P(x \in \xi_{t-\ln 2}^0).$$

This inequality comes from combining three facts. (i) The probability the particle at x survives from $t - \ln 2$ to t is 1/2. (ii) If it survives, then the probability it will give birth to at least one particle is $1 - \exp(-\lambda \ln 2) \geq 1/2$. (iii) If a particle is born the probability it will survive for at least $\ln 2$ units of time is 1/2.

Plugging (2) into (1) gives for $t \geq \ln 2$,

$$\frac{d}{dt}E|\xi_t^0| \leq (\lambda - 1)E|\xi_t^0| - \frac{1}{8v(M)}E|\xi_{t-\ln 2}^0|. \tag{3}$$

Comparing with a branching process [i.e., dropping the negative term on the right-hand side of (1)] gives

$$E|\xi_t^0| \leq e^{(\lambda-1)\ln 2}E|\xi_{t-\ln 2}^0|.$$

Putting this in (3) gives

$$\frac{d}{dt}E|\xi_t^0| \leq \left(\lambda - 1 - \frac{e^{-(\lambda-1)\ln 2}}{8v(M)}\right)E|\xi_t^0|. \tag{4}$$

If $(\lambda - 1) \leq 1/9v(M)$, then the term in parentheses in (4) is less than $-0.01/v(M)$ for large M. Comparison with a branching process gives

$$E|\xi_{\ln 2}^0| \leq e^{(\lambda-1)\ln 2}.$$

So, using (4), we see that if $(\lambda - 1) \leq 1/9v(M)$, then

$$E|\xi_t^0| \leq \exp\left(\frac{\ln 2}{9v(M)} - \frac{(0.01)(t - \ln 2)}{v(M)}\right). \tag{5}$$

From the last result we see that $P(|\xi_t^0| \geq 1) \leq E|\xi_t^0| \to 0$ as $t \to \infty$. □

We turn now to the result in $d = 1$.

Proposition 3.2. *In $d = 1$ there is a constant C so that if M is large*

$$\lambda_c(M) \geq 1 + C/M^{2/3}.$$

Proof. By (1.6) it is enough to show that there is a $\delta > 0$ so that if $M \leq \delta\varepsilon^{-3/2}$, then $E|\xi_t^0| \leq 0.9$ for some t. To prove this, it is enough to show

Lemma 3.3. *For small ε, if $1/10\varepsilon \leq t \leq 1/\varepsilon$ and $E|\xi_t^0| \geq 0.9$, then $(d/dt)E|\xi_t^0| \leq -\varepsilon$.*

Since

$$E|\xi_{1/10\varepsilon}^0| \leq E|Z_{1/10\varepsilon}^0| = e^{1/10} \leq 1.2$$

(the last inequality coming from Lemma 2.7), Lemma 3.3 implies that $E|\xi_t^0|$ will be less than or equal to 0.9 at some time less than or equal to $0.4/\varepsilon$.

Proof of Lemma 3.3. We let $\lambda = 1 + \varepsilon$ in (1) and use the fact that $v(M) = 2M$ and $\lambda \geq 1$ to get

$$\frac{d}{dt}E|\xi_t^0| \leq \varepsilon E|\xi_t^0| - \frac{1}{2M}\sum_x \sum_{y\sim x} P(x, y \in \xi_t^0). \tag{6}$$

To estimate the second or "interference" term, we let

$$I_t = [-Kt^{1/2}, Kt^{1/2}],$$
$$N(x) = |\{y \sim x: y \in \xi_t^0\}|,$$
$$\eta_t = \{x \in \xi_t^0: N(x) \geq 6\varepsilon M\} \cap I_t,$$
$$\zeta_t = (\xi_t^0 - \eta_t) \cap I_t.$$

The points in η_t are "too crowded." The birth rate per particle is $1 + \varepsilon$, but these particles "lose" at least 3ε of their birth rate by giving birth onto occupied sites. Using the facts

$$\frac{1}{2M} \sum_x \sum_{y \sim x} 1_{(x,\, y \in \xi_t^0)} \geq 3\varepsilon|\eta_t| \tag{7}$$

and

$$E|\xi_t^0 \cap I_t^c| + E|\zeta_t| + E|\eta_t| = E|\xi_t^0| \tag{8}$$

in (6) gives

$$\frac{d}{dt}E|\xi_t^0| \leq \varepsilon\left(E|\xi_t^0 \cap I_t^c| + E|\zeta_t| - 2E|\eta_t|\right). \tag{9}$$

To estimate the first term on the right-hand side of (9), we observe that by (2.1)

$$E|\xi_t^0 \cap I_t^c| \leq E|Z_t(I_t^c)| = e^{(\lambda-1)t}P(X_t^0 \in I_t^c), \tag{10}$$

where X_t^0 is a continuous time random walk which starts at 0, takes steps at rate λ and has jumps which are uniformly distributed on $\{-1, -1 + 1/M, \dots, 1\} - \{0\}$. To get a bound on the right-hand side of (10) which is independent of M, we observe that if $\varepsilon \to 0$ and $M_\varepsilon \to M \in \{1, 2, \dots\}$, the central limit theorem implies

$$\sup\{P(X_t^0 \in I_t^c): 1/10\varepsilon \leq t \leq 1/\varepsilon\} \to P(\sigma_M \chi \notin [-K, K]) \equiv \alpha(K).$$

Here χ is a normal random variable with mean 0 and variance 1. σ_M^2 is the variance defined near (2.6), and σ_∞^2 is the limit of σ_M^2 as $M \to \infty$. From the last observation, it follows there is a constant ε_0 so that if $\varepsilon < \varepsilon_0$ and $1/10\varepsilon \leq t \leq 1/\varepsilon$, then

$$E|\xi_t^0 \cap I_t^c| \leq 2e\alpha(K). \tag{11}$$

[If not there would be a sequence $\varepsilon_k \to 0$, $M_k \to M$ so that the sup of the left-hand side over $1/10\varepsilon \leq t \leq 1/\varepsilon$ is greater than $2e\alpha(K)$.]

To estimate the second term on the right-hand side of (9), we observe

$$|\zeta_t| \leq (6\varepsilon M)(2K\sqrt{t}) \tag{12}$$

since $|[i, i+1] \cap \zeta_t| \leq 6\varepsilon M$, and there are $2K\sqrt{t}$ such intervals. Using $t \leq 1/\varepsilon$ in (12) gives that

$$|\zeta_t| \leq 12K\varepsilon^{1/2}M. \tag{13}$$

The last result is not good enough since we want the right-hand side to be small when $M = \delta\varepsilon^{-3/2}$. To improve the estimate, we observe that

$$E|\zeta_t| = E\big(|\zeta_t|\,|\xi_t^0 \neq \varnothing\big)P\big(\xi_t^0 \neq \varnothing\big) \leq 12K\varepsilon^{1/2}MP\big(|Z_t^0| > 0\big) \tag{14}$$

by (13). Using Lemma 2.1 now, it follows that there are constants C, ε_1 (which do not depend on K) so that if $\lambda = 1 + \varepsilon$, $\varepsilon < \varepsilon_1$ and $t \geq 1/10\varepsilon$, then

$$E|\zeta_t| \leq CK\varepsilon^{3/2}M. \tag{15}$$

(15) is the last piece of the puzzle. To put things together, let $\varepsilon < \varepsilon_2 = \min(\varepsilon_0, \varepsilon_1)$. Pick K so that in (11)

$$E|\xi_t^0 \cap I_t^c| \leq 2e\alpha(K) \leq 0.1. \tag{16}$$

If $M \leq \delta\varepsilon^{-3/2}$ the right-hand side of (15) is at most $CK\delta$ so if we pick δ small

$$E|\zeta_t| \leq 0.1. \tag{17}$$

Using (8) now with (16) and (17), we see that if $E|\xi_t^0| \geq 0.9$, then

$$E|\eta_t| > 0.7. \tag{18}$$

Using (16)–(18) in (9) gives for $\varepsilon < \varepsilon_2$,

$$\frac{d}{dt}E|\xi_t^0| \leq -1.2\varepsilon,$$

a conclusion slightly better than the result claimed in Lemma 3.3. □

4. Lower bounds in $d = 2$.

In this section we will prove

PROPOSITION 4.1. *In $d = 2$ there is a constant C so that if M is large*

$$\lambda_c(M) \geq 1 + C(\log M)/M^2.$$

Let $B_{m,n} = [m, m+1) \times [n, n+1)$. The key to the proof will be upper and lower bounds on the moments of $Z_t^0(B_{m,n})$. Recall the notation of Lemma 2.2, $m_k(t, x, A) = E(Z_t^x(A)^k)$, and that we drop the subscript when $k = 1$.

LEMMA 4.2. *There are constants C_1, C_2, C_3 (which are independent of M) so that if $\lambda = 1 + \varepsilon$, $t \leq 1/\varepsilon$ and $\varepsilon \leq 1$ we have*

(a) $m(t, x, B_{m,n}) \leq C_1(1+t)^{-1}$,
(b) $m_2(t, x, B_{m,n}) \leq C_2(1+t)^{-1}\log(2+t)$,
(c) $m_3(t, x, B_{m,n}) \leq C_3(1+t)^{-1}\log^2(2+t)$.

PROOF. To prove (a), notice that

$$m(t, x, B_{m,n}) = e^{(\lambda-1)t}P(X_t^x \in B_{m,n}) \leq eC/(1+t)$$

by Lemma 2.4. The C here is not the same as the one in that theorem. In what follows we will put t or $1 + t$ in the denominator as convenience dictates.

To prove (b), we use Lemma 2.2 to write

$$\begin{aligned} m_2(t, x, B_{m,n}) &= m(t, x, B_{m,n}) \\ &\quad + \int \beta_M(dz) \int_0^t ds \int m(t-s, x, dy) \\ &\quad \times m(s, y, B_{m,n}) m(s, y+z, B_{m,n}). \end{aligned}$$

To estimate the integral, we observe

$$m(s, y+z, B_{m,n}) \le C_1(s+1)^{-1},$$

$$\int m(t-s, x, dy) m(s, y, B_{m,n}) = m(t, x, B_{m,n}) \le C_1(t+1)^{-1}.$$

So if $\lambda \le 2$,

$$\begin{aligned} m_2(t, x, B_{m,n}) &\le C_1(t+1)^{-1} + 2\left(\int_0^t ds\, C_1(s+1)^{-1} \right) C_1(t+1)^{-1} \\ &\le C_2(t+1)^{-1} \log(t+2) \end{aligned}$$

for suitable C_2. To prove (c), we use Lemma 2.2 again

$$\begin{aligned} m_3(t, x, B_{m,n}) = m(t, x, B_{m,n}) + \frac{3}{2} \int \beta_M(dz) \int_0^t ds \int m(t-s, x, dy) \\ \times \{ m(s, y, B_{m,n}) m_2(s, y+z, B_{m,n}) \\ + m_2(s, y, B_{m,n}) m(s, y+z, B_{m+n}) \}. \end{aligned}$$

To estimate the first integral, we observe

$$m_2(s, y+z, B_{m,n}) \le C_2(s+1)^{-1} \log(s+2),$$

$$\int m(t-s, x, dy) m(s, y, B_{m,n}) = m(t, x, B_{m,n}) \le C_1(t+1)^{-1},$$

$$\int_0^t C_2(s+1)^{-1} \log(s+2)\, ds \le \log(t+2) \int_0^t C_2(s+1)^{-1}\, ds \le C_2 \log^2(t+2).$$

The second integral is estimated in almost the same way. We begin with the observation

$$m_2(s, y, B_{m,n}) \le C_2(s+1)^{-1} \log(s+1),$$

but this time we use the fact that for $\|z\| \le 1$,

$$\begin{aligned} \int m(t-s, x, dy) m(s, y+z, B_{m,n}) &= \int m(t-s, x, dy) m(s, y, B_{mn} - z) \\ &\le m(t, x, B_{m,n}^{+1}), \end{aligned}$$

where $B_{m,n}^{+1} = [m-1, m+2) \times [n-1, n+2)$. Since $B_{m,n}^{+1}$ is the union of nine unit squares

$$m(t, x, B_{m,n}^{+1}) \le 9C_1(t+1)^{-1}.$$

Integrating from 0 to t gives the same result as before except for a factor 9. Putting things together gives for $\lambda \le 2$,

$$\begin{aligned} m_3(t, x, B_{m,n}) &\le C_1(t+1)^{-1} + 30C_1C_2(t+1)^{-1}\log^2(t+2) \\ &\le C_3(t+1)^{-1}\log^2(t+2) \end{aligned}$$

for suitable C_3, completing the proof of Lemma 4.2. □

To prove Proposition 4.1, we will need lower bounds on the first two moments which are of the same order as the upper bounds in Lemma 4.2.

LEMMA 4.3. *Let $\lambda = 1 + \varepsilon$. There are constants b_1, b_2, $t_0 \ge 1$ and t_1 so that if $(m, n) \in [-t^{1/2}, t^{1/2}]$, then for $t \ge t_0$,*

(a) $$m(t, 0, B_{m,n}) \ge b_1(1+t)^{-1}e^{(\lambda-1)t},$$

and for $t \ge t_1$,

(b) $$m_2(t, 0, B_{m,n}) \ge b_2(1+t)^{-1}\log(2+t).$$

REMARK. The two constants t_0 and t_1 could, of course, be combined into one. The formulation above is convenient for the proof.

SKETCH OF PROOF. (b) follows from (a) and Lemma 2.2. To prove (a), we begin by observing that

$$m(t, 0, B_{m,n}) = e^{(\lambda-1)t}P(X_t^0 \in B_{m,n}).$$

If M were fixed, (a) would follow from the local central theorem, but M is not fixed, so we have to prove a version which is uniform in M. That proof takes some work and interrupts the flow of ideas, so the proof of Lemma 4.3 is postponed to the end of the section.

PROOF OF PROPOSITION 4.1. As in the proof of Proposition 3.2 it is enough to show

LEMMA 4.4. *There is an $a \in (0,1)$ so that if $a/2\varepsilon \le t \le a/\varepsilon$ and $E|\xi_t^0| \ge 1 - a$, then for $\varepsilon < \varepsilon_0$ we have*

$$\frac{d}{dt}E|\xi_t^0| \le -3\varepsilon.$$

For once this result is established, we can use the fact that

$$E|\xi_{a/2\varepsilon}^0| \le E|Z_{a/2\varepsilon}| = e^{a/2} \le 1 + a$$

(the second inequality coming from Lemma 2.7) to conclude that if $E|\xi_t^0| \ge 1 - a$ for all $a/2\varepsilon \le t \le a/\varepsilon$, then

$$E|\xi_{a/\varepsilon}^0| \le 1 + a - 3a/2 = 1 - a/2 < 1,$$

and (1.6) implies that the contact process will die out.

We begin the proof of Lemma 4.4 with the following trivial, but useful, observation.

LEMMA 4.5. *If* $0 \le X \le Y$ *and* $EX \ge (1-\theta)EY$, *then*

$$EY^2 - EX^2 \le 2(\theta EY EY^3)^{1/2}.$$

PROOF. $Y^2 - X^2 = (Y-X)(Y+X)$. Applying the Cauchy–Schwarz inequality to $(Y-X)^{1/2}$ and $(Y-X)^{1/2}(Y+X) \le 2Y^{3/2}$ gives

$$E(Y^2 - X^2) \le \big(E(Y-X)E(4Y^3)\big)^{1/2},$$

proving the lemma. □

To prove Lemma 4.4, we pick θ so that $(b_2 - 2(\theta C_1 C_3)^{1/2}) > 0$ and then pick $a < 1/3$ so that $0 < e^a(1 - b_1\theta) < 1 - a$. Here b_1, b_2, C_1 and C_3 are the constants of Lemmas 4.3 and 4.2. The reader will see the reasons for these choices as the argument progresses. Let

$$V_{m,n}(t) = Z_t^0([m, m+1) \times [n, n+1)),$$
$$U_{m,n}(t) = \xi_t^0([m, m+1) \times [n, n+1)),$$
$$I_t = [-t^{1/2}, t^{1/2}]^2 \cap \mathbb{Z}^2.$$

Now if $a/2\varepsilon \le t \le a/\varepsilon$, we claim that for at least one-half of the $(m,n) \in I_t$,

$$EU_{m,n}(t)/EV_{m,n}(t) \ge 1 - \theta. \tag{1}$$

For otherwise

$$\begin{aligned}\sum_{(m,n)\in I_t} (EV_{m,n}(t) - EU_{m,n}(t)) &\ge \frac{1}{2}|I_t|\theta \inf_{(m,n)\in I_t} EV_{m,n}(t) \\ &\ge \frac{1}{2}|I_t|\theta b_1(1+t)^{-1}e^{(\lambda-1)t}\end{aligned} \tag{2}$$

by Lemma 4.3. Since $|I_t| = (2[t^{1/2}] + 1)^2 \ge 2t + 2$ for large t, our choice of a implies that

$$E|\xi_t^0| = \sum_{(m,n)} EU_{m,n}(t) \le e^{(\lambda-1)t}(1 - b_1\theta) \le e^a(1 - b_1\theta) < 1 - a, \tag{3}$$

contradicting the assumption that $E|\xi_t^0| \ge 1 - a$. Let

$$J_t = \{(m,n) \in I_t\colon EU_{m,n}(t)/EV_{m,n}(t) \ge 1 - \theta\}.$$

Applying Lemma 4.5 to $X = U_{m,n}(t)$ and $Y = V_{m,n}(t)$ and then using Lemma 4.2 gives

$$EV_{m,n}^2(t) - EU_{m,n}^2(t) \le 2(\theta C_1 C_3)^{1/2}\big((1+t)^{-1}\log(2+t)\big) \tag{4}$$

for $(m,n) \in J_t$. Lemma 4.3 tells us that

$$EV_{m,n}^2(t) \ge b_2(1+t)^{-1}\log(2+t) \tag{5}$$

for all $(m, n) \in I_t$, so using the fact that $|I_t| \geq 2t + 2$ for large t, we have

$$\sum_{(m,n)\in J_t} EU_{m,n}^2(t) \geq \frac{1}{2}(2t+2)\big(b_2 - 2(\theta C_1 C_3)^{1/2}\big)(1+t)^{-1}\log(2+t) \tag{6}$$
$$\geq \beta\big(\log \varepsilon^{-1} + \log(a/2)\big)$$

for $t \geq a/2\varepsilon$, if $\beta = (b_2 - 2(\theta C_1 C_3)^{1/2})$. β is positive by the choice of θ.

The interference term

$$\sum_{x,\, y\sim x} P\big(x, y \in \xi_t^0\big) \geq \sum_{m,n} E\big(U_{m,n}(t)(U_{m,n}(t) - 1)\big)$$
$$\geq \sum_{(m,n)\in J_t} EU_{m,n}^2(t) - \sum_{(m,n)} EV_{m,n}(t) \tag{7}$$
$$\geq \beta\big(\log \varepsilon^{-1} + \log(a/2)\big) - e^a$$

if $t \in [a/2\varepsilon, a/\varepsilon]$. Plugging the last bound into the differential equation (1.2) gives

$$\frac{d}{dt}E|\xi_t^0| \leq \varepsilon E|\xi_t^0| - \frac{\lambda}{v(M)}\big(\beta\big(\log \varepsilon^{-1} + \log(a/2)\big) - e^a\big). \tag{8}$$

Now if $t \leq a/\varepsilon$ and $a \leq 1$, then

$$E|\xi_t^0| \leq E|Z_t^0| \leq e^a \leq 1 + 2a \tag{9}$$

by Lemma 2.7. Combining (8) and (9) gives that if $v(M) \leq (\beta/5)(\varepsilon^{-1}\log \varepsilon^{-1})$,

$$\frac{d}{dt}E|\xi_t^0| \leq \varepsilon\big\{(1+2a) - 5 + \big(g(a)/\log \varepsilon^{-1}\big)\big\}, \tag{10}$$

where $g(a) = 5\log(a/2) + 5e^a/\beta$. For fixed a, if ε is small enough, $g(a)/\log(\varepsilon^{-1}) < 1/3$. Since a was chosen to be less than or equal to $1/3$, it follows that

$$\frac{d}{dt}E|\xi_t^0| \leq -3\varepsilon, \tag{11}$$

proving Lemma 4.4. □

The last detail remaining in the proof of Proposition 4.1 is the proof of Lemma 4.3. The key to the proof is a slight generalization of the local central limit theorem. Let X_t^M be a continuous-time random walk with $X_0^M = 0$, which jumps at rate 1 and has jumps uniformly distributed on $\{x \in Z^d/M\colon 0 < \|x\| \leq 1\}$. The last definition is inconsistent with our earlier practice of using superscripts to denote the starting point, but we will only use the new notation until the end of this section, and we will not use the old notation during that time.

LEMMA 4.6. *If $M \to \infty$ and $x_n/n^{1/2} \to x$ then for any Borel set with $|\partial B| = 0$, $|B| < \infty$,*

$$n^{d/2}P\big(X_n^M \in x_n + B\big) \to |B|n(x),$$

where

$$n(x) = (2\pi\sigma^2)^{-d/2}\exp(-|x|^2/2\sigma^2),$$

σ^2 *is the limit of the* σ_M^2 *and* σ_M^2 *is the variance of one component of* X_1^M.

PROOF. We will only prove the result for rectangles $B = [a_1, b_1] \times \cdots \times [a_d, b_d]$, since that is all we need. Since the components of X_n^M are independent, it suffices to prove the result in one dimension. To do this, we will use ideas from Sections 10.4 and 10.2 of Breiman (1968). Let

$$h_0(y) = \frac{1}{\pi}\frac{1-\cos y}{y^2} \quad \text{and} \quad h_\theta(y) = e^{i\theta y}h_0(y).$$

If we introduce the Fourier transform

$$\hat{g}(u) = \int e^{iuy}g(y)\,dy,$$

then

$$\hat{h}_0(u) = \begin{cases} 1-|u| & |u| \le 1, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\hat{h}_\theta(u) = \hat{h}_0(u+\theta).$$

We will show that for any θ,

$$\sqrt{n}\,Eh_\theta(X_n^M - x_n) \to n(x)\int h_\theta(y)\,dy. \tag{12}$$

Taking $\theta = 0$ and letting

$$\alpha_n = \sqrt{n}\,Eh_0(X_n^M - x_n) \quad \text{and} \quad \alpha = n(x)\int h_0(y)\,dy = n(x),$$

it follows that $\alpha_n \to \alpha$. Let

$$\mu_n(A) = \sqrt{n}\,P(X_n^M - x_n \in A) \quad \text{and} \quad \mu(A) = n(x)|A|,$$

where $|A|$ is the Lebesgue measure of A. Finally, define probability measures by

$$\nu_n(B) = \alpha_n^{-1}\int_B h_0(y)\mu_n(dy) \quad \text{and} \quad \nu(B) = \alpha^{-1}\int_B h_0(y)\mu(dy).$$

Written in terms of ν_n and ν, (12) becomes

$$\int e^{i\theta y}\nu_n(dy) \to \int e^{i\theta y}\nu(dy). \tag{13}$$

Since this holds for all θ, we have shown that the characteristic functions converge, and it follows that $\nu_n \Rightarrow \nu$, where $\Rightarrow$ denotes weak convergence. Now the function

$$k(y) = 1_{[a,b]}(y)/h_0(y)$$

is bounded and continuous a.s. with respect to ν, so it follows that

$$\int k(y)\nu_n(dy) \to \int k(y)\nu(dy).$$

Since $\alpha_n \to \alpha$, this implies the desired result:

$$\sqrt{n}\, P\big(X_n^M - x_n \in [a, b]\big) \to (b - a)\,n(x),$$

and it only remains to prove (12).

Let $\varphi_M(u) = E\exp(iuX_1^M)$, and let $J = [-j, j]$ be a closed interval. $|\varphi_M(u)| = 1$ if and only if $u = 2\pi Mn$ for some $n \in Z$, so if M is large $|\varphi_M(u)| \neq 1$ on $J - \{0\}$. Let h be a bounded continuous function such that

$$\hat{h}(u) = \int e^{iux} h(x)\, dx$$

vanishes on J^c, for example, each h_θ qualifies if $j \geq j(\theta)$. We will show that (12) holds for all such h. Inverting the Fourier transform gives

$$h(x) = \frac{1}{2\pi}\int e^{-iux}\hat{h}(u)\, du. \tag{14}$$

[See Chung (1974), page 155.] If we let F_n^M be the distribution of $X_n^M - x_n$, then

$$\begin{aligned} Eh\big(X_n^M - x_n\big) &= \frac{1}{2\pi}\int\int e^{-iux}\hat{h}(u)\, du\, dF_n^M(x) \\ &= \frac{1}{2\pi}\int \varphi_M^n(-u)\exp(ix_n u)\hat{h}(u)\, du. \end{aligned} \tag{15}$$

Expanding the characteristic function in a power series at 0 and using the fact that $E|X_1^M|^3 \leq C$ gives [see Billingsley (1978), formula (26.5), page 297]

$$\big|\varphi_M(u) - \big(1 - \sigma_M^2 u^2/2\big)\big| \leq 2\frac{C|u|^3}{3!}. \tag{16}$$

(Recall σ_M^2 = the variance of X_1^M.) As $M \to \infty$, $\sigma_M^2 \to \sigma^2 > 0$, so we can pick M large enough so that $|\sigma_M^2 - \sigma^2| < \sigma^2/2$, and then b small enough so that

$$2\frac{C|u|^3}{3!} \leq \frac{\sigma^2 u^2}{8} \quad \text{for } u \in (-b, b). \tag{17}$$

With this done, we have

$$0 \leq |\varphi_M(u)| \leq 1 - \frac{\sigma^2 u^2}{4} + \frac{\sigma^2 u^2}{8} \leq 1 - \frac{\sigma^2 u^2}{8} \leq \exp\left(-\frac{\sigma^2 u^2}{8}\right). \tag{18}$$

As $M \to \infty$, $X_1^M \Rightarrow X_1^\infty$, so $\varphi_M(u) \to \varphi_\infty(u)$ uniformly in every finite interval [Chung (1974), Theorem 6.3.1, page 160]. It follows from (18) that if M is large and $N = (-b, b)$, then for $u \in J - N$,

$$|\varphi_M(u)| \leq 1 - \beta \quad \text{with } \beta > 0. \tag{19}$$

From the last observation it follows that

$$\sqrt{n}\left|\frac{1}{2\pi}\int_{N^c}\varphi_M^n(-u)\exp(iux_n)\hat{h}(u)\, du\right| \leq \frac{2j}{2\pi}\sqrt{n}\,(1 - \beta)^n\|\hat{h}\|_\infty \to 0 \tag{20}$$

as $n \to \infty$. For the interval N, change variables $u = v/\sqrt{n}$ to get

$$(21)\quad \begin{aligned}&\frac{1}{2\pi}\int_N \varphi_M^n(-u)\exp(iux_n)\hat{h}(u)\,du\\ &= \frac{1}{2\pi}\int_{-b\sqrt{n}}^{b\sqrt{n}} \varphi_M^n(-v/\sqrt{n})\exp(ix_n v/\sqrt{n})\hat{h}(v/\sqrt{n})\,dv/\sqrt{n}.\end{aligned}$$

The central limit theorem implies that

$$(22)\qquad \varphi_M^n(-v/\sqrt{n}) \to \exp(-\sigma^2 v^2/2),$$

and by (18) the integrand is dominated by $\|\hat{h}\|_\infty \exp(-\sigma^2 V^2/8)$. So by (20)–(22)

$$(23)\qquad \frac{\sqrt{n}}{2\pi}\int \varphi_M^n(-u)e^{ix_n u}\hat{h}(u)\,du \to \frac{\hat{h}(0)}{2\pi}\int_{-\infty}^{\infty} e^{-\sigma^2 v^2/2}e^{ixv}\,dv.$$

To evaluate the right-hand side, we recall the definition of $\hat{h}(0)$ and do a little calculus (or apply the inversion formula to the normal distribution) to see that the limit is

$$\mathscr{n}(x)\int h(y)\,dy,$$

proving (12) and completing the proof of Lemma 4.6. □

REMARKS. (i) Above we have proved the result only for times $n \in \mathbb{Z}_+$. To handle a general t, observe that if $n \le t \le n+1$, X_t^M is the sum of n independent copies of $X^M(t/n)$ and this distribution converges weakly to X_1^∞ as $n \to \infty$. To handle jumps at rate $\lambda > 1$, just change time $t \to \lambda t$.

(ii) In the proof we used the fact that $E|X_1^M|^3 \le C$. By more careful use of the formula from Billingsley in the derivation of (16), this condition can be replaced by one of the Lindeberg type. We leave this extension to the reader who needs it.

With Lemma 4.6 established, the proof of (a) in Lemma 4.3 is trivial. As we observed earlier

$$m(t,0,B_{m,n}) = e^{(\lambda-1)t}P(X_t^M \in B_{m,n}).$$

Lemma 4.6 implies that if $t, M \to \infty$, $\lambda \to 1$ and $m_t/t^{1/2} \to x$, $n_t/t^{1/2} \to y$, then

$$(24)\qquad tP(X_t^M \in B(m_t, n_t)) \to \frac{1}{2\pi\sigma^2}e^{-(x^2+y^2)/2\sigma^2}.$$

From this it follows that for any k there are constants $b_1(k)$, M_0 and $t_0 \ge 1$ (which depend on k), so that if $M \ge M_0$, $t > t_0$ and $(m,n) \in [-(kt)^{1/2},(kt)^{1/2}]^2$, then

$$(25)\qquad m(t,0,B_{m,n}) \ge \frac{b_1(k)}{(1+t)}e^{(\lambda-1)t}.$$

The last result with $k = 1$ is (a). We will need the more general result in the proof of (b) which follows.

From Lemma 2.2, $m_2(t, 0, B_{m,n})$ is larger than

$$\int \beta_M(dz) \int_{t_0}^{t/2} ds \int_{\|y-(m,n)\| \le s^{1/2}} m(t-s, 0, dy) \times m(s, y, B_{m,n}) m(s, y+z, B_{m,n}). \tag{26}$$

Because of the restrictions on y and s, the last two terms are at least (recall $\lambda \ge 1$)

$$b_1(1)/(s+1). \tag{27}$$

Now $\|y - (m, n)\| \le s^{1/2} \le t^{1/2}$ and $\|(m, n)\| \le t^{1/2}$ imply $\|y\| \le 2t^{1/2} \le (8(t-s))^{1/2}$, so

$$\int_{\|y-(m,n)\| \le s^{1/2}} m(t-s, 0, dy) \ge 2sb_1(8)/(1+t). \tag{28}$$

Using (27) and (28) in (26) gives

$$\begin{aligned} m_2(t, 0, B_{m,n}) &\ge \int_{t_0}^{t/2} 2s \frac{b_1(8)}{(1+t)} \left(\frac{b_1(1)}{(1+s)} \right)^2 ds \\ &\ge \frac{b_1(8) b_1(1)^2}{(1+t)} \int_{t_0}^{t/2} (1+s)^{-1} ds \\ &= \frac{C}{1+t} \left(\log\left(1 + \frac{t}{2}\right) - \log(1 + t_0) \right) \ge b_2 (1+t)^{-1} \log(2+t) \end{aligned}$$

for $t \ge t_1$ and suitable b_2. □

5. A block construction for the contact process. Our strategy for proving the contact process survives is to show that in "space-time" it dominates two-dimensional site percolation with p close to 1. To explain our construction, we have to recall how the contact process can be constructed from a "graphical representation." For each $x \in \mathbb{Z}^d/M$ we have a rate 1 Poisson process $\{T_n^x, n = 1, 2, \dots\}$ and for all x, y with $0 < \|x - y\| \le 1$, we have a rate $\lambda/v(M)$ Poisson process $\{T_n^{(x,y)}, n = 1, 2, \dots\}$. At times T_n^x we write a δ at x to indicate that the particle at x dies (if one is present). At times $T_n^{x,y}$ we draw an arrow from x to y to indicate that if x is occupied and y is vacant there will be a birth at y.

To construct the process from the graphical representation we say there is a path from (x, s) to (y, t) if there is a sequence of times $s = s_0 < s_1 \cdots < s_{n+1} = t$ and a sequence of points $x = x_0, x_1, \dots, x_n = y$ so that (i) at time s_i, $1 \le i \le n$, there is an arrow from x_{i-1} to x_i and (ii) there is no δ at x_i during (s_i, s_{i+1}) for $0 \le i \le n$.

To construct the contact process starting from the set of sites A occupied at time s, let

$$\xi_t^{(A,s)} = \{y: \text{there is a path from } (x, s) \text{ to } (y, t) \text{ for some } x \in A\}.$$

It should be clear that this recipe constructs the contact process. If not, see Chapter VI of Liggett (1985) or Chapter 4 of Durrett (1988) for more details.

To describe the block construction, we need some notation. In what follows K_1, K_2 and K_3 are constants which will be chosen later. Let $\mathscr{L} = \{(j,k) \in \mathbb{Z}^2: j+k \text{ is even}\}$. Let $L = (K_2/\varepsilon)$, $l = L^{1/2}$, and let

$$I_j = [(2j-1)l, (2j+1)l] \times [-l, l]^{d-1}.$$

At the kth stage of the construction we will be given sets $A_{j,k} \subset I_j$ for $(j,k) \in \mathscr{L}$ with

$$\text{(I)} \qquad K_1/\varepsilon \le |A_{j,k}| \le 2K_1/\varepsilon,$$

$$\text{(II)} \qquad |A_{j,k} \cap B_n| \le 2K_3 h(\varepsilon) \quad \text{for all } n \in \mathbb{Z}^d,$$

where $B_n = [n_1 - 1, n_1 + 1) \times \cdots \times [n_d - 1, n_d + 1)$ and

$$h(\varepsilon) = \begin{cases} \varepsilon^{-1/2} & d = 1, \\ \log(\varepsilon^{-1}) & d = 2, \\ 1 & d \ge 3. \end{cases}$$

We will say that (j,k) is open, if a modified contact process $\bar{\xi}_t^{j,k}$ which starts with $A_{j,k}$ occupied at time kL and is not allowed to give birth outside $H_j = ((2j-2)l, (2j+2)l) \times \mathbb{R}^{d-1}$ survives up to time $(k+1)L$, and there are sets $A^i_{j,k+1} \subset \bar{\xi}^{j,k}_{(k+1)L} \cap I_{j+i}$ for $i = -1, 1$ which have

$$\text{(I}'\text{)} \qquad \left|A^i_{j,k+1}\right| = K_1/\varepsilon,$$

$$\text{(II}'\text{)} \qquad \left|A^i_{j,k+1} \cap B_n\right| \le K_3 h(\varepsilon) \quad \text{for all } n \in \mathbb{Z}^d.$$

(Here and in what follows we will assume K_1 is chosen so that K_1/ε is an integer.) If (j,k) is not open we set $A^i_{j,k+1} = \varnothing$. To continue the construction on level $k+1$, we let $A_{j,k+1} = A^{+1}_{j-1,k+1} \cup A^{-1}_{j+1,k+1}$.

A sequence of points $(m_0, n), (m_1, n+1), \dots, (m_k, n+k) \in \mathscr{L}$ is said to be an open path if all the points are open and for $0 \le i \le k-1$, $m_{i+1} \in \{m_i - 1, m_i + 1\}$. The construction is designed so that (a) if the contact process starts with $A_{0,0}$ satisfying (I) and (II) occupied and there is an infinite open path on $\mathscr{L}$ starting at $(0,0)$ (i.e., "percolation occurs"), then the contact process survives, and (b) the regions H_i, H_j are disjoint if $|i-j| \ge 2$. [Recall $\mathscr{L} = \{(m,n) \in Z^2: m+n \text{ is even}\}$ and look at Figure 1 which shows the boxes in $d = 1$.]

Property (b) implies that given what has happened at levels $k \le n-1$, the fates of sites (m,n) on a fixed level n are independent. Combining this observation with (a) and a result from Durrett (1984) [see (1) in Section 10] gives

LEMMA 5.1. *If for any $A_{j,k}$ satisfying* (I) *and* (II), *the probability (j,k) is open is greater than or equal to $1-\kappa$ where $\kappa < 1/81$, then the probability there*

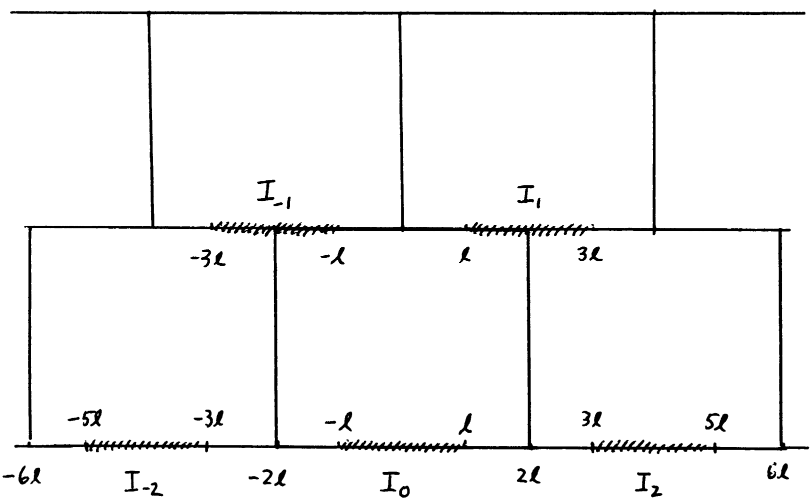

FIG. 1.

is no infinite open path from $(0,0)$ *on* $\mathscr{L}$ *is at most*

$$\sum_{m=4}^{\infty} 3^m \kappa^{m/4} = \frac{81\kappa}{1 - 3\kappa^{1/4}}.$$

PROOF. One can easily verify by induction that the open sites on $\mathscr{L}$ dominate oriented site percolation with parameter κ. The quantity given above is an upper bound for the probability that site percolation fails to percolate; see Durrett (1984), Section 10. □

6. Estimates for branching random walk. Let $\lambda = 1 + \varepsilon$, $L = (K_2/\varepsilon)$ and let

$$I_j = \left[(2j-1)L^{1/2}, (2j+1)L^{1/2}\right] \times \left[-L^{1/2}, L^{1/2}\right]^{d-1}.$$

Let $\bar{Z}_t^A$ denote the branching random walk starting from a set $A \subset I_0$ [i.e., $Z_0^A(x) = 1$ if $x \in A$, $= 0$, otherwise], and modified so that no births are allowed at points outside $(-2L^{1/2}, 2L^{1/2}) \times \mathbb{R}^{d-1}$. In this section we will prove

LEMMA 6.1. *If* $\delta > 0$, *there are constants* ε_0, K_1 *and* K_2 *so that if* $|A| \geq (K_1/\varepsilon)$, *then for* $\varepsilon < \varepsilon_0$,

$$P\left(\bar{Z}_L^A(I_j) \geq 3K_1/\varepsilon,\ j = -1 \text{ and } 1\right) > 1 - \delta.$$

PROOF. The proof of Lemma 6.1 is based on estimating the first and second moments of $\bar{Z}_t^A$. Let $\bar{Z}_t^x$ denote the modified branching random walk starting from $\bar{Z}_0^x = \{x\}$ (i.e., $\bar{Z}_t^A$ when $A = \{x\}$). First note that

$$E\bar{Z}_t^x(I_j) = e^{(\lambda-1)t} P\left(\bar{X}_t^x \in I_j\right), \tag{1}$$

where $\bar{X}_t^x$ denotes the random walk killed when it leaves $(-2L^{1/2}, 2L^{1/2}) \times \mathbb{R}^{d-1}$. The proof of (1) is the same as that of (2.1).

To estimate second moments, we use the trivial inequality

$$E\left(e^{-(\lambda-1)t}\bar{Z}_t^x(I_j)\right)^2 \le E\left(e^{-(\lambda-1)t}\bar{Z}_t^x(\mathbb{R}^d)\right)^2. \tag{2}$$

Setting $A = \mathbb{R}^d$ in Lemma 2.2 gives

$$E(Z_t^x)^2 = EZ_t^x + \lambda \int_0^t EZ_{t-s}^x (EZ_s^x)^2 \, ds.$$

So

$$\begin{aligned} E\left(e^{-(\lambda-1)t}Z_t^x\right)^2 &= e^{-(\lambda-1)t} + \lambda \int_0^t e^{-(\lambda-1)(t-s)} \, ds \\ &= e^{-(\lambda-1)t} + \frac{\lambda}{(\lambda-1)}\left(1 - e^{-(\lambda-1)t}\right) \le \frac{2}{\varepsilon} \end{aligned} \tag{3}$$

if $\varepsilon = \lambda - 1 \le 1$. Inequalities (2) and (3) imply that

$$\sigma^2\left(e^{-(\lambda-1)t}\bar{Z}_t^x(I_j)\right) \le 2/\varepsilon, \tag{4}$$

where $\sigma^2(Y)$ denotes the variance of Y.

The next step is to get a lower bound on $E\bar{Z}_t^x(I_j)$, or in view of (1), a lower bound on $P(\bar{X}_t^x \in I_j)$. To get a bound independent of M and $x \in I_0$, we observe that if $t = K_2/\varepsilon$ and $x = \theta_\varepsilon(K_2/\varepsilon)^{1/2}$ with $\theta_\varepsilon \to \theta \in [-1,1]^d$, and we let $M_\varepsilon \to M$ (possibly ∞) as $\varepsilon \to 0$, then

$$P\left(\bar{X}_{K_2/\varepsilon}^x \in I_1\right) \to P\left(\bar{B}_1^{M,\theta} \in [1,2] \times [-1,1]^{d-1}\right). \tag{5}$$

Here $\bar{B}_t^{M,\theta}$ is a Brownian motion with mean 0 and covariance $\sigma_M^2 tI$, starting from θ and killed when its first component exits $[-2,2]$, σ_M^2 is the variance defined near (2.6) and σ_∞^2 is the limit of the σ_M^2. Since $\sigma_M^2 > 0$ and $\sigma_\infty^2 = 1/3$, it follows that σ_M^2 is bounded and bounded away from 0, so the right-hand side of (5) is at least $\rho > 0$ for all $\theta \in [-1,1]^d$ and $M \in \{1, 2, \dots, \infty\}$. From this it follows that, for any value of K_2, we have for $\varepsilon < \varepsilon_0(K_2)$,

$$P\left(\bar{X}_{K_2/\varepsilon}^x \in I_1\right) \ge \rho/2 > 0 \tag{6}$$

for all $x \in I_0 = [-L, L]^d$ and $M \ge 1$.

Combining (6) and (1) and recalling (4) shows that the variable

$$V = e^{-(\lambda-1)L}\bar{Z}_L^A(I_1)$$

has (for small ε)

$$EV \ge \rho|A|/2, \tag{7}$$

$$\sigma^2(V) \le 2|A|/\varepsilon. \tag{8}$$

So Chebyshev's inequality implies

$$P\left(V \le \frac{\rho|A|}{4}\right) \le P\left(|V - EV| \ge \frac{\rho|A|}{4}\right) \le \frac{2|A|}{\varepsilon}\frac{16}{\rho^2|A|^2} = \frac{32}{\rho^2}\frac{1}{|A|\varepsilon}. \tag{9}$$

If we pick K_1 so that $32/\rho^2 K_1 \le \delta/2$ and pick K_2 so that $\exp(K_2)\rho/4 \ge 3$, then for $\varepsilon < \varepsilon_0(K_2)$ and $|A| \ge K_1/\varepsilon$ we have

$$P\left(\bar{Z}^A_{K_2/\varepsilon}(I_1) \le 3K_1/\varepsilon\right) \le \delta/2. \tag{10}$$

The same estimate applies to I_{-1}, and we have proved the desired result. □

7. Upper bounds. To prove the upper bounds in Theorem 1, it is enough to show that if the volume (defined in the Introduction) satisfies

$$v(M) \ge \begin{cases} K_4\varepsilon^{-3/2} & d = 1, \\ K_4\varepsilon^{-1}\log(\varepsilon^{-1}) & d = 2, \\ K_4\varepsilon^{-1} & d \ge 3, \end{cases} \tag{1}$$

then the contact process survives when K_4 is large. If we let

$$h(\varepsilon) = \begin{cases} \varepsilon^{-1/2} & d = 1, \\ \log(\varepsilon^{-1}) & d = 2, \\ 1 & d \ge 3, \end{cases} \tag{2}$$

then (1) can be written as

$$v(M) \ge K_4\varepsilon^{-1}h(\varepsilon). \tag{3}$$

Our first step is to thin $\bar{Z}^A_{K_2/\varepsilon}$ (the process defined in Section 6), so that there are at most $K_3 h(\varepsilon)$ particles in any set $B_n = [n_1 - 1, n_1 + 1] \times \cdots \times [n_d - 1, n_d + 1]$ with $n \in \mathbb{Z}^d$. We call the thinned process $\hat{Z}^A_{K_2/\varepsilon}$.

LEMMA 7.1. *Suppose $\delta > 0$ and we pick ε_0, K_1, K_2 so that Lemma 6.1 holds. Let I_j be as in Section 6, and let $A \subset I_0$ with $K_1/\varepsilon \le |A| \le 2K_1/\varepsilon$. If K_3 is large and $\varepsilon < \varepsilon_0$, then*

$$P\left(\hat{Z}^A_{K_2/\varepsilon}(I_j) > 2K_1/\varepsilon \text{ for } j = -1, 1\right) > 1 - 2\delta.$$

PROOF. Our approach will be to show that the unrestricted process has $Z^A_{K_2/\varepsilon}(B_n) \le K_3 h(\varepsilon)$ with high probability and then to apply Lemma 6.1. The key to the estimate for the unrestricted process is obtaining upper bounds on the second moments. Lemma 2.2 tells us that

$$\begin{aligned} E\left(Z_t^x(B_n)\right)^2 = m(t,x,B_n) + \int \beta_M(dz)\int_0^t ds \int m(t-s,x,dy) \\ \times m(s,y,B_n)m(s,y+z,B_n). \end{aligned} \tag{4}$$

By Lemma 2.4, the first term is less than or equal to $C(1+t)^{-d/2}$ and

$$m(s, y+z, B_n) \le C(1+s)^{-d/2}. \tag{5}$$

Now

$$\int m(t-s,x,dy)m(s,y,B_n) = m(t,x,B_n) \le C\varepsilon^{d/2} \tag{6}$$

if $t = K_2/\varepsilon$. Here and in what follows C is a constant which depends on d, K_1 and K_2 and will change from line to line. As in Section 4, we have replaced t by $t+1$ for convenience and will continue to do so. Integrating from 0 to t gives

$$(7) \qquad \int_0^t ds(1+s)^{-d/2} \le \begin{cases} C(1+t)^{1/2} & d=1, \\ C\log(1+t) & d=2, \\ C & d \ge 3. \end{cases}$$

Combining the results above gives for $t = K_2/\varepsilon$ and $\varepsilon \le 1/2$,

$$(8) \qquad E\left(Z_t^x(B_n)^2\right) \le \begin{cases} C & d=1, \\ C\varepsilon\log(\varepsilon^{-1}) & d=2, \\ C\varepsilon^{d/2} & d \ge 3. \end{cases}$$

Let $A \subset I_0$ be a set with $\le 2K_1/\varepsilon$ particles and write

$$(9) \qquad E\left(Z_t^A(B_n)^2\right) = \sum_{x\in A} E\left(Z_t^x(B_n)^2\right) + \sum_{x\in A}\sum_{\substack{y\in A\\ y\ne x}} EZ_t^x(B_n)EZ_t^y(B_n).$$

Using (8), the first term on the right-hand side of (9) is smaller than

$$(10) \qquad \begin{cases} C/\varepsilon & d=1, \\ C\log\varepsilon^{-1} & d=2, \\ C\varepsilon^{(d/2-1)} & d \ge 3. \end{cases}$$

The second term on the right-hand side of (9) is smaller than

$$(11) \qquad \begin{aligned} \Big(\sum_{x\in A} EZ_t^x(B_n)\Big)^2 &\le \Big(|A|e^{(\lambda-1)t}\sup_x P(X_t^x \in B_n)\Big)^2 \\ &\le \left(2K_1\varepsilon^{-1}C(1+t)^{-d/2}\right)^2 \le C\varepsilon^{d-2} \end{aligned}$$

for $t = K_2/\varepsilon$ and $\varepsilon \le 1/2$. Combining this with the estimate in (10) gives

$$(12) \qquad E\left(Z_t^A(B_n)^2\right) \le \begin{cases} C/\varepsilon & d=1, \\ C\log\varepsilon^{-1} & d=2, \\ C\varepsilon^{d/2-1} & d=3. \end{cases}$$

Let $V_n = Z_t^A(B_n)$. Recalling the definition of $h(\varepsilon)$ in (2), it follows that

$$(13) \qquad E(V_n;\, V_n > K_3h(\varepsilon)) \le \frac{EV_n^2}{K_3h(\varepsilon)} \le \frac{C}{K_3}\varepsilon^{d/2-1}.$$

If we let

$$D_j = \sum_n V_n 1_{(V_n > K_3h(\varepsilon))}$$

where the sum is over $n \in I_j$ and $j = -1$ or 1, then

$$(14) \qquad \bar{Z}_t^A(I_j) - \hat{Z}_t^A(I_j) \le D_j.$$

Since by the definition of I_j there are at most $2^d(K_2/\varepsilon)^{d/2}$ terms in the sum, (13) tells us that

$$ED_j \le \frac{C}{K_3}\varepsilon^{-1}. \tag{15}$$

Chebyshev's inequality gives

$$P\left(D_j > \frac{K_1}{\varepsilon}\right) \le \frac{ED_j}{K_1/\varepsilon} \ge \frac{C}{K_3}, \tag{16}$$

where C depends on K_1 and K_2. Picking K_3 to make the last quantity less than $\delta/2$ and using Lemma 6.1 proves Lemma 7.1. □

Having shown Lemma 7.1, the next step in our proof is to show that if K_4 is large, then the modified contact process considered in Section 5 does not differ much from the corresponding modified branching random walk. Let $\bar{\xi}_t^A$ denote the modified contact process in which births outside $(-2L^{1/2}, 2L^{1/2}) \times \mathbb{R}^{d-1}$ are not allowed, and suppose $A \subset I_0$ has

(I) $$K_1/\varepsilon \le |A| \le 2K_1/\varepsilon,$$

(II) $$|A \cap B_n| \le 2K_3 h(\varepsilon) \quad \text{for all } n \in \mathbb{Z}^d.$$

If we let $f_L(x)$ = the fraction of x's neighbors which are in $(-2L^{1/2}, 2L^{1/2}) \times \mathbb{R}^{d-1}$, then reasoning as we did for (1.2) gives

$$\begin{aligned}\frac{d}{dt}E|\bar{\xi}_t^A| &= (\lambda - 1)E\sum_x f_L(x)\bar{\xi}_t^A(x) - \frac{\lambda}{v(M)}\sum_x\sum_{y\sim x} P(x, y \in \bar{\xi}_t^A)\\ &\ge (\lambda - 1)E\sum_x f_L(x)\bar{\xi}_t^A(x) - \frac{\lambda}{v(M)}\sum_x\sum_{y\sim x} P(x, y \in Z_t^A)\end{aligned} \tag{17}$$

and

$$\frac{d}{dt}E|\bar{Z}_t^A| = (\lambda - 1)E\sum_x f_L(x)\bar{Z}_t^A(x). \tag{18}$$

If we let $\delta_t^A(x) = \bar{Z}_t^A(x) - \bar{\xi}_t^A(x)$ and subtract, then we have

$$\frac{d}{dt}E|\delta_t^A| \le (\lambda - 1)E|\delta_t^A| + \frac{\lambda}{v(M)}\sum_x\sum_{y\sim x} P(x, y \in Z_t^A), \tag{19}$$

where $|\delta_t^A| = \sum_x \delta_t^A(x)$. To analyze the last equation, we use the following lemma.

Lemma 7.2. *If $f(0) = 0$, $g(t) \ge 0$ and $(d/dt)f(t) \le \varepsilon f(t) + g(t)$, then for $t \le \tau/\varepsilon$,*

$$f(t) \le e^{\tau}\int_0^t g(s)\,ds.$$

PROOF.

$$h(t) = e^{\varepsilon t}\int_0^t e^{-\varepsilon s} g(s)\,ds$$

has $h(0) = 0$ and

$$\frac{d}{dt}h(t) = \varepsilon h(t) + g(t).$$

A standard comparison argument [look at $t_0 = \inf\{t\colon h(t) < f(t)\}$] shows that $h(t) \geq f(t)$. The result follows. □

Lemma 7.2 shows us that $E|\delta_t^A|$ can be estimated by bounding

$$g(t) = \frac{\lambda}{v(M)} \sum_x \sum_{y \sim x} P(x, y \in Z_t^A). \tag{20}$$

To do this, we will observe that the double sum is the number of pairs of particles $x, y \in Z_t^A$ with $0 < \|x - y\| \leq 1$. Now if $x \sim y$, then x, y are both in B_n where $n_i = 1 + \min([x_i], [y_i])$, so

$$g(t) \leq \frac{\lambda}{v(M)} \sum_n E\big(Z_t^A(B_n)^2\big). \tag{21}$$

As in the proof of Lemma 7.1 we begin by writing

$$\begin{aligned}\sum_n E\big(Z_t^x(B_n)^2\big) &= \sum_n m(t, x, B_n) \\ &\quad + \sum_n \int \beta_M(dz) \int_0^t ds \int m(t - s, x, dy) \\ &\quad \times m(s, y, B_n) m(s, y + z, B_n).\end{aligned} \tag{22}$$

The first term is less than or equal to $2^d E Z_t^x$. To estimate the second, we observe that if $t = K_2/\varepsilon$,

$$m(s, y + z, B_n) \leq C(1 + s)^{-d/2} \tag{23}$$

and

$$\sum_n \int m(t - s, x, dy) m(s, y, B_n) \leq 2^d E Z_t^x \leq C. \tag{24}$$

Here and in what follows, C is a constant which depends upon K_1 and K_2 (and later K_3), and will change from line to line. Integrating from 0 to t gives

$$\int_0^t ds(1 + s)^{-d/2} \leq \begin{cases} C(1 + t)^{1/2} & d = 1, \\ C \log(1 + t) & d = 2, \\ C & d \geq 3. \end{cases} \tag{25}$$

Combining (22)–(25) gives for $t = K_2/\varepsilon$ and $\varepsilon \le 1/2$,

$$\sum_n E\big(Z_t^x(B_n)^2\big) \le h(\varepsilon). \tag{26}$$

Continuing to follow the pattern of the proof of Lemma 7.1, we let $A \subset I_0$ be a set with at most $2K_1/\varepsilon$ particles, and write

$$\sum_n E\big(Z_t^A(B_n)^2\big) = \sum_n \sum_{x\in A} E\big(Z_t^x(B_n)^2\big) + \sum_n \sum_{x\in A} \sum_{\substack{y\in A\\ y\neq x}} EZ_t^x(B_n)EZ_t^y(B_n). \tag{27}$$

Using (26), we see that if $t \le K_2/\varepsilon$ and $|A| \le 2K_1/\varepsilon$ the first term on the right-hand side of (27) is smaller than

$$2K_1 h(\varepsilon)/\varepsilon. \tag{28}$$

To bound the second term on the right-hand side of (27), we observe that it is smaller than

$$\sum_n \Big(\sum_{x\in A} EZ_t^x(B_n)\Big)^2 = \sum_n \big(EZ_t^A(B_n)\big)^2 \le \Big(\sup_n EZ_t^A(B_n)\Big)\sum_n EZ_t^A(B_n). \tag{29}$$

For the second term on the right-hand side of (29), we observe that for $t \le K_2/\varepsilon$,

$$\sum_n EZ_t^A(B_n) \le 2^d EZ_t^A = 2^d|A|e^{(k-1)t} \le C/\varepsilon \tag{30}$$

since $|A| \le 2K_1/\varepsilon$. For the first term on the right-hand side of (29) we write (for $t \le K_2/\varepsilon$)

$$\begin{aligned} EZ_t^A(B_n) &= \sum_{x\in A} e^{(\lambda-1)t}P(X_t^x \in B_n) \\ &\le C \sum_{m\in(2Z)^d} |A \cap B_m| \sup_{x\in B_m} P(X_t^x \in B_n). \end{aligned} \tag{31}$$

Now $|A \cap B_m| \le 2K_3 h(\varepsilon)$ by hypothesis II. (This is the only time that assumption is used.) If $x \in B_m$,

$$P(X_t^x \in B_n) = P\big(X_t^0 \in B_n - x\big) \le P\big(X_t^0 \in B_{n-m}^{+1}\big),$$

where $B_k^{+1} = [k_1 - 2, k_1 + 2) \times \cdots \times [k_d - 2, k_d + 2)$ so the last expression in (31) is smaller than

$$C(2K_3 h(\varepsilon)) \sum_{m\in(2\mathbb{Z})^d} P\big(X_t^0 \in B_{n-m}^{+1}\big) \le Ch(\varepsilon) \tag{32}$$

since the sum over m is $\le 2^d$. Combining (32) and (30) gives

$$\sum_n \big(EZ_t^A(B_n)\big)^2 \le Ch(\varepsilon)/\varepsilon, \tag{33}$$

which when put together with (27) and (28) gives

$$\sum_n E\big(Z_t^A(B_n)^2\big) \le Ch(\varepsilon)/\varepsilon. \tag{34}$$

The rest is arithmetic. Using (34) with (20) and (21) gives that if $v(M) \geq K_4 h(\varepsilon)/\varepsilon$ and $\varepsilon \leq 1/2$ (recall $\lambda = 1 + \varepsilon$),

$$(35)\quad g(t) \equiv \frac{\lambda}{v(M)} \sum_x \sum_{y \sim x} P\big(x, y \in Z_t^A\big) \leq \frac{\lambda}{v(M)} \sum_n E\big(Z_t^A(B_n)^2\big) \leq \frac{C}{K_4}.$$

Using (35) and Lemma 7.2 on the differential equation (19), we see that if $t \leq K_2/\varepsilon$,

$$(36)\qquad E|\delta_t^A| \leq \frac{C}{K_4}\varepsilon^{-1}.$$

Using Chebyshev's inequality now gives

$$(37)\qquad P\big(|\delta_t^A| > K_1/\varepsilon\big) < \delta.$$

Combining the last result with Lemma 7.1, shows that if A satisfies assumptions (I) and (II) above, then with probability greater than $1 - 3\delta$, there are sets $A^j \subset \bar{\xi}_t^A \cap I_j$, $j = -1$ and 1, with

(I′) $$|A^j| = K_1/\varepsilon,$$

(II′) $$|A^j \cap B_n| \leq K_3 h(\varepsilon) \quad \text{for all } n \in \mathbb{Z}^d.$$

If $\delta < 1/3 \cdot (81)^2$ and the contact process starts from $A \subset I_0$ satisfying (I) and (II), then Lemma 5.1 implies that the process will survive with probability

$$(38)\qquad \geq 1 - \frac{81(3\delta)}{1 - (3/9)} > 1 - 365\delta > 0.$$

We apologize to the reader for the silly arithmetic. For the proof in Section 9 it will be important that the success probability approaches 1 as $\delta \to 0$.

The arguments above have shown that if

$$(39)\qquad M \geq \begin{cases} C\varepsilon^{-3/2} & d = 1,\\ C\big(\varepsilon^{-1}\log\varepsilon^{-1}\big)^{1/2} & d = 2,\\ C\varepsilon^{-1/d} & d \geq 3,\end{cases}$$

and C is large, then the contact process survives for all time with positive probability. Inverting gives Theorem 1.

8. Proof of Theorem 2. Let $\delta > 0$ and pick K_1, K_2, K_3 and K_4 so that if we use the construction developed in Sections 5–7 and have $v(M) \geq K_4\varepsilon^{-1}h(\varepsilon)$, then the probability of success is at least $1 - \delta$ when $\varepsilon < \varepsilon_0$. If we pick $\varepsilon < \lambda - 1$, then the contact process with parameter λ is larger than the one with parameter $1 + \varepsilon$. To prove Theorem 2 then, it is enough to show

LEMMA 8.1. *If $L = K_2/\varepsilon$ and ε is small, then with a probability ρ_0 close to $(\lambda - 1)/\lambda$, $\xi_{2L}^0 \cap [L^{1/2}, 3L^{1/2}] \times [-L^{1/2}, L^{1/2}]^{d-1}$ will contain a set with*

(i) $|A| = K_1/\varepsilon$ *and*
(ii) $|A \cap B_n| \leq K_3 h(\varepsilon)$ *for all* $n \in \mathbb{Z}^d$,

where

$$B_n = [n_1 - 1, n_1 - 1) \times \cdots \times [n_d - 1, n_d + 1).$$

For this will imply that $P(\Omega_\infty) \geq (1 - \delta)\rho_0$. To prove Lemma 8.1, we will do three things.

(a) Show that if ε is small, then with probability close to $(\lambda - 1)/\lambda$, $Z_L^0 \cap [-L^{1/2}, L^{1/2}]^d$ will contain a set A with $|A| = K_1/\varepsilon$.

(b) Use Lemma 7.1 to conclude that with probability close to $(\lambda - 1)/\lambda$, $Z_{2L}^0 \cap ([L^{1/2}, 3L^{1/2}] \times [-L^{1/2}, L^{1/2}]^{d-1})$ contains a set satisfying (i) and (ii).

(c) Show that if ε is fixed, then $P(|Z_{2L}^0| \neq |\xi_{2L}^0|) \to 0$ as $M \to \infty$.

If the random walk law (i.e., the distribution of X_t^0) were fixed, (a) would follow from the fact that in a supercritical branching random walk where the steps have mean 0 and covariance Σ, then as $t \to \infty$,

$$(1) \qquad e^{-(\lambda-1)t} Z_t\big([-t^{1/2}, t^{1/2}]^d\big) \to WP\big(\chi \in [-1,1]^d\big) \quad \text{in probability},$$

where W is the a.s. limit of $e^{-(\lambda-1)t}|Z_t|$ and χ is a normal random variable with mean 0 and covariance Σ. [See Ney (1965) or Asmussen and Kaplan (1976). For a local limit theorem, see Theorem 4.1 in Durrett (1979).]

Unfortunately, the random walk law is not fixed so we have to reprove (1) to show that the convergence occurs in L^2 uniformly in M. The first step is to introduce

LEMMA 8.2. *For any Borel set B,*

$$\begin{aligned} E\big(e^{-(\lambda-1)t} Z_t^0(B) - e^{-(\lambda-1)t}|Z_t^0| P(X_t^0 \in B)\big)^2 &= \int_0^t du\, e^{-(\lambda-1)u} \int P\big(X_u^0 \in dy\big) \int \beta_M(dz) \\ &\quad \times \big\{P\big(X_{t-u}^0 \in B - y\big) - P\big(X_t^0 \in B\big)\big\} \\ &\quad \times \big\{P\big(X_{t-u}^0 \in B - y - z\big) - P\big(X_t^0 \in B\big)\big\}. \end{aligned}$$

This is a special case of a formula proved in the appendix of Fleischman (1978). To get from Lemma 8.2 to (1), we need to estimate

$$(2) \qquad P\big(X_{t-u}^0 \in H_t - y\big) - P\big(X_t^0 \in H_t\big)$$

for $H_t = [-t^{1/2}, t^{1/2}]^d$, and for most y in the support of X_u^0. If $u = r_t t$ and $y = a_t t^{1/2}$ with $r_t \to r$, $a_t \to a$ and $M_t \to M$ (possibly ∞), then (2) approaches

$$(3) \qquad P\big(B_{1-r}^M \in J_1 - a\big) - P\big(B_1^M \in J_1\big),$$

where $J_1 = [-1, 1]^d$ and B_t^M is a d-dimensional Brownian motion with mean 0 and covariance $\lambda\sigma_M^2 tI$ starting from 0. Here σ_M^2 is the variance defined near (2.6) and σ_∞^2 is the limit of σ_M^2.

Let $\eta > 0$. By Chebyshev's inequality we can pick $K \geq 1$ so that for all M and $u \geq 0$,

$$(4) \qquad P\big(\|X_u^0\| \geq K(u^{1/2} + 1)\big) \leq \eta.$$

Now if $\|y\| \le K(u^{1/2}+2)$ (the extra 1 is included to handle $y+z$ with $\|z\| \le 1$) and $u \le r_0 t$, then

$$\|a\| \equiv \|y/t^{1/2}\| \le K\left(r_0^{1/2} + 2t^{-1/2}\right).$$

If r_0 is fixed, the second term is smaller than the first when t is large. This motivates picking r_0 so that

$$|P\left(B_{1-r}^M \in J_1 - a\right) - P\left(B_1^M \in J_1\right)| \le \eta^{1/2}$$

for all $r \le r_0$, $\|a\| \le 2Kr_0^{1/2}$ and all M. Combining the last fact with the convergence of $X_{st}^0/t^{1/2}$ to B_s^M, we see that if $t \ge t_0$,

$$(5)\qquad |P\left(X_{t-u}^0 \in H_t - y\right) - P\left(X_t^0 \in H_t\right)| \le 2\eta^{1/2}$$

for $u \le r_0 t$ and y with $\|y\| \le K(u^{1/2}+2)$.

REMARK. To prepare for developments in the next section, we would like the reader to observe that the argument above works if $H_t = [L^{1/2}, 3L^{1/2}] \times [-L^{1/2}, L^{1/2}]^{d-1}$ and $J_1 = [1,3] \times [-1,1]^{d-1}$.

Using (4) and (5), we can estimate the integrals in Lemma 8.2. There are two cases.

CASE 1. $u \ge r_0 t$. Estimating the two differences of probabilities by 1 and integrating gives an upper bound

$$(6)\qquad \lambda \int_{r_0 t}^{t} e^{-(\lambda-1)u}\,du \le \frac{\lambda}{\lambda-1} e^{-(\lambda-1)r_0 t},$$

the second bound resulting from replacing the upper limit by ∞ in the integral.

CASE 2. $u \le r_0 t$. We use (5) for y with $\|y\| \le K(u^{1/2}+1)$ and estimate the two differences by 1 when $\|y\| \ge K(u^{1/2}+1)$. Using (4), it follows that the integral is smaller than

$$(7)\qquad \lambda \int_0^{r_0 t} e^{-(\lambda-1)u} 5\eta\,du \le \eta \frac{5\lambda}{\lambda-1}.$$

Since $\eta > 0$ is arbitrary it follows from (6) and (7) that

$$(8)\qquad E\left(e^{-(\lambda-1)t} Z_t^0(A_t) - e^{-(\lambda-1)t}|Z_t^0| P\left(X_t^0 \in A_t\right)\right)^2 \to 0,$$

uniformly in M as $t \to \infty$. Since

$$(9)\qquad e^{-(\lambda-1)t}|Z_t^0| \to W \quad \text{a.s.},$$

$$(10)\qquad P(W>0) = P\left(|Z_t^0| > 0 \text{ for all } t\right) = \frac{\lambda-1}{\lambda}$$

and

$$(11)\qquad P\left(X_t^0 \in H_t\right) \to \text{a positive limit},$$

it follows that for any $M < \infty$,

(12) $$P\big(Z_t^0(H_t) \ge M\big) \to (\lambda - 1)/\lambda \quad \text{as } t \to \infty.$$

(11) is the result promised in (a). The last detail then is to prove (c), but this is easy. With ε fixed

(13) $$P\Big(\sup_{0 \le t \le 2K_1/\varepsilon} |Z_t^0| > M^{1/3}\Big) \to 0$$

as $M \to \infty$ (compare with a branching process with no deaths). If $|Z_t^0| \le M^{1/3}$ for all $0 \le t \le 2K_1/\varepsilon$, then even if all the particles were in $[0,1]^d$, the rate at which collisions occur (i.e., births from an occupied site which land on an occupied site) would be smaller than $M^{1/3}M^{1/3}/v(M)$. It follows from this that

(14) $$\begin{aligned} &P\big(|Z_t^0| \ne |\xi_t^0| \text{ for some } t \le 2K_1/\varepsilon\big) \\ &\quad \le 2K_1 M^{2/3}/\big(\varepsilon v(M)\big) \to 0 \quad \text{as } M \to \infty. \end{aligned}$$

[Recall $v(M) \approx M^d$.]

9. Proof of Theorem 3. In this section we return to the setup of Sections 5–7. $\lambda = 1 + \varepsilon$ and K_1–K_3 are the constants from the construction. (They will have to be adjusted below.) Let $L = K_2/\varepsilon$ and let

$$I_j = \big[(2j-1)L^{1/2}, (2j+1)L^{1/2}\big] \times \big[-L^{1/2}, L^{1/2}\big]^{d-1}.$$

As in Section 8, we want to show that if $\delta > 0$ and ε is small, then with a probability greater than or equal to $(1 - 5\delta)\varepsilon/(1+\varepsilon)$, $\xi_L^0 \cap I_1$ will contain a set with

(i) $|A| = K_1/\varepsilon$,
(ii) $|A \cap B_n| \le K_3 h(\varepsilon)$ for all $n \in \mathbb{Z}^d$,

where $B_n = [n_1 - 1, n_1 + 1) \times \cdots \times [n_d - 1, n_d + 1)$. The key to doing this is to prove

Lemma 9.1. *If $\delta > 0$ and K_1 are given, then we can pick K_2 and ε_0 so that for $\varepsilon < \varepsilon_0$,*

$$P\big(Z_L^0(I_1) \ge 3K_1/\varepsilon\big) \ge (1 - 3\delta)\varepsilon/(1+\varepsilon).$$

The last result should remind the reader of Lemma 6.1. There are two differences. (a) We are starting with one particle, so we cannot make the error probabilities small by making K_1 large. (b) We do not need to truncate the branching process.

Once Lemma 9.1 is proved, Theorem 3 follows easily using the methods of Section 7, so we give that part of the proof now. Pick K_1 and K_2 so that Lemma 6.1 can be applied and then increase K_2 so that Lemma 9.1 applies. From (7.8) we get

(1) $$E\big(Z_L^0(B_n)^2\big) \le \begin{cases} C & d = 1, \\ C\varepsilon\log(\varepsilon^{-1}) & d = 2, \\ C\varepsilon^{d/2} & d \ge 3. \end{cases}$$

So if we let $V_n = Z_L^0(B_n)$, and

$$h(\varepsilon) = \begin{cases} \varepsilon^{-1/2} & d = 1, \\ \log(\varepsilon^{-1}) & d = 2, \\ 1 & d \geq 3, \end{cases} \tag{2}$$

then

$$E(V_n;\, V_n > K_3 h(\varepsilon)) \leq \frac{EV_n^2}{K_3 h(\varepsilon)} \leq \frac{C}{K_3}\varepsilon^{d/2}. \tag{3}$$

If we let

$$D = \sum_n V_n 1_{(V_n > K_3 h(\varepsilon))},$$

where the sum is over $n \in I_1$, then it follows from (3) that

$$ED \leq \frac{C}{K_3}, \tag{4}$$

and Chebyshev's inequality gives

$$P\left(D > \frac{K_1}{\varepsilon}\right) \leq \frac{ED}{K_1/\varepsilon} \leq \frac{C}{K_3}\varepsilon. \tag{5}$$

Again, the next step is to thin Z_L^0 so that there are at most $K_3 h(\varepsilon)$ particles in any B_n, and call the thinned process $\hat{Z}_L^0$. If K_3 is large, then (5) implies

$$P\big(\hat{Z}_L^0(I_1) > 2K_1/\varepsilon\big) > (1 - 4\delta)\varepsilon/(1 + \varepsilon). \tag{6}$$

By increasing K_3, we can guarantee the conclusion of Lemma 7.1 also holds.

To estimate the difference between the contact process and the branching process, we write $|\delta_t^0| = |Z_t^0| - |\xi_t^0|$, and follow the derivation of (7.19) to get

$$\frac{d}{dt}E|\delta_t^0| \leq (\lambda - 1)E|\delta_t^0| + \frac{\lambda}{v(M)}\sum_x \sum_{y \sim x} P(x, y \in Z_t^0). \tag{7}$$

Using the reasoning which led to (7.26) now gives for $\varepsilon \leq 1/2$,

$$\sum_x \sum_{y \sim x} P(x, y \in Z_t^0) \leq \sum_n E\big(Z_t^0(B_n)^2\big) \leq Ch(\varepsilon). \tag{8}$$

The last result implies that if $v(M) \geq K_4 h(\varepsilon)/\varepsilon$, then for $t \leq K_2/\varepsilon$,

$$\frac{\lambda}{v(M)}\sum_x \sum_{y \sim x} P(x, y \in Z_t^0) \leq \frac{C}{K_4}\varepsilon. \tag{9}$$

Using (9) and Lemma 7.2 on (7), we see that if $t \leq K_2/\varepsilon$ and $v(M) \geq K_4\varepsilon^{-1}h(\varepsilon)$, then

$$E|\delta_t^0| \leq \frac{C}{K_4}. \tag{10}$$

Using Chebyshev's inequality now gives that if K_4 is large

$$P\big(|\delta_t^0| > K_1/\varepsilon\big) < \delta\varepsilon. \tag{11}$$

By picking K_4 larger, we can guarantee that if the contact process starts from $A \subset I_1$ satisfying (i) and (ii), then it survives with probability greater than $1 - C\delta$ [see (7.38)]. Combining (11) with (6) shows that with probability greater than $(1 - 5\delta)\varepsilon/(1 + \varepsilon)$ there is a set $A \subset \xi_L^0 \cap I_1$ with properties (I) and (II). This implies $P(\Omega_\infty) \geq (1 - C'\delta)\varepsilon/(1 + \varepsilon)$ and the proof of Theorem 3 is complete.

The key to the proof of Lemma 9.1 is the following result.

LEMMA 9.2. *Let $\lambda = (1 + \varepsilon)$ and suppose $\varepsilon \leq 1$. There is a constant $\delta(\tau)$ [independent of M and having $\delta(\tau) \to 0$ as $\tau \to \infty$], so that if $t = \tau/\varepsilon$ and $I_1 = [t^{1/2}, 3t^{1/2}] \times [-t^{1/2}, t^{1/2}]^{d-1}$, then*

$$E\left(e^{-(\lambda-1)t}Z_t^0(I_1) - e^{-(\lambda-1)t}|Z_t^0|P\left(X_t^0 \in I_1\right)\right)^2 \leq \frac{\delta(\tau)}{\varepsilon}.$$

REMARKS. (i) To see that the bound is the right order of magnitude, observe that Lemma 2.1 implies that $|Z_{\tau/\varepsilon}^0| > 0$ with probability approximately equal to ε, and

$$E\left(|Z_{\tau/\varepsilon}^0| \,\big|\, |Z_{\tau/\varepsilon}^0| > 0\right) \approx e^{\tau}/\varepsilon,$$

so (ignoring a few details)

$$E\left(e^{-(\lambda-1)\tau/\varepsilon}|Z_{\tau/\varepsilon}^0|\right)^2 \approx 1/\varepsilon.$$

(ii) The fact that $\delta(\tau) \to 0$ as $\tau \to \infty$ says that if τ is large and $|Z_{\tau/\varepsilon}^0| > 0$, then the particles are distributed in space as the central limit theorem dictates. Here, we have only proved the result for the one set we are interested in but the proof below generalizes easily.

To prove Lemma 9.2, we use Lemma 8.2 and apply (8.4) and (8.5) to estimate the terms on the right-hand side. As in Section 8, there are two cases.

CASE 1. $u \geq r_0 t$. Estimating the two differences of probabilities by 1 and integrating gives an upper bound

$$\lambda \int_{r_0 t}^{t} e^{-\varepsilon u}\, du \leq 2e^{-r_0\tau}/\varepsilon, \tag{12}$$

the second bound resulting from replacing the upper limit by ∞ in the integral.

CASE 2. $u \leq r_0 t$. We use (8.5) for y with $\|y\| \leq K(u^{1/2} + 1)$ and estimate the two differences by 1 when $\|y\| \geq K(u^{1/2} + 1)$. Using (8.4), it follows that the integral is smaller than

$$\lambda \int_0^\infty e^{-\varepsilon u} 5\eta\, du \leq 10\eta/\varepsilon. \tag{13}$$

Combining the bounds from the two cases gives

$$(2\exp(-r_0\tau) + 10\eta)/\varepsilon. \tag{14}$$

If we pick η so that $10\eta < \delta/2$, then for τ large, the upper bound is smaller than δ/ε proving Lemma 9.2.

We turn now to the proof of Lemma 9.1. (6.6) implies that there is an $\eta > 0$ so that if $H_n = [n^{1/2}, 3n^{1/2}] \times [-n^{1/2}, n^{1/2}]^{d-1}$ and $n \geq n_0$, then

$$P(X_n^0 \in H_n) \geq \eta > 0. \tag{15}$$

The next step is to apply Lemma 2.1 to the branching process with offspring distribution $p_k = P(|Z_1^0| = k)$. If we do this, then (in the notation of Lemma 2.1) $\mu = e^{\varepsilon}$, ν depends on ε but converges to a positive finite limit α as $\varepsilon \to 0$ and

$$c_n = \nu \frac{(1 - \mu^{-n})}{(\mu - 1)} = \nu \frac{1 - e^{-\varepsilon n}}{e^{\varepsilon} - 1} \tag{16}$$

since $\mu > 1$.

Writing Z_n for $|Z_n^0|$ to simplify notation and applying Lemma 2.1 now gives that as $n \to \infty$,

$$P(Z_n/c_n\mu^n > x | Z_n > 0) \to e^{-x}. \tag{17}$$

Replacing ν by α and $e^{\varepsilon} - 1$ by ε in the denominator of $c_n\mu^n$, the last conclusion can be written as

$$P\left(Z_n \frac{\varepsilon}{\alpha(e^{\varepsilon n} - 1)} > x | Z_n > 0\right) \to e^{-x}. \tag{18}$$

We want to combine the last result with Lemma 9.2 to prove Lemma 9.1. To facilitate this, we begin by observing that if $\rho < \eta$ (15) implies

$$\begin{aligned} &P\left(e^{-\varepsilon n} Z_n P(X_n^0 \in H_n) > \frac{2\rho^2}{\varepsilon} \middle| Z_n > 0\right) \\ &\quad \geq P\left(Z_n \frac{\varepsilon}{\alpha(e^{\varepsilon n} - 1)} > 2\rho \frac{e^{\varepsilon n}}{\alpha(e^{\varepsilon n} - 1)} \middle| Z_n > 0\right) \\ &\quad \geq P\left(Z_n \frac{\varepsilon}{\alpha(e^{\varepsilon n} - 1)} > \frac{4\rho}{\alpha} \middle| Z_n > 0\right) \end{aligned} \tag{19}$$

if $n \geq (\ln 2)/\varepsilon$. (So later we have to pick $K_2 \geq \ln 2$.) From (18), the last quantity approaches $e^{-4\rho/\alpha}$ as $n \to \infty$ and $\varepsilon \to 0$. If we pick $\rho < \eta$ so that $e^{-4\rho/\alpha} > 1 - \delta$, then we can pick ε_1, n_1 so that for $n \geq n_1$ and $\varepsilon < \varepsilon_1$,

$$P\left(e^{-\varepsilon n} Z_n P(X_n^0 \in H_n) > \frac{2\rho^2}{\varepsilon} \middle| Z_n > 0\right) \geq 1 - 2\delta. \tag{20}$$

Let

$$\Delta_n = e^{-\varepsilon n} Z_n^0(I_1) - e^{-\varepsilon n} |Z_n^0| P(X_n^0 \in I_1).$$

$\Delta_n = 0$ on $\{|Z_n^0| > 0\}$ and

$$P(|Z_n^0| > 0) \geq \varepsilon/(1 + \varepsilon) \geq \varepsilon/2 \tag{21}$$

for $\varepsilon \leq 1$. So if we let $n = L$ (which we assume is an integer) and recall Z_n is an abbreviation for $|Z_n^0|$, it follows from Lemma 9.2 that

$$(22) \qquad E\left(\Delta_L^2 \mid |Z_L^0| > 0\right) = \frac{E(\Delta_L^2)}{P(|Z_L^0| > 0)} \leq \frac{\delta(K_2)}{\varepsilon} \frac{2}{\varepsilon}.$$

Using Chebyshev's inequality now gives

$$(23) \qquad P\left(\Delta_L < -\rho^2/\varepsilon \mid |Z_L^0| > 0\right) \leq 2\delta(K_2)/\rho^4.$$

Pick $K_2 > \ln 2$ [for the last computation in (19)] and large enough so that

$$(24) \qquad 2\delta(K_2)/\rho^4 < \delta,$$

$$(25) \qquad \rho^2 \exp(K_2) > 3K_1.$$

With this choice of K_2, (20) and (23) give (observe that $H_L = I_1$)

$$(26) \qquad P\left(Z_L^0(I_1) > 3K_1/\varepsilon \mid |Z_L^0| > 0\right) \geq (1 - 3\delta),$$

which is the desired result since

$$P\left(|Z_L^0| > 0\right) \geq \frac{\varepsilon}{1 + \varepsilon} = \frac{\lambda - 1}{\lambda}.$$

Acknowledgments. The second author is responsible for any abuses of the English language which occur in this article. He would like to thank H. Kesten for several useful conversations about random walks and thank M. E. Fisher for explaining why Theorem 2 does not contradict the universality of the exponent β. Fisher's remark and his claim that "no one has ever rigorously computed a crossover exponent" were the motivation for Theorem 3.

REFERENCES

ASMUSSEN, S. and KAPLAN, N. (1976). Branching random walks. I. *Stochastic Process. Appl.* 4 1–13.

ATHREYA, K. B. and NEY, P. E. (1972). *Branching Processes*. Springer, New York.

BILLINGSLEY, P. (1978). *Probability and Measure*. Wiley, New York.

BREIMAN, L. (1968). *Probability*. Addison-Wesley, Reading, Mass.

CHUNG, K. L. (1974). *A Course in Probability Theory*, 2nd ed. Academic, New York.

DURRETT, R. (1979). An infinite particle system with additive interactions. *Adv. in Appl. Probab.* **11** 355–383.

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

DURRETT, R. (1988). *Lecture Notes on Particle Systems and Percolation*. Wadsworth, Belmont, Calif.

DURRETT, R. and SCHONMANN, R. H. (1987). Stochastic growth models. In *Percolation Theory and Ergodic Theory of Infinite Particle Systems. IMA Volumes in Mathematics and Its Applications* (H. Kesten, ed.) **8** 85–119. Springer, New York.

FLEISCHMAN, J. (1978). Limiting distributions for branching random fields. *Trans. Amer. Math. Soc.* **239** 353–390.

GRIFFEATH, D. (1983). The binary contact path process. *Ann. Probab.* **11** 692–705.

HOLLEY, R. and LIGGETT, T. (1981). Generalized potlatch and smoothing processes. *Z. Wahrsch. verw. Gebiete* **55** 165–195.

JAGERS, P. (1975). *Branching Processes with Biological Applications*. Wiley, New York.

KESTEN, H. (1969). A sharper form of the Doeblin–Lévy–Kolmogorov–Rogozin inequality for concentration functions. *Math. Scand.* **25** 133–144.

LIGGETT, T. (1985). *Interacting Particle Systems*. Springer, New York.

NEY, P. E. (1965). The convergence of a random distribution function associated with a branching process. *J. Math. Anal. Appl.* **12** 316–327.

SAWYER, S. (1976). Branching diffusion processes in population genetics. *Adv. in Appl. Probab.* **8** 659–689.

SCHONMANN, R. H. and VARES, M. E. (1986). The survival of the large dimensional basic contact process. *Probab. Theory Related Fields* **72** 387–393.

THOULESS, D. J. (1969). Critical region for the Ising model with a long range interaction. *Phys. Rev.* **181** 954–968.

M. BRAMSON
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN
MADISON, WISCONSIN 53706

R. DURRETT
DEPARTMENT OF MATHEMATICS
WHITE HALL
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853

G. SWINDLE
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
LOS ANGELES, CALIFORNIA 90024