The Annals of Probability
1988, Vol. 16, No. 3, 1158–1173

THE CONTACT PROCESS ON A FINITE SET

BY RICHARD DURRETT[1] AND XIU-FANG LIU[2]

Cornell University and Beijing Normal University

In this paper we show that the phase transition in the contact process manifests itself in the behavior of large finite systems. To be precise, if we let σ_N denote the time the process on $\{1,\dots,N\}$ first hits $\varnothing$ starting from all sites occupied, then there is a critical value λ_c so that (i) for $\lambda < \lambda_c$ there is a constant $\gamma(\lambda) \in (0,\infty)$ so that as $N \to \infty$, $\sigma_n/\log N \to 1/\gamma(\lambda)$ in probability and (ii) for $\lambda > \lambda_c$ there are constants $\alpha(\lambda), \beta(\lambda) \in (0,\infty)$ so that as $N \to \infty$,

$$P(\alpha(\lambda)/2 - \varepsilon \le (\log \sigma_N)/N \le \beta(\lambda) + \varepsilon) \to 1,$$

for all $\varepsilon > 0$. Our results improve upon an earlier work of Griffeath but as the reader can see the second one still needs improvement. To help decide what should be true for the contact process we also consider the analogous problem for the biased voter model. For this process we can show $(\log \sigma_N)/N \to \alpha(\lambda) = \beta(\lambda)$ in probability, and it seems likely that the same result is true for the contact process.

1. Introduction. Let ζ_t^N denote the basic contact process on $\{1,2,\dots,N\}$ starting from all sites occupied. That is, ζ_t^N is the Markov chain with state space = the set of all subsets of $\{1,\dots,N\}$ and transition rates (or q-matrix) given by

$$q(A, A - \{x\}) = 1, \qquad \text{if } x \in A,$$

$$q(A, A \cup \{x\}) = \lambda|A \cap \{x-1, x+1\}|, \quad \text{if } x \notin A,$$

where $|S|$ denotes the cardinality of S. ζ_t^N is a Markov chain with a finite state space and an absorbing state (the empty set) so at first glance it may seem there is nothing interesting to say about its limiting behavior: If $\sigma_N = \inf\{t: \zeta_t^N = \varnothing\}$, then $P(\sigma_N < \infty) = 1$ for all λ.

More interesting behavior, i.e., a "phase transition," appears if we consider what happens when $N \to \infty$. To state our results we have to introduce ξ_t, the contact process on Z. It is the Markov process with transition probabilities which as $t \to 0$ satisfy

$$P(x \notin \xi_t | \xi_0) \sim t, \qquad \text{if } x \in \xi_0,$$

$$P(x \in \xi_t | \xi_0) \sim \lambda t |\xi_0 \cap \{x-1, x+1\}|, \quad \text{if } x \notin \xi_0,$$

where $f(t) \sim g(t)$ means $f(t)/g(t) \to 1$ as $t \to 0$. It is by now well known that there is a unique Markov process with the properties given above and there are

Received July 1986.

[1]Partially supported by an NSF grant and the Army Research Office through the Mathematical Sciences Institute at Cornell University.

[2]This work was done while this author was visiting Cornell University and supported by the Chinese government.

AMS 1980 *subject classification.* 60K35.

Key words and phrases. Contact process, biased voter model.

several ways to construct it. See Liggett (1985), Chapter 6 [or Griffeath (1981)], for information on how to construct the process and the basic properties we use below if no reference is given.

Let ξ_t^0 denote the contact process on Z starting from $\xi_0^0 = \{0\}$ and λ_c be its critical value,

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

With this notation introduced we can state our first result.

THEOREM 1. *Let* $\lambda < \lambda_c$ *and*

$$\gamma_1(\lambda) = -\lim_{n\to\infty} \frac{1}{n} \log P(\xi_n^0 \neq \varnothing).$$

As $N \to \infty$,

$$\frac{\sigma_N}{\log N} \to \frac{1}{\gamma_1(\lambda)}$$

in probability.

The results of Griffeath (1981) show

$$\limsup_{n\to\infty} \frac{1}{n} \log P(\xi_n^0 \neq \varnothing) < 0,$$

so $\gamma_1(\lambda) > 0$. [$\gamma_1(\lambda)$ is finite since it is $\leq \gamma_1(0) = 1$.] The existence of the limit above was first observed in Durrett (1984). For completeness we give the simple proof of the existence of the limit in Section 2. As the reader can see from the proofs given there, the only hard part is guessing the constant to which $\sigma_N/(\log N)$ should converge (and proving it is > 0). Failure to find the right definition for $\lambda > \lambda_c$ probably explains the incompleteness of the next result.

THEOREM 2. *Let* $\lambda > \lambda_c$ *and let* $-\alpha_2(\lambda)$ *and* $-\beta_2(\lambda)$ *denote the* lim sup *and* lim inf *of*

$$\frac{1}{n} \log P(\hat{\tau}^{\{1,\dots,n\}} < \infty),$$

where $\hat{\tau}^A = \inf\{t: \hat{\xi}_t^A = \varnothing\}$ *and* $\hat{\xi}_t^A$ *is the contact process on the half line* $\{1,2,\dots\}$ *starting from* $\hat{\xi}_0^A = A$. *If* $\varepsilon > 0$, *then as* $N \to \infty$,

$$P\left(\frac{\log \sigma_N}{N} > \beta_2(\lambda) + \varepsilon\right) \to 0,$$

$$P\left(\frac{\log \sigma_N}{N} < \frac{\alpha_2(\lambda)}{2} - \varepsilon\right) \to 0.$$

In words our results say σ_N grows like $C \log N$ for $\lambda < \lambda_c$ and $\exp(CN)$ for $\lambda > \lambda_c$. [In Section 3 we will show $0 < \alpha_2(\lambda) \leq \beta_2(\lambda) < \infty$.] Our theorems sharpen a previous result of Griffeath [(1981), see Theorem 13, page 177 and the remark on page 183 for the extension to all $\lambda > \lambda_c$] and an earlier work of Stavskaya and Piatetski-Shapiro (1968) and Toom (1968). They did not prove the existence of the limit in Theorem 1 and had worse upper and lower bounds in Theorem 2. Our second result can also be improved. The first and most obvious problem is to prove that $\alpha_2(\lambda) = \beta_2(\lambda)$, i.e., the limit exists. After that there is the annoying fact that the upper and lower bounds differ by a factor of 2.

Strong evidence for the existence of a limit for $\log(\sigma_N)/N$ can be found in Schonmann (1985). He proved, after preliminary results of Cassandro, Galves, Oliveri and Vares (1984), that as $N \to \infty$, $\sigma_N/E\sigma_N$ converges in distribution to a mean one exponential. If we let $c_N = (1/N)\log E\sigma_N$ this implies

$$\left(\frac{1}{N}\log \bar{\sigma}_N\right) - c_N \to 0, \quad \text{in probability,}$$

so as $N \to \infty$, the distribution of $(\log \sigma_N)/N$ becomes concentrated at the point which, of course, may be wandering around in $[\alpha_2(\lambda)/2, \beta_2(\lambda)]$.

As the reader can guess from our notation we believe that the upper bound in Theorem 2 is correct. (This turns out to be true. See the Epilogue at the end of the paper.) To obtain some insight into this question we investigated the analogous problems for the biased voter model on $\{1, \dots, N\}$. Using ζ_t^N again to denote this process it may be described precisely as the Markov chain with state space = the set of all subsets of $\{1, \dots, N\}$ and transition rates (or q-matrix) given by

$$q(A, A - \{x\}) = |A^c \cap \{x-1, x+1\}|, \quad \text{if } x \in A,$$
$$q(A, A \cup \{x\}) = \lambda|A \cap \{x-1, x+1\}|, \quad \text{if } x \notin A.$$

While the formulas for the transition rates of this process are a little more complicated than those for the contact process, the biased voter model is much simpler. An occupied site with two occupied neighbors cannot die and it is easy to check that if we let L_t be a continuous time random walk on $\{1, 2, \dots\}$ which starts at 1 and makes transition $x \to x + 1$ at rate 1 and $x \to x - 1$ at rate λ (when $x > 1$) and let R_t^N be an independent process on $\{N, N - 1, \dots\}$, which has the same distribution as $(N + 1) - L_t$, then a realization of ζ_t^N can be constructed by setting

$$\zeta_t^N = \{L_t, \dots, R_t^N\}, \quad \text{for } t < \sigma_N,$$
$$= \varnothing, \quad \text{for } t \ge \sigma_N,$$

where $\sigma_N = \inf\{t\colon R_t^N < L_t\}$.

With the last construction in hand it is easy to prove

THEOREM 3. *If* $\lambda < 1$, *then as* $N \to \infty$,

$$\frac{1}{N}\bar{\sigma}_N \to \frac{1}{2}(1 - \lambda), \quad \textit{in probability.}$$

So we leave this as an exercise for the reader and turn our attention to the behavior for $\lambda > 1$. If we let l_t be a continuous time random walk on Z which starts at 1 and makes transitions $x \to x + 1$ at rate 1, $x \to x - 1$ at rate λ, and let r_t^N be an independent process which has the same distribution as $(N + 1) - l_t$, then we can construct the biased voter model on Z starting from $\{1, \dots, N\}$ by setting

$$\xi_t^{\{1,\dots,N\}} = \{l_t, \dots, r_t^N\}, \quad \text{for } t < \tau_N,$$
$$= \varnothing, \qquad \text{for } t \geq \tau_N,$$

where $\tau_N = \inf\{t: r_t^N < l_t\}$.

Since $S(t) \equiv r_t^N - l_t$ is a continuous time random walk it follows from the gambler's ruin formula (or the observation that $\lambda^{S(t)}$ is a martingale) that

$$P(\tau_N < \infty) = \lambda^{-N},$$

for $\lambda > 1$ so $\lambda_c = 1$. From the last result it is immediate that if we let $\tau^A = \inf\{t: \xi_t^A \neq \varnothing\}$, then

$$\lim_{n \to \infty} \frac{1}{n} \log P(\tau^{\{1,\dots,n\}} < \infty) = -\log \lambda.$$

The results in the last paragraph fix (for the biased voter model) one of the problems we had for the contact process. In Section 5 we show how to fix the other one and prove

THEOREM 4. *If $\lambda > 1$, then as $N \to \infty$,*

$$\frac{1}{N} \log \sigma_N \to \log \lambda, \quad \textit{in probability}.$$

Ironically, we have to work harder here to estimate $P(L_t > \theta N)$ than we do for the contact process, but once we have the desired estimate the fact that L_t and R_t^N are independent allows us to multiply our estimates and avoid the factor of 2 in Theorem 2.

The rest of the paper is devoted to the proofs of the results given above. To prepare the reader for this we need to explain the method behind our notation: ξ_t^A denotes the contact process (or biased voter model) on Z. When a $\hat{}$ is put on top we are considering the process restricted to $\{1, 2, \dots\}$ and when a $\tilde{}$ is put there (see the end of Section 3) the process is on $\{N, N-1, \dots\}$. Similar labeling conventions apply to the extinction times τ and the left and right edge processes l_t and r_t^N, with the exception that in the discussion of the biased voter model the processes which should be called $\hat{l}_t$ and $\tilde{r}_t^N$ are called L_t and R_t^N.

2. Proof of the results for $\lambda < \lambda_c$. We begin with the upper bound on σ_N. Let ξ_t^0 be the contact process on Z starting from $\xi_0^0 = \{0\}$. When $\xi_t^0 \neq \varnothing$, $|\xi_t^0| \geq 1$, and starting with more particles increases the survival probability of the contact process, so we have

$$P(\xi_{t+s}^0 \neq \varnothing \mid \xi_t^0 \neq \varnothing) \geq P(\xi_s^0 \neq \varnothing),$$

or rearranging terms,

$$P(\xi_{t+s}^0 \neq \varnothing) \geq P(\xi_t^0 \neq \varnothing) P(\xi_s^0 \neq \varnothing).$$

Taking logs and letting $a_t = \log P(\xi_t^0 \neq \varnothing)$, the last equation becomes

$$a_{t+s} \geq a_t + a_s,$$

i.e., $t \to a_t$ is superadditive. It is well known and easy to prove that this implies

$$\frac{1}{t} a_t \to \sup_{u>0} \frac{a_u}{u},$$

so if we let

$$\gamma_1(\lambda) = -\sup_{t>0} \frac{1}{t} \log P(\xi_t^0 \neq \varnothing),$$

it follows that

$$(1) \qquad \frac{1}{t} \log P(\xi_t^0 \neq \varnothing) \to -\gamma_1(\lambda)$$

and

$$(2) \qquad P(\xi_t^0 \neq \varnothing) \leq e^{-\gamma_1(\lambda)t}.$$

The reader should note that up to this point everything is valid for any λ. The fact that $\gamma_1(\lambda) > 0$ when $\lambda < \lambda_c$ is due to Griffeath (1981), Section 5. A proof can also be found in Durrett (1984).

The proof of the upper bound is an immediate consequence of (2).

$$P(\zeta_t^N \neq \varnothing) \leq NP(\xi_t^0 \neq \varnothing) \leq Ne^{-\gamma_1(\lambda)t},$$

so if we let $t = \log N/\gamma_1(\lambda) + K_N$, then

$$P(\zeta_t^N \neq \varnothing) \leq e^{-\gamma_1(\lambda)K_N} \to 0,$$

whenever $K_N \to \infty$.

The proof of the lower bound only requires a little more work. (1) implies that if t is large

$$P(\xi_t^0 \neq \varnothing) \geq e^{-(1+\varepsilon)\gamma_1(\lambda)t}.$$

If we let $t = (1-\varepsilon)\log N/\gamma_1(\lambda)$, then the probability of survival for each particle is $\geq N^{-(1-\varepsilon^2)}$, so to guarantee that at least one survives we need to give ourselves enough independent changes. To do this we consider the contact process on $\{1, \dots, N\}$ with a particle starting at each point of the form $(2k-1)C\log N$ for $1 \leq k \leq [N/(2C\log N)]$ and modified so that the points $2kC\log N$ are never allowed to become occupied. (Here C is a constant which depends upon λ and N and will be chosen later to be large enough and so that $C\log N$ is an integer.)

The last rule makes the behavior of the process in the intervals $(2kC\log N, (2k+2)C\log N)$ independent. The next computation shows that if C is large the restriction does not effect the growth of the individual processes very much. Let $r_t^0 = \sup \xi_t^0$. If we consider the contact process with no deaths, then the position of the right edge r_t^0 has a Poisson distribution with mean λt so if we let $S(\lambda t)$ be a random variable with that distribution

$$P\Big(\sup_{s\le t} r_s^0 \ge m\Big) \le P(S(\lambda t) \ge m).$$

Now

$$E(\exp(\theta S(\lambda t))) = \exp(\lambda t(e^\theta - 1))$$

and Chebyshev's inequality implies

$$P(S(\lambda t) > K\lambda t) \le \exp(\lambda t(e^\theta - 1 - \theta K)).$$

So if we let $\theta = 1$ and observe $e - 1 < 2$ the right-hand side becomes $\exp(\lambda t(2 - K))$. As we mentioned above we are interested in what happens when $t = (1-\varepsilon)\log N/\gamma_1(\lambda)$ and we can assume without loss of generality that $\varepsilon < \frac{1}{2}$, so if we let $K = 2 + (4\gamma_1(\lambda)/\lambda)$ and work back through the computations above we see that for this choice of K and t,

$$P\Big(\sup_{s\le t} r_s^0 \ge K(1-\varepsilon)\lambda \log N/\gamma_1(\lambda)\Big) \le N^{-2}.$$

Now if the process starting at $(2k-1)C\log N$ does not escape from $((2k-2)C\log N, 2kC\log N)$ by time t, then it evolves just like $\xi_s^0 0 \le s \le t$ so letting $C = K\lambda/\gamma_1(\lambda)$ (and enlarging C slightly to make $C\log N$ an integer) we see that if $t = (1-\varepsilon)\log N/\gamma_1(\lambda)$ and N is large

$$P(\zeta_t^N = \varnothing) \le \left(1 - N^{-(1-\varepsilon^2)} - 2N^{-2}\right)^{[N/2C\log N]},$$

which $\to 0$ as $N\to\infty$ because of the simply proven fact that if $\varepsilon_n \to 0$ and $N_n \to \infty$ in such a way that $N_n\varepsilon_n \to \infty$, then

$$(1-\varepsilon_n)^{N_n} \to 0.$$

Remark. The reader should take a moment to observe that the arguments above are very general. The existence of

$$\lim_{t\to\infty} \frac{1}{t}\log P(\xi_t^0 \ne \varnothing) = -\gamma_1(\lambda)$$

and the bound

$$P(\xi_t^0 \ne \varnothing) \le e^{-\gamma_1(\lambda)t}$$

follows form monotonicity and the Markov property, and the construction in the last part of the proof is valid for any interaction with finite range.

3. Proof of the results for $\lambda > \lambda_c$. We begin by recalling some facts about the contact process.

(1) As $n\to\infty$,

$$\frac{1}{n}\log P(\tau^{\{1,\dots,n\}} < \infty) \to -\gamma_2(p).$$

Proof. $\tau^{\{1,\ldots,n\}} < \infty$ is a decreasing event (fewer births in the graphical representation makes it more likely) so Harris' (1960) inequality implies

$$(*)\qquad \begin{aligned} &P\big(\tau^{\{1,\ldots,m\}} < \infty, \tau^{\{m+1,\ldots,m+n\}} < \infty\big) \\ &\quad \ge P\big(\tau^{\{1,\ldots,m\}} < \infty\big)P\big(\tau^{\{m+1,\ldots,m+n\}} < \infty\big) \\ &\quad = P\big(\tau^{\{1,\ldots,m\}} < \infty\big)P\big(\tau^{\{1,\ldots,n\}} < \infty\big). \end{aligned}$$

[Sticklers for details should notice that although Harris' theorem cannot be applied directly to the graphical representation, it can, however, be applied to the oriented percolation process which has connections from $(x, n\varepsilon)$ to $(x(n+1)\varepsilon)$ with probability $1-\varepsilon$ and from $(x, n\varepsilon)$ to $(x+1,(n+1)\varepsilon)$ and from $(x, n\varepsilon)$ to $(x-1,(n+1)\varepsilon)$ with probability $\lambda\varepsilon$. Applying Harris' theorem to the oriented percolation process and letting $\varepsilon \to 0$ shows that for any $T < \infty$,

$$\begin{aligned} &P\big(\tau^{\{1,\ldots,m\}} \le T, \tau^{\{m+1,\ldots,m+n\}} \le T\big) \\ &\quad \ge P\big(\tau^{\{1,\ldots,m\}} \le T\big)P\big(\tau^{\{m+1,\ldots,m+n\}} \le T\big) \end{aligned}$$

and letting $T \to \infty$ proves the desired result.] □

With $(*)$ established the rest of the proof is like the proof of (1) in Section 2. If we let $a_n = \log P(\tau^{\{1,\ldots,n\}} < \infty)$, then $(*)$ becomes $a_{m+n} \ge a_m + a_n$ so (1) holds with

$$-\gamma_2(\lambda) = \sup_n \frac{1}{n} \log P\big(\tau^{\{1,\ldots,n\}} < \infty\big).$$

It is unfortunate for us that the argument above does not work when τ is replaced by $\hat{\tau}$, the extinction time of the process in the half-space, because then the last $=$ becomes $\le$. Some information about $\hat{\tau}$ can, of course, be obtained form the last argument: If

$$-\beta_2(\lambda) = \liminf_{n\to\infty} \frac{1}{n} \log P\big(\hat{\tau}^{\{1,\ldots,n\}} < \infty\big),$$

then $\beta_2(\lambda) \le \gamma_2(\lambda)$ and it follows from the formula for $\gamma_2(\lambda)$ that $\beta_2(\lambda) < \infty$.

Fortunately for us, some things generalize in a straightforward way from the line to the half-space.

(2) If $\lambda > \lambda_c$, then

$$-\alpha_2(\lambda) = \limsup_{n\to\infty} \frac{1}{n} \log P\big(\hat{\tau}^{\{1,\ldots,n\}} < \infty\big) < 0.$$

(3) If $\lambda > \lambda_c$, then there are constants $C, \delta \in (0,\infty)$ (independent of A) so that

$$P\big(t \le \hat{\tau}^A < \infty\big) \le Ce^{-\delta t}.$$

The proofs given in Durrett (1984) on pages 1028–1029 and 1031–1032 for oriented percolation generalize easily to the present situation: All that is needed is that in oriented percolation there is percolation in a half-space when p is close to 1 and that this can be proved by the contour method in Section 10 of the paper cited.

At this point we have assembled all the ingredients and we can begin the proof of Theorem 2. We begin by observing that

$$P(\hat{\tau}^{\{1,\dots,n\}} \le n\theta) + P(n^{\theta} \le \hat{\tau}^{\{1,\dots,n\}} < \infty) = P(\hat{\tau}^{\{1,\dots,n\}} < \infty)$$

and $P(n\theta \le \hat{\tau}^{\{1,\dots,n\}} < \infty) \le Ce^{-\delta n\theta}$, where $C, \delta \in (0,\infty)$ are independent of n, so if θ is large

$$-\beta_2(\lambda) = \liminf_{n\to\infty} \frac{1}{n} \log P(\hat{\tau}^{\{1,\dots,n\}} \le n\theta)$$

[recall $\beta_2(\lambda)$ is defined by setting $\theta = \infty$ in the right-hand side].

Comparing σ_N with a process which is reset to $\{1,\dots,N\}$ at times $N\theta, 2N\theta,\dots$ gives

$$P(\sigma_N > t) \le P(\hat{\tau}^{\{1,\dots,N\}} > N\theta)^{[t/N\theta]},$$

where $[x]$ = the greatest integer $\le x$. If $\varepsilon > 0$ and N is large

$$P(\hat{\tau}^{\{1,\dots,N\}} \le N\theta) \ge \exp(-(1+\varepsilon)\beta_2(\lambda)N)$$

and hence

$$P(\sigma_N > t) \le (1 - \exp(-(1+\varepsilon)\beta_2(\lambda)N))^{[t/N\theta]}.$$

If we let $t = \exp((1+2\varepsilon)\beta_2(\lambda)N)$ and use the observation at the end of Section 2 we see that $P(\sigma_N > \exp((1+2\varepsilon)\beta_2(\lambda)N)) \to 0$, which proves the first half of Theorem 2. To prepare for a remark in Section 5 the reader should observe that the only property of the contact process used above is that it is a monotone or "attractive" process.

To prove the second half of Theorem 2 we let $\hat{l}_t = \inf \hat{\xi}_t^{\{1,2,\dots\}}$ and observe that the duality equation implies

$$P(\hat{l}_t > n) = P(\hat{\tau}^{\{1,\dots,n\}} \le t).$$

We want to estimate the probability that $\hat{l}_s$ exceeds $N/2$ at some time $\le \exp((1-2\varepsilon)\alpha_2(\lambda)N/2)$. To do this we observe that

$$E\left|\{s \le t\colon \hat{l}_s > N/2\}\right| \le tP(\hat{\tau}^{\{1,\dots,N/2\}} < \infty),$$

where $|\{\cdots\}|$ denotes the Lebesgue measure of the indicated set. Setting $t = \exp((1-2\varepsilon)\alpha_2(\lambda)N/2)$ we have for N sufficiently large

$$E\left|\{s \le \exp((1-2\varepsilon)\alpha_2(\lambda)N/2)\colon \hat{l}_s > N/2\}\right| \le \exp(-\varepsilon\alpha_2(\lambda)N/2).$$

To convert this into the bound we want, we consider the first time $\hat{l}_t > N/2$ and observe that with probability $\ge e^{-\lambda}$, $\hat{l}_t$ will remain $> N/2$ for at least 1 unit of time so

$$P\left(\hat{l}_t > N/2 \text{ for some } t \le e^{(1-2\varepsilon)\alpha_2(\lambda)N/2} - 1\right) \le e^{-\varepsilon\alpha_2(\lambda)N/2}/e^{-\lambda}.$$

A similar argument shows that if we consider the contact process on $\{N, N-1,\dots\}$, call it $\tilde{\xi}_t^N$, and let $\tilde{r}_t^N = \sup \tilde{\xi}_t^{\{N, N-1,\dots\}}$, then

$$P\left(\tilde{r}_t^N < N/2 \text{ for some } t < e^{(1-2\varepsilon)\alpha_2(\lambda)N/2} - 1\right) \le e^{-\varepsilon\alpha_2(\lambda)N/2}/e^{-\lambda}.$$

To combine this with the last inequality to give the desired lower bound on the exit time we let ζ_t^N be the contact process on $\{1,\dots,N\}$, let $l_t^N = \inf \zeta_t^N$, let $r_t^N = \sup \zeta_t^N$, and observe that on $\{\hat{l}_s \le \tilde{r}_s^N \text{ for all } s \le t\}$ we have $l_t^N = \hat{l}_t$, $r_t^N = \tilde{r}_t^N$ and

$$\zeta_t^N = \left[l_t^N, r_t^N\right] \cap \hat{\xi}_t = \left[l_t^N, r_t^N\right] \cap \tilde{\xi}_t^N. \tag{4}$$

[To prove this we check that while $\zeta_t^N \neq \varnothing$ every transition preserves the last two equalities (see Durrett (1984), Section 3, for a detailed proof of a similar result).] Combining the last observation with the inequalities above shows

$$P\big(\sigma_N \le \exp((1-2\varepsilon)N\alpha_2(\lambda)/2) - 1\big) \to 0$$

and proves the second half of Theorem 2.

4. Equality of constants for the biased voter model. In this section we will show that if $\hat{\xi}_t^{\{1,\dots,n\}}$ is the biased voter model in $\{1,2,\dots\}$, then

$$\lim_{n\to\infty} \frac{1}{n} \log P\big(\hat{\tau}^{\{1,\dots,n\}} < \infty\big) = -\log\lambda. \tag{1}$$

The first step in doing this is to observe that if L_t and r_t^N are the processes defined in the introduction (observe that one is capital and the other lowercase), then a realization of the process on the half-line can be constructed as follows:

$$\begin{aligned} \hat{\xi}_t^{\{1,\dots,N\}} &= \{L_t, r_t^N\}, && t < \hat{\tau}_N, \\ &= \varnothing, && t \ge \hat{\tau}_N, \end{aligned}$$

where $\hat{\tau}_n = \inf\{t\colon r_t^N < L_t\} = \hat{\tau}^{\{1,\dots,N\}}$.

As we observed in the Introduction

$$P\big(\tau^{\{1,\dots,N\}} < \infty\big) = \left(\frac{1}{\lambda}\right)^N, \tag{2}$$

and it is clear that

$$P\big(\hat{\tau}^{\{1,\dots,N\}} < \infty\big) \ge P\big(\tau^{\{1,\dots,N\}} < \infty\big),$$

so it suffices to show

$$\limsup_{N\to\infty} \frac{1}{N} P\big(\hat{\tau}^{\{1,\dots,N\}} < \infty\big) \le -\log\lambda;$$

i.e., we need upper bounds on the quantity in question.

Our first step is to prove:

(3) There is a constant C such that

$$P\Big(\max_{s\le t} L_s \ge k\Big) \le Ct\lambda^{-k},$$

for all $t \ge 1$ and $k \ge 1$.

Proof. Let $S_1, S_2, \dots$ be the times at which L_t returns to 1, i.e., we let $S_0 = 0$ and for $i \geq 1$,

$$U_i = \inf\{t > S_{i-1}: L_t = 2\}, \qquad S_i = \inf\{t > U_i: L_t = 1\},$$

$$\begin{aligned} P\Big(\max_{s \leq t} L_s \geq k\Big) &\leq \sum_{i=0}^{\infty} P\Big(\max_{S_i \leq u \leq S_{i+1}} L_u \geq k,\ S_i \leq t\Big) \\ &= \sum_{i=0}^{\infty} P\Big(\max_{S_i \leq u \leq S_{i+1}} L_u \geq k \,|\, S_i \leq t\Big) P(S_i \leq t) \\ &= \frac{\lambda^2 - \lambda}{\lambda^k - \lambda} \sum_{i=0}^{\infty} P(S_i \leq t) \\ &\leq C\lambda^{-k} ET(t), \end{aligned}$$

where $T(t) = 1 + \sup\{i: S_i \leq t\}$.

To estimate $ET(t)$ we observe that $P(U_0 > 1) = e^{-1}$ so $P(S_{i+1} - S_i > 1) \geq e^{-1}$ and a simple argument [see, e.g., Chung (1974), page 136] shows $ET(t) < \infty$ for all t. Considering the time of the first return after s and observing that $T(t)$ counts the return at time 0, shows that $ET(t) \geq E(T(s + t) - T(s))$ holds for all $s > 0$. Iterating the last relation with $t = 1$ gives

$$ET(t) \leq ([t] + 1)ET(1) \leq 2tET(1)$$

for $t \geq 1$ and we have proved (3). □

(4) Let $r_t = r_t^N - N$. Then for all $a > 0$ there is an $\varepsilon(a) > 0$ such that

$$P(r_t - (\lambda - 1)t \leq -at) \leq e^{-\varepsilon(a)t}.$$

Proof. This type of large deviation estimate is old [see Feller (1971)] and should be well known but for completeness we sketch the very easy proof. Let $s_t = r_t - (\lambda - 1)t$. If $\theta < 0$, then

$$e^{-a\theta t} P(s_t \leq -at) \leq Ee^{\theta s_t} = \exp(\varphi(\theta)t),$$

where $\varphi(\theta)$ is a function with $\varphi'(0) = Es_1 = 0$. The last inequality implies

$$P(s_t \leq -at) \leq \exp((\varphi(\theta) + a\theta)t)$$

and [since $\varphi'(0) = 0$, $a > 0$] if we pick $\theta < 0$ small $\varphi(\theta) + a\theta < 0$. □

Having assembled the necessary ingredients we are ready to do the

Proof of (1). As we remarked earlier it suffices to show

$$\limsup_{N \to \infty} \frac{1}{N} \log P(\hat{\tau}^{\{1,\dots,N\}} < \infty) \leq -\log \lambda.$$

For this purpose let $\rho_N = \inf\{t: |\hat{\xi}_t^{\{1,\dots,N\}}| = 1\}$, observe $\rho_N < \hat{\tau}^{\{1,\dots,N\}}$, and recall how we constructed $\hat{\xi}_t^{\{1,\dots,N\}}$ from two independent processes L_t and r_t^N. To estimate $P(\rho_N < \infty)$ we will decompose the event according to the value of

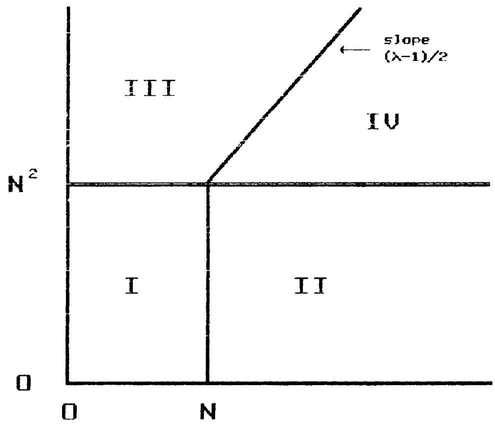

FIG. 1.

$\rho_N = t$, $\xi^{\{1,\dots,N\}}(\rho_N) = \{k\}$, and divide $\{(k, t)\colon k \in Z^+,\ 0 \le t < \infty\}$ into the four parts pictured in Figure 1. To simplify the notation we will let $\{\kappa_N\} = \xi^{\{1,\dots,N\}}(\rho_N)$.

$$\begin{aligned} P(\rho_N < \infty) &= \sum_{k=1}^{N} P(\kappa_N = k,\ \rho_N \le N^2) + P(\kappa_N > N,\ \rho_N \le N^2) \\ &\quad + P(\kappa_N - N \le (\rho_N - N^2)(\lambda - 1)/2,\ \rho_N > N^2) \\ &\quad + P(\kappa_N - N > (\rho_N - N^2)(\lambda - 1)/2,\ \rho_N > N^2) \\ &= \left(\sum_{k=1}^{N} \mathrm{I}_k\right) + \mathrm{II} + \mathrm{III} + \mathrm{IV}. \end{aligned}$$

The four terms above will be estimated in correspondingly numbered parts of the proof.

PART I. The gambler's ruin formula implies

$$P\left(\min_{0 \le t < \infty} r_t^N \le k\right) = \left(\frac{1}{\lambda}\right)^{N-k}. \tag{5}$$

Since L_t and r_t^N are independent processes

$$\begin{aligned} \mathrm{I}_k &= P(L_t = k,\ r_t^N = k \text{ for some } t \le N^2) \\ &\le P\left(\max_{0 \le t \le N^2} L_t \ge k,\ \min_{0 \le t < \infty} r_t^N \le k\right) \\ &= P\left(\max_{0 \le t \le N^2} L_t \ge k\right) P\left(\min_{0 \le t < \infty} r_t^N \le k\right) \\ &\le CN^2\lambda^{-k}\lambda^{-(N-k)} = CN^2\lambda^{-N}, \end{aligned}$$

by (3) and (5).

PART II. Using (3) again gives

$$\mathrm{II} = P\big(\kappa_N > N,\, \rho_N \le N^2\big) \le P\Big(\max_{0\le t\le N^2} L_t > N\Big) \le CN^2\lambda^{-N}.$$

PART III. Let $T = \inf\{t > N^2\colon r_t^N - N \le (t - N^2)(\lambda - 1)/2\}$. Then III $\le$ $P(T < \infty)$. To estimate the last probability we begin as in Section 3 by estimating (here $|\ |$ denotes Lebesgue measure)

$$\begin{aligned} &E\big|\{t\colon t \ge N^2;\ r_t^N - N \le (t - N^2)(\lambda - 1)/2\}\big| \\ &\quad= \int_{N^2}^{\infty} P\big(r_t^N - N \le (t - N^2)(\lambda - 1)/2\big)\, dt \\ &\quad\le \int_{N^2}^{\infty} P\big(r_t^N - N - (\lambda - 1)t \le -(\lambda - 1)t/2\big)\, dt, \end{aligned}$$

which by (4) is [here $\delta \equiv \varepsilon(\lambda - 1/2)$]

$$\le \int_{N^2}^{\infty} e^{-\delta t}\, dt = \delta^{-1}e^{-\delta N^2}.$$

To convert this into an estimate on $P(T < \infty)$ we observe that with probability $e^{-(\lambda+1)}$ we have $r_t^N = r_T^N$ for all $t \in [T, T+1]$ so

$$P(T < \infty)e^{-(\lambda+1)} \le E\big|\{t\colon t \ge N^2,\ r_t^N - N \le (t - N^2)(\lambda - 1)/2\}\big|.$$

PART IV. Breaking things up according to the interval $[N^2 + k - 1, N^2 + k]$ in which ρ_N is, we get

$$\begin{aligned} \mathrm{IV} &\le P\Big(\max_{0\le s\le \rho_N} L_s \ge N + \big(\rho_N - N^2\big)\frac{\lambda - 1}{2},\ \rho_N > N^2\Big) \\ &\le \sum_{k=1}^{\infty} P\Big(\max_{0\le s\le N^2+k} L_s \ge N + (k - 1)(\lambda - 1)/2\Big) \\ &\le \sum_{k=1}^{\infty} C(N^2 + k)\lambda^{-(N+(k-1)(\lambda-1)/2)}, \end{aligned}$$

by (3) and the observation that if $L_s > x$, then $L_s \ge [x] + 1$. Summing the series gives

$$\mathrm{IV} \ge (CN^2 + C)\lambda^{-N}$$

(here and in what follows C is a constant whose value is unimportant and which will change from line to line).

PART V. Having estimated the four terms all that remains is to add up the estimates to conclude

$$\begin{aligned} P(\rho_N < \infty) &\le \sum_{k=1}^{N} \mathrm{I}_k + \mathrm{II} + \mathrm{III} + \mathrm{IV} \\ &\le CN^3\lambda^{-N} + CN^2\lambda^{-N} + e^{-(\lambda+1)}\delta^{-1}e^{-\delta N^2} + (CN^2 + C)\lambda^{-N}, \end{aligned}$$

so

$$\limsup_{N\to\infty} \frac{1}{N}\log P(\rho_N < \infty) \le -\log\lambda,$$

and the proof of (1) is complete. □

5. Proof of Theorem 4. In this section we will prove

THEOREM 4. *If* $\lambda > 1$, *then as* $N \to \infty$,

$$\frac{1}{N}\log \sigma_N \to \log\lambda, \quad \textit{in probability}.$$

In the last section we showed

$$\lim_{N\to\infty} \frac{1}{N}\log P(\hat{\tau}^{\{1,\dots,N\}} < \infty) = -\log\lambda$$

or in the notation of Theorem 2, $\beta_2(\lambda) = -\log\lambda$, so repeating the proof of the first half of that result (which as we remarked in Section 3 is very general) shows that for all $\varepsilon > 0$,

$$P\left(\frac{\log\sigma_N}{N} > \beta_2(\lambda) + \varepsilon\right) \to 0$$

and all we have to show is

$$(*) \qquad P\left(\frac{\log\sigma_N}{N} < -(1-\varepsilon)\log\lambda\right) \to 0.$$

In the last section we showed [here the formula number indicates the result below was formula (3) in Section 4]:

(4.3) There is a constant C such that

$$P\left(\max_{s\le t} L_s \ge k\right) \le Ct\lambda^{-k},$$

for all $t \ge 1$ and $k \ge 1$.

The key to the proof of Theorem 4 is to improve the last result to:

(1) If $K < \infty$ there are constants $C, \delta \in (0,\infty)$ (which will depend upon K) so that if $t \le \exp(KN)$, then

$$P\left(\max_{t-1\le s\le t} L_s \ge k\right) \le CN^2\lambda^{-k} + Ce^{-\delta N^2}.$$

PROOF OF (1). Using the notation of Section 4, let $S_1, S_2, \dots$ be the times at which L_t returns to 1, i.e., we let $S_0 = 0$ and for $i \ge 1$,

$$U_i = \inf\{t > S_{i-1}: L_t = 2\}, \qquad S_i = \inf\{t > U_i: L_t = 1\}.$$

The first step in our proof is to prove:

(2) There are constants $C, \delta \in (0, \infty)$ so that

$$P(S_1 > t) \le Ce^{-\delta t}.$$

PROOF. Let $V_1 = S_1 - U_1$,

$$P(S_1 > t) \le P(U_1 > t/2) + P(V_1 > t/2).$$

Since U_1 is a mean-one exponential

$$P(U_1 > t/2) = e^{-t/2}.$$

To estimate V_1 observe that $L_t = 2$ for $t = U_1$ and for $U_1 \le t \le S_1$, L_t behaves like the unrestricted random walk l_t so

$$\begin{aligned} P(V_1 > t/2) &= P(l_s \ge 2 \text{ for } s \le t/2 | l_0 = 2) \\ &\le P(l_{t/2} - l_0 \ge 0) \\ &\le Ce^{-\varepsilon(\lambda-1)t/2}, \end{aligned}$$

by a trivial generalization of (4.4). Adding this to the first estimate proves (2).

We call the time intervals $[S_i, S_{i+1}]$, $i \ge 0$, excursion intervals and following the notation in the last section let $T(t)$ = the number of excursions which start in $[0, t]$. The last result says there are not too many long intervals, the next says there are not too many short ones.

(3) There are constants $C, \delta \in (0, \infty)$ so that for $t \ge 1$,

$$P(T(t) > 3t) \le Ce^{-\delta t}.$$

PROOF. Let $X_i = S_i - S_{i-1}$. The X_i are i.i.d. with $P(X_i > 1) \ge P(U_1 > 1) = e^{-1}$ and $e < 3$ so

$$P\left(S_k \le \tfrac{1}{3}k\right) \le P\left(|\{i \le k \colon X_i > 1\}| \le \tfrac{1}{3}k\right) \le Ce^{-\varepsilon k},$$

by the large deviations estimate used to prove (4.4). From the last inequality it follows that if t is an integer

$$P(T(t) > 3t) = P(S_{3t} \le t) \le Ce^{-3\varepsilon t},$$

and then the result given above follows easily from this.

With our preliminary results on excursion intervals established we are ready to begin the proof of (1). The second term on the right-hand side comes from:

(4) The probability that some excursion starting in $[0, e^{KN}]$ has length $\ge N^2$ is

$$\begin{aligned} &\le P\left(T(e^{KN}) \ge 3e^{KN}\right) + 3e^{KN}P\left(S_1 \ge N^2\right) \\ &\le Ce^{-\delta 3\exp(KN)} + 3e^{KN}Ce^{-\delta N^2} \\ &\le Ce^{-\delta N^2} \end{aligned}$$

(here as before and in what follows, C and δ are positive finite constants whose values change from line to line).

The first term on the right-hand side of (1) (and another piece the same size as the second term) comes from:

(5) The probability some excursion starting in $[t - N^2, t]$ has

$$\max_{S_t \le s \le S_{t+1}} L_s \ge k$$

is

$$\le P\big(T(t) - T(t - N^2) \ge 3N^2\big) + 3N^2 P\Big(\max_{s \le s_1} L_2 \ge k\Big)$$

$$\le Ce^{-\delta N^2} + 3N^2(1/\lambda)^{k-2}$$

[see the proof of (4.3) to bound the second term]. Adding (4) and (5) gives (1). □

With (1) established it is easy to prove Theorem 4. Pick $K > \log\lambda$ and apply (1) to conclude that for $t \le \lambda^N$,

$$P\Big(\max_{t-1 \le s \le t} L_s \ge k\Big) \le CN^2\lambda^{-k} + Ce^{-\delta N^2}$$

and

$$P\Big(\min_{t-1 \le s \le t} R_s^N \le j\Big) \le CN^2\lambda^{-(N-j)} + Ce^{-\delta N^2}.$$

Let $M = [2/\varepsilon] + 1$ and for $0 \le i \le M$ let $n_i = iN/M$. The n_i have $n_i - n_{i-1} \le \varepsilon N/2$ so recalling that in the Introduction we constructed the biased voter model on $\{1, \dots, N\}$ from independent companies of L_s and R_s^N gives

$$P\big(\sigma_N \le \lambda^{(1-\varepsilon)N}\big) \sum_{t=1}^{[\lambda^{(1-\varepsilon)N}]+1} \sum_{i=1}^{M} P\Big(\max_{t-1 \le s \le t} L_s \ge n_{i-1}\Big) P\Big(\max_{t-1 \le s \le t} R_s^N \le n_i\Big)$$

$$\le \big([\lambda^{(1-\varepsilon)N}] + 1\big) M \big(C^2 N^4 \lambda^{-N(1-\varepsilon/2)} + 3Ce^{-\delta N^2}\big),$$

where in the last line we have used the fact if $CN^2\lambda^{-k}$ or $Ce^{-\delta N^2} > 1$ we can always estimate the probability by 1. As $N \to \infty$ the right-hand side $\to 0$ so we have proved $(*)$ and hence Theorem 4.

Epilogue. Since the completion of this paper the first author and Roberto Schonmann have been able to show that $\alpha_2(\lambda) = \beta_2(\lambda)$ and the factor of 2 in the upper bound in Theorem 2 can be removed. The keys to doing this are (i) finding the right definition of the constant which appears in the limit theorem:

$$-\delta_2(\lambda) = \lim \frac{1}{n} \log P\big(\tau^{\text{eq} \cap \{1, \dots, n\}} < \infty\big),$$

where eq is short for the nontrivial equilibrium state for the contact process, and (ii) the use of a new "planar graph" duality [inspired by a similar idea of Dhar, Barma and Phani (1981) for oriented percolation] in which connections in the contact process graphical representation are dual to the contours of Griffeath

(1981), page 160. If we let

$$-\gamma_2(\lambda) = \lim \frac{1}{n} \log P(\tau^{\{1,\dots,n\}} < \infty),$$

then the duality mentioned above can be used to show that

$$\alpha_2(\lambda) = \beta_2(\lambda) = \delta_2(\lambda) = \gamma_2(\lambda)$$

and once this is done it is easy to prove the upper bound in Theorem 2 without the fact of 2. Details will appear in a future issue of *The Annals of Probability.*

REFERENCES

CASSANDRO, M., GALVES, A., OLIVIERI, E. and VARES, M. E. (1984). Metastable behavior of stochastic, dynamics: A pathwise approach. *J. Statist. Phys.* **35** 603–634.

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

DHAR, D., BARMA, M. and PHANI, M. (1981). Duality transformations for two-dimensional directed percolation and resistance problems. *Phys. Rev. Lett.* **47** 1238–1241.

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

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications* **2**, 2nd ed. Wiley, New York.

GRIFFEATH, D. (1981). The basic contact process. *Stochastic Process. Appl.* **11** 151–185.

HARRIS, T. E. (1960). A lower bound for the critical probability in a certain percolation process. *Proc. Cambridge Philos. Soc.* **56** 13–20.

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

SCHONMANN, R. H. (1985). Metastability for the contact process, *J. Statist. Phys.* **41** 445–464.

STAVSKAYA, O. N. and PIATETSKI-SHAPIRO, I. I. (1968). On homogeneous networks of spontaneously active elements. *Problemy Kibernet.* **20** 91–106.

TOOM, A. L. (1968). A family of uniform nets of formal systems. *Soviet Math. Dokl.* **9** 1338–1341.

DEPARTMENT OF MATHEMATICS
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853

DEPARTMENT OF MATHEMATICS
BEIJING NORMAL UNIVERSITY
BEIJING
CHINA