TEN LECTURES

ON PARTICLE SYSTEMS

Rick DURRETT

Preface. These lectures were written for the 1993 St. Flour Probability Summer School. Their aim is to introduce the reader to the mathematical techniques involved in proving results about interacting particle systems. Readers who are interested instead in using these models for biological applications should instead consult Durrett and Levin (1993).

In order that our survey is both broad and has some coherence, we have chosen to concentrate on the problem of proving the existence of nontrivial stationary distributions for interacting particle systems. This choice is dictated at least in part by the fact that we want to make propaganda for a general method of solving this problem invented in joint work with Maury Bramson (1988): comparison with oriented percolation. Personal motives aside, however, the question of the existence of nontrivial stationary distributions is the first that must be answered in the discussion of any model.

Our survey begins with an overview that describes most of the models we will consider and states the main results we will prove, so that the reader can get a sense of the forest before we start investigating the individual trees in detail. In Section 2 we lay the foundations for the work that follows by proving an existence theorem for particle systems with translation invariant finite range interactions and introducing some of the basic properties of the resulting processes. In Section 3 we give a second construction that applies to a special class of "additive" models, that makes connections with percolation processes and that allows us to define dual processes for these models.

The general method mentioned above makes its appearance in Section 4 (with its proofs hidden away in the appendix) and allows us to prove a very general result about the existence of stationary distributions for attractive systems with state space $\{0,1\}^S$. The comparison results in Section 4 are the key to our treatment of the threshold contact and voter models in Section 5, the cyclic systems in Section 6, the long range contact process in Section 7, and the predator prey system in 9.

In Section 7 we explore the first of two methods for simplifying interacting particle systems: assuming that the range of interaction is large. In Section 8 we meet the second: superimposing particle motion at a fast rate. The second simplification leads to a connection with reaction diffusion equations which we exploit in Section 9 to prove the existence of phase transitions for predator prey systems.

The quick sketch of the contents of these lectures in the last three paragraphs will be developed more fully in the overview. Turning to other formalities, I would like to thank the organizers of the summer school for this opportunity to speak and write about my favorite subject. Many of the results presented here were developed with the support of the National Science Foundation and the Army Research Office through the Mathematical Science Institute at Cornell University. During the Spring semester of 1993, I gave 10 one and a half hour lectures to practice for the summer school and to force myself to get the writing done on time. You should be grateful to the eight people who attended this dress rehersal: Hassan Allouba, Scott Arouh, Itai Benjamini, Carol Bezuidenhout, Elena Bobrovnikova, Sungchul Lee, Gang Ma, and Yuan-Chung Sheu, since their suffering has lessened yours.

Although it is not yet the end of the movie, I would like to thank the supporting cast now: Tom Liggett, who introduced me to this subject; Maury Bramson, the co-discoverer of the comparison method and long range limits, to whom I turn when my problems get too hard; David Griffeath, my electronic colleague who introduced me (and the rest of the world) to the beautiful world of the Greenberg Hastings and cyclic cellular automata; Claudia Neuhauser, my former student who constantly teachs me how to write; and Ted Cox, with whom I have written some of my best papers. The field of interacting particle systems has grown considerably since Liggett's 488 page book was published in 1985, so it is inevitable that more is left out than is covered in these notes. The most overlooked researcher in this treatment is Roberto Schonmann whose many results on the contact process, bootstrap percolation, and metastability in the Ising model did not fit into our plot.

1. Overview

In an interacting particle system, there is a countable set of spatial locations S called *sites*. In almost all of our applications $S = \mathbf{Z}^d$, the set of points in d dimensional space with integer coordinates. Each site can be in one of a finite set of *states* F, so the state of the system at time t is $\xi_t : S \to F$ with $\xi_t(x)$ giving the state of x. To describe the evolution of these models, we specify an *interaction neighborhood*

$$\mathcal{N} = \{z_0, z_1, \dots z_k\} \subset \mathbf{Z}^d$$

with $z_0 = 0$ and define *flip rates*

$$c_i(x,\xi) = g_i(\xi(x+z_0), \xi(x+z_1), \dots, \xi(x+z_k))$$

In words, the state of x flips to i at rate $c_i(x,\xi)$ when the state of the process is ξ. In symbols, if $\xi_t(x) \neq i$ then

$$\frac{P(\xi_{t+s}(x) = i | \xi_t = \xi)}{s} \to c_i(x,\xi) \quad \text{as } s \to 0$$

The formula for c_i indicates that our interaction is *finite range*, i.e., the flip rates depend only on the state of x and of a finite number of neighbors; and *translation invariant*, i.e., the rules applied at x are just a translation of those applied at 0.

To explain what we have in mind when making these definitions, we now describe two famous concrete examples. In this section and throughout these lectures (with the exception of Sections 2 and 3) we will suppose that

$$\mathcal{N} = \{x : \|x\|_p \le r\}$$

Here $r \ge 1$ is the *range* of the interaction and $\|x\|_p$ is the usual L^p norm on $\mathbf{R}^d$. That is, $\|x\|_p = (x_1^p + \ldots + x_d^p)^{1/p}$ when $1 \le p < \infty$ and $\|x\|_\infty = \sup_i |x_i|$. In most of our models the flip rates are based on the number of neighbors in state i, so we introduce the notation:

$$n_i(x,\xi) = |\{z \in \mathcal{N} : \xi(x+z) = i\}|$$

where $|A|$ is the number of points in A.

Example 1.1. The basic contact process. To model the spread of a plant species we think of each site x as representing a square area in space with $\xi_t(x) = 0$ if that area is vacant and $\xi_t(x) = 1$ if there is a plant there, and we formulate the dynamics as follows:

$$\begin{aligned} c_0(x,\xi) &= \delta \qquad \text{if } \xi(x) = 1 \\ c_1(x,\xi) &= \lambda n_1(x,\xi) \qquad \text{if } \xi(x) = 0 \end{aligned}$$

In words, plants die at rate δ independent of the state of their neighbors, while births at vacant sites occur at a rate proportional to the number of occupied neighbors. Note that flipping to i has no effect when $\xi(x) = i$ so the value of $c_i(x, \xi)$ on $\{\xi(x) = i\}$ is irrelevant and we could delete the qualifying phrases "if $\xi(x) = 1$" and "if $\xi(x) = 0$" if we wanted to.

Example 1.2. The basic voter model. This time we think of the sites in $\mathbf{Z}^d$ as representing an array of houses each of which is occupied by one individual who can be in favor of ($\xi_t(x) = 1$) or against ($\xi_t(x) = 0$) a particular issue or candidate. Our simple minded voters change their opinion to i at a rate that is equal to the number of neighbors with that opinion. That is,

$$c_i(x, \xi) = n_i(x, \xi)$$

The first question to be addressed for these models is:

Do the rates specify a unique Markov process?

There is something to be proved since there are infinitely many sites and hence no first jump, but for our finite range translation invariant models, a result of Harris (1972) allows us to easily show that the answer is Yes. (See Section 2.) The main question we will be interested in is:

When do interacting particle systems have a nontrivial stationary distributions?

To make this question precise we need a few definitions. The state space of our Markov process is F^S, the set of all functions $\xi : S \to F$. We let $\mathcal{F}$ = all subsets of F and equip F^S with the usual product σ-field $\mathcal{F}^S$, which is generated by the *finite dimensional sets*

$$\{\xi(y_1) = i_1, \ldots, \xi(y_k) = i_k\}$$

So any measure π on $\mathcal{F}^S$ can be described by giving its *finite dimensional distributions*

$$\pi(\xi(y_1) = i_1, \ldots, \xi(y_k) = i_k)$$

As in the theory of Markov chains, π is said to be a *stationary distribution* for the process if when we start from an initial state ξ_0 with distribution π (i.e., $\pi(A) = P(\xi_0 \in A)$ for $A \in \mathcal{F}^S$) then ξ_t has distribution π for all $t > 0$. Since our dynamics are translation invariant, we will have a special interest in stationary distributions that are *translation invariant*, i.e., ones in which the probabilities $\pi(\xi(x + y_1) = i_1, \ldots, \xi(x + y_k) = i_k)$ do not depend upon x.

To explain the term "nontrivial" we note that in Example 1.1 the "all 0" state ($\xi(x) \equiv 0$) and in Example 1.2 for any i the all i state are *absorbing states*. That is, once the process enters these states it cannot leave them. If S were finite this fact (and enough irreducibility, which is present in Examples 1.1 and 1.2) would imply that all stationary distributions were *trivial*, i.e., concentrated on absorbing states. However, when S is infinite this argument fails and indeed, as the next few results show it is possible to have a nontrivial stationary distributions.

Theorem 1. Consider the basic contact process with $\mathcal{N} = \{x : \|x\|_p \le r\}$ with $r \ge 1$. If $\lambda|\mathcal{N}| \le \delta$ then there is only the trivial stationary distribution. If $\delta/\lambda < \delta_0$ then there is a nontrivial translation invariant stationary distribution.

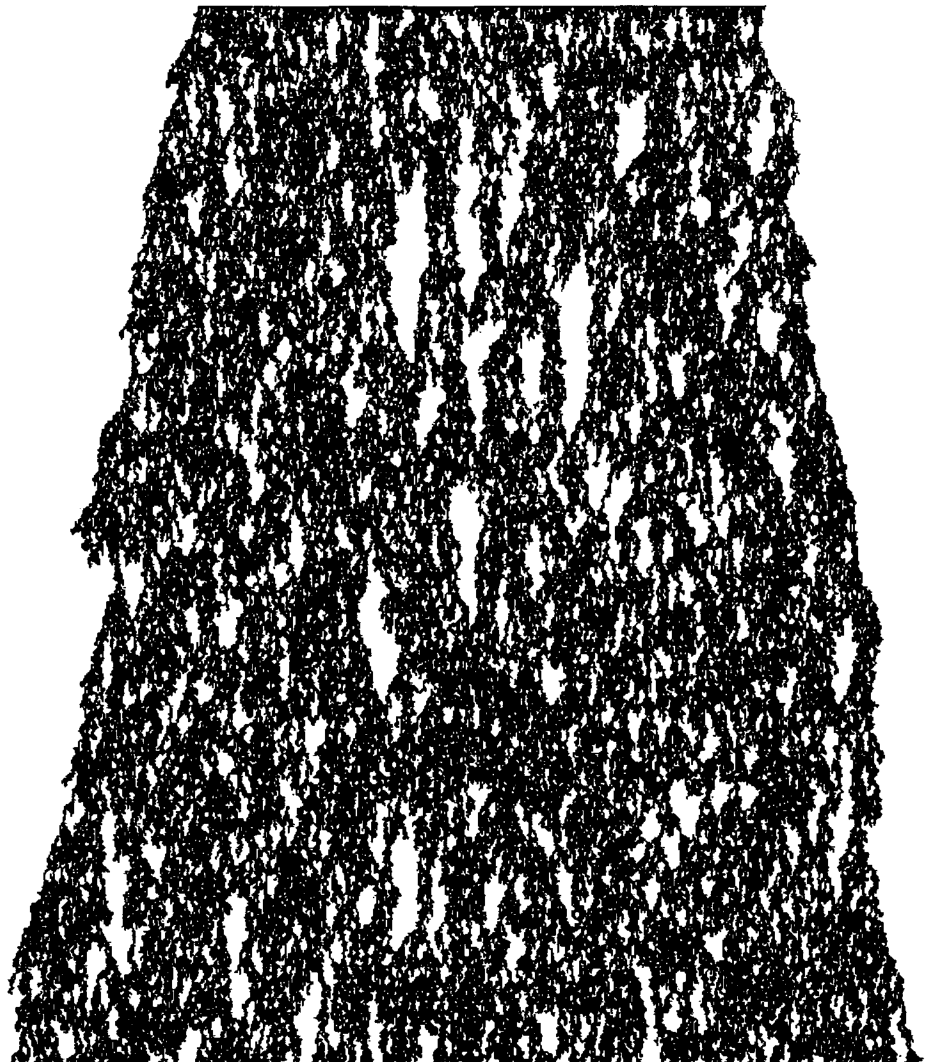


Figure 1.1. Nearest neighbor contact process in $d = 1$ with $\lambda = 2$.

The first result is easy to see. If the contact process has k particles then the number drops to $k-1$ at rate δk and increases to $k+1$ at rate $\leq \lambda|\mathcal{N}|$ with the upper bound achieved when all particles are isolated (i.e., no two particles are neighbors). The reader should attempt to prove the converse before we hit it with our sledgehammer in Section 4. By a simple comparison that you will learn about in Section 2, it is enough to prove the result when $\mathcal{N} = \{x : \|x\|_1 = 1\}$ and $d = 1$. A simulation of this case with $\lambda = 2$ is given in Figure 1.1. A result of Holley and Liggett (1978) implies that in this situation there is a nontrivial stationary distribution. In our simulation we have started with the interval $[180, 540]$ occupied at time 0 at the top of the page. As time runs down the page from 0 to 720, it is clear that the region occupied by particles is growing linearly, as predicted by a result of Durrett (1980).

Turning to the voter model, the classic paper of Holley and Liggett (1975) tells us that

Theorem 2A. *Clustering* occurs in $d \leq 2$. That is, for any ξ_0 and $x, y \in \mathbf{Z}^d$ we have

$$P(\xi_t(x) \neq \xi_t(y)) \to 0 \text{ as } t \to \infty$$

Theorem 2B. Let ξ_t^θ denote the process starting from an initial state in which the events $\{\xi_0^\theta(x) = 1\}$ are independent and have probability θ. In $d \geq 3$ as $t \to \infty$, $\xi_t^\theta \Rightarrow \xi_\infty^\theta$, a translation invariant stationary distribution in which $P(\xi_\infty^\theta(x) = 1) = \theta$.

Here $\Rightarrow$ denotes *weak convergence*, which in this setting is just convergence of finite dimensional distributions. That is, for any $x_1, \ldots x_m \in \mathbf{Z}^d$ and $i_1, \ldots, i_m \in \{1, 2 \ldots, \kappa\}$ we have

$$P(\xi_t^\theta(x_1) = i_1, \ldots \xi_t^\theta(x_m) = i_m) \to P(\xi_\infty^\theta(x_1) = i_1, \ldots \xi_\infty^\theta(x_m) = i_m)$$

We will say that *coexistence* occurs if there is a translation invariant stationary distribution in which each of the possible states in F has positive density. Theorems 2A and 2B say that in the voter model coexistence is possible in $d \geq 3$ but not in $d \leq 2$. We will see in Section 3 that this is a consequence of the fact that if we take two independent random walks with jumps uniformly distributed on $\mathcal{N}$ then they will hit with probability 1 in $d \leq 2$ but with probability < 1 in $d \geq 3$.

Figure 1.2 gives a simulation of a voter model with five opinions on $\{0, 1, \ldots, 119\}^2$. Here and in the next six simulations in this section, we use periodic boundary conditions. That is, sites on the top row are neighbors of those on the bottom row, and those on the left edge are neighbors of those on the right edge. We started at time 0 by assigning a randomly chosen opinion to each site. Figure 1.2 shows the state at time 500 suggesting that the clustering asserted in Theorem 2A occurs very slowly. Results of Cox (1988) imply that the expected time for our system to reach consensus is about

$$4 \ln(5/4) \cdot \frac{2}{\pi} (120)^2 \ln 120 = 39,173$$

The conclusions just derived for the voter depend on the fact that the flip rates are linear. Nonlinear flip rates can produce quite different behavior:

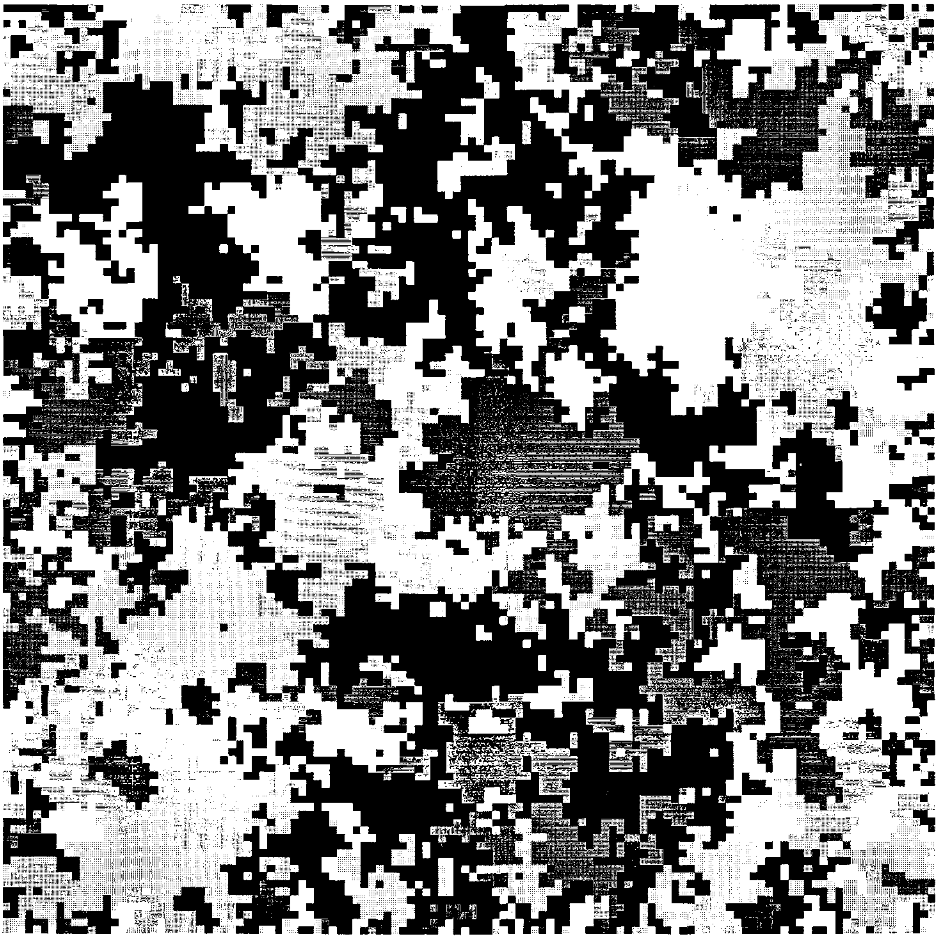


Figure 1.2. Five opinion two dimensional voter model at time 500

Example 1.3. The threshold voter model. Cox and Durrett (1991) introduced a modification of the voter model in which

$$c_i(x,\xi) = \begin{cases} 1 & \text{if } n_i(x,\xi) \geq \theta \\ 0 & \text{if } n_i(x,\xi) < \theta \end{cases}$$

In words, these voters change their opinion at rate 1 if at least θ neighbors disagree with them. This change in the rules changes the behavior of the model drastically.

We start with the case $\theta = 1$:

Theorem 3A. If $d = 1$ and $\mathcal{N} = \{-1, 1\}$ then clustering occurs.

Theorem 3B. In all other cases (recall we supposed that $\mathcal{N} = \{z : \|z\|_p \leq r\}$ with $r \geq 1$) we have coexistence. That is, there is a nontrivial translation invariant stationary distribution μ_{12} in which 1's and 2's each have density 1/2.

Here as in many other cases, the one dimensional nearest neighbor case is an exception. Cox and Durrett (1991) proved Theorem 3A and that coexistence occurs in some cases (e.g., $d = 1$ and $r \geq 7$) but the sharp Theorem 3B is due to Liggett (1992). Note that in the threshold voter model coexistence occurs in all but one case, while in the basic voter model coexistence occurs only in $d \geq 3$. A second difference is that when coexistence occurs the basic voter model has a one parameter family of nontrivial stationary distributions constructed in Theorem 2B but we believe

Conjecture 3C. When coexistence occurs in the threshold one voter model there is a unique spatially ergodic translation invariant stationary distribution in which 1's and 2's have positive density.

Here, we say that π on F^S is *spatially ergodic* if under π the family of random variables $\{\xi(x) : x \in \mathbf{Z}^d\}$ is an ergodic stationary sequence, i.e., the σ-field of events invariant under all spatial shifts is trivial. We need the assumption of spatial ergodicity to rule out nontrivial convex combinations

$$a\mu_1 + b\mu_2 + (1 - a - b)\mu_{12}$$

where μ_i is the point mass on the all i state, and μ_{12} is the measure constructed in Theorem 3B. In general, the set of translation invariant stationary distributions for an interacting particle system is a convex set and in most examples, the extreme points of the set are the stationary distributions that are spatially ergodic. However, there is no general result that shows this is true. See Problem 7 on page 178 of Liggett (1985).

While the threshold 1 case is fairly well understood, there are many open problems concerning higher thresholds. To illustrate these we observe that computer simulations suggest

Conjecture 3D. For the Moore neighborhood $\mathcal{N} = \{z : \|z\|_\infty = 1\}$ in $d = 2$ the threshold voter model has the following behaviors

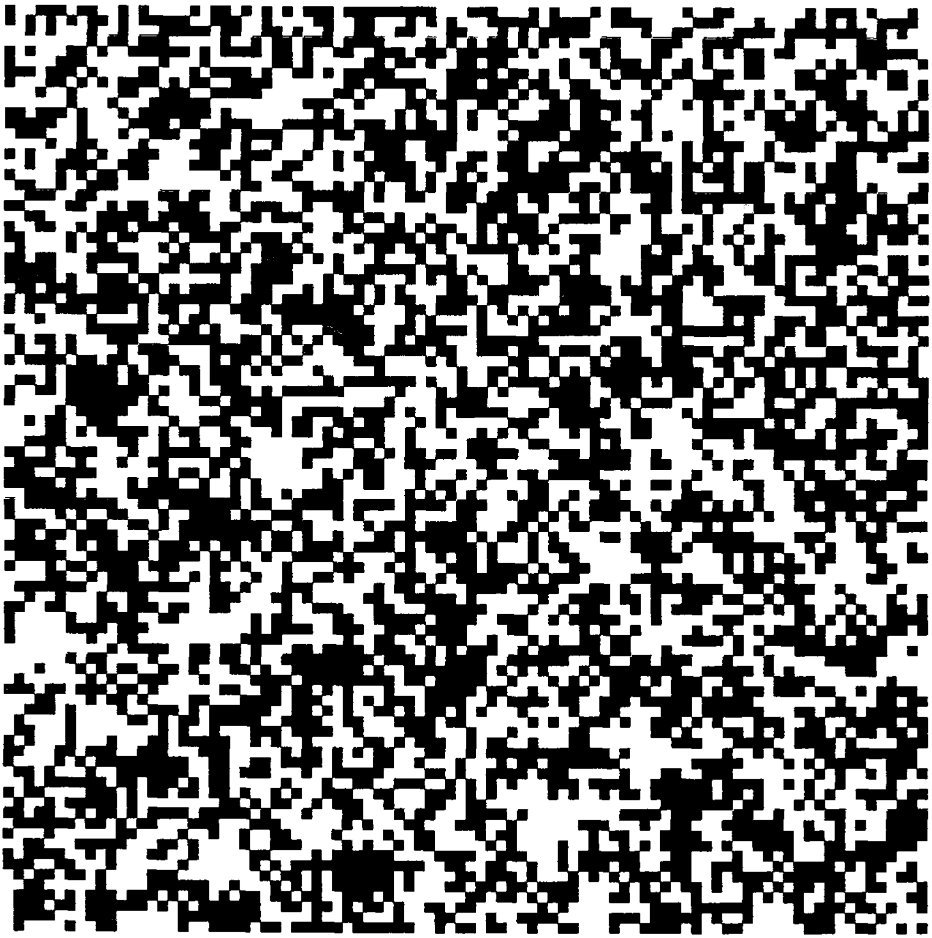


Figure 1.3. Threshold 2 voter model, Moore neighborhood

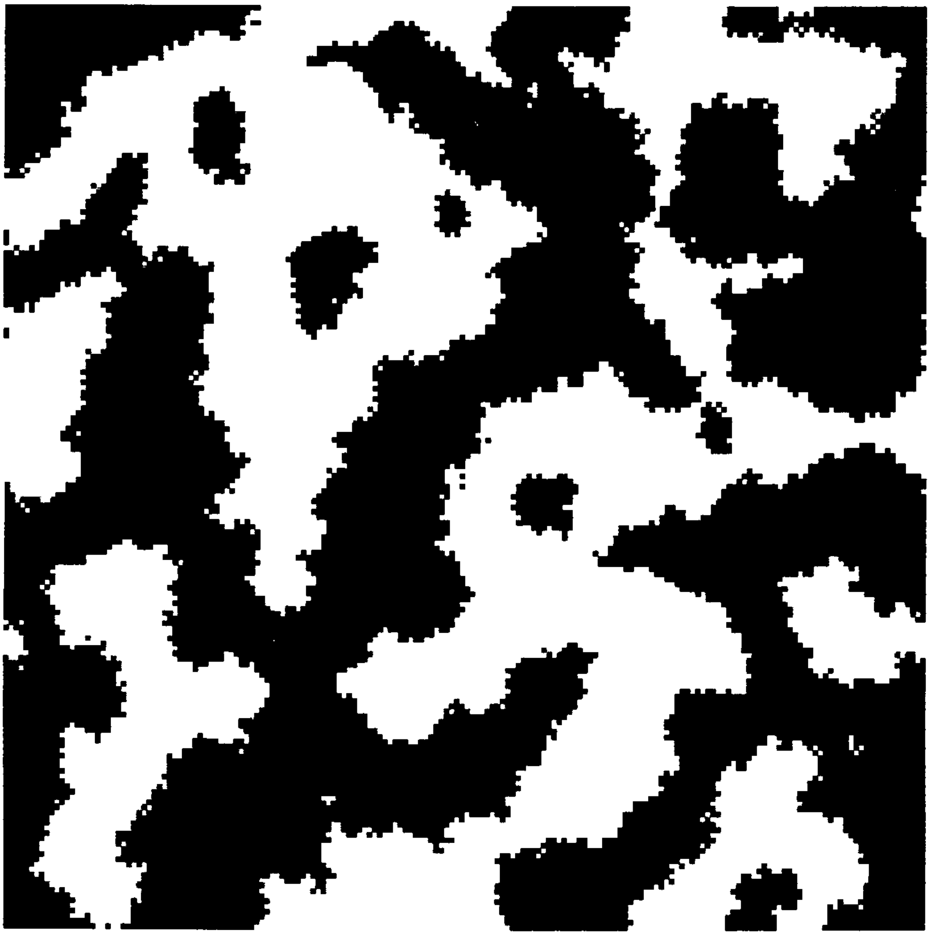


Figure 1.4. Threshold 3 voter model, Moore neighborhood

coexistence	$\theta = 1, 2$
clustering	$\theta = 3, 4$
fixation	$\theta \geq 5$

Here, *fixation* means that each sites flips only a finite number of times. To see that the last line is a reasonable guess note that an octagon of 1's cannot flip to 0 since each 1 has at most 4 neighbors that are 0

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{matrix}$$

We will prove the result about fixation for $\theta \geq 5$ and coexistence for $\theta = 1$ in Section 5. The other conclusions are open problems. In support of our conjectures we introduce Figures 1.3 and 1.4 which give simulations at time 50 of the case $\theta = 2$ on $\{0, 1, \ldots, 89\}^2$ and $\theta = 3$ on $\{0, 1, \ldots, 179\}^2$ starting from product measure with density 1/2.

Our next two systems model the competition of biological species. We begin with

Example 1.4. The multitype contact process. The set of states is $F = \{0, 1, \ldots, \kappa\}$, where 0 indicates a vacant site and $i > 0$ indicates a site occupied by one plant of type i. The flip rates are linear

$$\begin{aligned} c_0(x, \xi) &= \delta_{\xi(x)} \\ c_i(x, \xi) &= \lambda_i n_i(x, \xi) \quad \text{if } \xi(x) = 0 \end{aligned}$$

Here and in what follows the rates we do not mention are 0. Suppose for simplicity that $\kappa = 2$. Neuhauser (1992) has shown

Theorem 4A. Suppose $\delta_1 = \delta_2$ and $\lambda_1 > \lambda_2$. If ξ_0 is translation invariant and has a positive density of 1's then $P(\xi_t(x) = 2) \to 0$.

In words, the species with the higher birth rate wins out ("survival of the fittest"). The following stronger result should be true but Neuhauser's proof relies heavily on the assumption that $\delta_1 = \delta_2$.

Conjecture 4B. Suppose $\lambda_1/\delta_1 > \lambda_2/\delta_2$. If ξ_0 contains infinitely many 1's then $P(\xi_t(x) = 2) \to 0$.

When $\lambda_1 = \lambda_2$ and $\delta_1 = \delta_2$, Neuhauser showed that the multitype contact process behaves like the voter model.

Theorem 4C. *Clustering* occurs for translation invariant initial states in $d \leq 2$. That is, if ξ_0 is translation invariant, then for any $x, y \in \mathbf{Z}^d$, and $1 \leq i < j \leq \kappa$ we have

$$P(\xi_t(x) = i, \xi_t(y) = j) \to 0 \text{ as } t \to \infty$$

Theorem 4D. Let ξ_t^θ denote the process starting from an initial state in which the events $\{\xi_0^\theta(x) = i\}$ are independent and have probability θ_i with $\theta_2 = 1 - \theta_1$. In $d \geq 3$, as $t \to \infty$, $\xi_t^\theta \Rightarrow \xi_\infty^\theta$, a translation invariant stationary distribution in which

$$P(\xi_t(x) = i) = \begin{cases} 1 - \rho & \text{when } i = 0 \\ \rho\theta_i & \text{when } i > 0 \end{cases}$$

where ρ is the equilibrium density of occupied sites in the one type contact process.

The last result is a little disturbing for biological applications. It says that if species compete on an equal footing then coexistence is not possible in $d = 2$ even if the birth and death rates are exactly the same. (This situation may sound unlikely to occur in nature but it occurs, for example, if we look at the competition of *genets* genetically identical individuals of the same species.) Somewhat surprisingly, if species 2 dominates species 1, we get coexistence for an open set of parameter values.

Example 1.5. Successional dynamics. We suppose that the set of states at each site are 0 = grass, 1 = a bush, 2 = a tree and we formulate the dynamics as

$$\begin{aligned} c_0(x,\xi) &= \delta_{\xi(x,\xi)} \\ c_1(x,\xi) &= \lambda_1 n_1(x,\xi) \quad \text{if } \xi(x) = 0 \\ c_2(x,\xi) &= \lambda_2 n_2(x,\xi) \quad \text{if } \xi(x) \leq 1 \end{aligned}$$

The title of this example and its formulation are based on the observation that if an area of land is cleared by a fire, then regowth will occur in three stages: first grass appears then small bushes and finally trees, with each species growing up through and replacing the previous one. With this in mind, we allow each type to give birth onto sites occupied by lower numbered types. As in the threshold voter model, the one dimensional nearest neighbor case is an exception.

Theorem 5A. Coexistence is not possible in the one dimensional nearest neighbor case, i.e., $d = 1$, $\mathcal{N} = \{-1, 1\}$.

Conjecture 5B. In all other cases (recall we supposed that $\mathcal{N} = \{z : \|z\|_p \leq r\}$ with $r \geq 1$) we have coexistence for an open set of values $(\delta_1, \lambda_1, \delta_2, \lambda_2)$.

Figure 1.5 shows a simulation of the nearest neighbor model on $\{0, 1, \ldots, 89\}^2$ with parameters $\lambda_1 = 5/4$, $\delta_1 = 1$, $\lambda_2 = 1.9/4$, and $\delta_2 = 1$ run until time 100, which presumably represents the equilibrium state. Sites in state 1 are gray; those in state 2 are black.

Proving that coexistence occurs in the two dimensional nearest neighbor case of this model seems to be a difficult problem, since computer simulations indicate that the open set referred to in Conjecture 5B is rather small. However, if we assume that the range of interaction is large, we can get very accurate results about the coexistence region. Let $\beta_i = \lambda_i |\mathcal{N}|$.

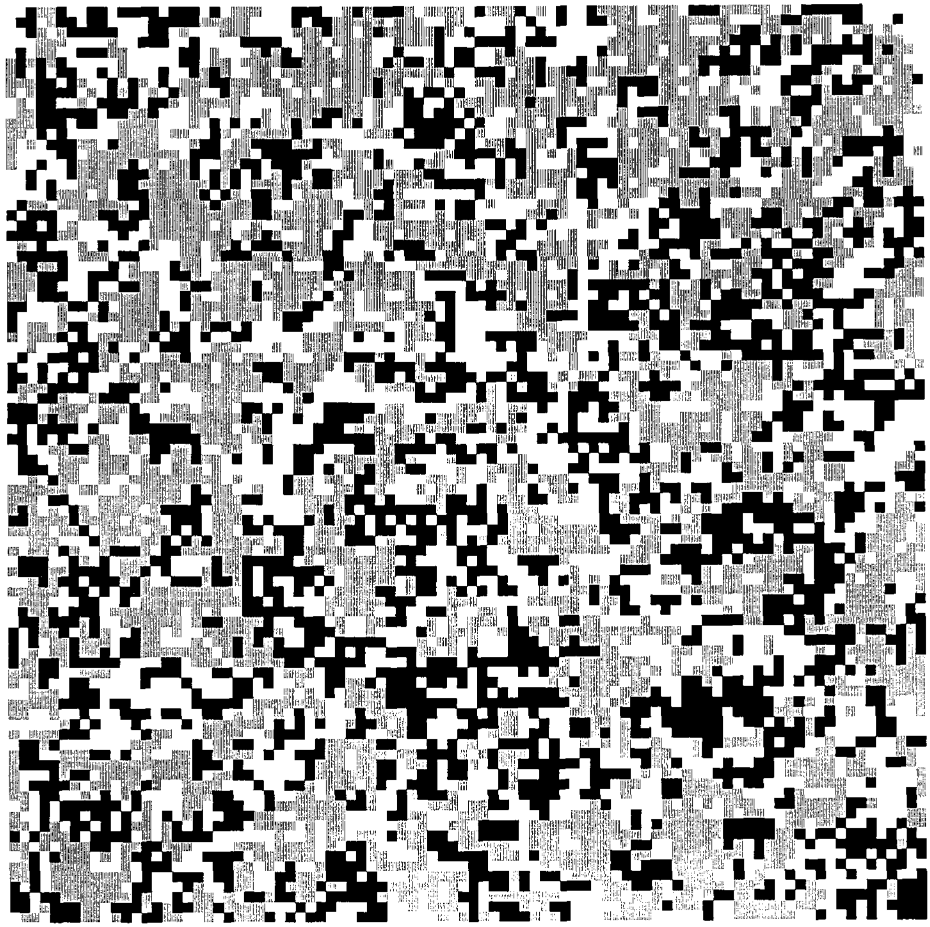


Figure 1.5. Two dimensional nearest neighbor succesional dyanmics, $\beta_1 = 5$, $\beta_2 = 1.9$

Theorem 5C. Suppose that

$$(\star) \qquad \beta_1 \cdot \frac{\delta_2}{\beta_2} > \delta_1 + \beta_2 \cdot \frac{\beta_2 - \delta_2}{\beta_2}$$

If r is large then coexistence occurs.

Theorem 5D. Suppose that

$$\beta_1 \cdot \frac{\delta_2}{\beta_2} < \delta_1 + \beta_2 \cdot \frac{\beta_2 - \delta_2}{\beta_2}$$

If r is large then coexistence is impossible.

Theorems 5A and 5C are due to Durrett and Swindle (1991), while the converse in 5D is due to Durrett and Schinazi (1993). To explain the condition in Theorems 5C and 5D, we begin by observing that if we assumed that $u(t) = P(\xi_t(x) = 2)$ does not depend on x and the states of neighboring sites were independent, then writing $y \sim x$ to denote "y is a neighbor of x"

$$(1.1) \qquad \begin{aligned} \frac{du}{dt} &= -\delta_2 P(\xi_t(x) = 2) + \sum_{y \sim x} \lambda_2 P(\xi_t(x) < 2, \xi_t(x) = 2) \\ &= -\delta_2 u + \beta_2 u(1-u) \end{aligned}$$

where the first equality is true in general and the second follows from our assumptions and the fact that $\beta_2 = |\mathcal{N}|\lambda_2$. Dropping the $-\beta_2 u^2$ term

$$\frac{du}{dt} \le (\beta_2 - \delta_2)u$$

so if $\delta_2 > \beta_2$ all solutions tend to 0 exponentially fast. If $\delta_2 < \beta_2$ and we let $u^* = (\beta_2 - \delta_2)/\beta_2$ then

$$-\delta_2 u + \beta_2 u(1-u) \begin{cases} > 0 & \text{for } 0 < u < u^* \\ < 0 & \text{for } u > u^* \end{cases}$$

so if $u(0) > 0$, $u(t) \to u^*$ as $t \to \infty$.

Applying the reasoning that led to (1.1) to $v(t) = P(\xi_t(x) = 1)$ we see that

$$(1.2) \qquad \begin{aligned} \frac{dv}{dt} &= -\delta_1 P(\xi_t(x) = 1) - \sum_{y \sim x} \lambda_2 P(\xi_t(x) = 1, \xi_t(y) = 2) \\ &\quad + \sum_{y \sim x} \lambda_1 P(\xi_t(x) = 0, \xi_t(y) = 1) \\ &= -\delta_1 v - \beta_2 vu + \beta_1(1-u-v)v \end{aligned}$$

where again the first equality is true in general and the second follows from our assumptions and the fact that $\beta_i = |\mathcal{N}|\lambda_i$. To analyze (1.2), we note that if the 2's are in equilibrium and the density of 1's is very small, then

$$u = (\beta_2 - \delta_2)/\beta_2 \qquad (1 - u - v) \approx \delta_2/\beta_2$$

$$\text{1's are born at rate } \approx \beta_1 \cdot \frac{\delta_2}{\beta_2} \cdot v$$

$$\text{1's die at rate } \approx \left(\delta_1 + \beta_2 \cdot \frac{\beta_2 - \delta_2}{\beta_2}\right) v$$

So if ($\star$) holds a small density of 1's will grow in time, while if we reverse the inequality in ($\star$) and use $(1 - u - v) \le (\beta_2 - \delta_2)/\beta_2$ then the birth rate always exceeds the death rate and $v(t) \to 0$.

The practice of calculating how densities evolve when we suppose that adjacent sites are independent is called *mean field theory*. Theorems 5C and 5D are one instance of the general principle that when the range of interaction is large mean field calculations are almost correct. A second method of making mean field calculations correct, which leads to connections with nonlinear partial differential equations, is to introduce particle motion at a fast rate.

Example 1.6. Predator prey systems. In this model we think of 0 = vacant, 1 = occupied by a fish, and 2 = occupied by a shark and we have the following flip rates

$$\begin{aligned} c_1(x,\xi) &= \beta_1 n_1(x,\xi)/2d \quad \text{if } \xi(x) = 0 \\ c_2(x,\xi) &= \beta_2 n_2(x,\xi)/2d \quad \text{if } \xi(x) = 1 \\ c_0(x,\xi) &= \begin{cases} \delta_1 & \text{if } \xi(x) = 1 \\ \delta_2 + (\gamma n_2(x,\xi)/2d) & \text{if } \xi(x) = 2 \end{cases} \end{aligned}$$

In words, fish die at rate δ_1 and are born at vacant sites at a rate proportional to the number of fish at neighboring sites. So in the absence of sharks, the fish are a contact process.

Sharks die of natural causes at rate δ_2 and kill a neighboring shark at rate $\gamma/2d$. The birth rate for sharks may look a little strange at first: fish turn into sharks at rate proportional to the number of shark neighbors. This is not what happens in the ocean but it does capture an essential feature of the interaction: when the density of fish is too low then the sharks die faster than they give birth. A second justification of this mechanism is that, as we will see in Section 9, in a suitable limit we get standard predator-prey equations.

Here $n_i(x,\xi) = |\{z \in \mathcal{N} : \xi(x+z) = i\}|$ as usual, but for reasons that will become clear in a moment we take $S = \epsilon\mathbf{Z}^d$ and $\mathcal{N} = \{z : |z| = \epsilon\}$ the nearest neighbors. We use a small lattice so that we can introduce *stirring* at a fast rate, i.e., for each $x, y \in \epsilon\mathbf{Z}^d$ with $|x - y| = \epsilon$ we exchange the values at x and y at rate ϵ^{-2}. That is, we change the configuration from ξ to $\xi^{x,y}$ defined by

$$\xi^{x,y}(x) = \xi(y), \quad \xi^{x,y}(y) = \xi(x), \quad \xi^{x,y}(z) = \xi(z) \text{ if } z \ne x, y$$

The combination of the space scale of ϵ and the time scale of ϵ^{-2} means that the individual values will perform Brownian motions in the limit $\epsilon \to 0$. The fast stirring keeps the states of neighboring sites independent, so using mean field reasoning leads to the following result due to DeMasi, Ferrari and Lebowitz (1986).

Theorem 6A. Suppose $\xi_0^\epsilon(x)$, $x \in \epsilon\mathbf{Z}^d$, are independent and let $u_i^\epsilon(t,x) = P(\xi_t(x) = i)$. If $u_i^\epsilon(0,x) = g_i(x)$ is continuous then as $\epsilon \to 0$, $u_i^\epsilon(t,x)$ converges to $u_i(t,x)$ the bounded solution of

$$\begin{aligned} \frac{\partial u_1}{\partial t} &= \Delta u_1 + \beta_1 u_1(1 - u_1 - u_2) - \beta_2 u_1 u_2 - \delta_1 u_1 \\ \frac{\partial u_2}{\partial t} &= \Delta u_2 + \beta_2 u_1 u_2 - \delta_2 u_2 - \gamma u_2^2 \end{aligned} \tag{1.3}$$

with $u_i(0,x) = g_i(x)$.

Here the Δu_i terms reflect the fact that in the limit the individual values are performing Brownian motions run at rate 2. The other terms can be seen by using the reasoning that led to (1.1) and (1.2).

If we suppose that the initial functions $g_i(x)$ are constant then this is true at later times $u_i(t,x) = v_i(t)$ and the v_i satisfy

$$\begin{aligned} \frac{\partial v_1}{\partial t} &= v_1((\beta_1 - \delta_1) - \beta_1 v_1 - (\beta_1 + \beta_2) v_2) \\ \frac{\partial v_2}{\partial t} &= v_2(-\delta_2 + \beta_2 v_1 - \gamma v_2) \end{aligned} \tag{1.4}$$

Here we have rearranged the right hand side to show that it is the standard predator-prey equations with limited growth. (See for example Hirsch and Smale (1974) p. 263.) To determine the conditions for coexistence, we start by finding the fixed points of the dynamical systems, i.e., points (ρ_1, ρ_2) so that $v_i(t) \equiv \rho_i$ is a solution of (1.4). There are three

(i) $\rho_1 = \rho_2 = 0$. No sharks or fish, the trivial equilibrium.

(ii) We have a solution with $\rho_2 = 0$ and $\rho_1 = (\beta_1 - \delta_1)/\beta_1$ if $\beta_1 > \beta_2$. This forumla is the same as the one in the last example because in the absence of sharks, fish are a contact process.

(iii) There is a fixed point with $\rho_i = \sigma_i > 0$ if and only if

$$\frac{\beta_1 - \delta_1}{\beta_1} > \frac{\delta_2}{\beta_2} \tag{1.5}$$

(which implies $\beta_1 > \delta_1$). We do not have an intuitive explanation for the last condition. It is simply what results when we solve the two equations in two unknowns.

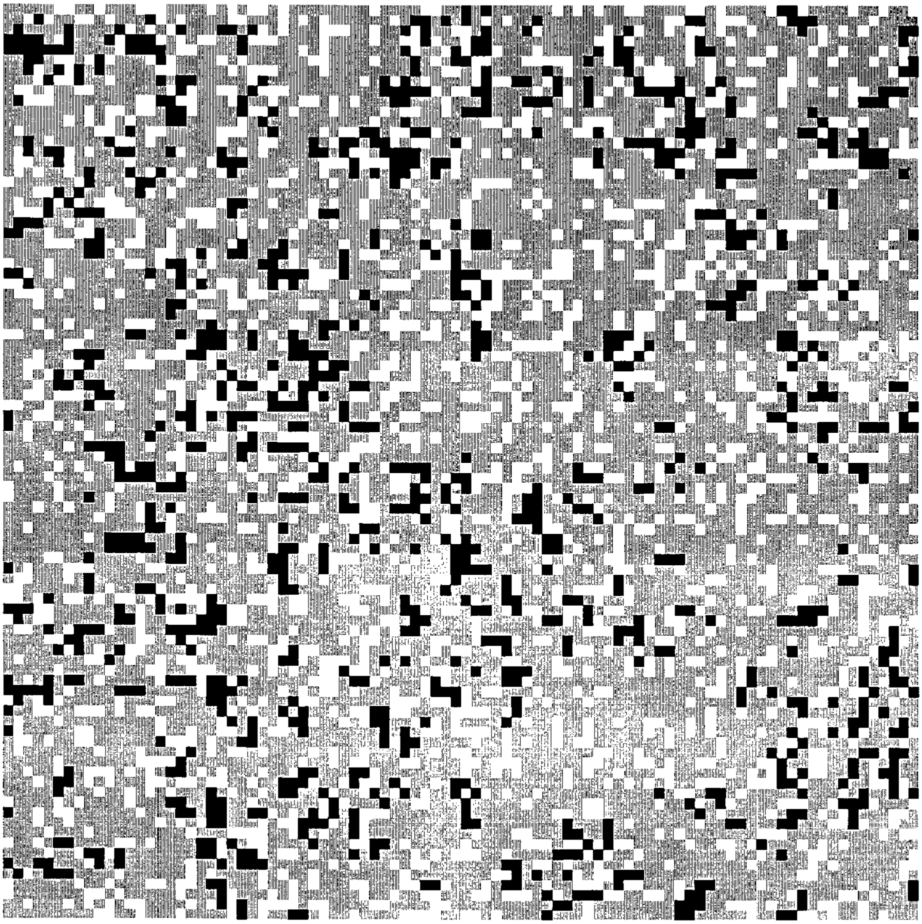


Figure 1.6. Predator prey model $\beta_1 = \beta_2 = 3$, $\delta_1 = \delta_2 = 1$, $\gamma = 1$

By exploiting the connection between the particle system and the partial differential equation given in Theorem 6A, we can prove

Theorem 6B. If (1.5) holds then for small ϵ coexistence occurs.

It would be nice to prove coexistence results without fast stirring. Figure 1.6 shows a simulation of the system on $\{0, 1, \ldots, 79\}^2$ at time 50 with $\beta_1 = \beta_2 = 3$, $\delta_1 = \delta_2 = 1$, $\gamma = 1$ and no stirring. Again sites in state 1 are gray; those in state 2 are black.

Example 1.7. Epidemic model. In this example, we think of $\mathbf{Z}^2$ as representing an array of houses each of which is occupied by one individual who can be (0) susceptible = healthy but capable of getting the disease, (1) infected with the disease, or (2) immune to further infection. The flip rates are

$$
\begin{aligned}
c_1(x,\xi) &= \lambda n_1(x,\xi) \quad \text{if } \xi(x) = 0 \\
c_2(x,\xi) &= \delta \quad \text{if } \xi(x) = 1 \\
c_0(x,\xi) &= \alpha \quad \text{if } \xi(x) = 2
\end{aligned}
$$

As usual, the rates we did not mention are 0. In words, a susceptible individual gets infected at a rate proportional to the number of infected neighbors. Infected individuals become removed at rate δ. Here $1/\delta$ is the mean duration of the disease and to obtain the Markov property we have assumed that the duration of the disease has an exponential distribution. If we want to model the short term behavior of a measles or flu epiemic then we set $\alpha = 0$ since recovered individuals are immune to the disease. If we want to examine longer time properties then immune individuals will die (or move out of town) and new susceptibles will be born (or move into town) so to keep a fixed population size of one individual per site, we combine the two transitions into one.

To describe the conditions for coexistence we begin with case $\alpha = 0$ and consider the behavior of the model starting from one infected individual at 0 in the midst of an otherwise susceptible population. Let $\eta_t = \{x : \xi_t(x) = 1\}$ be the set of the infected individuals at time t and let $\tau = \inf\{t : \eta_t = \emptyset\}$. We will have $\eta_t = \emptyset$ for all $t > \tau$ so we say the infection *dies out* at time τ. Let $\delta_c = \inf\{\delta : P(\tau = \infty) = 0\}$. The faster people recover the harder it is for the epidemic to propagate so we have $P(\tau = \infty) = 0$ for all $\delta > \delta_c$.

If we restrict our attention to the nearest neighbor case, then results of Cox and Durrett (1988) describe the asymptotic behavior of the epidemic when $\delta < \delta_c$ and $\tau = \infty$. Building on those results Durrett and Neuhauser (1991) have shown

Theorem 7. Suppose $d = 2$ and $\mathcal{N} = \{x : |x| = 1\}$. If $\delta < \delta_c$ and $\alpha > 0$ then coexistence occurs.

Zhang has generalized the results of Cox and Durrett (1988) to finite range interactions. Presumably one can also prove the result of Durrett and Neuhauser (1991) in that level of generality but no one has had the courage to try to write out all the details.

Closely related to the epidemic model is

Example 1.8. Greenberg Hastings Model. In this model, we think of having a neuron at each $x \in \mathbf{Z}^d$ that is connected to each of its neighbors. The states of each neuron are $F = \{0, 1, \dots, \kappa - 1\}$ where 1 is excited, $2, \dots, \kappa - 1$ are a sequence of recovery states, and 0 indicates a fully rested neuron that is capable of being excited. These interpretations motivate the following flip rates

$$\begin{aligned} c_1(x,\xi) &= 1 \quad \text{if } \xi(x) = 0 \text{ and } n_i(x,\xi) \geq \theta \\ c_i(x,\xi) &= 1 \quad \text{if } i \neq 1 \text{ and } \xi(x) = i - 1 \end{aligned}$$

Here arithmetic is done modulo κ so $0 - 1 = \kappa - 1$. The second rule says that once excited, the neuron progresses through the recovery states at rate 1 until it is fully rested; the first that a rested neuron becomes excited at rate 1 if the number of its neighbors that are excited is at least the threshold θ. The next result, due to Durrett (1992), gives a regime in which this model has (somewhat boring) stationary distributions.

Theorem 8A. Let $\epsilon > 0$ and suppose $\theta \leq (1 - \epsilon)|\mathcal{N}|/2\kappa$. If r is large then there is a stationary measure close to the uniform product measure.

Here the *uniform product measure* is the one in which the coordinates $\xi(x)$ are independent and $P(\xi(x) = i) = 1/\kappa$. Based on the analogy with the epidemic model where if $\delta < \delta_c$ there is a coexistence for any $\alpha > 0$, we expect that

Conjecture 8B. There is a constant $a > 0$ so that if $\theta \leq a|\mathcal{N}|$ then coexistence occurs for any κ.

Computer simulations indicate that in this regime the excitation sustains itself by producing moving fronts. See Figure 1.7 for a simulation of the system with $\mathcal{N} = \{x : \|x\|_\infty \leq 2\}$, threshold $\theta = 3$, and $\kappa = 8$. Excited states are black, rested sites are white, recovering sites are appropriate shades of gray.

The analogue of Conjecture 8B has been proved by Durrett and Griffeath (1993) for the Greenberg Hastings cellular automaton in which $\xi_{n+1}(x) = \xi_n(x) + 1$ if $\xi_n(x) > 0$ or $\xi_n(x) = 0$ and $n_i(x, \xi_n) \geq \theta$; $\xi_{n+1}(x) = \xi_n(x)$ otherwise. See Figure 1.8 for a simulation of the cellular automaton with the same color scheme and parameters: $\mathcal{N} = \{x : \|x\|_\infty \leq 2\}$, threshold $\theta = 3$, and $\kappa = 8$ run until it has become periodic with period 8. For more on the cellular automaton consult Fisch, Gravner and Griffeath (1991), (1992), (1993), and Gravner and Griffeath (1993).

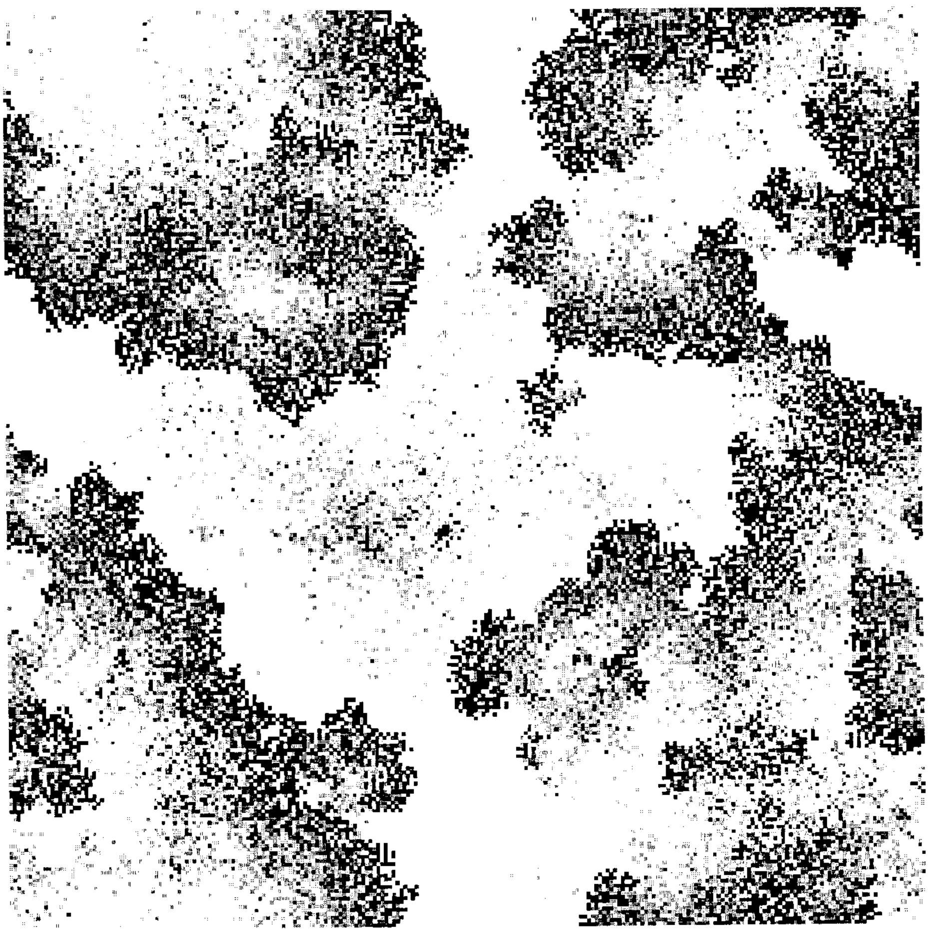


Figure 1.7. Greenberg Hastings model. $\mathcal{N} = \{x : \|x\|_\infty \leq 2\}$, $\theta = 3$, $\kappa = 8$

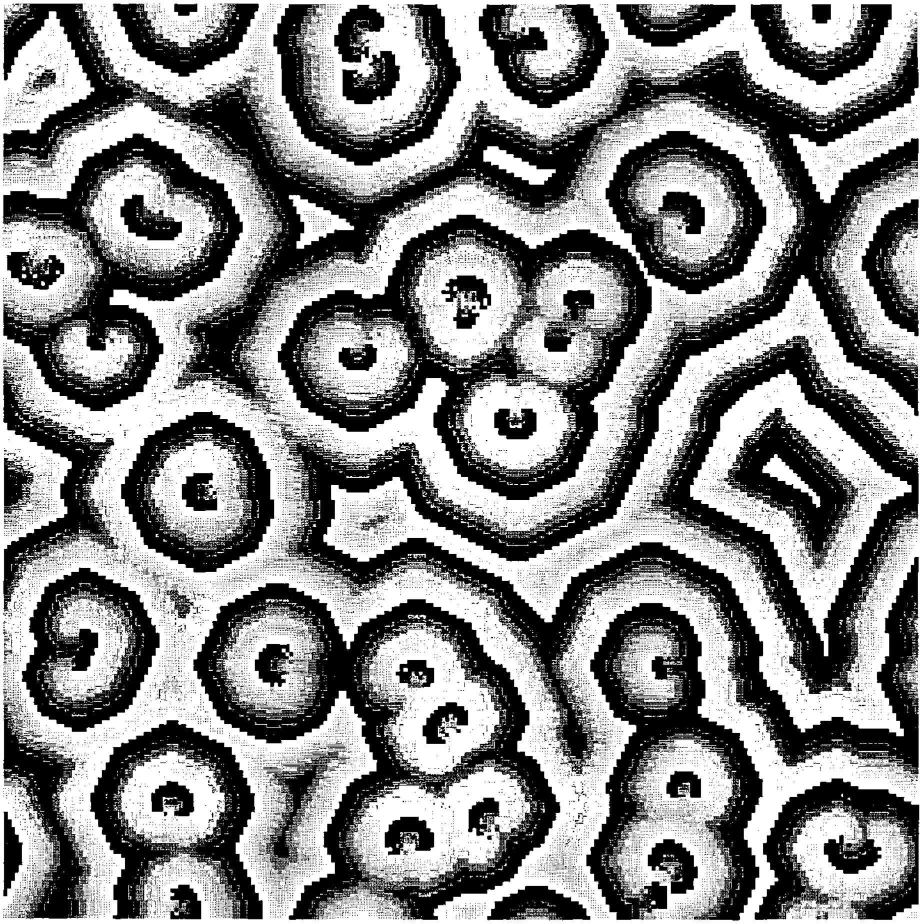


Figure 1.8. Greenberg Hastings cellular automaton. $\mathcal{N} = \{x : \|x\|_\infty \leq 2\}$, $\theta = 3$, $\kappa = 8$

2. Construction, Basic Properties

To construct an interacting particle system from given translation invariant finite flip rates

$$c_i(x,\xi) = g_i(\xi(x+z_0), \xi(x+z_1), \ldots, \xi(x+z_k))$$

based on a neighborhood set $\mathcal{N} = \{z_0, z_1, \ldots, z_n\}$ we can, by changing the time scale, assume that $c_i(x,\xi) \leq 1$. For each $x \in \mathbf{Z}^d$ and $i \in F$, let $\{T_n^{x,i} : n \geq 1\}$ be the arrival times of independent rate 1 Poisson processes (i.e., if we set $T_0^{x,i} = 0$ then the increments $T_n^{x,i} - T_{n-1}^{x,i}$ are independent and have an exponential distribution with mean 1) and let $U_n^{x,i}$ be independent and uniform on (0,1). At time $t = T_n^{x,i}$ site x will flip to state i if $U_n^{x,i} < c_i(x, \xi_{t-})$ and stay unchanged otherwise. To see that this recipe produces the desired flip rates recall the *thinnning property* of the Poisson process: if we keep the points from $\{T_n^{x,i} : n \geq 1\}$ that have $U_n^{x,i} < p$ then the result is a Poisson process with rate p.

Since there are infinitely many Poisson processes, and hence no first arrival, we have to show that we can use our recipe to compute the time evolution. To do this, we use an argument of Harris (1972). Let t_0 be a small positive number to be chosen later. We draw an unoriented arc between x and y if $y - x \in \mathcal{N}$ and for some i, $T_1^{x,i} < t_0$. The presence of an arc between x and y indicates that a Poisson arrival has caused x to look at y to see if it should flip or caused y to look at x. Conversely, if there is no arc between x and y then neither site has looked at the other. The last observation implies that the sites in two different components of the resulting random graph have not influenced each other by time t_0 and hence their evolutions can be computed separately. To finish the construction then it suffices to show

(2.1) **Theorem.** If t_0 is small enough then with probability one, all the connected components of our random graph are finite.

For then in each component there is a first flip and we can compute the effects of the changes sequentially. This allows us to construct the process up to time t_0 but t_0 is independent of the initial configuration, so iterating we can construct the process for all time.

PROOF OF (2.1): Let $\mathcal{N}^* = \{z_1, \ldots, z_k, -z_1, \ldots, -z_k\}$ be the set of possible displacements along edges of the graph. (In this section alone, we will allow $\mathcal{N}$ to be a general finite set not just $\{x : \|x\|_p \leq r\}$.) We say that $y_0, y_1, \ldots y_n$ is a *path* of length n if $y_m - y_{m-1} \in \mathcal{N}^*$ when $0 < m \leq n$. We call a path *self-avoiding* if $y_i \neq y_j$ when $0 \leq i < j \leq n$. Let $R = \max\{|z_i| : z_i \in \mathcal{N}\}$. (Here $|z| = \|z\|_2$.) We claim that

(a) If 0 is connected to some point with $|z| > M$ then there is a self-avoiding path of length $\geq M/R$ starting at 0.

To see this note that if there is a path from 0 to z, then by removing loops we can make it self-avoiding. Since each step along the path moves us a distance $\leq R$, there must be at least $|z|/R$ such steps. The next ingredient in the proof is

(b) If x, y, z, w are distinct, the presence of edges from x to y and from z to w are independent events.

To see this note that the presence of an edge from x to y is determined by the Poisson processes $T_n^{x,i}$ and $T_n^{y,i}$ with $i \in F$. From (b) it is easy to see

(c) Let $N = |\mathcal{N}^*|$ and $\kappa = |F|$. The probability of a self avoiding path of length $2n-1$ starting at a given point x is at most

$$N^{2n-1}(1-e^{-2\kappa t_0})^n$$

The first factor is the number of paths of length $2n-1$ and hence an upper bound on the number of self-avoiding paths. To see the second factor note that the presence of the edges $(z_0,z_1),(z_2,z_3),\dots(z_{2n-2},z_{2n-1})$ are independent events that have probability $1-e^{-2\kappa t_0}$ since the probability of no arrival by time t_0 in one of the 2κ Poisson processes $T_n^{x,i}$ and $T_n^{y,i}$ is $e^{-2\kappa t_0}$.

If we pick t_0 small enough then $N^2(1-e^{-2\kappa t_0}) \le 1/2$, so the probability of a self-avoiding path of length $2n-1$ decreases to 0 exponentially fast, and it follows from (a) that with probability 1 the cluster containing any given point x is finite. □

An immediate consequence of the construction is

(2.2) **Corollary.** If ξ_0 is translation invariant then ξ_t is.

PROOF: The family of Poisson processes is translation invariant, so if the initial state is, then so is the result of our computation. □

It should also be clear from the construction that ξ_t is a Markov process, i.e., if we know the state at time s, information about ξ_r for $r < s$ is irrelevant for computing the evolution for $t > s$. Being a Markov process there is an associated family of operators defined by

$$T_t f(\xi) = E_\xi f(\xi_t)$$

where E_ξ denotes the expected value starting from $\xi_0 = \xi$. The Markov property of ξ_t implies that the T_t form a *semigroup*. That is, $T_s T_t = T_{s+t}$. If you are not familiar with semi-groups don't worry. We will only use the most basic results that can be found in Chapter 1 of Dynkin (1965) or in Chapter ? of Revuz and Yor (1991), and we will only use those facts in this section. The first thing we want to prove is

(2.3) **Corollary.** T_t is a *Feller semigroup*, i.e., if f is continuous with respect to the product topology on F^S then $T_t f$ is continuous.

PROOF: Note that our construction defines on the same probability space the process starting from any initial configuration. If $t \le t_0$ then proof of (2.1) shows that up to time t_0, $\mathbf{Z}^d$ breaks up into a collection of finite non-interacting islands. From the last fact it follows easily that if $\xi_0^n \to \xi_0$, (which means that for each fixed x, $\xi_0^n(x) \to \xi_0(x)$) then $\xi_t^n \to \xi_t$ almost surely. If f is continuous it follows that $f(\xi_t^n) \to f(\xi_t)$ almost surely. Since F^S is compact in the product topology, any continuous function is necessarily bounded, and it follows from the bounded convergence theorem that $Ef(\xi_t^n) \to Ef(\xi_t)$. This proves the result for $t \le t_0$. Using the semigroup property $T_{t+s} = T_s T_t$, it follows that the result holds for $t \le 2t_0$, $t \le 3t_0$, and hence for all t. □

Our next step is to compute the generator of the semigroup. Let $\xi^{x,i}$ denote the configuration ξ flipped to i at x. That is,

$$\xi^{x,i}(x) = i \qquad \xi^{x,i}(y) = \xi(y) \quad \text{otherwise}$$

Suppose $f(\xi)$ only depends on the values of finitely many coordinates and let

$$Lf = \sum_{x \in \mathbf{Z}^d, i \in F} c_i(x,\xi)\left(f(\xi^{x,i}) - f(\xi)\right)$$

The sum converges since only finitely many terms are nonzero. Our next result says that L is the generator of T_t.

$$\text{(2.4)} \qquad \left.\frac{d}{dt} T_t f(\xi)\right|_{t=0} = Lf(\xi)$$

If you have seen the generator of a Markov process with a discrete state space the formula should not be surprising. The proof of (2.4) is much like the proof for that case so we will only give a quick sketch.

PROOF: Suppose f only depends on the values of ξ in $[-L, L]^d$ and recall we have defined $R = \max\{|z_i| : z_i \in \mathcal{N}\}$. If t is small then with high probability there is at most one site $x \in [-L - 2R, L + 2R]^d$ and one value of $i \in F$ with $T^{x,i} < t$. By considering the various possible values of x and i and noting that the probability that $\xi_0 = \xi$ changes to $\xi^{x,i}$ is $\sim tc_i(x,\xi)$, the result follows easily. □

For the rest of this section, we will restrict our attention to the case $F = \{0,1\}$, in which case we think of $1 =$ occupied by a particle and $0 =$ vacant. Since we think of 1's are particles we call $c_1(x,\xi)$ the *birth rates* and call $c_0(x,\xi)$ the *death rates* . We say that the birth rates $c_1(x,\xi)$ are *increasing* if

$$\xi(y) \le \zeta(y) \text{ for all } y \ne x \text{ and } \xi(x) = \zeta(x) = 0 \text{ implies } c_1(x,\xi) \le c_1(x,\zeta)$$

We say that *death rates* $c_0(x,\xi)$ are *decreasing* if

$$\xi(y) \le \zeta(y) \text{ for all } y \ne x \text{ and } \xi(x) = \zeta(x) = 1 \text{ implies } c_1(x,\xi) \ge c_1(x,\zeta)$$

A process with increasing birth rates and decreasing death rates is said to be *attractive.* The last term comes from analogies with the Ising model in statistical mechanics. This assumption is not very attractive for biological systems since there the death rate usually increases due to crowding, but the attractive property is what we need to prove the following useful result.

(2.5) **Theorem.** For an attractive process, if we are given initial configurations with $\xi_0(x) \le \zeta_0(x)$ for all x then the processes defined by our construction have $\xi_t(x) \le \zeta_t(x)$ for all x and t.

PROOF: Intuitively this is true since each flip preserves the inequality. To check this suppose that $\xi_{s-}(y) = 0$ and a birth event $T_n^{y,1}$ occurs at time s. If $\zeta_{s-}(y) = 1$ then $\zeta_s(y) = 1$ and the inequality will certainly hold after the flip. If $\zeta_{s-}(y) = 0$ and the inequality holds before the flip, then since our birth rates are increasing $c(y, \xi_{s-}) \le c(y, \zeta_{s-})$. By considering the possible values of $U_n^{y,i}$ se see that in all cases the inequality holds after the flip.

value of $U_n^{y,i}$	change in ξ	change in ζ
$[0, c(y, \xi_{s-}))$	flips to 1	flips to 1
$[c(y, \xi_{s-}), c(y, \zeta_{s-}))$	stays 0	flips to 1
$[c(y, \zeta_{s-}), 1)$	stays 0	stays 0

A similar argument applies if $\xi_{s-}(y) = 1$ and a death event $T_n^{y,0}$ occurs at time s.

To turn the intuitive argument in the last paragraph into a proof, suppose that the inequality fails at some point x at some time $t \le t_0$. Let $\mathcal{C}_x$ be the connected component containing x for the random graph defined in the proof of (3.1), and let $s > 0$ be the first time the property fails at some point $y \in \mathcal{C}_x$. By the definition of s the inequality holds on $\mathcal{C}_x$ before time s. Since $\mathcal{C}_x$ contains all the neighbors of any site in $\mathcal{C}_x$ that flips by time t_0 it follow from the argument in the last paragraph that the inequality will hold until the next flip after time s. Since $\mathcal{C}_x$ is a finite set, the next flip will occur at a time $> s$, contradicting the defintion of s and showing that the inequality must hold up to time t_0. Iterating the last conclusion we see that the result holds for all time. □

To explain our interest in (2.2), (2.4), and (2.5) we will now prove that

(2.6) **Theorem.** If $\lambda|\mathcal{N}| < \delta$ then the contact process has no nontrivial stationary distribution.

PROOF: Consider the contact process starting from all sites occupied, i.e., suppose $\xi_0^1(x) = 1$ for all x. It follows from (2.2) that $P(\xi_t^1(x) = 1)$ is independent of x, so writing $y \sim x$ for "y is a neighbor of x" and using $\frac{d}{dt} T_t f = T_t L f$ we have

$$\begin{aligned} \frac{d}{dt} P(\xi_t^1(x) = 1) &= -\delta P(\xi_t^1(x) = 1) + \sum_{y \sim x} \lambda P(\xi_t^1(x) = 0, \xi_t^1(y) = 1) \\ &\le -\delta P(\xi_t^1(x) = 1) + \lambda |\mathcal{N}| P(\xi_t^1(y) = 1) \end{aligned}$$

If $\lambda|\mathcal{N}| < \delta$ then the last inequality implies that $P(\xi_t^1(x) = 1) \to 0$ as $t \to \infty$. Now any initial configuration has $\xi_0(x) \le 1 = \xi_0^1(x)$ for all x, so by (2.5), we have $\xi_t(x) \le \xi_t^1(x)$ for all t and x and it follows that $P(\xi_t(x) = 1) \to 0$ for any initial configuration. If we pick ξ_0 to have a stationary distribution then $P(\xi_t(x) = 1)$ is independent of t, so the last conclusion implies this probability is 0 and the result follows. □

The last argument shows that if we start an attractive process with all sites occupied and find $P(\xi_t^1(x) = 1) \to 0$ then there is no nontrivial stationary distribution. Our next result proves the converse. Recall that $\Rightarrow$ denotes weak convergence, which in this setting is just convergence of finite dimensional distribution.

(2.7) **Theorem.** As $t \to \infty$, $\xi_t^1 \Rightarrow \xi_\infty^1$. The limit is a stationary distribution which is stochastically larger than any other stationary distribution and called the *upper invariant measure.*

PROOF: The key to the proof is the following observation:

(2.8) **Lemma.** For any set $A \subset \mathbf{Z}^d$, $t \to P(\xi_t^1(x) = 0 \text{ for all } x \in A)$ is increasing.

PROOF: Let $\zeta_0 = \xi_s^1$. Clearly, $\xi_0^1(x) \ge \zeta_0(x)$ for all x so (2.5) implies that for all t and x, $\xi_t^1(x) \ge \zeta_t(x)$. Since ζ_t has the same distribution as ξ_{s+t}^1 it follows that

$$P(\xi_t^1(x) = 0 \text{ for all } x \in A) \le P(\xi_{s+t}^1(x) = 0 \text{ for all } x \in A) \qquad \square$$

Let $\phi(A) = P(\xi(x) = 0 \text{ for all } x \in A)$ and $B = \{x_1, \dots, x_m\}$ Using the inclusion exclusion formula on the events $E_i = \{\xi_t(x) = 0\}$ on $A \cup \{x_i\}$, we can express any finite dimensional distribution in terms of the $\phi(C)$.

$$\begin{aligned} 1 - P(\xi(x) = 0 \text{ for all } x \in A, \xi(x) = 1 \text{ for all } x \in B) &= P(\cup_{i=1}^m E_i) \\ &= \sum_{i=1}^m \phi(A \cup \{x_i\}) - \sum_{i<j} \phi(A \cup \{x_i, x_j\}) + \dots + (-1)^{m+1}\phi(A \cup B) \end{aligned}$$

So (2.8) implies convergence of all finite dimensional distributions. $\square$

The fact that ξ_∞^1 is a stationary distribution follows from a general result.

(2.9) **Lemma.** Suppose the Markov process X has a Feller semigroup and $X_t \Rightarrow X_\infty$ then (the distribution of) X_∞ is a stationary distribution.

PROOF: Recall that if X_0 has distribution μ then the probability measure μT_t defined by

$$\int (\mu T_t)(dx) f(x) = \int \mu(dx) T_t f(x) = \int \mu(dx) E_x f(X_t)$$

for all bounded continuous functions f gives the distribution of X_t when X_0 has distribution μ. The key to the proof of (2.9) is the following general fact:

(2.10) If T_t is a Feller semigroup and $\mu_s \Rightarrow \mu$ then $\mu_s T_t \Rightarrow \mu T_t$.

To prove (2.10) we note that $T_t f$ is bounded and continuous

$$\begin{aligned} \lim_{s\to\infty} \int (\mu_s T_t)(dx) f(x) &= \lim_{s\to\infty} \int \mu_s(dx) T_t f(x) \\ &= \int \mu(dx) T_t f(x) = \int (\mu T_t)(dx) f(x) \end{aligned}$$

where the second inequality follows from the fact that $T_t f$ is continuous and $\mu_s \Rightarrow \mu$. To prove (2.9) now, let μ_s be the distribution of X_s and note that the Markov property implies $\mu_s T_t = \mu_{s+t}$. The right hand side converges to μ, and by (2.10) the left hand side converges to μT_t, so $\mu T_t = \mu$, i.e., μ is a stationary distribution. □

Finally we have to explain and show the claim "ξ^1_∞ is stochastically larger than any other stationary distribution π." By *stochastically larger* we mean that if f is any increasing function which depends on only finitely many cooordinates then

$$E f(\xi^1_\infty) \geq \int f(\xi) d\pi(\xi) \tag{2.11}$$

Here f is *increasing* means that if $\xi(x) \leq \zeta(x)$ for all x then $f(\xi(x)) \leq f(\zeta(x))$. To prove the claim let ζ_0 have distribution π. Clearly, $\xi^1_0(x) \geq \zeta_0(x)$ for all x so (3.5) implies that $\xi_t(x) \geq \zeta_t(x)$ for all t and x. Now if f is increasing

$$E f(\xi_t) \geq E f(\zeta_t) = \int f(\xi) d\pi(\xi)$$

since π is a stationary distribution. If f depends on only finitely many coordinates then it is continuous and

$$E f(\xi^1_t) \to E f(\xi^1_\infty)$$

Combining the last two conclusions, proves our claim and completes the proof of (3.7). □

(2.12) **Remark.** A result of Holley implies that since ξ^1_∞ is stochastically larger than π, we can define random variables ξ and ζ with these distributions on the same probability space so that $\xi(x) \geq \zeta(x)$.

Later we will need a variation of (2.9). The next result and (3.15) are not needed until Section 5, so I suggest that you wait until later to read the rest of this section.

(2.13) **Theorem.** Suppose the Markov process X has a compact state space Λ and a Feller semigroup T_t. Let μ_t be the distribution of X_t and ν_t the Cesaro average defined by

$$\nu_t(A) = \frac{1}{t} \int_0^t \mu_s(A)$$

If $t_k \to \infty$ and $\nu_{t_k} \Rightarrow \nu$ then ν is a stationary distribution.

(2.14) **Corollary.** Since the set of probability measures on Λ is compact in the weak topology, this implies in particular that stationary distributions exist.

PROOF: Since $\mu_s T_r = \mu_{s+r}$ we have

$$\begin{aligned} \nu_{t_k} T_r &= \frac{1}{t_k} \int_0^{t_k} \mu_s T_r ds = \frac{1}{t_k} \int_r^{r+t_k} \mu_s ds \\ &= \nu_{t_k} + \frac{1}{t_k} \int_{t_k}^{r+t_k} \mu_s ds - \frac{1}{t_k} \int_0^r \mu_s ds \end{aligned}$$

The two error terms on the right hand side have each total mass r/t_k and hence converge weakly to 0. Since $\nu_{t_k} \Rightarrow \nu$ it follows that $\nu_{t_k} T_r \Rightarrow \nu$. On the other hand it follows from (3.10) that $\nu_{t_k} T_r \Rightarrow \nu T_r$ so we have $\nu T_r = \nu$ as desired. □

In Section 5, we will also need the following result:

(2.15) **Theorem.** The upper invariant measure ξ^1_∞ is spatially ergodic.

PROOF: We begin with the observation that

(2.16) for each t, ξ^1_t is spatially ergodic.

To prove (2.16) we let $V^x = (\{T^{x,i}_n, n \geq 1\}, \{U^{x,i}_n, n \geq 1\}, i = 0, \ldots, \kappa - 1)$. $\{V^x, x \in \mathbf{Z}^d\}$ are i.i.d. and $\xi_t(x)$ is a function of the V_x so the result follows from a generalization of (1.3) in Chapter 6 of Durrett (1992). In words, functions of ergodic sequences are ergodic.

To let $t \to \infty$, we note that the proof of (2.8) shows ξ^1_t is stochastically larger than ξ^1_∞ so (2.12) implies that we can construct the two processes on the same space so that $\xi^1_t(x) \geq \xi^1_\infty(x)$ for all x. Let f be an increasing function that depends on only finitely many coordinates. The ergodic theorem implies that as $L \to \infty$

$$\frac{1}{(2L+1)^d} \sum_{x:\|x\|_\infty \leq L} \xi^1_t(x) \to Ef(\xi_t)$$
$$\frac{1}{(2L+1)^d} \sum_{x:\|x\|_\infty \leq L} \xi^1_\infty(x) \to E(f(\xi_\infty)|\mathcal{I})$$

The last result and our comparison imply that $E(f(\xi_\infty)|\mathcal{I}) \leq Ef(\xi_t)$ where $\mathcal{I}$ is the σ-field of shift invariant events. Letting $t \to \infty$ we have $E(f(\xi_\infty)|\mathcal{I}) \leq Ef(\xi_\infty)$ and since the left hand side has expected value $Ef(\xi_\infty)$, it follows that

$$E(f(\xi_\infty)|\mathcal{I}) = Ef(\xi_\infty) \qquad \text{a.s.} \tag{2.17}$$

At this point we have shown that (2.17) holds for increasing functions that depends on only finitely many coordinates. Now every function on $\{0,1\}^k$ is a difference of two increasing functions so (2.17) holds for any function of finitely many coordinates. Taking limits and using the inequality

$$E|E(X - Y|\mathcal{I})| \leq E(|X - Y||\mathcal{I}) = E|X - Y|$$

shows that (2.17) holds for all bounded f so $\mathcal{I}$ is trivial. □

3. Percolation Substructures, Duality

In this section we introduce a variation of the construction used in Section 2, due to Harris (1976) and Griffeath (1979), which applies to a special class of models with state space $\{0,1\}^S$ and leads to a "duality relationship." For these purposes it is convenient to write our systems as set valued processes in which the state at time t is the set of sites occupied by 1's. We begin with

Example 3.1. The basic contact process. We let $\mathcal{N}$ be a finite set of neighbors of 0, say that y is a neighbor of x if $y - x \in \mathcal{N}$, and formulate the dynamics as follows:

(i) Particles die at rate 1.

(ii) A particle is born at a vacant site x at rate λ times the number of occupied neighbors.

To construct the process we introduce independent Poisson processes $\{U_n^x, n \geq 1\}$ with rate 1 and $\{T_n^{x,y}, n \geq 1\}$ with rate λ for each $x, y \in \mathbf{Z}^d$ with $y - x \in \mathcal{N}$. At the space time points (x, U_n^x) we write a δ to indicate that a death will occur if x is occupied, and we draw an arrow from $(y, T_n^{x,y})$ to $(x, T_n^{x,y})$ to indicate that if y is occupied then there will be a birth from y to x.

Given the Poisson processes and forgetting about the special marks, we could construct the process using the algorithm described in the last section. We introduce the special marks to make contact with percolation: we imagine fluid entering the bottom of the picture at the points in ξ_0 and flowing up the structure. The δ's are dams, the arrows are pipes that allow the fluid to flow in the direction of the arrow, and ξ_t is the set of sites that are wet at time t.

An example of the *percolation substructure* and the corresponding realization of ξ_t starting from $\xi_0 = \{0,1\}$ is given in Figure 3.1. The thick lines indicate the sites that are occupied. To be able to define the dual process, we need an explicit recipe for constructing ξ_t from the picture. We say that there is a *path from* $(x,0)$ *to* (y,t) if there is a sequence of times $s_0 = 0 < s_1 < s_2 < s_n < s_{n+1} = t$ and spatial locations $x_0 = x, x_1, \ldots, x_n = y$ so that

(i) for $i = 1, 2, \ldots, n$ there is an arrow from x_{i-1} to x_i at time s_i

(ii) the vertical segments $\{x_i\} \times (s_i, s_{i+1})$, $i = 0, 1, \ldots n$ do not contain any δ's.

(Exercise: Find a path from $(2,0)$ to $(3,t)$ in Figure 3.1.) Intutitively the arrows are births that will occur if there are no δ's in the intervals in (ii), so to define the process starting from $\xi_0^A = A$ we let

$$\xi_t^A = \{y : \text{ for some } x \in A \text{ there is a path from } (x,0) \text{ to } (y,t)\} \tag{3.1}$$

It should be clear from the definitions that ξ_t^A is the contact process with one small modification: because of the open intervals in (ii) and the strict inequality in $s_n < s_{n+1} = t$, the process we have constructed is left continuous. For example, if there is a death at x at time t, the particle will not be dead at time t but it will be dead at time $t + \epsilon$ when ϵ

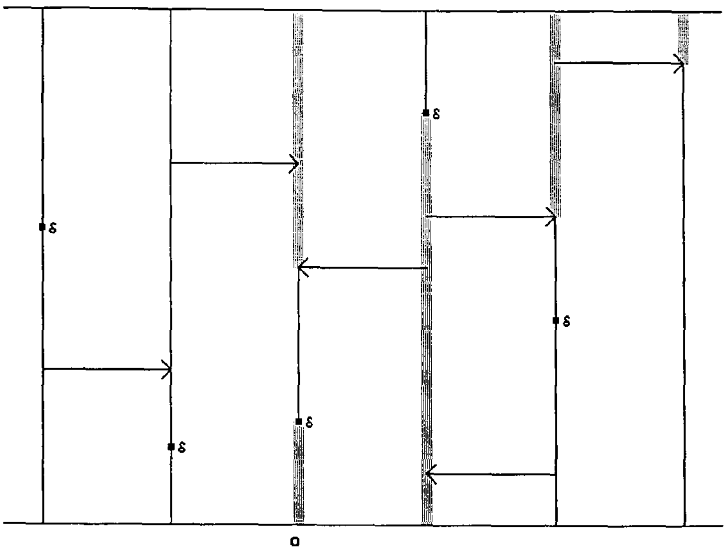


Figure 3.1. Contact process

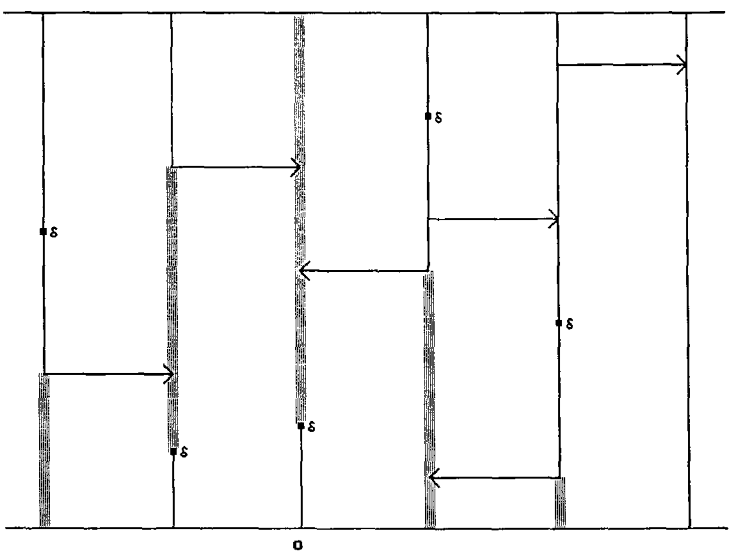


Figure 3.2. Dual of the contact process

is small.

Although left continuous versions of Markov processes are not the traditional ones, we will tolerate them in this section since our main goal is to define the dual process and derive the duality relation (3.2), which is a statement about the one dimensional distributions. (Note that there are only countably many jumps so the left and right continuous versions are equal almost surely at any fixed t.) To construct the dual process starting from time t, we say that there is a *path down from* (y,t) *to* $(x,t-r)$ if there is a sequence of times $s_0 = 0 < s_1 < s_2 < s_n < s_{n+1} = r$ and spatial locations $x_0 = y, x_1, \dots, x_n = x$ so that

(i) for $i = 1, 2, \dots, n$ there is an arrow from x_i to x_{i-1} at time $t - s_i$

(ii) the vertical segments $\{x_i\} \times (t - s_{i+1}, t - s_i)$, $i = 0, 1, \dots n$ do not contain any $\delta's$.

That is, we have to avoid δ's as before but this time we move across arrows in a direction opposite to their orientation. (Exercise: Find a path down from $(3,t)$ to $(2,0)$ in Figure 3.1.)

The last definition is chosen so that there is a path from $(x,0)$ to (y,t) if and only if there is a path down from (y,t) to $(x,0)$ and hence if we define

$$\hat{\xi}_s^{(B,t)} = \{x : \text{ for some } y \in B \text{ there is a path down from } (y,t) \text{ to } (x,t-s)\} \tag{3.2}$$

then $\{\xi_t^A \cap B \neq \emptyset\} = \{A \cap \hat{\xi}_t^{(B,t)} \neq \emptyset\}$. With a little more thought one sees that for any $0 \le s \le t$

$$\{\xi_t^A \cap B \neq \emptyset\} = \{\xi_s^A \cap \hat{\xi}_{t-s}^{(B,t)} \neq \emptyset\} = \{A \cap \hat{\xi}_t^{(B,t)} \neq \emptyset\} \tag{3.3}$$

Figure 3.2 shows a picture of the dual process $\hat{\xi}_s^{(\{0\},t)}$. To work with the dual, it is useful to define a process $\hat{\xi}_s^B$ so that for each t, $\{\hat{\xi}_s^B; 0 \le s \le t\}$ has the same distribution as $\{\hat{\xi}_s^{(B,t)} : 0 \le s \le t\}$. Comparing the definition of the original process and the dual shows that we can do this by reversing the direction of the arrows in the original percolation substructure and then applying the original definition. From this observation it should be clear that if ξ_t^A is a contact process with neighborhood set $\mathcal{N}$ then $\hat{\xi}_t^B$ is a contact process with neighborhood set $-\mathcal{N} = \{-x : x \in \mathcal{N}\}$. So if we use our favorite neighborhood $\mathcal{N} = \{x : \|x\|_p \le r\}$ then the contact process is *self-dual*, i.e., $\{\hat{\xi}_t^B, t \ge 0\}$ and $\{\xi_t^B, t \ge 0\}$ have the same distribution.

Example 3.2. The voter model. Recall that our simple minded voters have two opinions 0 or 1, and that a voter at x changes her opinion at a rate equal to the number of neighbors (i.e., y with $y - x \in \mathcal{N}$) with the opposite opinion. To make the percolation substructure we let $\{U_n^{x,y} : n \ge 1\}$ be independent Poisson processes with rate 1 when $x, y \in \mathbf{Z}^d$ with $y - x \in \mathcal{N}$, we draw an arrow from $(y, U_n^{x,y})$ to $(x, U_n^{x,y})$ and write a δ at $(x, U_n^{x,y})$. We define paths as before and use the paths to define a set valued process in which the state at time t is the set of sites with opinion 1. Writing 1 for occupied and 0 for vacant and thinking about the defintion it is easy to see that the effect of an "arrow-delta" from y to x is as follows:

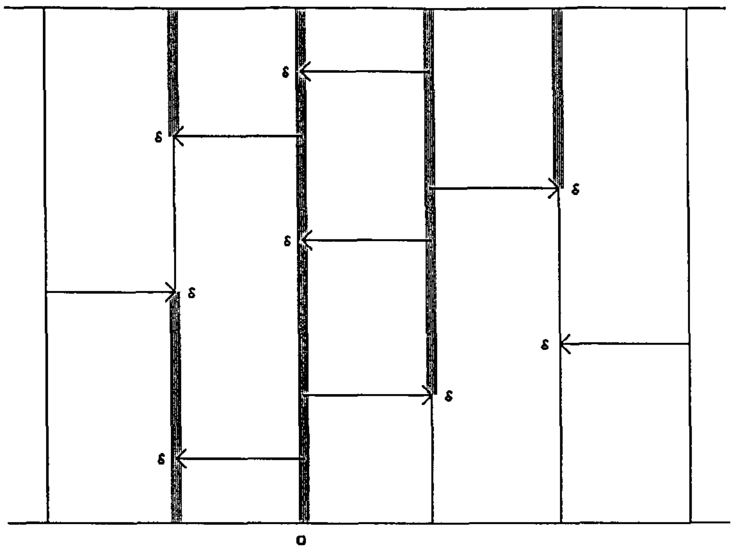


Figure 3.3. Voter model

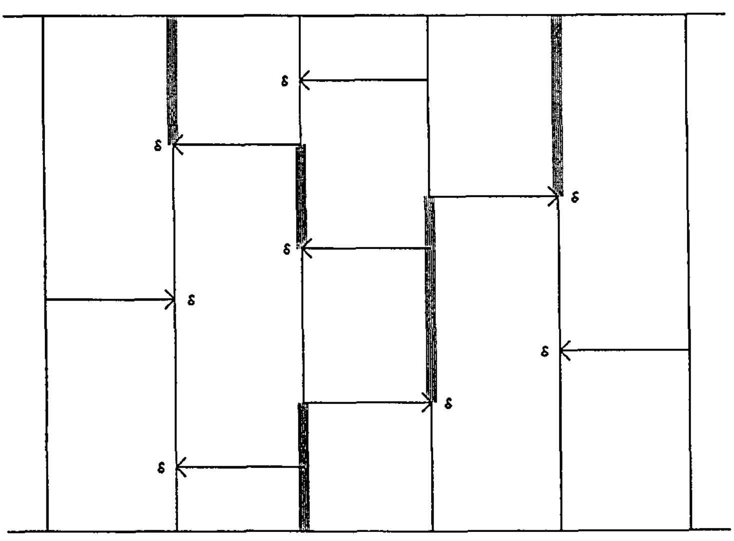
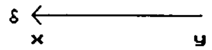


Figure 3.4. Dual of the voter model



before		after	
x	y	x	y
0	0	0	0
1	0	0	0
0	1	1	1
1	1	1	1

In words, because of the δ at x, x will occupied after the "arrow-delta" if and only if y is occupied. From the table (or from the verbal description) we see that the effect of an "arrow-delta" from y to x is to force the voter at x to imitate the voter at y, so the process defined by (3.1) is the voter model. Figure 3.3 gives an example of the construction with $\xi_0 = \{-1, 0\}$. Again the thick lines indicate occupied sites.

The motivation for this construction is that it allows us to define a dual process which in the case of the voter model is quite simple. Since dual paths cannot continue through δ's and can only move across arrows in a direction opposite their orientation, it is easy to check that $\hat{\xi}_s^{(\{x\},t)}$ is always a single site $S_s^{x,t}$, which has the interpretation that the voter at x at time t has the same opinion of the voter at $S_s^{x,t}$ at time $t-s$. See Figure 3.4 which shows $\hat{\xi}_s^{(\{x\},t)}$ for $x = -1$ and $x = 2$. In words, $S_s^{x,t}$ sits at a site y until $t - s = U_n^{y,z}$ for some z, indicating the voter at y imitated the one at z, at which time $S_s^{x,t}$ jumps from y to z. From the last description it should be clear that $S_s^{x,t}$ is a continuous time random walk that for each $w \in \mathcal{N}$ jumps from y to $y + w$ at rate 1.

To determine the behavior of the dual starting from more than one point, we note that it is constructed from a percolation structure with independent Poisson processes $\{U_n^{x,y} : n \geq 1\}$ for $x, y \in \mathbf{Z}^d$ with $y - x \in \mathcal{N}$ at which time we draw an arrow from $(x, U_n^{x,y})$ to $(y, U_n^{x,y})$ and write a δ at $(x, U_n^{x,y})$. From the definition it is easy to see that a "delta-arrows" from x to y has the following effect

before		after	
x	y	x	y
0	0	0	0
1	0	0	1
0	1	0	1
1	1	0	1

The δ at x makes it vacant while the arrow from x to y will make y occupied if there was a particle at y or at x. These are the transitions of a *coalescing random walk*. Particles move independently until they hit and then move together after that. The duality relationship (3.3) between the voter model and coalescing random walks leads easily to the results of Holley and Liggett (1975). These conclusions are true quite generally but we will state them only for our favorite neighborhoods $\{z : \|z\|_p \leq r\}$ with $r \geq 1$. To make the statements here match Theorems 2A and 2B in Section 1, we revert to coordinate notation: $\xi_t(x) = 1$ if and only if $x \in \xi_t$.

Theorem 3.1. *Clustering* occurs in $d \leq 2$. That is, for any ξ_0 and $x, y \in \mathbf{Z}^d$ we have

$$P(\xi_t(x) \neq \xi_t(y)) \to 0 \text{ as } t \to \infty$$

Theorem 3.2. Let ξ_t^θ denote the process starting from an initial state in which the events $\{\xi_0^\theta(x) = 1\}$ are independent and have probability θ. In $d \geq 3$ as $t \to \infty$, $\xi_t^\theta \Rightarrow \xi_\infty^\theta$, a translation invariant stationary distribution in which $P(\xi_\infty(x) = 1) = \theta$.

PROOF OF THEOREM 3.1. From our discussion of the dual it should be clear that

$$P(\xi_t(x) \neq \xi_t(y)) \leq P(S_t^{(x,t)} \neq S_t^{(y,t)})$$

since if the two sites x and y trace their opinions back to the same site at time 0 then they will certainly be equal at time t. Now the difference $S_s^{(x,t)} - S_s^{(y,t)}$ is a random walk stopped when it hits 0, and the random walk has jumps that have mean 0 and finite variance. Such random walks are *recurrent*, and since ours is also an irreducible Markov chain, it will eventually hit 0. Since 0 is an absorbing state for $S_s^{(x,t)} - S_s^{(y,t)}$ it follows that $P(S_t^{(x,t)} \neq S_t^{(y,t)}) \to 0$ and the proof is complete. □

Remark. The reader should not misinterpret Theorem 3.1 as saying that the voter model is boring in $d \leq 2$. Cox and Griffeath (1986) have proved a number of interesting results about the clustering in $d = 2$, which is rather exotic since two dimensional random walk is just barely recurrent.

PROOF OF THEOREM 3.2. From the proof of (2.8) we see that it is enough to prove the convergence of $P(\xi_t \cap B = \emptyset)$ for each B. To treat these probabilities we observe that

$$P(\xi_t \cap B = \emptyset) = E\{(1-\theta)^{|\hat{\xi}_t^{(B,t)}|}\}$$

since by duality there are no particles in B at time t if and only if none of the sites in $\hat{\xi}_t^{(B,t)}$ is occupied at time 0, an event with probability $(1-\theta)^{|\hat{\xi}_t^{(B,t)}|}$. To analyze the right hand side we note that $\hat{\xi}_t^{(B,t)}$ has the same distribution as $\hat{\xi}_t^B$ constructed from the percolation substructure that has the directions of all the arrows reversed. Since $\hat{\xi}_t^B$ is a coalescing random walk, $|\hat{\xi}_t^B|$ is a decreasing function of t and has a limit. Since $0 \leq (1-\theta)^{|\hat{\xi}_t^B|} \leq 1$ it follows from the bounded convergence theorem that

$$\lim_{t\to\infty} E\{(1-\theta)^{|\hat{\xi}_t^{(B,t)}|}\} \text{ exists}$$

and the proof is complete. □

Since the ξ_t^θ are translation invariant (by (2.2)), it follows that the limits ξ_∞^θ are.

$$P(x \in \xi_t^\theta) = P(S_t^{x,t} \in \xi_0^\theta) = \theta$$

for all t so $P(x \in \xi_\infty^\theta) = \theta$. Holley and Liggett (1975) showed that the ξ_∞^θ are spatially ergodic and give all the stationary distributions for the voter model. That is, all stationary distributions are a convex combination of the (distributions of the) ξ_∞^θ. For proofs of this result see the original paper by Holley and Liggett (1975) or Chapter V of Liggett (1985).

Using duality we can prove a convergence theorem due to Harris (1976) for a general class of processes that contains the contact process as a special case. We begin by introducing the models we will consider.

Additive processes. For each finite $A \subset \mathbf{Z}^d$ and $x \in \mathbf{Z}^d$ we introduce independent Poisson processes $\{T_n^{x,A}, n \geq 1\}$ and $\{U_n^{x,A}, n \geq 1\}$ with rates $\lambda(A)$ and $\delta(A)$. (To have a finite range interaction, we only allow finitely many of the rates to be nonzero.) At times $T_n^{x,A}$ we draw arrows from $x+z$ to x for all $z \in A$ and there will be a birth if some site in $x+A$ is occupied. At times $U_n^{x,A}$ we write a δ at x, draw arrows from $x+z$ to x for all $z \in A$, and there will be a death at x unless some point in $x+A$ is occupied. The process is then obtained from the percolation substructure by using (3.1). In the new notation our two examples may be written as (the rates we do not mention are 0):

The contact process. $\lambda(A) = \lambda$ if $A = \{x\}$ with $x \in \mathcal{N}$; $\delta(\emptyset) = 1$.

The voter model. $\delta(A) = 1$ if $A = \{x\}$ with $x \in \mathcal{N}$.

It should be clear that for any additive process the birth rates are increasing and the death rates are decreasing so these systems are attractive. To see that additive processes are a fairly small subclass of the attractive models, we will now consider

Example 3.3. Nonlinear Contact Processes. In these systems the flip rates are

$$c_0(x,\xi) = 1$$
$$c_1(x,\xi) = b(|\{y \in \mathcal{N} : \xi(x+y) = 1\}|)$$

where $b(0) = 0$. To get the desired death rates we set $\delta(\emptyset) = 1$ and $\delta(A) = 0$ otherwise. To see what birth rates we can create we begin with the special case

(i) $d = 1$, $\mathcal{N} = \{-1, 1\}$. In this situation we must have

$$\lambda(\{1\}) = \lambda(\{-1\}) = a_1 \qquad \lambda(\{1,-1\}) = a_2$$

and the other $\lambda(A) = 0$, so $b(1) = a_1 + a_2$ and $b(2) = 2a_1 + a_2$ which is possible with $a_1, a_2 \geq 0$ if and only if

$$b(1) \leq b(2) \leq 2b(1)$$

The extreme case $b(2) = 2b(1)$ is the basic contact process, the other extreme $b(2) = b(1) = b$ is called the *threshold contact process* because the birth rate is b if there is at least one occupied neighbor. An example of a system not covered by this construction is the *sexual reproduction model* which has $b(1) = 0$ and $b(2) = \lambda$.

(ii) Suppose $|\mathcal{N}| = 4$ and think about $\mathcal{N} = \{-2,-1,1,2\}$ in $d = 1$ or $\mathcal{N} = \{z : \|z\|_1 = 1\}$ in $d = 2$. (The geometry of the set $\mathcal{N}$ does not enter into the decision as to whether or not a system is additive.) In this case $\lambda(A) = a_i$ if $A \subset \mathcal{N}$ with $|A| = i$ (and 0 otherwise) so

$$\begin{aligned} b(1) &= a_1 + 3a_2 + 3a_3 + a_4 \\ b(2) &= 2a_1 + 5a_2 + 4a_3 + a_4 \\ b(3) &= 3a_1 + 6a_2 + 4a_3 + a_4 \\ b(4) &= 4a_1 + 6a_2 + 4a_3 + a_4 \end{aligned}$$

To see the equation of $b(2)$ say, note that any two element subset of $\mathcal{N}$ touches 2 of the singleton subsets of $\mathcal{N}$, all but one of the 6 two element subsets, all 4 of the three element subsets, and the four element subset. Subtracting the equations gives

$$\begin{aligned} b(4) - b(3) &= a_1 \\ b(3) - b(2) &= a_1 + a_2 \\ b(2) - b(1) &= a_1 + 2a_2 + a_3 \\ b(1) - b(0) &= a_1 + 3a_2 + 3a_3 + a_4 \end{aligned}$$

and taking differences again

$$\begin{aligned} a_1 &= b(4) - b(3) \\ a_2 &= (b(3) - b(2)) - (b(4) - b(3)) \\ a_3 &= (b(2) - b(1)) - 2(b(3) - b(2)) + (b(4) - b(3)) \\ a_4 &= ((b(1) - b(0)) - 3(b(2) - b(1)) + 3(b(3) - b(2)) - (b(4) - b(3)) \end{aligned}$$

The process is additive if and only if these quantities are nonnegative. These conditions are monotonicity and convexity properties of the sequence of birth rates $b(i)$. A result for general neighborhoods can be found in Harris (1976), see (6.4) on page 184. The conclusions we would like the reader to draw from this computation are that (i) the additive processes are a small subset of the attractive processes but (ii) when we consider nonlinear contact processes with $|\mathcal{N}| = 4$ additive processes are a four dimensional subset of the four dimensional set of models.

Harris' convergence theorem for additive processes. Before getting started we need to introduce a technical condition. Let ξ_t^0 denote the process starting from a single particle at the origin. We say ξ_t is *irreducible* if for any x and $t > 0$ $P(x \in \xi_t) > 0$. Recall that in Section 2, we let ξ_t^1 denote the process starting from $\xi_0^1 = \mathbf{Z}^d$ and showed that for any attractive process $\xi_t^1 \Rightarrow \xi_\infty^1$, a translation invariant stationary distribution.

Theorem 3.3. Suppose ξ_t is an irreducible additive process with $\delta(\emptyset) > 0$. If ξ_0 is translation invariant and assigns 0 probability to the empty configuration then $\xi_t \Rightarrow \xi_\infty^1$ as $t \to \infty$.

Corollary. ξ_∞^1 is the only translation invariant stationary distribution that assigns 0 probability to the empty configuration.

Remarks. The condition $\delta(\emptyset) = 0$ eliminates the voter model for which the conclusion of Theorem 3.3 is always false. Our result is only for translation invariant initial distributions. With a lot more work one can prove a *complete convergence theorem*:

Theorem 3.4 Suppose ξ_t is an irreducible additive process with $\delta(\emptyset) > 0$. Then for any A,

$$\xi_t^A \Rightarrow P(\tau^A < \infty)\delta_\emptyset + P(\tau^A = \infty)\xi_\infty^1$$

where $\delta_\emptyset$ denotes the pointmass on the emptyset and we are using ξ^1_∞ to denote its distribution.

In words, if the process does not die out, then at large times it looks like the process starting from all 1's. This implies that all stationary distributions have the form $\theta\delta_\emptyset + (1-\theta)\xi^1_\infty$. For the contact process, this result is due to Bezuidenhout and Grimmett (1990). To prove this in the general case you will need to consult Bezuidenhout and Gray (1993).

PROOF OF THEOREM 3.3. To begin we note that the duality equation (3.3) implies

$$\begin{aligned} P(\xi^1_t \cap B \neq \emptyset) &= P(\hat{\xi}^{(B,t)}_t \cap \xi^1_0 \neq \emptyset) \\ &= P(\hat{\xi}^B_t \neq \emptyset) \to P(\hat{\tau}^B = \infty) \end{aligned}$$

as $t \to \infty$. As in the proof of Theorem 3.2, the argument in (2.8) shows that it is enough to prove $P(\xi_t \cap B \neq \emptyset) \to P(\hat{\tau}^B = \infty)$. Half of this is very easy. By duality and the fact that $\xi_0 \subset \mathbf{Z}^d$

$$P(\xi_t \cap B \neq \emptyset) = P(\xi_0 \cap \hat{\xi}^{(B,t)}_t \neq \emptyset) \leq P(\hat{\tau}^B > t)$$

so

$$\limsup_{t\to\infty} P(\xi_t \cap B \neq \emptyset) \leq P(\hat{\tau}^B = \infty)$$

To prove the other direction, we let t_0 be the constant in (2.1) and observe that (3.3) implies

$$P(\xi_{t+t_0} \cap B \neq \emptyset) = P(\xi_{t_0} \cap \hat{\xi}^{(B,t+t_0)}_t \neq \emptyset)$$

To get the right hand side to converge to $P(\hat{\tau}^B = \infty)$ we need to show that when $\hat{\xi}^{(B,t+t_0)}_t \neq \emptyset$ then it will intersect ξ_{t_0} with high probability. The first step in doing this is to show that when $\hat{\xi}^{(B,t+t_0)}_t \neq \emptyset$, it will contain a large number of points with high probability. To do this, let

$$\Lambda = \sum_A |A|(\lambda(A) + \delta(A))$$

be the rate at which an isolated particle gives birth to a new particle and let $\alpha = (1 - e^{-\delta(\emptyset)})e^{-\Lambda}$ be a lower bound on the probability that in one unit of time an isolated particle is killed and does not give birth. Now for any K

$$P(t < \hat{\tau}^B \leq t+1) \geq \alpha^K P(0 < |\hat{\xi}^{(B,t+t_0)}_t| \leq K)$$

To see this note that the events that each particle is killed by a δ are independent, and write the statement that no particle gives birth in terms of Poisson processes in the percolation substructure. Since $P(t < \hat{\tau}^B \leq t+1) \to 0$ as $t \to \infty$, and α^K is a positive constant, it follows that

$$P(0 < |\hat{\xi}^{(B,t+t_0)}_t| \leq K) \to 0 \tag{3.4}$$

To complete the proof now it suffices to show

(3.5) **Lemma.** If $\epsilon > 0$ then we can pick K large enough so that if $|A| \geq K$ then $P(\xi_{t_0} \cap A = \emptyset) \leq 3\epsilon$.

For then it follows that from (3.5) and (3.4) that

$$\liminf_{t\to\infty} P(\xi_{t_0} \cap \hat{\xi}_t^{(B,t+t_0)} \neq \emptyset) \geq (1-3\epsilon) \liminf_{t\to\infty} P(|\hat{\xi}_t^{(B,t+t_0)}| \geq K)$$
$$\geq (1-3\epsilon)P(\hat{\tau}^B > t)$$

Remark. For the conclusion in (3.5) it is important that we let the process run for a positive amount of time. The initial configuration ξ_0 that is $2\mathbf{Z}$ with probability 1/2 and $2\mathbf{Z}+1$ with probability 1/2 is translation invariant but $P(\xi_0 \cap \{2,4,\ldots,2K\}) = 1/2$ for all K.

PROOF OF (3.5): For this proof it is convenient to use the coordinate representation of the process, i.e., $\xi_t(x) = 1$ if x is occupied at time t and 0 otherwise. Let μ be the distribution of ξ_0 (i.e., the induced measure on $\{0,1\}^S$) and use P_ξ to denote the probability law for ξ_t when $\xi_0 = \xi$. Our assumption of irreducibility and attractiveness imply that $P_\xi(\xi_{t_0}(x) = 1) > 0$ unless $\xi \equiv 0$, an event that by assumption has probability 0, so

(3.6) For any $\epsilon > 0$ there is a $\rho < 1$ so that

$$\mu(\{\xi : P_\xi(\xi_{t_0}(x) = 0) > \rho\}) \leq \epsilon$$

Here we need translation invariance to conclude that the left hand side does not depend on x. The second ingredient is to note repeated use of Hölder's inequality gives

$$E(X_1 \cdots X_k) \leq (E|X_1^k|)^{1/k} \cdots (E|X_k^k|)^{1/k}$$

which in turn implies

(3.7) Let $X_1, \ldots, X_k$ be random variables so that $0 \leq X_i \leq 1$ and $P(X_i > \rho) \leq \epsilon$. Then

$$E(X_1 \cdots X_k) \leq \rho^k + \epsilon$$

Pick J so that $\rho_\epsilon^J \leq \epsilon$. Our proof of the next result explains why we chose the time t_0. The result is valid for any time t, see Holley (1972).

(3.8) Given $\epsilon > 0$ and J, we can pick L so that if $B \subset \mathbf{Z}^d$ with $|B| = J$ and $\|x-y\|_\infty > 2L$ whenever $x, y \in B$ with $x \neq y$ then

$$\left| E_\xi \left\{ \prod_{x\in B} (1 - \xi_{t_0}(x)) \right\} - \prod_{x\in B} \{E_\xi(1 - \xi_{t_0}(x))\} \right| \leq \epsilon$$

PROOF OF (3.8): First we compute the value of each $\xi_{t_0}(x)$ with $x \in B$ by using an independent copy of the percolation substructure $\mathcal{P}_x$. The second step is to combine all these independent substructures to make a new one $\mathcal{P}_{all}$ by taking $T_n^{y,A}$ and $U_n^{y,A}$ from $\mathcal{P}_x$ if and only if $y + A \subset D(x, L) = \{z : \|x - z\|_\infty \le L\}$ and then using another independent percolation substructure $\mathcal{P}^*$ to fill in the missing Poisson processes. Let R be the largest value of $\|x\|_\infty$ for a point in some set A with $\lambda(A)$ or $\delta(A) > 0$. R is the range of the interaction. If the cluster containing x in $\mathcal{P}_x$ defined in the proof of (2.1) lies inside $D(x, L - R)$ then it is identical with the cluster containing x in $\mathcal{P}_{all}$ and the values computed for ξ_{t_0} are the same. Since the states of x in the processes on $\mathcal{P}_x$ are independent, it follows from the proof of (2.1) that if L is large the random variables $1 - \xi_{t_0}(x)$ on $\mathcal{P}_{all}$ are equal with high probability to independent random variables and (3.8) follows. □

To complete the proof of (3.5) now, we observe that

(3.9) If $B \subset \mathbf{Z}^d$ with $|B| = J$ and if $\|x - y\|_\infty > 2L$ whenever $x, y \in B$ with $x \ne y$ then

$$P(\xi_{t_0}(x) = 0 \text{ for all } x \in B) = \int \mu(d\xi)\, E_\xi \prod_{x \in B} (1 - \xi_{t_0}(x))$$

$$\le \epsilon + \int \mu(d\xi) \prod_{x \in B} E_\xi(1 - \xi_{t_0}(x)) \le 2\epsilon + \rho_\epsilon^J \le 3\epsilon$$

by (3.8), (3.6), (3.7), and the choice of J. To get from the last result to the desired conclusion we let $K = (4L + 1)^d J$ and observe that if $|A| \ge K$ we can find a subset B with $|B| = J$ that sastisfies the hypotheses of (3.9). □

Example 3.4. Multitype contact processes, defined in Section 1, have state space $\{0, 1, \kappa - 1\}^S$ where 0 indicates a vacant site and $i > 0$ indicates a site occupied by one plant of type i, and have flip rates that are linear:

$$c_0(x, \xi) = \delta_{\xi(x)}$$
$$c_i(x, \xi) = \lambda_i n_i(x, \xi) \qquad \text{if } \xi(x) = 0$$

When $\lambda_i = \lambda$ and $\delta_i = \delta$, this process can be studied by using a duality that is a hybrid of the one for the contact process and for the voter model. The first step is to construct the process as we did the contact process. We introduce independent Poisson processes $\{U_n^x, n \ge 1\}$ with rate δ and $\{T_n^{x,y}, n \ge 1\}$ with rate λ for each $x, y \in \mathbf{Z}^d$ with $y - x \in \mathcal{N}$. As before, we write a δ at (x, U_n^x) to indicate that a death will occur if x is occupied by a particle of either type, and we draw an arrow from $(y, T_n^{x,y})$ to $(x, T_n^{x,y})$ to indicate that if x is vacant and y is occupied then there will be a birth from y to x.

If we define the dual process as in (3.2) then reasoning as before we see that x will be occupied at time t if and only if some site in $\hat{\xi}_t^{(\{x\},t)}$ is occupied in ξ_0. The dual for the mutlitype contact process is the set $\hat{\xi}_t^{(\{x\},t)}$ plus an ordering of that set with the interpretation that the type of x is that of the first occupied site in the ordering. For example in the realization drawn in Figure 3.2, the ordering is $1 > 2 > -2$

The first site in $\hat{\xi}_s^{(\{x\},t)}$ in this ordering is called the ***distinguished particle***. Results of Neuhauser (1992) show that the movements of the distinguished particle are enough like those of a random walk to conclude that in $d \leq 2$ the distinguished particles for the duals of two different sites will eventually be equal for large t. This is the key idea in proving Theorems 4C and 4D in Section 1. In the two type case, when $\delta_1 = \delta_2$ and $\lambda_1 < \lambda_2$ we can augment the construction above with Poisson processes of arrows that only allow the births of 2's and an easy argument gives Theorem 4A. However such an approach will never give us Conjecture 4B.

4. A Comparison Theorem

In this section we will introduce a comparison theorem that is very useful in proving the existence of nontrivial translation invariant stationary distributions. At this point we have to ask for the reader's patience: the result given in Theorem 4.3 is powerful but you will need to see a few applications to understand how it works.

Our general method for proving the existence of stationary distributions is to compare the process of interest with oriented percolation, so our first step is to introduce oriented percolation and state some of its basic properties, the proofs of which are hidden away in the appendix. Let

$$\mathcal{L}_0 = \{(x,n) \in \mathbf{Z}^2 : x+n \text{ is even}, n \geq 0\}$$

and make $\mathcal{L}_0$ into a graph by drawing oriented edges from (x,n) to $(x+1,n+1)$ and from (x,n) to $(x-1,n+1)$. Given random variables $\omega(x,n)$ that indicate whether the sites are open (1) or closed (0), we say that (y,n) can be reached from (x,m) and write $(x,m) \to (y,n)$ if there is a sequence of points $x = x_m, \ldots, x_n = y$ so that $|x_k - x_{k-1}| = 1$ for $m < k \leq n$ and $\omega(x_k,k) = 1$ for $m \leq k \leq n$. In the standard oriented percolation model the variables $\omega(x,n)$ are independent, but in almost all cases our comparisons will introduce dependencies between the $\omega(x,n)$, so we need a more general set-up. We say that the $\omega(x,n)$ are "*M dependent with density at least* $1-\gamma$" if whenever (x_i,n_i), $1 \leq i \leq I$ is a sequence with $\|(x_i,n_i)-(x_j,n_j)\|_\infty > M$ if $i \neq j$ then

$$P(\omega(x_i,n_i) = 0 \text{ for } 1 \leq i \leq I) \leq \gamma^I \tag{4.1}$$

Note: Classical M-dependence would require that the $\omega(x_i,n_i)$ considered above are independent. However the probability in (4.1) is the only one we need to control and hence the only thing we assume.

Given an initial condition $W_0 \subset 2\mathbf{Z} = \{x : (x,0) \in \mathcal{L}_0\}$, we can define a process by

$$W_n = \{y : (x,0) \to (y,n) \text{ for some } x \in W_0\}$$

In words, the sites W_n are those that are wet at time n. To keep the terminology straight, think of open sites as air spaces in a rock, and the sites in W_n as the ones that the fluid can reach (and hence wet) at level n. We use W_n^0 to denote the process that results when $W_0^0 = \{0\}$ and we let

$$\mathcal{C}_0 = \{(y,n) : (0,0) \to (y,n)\}$$

be the set of all points in space-time that can be reached by a path from $(0,0)$. (When $(0,0)$ is open $\mathcal{C}_0 = \cup_n(W_n^0 \times \{n\})$.) $\mathcal{C}_0$ is called the *cluster containing the origin*. Figure 4.1 shows a simulation of the independent oriented percolation process in which sites are open (indicated by black dots) with probability $p = 0.6$. Time goes up the page and lines connect the points of $\mathcal{C}_t$.

When the cluster containing the origin is infinite, i.e., $\{|\mathcal{C}_0| = \infty\}$ we say that *percolation occurs*. Our first result shows that if the density of open sites is high enough then percolation occurs. All that is important about the upper bound is that it is < 1 for small γ and converges to 0 as $\gamma \to 0$.

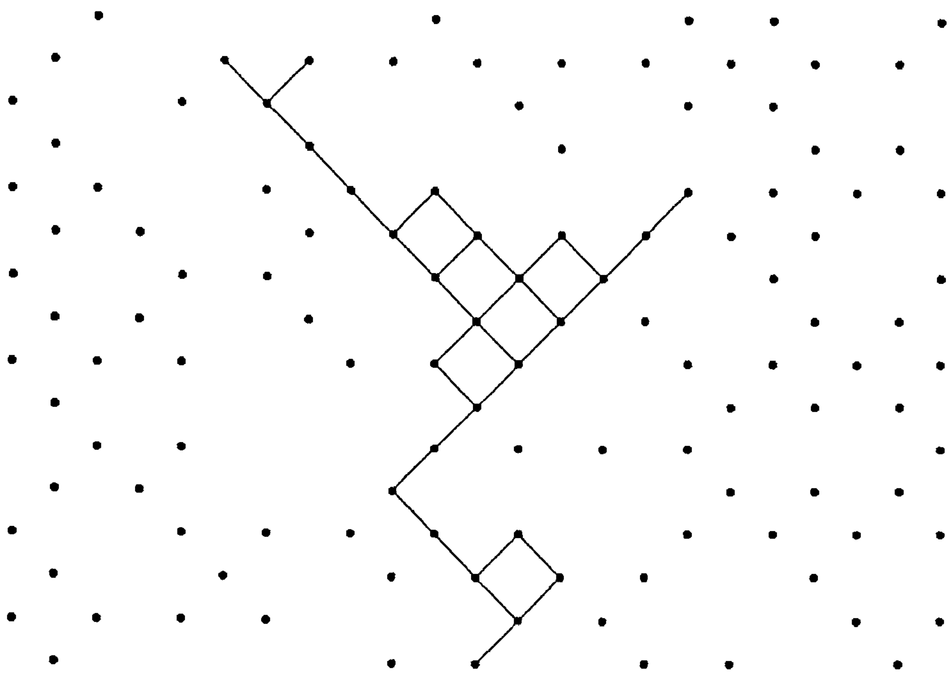


Figure 4.1

Theorem 4.1. If $\gamma \leq 6^{-4(2M+1)^2}$ then

$$P(|\mathcal{C}_0| < \infty) \leq 55\,\theta^{1/(2M+1)^2} \leq 1/20$$

In order to prove the existence of stationary distributions we need results about M dependent oriented percolation starting from the initial configuration W_0^p in which the events $\{x \in W_0^p\}$, $x \in 2\mathbf{Z}$ are independent and have probability p. We will sometimes call this a *Bernoulli random set with density* p. Taking $p = 1$ (i.e., all sites wet initially) corresponds to computing the upper invariant measure for oriented percolation, but for some of the proofs below we will need to allow $p < 1$. Note that the estimate on the lim inf is independent of p and is 1 minus the upper bound in Theorem 4.1.

Theorem 4.2. If $p > 0$ and $\gamma \leq 6^{-4(2M+1)^2}$ then

$$\liminf_{n\to\infty} P(0 \in W_{2n}^p) \geq 1 - 55\,\gamma^{1/(2M+1)^2} \geq 19/20$$

The last result shows that if the density of open sites in oriented percolation is sufficiently high and if we start with from a Bernoulli random set with density p then the probability 0 is wet at time t does not go to 0. This result will allow us to prove in a number of situations that if we start from a suitably chosen translation invariant initial distribution, then the density of sites of type i does not go to 0 and then using (2.7) or (2.11) that a nontrivial translation invariant stationary distribution exists. The missing link is provided by Theorem 4.3, which gives general conditions that guarantee a process dominates oriented percolation. This is the result we warned the reader about at the beginning of the section – it does not look pretty but it is very useful in a number of situations.

Comparison Assumptions. We suppose given the following ingredients: a translation invariant finite range process $\xi_t : \mathbf{Z}^d \to \{0, 1, \ldots \kappa - 1\}$ that is constructed from the graphical representation given in Section 2, an integer L, and a collection H of configurations determined by the values of ξ on $[-L, L]^d$ with the following property:

if $\xi \in H$ then there is an event G_ξ measurable with respect to the graphical representation in $[-k_0 L, k_0 L]^d \times [0, j_0 T]$ and with $P(G_\xi) \geq (1 - \gamma)$ so that if $\xi_0 = \xi$ then on G_ξ, ξ_T lies in $\sigma_{2Le_1} H$ and in $\sigma_{-2Le_1} H$.

Here $(\sigma_y \xi)(x) = \xi(x + y)$ denotes the translation (or shift) of ξ by y and $\sigma_y H = \{\sigma_y \xi : \xi \in H\}$. If we let $M = \max\{j_0, k_0\}$ then the space time regions

$$\mathcal{R}_{m,n} = (m2Le_1, nT) + \{[-k_0 L, k_0 L]^d \times [0, j_0 T]\}$$

that correspond to points $(m, n), (m', n') \in \mathcal{L}$ with $\|(m, n) - (m', n')\|_\infty > M$ are disjoint.

For a concrete instance of the comparison assumptions consider the applications we will make to the threshold contact process in Section 5 and to the basic contact process in Section 7. In both cases $\kappa = 2$, and H is the set of configurations with at least K 1's in $[-L, L]^d$, $k_0 = 4$, and $j_0 = 1$. In words, we show that if there is a "pile" of at least K particles in $[-L, L]^d$ then with high probability there will be piles of at least K particles in $-2Le_1 + [-L, L]^d$ and in $2Le_1 + [-L, L]^d$ at time T, and the event that guarantees this is measurable with respect to the graphical representation in $[-4L, 4L]^d \times [0, T]$. Figure 4.2 below gives a picture of the event.

Using words inspired by the contact process example, our comparison assumptions say that if we have a "pile of particles" in $I_m = m2Le_1 + [-L, L]^d$ at time nT (i.e., $\xi_{nT} \in \sigma_{m2Le_1} H$) then with high probability we will have piles of particles in I_{m-1} and in I_{m-1} at time $(n + 1)T$, and the event that guarantees this is measurable with respect to the graphical representation in $\mathcal{R}_{m,n}$. If we think of drawing arrows from (m, n) to $(m + 1, n + 1)$ and to $(m - 1, n + 1)$ whenever the good event in $\mathcal{R}_{m,n}$ occurs then the connection with oriented percolation should be clear.

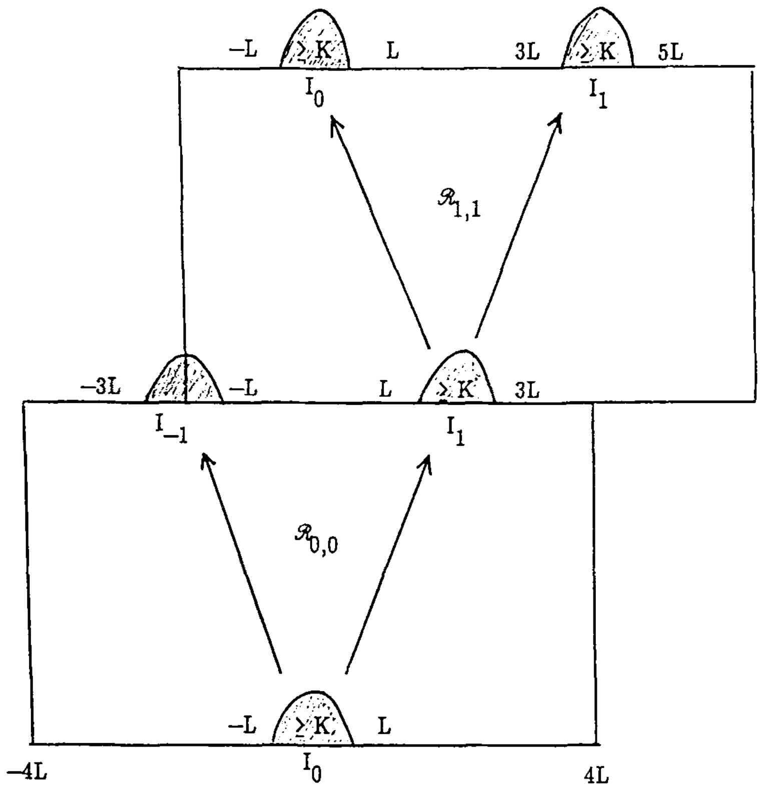


Figure 4.2

To formulate our theorem we let $X_n = \{m : (m,n) \in \mathcal{L}_0, \xi_{nT} \in \sigma_{m2Le_1} H\}$. Intuitively, $m \in X_n$ if there is a pile of particles in I_m at time nT

Theorem 4.3. If the comparison assumptions hold then we can define random variables $\omega(x,n)$ so that X_n dominates an M dependent oriented percolation process with initial configuration $W_0 = X_0$ and density at least $1-\gamma$, i.e., $X_n \supset W_n$ for all n.

Again the details are hidden away in the appendix so that they can be digested after the reader has seen that this is a useful result.

Our first indication that Theorems 4.1–4.3 are useful is a simple proof of a general result about the existence of stationary distributions, which contains as a special case a number of earlier results. To formulate our result we will consider a fixed set of increasing birth rates $c_1(x,\xi)$ and introduce death rates $c_0(x,\xi) \equiv \epsilon$. We say that the birth rates are *robust* if there is an $\epsilon_0 > 0$ so that there is a translation invariant stationary distribution with a positive density of 1's for $\epsilon < \epsilon_0$. Our next result gives a sufficient condition for robustness. It may look a little strange at first but it has been formulated to be easy to prove and to apply.

Theorem 4.4. Let $\bar{\xi}_t^{L,\rho}$ denote the process with no deaths, i.e., $\epsilon = 0$, starting from $\bar{\xi}_0^{L,\rho}(x) = 1$ for $x \in [-L,L]^d$, $= 0$ otherwise and modified so that no births are allowed outside $[-\rho L, \rho L]^d$. Suppose that we can pick $\rho \geq 3$ so that for any $\delta > 0$ we can pick L and $T < \infty$ so that

$$P(\bar{\xi}_T^{L,\rho}(x) = 1 \text{ for all } x \in [-3L, 3L]^d) \geq 1 - \delta$$

Then the birth rates are robust (and fertile).

Ignoring the undefined term in parentheses, this theorem says that if, in the absence of deaths, the birth mechanism can triple the size of a cube $[-L,L]^d$ with high probability, then there is a nontrivial translation invariant stationary distribution when the death rate $c_0(x,\xi) \equiv \epsilon$ is small. The requirement that this can be done when the model is "modified so that no births are allowed outside $[-\rho L, \rho L]^d$" is a technical condition that is usually satisfied with $\rho = 3$.

PROOF OF THEOREM 4.4: If we let $K = \rho$, $J = 1$ and $H = \{\xi : \xi(x) = 1 \text{ for all } x \in [-L,L]^d\}$. then the hypotheses of Theorem 4.4 are that the comparison assumptions hold for the system with $\epsilon = 0$. However, once L and T are fixed it follows that for $\epsilon \leq \epsilon_0$, the good event G_ξ for the one configuration in H has probability at least $1 - 2\delta$, since the probability a death occurs at some site in the space time box $[-3L, 3L]^d \times [0,T]$ is less than δ when ϵ_0 is sufficiently small.

To construct our stationary distribution, we consider the process ξ_t^1 starting from $\xi_0^1(x) = 1$ for all x. In this case $X_0 = 2\mathbf{Z}$ so using Theorems 4.3 and 4.2 with $p = 1$, it follows that if $\epsilon \leq \epsilon_0$ then

$$\liminf_{n \to \infty} P(\xi_{nT}^1(0) = 1) \geq 19/20$$

Using (2.7) now it follows that there is a nontrivial stationary distribution. □

We will now give three examples to shows that is easy to check the conditions of Theorem 4.4.

Corollary 4.5. If we fix $\lambda = 1$ in the contact process with neighborhood $\mathcal{N} = \{x : \|x\|_p \leq r\}$ where $r \geq 1$ then there is a nontrivial stationary distribution when the death rate $\delta < \delta_0$.

PROOF: Take $\rho = 3$ and $L = 1$. Since 1's can never flip to 0 it is easy to see that

$$\lim_{T\to\infty} P(\bar{\xi}_T^{L,\rho}(x) = 1 \text{ for all } x \in [-3L, 3L]^d) = 1$$

so the hypotheses of Theorem 4.4 are satisfied. □

Example 4.1. One Dimensional Counting Rules. Suppose $d = 1$, $\mathcal{N} = \{z : |z| \le k\}$, and let

$$n_1(x,\xi) = |\{z \in \mathcal{N} : \xi(x+z) = 1\}|$$

be the number of neighbors of x that are 1. We call a birth rate $c_1(x,\xi)$ a *counting rule* if it only depends on the number of 1's in the neighborhood, i.e., $c_1(x,\xi) = b(n_1(x,\xi))$ Clearly a counting rule birth rate is increasing if and only if $j \to b(j)$ is nondecreasing. Let $j_0 = \min\{j : b_j > 0\}$ and call j_0 the *order* of the birth rate. The next result is due to Mityugin.

Corollary 4.6. When $d = 1$ and $\mathcal{N} = \{j : |j| \le k\}$, increasing counting rule birth rates are robust if and only if their order $j_0 \le k$.

PROOF: If $j_0 > k$ then a string of at least $k+1$ consecutive 0's can never flip back to 1 even if all the other sites are 1. If $c_0(x,\xi) \equiv \epsilon > 0$ then such a string will eventually be created and grow to cover the whole line, so there cannot be a nontrivial stationary distribution.

If $j_0 \le k$, we take $\rho = 3$ and choose L so that $2L + 1 \ge k$. When $\epsilon = 0$ the 1's never flip back to 0. The 0 at $L+1$ has k neighbors that are 1 and hence flips to 1 at rate $b(k) \ge b(j_0) > 0$. Once the 0 at $L+1$ flips to 1, the 0 at $L+2$ will flip to 1 at rate $b(k)$, so

$$\lim_{T\to\infty} P(\bar{\xi}_T^{L,\rho}(x) = 1 \text{ for all } x \in [-3L, 3L]) = 1$$

and the hypotheses of Theorem 4.4 are satisfied. □

Things get more interesting in two dimensions.

Example 4.2. Two Dimensional Threshold Birth Rates. Suppose $d = 2$ and $\mathcal{N} = \{z : \|z\|_\infty = 1\}$, i.e., in addition to the four nearest neighbors we use the four diagonally adjacent points:

$$\mathcal{N} = \left\{ \begin{array}{ccc} (-1,1) & (0,1) & (1,1) \\ (-1,0) & & (1,0) \\ (-1,-1) & (0,-1) & (1,-1) \end{array} \right\}$$

This is sometimes called the *Moore neighborhood* in honor of one of the pioneers in the field of cellular automata. Let $n_1(x,\xi) = |\{j \in \mathcal{N} : \xi(x) = 1\}$ be the number of neighbors in state 1 and let

$$c_1(x,\xi) = \begin{cases} 1 & \text{if } n_1(x,\xi) \ge \theta \\ 0 & \text{if } n_1(x,\xi) < \theta \end{cases}$$

This is called a threshold θ since the birth rate 1 if there are at least θ 1's in the neighborhood then the birth rate is 1, and otherwise it is 0. From Theorem 4.4 we get easily that

Corollary 4.7. Two dimensional threshold birth rates for the Moore neighborhood in two dimension are robust if $\theta \leq 3$.

PROOF: Take $\rho = 3$, $L = 1$, and draw a picture.

4	3	2	3	4
3	1	1	1	3
2	1	1	1	2
3	1	1	1	3
4	3	2	3	4

We start with the 3×3 square of 1's occupied by 1's. If $\theta \leq 3$ then the four sites marked with 2's have birth rate 1 and will eventually become occupied. Once they do, the eight sites marked 3 have three occupied neighbors and will become occupied. Finally the four sites marked 4 will become occupied. At this point we have shown how the process can fill up $[-2,2]^2$. Repeating the argument, it is easy to see that

$$\lim_{T\to\infty} P(\bar{\xi}_T^{L,\rho}(x) = 1 \text{ for all } x \in [-3,3]^2) = 1$$

the hypothesis of Theorem 4.4 is satisfied and the result follows. □

In the last argument it was important that we used the Moore neighborhood, instead of the usual nearest neighbors $\{z : |z| = 1\}$. If we use the nearest neighbors then, no matter how big L, is if we start with $[-L,L]^2$ occupied nothing happens since any site outside $[-L,L]^2$ has at most one occupied neighbor.

0	0	0	0	0
0	0	x	0	0
1	1	1	1	1

Since births are impossible outside any rectangle containing the 1's in the initial configuration, it is clear that the threshold two birth rate for the nearest neighbors *dies out* whenever the death rate is $c_0(x,\xi) \equiv \epsilon > 0$. That is, if there are only finitely many 1's in ξ_0, then

$$P(\xi_t \not\equiv 0) \to 0 \quad \text{as } t \to \infty$$

Here $\xi_t \equiv 0$ is short for $\xi_t(x) = 0$ for all x. Note that the all 0's state is absorbing so $t \to P(\xi_t \not\equiv 0)$ is decreasing. The opposite of dies out is *survives*. That is, if L is large enough and we start with 1's on $[-L,L]^d$ then

$$\lim_{t\to\infty} P(\xi_t \not\equiv 0) > 0 \quad \text{as } t \to \infty$$

We say that a birth rate is *fertile* if it survives when $c_0(x, \xi) = \epsilon$ and $\epsilon < \epsilon_0$. As the parenthetical phrase in Theorem 4.4 indicates, our sufficient conditions for robustness are also sufficient for fertility.

Having two notions of what it means for birth rates to be large enough, fertility and robustness, it is natural ask what is the relationship between these two notions:

1. Results of Bezuidenhout and Gray imply that increasing birth rates that are fertile are also robust, but the two notions are not equivalent.

2. As we have shown the two dimensional threshold two system using the nearest neighbors is not fertile. However, Bramson and Gray (1991) have shown that it is robust. Intuitively the process cannot grow outside of a rectangle but it is good at filling in holes that develop so it can have a nontrivial stationary distribution when ϵ is small.

In the case of the Moore neighborhood in two dimensions, it is easy to see that the threshold 4 system is not fertile but techniques of Bramson and Gray (1991) can be used to show that it is robust. The threshold 5 system has finite configurations of 0's that cannot be filled in

$$\begin{array}{cccc} & 0 & 0 & \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ & 0 & 0 & \end{array}$$

so an easy argument shows that it is not robust. An interesting open problem is to look at the neighborhoods $\mathcal{N} = \{z : \|z\|_p \leq r\}$ (or even just take $p = \infty$) and find the largest thresholds for which the threshold θ birth rule on that neighborhood is robust (resp. fertile).

Further results. There are many other results proving the existence of phase transitions for processes with state space $\{0,1\}^S$. Gray and Griffeath (1982) proved a "stability theorem for attractive nearest neighbor spin systems on $\mathbf{Z}$" by the contour method, a result which was reproved by the methods of this section by Bramson and Durrett (1988). Gray (1987) proved results for the one dimensional majority vote model. Chen (1992) used ideas from bootstrap percolation to study a model with sexual reproduction. In general the numerical bounds on critical values from this method are terrible but Durrett (1992c) has shown that in some cases you can get good bounds.

Bramson and Neuhauser (1993) studied perturbations of one dimensional cellular automata. Their results are exciting because they apply to a number of examples that are not attractive. An important special case is that if one considers the addition mod 2 automaton:

$$\eta_{n+1}(x) = (\eta_n(x-1) + \eta_n(x+1)) \pmod 2$$

and adds spontaneous deaths at a small rate ϵ then there is a stationary distribution close to product measure with density 1/2. Figure 4.3 shows the cellular automaton starting from a single 1 at 0, which generates a discrete version of the Sierpinski gasket. Figure 4.4 shows what happens when we introduce spontaneous deaths at rate $\epsilon = 0.01$. Note that there are many more occupied sites in the model with extra deaths.

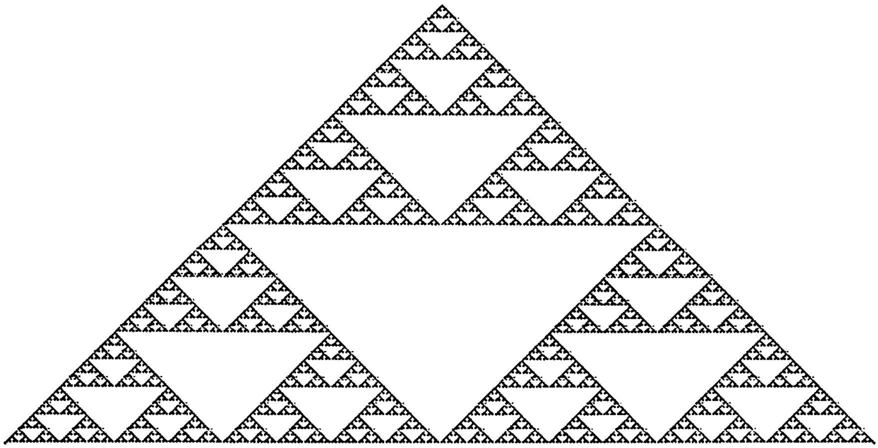


Figure 4.3. Pascal's triangle mod 2

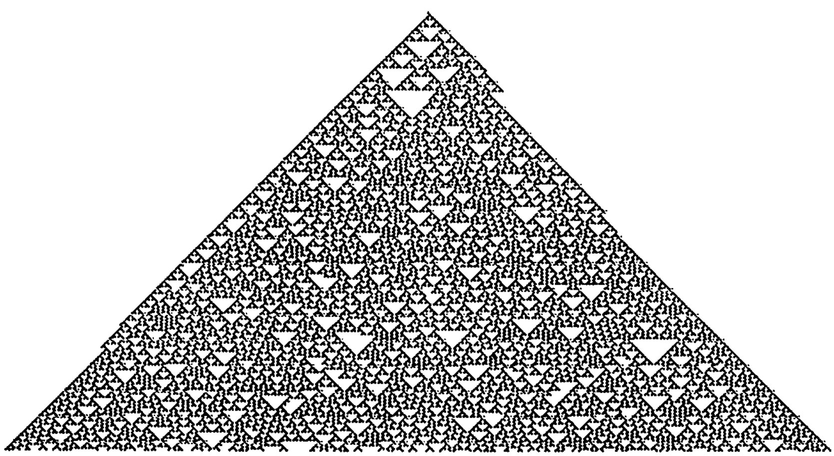


Figure 4.4. Plus spontaneous deaths with probability 0.01

5. Threshold Models

We begin by recalling a definition given in Section 1.

Example 5.1. The threshold voter model. The state space is $\{0,1\}^S$ and the flip rates are

$$c_i(x,\xi) = \begin{cases} 1 & \text{if } n_i(x,\xi) \geq \theta \\ 0 & \text{if } n_i(x,\xi) < \theta \end{cases}$$

Here, as usual, $n_i(x,\xi) = |\{y \in \mathcal{N} : \xi(x+y) = i\}|$ is the number of neighbors of type i and we assume $\mathcal{N} = \{y : \|y\|_p \leq r\}$ for some $1 \leq p \leq \infty$ and $r \geq 1$.

Our first goal is to show that the behavior of the threshold 1 voter model is much different from that of the basic voter model. We begin with one case in which the behavior is the same.

Theorem 5.1. Suppose $d = 1$ and $\mathcal{N} = \{-1,1\}$. Then the threshold 1 voter model clusters starting from any translation invariant initial state ξ_0. That is, for any $x \neq y$ we have $P(\xi_t(x) \neq \xi_t(y)) \to 0$.

PROOF: To motivate the proof, take a look at Figure 5.1 which shows a simulation of the system on $\{0,1,\ldots,719\}$ with periodic boundary conditions (i.e., 0 and 719 are neighbors). The initial configuration at the top of the page is product measure with density 1/2. As we go down the page from time 0 at the top to time 690 at the bottom, it should be clear that intervals of sites with the same opinion can be destroyed but cannot be created. Thus the number of intervals per unit distance will go to 0, i.e., the system clusters.

To turn the last paragraph into a proof, we define a process on $1/2 + \mathbf{Z}$ so that

$$\zeta_t(x) = |\xi_t(x-1/2) - \xi_t(x+1/2)|$$

In words, there is a 1 at x if and only if $\xi_t(x-1/2) \neq \xi_t(x+1/2)$. ζ is called the *boundary process of* ξ since the 1's mark the boundaries between clusters of the voters with the same opinion. To see how ζ evolves consider the following picture

ξ	1		1		0		1		1		0		0
ζ		0		1		1		0		1		0	
x	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6	6.5	7

Isolated 1's in ζ like the one at 5.5 perform random walks: the 1 at 5 flips to 0 at rate one and when this occurs the boundary jumps from 5.5 to 4.5; similarly, the 0 at 6 flips to 1 at rate one and when this occurs the boundary jumps from 5.5 to 6.5. When a 1 is adjacent to another 1 (like those at 2.5 and 3.5) they annihilate at rate 1, since when the 0 at 3 flips to 1 the two boundaries disappear.

Let $u(t) = P(\zeta_t(x) = 1)$, which is independent of x since we have supposed that ξ_0 is translation invariant. Since 1's can be destroyed in ζ but cannot be created, it should not be surprising that $u(t) \to 0$ as $t \to \infty$. To prove this, we note that

$$\text{(5.1)} \qquad \frac{du}{dt} = -P(\zeta_t(x) = 1, \zeta_t(x-1) = 1) - P(\zeta_t(x) = 1, \zeta_t(x+1) = 1)$$

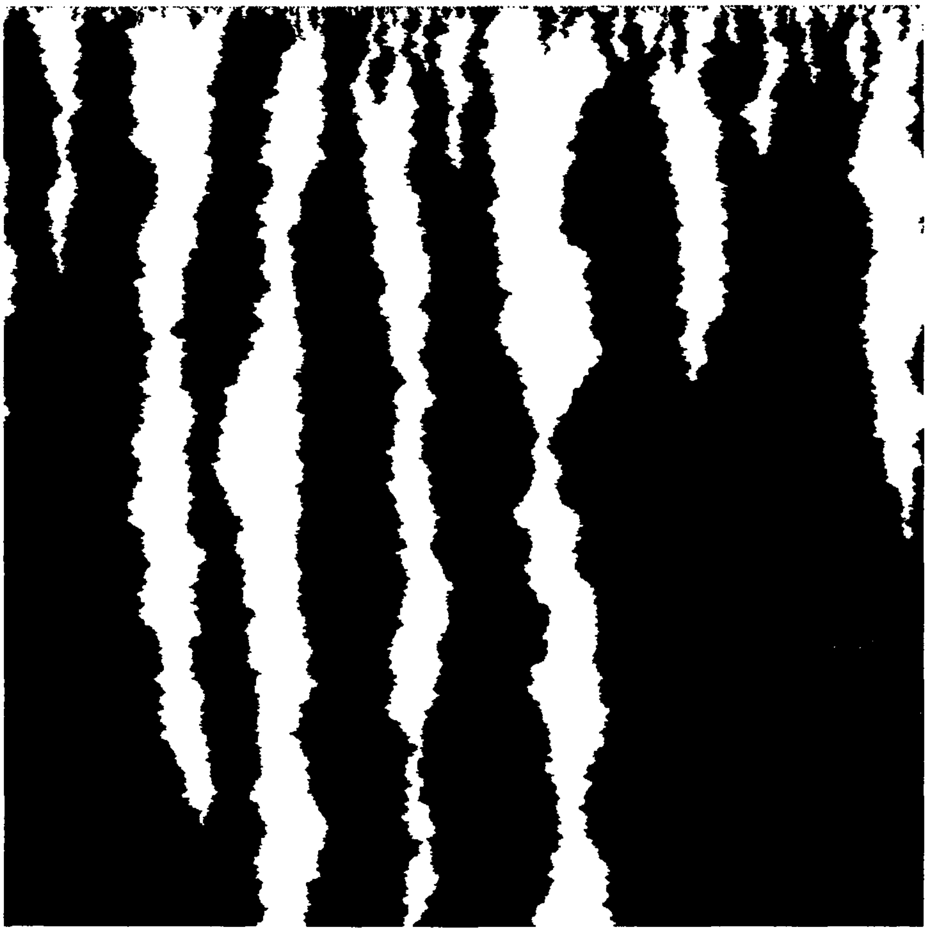


Figure 5.1. One dimensional nearest neighbor threshold voter model.

This can be proved by using $\frac{d}{dt}T_t f = T_t L f$ or more intuitively by noting that the right hand side gives the two ways that a 1 at x can be destroyed. The terms that involve a 1 moving to x or moving away from x cancel.

Translation invariance implies that the right hand side of (5.1) is

$$-2P(\zeta_t(0.5) = 1, \zeta_t(-0.5) = 1) \equiv -v(t)$$

The first step in proving $u(t) \to 0$ is to show that if $t \geq 1$

$$v(t) \geq g(u(t-1)) \text{ where } g(x) > 0 \text{ when } x > 0 \tag{5.2}$$

To do this we note that if $u(s) \geq 1/L$ where L is an integer then

$$P(\zeta_s \text{ has at least two 1's in } (-L, L]) \geq \frac{1}{2L-1} \tag{5.3}$$

for otherwise we get a contradiction

$$2 \leq 2Lu(s) = E \sum_{x \in (-L,L]} \zeta_s(x) < 1 \cdot \frac{2L-2}{2L-1} + 2L \cdot \frac{1}{2L-1} = \frac{4L-2}{2L-1} = 2$$

Now if we have an initial configuration in which there are at least two 1's in $(-L, L]$ there is a probability $\geq \epsilon_L > 0$ that no particles will enter $(-L, L]$ before time 1, the two particles closest to 0 will move to 0.5 and -0.5, and none of the other particles in $(-L, L]$ will move. Combining this observation with (5.3) proves (5.2). To complete the proof of Theorem 5.1 now, we observe that $u(t)$ is decreasing so $u(t) \to u(\infty) \geq 0$ as $\to \infty$. If $u(\infty) > 0$ then for all t we have

$$\frac{du}{dt} = -v(t) \leq -g(u(\infty)) < 0$$

so integrating we find $u(t) \to -\infty$ a contradiction. □

Remark. The argument above applies to any one dimensional nearest neighbor system in which $c_i(x, \xi) = f(n_i(x, \xi))$ with $f(0) = 0$, the so-called *nonlinear voter models*. In the case of the basic voter model, i.e., $f(2) = 2f(1)$ the boundary process is an *annihilating random walk*. That is, particles perform independent random walks until they hit at which time the two particles annihilate. Theorem 3.1 shows that for the basic voter model clustering occurs for any initial configuration. Theorem 4 in Cox and Durrett shows that for the threshold voter model clustering occurs for any initial configuration. We

Conjecture 5.1. In any one dimensional nearest neighbor nonlinear voter model clustering occurs for any initial configuration.

Our next goal is to show that coexistence is possible in the threshold 1 voter model even in one dimension. To do this we will use some ideas from Liggett (1993) to compare with

Example 5.2. The threshold contact process. The state space is $\{0,1\}^S$ and the flip rates are

$$c_1(x,\xi) = \begin{cases} \lambda & \text{if } n_1(x,\xi) \ge \theta \\ 0 & \text{if } n_1(x,\xi) < \theta \end{cases}$$
$$c_0(x,\xi) = 1$$

Here $c_1(x,\xi)$ is the same as in the threshold voter model but we have set $c_0(x,\xi) \equiv 1$.

(5.4) **Lemma.** If the threshold θ contact process with $\lambda = 1$ has a nontrivial stationary distribution then so does the threshold θ voter model.

PROOF: To construct the stationary distribution we will start the threshold voter model ξ from $\nu_{1/2}$, product measure with density 1/2, and compare with the threshold contact process ζ to show that clustering does not occur.

The first step in doing this is to show that the upper invariant measure π for the threshold voter model with $\lambda = 1$ is stochastically smaller than $\nu_{1/2}$ To do this we compare the threshold contact process ζ with the "independent flips process" η_t in which $c_i(x,\eta) \equiv 1$, i.e., each site flips at rate 1 independently of the others. Since sites in η flip to 1 at rate one independent of what is around them, if we start ζ and η with $\zeta_0 = \eta_0$ having distribution π and construct the two processes using the recipe in Section 2 then $\zeta_t(x) \le \eta_t(x)$ for all t and x. This is true since 1's flip to 0 at rate 1 in both processes while 0's flip to 1 at rate 1 always in η, but at rate 1 in ζ only if there are enough 1 neighbors. On the graphical representation then we find that each flip preserves the inequality and the result can be proved like (2.5).

Now since the sites in η flip independently it is easy to see that as $t \to \infty$ η_t converges to $\nu_{1/2}$. The inequality $\zeta_t(x) \le \eta_t(x)$ and the fact that ζ_t always has distribution π imply that π is stochastically smaller than $\nu_{1/2}$. To prove this we observe that if f is increasing and depends on only finitely many coordinates then $Ef(\zeta_t) \le Ef(\eta_t)$ and since any such f is bounded and continuous letting $t \to \infty$ gives

$$\int f(\xi) d\pi \le \int f(\xi) d\pi$$

checking the definition we gave in (2.11).

Now the result of Holley in the remark (2.12) implies that we can define ξ_0 with distribution $\nu_{1/2}$ and ζ_0 with distribution π, so that that $\xi_0(x) \ge \zeta_0(x)$ for all x. Since sites in ζ flip to 0 at rate one, while those in ξ only flip to 0 at rate one when there are enough 0 neighbors, and the rates of flipping to 1 are the same, if we construct the two processes using the recipe in Section 2 then $\xi_t(x) \ge \zeta_t(x)$ for all x and t. To construct a stationary distribution for ξ, let μ_t be the distribution of ξ_t, form the Cesaro average

$$\bar{\mu}_T = \frac{1}{T} \int_0^T \mu_t dt$$

and let $\bar{\mu}_\infty$ be the limit of a weakly convergent subsequence. It follows from (2.13) that $\bar{\mu}_\infty$ is a stationary distribution. To see that it concentrates on configurations with infinitely many 1's we note that the inequality $\xi_t(x) \geq \zeta_t(x)$ implies that $\bar{\mu}_\infty$ is larger than the upper invariant measure π, which is spatially ergodic by (2.15) and hence concentrates on configurations with infinitely many 1's. To see that $\bar{\mu}_\infty$ concentrates on configurations with infinitely many 0's, note that the initial distribution $\nu_{1/2}$ and the threshold voter model are symmetric under the interchange of 0's and 1's, so the limit measure $\bar{\mu}_\infty$ is as well. □

Liggett (1993) has shown

Theorem 5.2. When $d = 1$ and $\mathcal{N} = \{-2, -1, 1, 2\}$ or $d = 2$ and $\mathcal{N} = \{y : \|y\|_1 = 1\}$ the threshold 1 contact process with $\lambda = 1$ has a nontrivial stationary distribution.

Since enlarging the neighborhood $\mathcal{N}$ makes it easier for the threshold 1 contact process to have a nontrivial stationary distribution, it follows from (5.4) and Theorem 5.2 that

Theorem 5.3. Suppose $\mathcal{N} = \{z : \|z\|_p \leq r\}$ with $1 \leq p \leq \infty$ and $r \geq 1$. With the exception of the one dimensional nearest neighbor case, the threshold one voter model always has a nontrivial stationary distribution.

By another comparison argument Liggett shows that to prove Theorem 5.3 it is enough to consider the case $d = 1$ and $\mathcal{N} = \{-2, -1, 1, 2\}$ – map $\mathbf{Z}^2$ to $\mathbf{Z}$ by $(x, y) \to x + 2y$ and notice that the image of the two dimensional threshold contact process dominates the one dimensional one. A simulation of the case $d = 1$ and $\mathcal{N} = \{-2, -1, 1, 2\}$ given in Figure 5.2, which parallels the one for the nearest neighbor case in Figure 5.1, makes it clear that Theorem 5.3 is true. However, the proof of Theorem 5.2 (which implies 5.3) requires a tricky generalization of the result Holley and Liggett (1978) that the one dimensional nearest neighbor contact process has $\lambda_c \leq 2$. Therefore we content ourselves to prove less (and more).

Theorem 5.4. Suppose $\mathcal{N} = \{y : \|y\|_p \leq r\}$ with $r \geq 1$. For any threshold θ if $r \geq r_0(d, \theta)$ then there is a nontrivial stationary distribution for threshold θ contact process with $\lambda = 1$ and hence also for the threshold θ voter model.

PROOF: We will use the comparison theorem from Section 4. To do this, it is convenient to suppose that ξ has been constructed from a percolation substructure with rate 1 Poisson processes $\{T_n^x, n \geq 1\}$ at which times we draw arrows from $y + x$ to x for all $y \in \mathcal{N}$, and rate 1 Poisson processes $\{U_n^x, n \geq 1\}$ at which times we write a δ at x.

Exercise. This shows that the threshold contact process can be constructed from a percolation substructure defined in Section 3. What is the dual process?

Suppose $r = (2d + 2)L$. To check the comparison assumptions, let H be the configurations that have at least θ 1's in $[-L, L]^d$. Let $\gamma > 0$. If T is small enough then the probability that $U_1^x > T$ for all of our θ 1's, is $e^{-\theta T} > 1 - \gamma/5$. Now since $r = (2d + 2)L$, the neighborhood of each site in $I_1 = [L, 3L] \times [-L, L]^{d-1}$ contains all the sites in $[-L, L]^d$

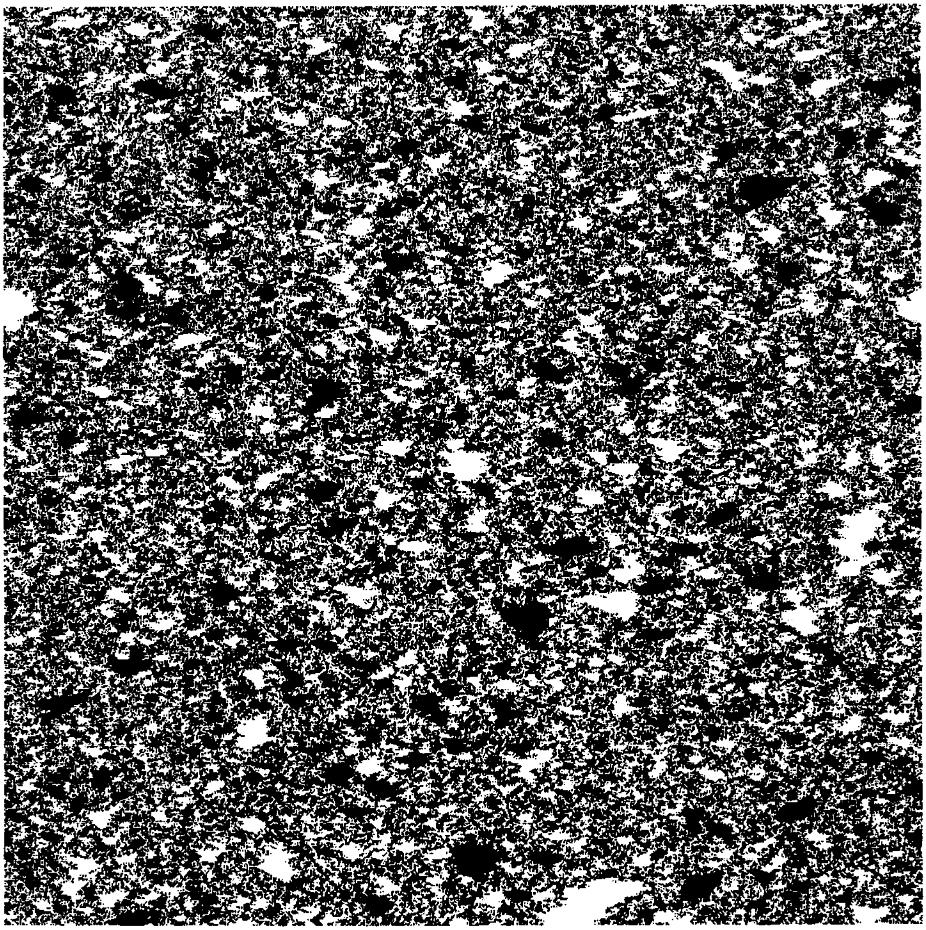


Figure 5.2. One dimensional threshold voter model, range two.

(distances are largest for the L^1 norm and for the points $(3L, L, \dots, L)$ and $(-L, L, \dots, L)$). Now as long as there are at least θ 1's in $[-L, L]^d$, each site in $[L, 3L]$ will flip to 1 at rate 1. If r and hence L is sufficiently large then with probability at least $1-\gamma/5$ at least θ sites will flip to 1 by time T. A similar remark applies to the sites in $I_{-1} = [-3L, -L] \times [-L, L]^{d-1}$, and our first estimate implies that in each case the probability one of our θ 1's flips back to 0 by time T is $\le \gamma/5$.

The results in the last paragraph show that if we start with θ 1's in $I_0 = [-L, L]^d$ then with probability at least $1-\gamma$ there will be at least θ 1's in I_1 and in I_{-1} at time T. Our good event is measurable with respect to the graphical representation in $[-3L, 3L]^d$ so we have checked the comparison assumptions of Section 4 with $k_0 = 3$ and $j_0 = 1$. If we start the threshold contact process with all sites occupied then Theorem 4.3 implies our process dominates an oriented percolation starting with all sites wet, so Theorem 4.2 shows

$$\liminf_{n\to\infty} P(0 \in X_{2n}) \ge 19/20$$

Now $0 \in X_{2n}$ means that there are at least θ 1's in $[-L, L]^d$ at time $2nT$ and ξ_{2nT} is translation invariant so it follows that

$$\begin{aligned}\liminf_{n\to\infty} P(\xi_{2nT}(0) = 1) &= \liminf_{n\to\infty} \frac{1}{(2L+1)^d} \sum_{x\in[-L,L]^d} P(\xi_{2nT}(x) = 1) \\ &\ge \frac{1}{(2L+1)^d} \cdot \theta \cdot \frac{19}{20} > 0\end{aligned}$$

To pass from this result to the whole sequence we notice that since a 1 survives for t units of time with probability e^{-t}, $P(\xi_{2nT+t}(0) = 1) \ge e^{-t} P(\xi_{2nT}(0) = 1)$. Combined with the last result this implies

$$\liminf_{T\to\infty} \frac{1}{T} \int_0^T P(\xi_s(0) = 1)\, ds > 0$$

and it follows from (2.13) that there is a nontrivial stationary distribution. □

The last result shows that if the threshold is small compared to the number of neighbors then *coexistence* occurs in the threshold voter model, i.e. there is a stationary distribution that concentrates on configurations with infinitely 1's and infinitely many 0's. Our next result due to Durrett and Steif (1993) shows that if the threshold is too large the system *fixates*, i.e., with probability one each site changes its state only finitely many times.

Theorem 5.5. Suppose $\mathcal{N} = \{y : \|y\|_p \le r\}$. If $\theta > (|\mathcal{N}| - 1)/2$ then the system fixates.

The borderline case in this result, $\theta = (|\mathcal{N}|+1)/2$ ($|\mathcal{N}|$ is always odd), is called *the majority vote process*, since you change your mind if you are in the minority in your neighborhood.

PROOF: Our proof is based on an idea of Grannan and Swindle. Let $\delta_{x,y}(t)$ be 1 if $\xi_t(x) \ne \xi_t(y)$, 0 otherwise, and define the *energy* at time t to be

$$\mathcal{E}_t = \sum_{x,y:y-x\in\mathcal{N}} e^{-\epsilon\|x+y\|_2} \delta_{x,y}(t)$$

where $\epsilon > 0$ is to be chosen later. Since $0 \le \mathcal{E}_0 < \infty$, we can prove Theorem 5.5 by showing

(5.5) If $\theta > (|\mathcal{N}| - 1)/2$ and ϵ is small then a flip at x decreases the energy by at least $\gamma(x) > 0$.

To prove (5.5) we note that if $\alpha = |\{y \in x+\mathcal{N} : \xi_t(y) \ne \xi_t(x)\}|$ and $N = \sup\{\|x\|_2 : x \in \mathcal{N}\}$ then the drop in energy due to a flip at x is at least

$$e^{-2\epsilon\|x\|_2}\left[e^{-\epsilon N}\alpha - e^{\epsilon N}(|\mathcal{N}| - 1 - \alpha)\right] \tag{5.6}$$

since (i) the site x now agrees with the α sites it used to disagree with and now disagrees with the other $|\mathcal{N}| - 1 - \alpha$ neighbors and (ii) even in the worst case all the points in $\{y \in x + \mathcal{N} : \xi_t(y) \ne \xi_t(x)\}$ have $\|x + y\|_2 \le 2\|x\|_2 + N$ and the other points $y \in x + \mathcal{N}$ have $\|x + y\|_2 \ge 2\|x\|_2 - N$. In order for a flip to occur we must have $\alpha \ge \theta > (|\mathcal{N}| - 1)/2$ and hence $|\mathcal{N}| - 1 - \alpha < \alpha$. Since the last two number are integers smaller than $|\mathcal{N}|$, (5.5) follows from (5.6). □

Refinements of Theorem 5.4. Before we stated Theorem 5.5, we said "if the threshold is small compared to the number of neighbors" then the threshold contact process with $\lambda = 1$ has a nontrivial stationary distribution (and hence there is coexistence in the threshold voter model). What we would like to concentrate on now is:

How large can θ be when the range is r?

The comparison theorem involves obnoxiously small constants (when $M = 1$ Theorems 4.1 and 4.2 require $\gamma \le 6^{-100}$). So we cannot hope to get a nontrivial result for $r = 10$, or even $r = 10,000$, but it is not unreasonable to look at how θ behaves asymptotically with r. The results were are about to give foreshadow the developments in the next section, but are not needed for them, or for any subsequent section, and can be skipped without loss.

Here and until the end of the section we suppose $\mathcal{N} = \{z : \|z\|_p \le r\}$, let $N = |\mathcal{N}|$, and we investigate what happens for fixed p as $r \to \infty$ First let's see what we get when we follow the proof of Theorem 5.4.

(5.7) There is a $c_p > 0$ so that if $\theta \le c_p\sqrt{N}$ and if r (and hence N) is large then the threshold θ contact process with $\lambda = 1$ has a nontrivial translation invariant stationary distribution.

Proof: Taking $T = \gamma/5\theta$ gives $e^{-\theta T} = e^{-\gamma/5} \ge 1 - \gamma/5$. Having fixed the time, the number of sites in $[L, 3L] \times [-L, L]^{d-1}$ that flip to 1 by time T has a binomial distribution with parameters $n = (2L + 1)^d$ and $p = 1 - e^{-T} \ge \gamma/6\theta$ when θ is large. If we let Z be the number of sites in $[L, 3L] \times [-L, L]^{d-1}$ that flip to 1 by time T then Z has mean $\ge (2L+1)^d\gamma/6\theta$ and variance $\le (2L+1)^d\gamma/6\theta$ so if we set $(2L+1)^d\gamma/6\theta = 2\theta$ (sticklers for details should take the smallest integer L so that $\ge$ holds) Chebyshev's inequality implies that

$$P(Z \le \theta) \le \frac{(2L + 1)^d\gamma/6\theta}{\theta^2} \le \frac{2}{\theta} \to 0$$

as $\theta \to \infty$. Now $\theta^2 = (2L+1)^d\gamma/12 \geq c_p N$ since $r = (2d+2)L$ and the result follows. □

By choosing a more intelligent block event we can get

(5.8) There is a $c_p > 0$ so that if $\theta \leq c_p N$ and if r (and hence N) is large then the threshold θ contact process with $\lambda = 1$ has a nontrivial translation invariant stationary distribution.

PROOF: Let $\theta = (2L+1)^d/5$ and let H be the configurations that have at least $(2L+1)^d/4$ 1's in $[-L, L]^d$. If we pick $r = (2d+2)L$ then $\theta \geq c_p N$ for all r and as long as there are at least θ 1's in $[-L, L]^d$ the number of 1's in $[-L, L]^d$ (or in $[L, 3L] \times [-L, L]^{d-1}$), behaves like a Markov chain that jumps $k \to k+1$ at rate $(2L+1)^d - k$ and $k \to k-1$ at rate k. Now when $k \leq (2L+1)^d/3$ this chain jumps at rate $(2L+1)^d$ moving up with probability at least 2/3 and down with probability at most 1/3. A comparison with asymmetric simple random walk shows

(i) with high probability it will take a long time (i.e., at least $e^{c(2L+1)^d}$ for some $c > 0$) for the total number of 1's in $[-L, L]^d$ to go below θ

(ii) we can pick a large time T (that is independent of L) so that if L is large then with high probability the number of 1's in $[L, 3L] \times [-L, L]^{d-1}$ and in $[-3L, -L] \times [-L, L]^{d-1}$ at time T will be at least $(2L+1)^d/4$

We leave it to the reader to fill in the missing details since we know how to prove a sharp result:

(5.9) Let $c < 1/4$. If $\theta \leq cN$ and if r (and hence N) is large then the threshold θ contact process with $\lambda = 1$ has a nontrivial translation invariant stationary distribution.
Let $c > 1/4$. If $\theta \geq cN$ and if r (and hence N) is large then the threshold θ contact process with $\lambda = 1$ has only the trivial stationary distribution.

The proof of the first conclusion is closely related to that of Theorem 6.1. For details and the proof of the converse see Durrett (1992).

6. Cyclic Models

As already suggested by our remarks on refinements in the last section, we can considerably close the gap between Theorems 5.4 and 5.5 if we look at systems with large range. The proof of our main result, Theorem 6.1, is no harder for a class of models that includes a multicolor version of the threshold voter model, so we formulate the result in that generality.

Example 6.1. Cyclic Color Model. The states of each site are $\{0, 1, \ldots, \kappa - 1\}$ and the flip rates are

$$c_i(x,\xi) = \begin{cases} 1 & \text{if } \xi(x) = i-1 \text{ and } n_i(x,\xi) \geq \theta \\ 0 & \text{otherwise} \end{cases}$$

Here and throughout this section, arithemtic is done modulo κ so $0 - 1 = \kappa - 1$. When $\kappa = 2$ the last definition reduces to the threshold voter model. The dynamics here were invented by David Griffeath as a generalization of the voter model. The cyclic color model is closely related to the hypercycle of evolutionary biology. See Eigen and Schuster (1979) and Boerlijst and Hogeweg (1991).

Our main result also applies to two other examples

Example 6.2. Greenberg Hastings Model. The states of each site are $\{0, 1, \ldots, \kappa - 1\}$ and the flip rates are

$$\begin{aligned} c_1(x,\xi) &= 1 \qquad && \text{if } \xi(x) = 0 \text{ and } n_1(x,\xi) \geq \theta \\ c_i(x,\xi) &= 1 && \text{if } \xi(x) = i-1 \end{aligned}$$

In words, we need an above threshold number of 1's to make the transition from $0 \to 1$ but then the rest of the transitions happen at rate 1. When $\kappa = 2$ this reduces to the threshold contact process with $\lambda = 1$.

Example 6.3. Host Parasitoid Interactions. Insect parasitoids lay their eggs on or in the bodies of other arthropods, and the parasitoid larvae kill their host as they feed on it. Hassell, Comins, and May (1991) introduced a cellular automaton model for this system. The corresponding particle system model has nine states $\{0, 1, \ldots 8\}$ and makes transitions as follows:

$$\begin{aligned} c_1(x,\xi) &= 1 \qquad && \text{if } \xi(x) = 0 \text{ and } n_1(x,\xi) \geq \theta \\ c_4(x,\xi) &= 1 && \text{if } \xi(x) = 3 \text{ and } n_5(x,\xi) \geq \theta \\ c_i(x,\xi) &= 1 && \text{if } i \neq 1, 4 \text{ and } \xi(x) = i-1 \end{aligned}$$

As they explain on page 256, the first transition corresponds to colonization of empty sites (state 0) by the host, the second to a mature parasitoid (state 5) colonizing a mature host (state 3), and the others to the aging and/or death of host and parasitoid.

To indicate what common features of the last three models are needed to apply Theorem 6.1, we say that ξ is a *cyclic model* if the states of each site are $\{0, 1, \ldots, \kappa - 1\}$ and makes transitions as follows:

$$c_i(x,\xi) = 1 \qquad \text{if } \xi(x) = i \text{ and } n_{g(i)}(x,\xi) \geq \theta_i$$

Here $g(i) \in \{0, 1, \ldots, \kappa - 1\}$ and we set $\theta_i = 0$ if the transition happens at rate 1 independent of the states of the neighbors. Let $\theta = \max_i \theta_i$.

Theorem 6.1. Let $\epsilon > 0$ and suppose $\theta \le (1 - \epsilon)|\mathcal{N}|/2\kappa$. If $r \ge R_\epsilon$ then there is a stationary distribution close to the uniform product measure.

Recall that we suppose $\mathcal{N} = \{y : \|y\|_p \le r\}$ and that the uniform product measure is the one in which the coordinates are independent and have $P(\xi(x) = i) = 1/\kappa$. When $\kappa = 2$ this says that for thresholds $a|\mathcal{N}|$ with $a < 1/4$ there is coexistence for large r. (This result was stated in (5.9).) In contrast Theorem 5.4 says that when $a \ge 1/2$ the system fixates for any r. We

Conjecture 6.1. When $\theta = a|\mathcal{N}|$ in the threshold voter model and $1/4 < a < 1/2$, clustering occurs for large r.

We will explain our reasons after we give the proof. Theorem 6.1 concentrates on the behavior for large range. For results about the one dimensional cyclic color model, see Bramson and Griffeath (1987) (1989), or for a treatment of the corresponding cellular automaton, see Fisch (1990a), (1990b), (1991).

Proof in $d = 1$: Let $a = \theta/|\mathcal{N}|$. By assumption $a \le (1 - \epsilon)/2\kappa$. Pick $\beta \in (0, \epsilon/4]$ so that $B = 1/\beta$ is an integer, pick $\rho < \sigma < 1/\kappa$ so that $(1 - 2\beta)\rho \ge (1 - \epsilon)/\kappa$, then pick r large enough so that

$$(1 - \beta)\rho \cdot \frac{2r}{2r + 1} \ge (1 - \epsilon)/\kappa$$

Let $K = \beta r$ and note that $BK = r$. For each $m \in \mathbf{Z}$, we call $[mK, (m + 1)K)$ a ***house***. We say that a house is *good* at time 0 if it contains at least σK sites in each of the states $0, 1, \ldots, \kappa - 1$. We say that the interval $[-r, r)$ is *good* at time 0 if all the houses it contains are good. This will be our event H when we apply Theorem 4.3.

We have chosen our constants so that as long as each house in $[-r, r)$ is *reasonable* i.e., contains at least ρK sites of each color, each site in $[-r - K, r + K)$ will see at least θ sites of each color. To check this, note that the worst case occurs when $x \in [r, r + K)$, but even in this case all the sites in $[K, r)$ are in its neighborhood and if all of the houses in $[K, r)$ are reasonable then the number of sites of a given color in x's neighborhood will be at least

$$\rho(r - K) = \rho r(1 - \beta) = \frac{\rho(1 - \beta)}{2} \cdot \frac{2r}{2r + 1} \cdot 2r + 1$$
$$\ge \frac{(1 - \epsilon)}{2\kappa} \cdot (2r + 1) \ge \theta$$

So as long as each house in $[-r, r)$ stays reasonable, the sites in $[-r - K, r + K)$ flip from i to $i + 1$ at rate 1 (here $(\kappa - 1) + 1 = 0$) and hence behave like independent Markov chains. These "single site" Markov chains are irreducible on $\{0, 1, \ldots, \kappa - 1\}$ and hence converge to the equilibirum distribution, which assigns probability $1/\kappa$ to each state. Let $p_t(i, j)$ be the transition probability of the single site Markov chain, let $\sigma' \in (\sigma, 1/\kappa)$ and pick S so that $p_S(0,i) \geq \sigma'$ for all i. Let $T = 2BS$. By using a simple large deviations result (see (6.2) below) it is easy to show that with high probability

(a) All the houses in $[-r, r)$ stay reasonable until time T.

(b) The houses $[r + (j-1)K, r + jK)$ and $[-r - jK, -r - (j-1)K)$ will be good at time jS and stay reasonable to time T.

(c) All the houses in $[r, 3r)$ and $[-3r, -r)$ will be good at time T.

Figure 6.1 gives a picture of this expansion. The gray shaded area gives the space time region occupied by reasonable houses.

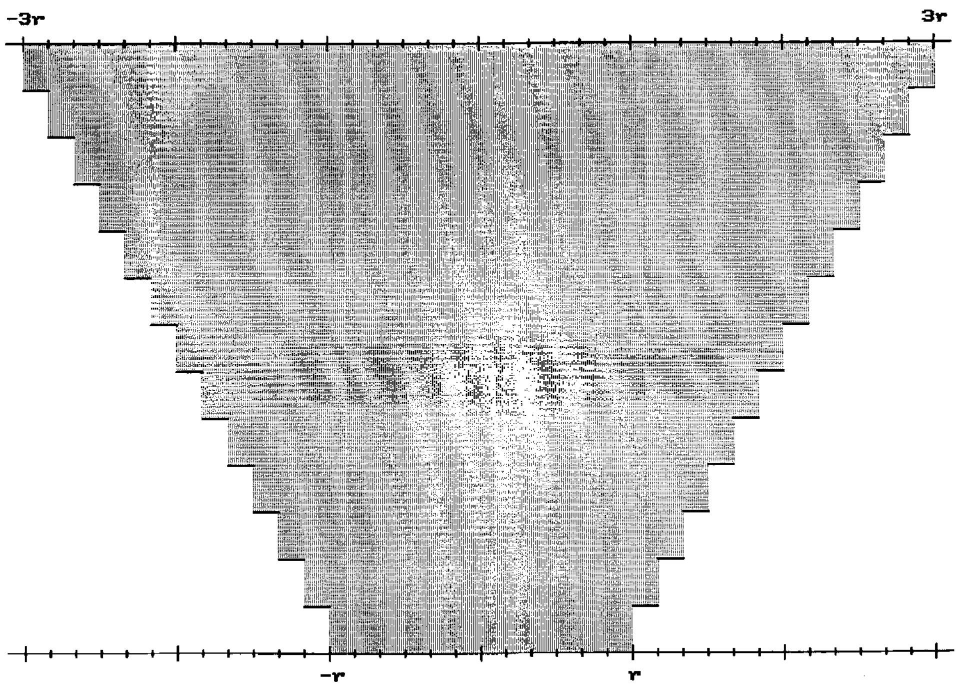


Figure 6.1

Since our good event is measurable with respect to the Poisson processes in $[-3r, 3r) \times [0, T)$ we have verified the comparison assumptions with $L = r$, $K = 3$, $J = 1$. If we start our cyclic system from uniform product measure then X_0 is a Bernoulli set with density $p > 0$. (p is close to 1 if L is large but we do not need that.) Applying Theorems 4.3 and 4.2 now it follows that

$$\liminf_{n \to \infty} P(0 \in X_n) \geq 19/20$$

Arguing as in the end of the proof of Theorem 5.4 it is easy to improve this conclusion to

$$\liminf_{n\to\infty} P(\text{ all } \kappa \text{ colors are in } [-r,r)) > 0$$

and it follows from (2.13) that there is anontrivial stationary distribution. By using an improvement of Theorem 4.2 given in the appendix (see Theorem A.3)

(6.1) **Lemma.** If $p > 0$ and $\gamma \le 6^{-4(2M+1)^2}$ then

$$\liminf_{n\to\infty} P(\{-2K,\dots,2K\} \cap W^p_{2n} \ne \emptyset) \ge 1 - \epsilon_K$$

where $\epsilon_K \to 0$ as $K \to \infty$

we can show that the stationary distribution we constructed concentrates on configurations in which there are infinitely many sites in each state. ((6.1) shows directly that with probability one each state appears somewhere in the configuration, but the distribution is stationary so if there were only finitely many sites in some state we would have positive probability of having 0 in that state a contradiction.)

By the arguments in the last paragraph it is enough to show that (a), (b), and (c) hold. The first step is proving the large deviations estimate.

(6.2) **Lemma.** *Let $X_1, \dots X_n$ be i.i.d. with $P(X_i = 1) = p$, $P(X_i = 0) = 1 - p$. Then*

$$P(X_1 + \dots + X_n \le n(p - \epsilon)) \le \exp(-\epsilon^2 n/2).$$

Remark. This result and its proof are standard but we need to know that the right hand side does not depend on p.

Proof: If $\alpha > 0$ then

$$P(X_1 + \dots + X_n \le n(p-\epsilon))\, e^{-\alpha n(p-\epsilon)} \le (pe^{-\alpha} + (1-p))^n$$

Taking log's, dividing by n, rearranging and then using $\log(1+x) \le x$ we have

$$\begin{aligned}\frac{1}{n}\log P(X_1 + \dots + X_n \le n(p-\epsilon)) &\le \alpha(p-\epsilon) + \log(1 + p(e^{-\alpha} - 1)) \\ &\le \alpha(p-\epsilon) + p(e^{-\alpha} - 1) = -\alpha\epsilon + p(e^{-\alpha} - 1 + \alpha)\end{aligned}$$

Now $e^{-\alpha} - 1 + \alpha = \alpha^2/2! - \alpha^3/3! + \dots \le \alpha^2/2$ for $0 < \alpha < 1$, so taking $\alpha = \epsilon$ and using $p \le 1$ gives

$$P(X_1 + \dots + X_n \le n(p-\epsilon)) \le \exp(-\epsilon^2 n/2)$$

and completes the proof of (6.2). □

Let Z_t be a copy of the single site Markov chain, let $p_t(i,j) = P_i(Z_t = j)$, and observe that $p_t(i,j) = p_t(0, j-i)$. Until the first time some house in $[-r,r)$ becomes unreasonable, the sites in each house in $[-r,r)$ behave like independent copies of the single site Markov chain so we consider a collection of $K = r\beta$ independent copies of Z_t and let v_i be the number of "sites" in state i at time 0. The expected number of sites in state j at time t is $w_j(t) = \sum_i v_i p_t(i,j)$. To prove (a) we apply (6.2) with $n = v_i \geq \sigma K$ to the sites that start in state i to see that with probability at least $1 - \exp(-\epsilon^2 \sigma K/2)$, at least $v_i(p_t(i,j) - \epsilon)$ of the sites that start in state i will be in state j at time t. Taking $\epsilon = (\sigma - \rho)$ and summing over i gives

$$\sum_i v_i(p_t(i,j) - \epsilon) \geq \sigma K \sum_i p_t(i,j) - K\epsilon \geq (\sigma - \epsilon)K = \rho K$$

since $\sum v_i = K$ and $\sum_i p_t(i,j) = \sum_i p_t(0, j-i) = 1$. So with probability at least $1 - \kappa \exp(-\epsilon^2 \sigma K/2)$, at least ρK sites will be in state j at time t.

The last bound is for a fixed time but it is easy to extend it to cover the interval $[0,T]$. Let $\delta = \epsilon^2 \sigma/2$, let $J = \exp(\delta K/2)$, and $t_k = k/J$ for $1 \leq k \leq JT$. The probability that the number of sites in state i is less than ρK at some time t_k is at most

$$\kappa JT \exp(-\epsilon^2 \sigma K/2) = \kappa T \exp(-\delta K/2)$$

The probability that two sites flip in some interval (t_{k-1}, t_k) is at most

$$JT\binom{K}{2} J^{-2} \leq K^3 T \exp(-\delta K/2).$$

When we never have two flips in any interval, the state at each $t \in (t_{k-1}, t_k)$ agrees with the state at one of the two endpoints. Combining the last two estimates we have that the probability a collection of K independent single site chains becomes unreasonable before time T

$$\leq (\kappa + K^3)T \exp(-\delta K/2)$$

Since the sites in $[-r,r)$ behave like independent single site chains until some house becomes unreasonable, the probability of the event in (a) is at least

$$1 - 2B(\kappa + K^3)T \exp(-\delta K/2)$$

The proof that the house $[r, r+K)$ will be good at time S is similar but simpler. If all the houses in $[-r,r)$ stay reasonable until time S then each site in $[r, r+K)$ always sees an above threshold number of sites of each color and flips to the next color at rate 1. We again consider a collection of K independent single site chains but this time starting from an arbitrary initial configuration. The choice of S guarantees that $p_S(i,j) \geq \sigma'$ so applying (6.2) to K i.i.d. random variables with $p = \sigma'$ we conclude that the fraction of sites in state j is at least σK with probability at least $1 - \kappa \exp(-(\sigma' - \sigma)^2 K/2)$. Once we know that with high probability $[r, r+K)$ is good at time S and all the houses in $[-r,r)$ are reasonable at all times in $[0,T]$, we can repeat the proof of (a) to conclude that the house $[r, r+K)$ stays reasonable at all times in $[S,T]$. This verifies (b) when $j = 1$ but by continuing in the same way we can prove the result for $2 \leq j \leq 2B$. Now (b) implies that all the houses in $[-3r+K, 3r-K)$ are reasonable at time $T-S$ we can repeat the proof that the house $[r, r+K)$ is good at time S to conclude that all the houses in $[-3r, 3r)$ are good at time T and the proof is complete. □

PROOF IN $d > 1$: Let $B_p(x,r) = \{y : \|x-y\|_p \leq r\}$. The key to the proof is the following fact, which basically says that large balls are almost flat.

(6.3) **Lemma.** *Suppose $\lambda < 1/2$. There are constants R_0, δ, and M_0, so that if $M \geq M_0$ and $R \geq R_0$ then for $x \in B_2(0,(R+\delta)M)$.*

$$|B_2(0, RM) \cap B_p(x, M)| \geq \lambda |B_p(x, M)|$$

PROOF: In one dimension we can take $R_0 = 1$ and $\delta = 1 - 2\lambda$. Turning to dimensions $d > 1$, let $Q = \{x \in R^d : \|x\|_p \leq 1\}$ and let q be its volume. To prove the result it is convenient to scale space by $1/M$ and translate so that x/M sits at the origin. Any $d-1$ dimensional hyperplane through the origin divides Q into two pieces with volume $q/2$. For $i = 1,2,3$ let $\lambda < \lambda_3 < \lambda_2 < \lambda_1 < 1/2$. By continuity, there is a $\delta > 0$ so that if a hyperplane passes within a distance δ of the origin then it divides Q into two pieces each of which has volume at least $q\lambda_1$. Another application of continuity shows that if R_0 is large and $D = B_2(y,r)$ with $r \geq R_0$ and $B_2(y,r) \cap B_2(0,\Delta) \neq \emptyset$ then the volume of $D \cap Q$ is at least $q\lambda_2$.

The last step is to argue that if M is large then the lattice behaves like the "continuum limit" considered above. Pick $\epsilon > 0$ so that if $D = B_2(y,r)$ is as above then the volume of $B_2(y, r-\epsilon) \cap (1-\epsilon)Q$ is always larger than $q\lambda_3$. Then pick M_0 so that $1/M_0 < \epsilon$ and if $M \geq M_0$ then $|B_p(0,M)|/qM^d < \lambda_3/\lambda$. Let $\mathcal{X} = (Z^d/M) \cap D \cap Q$. The first part of the choice of M_0 implies that if $M \geq M_0$ then

$$B_2(y, r-\epsilon) \cap Q(1-\epsilon) \subset \cup_{x \in \mathcal{X}}\ x + \left[\frac{-1}{2M}, \frac{1}{2M}\right]^d$$

so $M^{-d}|\mathcal{X}| \geq q\lambda_3 \geq \lambda |B_p(0,M)| M^{-d}$, by the second part of the choice of M_0 and the proof is complete. □

To use this lemma we pick $\lambda < 1/2$ and $\rho < 1/\kappa$ so that $\lambda\rho > a$, use (6.3) to pick R_0, Δ, M_0, and then pick $M_1 \geq M_0$ so that

(6.4) $\qquad \lambda\rho K^d |B_p(0, M_1)| > a|B_p(0, K(M_1+d))|$ holds for large K

Let $\sigma \in (\rho, 1/\kappa)$ and suppose that the range of interaction is $r = K(M_1+d)$. For $z \in \mathbf{Z}^d$ let

$$I_z = [z_1 K, (z_1+1)K) \times \cdots \times [z_d K, (z_d+1)K)$$

and call I_z a *house*. We say that a house is *good* at time 0 if it contains at least σK^d sites of each color. We say that ξ_0 is good if all the houses I_z, $z \in B_2(0, R_0 M_1)$ are good. This will be our event H when we apply Theorem 4.3.

We have set things up so that as long as each house in $B_2(0, R_0M_1)$ is *reasonable*, i.e. contains at least ρK^d sites of each color, each site in each house in $B_2(0,(R_0+\delta)M_1)$ sees at least θ sites of each color. To check this note if $z \in B_2(0,(R_0+\delta)M_1)$ then all the sites in any house I_w with $w \in B_2(0,R_0M_1) \cap B_p(z,M_1)$ are within p-norm distance $r=(M_1+d)K$ of each site in I_z. (To see note that $\|z-w\|_p \le M_1$ so $\|zK-wK\|_p \le M_1K$ and if we use 1 to denote a vector of 1's then $\|zK-(w+1)K\|_p \le (M_1+d)K$, with $p=1$ being the worst case.) By (6.3)

$$|B_2(0,R_0M_1) \cap B_p(x,M_1)| \ge \lambda |B_p(x,M_1)|$$

Multiplying the last inequality by ρK^d and using the choice of M_1 and K in (6.4) that

$$\begin{aligned}\rho K^d |B_2(0,R_0M_1) \cap B_p(x,M_1)| &\ge \lambda \rho K^d |B_p(x,M)| \\ &\ge a|B_p(x,K(M_1+d)| = \theta\end{aligned}$$

Pick B so that $B\delta > 2R_0$ and hence

$$B_p(x,(R_0+B\delta)M_1) \supset B_p(x,3R_0M_1)$$

Let $\sigma' \in (\sigma, 1/\kappa)$, choose S so that $p_S(0,i) \ge \sigma'$ for all i, and let $T = BS$. Let $D_j = B_2(0,(R_0+j\delta))$ (D is for disk) and $A_j = D_j - D_{j-1}$ (A is for annulus). By repeating the one dimensional proof we can show that with high probability

(a) All the houses in D_0 stay reasonable until time T.

(b) The houses in A_j will be good at time jS and stay reasonable to time T.

(c) All the houses in $D_B \supset B_2(0,3R_0)$ will be good at time T.

and the desired result follows from an application of Theorems 4.3 and 4.2 as before. □

We will now give the promised explanation of the conjecture for the case $\kappa = 2$. First consider the situation in $d=1$ and for ease of exposition call the two states "yellow" and "blue". As our proof shows if we have a sufficiently large interval of sites in which two colors occur with approximately equal frequency then the distribution of colors in this region will quickly converge to a product measure with density 1/2 and the region will expand, no matter what it encounters outside. For the region to expand we need $\theta = a|\mathcal{N}|$ with $a < 1/4$ for if $a > 1/4$ and all sites in $[r,2r)$ are yellow then the random region cannot expand since the site at r will have about $r/2 < a(2r+1)$ blue sites in its neighborhood. Applying the same reasoning to yellow sites in $x \in [br,r)$, who have about $(2-b)r/2$ blue sites in their neighborhood, we see that if $(2-b)/2 < 2a$, i.e. $b > 2-4a$ then the yellow sites in $[br,r)$ will not flip to blue but since $a < 1/2$ the blue sites will flip to yellow at rate one.

Similar reasoning applies to the system in $d > 1$ with $1/4 < a < 1/2$ and shows that a large enough ball of yellow sites will expand through a random region. The trouble with turning this into a proof is that we cannot guarantee that the blob will always find itself in competion with a random region. Indeed in a deterministic version of the threshold voter model in $d = 1$ (see Durrett and Steif (1993)) this naive picture is not correct since there are "blockades" that in some circumstances will stop the advance of blobs. However, we believe that this will not happen in random systems or in $d > 1$. In support of this claim, we note that Andjel, Mountford, and Liggett (1992) have shown that clustering occurs in $d = 1$ when $\mathcal{N} = \{-k, \ldots, k\}$ and $\theta = k$. The important special property of this example is that if an interval of 1's (or 0's) is long enough only the site on either end can flip.

7. Long Range Limits

In the last section, we saw that the cyclic color model and Greenberg Hastings models simplified considerably when the range of interaction was large. In this section we show that the contact process also simplifies in this way.

Example 7.1. The basic contact process. As usual the neighborhood is $\mathcal{N} = \{x : \|x\|_p \leq r\}$. We will write the contact process as a set valued process with the state = the set of sites occupied by particles and formulate the dynamics as follows:

(i) Each particle dies at rate 1, and gives birth at rate β.

(ii) A particle born at x is sent to a site y chosen at random from $x + \mathcal{N}$.

(iii) If y is vacant, it becomes occupied. If y is already occupied the birth has no effect.

If r is large and the contact process starts from a single occupied site then at least until the number of particles is a significant fraction of $|\mathcal{N}|$, the contact process will behave like a *branching random walk*, i.e., the process that obeys (i) and (ii) but allows any number of particles per site.

The total number of particles at time t in a branching random walk is a *branching process* – a Markov chain Z_t in which transitions from k to $k+1$ occur at rate $k\beta$ and transitions k to $k-1$ occur at rate k. Let $T_y = \inf\{t : Z_t = y\}$ and use P_x to denote the law of the branching process with $Z_0 = x$. Well known properties of the exponential distribution imply that

$$P_k(T_{k+1} < T_{k-1}) = \frac{\beta}{\beta+1} \quad \text{for } k > 0$$

so Z_t is a time change of an asymmetric random walk S_n that, when $k > 0$, makes transitions $k \to k+1$ with probability $\beta/(\beta+1)$ and $k \to k-1$ with probability $1/(\beta+1)$ and has 0 as an absorbing state, i.e., once $S_n = 0$ we will have $S_m = 0$ for all $m > n$. Using this observation and well known formulas for simple random walk it follows that

$$P_1(T_0 < \infty) = \begin{cases} 1 & \text{if } \beta \leq 1 \\ 1/\beta & \text{if } \beta > 1 \end{cases}$$

so the critical value of β for the survival of the branching process is 1.

The main result in this section is that as the range $r \to \infty$ the critical value for survival of the contact process converges to that of the branching process. Let $\tau^0 = \inf\{t : \xi_t^0 = \emptyset\}$ where ξ_t^0 denotes the contact process starting from a single particle at the origin, i.e., $\xi_0^0 = \{0\}$. Let $\beta_c = \inf\{\beta : P(\tau^0 = \infty) > 0\}$.

Theorem 7.1. As $r \to \infty$, $\beta_c \to 1$ and if $\beta > 1$

$$P(x \in \xi_\infty^1) \to \frac{\beta - 1}{\beta}$$

Remark. Schonmann and Vares (1986) have shown that if we consider the basic contact process in d dimensions with $\mathcal{N} = \{x : \|x\|_1 = 1\}$ and we let $\beta = 2d\lambda$ then the conclusions of Theorem 7.1 and (7.18) below hold.

PROOF: To begin we note that we can construct the contact process from a branching random walk by suppressing births onto occupied sites. So we can define the contact process and the branching random walk on the same space so that the branching random walk always has more particles than the contaact process, and it follows that $\beta_c \geq 1$ for all r. To prove the rest of the result we note that taking $A = \mathbf{Z}^d$ and $B = \{0\}$ in the duality equation (5.3) gives

$$P(\xi_t^1 \cap \{0\} \neq \emptyset) = P(\xi_t^0 \cap \mathbf{Z}^d \neq \emptyset) = P(\tau^0 > t)$$

Letting $t \to \infty$ we have

$$P(0 \in \xi_\infty^1) = P(\tau^0 = \infty) \tag{7.1}$$

So to prove Theorem 8.1 it suffices to show that

(7.2) If $\beta > 1$ then $P(\tau^0 = \infty) \to (\beta - 1)/\beta$

for this implies that $\limsup_{r\to\infty} \beta_c \leq 1$. To prove (8.2) we scale space by dividing by r and consider the contact process on $\mathbf{Z}^d/r$ to facilitate taking the limit $r \to \infty$. Our approach will be to use the comparison theorem, so we let $I_k = k2Le_1 + [-L, L]^d$ and consider a modification of the contact process $\bar{\xi}_t$ in which births are not allowed outside $(-4L, 4L)^d$. The two key ingredients in the proof are

(7.3) Let $\delta > 0$. If we pick L large, set $T = L^2$, and pick K large then for $r \geq r_0$, $\bar{\xi}_T$ will have at least K particles in I_1 and in I_{-1} with probability at least $1 - \delta$ whenever $\bar{\xi}_0$ has at least K particles in I_0

(7.4) Consider the process starting from $\xi_0^0 = \{0\}$. If we pick S large then for $r \geq r_1 \geq r_0$, ξ_S^0 will have at least K particles in I_0 with with probability at least $((\beta - 1)/\beta) - \delta$

Once this is done (7.2) follows by using Theorem 4.3 to compare

$$X_n = \{m : |\xi_{S+nT}^0 \cap I_m| \geq K\}$$

with a one-dependent oriented percolation with density $\geq 1 - \delta$ and Theorem 4.1 to conclude that the cluster containing $(0,0)$ in the percolation model will be infinite with probability at least $1 - 55\delta^{1/9}$. For these two facts imply that

$$P(\tau^0 = \infty) \geq \frac{\beta - 1}{\beta} - \delta - 55\delta^{1/9}$$

PROOF OF (7.3): The starting point is the observation that if we let $r \to \infty$ then the contact process on $\mathbf{Z}^d/r$ converges to a branching random walk η_t in which

(i) Each particle dies at rate 1, and gives birth at rate β.

(ii) A particle born at x is sent to a point y chosen at random from $\{y : \|y - x\|_p \le 1\}$.

This should be intutively clear since if we start with one particle at 0, fix T and let $r \to \infty$ then the probability of a *collision* (birth onto an occupied site) by time T goes to 0 as $r \to \infty$, and the displacements of the individual particles converge to a uniform distribution on $\{y : \|y\|_p \le 1\}$.

We will prove the convergence of the contact process on $\mathbf{Z}^d/r$ to the branching random walk later (see the "continuity argument" below). We have introduced this result now to motivate the first step of the proof, which is to prove the analogue of (7.3) for the branching random walk η_t, which is given in (7.12) below. Let η_t^x denote the branching random walk starting from $\eta_0^x = \{x\}$. To leave room for the limit $r \to \infty$ we consider $\bar\eta_t^x$ a modification of η_t^x in which particles that land outside $(-4L+1, 4L-1)^d$ are killed. Let $m(t,x,A) = E|\bar\eta_t^x \cap A|$ be the mean number of particles in A at time t for the modified branching random walk starting with a single particle at x. We claim that

(7.5) $$m(t,x,A) = e^{(\beta-1)t} P(\bar W_t^x \in A)$$

where $\bar W_t^x$ is a random walk that starts at x, jumps at rate β, has jumps that are uniform on $\{y : \|y\|_p \le 1\}$, and is killed when it lands outside $(-4L+1, 4L-1)^d$. To check this claim note that both sides of (7.5) satisfy the same differential equation: if $A \subset (-4L+1, 4L-1)^d$ then

$$\frac{dm(t,x,A)}{dt} = -m(t,x,A) + \int m(t,x,dy)\nu(A-y)$$

where $A - y = \{x - y : x \in A\}$ and ν is the uniform probability measure on $\{y : \|y\|_p \le 1\}$.

Let $I_1' = 2Le_1 + [-L+1, L-1]^d$, i.e. I_1 shrunk by a little bit. Donsker's theorem implies that if $T = L^2$ and $x/L \to \theta \in [-1,1]^d$

(7.6) $$P(\bar W_T^x \in I_1') \to \psi(\theta)$$

where $\psi(\theta) = P_\theta(B_t \in [-4,4]^d \text{ for } t \le 1,\ B_1 \in 2e_1 + [-1,1]^d)$ and B_t is a constant multiple of d-dimensional Brownian motion. $\psi(\theta) > 0$ and is continuous, so a simple argument (suppose not and extract a convergent subsequence) shows

(7.7) $$\liminf_{L\to\infty} \left[\inf_{x\in[-L,L]^d} P(\bar W_T^x \in I_1') \right] \ge \inf_{\theta\in[-1,1]^d} \psi(\theta) > 0.$$

It follows from (7.5)-(7.7) that we can pick L large enough so that

(7.8) $$\inf_{x\in[-L,L]^d} E|\bar\eta_T^x \cap I_1'| \ge 2.$$

Let $\bar\eta_t^A$ denote the modified branching random walk with $\bar\eta_0^A = A$. (7.8) implies

(7.9) $$E|\bar\eta_T^A \cap I_1'| \ge 2|A|$$

while an obvious comparison and a well known fact about branching processes (see Athreya and Ney (1972) for this and other facts about branching processes we will use) implies

$$(7.10)\qquad \mathrm{var}(\bar\eta_T^x \cap I_1') \le E|\bar\eta_T^x \cap I_1'|^2 \le E|\eta_T^x|^2 = C_T < \infty$$

Combining the last two conclusions and using Chebyshev's inequality it follows that if $A \subset [-L,L]^d$ has $|A| = K$ then

$$(7.11)\qquad P(|\bar\eta_T^A \cap I_1'| < K) \le \frac{\mathrm{var}(|\bar Z_T^A \cap I_1'|)}{(2|A|-K)^2} \le \frac{K \sup_x \mathrm{var}(|\bar Z_T^x \cap I_1'|)}{K^2} \le \frac{C_T}{K}$$

From the last result it follows that

(7.12) If $\delta > 0$ and K is large then for any $A \subset [-L,L]^d$ with $|A| = K$.

$$P(|\bar\eta_T^A \cap I_1'| < K) \le \delta/10$$

Continuity Argument. (7.12) shows that if $A \subset I_0$ has $|A| = K$ then with high probability $\bar\eta_T^A$ will have at least K particles in I_{-1} and in I_1. The next step is to prove the corresponding result for the contact process. To avoid some technicalities we will give the details only for the case in which $\mathcal{N} = \{z : \|z\|_\infty \le r\}$ and then indicate the extension to $p < \infty$ in a remark after the proof.

Let $\bar\xi_t^A$ be a modification of the contact process with $\bar\xi_0^A = A$ in which births outside $(-4L,4L)^d$ are not allowed. We begin by observing that the number of births up to time t in the contact process, V_t, is dominated by a branching process $\bar V_t$ in which births occur at rate β and deaths occur at rate 0. If $|A| = K$ then $E\bar V_t = Ke^{\beta t} < \infty$, so our comparison and Chebyshev's inequality imply

$$(7.13)\qquad P(V_T > r^{1/3}) \le P(\bar V_T > r^{1/3}) \le \frac{Ke^{\beta T}}{r^{1/3}} \to 0$$

since T is fixed and $r \to \infty$.

Let $G_1 = \{V_t \le r^{1/3}\}$. Here G is for good event and the subscript indicates it is the first of several we will consider. When G_1 occurs, the probability of having a birth land on an occupied site (a "collision") is

$$(7.14)\qquad \le r^{1/3}\frac{r^{1/3}}{(2r+1)^d} \to 0$$

since there are at most $r^{1/3}$ births and even if all the particles are in $\{x : \|x\|_\infty \le 1\}$ (on $\mathbf{Z}^d/r$) each birth has probability at most $r^{1/3}/(2r+1)^d$ of landing on an occupied site. Let G_2 be the event that there are no collisions by time t.

To deal with the spatial location of particles, we will create a coupling of the displacements of the particles in the branching random walk to those of particles in the contact process. To couple the displacements we observe that if U is uniform on $\{y : \|y\|_\infty \le 1\}$

and $\pi_r(x)$ is the closest point in $\mathbf{Z}/r^d$ to x (with some convention for breaking ties) then $U^r = \pi_r(U(1+1/2r))$ is uniform on $\mathcal{N}/r$.

Now if the U_i are the displacements of particles in the branching random walk, we will use the U_i^r for the displacements in the contact process. When our good events G_1 and G_2 occur, we have $G_3 =$ all of the points in the contact process ξ_s^A are within $r^{1/3}/r$ (in $\| \|_\infty$) of their counterparts in the branching process η_s^A. Passing to the truncated processes and noting that the branching particles are required to stay in $(-4L+1, 4L-1)^d$ for $0 \le s \le T$, while the contact process particles are required to stay in $(-4L, 4L)^d$, it follows that on G_3 we have $|\bar{\xi}_T^A \cap I_1| \ge |\bar{\eta}_T^A \cap I_1'|$ Combining the last observation with (7.12) gives (7.3). □

Remark. If $p < \infty$ then $U^r = \pi_r((1+1/2r)U)$ is not uniform on $\mathcal{N}/r$ but is within C/r of uniform in the total variation norm. In the last paragraph of the proof we then have $P(\|U_i - U_i^r\|_\infty > 1/r) \le C/r$, which since there are at most $r^{1/3}$ transitions on G_1, is good enough for the proof.

PROOF OF (7.4): By the continuity argument it is enough to show that we can pick S so that η_S^0 will have at least K particles in I_0 with probability at least $((\beta-1)/\beta) - \delta/2$. However, this follows from

(7.16) If Ω_∞ is the event that the branching process does not die out, then for any $L > 0$ and $K < \infty$,

$$P(|\eta_t^0 \cap [-L, L]^d| < K,\, \Omega_\infty) \to 0$$

Indeed as Asmussen and Kaplan (1976) have shown (see Theorem 2 on p. 5)

(7.17) There is a constant $\sigma > 0$ so that

$$\sqrt{t}e^{-(\beta-1)t}|\eta_t^0 \cap [-L, L]^d \to W \cdot \frac{(2L+1)^d}{(2\pi\sigma^2)^{d/2}}$$

where $W = \lim_{t\to\infty} e^{-(\beta-1)t}|\eta_t^0| > 0$ a.s. on Ω_∞

This completes the proof of (7.4) and hence of (7.2). □

The argument just used on the long range contact process can also be applied to

Example 7.2. Succesional dynamics. We suppose that the set of states at each site are 0 = grass, 1 = a bush, 2 = a tree and formulate the dynamics as

$$\begin{aligned} c_0(x,\xi) &= \delta_{\xi(x)} \\ c_1(x,\xi) &= \lambda_1 n_1(x) \quad \text{if } c_i(x) = 0 \\ c_2(x,\xi) &= \lambda_2 n_2(x) \quad \text{if } c_i(x) \le 1 \end{aligned}$$

The title of this example and its formulation are based on the observation that if an area of land is cleared by a fire, then regowth will occur in three stages: first grass appears then small bushes and finally trees, with each species growing up through and replacing the previous one. With this in mind, we allow each type to give birth onto sites occupied by lower numbered types.

Theorem 7.2. Let $\beta_i = \lambda_i|\mathcal{N}|$. Suppose that $\beta_2 > \delta_2$ and

$$(\star) \qquad \beta_1 \cdot \frac{\delta_2}{\beta_2} > \delta_1 + \beta_2 \cdot \frac{\beta_2 - \delta_2}{\beta_2}$$

If r is large then there is a nontrivial translation invariant stationary distribution in which all three types have positive density.

Sketch of Proof: The fact that the 2's do not feel the presence of the 1's implies that the set of sites occupied by 2's is a contact process. To construct a stationary distribution we start with the 2's in their upper invariant measure and we put 1's at all sites not occupied by 1's to get a process ξ_t^{12}. This is the analogue of starting an attractive system from all 1's and a result of Durrett and Moller imples that as $t \to \infty$, $\xi_t^{12} \Rightarrow \xi_\infty^{12}$ a translation invariant stationary distribution.

To prove that ξ_∞^{12} is nontrivial we will prove an analogue of (7.3). The first step is to prove the following result about the long range contact process (which is here considered as a subset of $\mathbf{Z}^d$)

(7.18) If $\beta > 1$ and $x \neq y$ then as $r \to \infty$

$$P(x, y \in \xi_\infty^1) \to \left(\frac{\beta - 1}{\beta}\right)^2$$

In words, the equilibrium distribution converges to a product measure as $r \to \infty$. Of course, the last conclusion only says that the sites are asymptotically pairwise independent, but the argument can easily be generalized to a finite number of x's.

Proof: By duality (see the proof of (7.1))

$$P(x, y \in \xi_\infty^1) = P(\tau^x = \infty, \tau^y = \infty)$$

Our comparison of the contact process with a branching process at the beginning of the proof of Theorem 8.1 shows that $P(\tau^x = \infty) \leq (\beta - 1)/\beta$ for all r. If we pick K and L as in (7.3) and then pick S large as in (7.4) then for $r \geq r_1$ we have

$$\begin{aligned} \frac{\beta - 1}{\beta} + \delta &\geq P(\tau^x > S) \\ &\geq P(|\xi_S^x \cap [-L, L]| > K) \geq \frac{\beta - 1}{\beta} - \delta \end{aligned}$$

Our choice of K and L and the comparison with oriented percolation shows that

$$P(|\xi_S^x \cap [-L, L]| > K, \tau^x < \infty) \leq 55\delta^{1/9}$$

Combining the last two estimates shows

$$|P(\tau^x = \infty) - P(\tau^x > S)| \le \delta + 55\delta^{1/9}$$

With this in hand the desired result follows easily since continuity argument shows that for any fixed S as $r \to \infty$

$$P(\tau^x > S, \tau^y > S) \to P(\eta_S^0 \ne \emptyset)^2$$ □

Turning now to the heart of the proof we will again scale space by dividing by r and consider the contact process on $\mathbf{Z}^d/r$ to facilitate taking the limit. The approach we will take is a combination of that of Durrett and Swindle (1991) and Durrett and Schinazi (1993). We will concentrate on explaining the main ideas and refer the reader to those papers for the details. Pick $\rho > (\beta_2 - \delta_2)/\beta_2$ so that

$$(\star\star) \qquad \beta_1(1-\rho) > \delta_1 + \beta_2\rho$$

By dividing space into cubes of side δr then using (7.18) and the weak law one can prove that with high probability all sites in our space time box have at most $\rho|\mathcal{N}|$ neighbors in state 2. (Recall that the set of 2's at any time is distributed according to the upper invariant measure.) This means that a single 1 will have births that land on an occupied site at rate $\ge \beta_1(1-\rho)$ while it dies at rate δ_1 and is smothered by a 2 at rate $\le \beta_2\rho$.

The inequality $(\star\star)$ implies that a single particle gives birth faster than it dies. If we start with a fixed number of 1's then in the limit $r \to \infty$ the 1's dominate a supercritical branching random walk. If this fixed number K is large and L and $T = L^2$ are chosen appropriately then for large r a truncated version of the process which is not allowed to give birth outside $(-4L, 4L)^d$ will with high probability have at least K particles in I_1 and in I_{-1} whenver the initial configuration has at least K particles in I_0.

The last result is an analogue of (8.3) but there is one problem. The event that $\xi_t(x) = 2$, which is the same as the survival of the dual contact process of 2's starting from (x,t), does not have a finite range of dependence. To avoid this problem we adopt the more liberal viewpoint that x is occupied by a 2 at time t if the dual process escapes from a certain space-time box. If the box is large enough the liberalization of the definition does not increase the density of 2's by enough to violate $(\star\star)$, we can verify the comparison assumptions of Theorem 4.3 and the desired result follows from Theorem 8.2.

8. Rapid Stirring Limits

The point of this section is that if we take a fixed interacting particle system, scale space by ϵ and "stir" the particles at rate ϵ^{-2} then as $\epsilon \to 0$ the particle system converges to the solution of a reaction diffusion equation. To be precise, we consider processes $\xi_t^\epsilon : \epsilon \mathbf{Z}^d \to \{0, 1, \ldots, \kappa - 1\}$ that evolve as follows

(i) there are *translation invariant finite range flip rates*

$$c_i(x, \xi) = h_i(\xi(x), \xi(x + \epsilon y_1), \ldots, \xi(x + \epsilon y_N))$$

(ii) *rapid stirring:* for each $x, y \in \epsilon \mathbf{Z}^d$ with $\|x - y\|_1 = \epsilon$ we exchange the values at x and y at rate ϵ^{-2}. That is, we change the configuration from ξ to $\xi^{x,y}$ where

$$\xi^{x,y}(y) = \xi(x) \qquad \xi^{x,y}(x) = \xi(y) \qquad \xi^{x,y}(z) = \xi(z) \quad z \neq x, y$$

The reader should note that in (i) changing ϵ scales the lattice but does not change the interaction between the sites. In (ii) we superimpose stirring in such a way that the individual values will be moving according to Brownian motions (run at rate 2) in the limit. The motivation for modifying the system in this way comes from the following *mean field limit theorem* of De Masi, Ferrari, and Lebowitz (1986). The derivation of such "hydrodynamic limits" has become a major enterprise (see e.g., Spohn (1991) or DeMasi and Presutti (1992)) but this particular result is rather easy to establish.

Theorem 8.1. Suppose $\xi_0^\epsilon(x)$ are independent and let $u_i^\epsilon(t, x) = P(\xi_t^\epsilon(x) = i)$. If $u_i^\epsilon(0, x) = g_i(x)$ is continuous then as $\epsilon \to 0$, $u_i^\epsilon(t, x) \to u_i(t, x)$ the bounded solution of

$$\partial u_i / \partial t = \Delta u_i + f_i(u) \qquad u_i(0, x) = g_i(x) \tag{8.1}$$

where

$$f_i(u) = < c_i(0, \xi) 1_{(\xi(0) \neq i)} >_u - \sum_{j \neq i} < c_j(0, \xi) 1_{(\xi(0) = i)} >_u \tag{8.2}$$

and $< \phi(\xi) >_u$ denotes the expected value of $\phi(\xi)$ under the product measure in which state j has density u_j, i.e., when $\xi(x)$ are i.i.d. with $P(\xi(x) = j) = u_j$.

Theorem 8.1 is easy to understand. The stirring mechanism (i.e., (ii)) has product measures as its stationary distributions. See Griffeath (1979), Section II.10. When ϵ is small, stirring operates at a fast rate and keeps the system close to a product measure. The rate of change of the densities can then be computed assuming adjacent sites are independent. To help explain the somewhat ugly formula in (8.2) we will now consider two concrete examples.

Example 8.1. The basic contact process. In this case $c_0(x, \xi) = 1$ and $c_1(x, \xi) = \lambda n_1(x, \xi)$ where $n_i(x, \xi) = |\{y \in \mathcal{N} : \xi(x + y) = i\}$ is the number of neighbors in state i.

We claim that when $|\mathcal{N}| = N$ the equation in (9.1) becomes (we do not need an equation for $u_0 = 1 - u_1$)

$$\partial u_1/\partial t = \Delta u_1 - u_1 + N\lambda(1 - u_1)u_1$$

To see the second term on the right hand side the equation, we note that particles die at rate 1 independent of the state of neighbors. For the third, we note that if we assume all sites are independent then the probability x is vacant and $y \in x + \mathcal{N}$ is occupied is $(1 - u_1)u_1$. Each such pair produces a new particle at rate λ and there are N such pairs, so the total rate at which new particles are created (assuming that adjacent sites are independent) is $N\lambda(1 - u_1)u_1$.

The equation in the last example is just the mean field equation for the contact process that we have seen several times before. To see something new we look at

Example 8.2. The threshold one voter model. In this case

$$c_i(x, \xi) = 1 \quad \text{if } n_i(x, \xi) \geq 1$$

and if we assume $|\mathcal{N}| = N$ then the limiting equation is (again we do not need an equation for $u_0 = 1 - u_1$)

$$\partial u_1/\partial t = \Delta u_1 - u_1(1 - u_1^N) + (1 - u_1)(1 - (1 - u_1)^N)$$

To see this note that if all sites are independent then the probability x is occupied and at least one neighbor is vacant is $u_1(1 - u_1^N)$ and this is the rate at which 1's are destroyed. Intercahnging the roles of vacant and occupied in the last sentence gives the third term.

Having explained the formula in (8.2) we turn now to a result that extends Theorem 8.1 by showing that the particle system itself, not just its expected values are close to the p.d.e. To motivate the statement we note that the states of the sites in the model become independent in the limit $\epsilon \to 0$ and the number of sites per unit volume becomes large so it should not be surprising that in the limit $\xi_t^\epsilon(x)$ becomes deterministic.

Theorem 8.2. Let $\phi(x)$ be a smooth function with compact support. As $\epsilon \to 0$

$$\epsilon^d \sum_{y \in \epsilon \mathbf{Z}^d} \phi(y) 1_{(\xi_t^\epsilon(y)=i)} \to \int \phi(y) u_i(t, y)\, dy$$

in probability.

Although the indicator function of a bounded open set G is not continuous, this should be thought of as saying that

$$\epsilon^d \sum_{y \in \epsilon \mathbf{Z}^d \cap G} 1_{(\xi_t^\epsilon(y)=i)} \to \int_G u_i(t, y)\, dy$$

or more intuitively that the fraction of sites near y that are in state i converges to $u_i(t,y)$. The result for an open set G is also true, but is a little more difficult to prove precisely because 1_G is not continuous.

Theorem 8.2 provides a link between the particle system with fast stirring that we will exploit in the next lecture to prove the existence of stationary distributions for a predator prey model with fast stirring. Once Theorem 8.1 is established, the proof of Theorem 8.2 is easy: compute second moments and use Chebyshev's inequality. So we will concentrate on the proof of Theorem 8.1. The ideas behind the proof are simple: we will give an explicit construction of the process that allows us to define a dual processes by asking the question: "What is the state of x at time t?" and working backwards in time. The answer to this question can be determined by looking at the states of the sites in the "dual process" $I_\epsilon^{x,t}(s)$ at time $t-s$. The particles in $I_\epsilon^{x,t}(s)$ move according to stirring at a fast rate and give birth to new particles at rate

$$c^* = \sup_\xi \sum_i c_i(x,\xi)$$

We will show that for small ϵ the dual process is almost a branching random walk and converges to a branching Brownian motion as $\epsilon \to 0$. The proof of the last result leads easily to the conclusion that two dual processes $I_\epsilon^{x,t}(s)$ and $I_\epsilon^{y,t}(s)$ are asymptotically independent which gives the asymptotic independence of the sites in the parrticle systems. The convergence of the dual process to branching Brownian motion leads in a straightforward way to the convergence of the $u_i^\epsilon(t,x)$ to limits $u_i(t,x)$ and the asymptotic independence of adjacent sites implies that the $u_i(t,x)$ satisfy the limiting equations.

a. The dual process. The first step in the proof is to construct the process from a number of Poisson processes, all of which are assumed to be independent. The construction is similar in spirit to the one in Section 3 but it is convenient to do the details in a slightly different way. For each $x \in \epsilon\mathbf{Z}^d$, let $\{T_n^x, n \ge 1\}$, be a Poisson process with rate c^* and let $\{U_n^x, n \ge 1\}$ be a sequence of independent random variables that are uniform on $(0,1)$. At time T_n^x we compute the flip rates $r_i = c_i(x, \xi(T_n^x))$ and use U_n^x to determine what (if any) flip should occur at x at time T_n^x. To be precise we let $p_i = \sum_{j\le i} r_j/c^*$ for $i = 0,\dots,\kappa-1$ with $p_{-1} = 0$ and flip to i if $U_n^x \in (p_{i-1}, p_i)$. If $U_n^x \in (p_{\kappa-1}, 1)$ no flip occurs. To move the particles around, we let $\{S_n^{x,y}, n \ge 1\}$ be Poisson processes with rate ϵ^{-2} when $x, y \in \epsilon\mathbf{Z}^d$ with $\|x-y\|_1 = \epsilon$, and we declare that at time $S_n^{x,y}$ the values at x and y are exchanged.

The dual process $I_\epsilon^{x,t}(s)$ is naturally defined only for $0 \le s \le t$ but for a number of reasons, it is convenient to assume that the Poisson processes and uniform random variables in the construction are defined for negative times and define $I_\epsilon^{x,t}(s)$ for all $s \ge 0$. Let $\mathcal{N} = \{\epsilon y_1, \dots, \epsilon y_N\}$ be the set of neighbors of 0. The dual process makes transitions as follows:

If $y \in I_\epsilon^{x,t}(s)$ and $T_n^y = t-s$ then we add all the points of $y + \mathcal{N}$ to $I_\epsilon^{x,t}(s)$.

If $y \in I_\epsilon^{x,t}(s)$ and $S_n^{y,z} = t-s$ then we move the particle at y to z.

For a picture of (a rather unlikely sample path for) the dual when $d=1$ and $\mathcal{N} = \{-1,0,1\}$ see Figure 8.1

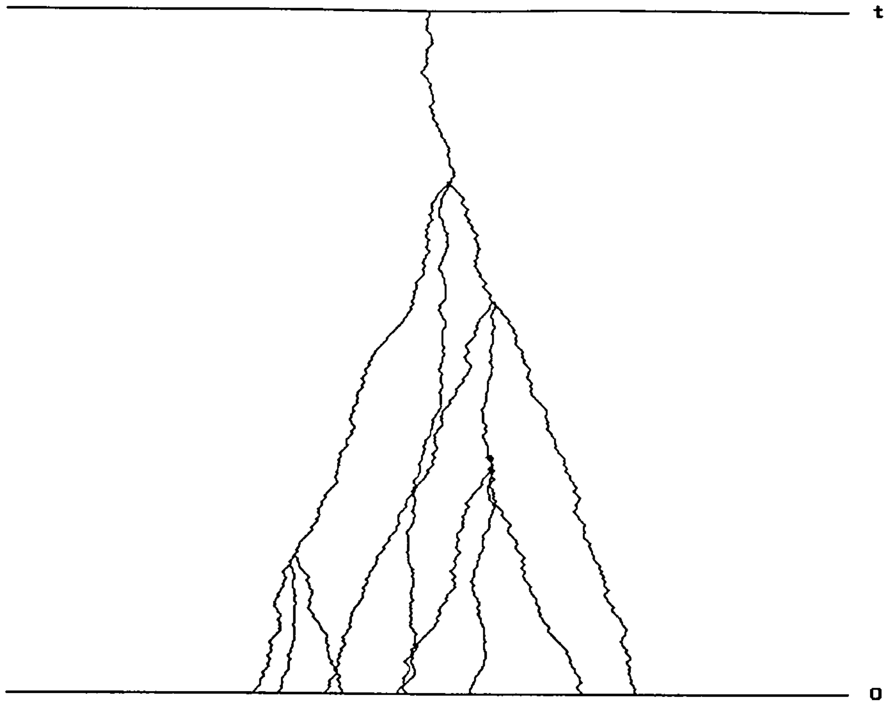


Figure 8.1

It is easy to see that we can compute the state of x at time t by knowing the states of the y in $I_\epsilon^{x,t}(s)$ at time $t-s$. We start with the values in $I_\epsilon^{x,t}(s)$ at time $t-s$ and work up to time t. At S arrivals we perform the indicated stirrings. When an arrival T_n^y occurs at a point of the dual, we look at the value of the process on $y+\mathcal{N}$, compute the flip rates r_i, and use U_n^x to determine what (if any) flip should occur.

To prepare for the proof of the convergence of $u_i^\epsilon(t,x)$ we will now give a more detailed description of $I_\epsilon^{x,t}(s)$. Let $X_\epsilon^0(0)=x$, let R_ϵ^1 be the smallest value of s so that we have a T arrival at $X_\epsilon^0(s)$ at time $t-s$, and set $X_\epsilon^i(s)=\epsilon y_i + X_\epsilon^0(s)$ for $1 \le i \le N$. Finally, we set $\mu_\epsilon^1 = 0$ to indicate that 0 is the mother of the N particles created at time R_ϵ^1. Passing now to the inductive step of the definition, suppose that we have defined the process up to time R_ϵ^m with $m \ge 1$. The $mN+1$ existing particles move as dictated by stirring until R_ϵ^{m+1}, the first time $s > R_\epsilon^m$ that a T arrival occurs at the location of one of our moving particles $X_\epsilon^k(s)$ and then we set $X_\epsilon^{mN+i}(s) = \epsilon y_i + X_\epsilon^k(s)$ for $1 \le i \le N$, and $\mu_\epsilon^{m+1} = k$. The new particles may be created at the locations of existing particles. If so we say that a *collision* occurs and call the new particle *fictitious*. We will prove later that the probability of a collision tends to 0 as $\epsilon \to 0$, but for proving the convergence of $u_i^\epsilon(t,x)$, it is convenient to allow the fictitious particles to move and give birth like other particles, so for each $m \geq 1$ we define an independent copy of the graphical representation which we use for the births and movement of the mth particle if it is fictitious. By definition all the offspring of fictitious particles are also fictitious.

b. The dual process is almost a branching random walk. The point of introducing fictitious particles is that $\mathcal{K}_t = mN + 1$ for $t \in [R^\epsilon_m, R^\epsilon_{m+1})$ defines a branching process in which each particle gives birth to N additional particles at rate c^*. Our next goal is to show that if ϵ is small then $I^{x,t}_\epsilon(s)$ is almost a branching random walk in which particles jump to a randomly chosen neighbor at rate $2d\epsilon^{-2}$ and give birth as above. To do this we will couple X^k_ϵ to independent random walks Y^k_ϵ that start at the same location at time β_k = the birth time of X^k_ϵ, and jump to a randomly chosen neighbor at rate $2d\epsilon^{-2}$.

We say X^k_ϵ is *crowded* at time s if for some $j \neq k$ $\|X^j_\epsilon(s) - X^k_\epsilon(s)\|_1 \leq \epsilon$. When X^k_ϵ is not crowded, we define the displacements of Y^k_ϵ to be equal to those of X^k_ϵ. When X^k_ϵ is crowded we use independent Poisson processes to determine the jumps of Y^k_ϵ. To estimate the difference between X^k_ϵ and Y^k_ϵ, we need to estimate the amount of time X^k_ϵ is crowded. Let $j \neq k$, $V^\epsilon_s = X^k_\epsilon(s) - X^j_\epsilon(s)$ and W^ϵ_s be a random walk that jumps to a randomly chosen neighbor at rate $4d\epsilon^{-2}$. (Notice that V^ϵ_s is the difference of two random walks and hence jumps at rate $4d\epsilon^2$. The transition probabilities of V^ϵ_s differ slightly from those of W^ϵ_s when $\|x\|_1 = \epsilon$. Here y denotes any point $\neq -x$ with $\|y\|_1 = \epsilon$.

jumps from x to	rate in V	rate in W
$-x$	ϵ^{-2}	0
0	0	$2\epsilon^{-2}$
$x+y$	$2\epsilon^{-2}$	$2\epsilon^{-2}$

From the last table it should be clear that $\|W^\epsilon_s\|_1$ is stochastically smaller than $\|V^\epsilon_s\|_1$, i.e., the two random variables can be constructed on the same space so that $\|W^\epsilon_s\|_1 \leq \|V^\epsilon_s\|_1$ for all s. To check this note that all the transition of V and W can be coupled except those in the first two lines of the table, but there $\|W\|_1$ jumps from 1 to 0 at rate ϵ^{-2} while $\|V\|_1$ jumps from 1 to 1 at rate $\epsilon^2/2$.

From the last comparison of $\|V\|_1$ and $\|W\|_1$ it follows that for any integer $M \geq 1$, $v^{M\epsilon}_t = |\{s \leq t : \|V^\epsilon_s\|_1 \leq M\epsilon\}|$ is stochastically smaller than $w^{M\epsilon}_t = |\{s \leq t : \|W^\epsilon_s\|_1 \leq M\epsilon\}|$. Well known asymptotic results for random walks imply that when $t\epsilon^{-2} \geq 2$

$$Ew^{M\epsilon}_t \leq \begin{cases} CM^d\epsilon^2 & d \geq 3 \\ CM^2\epsilon^2\log(t\epsilon^{-2}) & d = 2 \\ CM\epsilon t^{1/2} & d = 1 \end{cases} \tag{8.3}$$

To see this note that $w^{M\epsilon}_t$ has the same distribution as $\epsilon^2 w^M_{t\epsilon^{-2}}$ and the last line is equal to $CM\epsilon^2(t\epsilon^{-2})^{1/2}$.

Let $\chi^k_\epsilon(t)$ be the amount of time X^k_ϵ is crowded in $[0,t]$. It is easy to see that

$$E(\chi^k_\epsilon(t)|\mathcal{K}_t = K) \leq KEw^\epsilon_t \tag{8.4}$$

$$E\mathcal{K}_t = \exp(\nu t) \text{ where } \nu = c^*N \tag{8.5}$$

$$E(\chi^k_\epsilon(t)) \leq \exp(\nu t)Ew^\epsilon_t \tag{8.6}$$

To estimate the difference between $X_\epsilon^k(s)$ and $Y_\epsilon^k(s)$ we observe that if $\chi_\epsilon^k(t) = \tau$ then the number of "independent jumps" in the ith coordinate of Y_ϵ^k that occur in $[0,t]$ has a Poisson distribution with mean $\epsilon^{-2}\tau$. Let $\Delta_Y^i(s)$ be the net effect of the independent jumps on coordinate i up to time s. Recalling that changes in the ith coordinate of Y_ϵ^k have mean 0 and variance ϵ^2, it follows that $E\Delta_Y^i(s) = 0$ and

$$E(\Delta_Y^i(s)^2) = E\chi_\epsilon^k(s) \tag{8.7}$$

Since $\Delta_Y^i(s)$ is a martingale, Kolmogorov's inequality implies

$$E\left(\max_{0\le s\le t} \Delta_Y^i(s)^2\right) \le 4E(\Delta_Y^i(t)^2) \tag{8.8}$$

Using Markov's inequality (i.e., if $X \ge 0$ then $P(X > x) \le EX^r/x^r$) then (8.8), (8.7), (8.6), and (8.3) (noting that the worst case is $d = 1$) gives

$$P\left(\max_{0\le s\le t} |\Delta_Y^i(s)| \ge \epsilon^{.3}\right) \le \epsilon^{-.6} E\left(\max_{0\le s\le t} \Delta_Y^i(s)^2\right) \le C\epsilon^{.4}t^{1/2}\exp(\nu t) \tag{8.9}$$

Here and in what follows C dentores a constant whose value is unimportant and that will change from line to line. The arguments leading to the last inequality also apply to $\Delta_X^i(t)$, the net effect of jumps in $[0,t]$ while X_ϵ^k is crowded, so

$$P\left(\max_{0\le s\le t} \|X_\epsilon^k(s) - Y_\epsilon^k(s)\|_\infty \ge 2\epsilon^{.3}\right) \le C\epsilon^{.4}t^{1/2}\exp(\nu t) \tag{8.10}$$

The estimate in (8.10) shows that the X_ϵ^k are close to independent random walks. To see that with high probability no collisions occur, we pick M large enough so that $\|x\|_1 \le M$ for all $x \in \mathcal{N}$ and repeat the derivation of (8.6) with ϵ replaced by $M\epsilon$ to conclude that the expected number of births from X_ϵ^k while there is some other X_ϵ^j in $X_\epsilon^k + \mathcal{N}$ is smaller than

$$C\epsilon t^{1/2}\exp(\nu t) \tag{8.11}$$

(8.5) and Markov's inequality imply that

$$P(\mathcal{K}_t > K) \le K^{-1}\exp(\nu t) \tag{8.12}$$

When $\mathcal{K}_t \le K$, (8.11) implies that the expected number of collisions is smaller than

$$KC\epsilon t^{1/2}\exp(\nu t) \tag{8.13}$$

Combining the last two results and setting $K = \epsilon^{-.5}$ shows that the probability of a collision is smaller than

$$C\epsilon^{.5}(1+t)^{1/2}\exp(\nu t) \tag{8.14}$$

Having shown that collisions are unlikely we no longer have to worry about the labels μ_m^ϵ that tell us the mother of the N particles created at time R_m^ϵ since this will be clear from the evolution of the dual. A more significant consequence of the results in this subsection is that dual processes for different sites are asymptotically independent. To argue this, we say the two duals *collide* if a particle in one dual gives birth when crowded by a particle in the other one. The arguments leading to (8.14) show that with high probability two duals do not collide, and (8.10) implies that the movements of all the particles can be coupled to independent random walks.

c. Convergence of $u_i^\epsilon(t,x)$. The next step is to show that as $\epsilon \to 0$ the branching random walk Y converges to a branching Brownian motion Z. To do this we use Skorokhod's trick to embed the ith component of the kth walk, $Y_s^{k,i}$ in a a Brownian motion $Z_s^{k,i}$. Using some standard estimates (see Durrett and Neuhauser for details) it follows that

$$(8.15)\qquad P\left(\max_{0\le s\le t} \|Y_\epsilon^k(s) - Z^k(s)\|_\infty > 4\epsilon^{.3} \text{ for some } k \le K\right) \le KC\epsilon^{.32}(1+t)$$

To compute the state of x at time t, we need not only the dual process $I_\epsilon^{x,t}(s)$, $s \le t$ but also the labels μ_n^ϵ and the uniform random variables U_n^x. However, the uniform random variables are independent of the dual process and, as we pointed out in a remark after (8.14), the μ_n^ϵ are only needed when a collision occurs.

As we will now explain, the results in the last paragraph make it easy to show that $u_a^\epsilon(t,x) \to u_a(t,x)$ as $\epsilon \to 0$. Here and in what follows we will use a and b to denote possible states of the sites to ease the burden on the middle of the alphabet. The first step is to describe $u_a(t,x)$. Let Z_s be a branching Brownian motion starting with a single particle at x and let $\mathcal{K}_t$ be the number of particles at time t. For $0 \le k < \mathcal{K}_t$, we let $\zeta_0(k)$ be independent and $= a$ with probability $\phi_a(Z_t^k)$. Once the ζ_0 are defined, we work up the space time set $\{Z_{t-s}^k\} \times \{s\}$. The values of $\zeta_s(k)$, the state of Z_{t-s}^k at time s, stay constant as long as only stirring occurs. When $N+1$ branches Z_{t-s}^i, $Z_{t-s}^{kN+1}, \ldots Z_{t-s}^{(k+1)N}$ come together (corresponding to a birth in the dual), we compute the flip rate at Z_{t-s}^i assuming it is in state $\zeta_s(i)$ and its neighbors are in states $\zeta_s(kN+j), 1 \le j \le N$. We generate an independent random variable uniform on $(0,1)$ to determine what (if any) flip should occur at Z_{t-s}^i. After we decide if we should change $\zeta_s(i)$, we can ignore $\zeta_s(kN+j)$ for $1 \le j \le N$. When we reach time t we will only be looking at the value at $\zeta_t(0)$. We call this value, the *result of the computation* and let $u_a(t,x) = P(\zeta_t(0) = a)$.

The description in the last paragraph is much like the one given earlier for the dual with one exception: the uniform random variables come from an auxiliary i.i.d. sequence instead of being read off the graphical representation. When there are no collisions in the dual, then the family structure of the influence set and the branching Brownian motion are the same. In this case if the inputs $\zeta_0(k)$ and the uniform random variables used are the same, the two computations have the same result. We have supposed that the initial functions $\phi_b(x)$ are continuous so (2.19) implies that as $\epsilon \to 0$,

$$\max_k |\phi_b(X_\epsilon^k(t)) - \phi_b(Z^k(t))| \to 0$$

where the maximum is taken over particles alive at time t. The last observation implies that we can with high probability arrange for all the inputs to be the same and it follows that $u_a^\epsilon(t,x) \to u_a(t,x)$. The last proof extends trivially to show that if $x_\epsilon \to x$ then $u_a^\epsilon(t,x_\epsilon) \to u_a(t,x)$. At the end of subsection b, we observed that the influence sets from different points are asymptotically independent. Combining that observation with the proofs in this subsection implies that if $x_\epsilon \to x$ then

$$(8.16) \qquad P(\xi_t^\epsilon(x_\epsilon + \epsilon y_j) = c_j, 0 \le j \le N) \to \prod_{j=0}^{N} u_{c_j}(t,x)$$

We are interested in statements that allow $x_\epsilon \to x$ since this form of the conclusion implies that convergence occurs uniformly on compact sets.

d. The limit satisfies the p.d.e. The first step is to write the limiting equation in integral form.

(8.17) **Lemma.** Suppose $f_a, 0 \le a < \kappa$ are continuous and $g_a, 0 \le a < \kappa$ are bounded and continuous. The following statements are equivalent:
(i) The functions $u_a(t,x)$ are a classical solution of

$$\frac{\partial u_a}{\partial t} = \Delta u_a - f_a(u) \qquad u_a(0,x) = g_a(x)$$

i.e., the indicated derivatives exist and are continuous.
(ii) The functions $u_a(t,x)$ are bounded and satisfy

$$u_a(t,x) = \int p_t(x,y) g_a(y)\,dy + \int_0^t ds \int p_s(x,y) f_a(u(t-s,y))\,dy$$

where $p_t(x,y)$ is the transition probability for Brownian motion run at rate 2.

Proof: (i) implies that $Z_s^a \equiv u_a(t-s,B_s) - \int_0^s f_a(u(t-r,B_r))dr$ is a bounded martingale, so $Z_0^a = EZ_t^a$ and (ii) follows from Fubini's theorem. To prove the converse, we begin by observing that if (ii) holds then $u_a(t,x)$ has the necessary derivatives and Z_s^a is a martingale, so (i) follows from Itô's formula. □

To get (ii) we will use the integration by parts formula. Let S_t^ϵ be the semigroup for the stirred particle system and T_t^ϵ be the semigroup for pure stirring. The integration by parts formula implies that for nice functions ψ we have

$$(8.18) \qquad S_t^\epsilon \psi(\xi) = T_t^\epsilon \psi(\xi) + \int_0^t ds\, S_{t-s}^\epsilon L T_s^\epsilon \psi(\xi)$$

where L is the generator for the particle system with no stirring. We use (8.18) with $\psi_{x,a}(\xi) = 1$ if $\xi(x) = a$ and 0 otherwise. Now for this choice of ψ

$$(8.19) \qquad T_s^\epsilon \psi_{x,a}(\xi) = \sum_y p_s^\epsilon(x,y)\psi_{y,a}(\xi)$$

where $p_s^\epsilon(x,y)$ is the transition probability of a random walk that jumps from y to z at rate $\epsilon^{-2}/2$ if $\|y-z\|_1 = \epsilon$. Now if $c_b(y,\xi) = h_b(\xi(y+\epsilon y_0),\dots,\xi(y+\epsilon y_N))$ then

$$L\psi_{y,a} = -\sum_b h_{b_0}(a,b_1,\dots,b_N)\psi_{y,a}\prod_{j=1}^N \psi_{y+\epsilon y_j,b_j} + \sum_b h_a(b_0,b_1,\dots,b_N)\psi_{y,b_0}\prod_{j=1}^N \psi_{y+\epsilon y_j,b_j} \tag{8.20}$$

where the sums are over $b_0,\dots,b_N \in \{0,1,\dots,\kappa-1\}$. Substituting (8.19) and (8.20) into (8.18) gives

$$\begin{aligned} P(\xi_t^\epsilon(x)=a) &= \sum_y p_t^\epsilon(x,y)g_a(y) \\ &+ \int_0^t ds \sum_y p_s^\epsilon(x,y) E\Bigg\{ -\sum_b h_{b_0}(a,b_1,\dots,b_N)\psi_{y,a}(\xi_{t-s}^\epsilon)\prod_{j=1}^N \psi_{y+\epsilon y_j,b_j}(\xi_{t-s}^\epsilon) \\ &\qquad + \sum_b h_a(b_0,b_1,\dots,b_N)\psi_{y,b_0}(\xi_{t-s}^\epsilon)\prod_{j=1}^N \psi_{y+\epsilon y_j,b_j}(\xi_{t-s}^\epsilon)\Bigg\} \end{aligned} \tag{8.21}$$

The local central limit theorem implies

$$\sum_y |\epsilon^d p_s(x,y) - p_s^\epsilon(x,y)| \to 0 \tag{8.22}$$

as $\epsilon \to 0$. As we observed at the end of subsection c,

$$E\psi_{y,c_0}(\xi_{t-s}^\epsilon)\prod_{j=1}^N \psi_{y+\epsilon y_j,c_j}(\xi_{t-s}^\epsilon) \to \prod_{j=0}^N u_{c_j}(t-s,y)$$

and this convergence occurs uniformly on compact sets. Using (8.21), (8.22), and the dominated convergence theorem, gives

$$\begin{aligned} u_a(t,x) &= \int p_t(x,y)g_a(y)\,dy \\ &+ \int_0^t ds \int dy\, p_s(x,y)\Bigg\{ -\sum_b h_{b_0}(a,b_1,\dots,b_N)u_a(t-s,y)\prod_{j=1}^N u_{b_j}(t-s,y) \\ &\qquad + \sum_b h_a(b_0,b_1,\dots,b_N)u_{b_0}(t-s,y)\prod_{j=1}^N u_{b_j}(t-s,y)\Bigg\} \end{aligned} \tag{8.23}$$

The term in braces is

$$-\sum_{b\neq a} < c_b(0,\xi)1_{\{\xi(0)=a\}} >_{u(t-s,y)} + < c_a(0,\xi) >_{u(t-s,y)} = f_a(u(t-s,y)) \tag{9.24}$$

Combining this with (8.17) gives the conclusion of Theorem 8.1.

9. Predator Prey Systems

In this section we will show that if you "know enough" about the limiting p.d.e. in Theorem 8.1 then you can prove results about the existence of stationary distributions for the system with fast stirring. For our approach, what you need to know about the p.d.e. is the following:

($\star$) There are constants $A_i < a_i < b_i < B_i$, L, and T so that if $u_i(0,x) \in (A_i, B_i)$ when $x \in [-L,L]^d$ then $u_i(x,T) \in (a_i, b_i)$ when $x \in [-3L, 3L]^d$.

Theorem 9.1. If ($\star$) holds then there is a nontrivial translation invariant stationary distribution for the process with fast stirring.

As the reader can probably guess, ($\star$) and Theorem 9.2 combine to produce a block event that turns one "pile of particles" into two and has high probability when ϵ is small and then the result follows from our comparison theorem. The details are somewhat technical so we refer the reader to Section 3 of Durrett and Neuhauser (1993) and turn to the problem of checking that ($\star$) holds in one particular example. For other applications of this technique see Durrett and Neuhauser (1993) or Durrett and Swindle (1993).

Example 9.1. Predator Prey Systems. The state at time t is $\xi_t^\epsilon : \epsilon \mathbf{Z}^d \to \{0,1,2\}$. We think of 0 as vacant, 1 and 2 as occupied by a fish and shark respectively. As usual, $n_i(x,\xi)$ is the number of neighbors of x (i.e., y with $\|y-x\|_1 = \epsilon$) that are in state i. The system changes states at the following rates:

$$\begin{array}{ll} c_1(x,\xi) = \beta_1 n_1(x,\xi)/2d & \text{if } \xi(x) = 0 \\ c_0(x,\xi) = \delta_1 & \text{if } \xi(x) = 1 \\ c_2(x,\xi) = \beta_2 n_2(x,\xi)/2d & \text{if } \xi(x) = 1 \\ c_0(x,\xi) = \delta_2 + \gamma n_2(x,\xi)/2d & \text{if } \xi(x) = 2 \end{array}$$

The first two rates say that fish repopulate vacant sites at a rate proportional to the number of fish at neighboring sites and die at rate δ_1. That is, in the absence of sharks, the fish are a contact process. The third rate says that sharks reproduce when they eat fish. This transition is a little strange from a biological point of view, but it has the desirable property that sharks will die out when the density of fish is too small. The last rate says that sharks die at rate δ_2 when they are isolated and the rate increases linearly with crowding. Finally, the sharks and fish swim around: for each pair of neighbors x and y stirring occurs at rate ϵ^{-2}, i.e., the values at x and y are exchanged. Applying Theorem 8.1 gives

Theorem 9.2. *Suppose that $\xi_0^\epsilon(x)$, $x \in \epsilon\mathbf{Z}^d$ are independent and $u_i^\epsilon(t,x) = P(\xi_t^\epsilon(x) = i)$ for $i = 1,2$. If $u_i^\epsilon(0,x) = \phi_i(x)$, which is continuous, then as $\epsilon \to 0$, $u_i^\epsilon(t,x) \to u_i(t,x)$ the bounded solution of*

$$\begin{aligned} \frac{\partial u_1}{\partial t} &= \Delta u_1 + \beta_1 u_1(1 - u_1 - u_2) - \beta_2 u_1 u_2 - \delta_1 u_1 \\ \frac{\partial u_2}{\partial t} &= \Delta u_2 + \beta_2 u_1 u_2 - u_2(\delta_2 + \gamma u_2) \end{aligned} \tag{9.1}$$

with $u_i(0,x) = \phi_i(x)$.

As in the two examples in Section 8, the reaction terms are computed by assuming that adjacent sites are independent. To get $\beta_1 u_1(1-u_1-u_2)$ for example we note that if x is vacant and neighbor y is occupied by a fish, an event of probability $(1-u_1-u_2)u_1$ when sites are independent, births from y to x occur at rate $\beta_1/2d$ and there are $2d$ such pairs.

When the initial functions $\phi_i(x)$ are constant, $u_i(t,x) = v_i(t)$ and the v_i's satisfy

$$(9.2) \qquad \begin{aligned} \frac{dv_1}{dt} &= v_1((\beta_1 - \delta_1) - \beta_1 v_1 - (\beta_1 + \beta_2)v_2) \\ \frac{dv_2}{dt} &= v_2(-\delta_2 + \beta_2 v_1 - \gamma v_2) \end{aligned}$$

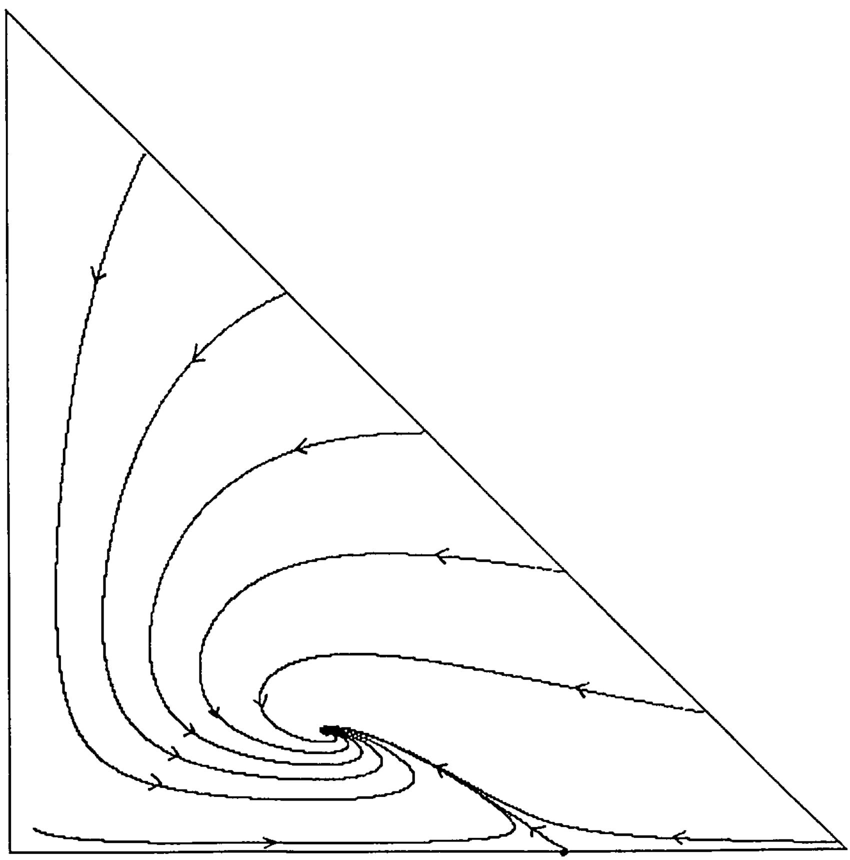


Figure 9.1

Here we have re-arranged the right hand side to show that the system is the standard predator–prey equations for species with limited growth. (See e.g., page 263 of Hirsch and Smale (1974).) Before we plunge into the details of analyzing (9.2), the reader should look at Figure 9.1, which gives some solutions of (9.2) with

$$\beta_1 = 3, \delta_1 = 1, \quad \beta_2 = 3, \delta_2 = 1, \quad \gamma = 1$$

In this case, as we will prove later, there is a fixed point at $(8/21, 3/21)$ that is globally attracting.

The first step in understanding (9.1) is to look at (9.2) and ask: "What are the fixed points, i.e., solutions of the form $v_i(t) = \rho_i$?" It is easy to solve for the ρ_i. There is always the trivial solution $\rho_1 = \rho_2 = 0$. In the absence of sharks the fish are a contact process. So if $\beta_1 > \delta_1$ there is a solution $\rho_1 = (\beta_1 - \delta_1)/\beta_1$, $\rho_2 = 0$. If we impose the stronger condition

$$(9.3) \qquad (\beta_1 - \delta_1)/\beta_1 > \delta_2/\beta_2$$

there is exactly one solution with $\rho_2 > 0$:

$$\rho_1 = \frac{(\beta_1 - \delta_1)\gamma + \delta_2(\beta_1 + \beta_2)}{\beta_1\gamma + \beta_2(\beta_1 + \beta_2)} \qquad \rho_2 = \frac{(\beta_1 - \delta_1)\beta_2 - \delta_2\beta_1}{\beta_1\gamma + (\beta_1 + \beta_2)\beta_2}$$

The condition $\beta_1 > \delta_1$ is an obvious necessary condition for the fish to survive in the absence of sharks. The condition (9.3) is not so intuitive but turns out to be sufficient for the existence of nontrivial stationary distributions for small ϵ.

Theorem 9.3. *Suppose that $(\beta_1 - \delta_1)/\beta_1 > \delta_2/\beta_2$ holds. If ϵ is small there is a nontrivial translation invariant stationary distribution in which the density of sites of type i is close to ρ_i.*

In view of Theorem 9.1 it suffices to prove $(\star)$, which is a consequence of the following convergence theorem.

Theorem 9.4. *Suppose that $(\beta_1 - \delta_1)/\beta_1 > \delta_2/\beta_2$ holds and the u_i solve (9.1) for continuous nonnegative $\phi_i(x)$ with $\phi_1(x) + \phi_2(x) \le 1$ and $\phi_i(x_i) > 0$ for some x_i. Then there is a $\sigma > 0$ so that as $t \to \infty$,*

$$\sup_{\|x\| \le \sigma t} |u_i(t,x) - \rho_i| \to 0.$$

PROOF: The proof is based on a simple idea due to Redheffer, Redlinger, and Walter (1988): the existence of a convex strict Lyapunov function for the dynamical system (10.2) plus two technical conditions we will identify in the proof, give a convergence theorem for the reaction diffusion equation. In this case the desired function is

$$H(v_1, v_2) = \beta_2(v_1 - \rho_1 \log v_1) + (\beta_1 + \beta_2)(v_2 - \rho_2 \log v_2)$$

Being the sum of four convex functions, H is clearly convex. The next step is to check that it is a strict Lyapunov function: if (v_1, v_2) is a solution of the dynamical system that does not start at the fixed point then $\partial H(v_1, v_2)/\partial t < 0$. Differentiating gives

$$\frac{\partial H}{\partial v_1} = \beta_2 \left(1 - \frac{\rho_1}{v_1}\right) \qquad \frac{\partial H}{\partial v_2} = (\beta_1 + \beta_2)\left(1 - \frac{\rho_2}{v_2}\right)$$

So using the chain rule and (9.2)

$$\begin{aligned} H_t \equiv \frac{\partial H(v_1, v_2)}{\partial t} &= \beta_2(v_1 - \rho_1)\{(\beta_1 - \delta_1) - \beta_1 v_1 - (\beta_1 + \beta_2)v_2\} \\ &\quad + (\beta_1 + \beta_2)(v_2 - \rho_2)\{-\delta_2 + \beta_2 v_1 - \gamma v_2\} \end{aligned}$$

Using the next two identities to subtract 0 from each term in braces

$$\begin{aligned} 0 &= (\beta_1 - \delta_1) - \beta_1\rho_1 - (\beta_1 + \beta_2)\rho_2 \\ 0 &= -\delta_2 + \beta_2\rho_1 - \gamma\rho_2 \end{aligned}$$

gives

$$\begin{aligned} H_t &= \beta_2(v_1 - \rho_1)\{-\beta_1(v_1 - \rho_1) - (\beta_1 + \beta_2)(v_2 - \rho_2)\} \\ &\quad + (\beta_1 + \beta_2)(v_2 - \rho_2)\{\beta_2(v_1 - \rho_1) - \gamma(v_2 - \rho_2)\} \\ &= -\beta_1\beta_2(v_1 - \rho_1)^2 - \gamma(\beta_1 + \beta_2)(v_2 - \rho_2)^2 \le 0 \end{aligned} \tag{9.4}$$

with strict inequality for $(v_1, v_2) \neq (\rho_1, \rho_2)$. The importance of the last conclusion is that $H(v_1(t), v_2(t))$ is strictly decreasing in t and hence all trajectories that begin in $(0, \infty)^2$ must end at the minimum of H, (ρ_1, ρ_2). For later purposes we would like to note that the level curves $H_t = -r$ are concentric ellipses.

The above computations that show H is a Lyapunov function obviously depend on the special form of (9.2). To prepare for other applications at the end of this section, we would like the reader to check that in what follows only equations (9.6) and (9.10) depend on the special form of H.

Since composing the Lyapunov function with solutions of the dynamical system shows that they converge to the fixed point, it is natural to look at $h(t, x) = H(u_1(t, x), u_2(t, x)) - H(\rho_1, \rho_2) \ge 0$ when u is a solution of (9.1). (Here we have subtracted the value of H at its minimum to make the minimum value 0.) To show the generality of this computation and to simplify notation we will write (9.1) as

$$\frac{\partial u_i}{\partial t} = \Delta u_i + f_i(u).$$

Differentiating and using the previous equation gives

$$\begin{aligned} \frac{\partial h}{\partial t} &= \sum_i \frac{\partial H}{\partial u_i}\frac{\partial u_i}{\partial t} = \sum_i \frac{\partial H}{\partial u_i} \cdot (\Delta u_i + f_i(u)) \\ \frac{\partial^2 h}{\partial x_m^2} &= \sum_i \frac{\partial H}{\partial u_i}\frac{\partial^2 u_i}{\partial x_m^2} + \sum_{i,j} \frac{\partial^2 H}{\partial u_i \partial u_j}\frac{\partial u_i}{\partial x_m}\frac{\partial u_j}{\partial x_m} \end{aligned}$$

Here and in what follows the indices i and j are summed from 1 to 2. Summing the second equation from $m = 1$ to d gives

$$\Delta h = \sum_i \frac{\partial H}{\partial u_i} \Delta u_i + \sum_{m,i,j} \frac{\partial^2 H}{\partial u_i \partial u_j} \frac{\partial u_i}{\partial x_m} \frac{\partial u_j}{\partial x_m}$$

so using $H_t = \sum_i \frac{\partial H}{\partial u_i} f_i(u)$ gives

$$\frac{\partial h}{\partial t} = \Delta h + H_t - \sum_{m,i,j} \frac{\partial^2 H}{\partial u_i \partial u_j} \frac{\partial u_i}{\partial x_m} \frac{\partial u_j}{\partial x_m}$$

Since H is convex the last term (including the minus sign) is nonpositive and we have

$$\frac{\partial h}{\partial t} \le \Delta h + H_t \tag{9.5}$$

To prove Theorem 9.4, we will use (9.5) to conclude

$$\sup_{\|x\| \le \epsilon t} h(t,x) \to 0 \tag{9.6}$$

If we were on a bounded set with Neumann boundary conditions this would be easy, since in this case $\inf u_i(t,x) > 0$ at positive times and thus $h(t,x)$ is bounded. If we let x_t be a place where $m_t = \max_x h(t,x)$ is attained then $\Delta h(t,x_t) \le 0$ so

$$\frac{dm_t}{dt} \le H_t \le \sup\{H_t(v) : H(v) = m_t\} < 0$$

and an argument like the one in the proof of Theorem 5.1 shows (9.6).

To prove (9.6) on $\mathbf{R}^d$ we have to deal with the fact that $h(t,x)$ may be unbounded. To do this we first get bounds on how fast the H_t will push h to 0 and then get *a priori* bounds on $h(t,x)$ inside $\|x\| \le at$ that will allow us to drive h to 0. To get upper bounds on H_t (recall it is ≤ 0), we let

$$g(h) = \inf\{-H_t(v_1,v_2) : H(v_1,v_2) \ge h\}$$

We have defined g this way to make it clear that $h \to g(h)$ is increasing. To determine the behavior as $h \to 0$ we observe that at (ρ_1,ρ_2) $\partial H/\partial v_i = 0$ and

$$\frac{\partial^2 H}{\partial v_1^2} = \frac{\beta_2 \rho_1}{v_1^2} \qquad \frac{\partial^2 H}{\partial v_2^2} = \frac{(\beta_1+\beta_2)\rho_2}{v_2^2} \qquad \frac{\partial^2 H}{\partial v_1 \partial v_2} = 0$$

So near (ρ_1,ρ_2)

$$H(v_1,v_2) - H(\rho_1,\rho_2) \approx \frac{\beta_2}{\rho_1}(v_1-\rho_1)^2 + \frac{(\beta_1+\beta_2)}{\rho_2}(v_2-\rho_2)^2$$

and it follows from (9.4) that $g(h) \sim Bh$. Since $g(h)$ is increasing we have

(9.7) $$g(h) \geq \alpha h/(1+h) \quad \text{for some } \alpha > 0$$

The next step in bounding $h(t,x)$ is to examine the behavior of

$$w' = \frac{-\alpha w}{1+w} \qquad w(0) = W$$

Since $w(t) \geq W - \alpha t$ the time to reach $\eta > 0$ is at least $(W-\eta)/\alpha$. To see this estimate is fairly sharp observe that while $w(t) \geq W^{1/2}-1$ we have $w' \leq -\alpha(1-W^{-1/2})$ so $w(t)$ reaches $W^{1/2}-1$ at time $\leq W/(\alpha(1-W^{-1/2}))$. When $w(t) \leq W^{1/2}-1$ we have $w' \leq -\alpha w/W^{1/2}$ so the time to go from $W^{1/2}-1$ to η is at most $W^{1/2}\alpha^{-1}(\log(W^{1/2}) - \log\eta)$. Adding the two estimates we see that

(9.8) *For $W \geq 4$ and $\eta < 1$ the time to reach η is smaller than*

$$2\alpha^{-1}W + CW^{1/2}(\log W - \log\eta).$$

To get *a priori* bounds on h note that our hypotheses imply that $u_i(1,x)$ is positive and continuous so there are constants μ_i so that $u_i(1,x) \geq \mu_i$ for all x with $\|x\|_2 \leq 1$, and we can without loss of generality assume that the last conclusion holds at time 0.

(9.9) **Lemma.** *There is a constant K so that if $\|x\|_2 \leq at$ and $t \geq 1$ then $h(t,x) \leq Kt$.*

Proof: We have supposed that $\phi_1(x)+\phi_2(x) \leq 1$ so the probabilistic interpretation implies $u_1(t,x) + u_2(t,x) \leq 1$ for all t and x and it follows that

$$\frac{\partial u_1}{\partial t} \geq \Delta u_1 - (\beta_1 + \beta_2)u_1$$
$$\frac{\partial u_2}{\partial t} \geq \Delta u_2 - (\delta_2 + \gamma)u_2$$

To see these inequalities it is convenient to write the right–hand side in the form in (9.2). (Recall that our main assumption (9.3) implies $\beta_1 > \delta_1$.) Let $c_1 = (\beta_1+\beta_2)$ and $c_2 = \delta_2+\gamma$. Recalling that solutions of

$$\frac{\partial u}{\partial t} = \Delta v - cv \qquad u(0,x) = \phi(x)$$

are given by

$$u(t,x) = e^{-ct}\int (4\pi t)^{-d/2} e^{-\|x-y\|_2^2/4t}\phi(y)\,dy$$

and using the maximum principle (see (9.11) at the end of this section) we have that when $\|x\|_2 \leq at$

$$\begin{aligned} u_i(t,x) &\geq e^{-c_i t}\mu_i \int_{\|y\|\leq 1} (4\pi t)^{-d/2} e^{-(\|x\|_2+1)^2/4t} dy \\ &\geq C_d\mu_i(4\pi t)^{-d/2}\exp(-(c_i + (a^2/4))t - a/2 - (1/4t)) \end{aligned}$$

where C_d is the volume of $\{y : \|y\|_2 \leq 1\}$. Combining the last expression with the fact that

$$h(x) \leq C(1 - \log(\min_i x_i)) \tag{9.10}$$

completes the proof of (9.9) □

Let $a > 0$ be chosen so that $3\alpha^{-1}aK < (1-a)$, i.e., so that if $\omega(t)$ solves

$$\omega' = -\alpha\omega/(1+\omega) \quad \text{with} \quad \omega(0) = Kat$$

then for any $\eta > 0$ when $t > T_\eta$, $\omega((1-a)t) < \eta$. We will prove Theorem 9.4 with $\sigma = a/2$. Let $D_r = \{y : \|y\|_2 < r\}$ and define $h_1^t(t,x)$ to be the solution of

$$\begin{aligned} \frac{\partial h}{\partial t} &= \Delta h - \alpha h/(1+h) \quad \text{in } \mathcal{D}_t \equiv [at,t] \times D_{at} \\ h(s,x) &= Ks \quad \text{if } s = at, \text{ or } x \in \partial D_{at} \end{aligned}$$

Since $h(s,x) \leq h_1^t(s,x)$ when $s = at$ or $x \in \partial D_{at}$, and $g(h) \geq \alpha h/(1+h)$ it follows from the maximum principle that $h(s,x) \leq h_1^t(s,x)$ for $(s,x) \in \mathcal{D}_t$.

To bound $h_1^t(t,x)$ we will use $h_2^t(s,x) = \omega(s - at)$. Another use of the maximum principle shows $h_1^t(s,x) \geq h_2^t(s,x)$ in $\mathcal{D}_t$. The last inequality is the opposite of the one we want but we will turn it around by showing that the difference is small when $\|x\|_2 \leq at/2$. Intuitively the difference is only due to paths $(t-s, B_s)$ that escape from the space time cylinder $[at,t] \times D_{at}$ on the side. Here B_s is a Brownian motion run at twice its usual speed. When the starting point $\|B_0\|_2 \leq at/2$ this event has exponentially small probability and brings a "reward" $\leq Kt$ so the difference $h_1^t - h_2^t$ goes to 0 exponentially fast as $t \to \infty$.

To begin to turn our intuition into a proof, we let $\bar{g}(x) = \alpha x/(1+x)$ and observe that Itô's formula implies that if $\tau = \inf\{s : B_s \notin D_{at}\}$ then

$$h_i^t(t - (s \wedge \tau), B_{s\wedge\tau}) - \int_0^{s\wedge\tau} \bar{g}(h_i^t(t-r, B_r))\, dr \qquad s \leq (1-a)t$$

is a bounded martingale. Using the martingale property at time $s = (1-a)t$ gives

$$\begin{aligned} h_i^t(t,x) = E_x &\left(h_i^t(at, B_{(1-a)t}) - \int_0^{(1-a)t} \bar{g}(h_i^t(t-r, B_r))\, dr \,; \tau > (1-a)t \right) \\ &+ E_x \left(h_i^t(t-\tau, B_\tau) - \int_0^{\tau} \bar{g}(h_i^t(t-r, B_r))\, dr \,; \tau \leq (1-a)t \right) \end{aligned}$$

Since $h_1^t(at,x) = h_2^t(at,x)$, $0 \leq h_2^t(t-r, B_r) \leq h_1^t(t-r, B_r)$ when $\tau \geq r$, and $\bar{g}$ is increasing, it follows that on $\{\tau > (1-a)t\}$ we have

$$\begin{aligned} &h_1^t(at, B_{(1-a)t}) - \int_0^{(1-a)t} \bar{g}(h_1^t(t-r, B_r))\, dr \\ &\qquad \leq h_2^t(at, B_{(1-a)t}) - \int_0^{(1-a)t} \bar{g}(h_2^t(t-r, B_r))\, dr \end{aligned}$$

Subtracting the two expressions for $h_i^t(t,x)$ and recalling $h_i^t \geq 0$ and $\bar{g} \geq 0$ gives

$$h_1^t(t,x) - h_2^t(t,x) \leq 0 + E_x\left(h_1^t(t-\tau, B_\tau) + \int_0^\tau \bar{g}(h_2^t(t-r, B_r))\,dr; \tau \leq (1-a)t\right)$$
$$\leq (Kt + \alpha t)P_x(\tau \leq (1-a)t)$$

since $h_1^t(s,x) = Ks$ when $x \in \partial D_{at}$ and $0 \leq \bar{g} \leq \alpha$. Standard large deviations estimates for Brownian motion imply that for $\|x\|_2 \leq at/2$, $P_x(\tau < (1-a)t) \leq C\exp(-\delta t)$, so

$$\sup_{\|x\| \leq at/2} |h_1^t(t,x) - h_2^t(t,x)| \to 0 \quad \text{as } t \to \infty$$

Since $h_1^t(t,x) \geq h(t,x)$ for $x \in D_{at}$ and $h_2^t(t,x) = \omega((1-a)t) < \eta$ for $t > T_\eta$, Theorem 9.4 follows. □

For completeness we give a proof of

(9.11) **Maximum Principle.** *Suppose $f_1(h) \geq f_2(h)$ and the h_i solve*

$$\frac{\partial h_i}{\partial t} = \Delta h_i - f_i(h_i) \quad \text{in } \mathcal{D}_t$$

with $h_1(s,x) \leq h_2(s,x)$ if $s = at$, or $x \in \partial D_{at}$ then $h_1(s,x) \leq h_2(s,x)$ in $\mathcal{D}_t$.

PROOF: This is easier to prove than to find in the library. Suppose first that $f_1(h) > f_2(h)$ and $h_1(s,x) < h_2(s,x)$ if $s = at$, or $x \in \partial D_{at}$. Let s_0 be the smallest value of s for which there is an x with $h_1(s,x) \geq h_2(s,x)$. Continuity of the h_i implies that we can find an x_0 so that $h_1(s_0,x_0) = h_2(s_0,x_0)$. The strict inequality between the h_i on the boundary implies $x_0 \in D_{at}$ and $s_0 > 0$. The definition of s_0 implies that $h_1(s_0,x) \leq h_2(s_0,x)$ for all x. Since $h_1(s_0,x_0) = h_2(s_0,x_0)$, we must have $\nabla h_1(s_0,x_0) = \nabla h_2(s_0,x_0)$ and $\Delta h_1(s_0,x_0) \leq \Delta h_2(s_0,x_0)$. Using the last fact and $f_1(h) > f_2(h)$ it follows that at (s_0,x_0)

$$\frac{\partial h_1}{\partial t} = \Delta h_1 - f_1(h_1) < \Delta h_2 - f_2(h_2) = \frac{\partial h_2}{\partial t}$$

However this implies that $h_1(s_0 - \epsilon, x) > h_2(s_0 - \epsilon, x_0)$ for small ϵ contradicting the definition of s_0, so we must have $h_1(s,x) < h_2(s,x)$ for all $(s,x) \in \mathcal{D}_{at}$. To prove the result in (2.4) now let $f_0(h) = f_1(h) + \epsilon$ and change the boundary values to $h_1(s,x) - \epsilon$. The new solution $h_0^\epsilon(s,x) < h_2(s,x)$ and converges pointwise to $h_1(s,x)$ as $\epsilon \to 0$. □

The main reason for interest in Theorem 9.4 is that it applies to systems of equations. However, as the next two examples suggest we also get interesting information when we apply it to a single equation.

Example 9.2. The basic contact process. If we let $\beta = \lambda N$ where N is the number of neighbors and write u for u_1 then the equation in Example 8.1 can be written as

(9.12) $$\frac{\partial u}{\partial t} = \Delta u - u + \beta(1-u)u \qquad u(0,x) = \phi(x)$$

To find a Lyapunov function we let $\rho = (\beta - 1)/\beta$ and write the dyanmical system as

$$\frac{dv}{dt} = v(-1 + \beta(1 - v)) = \beta v(\rho - v)$$

Taking $H(v) = v - \rho \log v$ and noticing $h'(v) = 1 - (\rho/v)$ we have

$$\frac{dH(v(t))}{dt} = -\beta(v - \rho)^2$$

Clearly H satisfies (9.10). Since $H'(\rho) = 0$ and $H''(v) = \rho/v^2$, repeating the proof of (9.7) shows it is satisfied. Since H is convex we get a convergence result like Theorem 9.4

Theorem 9.5. *Suppose that $\beta > 1$ u solves (9.12) for continuous $0 \le \phi(x) \le 1$ with $\phi(x_0) > 0$ for some x_0. Then there is a $\sigma > 0$ so that as $t \to \infty$,*

$$\sup_{\|x\| \le \sigma t} \left| u(t,x) - \frac{\beta - 1}{\beta} \right| \to 0.$$

Much better convergence results than this are known for this equation (see Aronson and Weinberger (1978) for more general results and Bramson (1983) for more detailed information), but the last result shows that $(\star)$ holds and we have

Theorem 9.6 *Suppose $\beta > 1$. If ϵ is small then the contact process with strring at rate ϵ^{-2} has a translation invariant stationary distribution in which the density of 1's is close to $(\beta - 1)/\beta$.*

Example 9.3. The threshold voter model. In this case if N is the number of neighbors and we write u for u_1 then the limiting equation in Example 8.2 is

$$\text{(9.13)} \qquad \frac{\partial u}{\partial t} = \Delta u - u(1 - u^N) + (1 - u)(1 - (1 - u)^N) \qquad u(0, x) = \phi(x)$$

When $N = 1$ the last two terms on the right hand side cancel so we will suppose that $N \ge 2$. For our Lyapunov function we take $H(v) = -\log v - \log(1 - v)$, which has

$$H'(v) = -\frac{1}{v} + \frac{1}{1 - v} = \frac{2v - 1}{v(1 - v)}$$

$$\begin{aligned} \frac{dH(v(t))}{dt} &= \frac{2v - 1}{v(1 - v)} \left\{ -v(1 - v^N) + (1 - v)(1 - (1 - v)^N \right\} \\ &= (2v - 1) \left\{ -(1 + v + \cdots + v^{N-1}) + (1 + (1 - v) + \cdots + (1 - v)^{N-1} \right\} \\ &= -(2v - 1)^2 \left\{ 1 + \sum_{j=2}^{N-1} \frac{(1 - v)^j - v^j}{1 - 2v} \right\} \end{aligned}$$

where the sum is 0 if $N = 2$. Since $(1-v)^j - v^j$ and $1-2v$ are both positive on $v < 1/2$ and negative on $v > 1/2$ their quotient is always positive. To compute the value at $v = 1/2$ we note that L'Hopital's rule implies that

$$\lim_{v \to 1/2} \frac{(1-v)^j - v^j}{1-2v} = \lim_{v \to 1/2} \frac{-j(1-v)^{j-1} - jv^{j-1}}{-2v} = j2^{-(j-1)}$$

so the term in braces is bounded away from 0 and ∞.

Since $H'(1/2) = 0$ and $H''(1/2) > 0$ it is easy to see as before that (9.7) holds. The other condition (9.10) does not hold as stated since $H(1) = \infty$. However it is easy to see that under suitable assumptions (9.9) holds and we have

Theorem 9.7. *Suppose that $N \geq 2$ u solves (9.13) for continuous $0 \leq \phi(x) \leq 1$ with $\phi(x_0) > 0$ for some x_0 and $\phi(x_1) < 1$ for some x_1. Then there is a $\sigma > 0$ so that as $t \to \infty$,*

$$\sup_{\|x\| \leq \sigma t} |u(t,x) - 1/2| \to 0.$$

Again better convergence results than this are known for this equation (see Aronson and Weinberger (1978) and Fife and McLeod (1977)) but the last result shows that $(\star)$ holds and we have

Theorem 9.8. *Suppose $N \geq 2$. If ϵ is small then the contact process with strring at rate ϵ^{-2} has a translation invariant stationary distribution in which the density of 1's is equal to 1/2.*

We get "equal to 1/2" rather than just "close to 1/2" by starting from product measure with density 1/2 and using the symmetry of the dynamics under interchange of 0's and 1's. Comparing this with Theorems 5.1 and 5.3, the only surprise is that in the nearest neighbor case there is a stationary distribution with fast stirring. We conjecture that the presence of stirring at any positive rate, there is a nontrivial stationary distribution. In support of this conjecture, Figure 9.2 shows a simulation of the nearest neighbor case with stirring rate = 3.

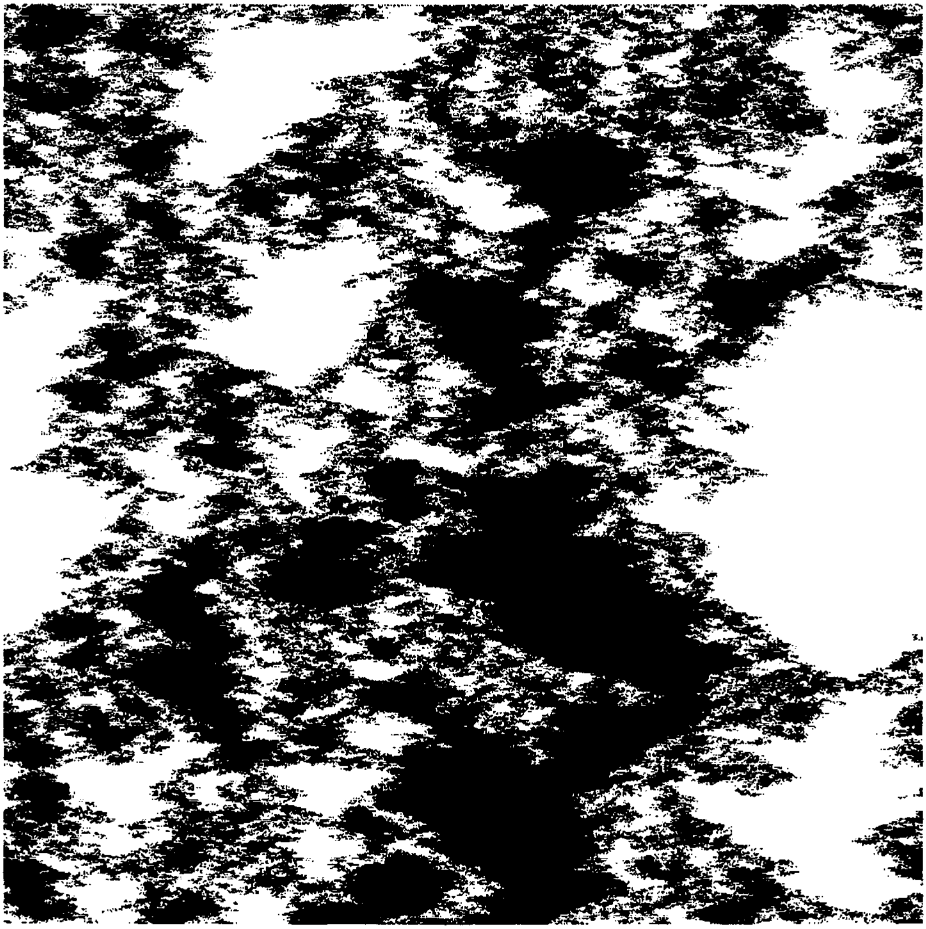


Figure 9.2. Threshold voter model, $d = 1$, $\mathcal{N} = \{-1, 1\}$, with stirring at rate 3

Appendix. Proofs of the Comparison Results

In this section we will prove Theorems 4.1, 4.2, and 4.3. The proofs are not beautiful but by now the reader has hopefully been convinced that they are useful. We begin by recalling the set-up and repeating some definitions that were more fully explained in Section 4. Let

$$\mathcal{L}_0 = \{(x,n) \in \mathbf{Z}^2 : x+n \text{ is even, } n \geq 0\}$$

and make $\mathcal{L}_0$ into a graph by drawing oriented edges from (x,n) to $(x+1,n+1)$ and from (x,n) to $(x-1,n+1)$. Given random variables $\omega(x,n)$ that indicate whether the sites are open (1) or closed (0), we say that (y,n) can be reached from (x,m) and write $(x,m) \to (y,n)$ if there is a sequence of points $x = x_m, \ldots, x_n = y$ so that $|x_k - x_{k-1}| = 1$ for $m < k \leq n$ and $\omega(x_k,k) = 1$ for $m \leq k \leq n$. We say that the $\omega(x,n)$ are "*M dependent with density at least* $1-\gamma$" if whenever (x_i,n_i), $1 \leq i \leq I$ is a sequence with $\|(x_i,m_i)-(x_j,m_j)\|_\infty > M$ if $i \neq j$ then

$$(A.1) \qquad P(\omega(x_i,n_i) = 0 \text{ for } 1 \leq i \leq I) \leq \gamma^I$$

Let $\mathcal{C}_0 = \{(y,n) : (0,0) \to (y,n)\}$ be the set of all points in space-time that can be reached by a path from $(0,0)$. $\mathcal{C}_0$ is called the *cluster containing the origin*. When the cluster is infinite, i.e., $\{|\mathcal{C}_0| = \infty\}$ we say that *percolation occurs*. Our first result shows that if the density of open sites is high enough then percolation occurs.

Theorem A.1. If $\theta \leq 6^{-4(2M+1)^2}$ then $P(|\mathcal{C}_0| < \infty) \leq 55\,\theta^{1/(2M+1)^2} \leq 1/20$.

PROOF: The proof is by the contour method. Even though the argument is messy to write down, the idea is simple: if $|\mathcal{C}_0| < \infty$ then there is a "contour" of closed sites that stops the percolation from occuring. As we will show, the probability of a specific contour of length n is $\leq g(\theta)^n$ where $g(\theta) \to 0$ as $\theta \to 0$ and the number of contours of length n is $\leq 3^n$ so by summing a geometric series we see that the existence of a contour is unlikely if θ is small.

Most of the work goes into defining the contour. Before starting on this we have to discard a trivial case: if $(0,0)$ is closed, an event with probability $\leq \gamma$, then $\mathcal{C}_0 = \emptyset$. For the rest of the proof we will concentrate on the case in which $(0,0)$ is open and hence $(0,0) \in \mathcal{C}_0$. Let $D = \{z \in \mathbf{R}^2 : \|z\|_1 \leq 1\}$, where D is for diamond. To turn the cluster $\mathcal{C}_0$ into a solid blob, we look at

$$\mathcal{D}_0 = \cup_{(m,n)\in\mathcal{C}_0} ((m,n)+D)$$

where $(m,n)+D = \{(m,n)+z : z \in D\}$ is the set D translated by (m,n). When $(0,0) \in \mathcal{C}_0$, the lowest point in $\mathcal{D}_0$ is $(0,-1)$. If $|\mathcal{C}_0| < \infty$, then the open set

$$G = \{\mathbf{R} \times (-1,\infty)\} - \mathcal{D}_0$$

has exactly one unbounded component U. We call $\Gamma = \partial U \cap \mathcal{D}_0$ the *contour* associated with $\mathcal{C}_0$, and orient it so that the segment $(0,-1) \to (1,0)$, which is always present, is oriented in the direction indicated. For an example see Figure A.1.

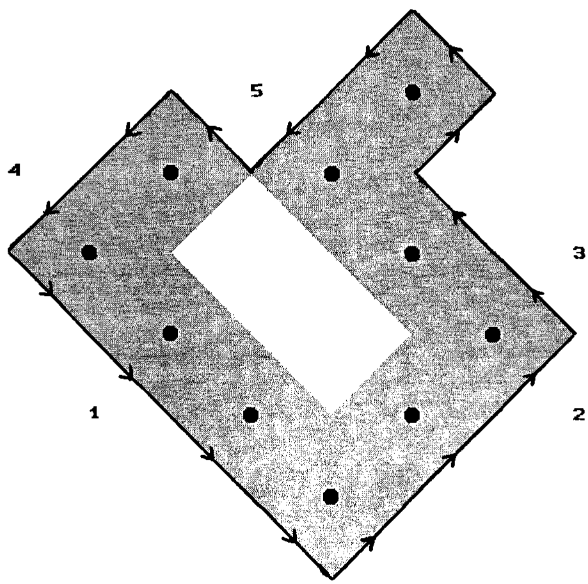


Figure A.1

The contour is made up of segments that are translates of the four sides of D

type	1	2	3	4
translate of	$(-1,0) \to (0,-1)$	$(0,-1) \to (1,0)$	$(1,0) \to (0,1)$	$(0,1) \to (-1,0)$

As we walk along the contour in the direction of the orientation, our left hand is always touching $\mathcal{D}_0$ and our right is always touching U. If we stand at the midpoint of one of the segments that make up Γ then the site in

$$\mathcal{L} = \{(m,n) \in \mathbf{Z}^2 : m+n \text{ is even}\}$$

closest to our right hand is called the *site associated with the segment.* A glance at Figure A.1 reveals that the sites associated with segments of types 3 and 4 must be closed but those associated with types 1 and 2 may be open or closed. Let n_i be the number of segments of type i in the contour. The segments of types 1 and 2 increase the x coordinate by 1, while those of types 3 and 4 decrease the x coordinate by 1. The contour ends where it begins so $n_3 + n_4 = n_1 + n_2$ and hence if the contour is composed of n segments we must have $n_3 + n_4 = n/2$. Now a closed site may be associated with one type 3 and one type 4 segement (see 5 on Figure A.1) but cannot be associated with more than one segment of each type, so if there are n segments in the contour there must be at least $n/4$ closed sites along it.

To count the number of contours of length n we note that the first segment is always $(0,-1) \to (1,0)$ and after that there are at most 3 choices at each stage (since we cannot retrace the step just made), so there are at most 3^{n-1} contours of length n. Suppose for the moment that the states of the sites are independent and open with probability $1-\gamma$. Noting that the length of the contour is ≥ 4, it follows that

$$P(0 < |\mathcal{C}_0| < \infty) \leq \sum_{n=4}^{\infty} 3^{n-1}\gamma^{n/4} = \frac{1}{3} \cdot \frac{(3\gamma^{1/4})^4}{1-3\gamma^{1/4}} = \frac{27\gamma}{1-3\gamma^{1/4}}$$

which is $< 1-\gamma$ if γ is small enough. (Recall $P(\mathcal{C}_0 = \emptyset) \leq \gamma$.) To extend the last result to the M dependent case, note that we can find a subset of the closed sites along the contour of size at least $n/4(2M+1)^2$ so that for each $z \neq w$ in this set $\|z-w\|_\infty > M$. (Pick any closed site to start, then throw out the $\leq (2M+1)^2 - 1$ closed sites in our set that are too close to the first one, pick another site, throw out the closed sites too close to it ...) Using (4.1) and noting our assumption on γ implies $3\gamma^{1/4(2M+1)^2} \leq 1/2$ we have

$$\begin{aligned} P(0 < |\mathcal{C}_0| < \infty) &\leq \sum_{n=4}^{\infty} 3^{n-1}\, \gamma^{n/4(2M+1)^2} \\ &= \frac{1}{3} \cdot \frac{(3\gamma^{1/4(2M+1)^2})^4}{1-3\gamma^{1/4(2M+1)^2}} \leq 54\,\gamma^{1/(2M+1)^2} \end{aligned}$$

Recalling now that $P(\mathcal{C}_0 = \emptyset) \leq \gamma \leq \gamma^{1/(2M+1)^2}$, we have proved Theorem 4.1. □

From the last proof it follows immediately that if we let $|\Gamma|$ denote the number of segments in the contour and assume $\gamma \leq 6^{-4(2M+1)^2}$ then

$$(A.2) \qquad P(L \leq |\Gamma| < \infty) \leq \sum_{n=L}^{\infty} 3^{n-1}\gamma^{n/4(2M+1)^2} = \frac{1}{3} \cdot \frac{(3\gamma^{1/4(2M+1)^2})^L}{1-3\gamma^{1/4(2M+1)^2}} \leq 2^{-L}$$

In order to prove the existence of stationary distributions we need results about M dependent oriented percolation starting from the initial configuration W_0^p in which the events $\{x \in W_0^p\}$, $x \in 2\mathbf{Z}$ are independent and have probability p. Let

$$W_n^p = \{y : (x,0) \to (y,n) \text{ for some } x \in W_0^p\}$$

Theorem A.2. *If $p > 0$ and $\gamma \leq 6^{-4(2M+1)^2}$ then*

$$\liminf_{n\to\infty} P(0 \in W_{2n}^p) \geq 1 - 55\,\theta^{1/(2M+1)^2} \geq 19/20$$

Proof: The first step is to look backwards in time to reduce the new problem to the old one solved in (A.2). This is the discrete time version of the duality considered in Section 3. To have the dual process defined for all time, it is convenient to introduce independent random variables $\omega(x,n)$ for $n < 0$ that have $P(\omega(x,n) = 1) = 1-\gamma$ and look at the percolation process on $\mathcal{L} = \{(x,n) \in \mathbf{Z}^2 : x + n \text{ is even}\}$. Later in the proof we will want to use the fact that $\gamma > 0$, so you should observe that the desired conclusion is trivial when $\gamma = 0$, i.e., all sites are open.

We say that (x,m) can be reached from (y,n) by a dual path (and write $(y,n) \to_* (x,m)$) if there is a sequence of points $x = x_m, \ldots, x_n = y$ so that $|x_k - x_{k-1}| = 1$ for $m < k \le n$ and $\omega(x_k,k) = 1$ for $m \le k \le n$. It should be clear from the definition that $(x,m) \to (y,n)$ if and only if $(y,n) \to_* (x,m)$, so

$$\{0 \in W_{2n}^p\} = \{(0,2n) \to_* (x,0) \text{ for some } x \in W_0^p\}$$

To estimate the right hand side it is convenient to introduce

$$\hat{W}_m^{2n} = \{x : (0,2n) \to_* (x, 2n-m)\}$$
$$\hat{C}_{(0,2n)} = \{(x,t) : (0,2n) \to_* (x,t)\}$$

By conditioning on the value of $\hat{W}_{2n}^{2n}$, it is easy to see that

$$(A.3) \qquad P(0 \in W_{2n}^p) = 1 - E\left\{(1-p)^{|\hat{W}_{2n}^{2n}|}\right\}$$

so to complete the proof we want to show that if n is large and $\hat{W}_{2n}^{2n} \neq \emptyset$ then $|\hat{W}_{2n}^{2n}|$ is large with high probability. The process $\hat{W}_m^{2n}$ comes from random variables $\omega(x,n)$ that have property $(A.1)$, and the event on the left hand side of $(A.4)$ below implies that the contour associated with $\hat{C}_{(0,2n)}$ has length at least $4n$, so (A.2) implies

$$(A.4) \qquad P(\hat{W}_{2n}^{2n} \neq \emptyset, |\hat{C}_{(0,2n)}| < \infty) \le P(4n \le |\Gamma| < \infty) \le 2^{-4n}$$

Now the sites $(x,-1) \in \mathcal{L}$ are independent of those in $\mathcal{L}_0$ and are closed with probability γ so

$$(A.5) \qquad P\left(\hat{W}_{2n+1}^{2n} = \emptyset \,\middle|\, 0 < |\hat{W}_{2n}^{2n}| \le \sqrt{n}\right) \ge \theta^{2\sqrt{n}}$$

Combining $(A.4)$ and $(A.5)$ gives

$$(A.6) \qquad P(0 < |\hat{W}_{2n}^{2n}| \le \sqrt{n}) \le \frac{P(\hat{W}_{2n}^{2n} \neq \emptyset, |\hat{C}_{(0,2n)}| < \infty)}{P\left(\hat{W}_{2n+1}^{2n} = \emptyset \,\middle|\, 0 < |\hat{W}_{2n}^{2n}| \le \sqrt{n}\right)} \le 2^{-4n}\gamma^{-2\sqrt{n}}$$

Using $(A.3)$ in the first step; then $(A.6)$ and $P(|\hat{W}_{2n}^{2n}| > 0) \ge P(|\hat{C}_{(0,2n)}| = \infty)$ in the second; and finally, Theorem 4.1 in the third we have

$$\begin{aligned} P(0 \in W_{2n}^p) &\ge \left\{1 - (1-p)^{\sqrt{n}}\right\} P(|\hat{W}_{2n}^{2n}| \ge \sqrt{n}) \\ &\ge \left\{1 - (1-p)^{\sqrt{n}}\right\} \left(P(|\hat{C}_{(0,2n)}| = \infty) - 2^{-4n}\gamma^{-\sqrt{n}}\right) \\ &\ge \left\{1 - (1-p)^{\sqrt{n}}\right\} \left(1 - 55\,\gamma^{1/(2M+1)^2} - 2^{-4n}\gamma^{-\sqrt{n}}\right) \end{aligned}$$

which proves the desired result. □

The arguments for the last two results can be extended easily to give the conclusion quoted in Section 6 as (6.1):

Theorem A.3. If $p > 0$ then

$$\liminf_{n\to\infty} P(\{-2K,\dots,2K\} \cap W^p_{2n} \neq \emptyset) \geq 1 - \epsilon_K$$

where $\epsilon_K \to 0$ as $K \to \infty$.

PROOF: By the reasoning in the proof of Theorem A.2, we have $\{-2K,\dots,2K\} \cap W^p_{2n} \neq \emptyset$ if and only if there is a path down from some (x,n) with $|x| \leq 2K$ to $(y,0)$ for some $y \in W^p_{2n}$. To estimate the probability that this occurs we suppose that all the sites $\{-2K+1, -2K+3, \dots, 2K-1\}$ are open at time $2n+1$, let

$$\hat{\mathcal{C}} = \{(x,t) : (y, 2n+1) \to_* (x,t) \text{ for some } |y| \leq 2K-1\}$$

and turn the cluster $\hat{\mathcal{C}}$ into a solid blob by looking at

$$\hat{\mathcal{D}} = \cup_{(m,n)\in\hat{\mathcal{C}}} (m,n) + D$$

where $D = \{z \in \mathbf{R}^2 : \|z\|_1 \leq 1\}$. As in the proof of Theroem $A.1$ when $|\hat{\mathcal{C}}| < \infty$ we can define a contour associated with the cluster, and when the contour has length n there will be at least $n/4(2M+1)^d$ closed sites so that for each $z \neq w$ in this set $\|z-w\|_\infty > M$. Since this time the shortest contour has length $8K$ using $(A.2)$ gives

$$P(|\hat{\mathcal{C}}| < \infty) \leq 2^{-8K}$$

If we let

$$\hat{W}^{K,2n+1}_m = \{y : (x, 2n+1) \to_* (y, 2n-m) \text{ for some } |x| \leq 2K-1\}$$

then the argument in the proof of Theorem $A.2$ shows that

$$P(0 < |\hat{W}^{K,2n+1}_m| \leq \sqrt{n}) \leq 2^{-4n}\gamma^{-\sqrt{n}}$$

So repeating the last computation in the proof of Theorem A.2 proves the result with $\epsilon_K = 2^{-8K}$ □

Our last task is to prove Theorem 4.3. We begin by recalling the

Comparison Assumptions. We suppose given the following ingredients: a translation invariant finite range process $\xi_t : \mathbf{Z}^d \to \{0,1,\dots \kappa-1\}$ that is constructed from the graphical representation given in Section 2, an integer L, and a collection H of configurations determined by the values of ξ on $[-L,L]^d$ with the following property:

if $\xi \in H$ then there is an event G_ξ measurable with respect to the graphical representation in $[-k_0L, k_0L]^d \times [0, j_0T]$ and with $P(G_\xi) \geq (1-\theta)$ so that if $\xi_0 = \xi$ then on G_ξ, ξ_T lies in $\sigma_{2Le_1}H$ and in $\sigma_{-2Le_1}H$.

Here $(\sigma_y\xi)(x) = \xi(x+y)$ denote the translation (or shift) of ξ by y and $\sigma_y H = \{\sigma_y\xi : \xi \in H\}$. If we let $M = \max\{j_0, k_0\}$ then the space time regions

$$\mathcal{R}_{m,n} = (m2Le_1, nT) + \left\{[-k_0L, k_0L]^d \times [0, j_0T]\right\}$$

that correspond to points $(m,n), (m',n') \in \mathcal{L}$ with $\|(m,n)-(m',n')\|_\infty > M$ are disjoint.

Theorem A.4. If the comparison assumptions hold then we can define random variables $\omega(x,n)$ so that $X_n = \{m : (m,n) \in \mathcal{L}_0, \xi_{nT} \in \sigma_{m2Le_1}H\}$ dominates an M dependent oriented percolation process with initial configuration $W_0 = X_0$ and density at least $1-\gamma$, i.e., $X_n \supset W_n$ for all n.

PROOF: We will define the $\omega(x,n)$ in the oriented percolation by induction. We begin by setting $V_0 = X_0$ and defining a slightly enlarged version of the percolation process V_n consisting of all the y so that can be reached from some $(x_0, 0)$ with $x_0 \in V_0$ by a sequence $x_0, x_1, \ldots x_n = y$ so that $|x_k - x_{k-1}| = 1$ for $1 \leq k \leq n$ and $\omega(x_k, k) = 1$ for $0 \leq k < n$, i.e., the last point in the sequence does not have to be open. Since $V_n \supset W_n$ it suffices to show that $X_n \supset V_n$.

Let $n \geq 0$ and suppose that V_n and the $\omega(x,\ell)$ with $\ell < n$ have been defined so that $X_n \supset V_n$. To define the $\omega(m,n)$, and hence V_{n+1}, we consider two cases.

CASE 1. $m \in V_n \subset X_n$. We set $\omega(m,n) = 1$ if $G_{\sigma_{-m2Le_1}\xi_{nT}}$ occurs in the graphical representation translated by $-m2Le_1$ in space and $-nT$ in time, 0 otherwise. By assumption this event is determined by the Poisson points in $\mathcal{R}_{m,n}$, has probability at least $1-\gamma$, and guarantees that $(m+1), (m-1) \in X_{n+1}$.

CASE 2. $m \notin V_n$. In this case, the value of $\omega(m,n)$ is not important for the evolution of the percolation process so we set $\omega(m,n)$ equal to an independent random variable that is 1 with probability $1-\gamma$ and 0 with probability γ.

If $m \in V_{n+1}$ then either $m-1 \in V_n$ and $\omega(m-1,n) = 1$ or $m+1 \in V_n$ and $\omega(m+1,n) = 1$. In either case the observation in Case 1 implies that $m \in X_{n+1}$. The last conclusion and induction imply that $X_n \supset V_n$ for all n. The last detail to check is that the $\omega(m,n)$ satisfy $(A.1)$ and again we use induction. If $I = 1$ the conclusion is true, so suppose now that $k > 1$ and that the conclusion is true for $I = k-1$. Let (x_i, n_i) $1 \leq i \leq k$ be a seqeunce of points with $\|(x_i,n_i) - (x_j,n_j)\|_\infty > M$ if $i \neq j$ and suppose that the sequence has been indexed so that $n_k \geq n_j$ for all $j < k$. Let $\mathcal{F}$ be the information contained in the graphical representation up to time n_kT or in one of the space time boxes $\mathcal{R}_{m_i,n_i}$ with $i < k$. The comparison assumptions and the fact that $n_k \geq n_j$ for $j < k$ imply that

$$P(\omega(m_k, n_k) = 0 | \mathcal{F}) \leq \gamma$$

Integrating the last inequality over $E_{k-1} = \{\omega(m_i, n_i) = 0 \text{ for } i \leq k-1\} \in \mathcal{F}$ which by induction has probability smaller than γ^{k-1} gives

$$\gamma^k \geq \int_{E_{k-1}} \gamma \, dP \geq \int_{E_{k-1}} P(\omega(m_k, n_k) = 0 | \mathcal{F}) \, dP$$
$$= P(E_{k-1} \cap \{\omega(m_k, n_k) = 0\}) = P(E_k)$$

which verifies $(A.1)$ and completes the proof. □

REFERENCES

Andjel, E.D. T.M. Liggett, and T. Mountford (1992) Clustering in one dimensional threshold voter models. *Stoch. Processes Appl.* **42**, 73–90

Aronson, D.G. and H.F. Weinberger (1978) Multidimensional diffusion equations arising in population genetics. *Advances in Math.* **30**, 33–76

Asmussen, S. and N. Kaplan (1976) Branching random walks, I. *Stoch. Processes Appl.* **4**, 1–13

Bezuidenhout, C. and L. Gray (1993) Critical attractive spin systems. *Ann. Probab.*, to appear

Bezuidenhout, C. and G. Grimmett (1990) The critical contact process dies out. *Ann. Probab.* **18**, 1462–1482

Bezuidenhout, C. and G. Grimmett (1991) Exponential decay for subcritical contact and percolation processes. *Ann. Probab.* **19**, 984–1009

Boerlijst, M.C. and P. Hogeweg (1991) Spiral wave structure in pre-biotic evolution: hypercycles stable against parasites. *Physica D* **48**, 17–28

Bramson, M. (1983) Convergence of solutions of the Kolmogorov equation to travelling waves. *Memoirs of the AMS*, **285**

Bramson, M. and R. Durrett (1988) A simple proof of the stability theorem of Gray and Griffeath. *Probab. Th. Rel. Fields* **80**, 293–298

Bramson, M., R. Durrett, and G. Swindle (1989) Statistical mechanics of Crabgrass. *Ann. Prob.* **17**, 444–481

Bramson, M. and L. Gray (1992) A useful renormalization argument. Pages ??? in *Random Walks, Brownian Motion, and Interacting Particle Systems*, edited by R. Durrett and H. Kesten, Birkhauser, Boston

Bramson, M. and D. Griffeath (1987) Survival of cyclic particle systems. Pages 21–30 in *Percolation Theory and Ergodic Theory of Infinite Particle Systems* edited by H. Kesten, IMA Vol. 8, Springer

Bramson, M. and D. Griffeath (1989) Flux and fixation in cyclic particle systems *Ann. Probab.* **17**, 26–45

Bramson, M. and C. Neuhauser (1993) Survival of one dimensional cellular automata. Preprint

Chen, H.N. (1992) On the stability of a population growth model with sexual reproduction in $\mathbf{Z}^2$. *Ann. Probab.* **20**, 232–285

Cox, J.T. (1988) Coalescing random walks and voter model consensus times on the torus in $\mathbf{Z}^d$. *Ann. Probab.*

Cox, J.T. and R. Durrett (1988) Limit theorems for the spread of epidemics and forest fires. *Stoch. Processes Appl.* **30**, 171–191

Cox, J.T. and R. Durrett (1992) Nonlinear voter models. Pages 189–202 in *Random Walks, Brownian Motion, and Interacting Particle Systems*, edited by R. Durrett and H. Kesten, Birkhauser, Boston

Cox, J.T. and D. Griffeath (1986) Diffusive clustering in the two dimensional voter model. *Ann. Probab.* **14**, 347–370

DeMasi, A., P. Ferrari, and J. Lebowitz (1986) Reaction diffusion equations for interacting particle systems. *J. Stat. Phys.* **44**, 589–644

DeMasi, A. and E. Presutti (1991) *Mathematical Methods for Hydrodynamic Limits*. Lecture Notes in Math **1501**, Springer, New York

Durrett, R. (1980) On the growth of one dimensional contact processes. *Ann. Probab.* **8**, 890-907

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, CA

Durrett, R. (1991a) Stochastic models of growth and competition. Pages 1049–1056 in *Proceedings of the International Congress of Mathematicians, Kyoto*, Springer, New York

Durrett, R. (1991b) The contact process, 1974-1989. Pages 1-18 in *Proceedings of the AMS Summer seminar on Random Media.* Lectures in Applied Math **27**, AMS, Providence, RI

Durrett, R. (1991c) Some new games for your computer. *Nonlinear Science Today* Vol. 1, No. 4, 1-7

Durrett, R. (1992a) Multicolor particle systems with large threshold and range. *J. Theoretical Prob.*, 5 (1992), 127–152

Durrett, R. (1992b) A new method for proving the existence of phase transitions. Pages 141-170 in *Spatial Stochastic Processes*, edited by K.S. Alexander and J.C. Watkins, Birkhauser, Boston

Durrett, R. (1992c) Stochastic growth models: bounds on critical values. *J. Appl. Prob.* **29**

Durrett, R. (1992d) *Probability: Theory and Examples.* Wadsworth, Belmont, CA

Durrett, R. (1993) Predator-prey systems. Pages 37–58 in *Asymptotic Problems in Probability Theory: Stochastic Models and Diffusions on Fractals*, edited by K.D. Elworthy and N. Ikeda, Pitman Research Notes in Math 283, Longman, Essex, England

Durrett, R. and D. Griffeath (1993) Asymptotic behavior of excitable cellular automata. Preprint

Durrett, R. and S. Levin (1993) Stochastic spatial models: A user's guide to ecological applications. *Phil. Trans. Roy. Soc. B*, to appear

Durrett, R. and A.M. Moller (1991) Complete convergence theorem for a competition model. *Probab. Th. Rel. Fields* **88**, 121–136

Durrett, R. and C. Neuhauser (1991) Epidemics with recovery in $d = 2$. *Ann. Applied Probab.* **1**, 189-206

Durrett, R. and C. Neuhauser (1993) Particle systems and reaction diffusion equations. *Ann. Probab.*, to appear

Durrett, R. and R. Schinazi (1993) Asymptotic critical value for a competition model. *Ann. Applied Probab.*, to appear

Durrett, R. and J. Steif (1993) Fixation results for threshold voter models. *Ann. Probab.*, to appear

Durrett, R. and G. Swindle (1991) Are there bushes in a forest? *Stoch. Proc. Appl.* **37**, 19–31

Durrett, R. and G. Swindle (1993) Coexistence results for catalysts. Preprint

Eigen, M. and P. Schuster (1979) *The Hypercycle: A Principle of Natural Self-Organization*, Springer, New York

Fife, P.C. and J.B. McLeod (1977) The approach of solutions of nonlinear diffusion equations to travelling front solutions. *Arch. Rat. Mech. Anal.* **65**, 335–361

Fisch, R. (1990a) The one dimensional cyclic cellular automaton: a system with deterministic dynamics which emulates a particle system with stochastic dynamics. *J. Theor. Prob.* **3**, 311–338

Fisch, R. (1990b) Cyclic cellular automata and related processes. *Physica D* **45**, 19–25

Fisch, R. (1992) Clustering in the one dimensional 3-color cyclic cellular automaton. *Ann. Probab.* **20**, 1528–1548

Fisch, R., J. Gravner, and D. Griffeath (1991) Threshold range scaling of excitable cellular automata. *Statistics and Computing* **1**, 23–39

Fisch, R., J. Gravner, and D. Griffeath (1992) Cyclic cellular automata in two dimensions. In *Spatial Stochastic Processes* edited by K. Alexander and J. Watkins, Birkhauser, Boston

Fisch, R., J. Gravner, and D. Griffeath (1993) Metastability in the Greenberg Hastings model. *Ann. Applied. Probab.*, to appear

Grannan, E. and G. Swindle (1991) Rigorous results on mathematical models of catalyst surfaces. *J. Stat. Phys.* **61**, 1085–1103

Gravner, J. and D. Griffeath (1993) Threshold growth dyanmics. *Transactions A.M.S.*, to appear

Gray, L. and D. Griffeath (1982) A stability criterion for attractive nearest neighbor spin systems on Z. *Ann. Probab.* **10**, 67–85

Gray, L. (1987) Behavior of processes with statistical mechanical properties. Pages 131–168 in *Percolation Theory and Ergodic Theory of Infinite Particle Systems* edited by H. Kesten, IMA Vol. 8, Springer

Griffeath, D. (1979) *Additive and Cancellative Interacting Particle Systems.* Lecture Notes in Math **724**, Springer

Harris, T.E. (1972) Nearest neighbor Markov interaction processes on multidimensional lattices. *Adv. in Math.* **9**, 66–89

Harris, T.E. (1976) On a class of set valued Markov processes. *Ann. Probab.* **4**, 175–194

Hassell, M.P., H.N. Comins, and R.M. May (1991) Spatial structure and chaos in insect population dynamics. *Nature* **353**, 255–258

Hirsch, M.W. and S. Smale (1974) *Differential Equations, Dynamical Systems, and Linear Algebra*, Academic Press, New York

Holley, R.A. (1972) Markovian interaction processes with finite range interactions. *Ann. Math. Stat.* **43**, 1961–1967

Holley, R.A. (1974) Remarks on the FKG inequalities. *Commun. Math. Phys.* **36**, 227–231

Holley, R.A., and T.M. Liggett (1975) Ergodic theorems for weakly interacting systems and the voter model. *Ann. Probab.* **3**, 643–663

Holley, R.A., and T.M. Liggett (1978) The survival of contact processes. *Ann. Probab.* **6**, 198–206

Kinzel, W. and J. Yeomans (1981) Directed percolation: a finite size renormalization approach. *J. Phys. A* **14**, L163–L168

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

Liggett, T.M. (1993) Coexistence in threshold voter models. *Ann. Probab.*, to appear

Neuhauser, C. (1992) Ergodic theorems for the multitype contact process. ***Probab. Theory Rel. Fields.*** **91**, 467–506

Redheffer, R., R. Redlinger, and W. Walter (1988) A theorem of La Salle-Lyapunov type for parabolic systems. *SIAM J. Math. Anal.* **19**, 121–132

Schonmann, R.H. and M.E. Vares (1986) The survival of the large dimensional basic contact process. *Probab. Th. Rel. Fields* **72**, 387–393

Spohn, H. (1991) *Large Scale Dynamics of Interacting Particle Systems*, Springer, New York

Zhang, Yu (1992) A shape theorem for epidemics and forest fires with finite range interactions. Preprint