

Interacting Particle Systems –
An Introduction

Thomas M. Liggett*

Mathematics Department, University of California, Los Angeles, USA

*Lectures given at the
School and Conference on Probability Theory
Trieste, 13 - 31 May, 2002*

LNS0417001

*tmlmath.ucla.edu

Abstract

Interacting particle systems is a large and growing field of probability theory that is devoted to the rigorous analysis of certain types of models that arise in statistical physics, biology, economics, and other fields. In these notes, we provide an introduction to some of these models, give some basic results about them, and explain how certain important tools are used in their study. The first chapter describes contact, voter and exclusion processes, and introduces the tools of coupling and duality. Chapter 2 is devoted to an analysis of translation invariant linear voter models, using primarily the duality that is available in that case. Chapters 3-5 are concerned with the exclusion process, beginning with the symmetric case, in which one can also use duality, and following with asymmetric systems, which are studied using coupling and other monotonicity techniques. At the end, we report on some very recent work on the stationary distributions of one dimensional systems with positive drift.

Contents

1	Introduction	5
1.1	Examples	6
1.2	Coupling	10
1.3	Duality	15
2	Linear Voter Models	20
2.1	The recurrent case	22
2.2	The transient case	23
3	Symmetric Exclusion Processes	32
3.1	The recurrent case	33
3.2	The transient case	34
3.3	The general (non-translation invariant) case	38
4	Translation Invariant Exclusion Processes	40
4.1	Stationary product measures	40
4.2	Coupling	42
4.3	Shift invariant stationary measures	43
4.4	Some extensions	46
5	One Dimensional Exclusion Processes	48
5.1	The nearest neighbor case	48
5.2	The general case	51
	References	56

1 Introduction

These notes are intended to provide a first exposure to the area of interacting particle systems. More comprehensive treatments of this field can be found in the author's 1985 and 1999 books, as well as in the references in the latter book. The main prerequisite for reading these notes is a good background in measure theoretic probability theory.

The processes we will discuss here are continuous time Feller processes η_t on the compact configuration space $\{0, 1\}^S$, where S is a countable set. (Recall that a Feller process is a strong Markov process whose transition measures are weakly continuous in the initial state.) A very simple such process is that in which the coordinates $\eta_t(x)$ evolve according to independent two state Markov chains with transitions from 0 to 1 and from 1 to 0 at rate 1. Even this simple example illustrates some of the important differences between particle systems and more classical Feller processes such as Brownian motion on R^d . The distribution of Brownian motion at any positive time is equivalent to Lebesgue measure, and hence these distributions at different positive times are equivalent to each other. In our example of independent two state Markov chains, the situation is entirely different. For example, if $\eta_0 \equiv 1$, then the distribution at every time is a homogeneous product measure with time dependent density, so the distributions of the process at different times are mutually singular with respect to one another.

The main issues to be considered here are:

(a) a description of the class \mathcal{I} of stationary distributions for the process, and

(b) limit theorems for the distribution of η_t as $t \rightarrow \infty$.

Settling these issues is naturally very easy in the example of independent two state Markov chains. The only stationary distribution is the product measure with density $\frac{1}{2}$, and there is convergence to this limit for every initial distribution.

In general, all that can be said is that since $\{0, 1\}^S$ is compact and the process is Feller, \mathcal{I} is always nonempty. Typically, there will be some stationary distributions that can be written down explicitly, and the task is to determine whether or not these exhaust all of \mathcal{I} . If not, these additional stationary distributions must usually be constructed by some type of limiting argument. The set \mathcal{I} is convex, and we will denote its extreme points by \mathcal{I}_e .

The process η_t is usually described by specifying the rates at which transitions occur. If S is finite, saying that the transition $\eta \rightarrow \zeta$ (for $\eta \neq \zeta$)

occurs at rate c means that

$$P^\eta(\eta_t = \zeta) = ct + o(t) \quad (1)$$

as $t \downarrow 0$. If S is infinite, the probability on the left of (1) is typically 0, but the intuitive meaning of “rate” is similar. One simply replaces the event on the left of (1) by $\{\eta_t = \zeta \text{ on } A\}$ for large finite sets $A \subset S$.

The precise relation between the process and its transition rates is provided by the infinitesimal generator Ω of η_t . It is a (typically unbounded) operator defined on an appropriate dense subset of $C(\{0, 1\}^S)$; Ω is determined by its values on the cylinder functions, i.e., functions that depend on finitely many coordinates. For cylinder functions, it takes the form

$$\Omega f(\eta) = \sum_{\zeta} c(\eta, \zeta) [f(\zeta) - f(\eta)], \quad (2)$$

where $c(\eta, \zeta)$ is the rate at which transitions occur from η to ζ . Our choices of rates will guarantee that the series in (2) converges uniformly for cylinder f 's. The stationary distributions of the process are determined by the generator as follows:

Theorem 1. *A probability measure μ on $\{0, 1\}^S$ is stationary for the process η_t if and only if*

$$\int \Omega f d\mu = 0$$

for all cylinder functions f on $\{0, 1\}^S$.

Of course, assumptions on the transition rates must be made in order for statements of this sort to hold. General conditions under which they do hold can be found in Chapter I of Liggett (1985). Our approach in these notes is to make assumptions for each model considered which will guarantee that statements such as those in Theorem 1 are correct.

1.1 Examples

We next turn to some of the most important models considered in this field. To describe their transition rates, we need the following notation: If $\eta \in \{0, 1\}^S$ and $x, y \in S$, then $\eta_x, \eta_{x,y} \in \{0, 1\}^S$ are defined by

$$\eta_x(u) = \begin{cases} \eta(u) & \text{if } u \neq x, \\ 1 - \eta(u) & \text{if } u = x, \end{cases} \quad \text{and} \quad \eta_{x,y}(u) = \begin{cases} \eta(u) & \text{if } u \neq x, y, \\ \eta(y) & \text{if } u = x, \\ \eta(x) & \text{if } u = y. \end{cases}$$

Thus η_x is obtained from η by changing its value at x , while $\eta_{x,y}$ is obtained by interchanging the values at x and y . In the latter case, if $\eta(x) \neq \eta(y)$, the transition $\eta \rightarrow \eta_{x,y}$ can be interpreted as moving a particle from x to y or vice versa.

Example 1. The contact process. Here S is a graph whose vertices have bounded degree, and λ is a positive parameter. Use the notation $x \sim y$ to mean that the vertices x and y are connected by an edge. Then for each $x \in S$,

$$\eta \rightarrow \eta_x \text{ at rate } \begin{cases} 1 & \text{if } \eta(x) = 1, \\ \lambda |\{y \sim x : \eta(y) = 1\}| & \text{if } \eta(x) = 0. \end{cases}$$

Here $|A|$ denotes the cardinality of the set A . The interpretation is that sites with $\eta(x) = 1$ are infected, while sites with $\eta(x) = 0$ are healthy. Infected sites recover from the infection after an exponential time of rate 1, while healthy sites become infected at a rate proportional to the number of infected neighbors. The only “trivial” (in the sense that it can be found by inspection) stationary distribution is the pointmass δ_0 on the configuration $\eta \equiv 0$.

While we have described the contact process as a model for the spread of infection, it arises in other contexts as well. For example, it is related to Reggeon Field Theory in high energy physics, and it is a building block for more complex models in biology.

Here are some of the answers to our basic questions in the case of the contact process. They illustrate some of the variety of behavior that even relatively simple systems can exhibit.

If S is finite, then $\mathcal{I} = \{\delta_0\}$, and η_t is eventually $\equiv 0$ for any initial configuration. This follows from finite state Markov chain theory.

If $S = \mathbb{Z}^d$, the d -dimensional integer lattice, then there is a critical value $\lambda(d) \in (\frac{1}{2d-1}, \frac{2}{d})$ so that

(i) $\lambda \leq \lambda(d)$ implies that $\mathcal{I} = \{\delta_0\}$ and $\eta_t \rightarrow \delta_0$ weakly for any initial configuration,

and

(ii) $\lambda > \lambda(d)$ implies that $\mathcal{I}_e = \{\delta_0, \nu\}$ for some $\nu \neq \delta_0$, and $\eta_t \rightarrow \nu$ weakly for any initial configuration with infinitely many infected sites.

Remark. If $d = 1$, $\lambda = \lambda(1)$, and $\eta_0 \equiv 1$, then $\eta_t \rightarrow \delta_0$ weakly, yet with probability 1, for every x , $\eta_t(x) = 1$ for arbitrarily large t 's – see Theorem

3.10 of Chapter VI of Liggett (1985) for the latter statement. This illustrates the fact that weak convergence results often miss interesting a.s. behavior.

If $S = T_d$ ($d \geq 2$), the tree in which every vertex has $d + 1$ neighbors, then there are two critical values satisfying

$$\lambda_1(d) < \lambda_2(d), \quad \frac{1}{d+1} \leq \lambda_1(d) \leq \frac{1}{d-1}, \quad \frac{1}{2\sqrt{d}} \leq \lambda_2(d) \leq \frac{1}{\sqrt{d-1}},$$

so that

(i) $\lambda \leq \lambda_1(d)$ implies that $\mathcal{I} = \{\delta_0\}$ and $\eta_t \rightarrow \delta_0$ weakly for any initial configuration,

(ii) $\lambda_1(d) < \lambda \leq \lambda_2(d)$ implies that \mathcal{I}_e is infinite, and

(iii) $\lambda > \lambda_2(d)$ implies that $\mathcal{I}_e = \{\delta_0, \nu\}$ for some $\nu \neq \delta_0$, and $\eta_t \rightarrow \nu$ weakly for any initial configuration with infinitely many infected sites.

In case (ii), if the initial configuration η has finitely many infected sites, then

$$P^\eta(\eta_t \neq 0 \forall t > 0) > 0, \text{ but } \forall x \in S, P^\eta(\exists T \text{ so that } \eta_t(x) = 0 \forall t \geq T) = 1.$$

That is, the set of infected sites does not become empty (with positive probability), but it “wanders out to ∞ ”, so that every site is eventually and thereafter healthy.

The proofs of these results when $S = Z^d$ or $S = T_d$ are quite involved, and can be found in Part I of Liggett (1999). After describing two more examples, we will return to the contact process to illustrate the use of two important techniques: coupling and duality.

Example 2. The linear voter model. Here S is an arbitrary countable set, and $p(x, y)$ are the transition probabilities for a Markov chain on S :

$$p(x, y) \geq 0 \quad \text{and} \quad \sum_y p(x, y) = 1.$$

The transition rates are now given by

$$\eta \rightarrow \eta_x \text{ at rate } \sum_{y: \eta(y) \neq \eta(x)} p(x, y).$$

The interpretation is that sites are individuals who at any time can have one of two opinions (denoted by 0 and 1) on an issue. At exponential times of rate 1, the individual at x chooses a y with probability $p(x, y)$, and adopts

y 's opinion. An alternate interpretation is in terms of spatial conflict. Two nations control the areas $\{x : \eta(x) = 0\}$ and $\{x : \eta(x) = 1\}$ respectively. A flip from 0 to 1 at x , for example, then represents an invasion of x by the second nation.

The trivial stationary distributions for the linear voter model are the pointmasses δ_0 and δ_1 on $\eta \equiv 0$ and $\eta \equiv 1$ respectively. We will see in Chapter 2 that there is a close connection between the existence of nontrivial stationary distributions for the voter model and the recurrence properties and harmonic functions of the Markov chain with transition probabilities $p(x, y)$.

Example 3. The exclusion process. Again S is a general countable set and $p(x, y)$ are the transition probabilities for a Markov chain on S . Now we must also assume that $p(x, y)$ satisfies

$$\sup_y \sum_x p(x, y) < \infty \quad (3)$$

in order to guarantee that η_t is well defined as a Feller process. The transition rates are given by

$$\eta \rightarrow \eta_{x,y} \quad \text{at rate} \quad p(x, y) \quad \text{if } \eta(x) = 1, \eta(y) = 0.$$

The interpretation of $\eta(x) = 0$ or 1 is that site x is vacant in the first case, and is occupied by a particle in the second case. In the evolution, a particle at x waits a unit exponential time, and then chooses a y with probability $p(x, y)$. If y is vacant, it moves to y , while if y is occupied, the particle remains at x . The generator takes the following form for cylinder functions f :

$$\Omega f(\eta) = \sum_{\eta(x)=1, \eta(y)=0} p(x, y) [f(\eta_{x,y}) - f(\eta)].$$

Note that (3) implies that this series converges uniformly, and hence defines a continuous function on $\{0, 1\}^S$.

Again, δ_0 and δ_1 are stationary, and it is often possible to produce other stationary distributions explicitly. To see that one should expect to have more stationary distributions for the exclusion process than for the linear voter model, consider the case in which S is finite. Assuming irreducibility of the Markov chain with transition probabilities $p(x, y)$, it is easy to see that $\mathcal{I}_e = \{\delta_0, \delta_1\}$ in the case of the linear voter model, while for the exclusion process, \mathcal{I}_e has one element for each $0 \leq n \leq |S|$ — it is the stationary

distribution for the exclusion process restricted to configurations of n particles. Chapters 3–5 are devoted to a description of \mathcal{I} in many cases, but there are still many interesting open problems in this context. For example, what analogues of Theorems 3–5 of Chapter 5 can be proved in dimensions greater than 1? See the discussion at the end of Chapter 4 and the paragraph following the statement of Theorem 4 of Chapter 5 as well.

The remainder of this chapter is devoted to a discussion of two important tools that we will use repeatedly, coupling and duality. Their use will be illustrated in two contexts – Markov chains and the contact process. At the end, we will derive the duality relations for the voter and exclusion processes that we will use in the later chapters.

1.2 Coupling

A coupling is a joint definition of two or more stochastic processes on a common probability space. It is a surprisingly powerful tool. Before discussing its use in the context of particle systems, we consider two applications of this technique to Markov chains.

1.2a. Coupling for Markov chains. Suppose first that X_t is a birth and death chain on the integers. This is a Markov chain that can only move to nearest neighbors. Let (X_t, Y_t) be two copies of this chain that are coupled in the following way:

- (a) X_t and Y_t move independently until the first time τ (if ever) that $X_t = Y_t$, and
- (b) X_t and Y_t move together after time τ .

This coupling has the property that $X_0 \leq Y_0$ implies that $X_t \leq Y_t$ for all $t \geq 0$. It follows that if f is any bounded increasing function on the integers, then

$$x \rightarrow E^x f(X_t)$$

is again an increasing function of x . We will exploit the analogous property for particle systems many times, beginning with our discussion of the contact process below.

Since harmonic functions will play an important role in Chapters 2 and 3, our second example is one in which coupling is used to prove that certain Markov chains have no nonconstant bounded harmonic functions. Recall that a function $f : S \rightarrow R^1$ is said to be harmonic for the Markov chain X_t if $E^x f(X_t) = f(x)$ for all $x \in S$ and all $t \geq 0$. Implicit in this definition is the requirement that the expected value be well defined, e.g., that f be bounded or nonnegative.

Theorem 2. *Suppose that X_t^1, \dots, X_t^n are independent irreducible Markov chains on the countable set S with the following property: For each i , two independent copies of X_t^i will eventually meet with probability 1. If f is a bounded harmonic function for the Markov chain $X_t = (X_t^1, \dots, X_t^n)$ on S^n , then f is constant.*

Proof. Define a coupling of two copies, X_t and Y_t , of the Markov chain in the following way:

(a) X_t and Y_t evolve independently until the first time τ_1 that $X_t^1 = Y_t^1$. (This will happen eventually by our assumption.)

(b) For $t \geq \tau_1$, X_t^1 and Y_t^1 run together, and the other $n - 1$ coordinates run independently, until the first time τ_2 that $X_t^2 = Y_t^2$.

(c) Continuing in this way, we find a finite stopping time $\tau = \tau_n$ so that for all $t \geq \tau$, $X_t = Y_t$.

Now fix $x, y \in S^n$ and use the initial states $X_0 = x$ and $Y_0 = y$. It follows that

$$|f(x) - f(y)| = |Ef(X_t) - Ef(Y_t)| \leq E|f(X_t) - f(Y_t)| \leq 2\|f\|_\infty P(\tau > t).$$

To complete the proof, let $t \rightarrow \infty$.

Remarks. (a) The assumption of Theorem 2 is satisfied, for example, by any irreducible recurrent random walk on Z^1 or Z^2 . To see this, let $p_t(x, y)$ be the transition probabilities for X_t^i , let X_t^i and Y_t^i be two independent copies of the random walk, and write

$$\begin{aligned} p_{2t}(0, 0) &= \sum_u p_t(0, u)p_t(u, 0) \\ &\leq \sqrt{\sum_u p_t^2(0, u) \sum_u p_t^2(u, 0)} \\ &= \sum_u p_t^2(0, u) = P^{(0,0)}(X_t^i = Y_t^i). \end{aligned}$$

By recurrence, the integral of the left side above is infinite. Therefore, the integral of the right side is infinite, and hence the random walk $X_t^i - Y_t^i$ is recurrent.

(b) We will use repeatedly the fact that irreducible random walks on Z^d have no nonconstant bounded harmonic functions. For many random walks, this fact follows from Theorem 2. A coupling proof for general random walks can be found on pages 69–70 of Liggett (1985).

1.2b. Coupling for the contact process. Turning to applications to particle systems, define a partial order on the set of probability measures on $\{0, 1\}^S$ by saying that $\mu \leq \nu$ if

$$\int f d\mu \leq \int f d\nu$$

for all bounded increasing functions f on $\{0, 1\}^S$. This is equivalent to the existence of a probability measure γ on $\{0, 1\}^S \times \{0, 1\}^S$ with marginals μ and ν that satisfies

$$\gamma(\{(\eta, \zeta) : \eta \leq \zeta\}) = 1$$

(Theorem 2.4 in Chapter II of Liggett (1985)). The partial order on $\{0, 1\}^S$ that we are using here is coordinatewise: $\eta \leq \zeta$ iff $\eta(x) \leq \zeta(x)$ for all $x \in S$.

The contact process is an example of an *attractive* process, i.e., one for which two copies started from η and ζ with $\eta \leq \zeta$ can be coupled in such a way that $\eta_t \leq \zeta_t$ for all $t \geq 0$. (We saw earlier that birth and death chains are attractive.) The property of the rates that allows for this coupling is the following: The rate of the transition $\eta \rightarrow \eta_x$ is an increasing function of η if $\eta(x) = 0$ and a decreasing function if $\eta(x) = 1$. Specifically, writing the ζ coordinate above the η coordinate, we can use the coupling that has the following transitions at site x :

$$\begin{aligned} \begin{pmatrix} 1 \\ 1 \end{pmatrix} &\rightarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ at rate } 1, \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} &\rightarrow \begin{cases} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ at rate } \lambda|\{y \sim x : \eta(y) = 1\}| \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ at rate } 1, \end{cases} \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} &\rightarrow \begin{cases} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ at rate } \lambda|\{y \sim x : \eta(y) = 1\}| \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ at rate } \lambda|\{y \sim x : \zeta(y) = 1, \eta(y) = 0\}|. \end{cases} \end{aligned}$$

It follows from the above coupling that if f is increasing on $\{0, 1\}^S$, then so is the function $\eta \rightarrow E^\eta f(\eta_t)$ for any $t \geq 0$. To see this, take $\eta \leq \zeta$ and use the coupling to write

$$\eta_t \leq \zeta_t \quad \Rightarrow \quad f(\eta_t) \leq f(\zeta_t) \quad \Rightarrow \quad E f(\eta_t) \leq E f(\zeta_t).$$

A consequence of this is that the relation $\mu \leq \nu$ is preserved by the evolution: If μ_t and ν_t are the distributions of the process at time t with initial distributions μ_0 and ν_0 respectively, then

$$\mu_0 \leq \nu_0 \quad \Rightarrow \quad \mu_t \leq \nu_t \quad (4)$$

for all $t \geq 0$. Again this is immediate, since for increasing f ,

$$\int f d\mu_t = \int E^n f(\eta_t) d\mu_0 \leq \int E^n f(\eta_t) d\nu_0 = \int f d\nu_t.$$

In order to develop some consequences of these ideas, it is convenient to introduce some notation. We will write $\mu S(t)$ for the distribution of the process at time t when the initial distribution is μ . The reason for this choice is the following. The semigroup $S(t)$ of the process η_t is defined for continuous functions f by $S(t)f(\eta) = E^n f(\eta_t)$. We then have

$$E^\mu f(\eta_t) = \int E^n f(\eta_t) d\mu = \int S(t)f d\mu = \int f d[\mu S(t)].$$

Therefore we use the same notation $S(t)$ for the dual objects that operate on functions or measures respectively.

Since any probability measure μ satisfies $\mu \leq \delta_1$, it follows from (4) that $\mu S(t) \leq \delta_1 S(t)$. Applying this to $\mu = \delta_1 S(s)$ and using the Markov property, it follows that

$$\delta_1 S(t+s) = \delta_1 S(s)S(t) \leq \delta_1 S(t).$$

Therefore $\delta_1 S(t) \downarrow$ in t , so that the weak limit

$$\nu = \lim_{t \rightarrow \infty} \delta_1 S(t) \quad (5)$$

exists. (For this, we use the fact that every cylinder function can be written as a finite linear combination of increasing cylinder functions.) This ν is the distribution that is referred to in our discussion of Example 1 above. It is often called the *upper invariant measure*, since it is the largest possible stationary distribution.

Of course, it is possible that $\nu = \delta_0$. In fact, that is what happens for small values of λ , as we will see below in Theorem 3. If that is the case, another application of coupling gives the following limit theorem:

$$\mu S(t) \rightarrow \delta_0$$

as $t \uparrow \infty$ for all initial distributions μ . To see this, use (4) to write

$$\delta_0 \leq \mu \leq \delta_1 \quad \Rightarrow \quad \delta_0 = \delta_0 S(t) \leq \mu S(t) \leq \delta_1 S(t),$$

and pass to the limit as $t \uparrow \infty$.

We would like to know that the set of λ 's for which $\nu = \delta_0$ is an interval, so that we can define the critical value λ_c as the right endpoint of that interval. But this again an easy consequence of coupling. To see this, write ν_λ for the limit in (5). We would like to know that ν_λ is increasing in λ . So, take $\lambda_1 < \lambda_2$, and couple copies η_t and ζ_t of the processes with these parameter values respectively, and initial distribution δ_1 so that $\eta_t \leq \zeta_t$ for all $t \geq 0$. (This is possible because the processes are attractive, and the transition rates are monotone functions of λ .) It follows that the distributions of the two processes at time t are ordered, and hence so are their limits as $t \uparrow \infty$: $\nu_{\lambda_1} \leq \nu_{\lambda_2}$.

As a final application of coupling to the contact process, we will prove the following result. Note that it gives the lower bound for $\lambda_1(d)$ for the process on T_d that was stated in Example 1 above. For the process on Z^d , it gives a slightly weaker lower bound for $\lambda(d)$.

Theorem 3. *Suppose the degree of every vertex of the graph S is at most d . Then the critical value λ_c satisfies*

$$\lambda_c \geq \frac{1}{d}.$$

Proof. Let η_t be the contact process on S , and let ζ_t be the branching random walk on S that is defined as follows. The value $\zeta_t(x)$ is a nonnegative integer that represents the number of particles at x at time t . The possible transitions are

$$\zeta(x) \rightarrow \zeta(x) - 1 \text{ at rate } \zeta(x),$$

and

$$\zeta(x) \rightarrow \zeta(x) + 1 \text{ at rate } \lambda \sum_{y \sim x} \zeta(y).$$

In other words, each particle dies at rate 1, and for neighboring vertices x and y , each particle at y gives birth to a new particle at x at rate λ . Now couple these processes with initial configurations $\eta_0 \equiv 1$ and $\zeta_0 \equiv 1$, so that $\eta_t(x) \leq \zeta_t(x)$ for all $t \geq 0$ and $x \in S$. Then

$$\frac{d}{dt} E\zeta_t(x) = \lambda \sum_{y \sim x} E\zeta_t(y) - E\zeta_t(x),$$

which, as we will see, implies that $E\zeta_t(x) \leq e^{(\lambda d - 1)t}$ for all $x \in S$ and $t \geq 0$. (Note that this is true with equality on a homogeneous graph of constant degree d , since then $f(t) = E\zeta_t(x)$ is independent of x and satisfies $f'(t) = (\lambda d - 1)f(t)$.) So, if $\lambda d < 1$, it follows that $E\zeta_t(x)$, and hence $E\eta_t(x)$ tends to 0 as $t \uparrow \infty$.

To check $E\zeta_t(x) \leq e^{(\lambda d - 1)t}$, one can proceed as follows. Rewrite the expression for the derivative above as

$$\frac{d}{dt} e^t E\zeta_t(x) = \lambda e^t \sum_{y \sim x} E\zeta_t(y).$$

Integrate this from 0 to t , and then take suprema appropriately to get

$$M(t) \leq e^{-t} + \lambda d \int_0^t e^{-(t-s)} M(s) ds,$$

where

$$M(t) = \sup_x E\zeta_t(x).$$

We need to show that $M(t) \leq e^{(\lambda d - 1)t}$ for all $t \geq 0$. Elementary estimates imply that $M(t)$ is bounded on bounded t sets. Suppose that $C \geq 1$ and $T > 0$ satisfy

$$M(t) \leq C e^{(\lambda d - 1)t}$$

for $0 \leq t \leq T$. (By local boundedness, such a C exists for every choice of T .) Using the above integral inequality, it follows that

$$M(t) \leq C' e^{(\lambda d - 1)t}$$

for $0 \leq t \leq T$, with $C' = C - (C - 1)e^{-\lambda d T}$. Iterating this argument gives

$$M(t) \leq e^{(\lambda d - 1)t}$$

for $0 \leq t \leq T$ as required.

1.3 Duality

Suppose $H(\eta, \zeta)$ is a nonnegative continuous function of two variables. The Markov processes η_t and ζ_t are said to be dual with respect to H if

$$E^\eta H(\eta_t, \zeta) = E^\zeta H(\eta, \zeta_t) \tag{6}$$

for all η, ζ and all $t \geq 0$. For reasonable choices of H , this relation means that probabilities related to one of the processes can be expressed in terms

of probabilities related to the other. Subtracting $H(\eta, \zeta)$ from both sides of (6), dividing by t and letting $t \downarrow 0$, one formally obtains the following relation between the generators Ω_1 and Ω_2 of the processes η_t and ζ_t :

$$\Omega_1 H(\cdot, \zeta)(\eta) = \Omega_2 H(\eta, \cdot)(\zeta). \quad (7)$$

Under weak conditions, which are satisfied by all the examples we will consider, one can prove that (6) and (7) are equivalent.

1.3a. Duality for Markov chains. To illustrate the use of duality, suppose that η_t is a birth and death chain on $\{0, 1, 2, \dots\}$ that is irreducible, except for the fact that 0 is a trap. In other words, η_t has transitions

$$n \rightarrow n + 1 \quad \text{at rate} \quad \beta(n)$$

and

$$n \rightarrow n - 1 \quad \text{at rate} \quad \delta(n)$$

for $n \geq 1$, where $\beta(n) > 0$ and $\delta(n) > 0$ for all $n \geq 1$. Letting $H(\eta, \zeta) = 1_{\{\eta \leq \zeta\}}$, we can compute the left side of (7) as follows:

$$\begin{aligned} \Omega_1 H(\cdot, \zeta)(\eta) &= \beta(\eta) [H(\eta + 1, \zeta) - H(\eta, \zeta)] + \delta(\eta) [H(\eta - 1, \zeta) - H(\eta, \zeta)] \\ &= \beta(\zeta) [H(\eta, \zeta - 1) - H(\eta, \zeta)] + \delta(\zeta + 1) [H(\eta, \zeta + 1) - H(\eta, \zeta)] \\ &= \Omega_2 H(\eta, \cdot)(\zeta), \end{aligned} \quad (8)$$

where Ω_2 is the generator of the birth and death process with transitions

$$n \rightarrow n - 1 \quad \text{at rate} \quad \beta(n)$$

and

$$n \rightarrow n + 1 \quad \text{at rate} \quad \delta(n + 1).$$

Note that this chain is irreducible, and the roles of the β 's and δ 's have been reversed. Writing (6) as

$$P^\eta(\eta_t \leq \zeta) = P^\zeta(\eta \leq \zeta_t)$$

and passing to the limit as $t \uparrow \infty$, we see that η_t has positive probability of escaping to ∞ if and only if ζ_t is positive recurrent, and in this case,

$$P^\eta(\eta_t \text{ is absorbed at } 0) = \sum_{\zeta \geq \eta} \pi(\zeta),$$

where π is the stationary distribution of the chain ζ_t .

1.3b. Duality for the contact process. The computation in (8) can be repeated in other contexts. The trick is to find an appropriate duality function H . One should look for relatively simple ones, since otherwise it is difficult to compute with them.

For example, consider the contact process η_t on the graph S , and let Ω be its generator. For $\eta \in \{0, 1\}^S$ and a finite subset A of S , let

$$H(\eta, A) = \prod_{x \in A} [1 - \eta(x)] = \begin{cases} 1 & \text{if } \eta \equiv 0 \text{ on } A, \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$H(\eta_x, A) - H(\eta, A) = \begin{cases} H(\eta, A \setminus \{x\}) & \text{if } x \in A \text{ and } \eta(x) = 1, \\ -H(\eta, A) & \text{if } x \in A \text{ and } \eta(x) = 0, \\ 0 & \text{if } x \notin A, \end{cases}$$

so

$$\begin{aligned} \Omega H(\cdot, A)(\eta) &= \sum_{\eta(x)=1} [H(\eta_x, A) - H(\eta, A)] \\ &\quad + \lambda \sum_{\substack{\eta(x)=0, \eta(y)=1 \\ x \sim y}} [H(\eta_x, A) - H(\eta, A)] \\ &= \sum_{x \in A, \eta(x)=1} H(\eta, A \setminus \{x\}) - \lambda \sum_{\substack{\eta(x)=0, \eta(y)=1 \\ x \sim y, x \in A}} H(\eta, A) \quad (9) \\ &= \sum_{x \in A} [H(\eta, A \setminus \{x\}) - H(\eta, A)] \\ &\quad + \lambda \sum_{\substack{x \sim y \\ x \in A}} [H(\eta, A \cup \{y\}) - H(\eta, A)]. \end{aligned}$$

Note that the right side of (9) is simply the generator of the process with transitions

$$\begin{aligned} A &\rightarrow A \setminus \{x\} && \text{at rate 1 if } x \in A, \\ A &\rightarrow A \cup \{y\} && \text{at rate } \lambda |\{x \in A : x \sim y\}| \text{ if } y \notin A \end{aligned} \quad (10)$$

applied to the function $H(\eta, \cdot)$. But this process, A_t , can be thought of as the contact process itself on configurations with finitely many infected sites,

with the identification $A = \{x : \eta(x) = 1\}$. Thus the contact process is self-dual, and (6) becomes

$$P^\eta(\eta_t \equiv 0 \text{ on } A) = P^A(\eta \equiv 0 \text{ on } A_t). \quad (11)$$

One might think that is is not useful, since it simply relates the contact process to itself. However, (11) has important implications. For example, take $\eta \equiv 1$ in (11), and pass to the limit as $t \uparrow \infty$ to obtain

$$\nu\{\eta : \eta \equiv 0 \text{ on } A\} = \lim_{t \rightarrow \infty} P^A(A_t = \emptyset) = P^A(A_t = \emptyset \text{ eventually}),$$

where ν is the upper invariant measure. In particular, $\nu = \delta_0$ if and only if the “finite” contact process A_t is absorbed at the empty set a.s. for any initial state A . This property might appear to be obvious, but it does not hold for some relatives of the contact process, such as critical, reversible nearest particle systems. (See Chapter VII of Liggett (1985).)

1.3c. Duality for the linear voter model. This time, for $\eta \in \{0, 1\}^S$ and finite $A \subset S$, let

$$H(\eta, A) = \prod_{x \in A} \eta(x) = \begin{cases} 1 & \text{if } \eta \equiv 1 \text{ on } A, \\ 0 & \text{otherwise,} \end{cases} \quad (12)$$

let Ω be the generator of the linear voter model, and write

$$\begin{aligned} \Omega H(\cdot, A)(\eta) &= \sum_{x, y: \eta(x) \neq \eta(y)} p(x, y) [H(\eta_x, A) - H(\eta, A)] \\ &= \sum_{x \in A, y \in S} p(x, y) H(\eta, A \setminus \{x\}) [1 - 2\eta(x)] \{ \eta(x)[1 - \eta(y)] + \eta(y)[1 - \eta(x)] \} \\ &= \sum_{x \in A, y \in S} p(x, y) H(\eta, A \setminus \{x\}) [\eta(y) - \eta(x)] \\ &= \sum_{x \in A, y \in S} p(x, y) [H(\eta, (A \setminus \{x\}) \cup \{y\}) - H(\eta, A)]. \end{aligned}$$

The right side above can be seen as the result of applying to $H(\eta, A)$, as a function of A , the generator of the process in which points in A_t move according to the Markov chain with transition probabilities $p(x, y)$, with the proviso that a point that moves to an occupied site coalesces with the point at that site. Thus A_t is called the *coalescing Markov chain* process. The duality relation (6) becomes

$$P^\eta(\eta_t \equiv 1 \text{ on } A) = P^A(\eta \equiv 1 \text{ on } A_t). \quad (13)$$

This will play an essential role in our analysis of the linear voter model in the next chapter.

3d. Duality for the symmetric exclusion process. Again take $H(\eta, A)$ as in (12), and let Ω be the generator of the exclusion process. Let

$$A_{x,y} = \begin{cases} A & \text{if } x, y \in A \text{ or } x, y \notin A, \\ A \setminus \{x\} \cup \{y\} & \text{if } x \in A, y \notin A, \\ A \setminus \{y\} \cup \{x\} & \text{if } y \in A, x \notin A, \end{cases}$$

and write

$$\begin{aligned} \Omega H(\cdot, A)(\eta) &= \sum_{x,y \in S} \eta(x)[1 - \eta(y)]p(x,y) [H(\eta_{x,y}, A) - H(\eta, A)] \\ &= \sum_{x,y \in S} \eta(x)[1 - \eta(y)]p(x,y) [H(\eta, A_{x,y}) - H(\eta, A)] \\ &= \sum_{x \notin A, y \in A} p(x,y)[1 - \eta(y)]H(\eta, A_{x,y}) - \sum_{x \in A, y \notin A} p(x,y)[1 - \eta(y)]H(\eta, A) \\ &= \sum_{x \in A, y \notin A} \{p(y,x)H(\eta, A_{x,y}) - p(x,y)H(\eta, A) \\ &\quad + [p(y,x) - p(x,y)]H(\eta, A \cup \{y\})\}. \end{aligned}$$

In order to get a duality statement, we need to recognize the right side above as the generator of a Markov process, applied to $H(\eta, A)$ as a function of A . This is easy to do if $p(x,y) = p(y,x)$ for all x,y , and in this case, we conclude that

$$\Omega H(\cdot, A)(\eta) = \sum_{x \in A, y \notin A} p(x,y) [H(\eta, A_{x,y}) - H(\eta, A)].$$

Thus we see that in the symmetric case, the exclusion process is self-dual, and hence

$$P^\eta(\eta_t \equiv 1 \text{ on } A) = P^A(\eta \equiv 1 \text{ on } A_t), \quad (14)$$

where A_t is the finite version of the exclusion process. We will use (14) extensively in Chapter 3. The lack of duality in the asymmetric case makes the process more challenging to analyze, but we are rewarded with more interesting and diverse behavior. We will address the asymmetric exclusion process in Chapters 4 and 5.

Considering the nature of the duals in the three cases, we see that for the symmetric exclusion process, the size of the dual does not change with time. By contrast, the dual of the linear voter model has a cardinality that can decrease with time, while in the case of the contact process, the cardinality of the dual can both increase and decrease. These differences have a large effect on how the duality relation is exploited.

2 Linear Voter Models

Linear voter models were introduced in Example 2 of Chapter 1. In this chapter, we will explain how duality can be used to give a complete analysis of the limiting behavior of this process, mainly following the approach of Holley and Liggett (1975). In order to simplify the presentation, we will restrict our discussion to the case in which $S = Z^d$, and $p(x, y) = p(0, y - x)$ are the transition probabilities for an irreducible random walk on Z^d . The general case is treated in Chapter V of Liggett (1985).

We will discover that the behavior of the process depends heavily on the recurrence or transience properties of the symmetrized random walk $X(t) - Y(t)$, where $X(t)$ and $Y(t)$ are independent random walks with unit exponential jump times and transition probabilities $p(\cdot, \cdot)$. We begin with an informal description of the way in which this comes about. Suppose we want to determine the opinion $\eta_t(x)$ of the voter at x at a large time t . His opinion came from some other voter at x_1 at some earlier time t_1 . Continuing backwards in this way, we find that $\eta_t(x) = \eta_0(X(t))$ for some random $X(t)$. Note that the process $X(t)$ is simply a random walk with transition probabilities $p(\cdot, \cdot)$ and initial point $X(0) = x$.

We can do a similar backward construction for the evolution of the opinion $\eta_t(y)$ for $y \neq x$, and find that $\eta_t(y) = \eta_0(Y(t))$ for a random walk $Y(s)$ with $Y(0) = y$. However, $X(s)$ and $Y(s)$ are no longer independent – they are independent until the first time τ at which $X(\tau) = Y(\tau)$, but after that time, they evolve together: $Y(s) = X(s)$ for $s > \tau$. (These are the coalescing random walks that we found in our discussion of the voter model duality in Chapter 1.) Thus, $\eta_t(x)$ and $\eta_t(y)$ can agree for two different reasons: (i) $t > \tau$, or (ii) $t < \tau$ and $\eta_0(X(t)) = \eta_0(Y(t))$. If for the independent random walks, $X(s) - Y(s)$ is recurrent, then the coalescing random walks will agree eventually with probability 1, and hence $\eta_t(x) = \eta_t(y)$ with large probability for large t .

Remark. It might appear that this would show that for each x and y , $\eta_t(x) = \eta_t(y)$ from some time on. This is not necessarily the case, since changing the t in the argument changes the random walks $X(s)$ and $Y(s)$. An easy example is the nearest neighbor symmetric linear voter model in 1 dimension. If the initial configuration is $\cdots 1 1 1 0 0 0 \cdots$, for example, then the configuration at later times has the same form, with the boundary between 1's and 0's moving like a simple symmetric random walk. Therefore, every site changes opinion infinitely often.

We will now formalize the above argument using duality directly. Let A_t be the dual process of coalescing random walks. It will be convenient to define

$$g(A) = P^A(|A_t| < |A| \text{ for some } t > 0).$$

We will regard $g(A)$ as a measure of how spread apart the points in A are — $g(A)$ small corresponds to A being spread out.

Lemma 1. (a) If $A \subset B$, then $g(A) \leq g(B)$.

(b) For any nonempty A ,

$$g(A) \leq \sum_{B \subset A, |B|=2} g(B).$$

Proof. Given $A \subset B$, A_t and B_t with $A_0 = A$ and $B_0 = B$ can be coupled together in such a way that $A_t \subset B_t$ for all t . To do so, if $x \in A_t$ use the same exponential times and the same jump choices for x viewed as a point in A_t and for x viewed as a point in B_t . With this coupling $|A_t| < |A_0|$ implies $|B_t| < |B_0|$, and hence $g(A) \leq g(B)$. This proves part (a).

For part (b), let $X_1(t), \dots, X_{|A|}(t)$ be independent random walks with the distribution of $X(t)$, and initial states $\{X_1(0), \dots, X_{|A|}(0)\} = A$. Then

$$\begin{aligned} g(A) &= P(X_i(t) = X_j(t) \text{ for some } i \neq j, t > 0) \\ &\leq \sum_{1 \leq i < j \leq |A|} P(X_i(t) = X_j(t) \text{ for some } t > 0) \\ &= \sum_{1 \leq i < j \leq |A|} g(\{X_i(0), X_j(0)\}). \end{aligned}$$

Exercise. Write down the transition rates for the Markov chain (A_t, B_t) on $\{(A, B) : A \subset B\}$ used in the proof of part (a) of Lemma 1.

2.1 The recurrent case

In this section, we assume that for the independent random walks $X(t)$ and $Y(t)$, $X(t) - Y(t)$ is recurrent. As we will see, it follows that the process tends to a consensus in this case.

Lemma 2. *For every finite, nonempty $A \subset \mathbb{Z}^d$,*

$$P^A(|A_t| = 1 \text{ eventually}) = 1. \quad (1)$$

Proof. The recurrence assumption implies that $g(A) = 1$ if $|A| = 2$. In particular, (1) holds for all A of size 2. We will now argue by induction on the size of A . By Lemma 1(a), $g(A) = 1$ for all A with at least two points. Now take A with $|A| > 1$, and let $\tau = \inf\{t > 0 : |A_t| < |A|\}$, which is then finite a.s. By the strong Markov property,

$$P^A(|A_t| = 1 \text{ eventually}) = E^A P^{A_\tau}(|A_t| = 1 \text{ eventually}).$$

Since $|A_\tau| = |A| - 1$, if (1) holds for all A of size n , it holds for all A of size $n + 1$ (with $n \geq 2$).

Theorem 3. (a) *For every $\eta \in \{0, 1\}^S$ and every $x, y \in S$,*

$$\lim_{t \rightarrow \infty} P^\eta(\eta_t(x) \neq \eta_t(y)) = 0.$$

(b) $\mathcal{I}_e = \{\delta_0, \delta_1\}$.

(c) *If $\mu\{\eta : \eta(x) = 1\} = \lambda$ for all $x \in S$, then*

$$\lim_{t \rightarrow \infty} \mu S(t) = \lambda \delta_1 + (1 - \lambda) \delta_0.$$

Proof. For part (a), let $X(t)$ and $Y(t)$ be coalescing random walks starting at x and y respectively, and let τ be the coalescing time. By duality (equation (13) of Chapter 1),

$$P^\eta(\eta_t(x) \neq \eta_t(y)) = P(\eta(X(t)) \neq \eta(Y(t))) \leq P(\tau > t) \rightarrow 0,$$

since $\tau < \infty$ a.s.

For part (b), note that $\mu \in \mathcal{I}$ implies that

$$\mu\{\eta : \eta(x) \neq \eta(y)\} = P^\mu(\eta_t(x) \neq \eta_t(y)) = \int P^\eta(\eta_t(x) \neq \eta_t(y)) \mu(d\eta),$$

which tends to 0 as $t \rightarrow \infty$ by part (a). Therefore μ concentrates on the two constant configurations $\eta \equiv 0$ and $\eta \equiv 1$.

For part (c), use duality again to write for finite $A \neq \emptyset$

$$\begin{aligned} \mu S(t)\{\eta : \eta \equiv 1 \text{ on } A\} &= \int P^\eta(\eta_t \equiv 1 \text{ on } A)\mu(d\eta) = \int P^A(\eta \equiv 1 \text{ on } A_t)\mu(d\eta) \\ &= \int P^A(\eta \equiv 1 \text{ on } A_t, |A_t| > 1)\mu(d\eta) + \sum_y P^A(A_t = \{y\})\mu\{\eta : \eta(y) = 1\}. \end{aligned}$$

Therefore,

$$|\mu S(t)\{\eta : \eta \equiv 1 \text{ on } A\} - \lambda| \leq 2P^A(|A_t| > 1),$$

which tends to 0 as $t \rightarrow \infty$ by Lemma 2.

Exercise. Is it the case that the weak limit of $\mu S(t)$ exists for every initial distribution μ ?

2.2 The transient case

In this section, we assume that for the independent random walks $X(t)$ and $Y(t)$, $X(t) - Y(t)$ is transient. In this case, the behavior of the process η_t is quite different. We begin by showing how to construct a nontrivial stationary distribution μ_ρ for each density $0 \leq \rho \leq 1$. Let ν_ρ be the homogeneous product measure on $\{0, 1\}^S$ with density ρ :

$$\nu_\rho\{\eta : \eta \equiv 1 \text{ on } A\} = \rho^{|A|}$$

for each finite $A \subset S$.

By duality,

$$\begin{aligned} \nu_\rho S(t)\{\eta : \eta \equiv 1 \text{ on } A\} &= \int P^\eta(\eta_t \equiv 1 \text{ on } A)d\nu_\rho = \int P^A(\eta \equiv 1 \text{ on } A_t)d\nu_\rho \\ &= \sum_B P^A(A_t = B) \int \prod_{x \in B} \eta(x) d\nu_\rho = E^A \rho^{|A_t|}. \end{aligned} \quad (2)$$

Since $|A_t|$ is nonincreasing in t , it follows that the limit of $\nu_\rho S(t)\{\eta : \eta \equiv 1 \text{ on } A\}$ exists for every finite $A \subset S$, and hence the weak limit

$$\mu_\rho = \lim_{t \rightarrow \infty} \nu_\rho S(t) \quad (3)$$

exists. Since μ_ρ is the limit of the distribution of the process at time t as $t \uparrow \infty$, it is automatically stationary, which we see as follows:

$$\begin{aligned} \int f d\mu_\rho S(t) &= \int S(t) f d\mu_\rho = \lim_{s \rightarrow \infty} \int S(t) f d\nu_\rho S(s) \\ &= \lim_{s \rightarrow \infty} \int f d\nu_\rho S(t+s) = \int f d\mu_\rho. \end{aligned}$$

Before going further, we need to develop some properties of the function g . Recall that $g(A+x) = g(A)$ for all $x \in S$.

Lemma 4. (a) $\lim_{x \rightarrow \infty} g(\{0, x\}) = 0$.

(b) If $X_1(t), \dots, X_n(t)$ are independent copies of the random walk $X(t)$, then

$$g(\{X_1(t), \dots, X_n(t)\}) \rightarrow 0 \text{ a.s.}$$

as $t \rightarrow \infty$.

(c) $g(A_t) \rightarrow 0$ a.s. for any initial set $A \neq \emptyset$.

Proof. For part (a), let $Z(t) = X(t) - Y(t)$, which is a symmetric random walk. Using this symmetry, the Markov property, and the Schwarz inequality, write

$$\begin{aligned} P^x(Z(2t) = y) &= \sum_z P^x(Z(t) = z)P^z(Z(t) = y) \\ &\leq \left[\sum_z [P^x(Z(t) = z)]^2 \right]^{1/2} \left[\sum_z [P^z(Z(t) = y)]^2 \right]^{1/2} \quad (4) \\ &= [P^x(Z(2t) = x)P^y(Z(2t) = y)]^{1/2} = P^0(Z(2t) = 0). \end{aligned}$$

Now let

$$G(x, y) = \int_0^\infty P^x(Z(t) = y) dt$$

be the Green function for the random walk $Z(t)$. By the strong Markov property applied at the hitting time of y ,

$$G(x, y) = P^x(Z(t) = y \text{ for some } t > 0)G(y, y).$$

Similarly, for $T > 0$,

$$\int_T^\infty P^x(Z(t) = 0) dt = P^x(Z(t) = 0 \text{ for some } t > T)G(0, 0),$$

which by (4), takes its largest value at $x = 0$. Therefore,

$$\begin{aligned} g(\{0, x\}) &= P^x(Z(t) = 0 \text{ for some } t \geq 0) \\ &\leq P^x(Z(t) = 0 \text{ for some } t \leq T) + P^x(Z(t) = 0 \text{ for some } t > T) \\ &\leq P^x(Z(t) = 0 \text{ for some } t \leq T) + P^0(Z(t) = 0 \text{ for some } t > T). \end{aligned}$$

Since $Z(t)$ is transient, the second term on the right above can be made small by taking T large. Then the first term on the right can be made small by taking $|x|$ large.

For part (b), use part (a) and Lemma 1(b), together with the fact that $X_i(t) - X_j(t) \rightarrow \infty$ a.s. as $t \rightarrow \infty$ for $i \neq j$ by the transience assumption. Part (c) follows from part (b) and Lemma 1(a), since A_t and $(X_1(t), \dots, X_n(t))$ can be coupled together in such a way that the containment

$$A_t \subset \{X_1(t), \dots, X_n(t)\}$$

is maintained.

We are now in a position to develop some properties of the stationary distributions μ_ρ we have constructed. For the statement of part (b) below, recall that a shift invariant probability measure on $\{0, 1\}^{\mathbb{Z}^d}$ is said to be ergodic if it assigns probability 0 or 1 to every shift invariant subset of $\{0, 1\}^{\mathbb{Z}^d}$. These turn out to be the extremal shift invariant distributions.

Theorem 5. *For each $0 \leq \rho \leq 1$, the measure μ_ρ defined in (3) has the following properties:*

(a) *The coordinates are asymptotically independent in the sense that*

$$|\mu_\rho\{\eta : \eta \equiv 1 \text{ on } A\} - \rho^{|A|}| \leq g(A)$$

for all A ,

(b) μ_ρ *is translation invariant and spatially ergodic,*

(c) $\mu_\rho\{\eta : \eta(x) = 1\} = \rho$,

and

(d)

$$\text{Cov}_{\mu_\rho}[\eta(x), \eta(y)] = \rho(1 - \rho) \frac{G(x, y)}{G(0, 0)},$$

where G is the Green function of $Z(t)$.

Proof. Passing to the limit in (2), we see that

$$\mu_\rho\{\eta : \eta \equiv 1 \text{ on } A\} = E\rho^{|A_\infty|}. \quad (5)$$

(Of course, this is an abuse of notation since the limit of A_t does not exist; by $|A_\infty|$, we mean the limit of $|A_t|$, which does exist by monotonicity.) Therefore, part (a) follows from

$$|E^A \rho^{|A_\infty|} - \rho^{|A|}| = E^A(\rho^{|A_\infty|} - \rho^{|A|}, |A_\infty| < |A|) \leq g(A).$$

The translation invariance statement in part (b) is immediate, since $\nu_\rho S(t)$ is translation invariant for every t . For the ergodicity statement, let A_t, A_t^\dagger

and A_t^2 be copies of the coalescing random walks process that are coupled so that

(a) A_t^1 and A_t^2 are independent,

and

(b) $A_t = A_t^1 \cup A_t^2$ for $t \leq \tau = \inf\{s > 0 : A_s^1 \cap A_s^2 \neq \emptyset\}$.

Given disjoint sets A^1 and A^2 , start the coupled chains with initial states $A_0^1 = A^1$, $A_0^2 = A^2$, and $A_0 = A^1 \cup A^2$. Since $|A_\infty| = |A_\infty^1| + |A_\infty^2|$ on the event $\{\tau = \infty\}$, (5) implies that

$$\begin{aligned} & |\mu_\rho\{\eta : \eta \equiv 1 \text{ on } A^1 \cup A^2\} - \mu_\rho\{\eta : \eta \equiv 1 \text{ on } A^1\}\mu_\rho\{\eta : \eta \equiv 1 \text{ on } A^2\}| \\ &= |E[\rho^{|A_\infty|} - \rho^{|A_\infty^1|+|A_\infty^2|}]| \leq P(\tau < \infty) \leq \sum_{x \in A^1, y \in A^2} g(\{x, y\}). \end{aligned}$$

Replacing A^2 by $A^2 + z$, it then follows from Lemma 4(a) that

$$\lim_{z \rightarrow \infty} \mu_\rho\{\eta : \eta \equiv 1 \text{ on } A^1 \cup (A^2 + z)\} = \mu_\rho\{\eta : \eta \equiv 1 \text{ on } A^1\}\mu_\rho\{\eta : \eta \equiv 1 \text{ on } A^2\},$$

and this implies that μ_ρ is spatially ergodic.

Part (c) is an immediate consequence of (5). Part (d) also follows from (5) as follows:

$$\begin{aligned} \text{Cov}_\mu[\eta(x), \eta(y)] &= \int \eta(x)\eta(y)d\mu - \rho^2 = E^{\{x,y\}}\rho^{|A_\infty|} - \rho^2 \\ &= \rho(1 - \rho)P^{\{x,y\}}(|A_\infty| = 1) \\ &= \rho(1 - \rho)P^{x-y}(Z(t) = 0 \text{ for some } t \geq 0) \\ &= \rho(1 - \rho)\frac{G(x-y, 0)}{G(0, 0)} = \rho(1 - \rho)\frac{G(x, y)}{G(0, 0)}. \end{aligned}$$

Exercise. Write down the transition rates for the Markov chain (A_t^1, A_t^2, A_t) used in the proof of part (b) of Theorem 5.

Next we check that we have captured all the extremal stationary distributions via the construction in Theorem 5.

Theorem 6. (a) \mathcal{I} is the closed convex hull of $\{\mu_\rho : 0 \leq \rho \leq 1\}$.

(b) $\mathcal{I}_e = \{\mu_\rho : 0 \leq \rho \leq 1\}$.

Proof. One containment is clear, since $\mu_\rho \in \mathcal{I}$ and \mathcal{I} is convex and closed. For the other, take $\mu \in \mathcal{I}$. Then by duality,

$$\begin{aligned} \mu\{\eta : \eta \equiv 1 \text{ on } A\} &= \mu S(t)\{\eta : \eta \equiv 1 \text{ on } A\} = \int P^\eta(\eta_t \equiv 1 \text{ on } A) d\mu \\ &= \int P^A(\eta \equiv 1 \text{ on } A_t) d\mu = \sum_B P^A(A_t = B) \mu\{\eta : \eta \equiv 1 \text{ on } B\}. \end{aligned} \quad (6)$$

To simplify the notation, let $h(A) = \mu\{\eta : \eta \equiv 1 \text{ on } A\}$,

$$V_t f(A) = E^A f(A_t), \text{ and } U_t f(x_1, \dots, x_n) = E^{x_1, \dots, x_n} f(X_1(t), \dots, X_n(t)),$$

where $X_1(t), \dots, X_n(t)$ are independent versions of our basic random walk. Then (6) can be written as $h = V_t h$. When applied to a function of a set A instead of a vector (x_1, \dots, x_n) , $U_t f(A)$ is interpreted as $E^{x_1, \dots, x_n} f(\{X_1(t), \dots, X_n(t)\})$, where $A = \{x_1, \dots, x_n\}$.

Using the coupling between A_t and $\{X_1(t), \dots, X_n(t)\}$ that we used in the proof of Lemma 4(b), we see that for any function f satisfying $|f(A)| \leq 1$ for all A ,

$$|V_t f(A) - U_t f(A)| \leq g(A). \quad (7)$$

Applying this to h gives

$$|h(A) - U_t h(A)| \leq g(A). \quad (8)$$

Therefore,

$$|U_s h(A) - U_{t+s} h(A)| \leq U_s g(A).$$

The right side tends to zero as $s \rightarrow \infty$ by Lemma 4(b), so $\lim_{s \rightarrow \infty} U_s h$ exists, and is harmonic for the irreducible random walk $(X_1(t), \dots, X_n(t))$ on S^n . Such harmonic functions are constant (see the second remark following the proof of Theorem 2 of Chapter 1), so we conclude that there are constants c_n so that

$$\lim_{s \rightarrow \infty} U_s h(A) = c_{|A|}$$

for every A . Of course, c_n depends on h , and therefore on the stationary measure μ . Passing to the limit in (8), we conclude that

$$|h(A) - c_{|A|}| \leq g(A). \quad (9)$$

Note that by Lemmas 1(b) and 4(a), $g(A)$ is small if A is spread out. Therefore, (9) says that if A is spread out and has cardinality n , then $h(A)$

is approximately equal to c_n . We are trying to show that μ is a mixture of the μ_ρ 's. By Theorem 5(a), $\mu_\rho\{\eta : \eta \equiv 1 \text{ on } A\}$ and $\nu_\rho\{\eta : \eta \equiv 1 \text{ on } A\}$ are approximately equal if A is spread out. Therefore, we would expect c_n to be a mixture of $\nu_\rho\{\eta : \eta \equiv 1 \text{ on } A\} = \rho^n$ for $\rho \in [0, 1]$.

Now we will carry out the program described in the last paragraph. To do so we need to show that $\{c_n\}$ is a moment sequence, i.e., there exists a probability measure γ on $[0, 1]$ so that

$$c_n = \int_0^1 \rho^n \gamma(d\rho).$$

A necessary and sufficient condition for this is that

$$\sum_{k=0}^n \binom{n}{k} (-1)^k c_{k+m} \geq 0 \quad (10)$$

for all nonnegative integers n, m . (See Theorem 2 of Section VII.3 of Feller (1966).) Note that the necessity is clear from

$$\sum_{k=0}^n \binom{n}{k} (-1)^k \int_0^1 \rho^{k+m} \gamma(d\rho) = \int_0^1 \rho^m (1 - \rho)^n \gamma(d\rho).$$

To check (10), fix m, n , and let A_i be a sequence of sets of size $m + n$ so that $g(A_i) \rightarrow 0$. Such a sequence exists by Lemma 4(b). Write $A_i = B_i \cup C_i$, where $|B_i| = m$ and $|C_i| = n$. Note that by Lemma 1(a), $g(B_i) \rightarrow 0$ and $g(C_i) \rightarrow 0$ as well, since these are subsets of A_i . By inclusion-exclusion,

$$\mu\{\eta : \eta \equiv 1 \text{ on } B_i, \eta \equiv 0 \text{ on } C_i\} = \sum_{F \subset C_i} (-1)^{|F|} h(B_i \cup F).$$

Applying (9), it then follows that

$$\lim_{i \rightarrow \infty} \mu\{\eta : \eta \equiv 1 \text{ on } B_i, \eta \equiv 0 \text{ on } C_i\} = \sum_{k=0}^n \binom{n}{k} (-1)^k c_{k+m},$$

which gives (10).

Now define

$$\mu^* = \int_0^1 \mu_\rho \gamma(d\rho) \quad \text{and} \quad h^*(A) = \mu^*\{\eta : \eta \equiv 1 \text{ on } A\}.$$

Since $\mu^* \in \mathcal{I}$, (6) can be applied to μ^* to conclude that $h^* = V_t h^*$. By Theorem 5(a),

$$|h^*(A) - c_{|A}| \leq g(A),$$

and comparing with (9),

$$|h^*(A) - h(A)| \leq 2g(A).$$

Applying V_t to this and using the harmonicity for V_t of both h and h^* gives

$$|h^*(A) - h(A)| \leq 2V_t g(A).$$

Letting $t \rightarrow \infty$ and using Lemma 4(c) gives $h^* \equiv h$, and hence $\mu^* = \mu$. Thus

$$\mu = \int_0^1 \mu_\rho \gamma(d\rho), \tag{11}$$

which completes the proof of part (a).

To prove part (b), note first that $\mu_\rho \in \mathcal{I}_e$ for each ρ by Theorem 5(b), since the ergodic measures are extremal in the class of all translation invariant measures. So,

$$\{\mu_\rho : 0 \leq \rho \leq 1\} \subset \mathcal{I}_e.$$

For the reverse containment, take $\mu \in \mathcal{I}_e$. By part (a), μ is a mixture of μ_ρ 's. Since μ is extremal, it must be one of the μ_ρ 's.

Exercise. Use the fact that (10) is a necessary and sufficient condition for c_n to be a moment sequence to prove de Finetti's Theorem, which states that every exchangeable measure μ on $\{0, 1\}^S$, where S is infinite, can be written as a mixture of the homogeneous product measures ν_ρ . (See Section VII.4 of Feller (1966). An exchangeable measure is one for which $\mu\{\eta : \eta \equiv 1 \text{ on } A\}$ depends on A only through its cardinality.) We will use de Finetti's Theorem in Chapters 3 and 4.

Next is a convergence theorem for translation invariant initial distributions.

Theorem 7. *Suppose μ is translation invariant and spatially ergodic. Then*

$$\lim_{t \rightarrow \infty} \mu S(t) = \mu_\rho,$$

where $\rho = \mu\{\eta : \eta(0) = 1\}$.

Proof. Let $h(A) = \mu\{\eta : \eta \equiv 1 \text{ on } A\}$, $h_1(A) = \mu_\rho\{\eta : \eta \equiv 1 \text{ on } A\}$, and

$$h_2(A) = \nu_\rho\{\eta : \eta \equiv 1 \text{ on } A\} = \rho^{|A|}.$$

Define V_t and U_t as in the proof of Theorem 6. By duality, we need to show that

$$V_t h(A) \rightarrow h_1(A) \tag{12}$$

for each A . We will prove instead that

$$U_t h(x) \rightarrow \rho^n \tag{13}$$

for each $x \in S^n$. To see that (13) implies (12), use (7) to write

$$\begin{aligned} |V_{t+s} h(A) - h_1(A)| &\leq |V_s(V_t h - U_t h)(A)| + |V_s U_t h(A) - h_1(A)| \\ &\leq V_s g(A) + |V_s U_t h(A) - h_1(A)| \end{aligned}$$

which implies by (13) that

$$\limsup_{t \rightarrow \infty} |V_t h(A) - h_1(A)| \leq V_s g(A) + |V_s h_2(A) - h_1(A)|. \tag{14}$$

Since $V_s h_2 \rightarrow h_1$ as $s \rightarrow \infty$ by (2) and (3), and $V_s g \rightarrow 0$ by Lemma 4(c), the right side of (14) tends to zero, and hence (12) holds.

Now we will check (13) using some elementary Fourier analysis. Since the covariance of the random field with distribution μ is positive definite, Böchner's theorem implies that there is a measure γ on $[-\pi, \pi]^d$ so that

$$\mu\{\eta : \eta(x) = 1, \eta(y) = 1\} - \rho^2 = \int e^{i\langle y-x, \theta \rangle} \gamma(d\theta).$$

($\langle \cdot, \cdot \rangle$ is the usual R^d inner product.) Let

$$\phi(\theta) = \sum_x p(0, x) e^{i\langle x, \theta \rangle}$$

be the characteristic function of the random walk jumps. Since the random walk is irreducible, $|\phi(\theta)| = 1$ if and only if $\theta = 0$. Note that

$$E^x e^{i\langle X_t, \theta \rangle} = e^{i\langle x, \theta \rangle} e^{-t[1-\phi(\theta)]}.$$

Now define

$$W_t(x, \eta) = \sum_y P^x(X(t) = y) \eta(y).$$

Then

$$\int [W_t(x, \eta) - W_s(x, \eta)]^2 \mu(d\eta) = \int \left| e^{-t[1-\phi(\theta)]} - e^{-s[1-\phi(\theta)]} \right|^2 \gamma(d\theta),$$

which tends to zero as $s, t \rightarrow \infty$. Therefore

$$W(x, \eta) = \lim_{t \rightarrow \infty} W_t(x, \eta)$$

exists in $L^2(\mu)$ for each x .

By the definition of $W_t(x, \eta)$ and the Chapman-Kolmogorov equations,

$$W_{t+s}(x, \eta) = \sum_y P^x(X(t) = y) W_s(y, \eta),$$

so that by passing to the limit as $s \rightarrow \infty$, we have

$$W(x, \eta) = \sum_y P^x(X(t) = y) W(y, \eta) \quad \text{a.s.}(\mu).$$

Since all bounded harmonic functions for the random walk are constant, we conclude that

$$W(x, \eta) = W(0, \eta) \quad \text{a.s.}(\mu) \tag{15}$$

for each x . Using the definition of $W_t(x, \eta)$ again, we have

$$W_t(x + u, \eta) = W_t(x, \tau_u \eta),$$

where τ_u is the shift: $\tau_u \eta(y) = \eta(y + u)$. Therefore, $W(x + u, \eta) = W(x, \tau_u \eta)$ a.s., so by (15), $W(0, \eta)$ is an a.s. invariant random variable. Since μ is ergodic, $W(0, \eta)$ is a.s. constant, and since it has mean ρ , we conclude that

$$W_t(x, \eta) \rightarrow \rho$$

in $L^2(\mu)$ as $t \rightarrow \infty$ for every x . But

$$U_t h(\{x_1, \dots, x_n\}) = \int \prod_{i=1}^n W_t(x_i, \eta) d\mu \rightarrow \rho^n$$

as $t \rightarrow \infty$. This verifies (13) as required.

3 Symmetric Exclusion Processes

Exclusion processes were introduced in Example 3 of Chapter 1. In this chapter, we will explain how coupling and duality can be used to give a complete analysis of the limiting behavior of this process in the symmetric case. The main reason for the symmetry assumption is that there is no useful duality for asymmetric systems. As in the previous chapter, we will restrict our discussion to the translation invariant case. Thus we assume that $S = Z^d$, and $p(x, y) = p(0, y - x) = p(y, x)$ are the transition probabilities for a symmetric irreducible random walk on Z^d . The general case is treated in Chapter VIII of Liggett (1985) – the main results of that chapter are stated without proof in Section 3 of this chapter.

The principal result of this chapter is the following.

Theorem 1. (a) $\mathcal{I}_e = \{\nu_\rho : 0 \leq \rho \leq 1\}$.

(b) If μ is translation and spatially ergodic, then

$$\lim_{t \rightarrow \infty} \mu S(t) = \nu_\rho,$$

where $\rho = \mu\{\eta : \eta(0) = 1\}$.

We will only prove part (a), since the proof of part (b) is similar to the proof of Theorem 7 of Chapter 2. Using duality, part (a) of Theorem 1 follows from the following statement.

Theorem 2. If f is a bounded harmonic function for the finite exclusion process A_t , then f is constant on $\{A : |A| = n\}$ for each $n \geq 1$.

Proof of Theorem 1(a). Recall that a probability measure μ on $\{0, 1\}^S$ is said to be exchangeable if

$$\mu\{\eta : \eta \equiv 1 \text{ on } A\}$$

depends on A only through its cardinality. By de Finetti's Theorem (see the exercise preceding Theorem 7 in Chapter 2), if S is infinite, then every exchangeable measure is a mixture of the homogeneous product measures ν_ρ . Therefore, Theorem 1(a) is equivalent to the statement that \mathcal{I} agrees with the set of exchangeable probability measures. To check this, use duality (equation (14) of Chapter 1) to write

$$\begin{aligned} \mu S(t)\{\eta : \eta \equiv 1 \text{ on } A\} &= \int P^\eta(\eta_t \equiv 1 \text{ on } A) d\mu = \int P^A(\eta \equiv 1 \text{ on } A_t) d\mu \\ &= \sum_B P^A(A_t = B) \mu\{\eta : \eta \equiv 1 \text{ on } B\}. \end{aligned} \quad (1)$$

Since the cardinality of A_t does not change with t , this shows that every exchangeable measure is stationary. Conversely, suppose μ is stationary. Then $\mu S(t) = \mu$ for each t , so (1) implies that the function $f(A) = \mu\{\eta : \eta \equiv 1 \text{ on } A\}$ is harmonic for A_t . By Theorem 2, f depends on A only through its cardinality, so μ is exchangeable.

The proof of Theorem 2 depends very much on whether the random walk $X(t)$ with transition probabilities $p(x, y)$ is transient or recurrent. We will explain the proof in these two cases in the next two sections.

3.1 The recurrent case

To prove Theorem 2 in this case, it suffices to construct a coupling of two copies A_t and B_t of the finite exclusion process for any initial states A_0, B_0 of cardinality n so that $|A_0 \cap B_0| = n - 1$ with the following property:

$$P(A_t = B_t \text{ for all } t \text{ beyond some point}) = 1. \quad (2)$$

(A coupling with property (2) is said to be successful.) To see that this is sufficient, we argue as in the proof of Theorem 2 of Chapter 1. Suppose that f is bounded and harmonic for the finite exclusion process A_t . Then using the coupling,

$$|f(A_0) - f(B_0)| = |Ef(A_t) - Ef(B_t)| \leq E|f(A_t) - f(B_t)| \leq 2\|f\|P(A_t \neq B_t).$$

The right side of this expression tends to zero as $t \rightarrow \infty$ by (2), and hence $f(A_0) = f(B_0)$. Since any two sets of the same cardinality have the property that one can be transformed into the other by successively moving one site at a time, it follows that $f(A) = f(B)$ for any two sets A, B with the same cardinality.

Next, we must construct a coupling with property (2). If $n = 1$, this is very simple: simply let A_t and B_t evolve independently until they meet, and make them evolve together thereafter. To see that they will eventually meet, write $A_t = \{X(t)\}$ and $B_t = \{Y(t)\}$. Then $Z(t) = X(t) - Y(t)$ is a random walk with exponential holding times of rate 2 and transition probabilities $p(x, y)$, until it first hits 0. ($Z(t)$ moves from u to v if either $X(t)$ has an increment of $v - u$ or $Y(t)$ has an increment of $u - v$; the rate of this is $p(0, v - u) + p(0, u - v)$. Now use symmetry.) Therefore $Z(t)$ is recurrent, and hence it will hit 0 eventually.

For general n , the proof is similar, but relies on a crucial observation. After A_t and B_t agree, they will move together. Until that time, call it τ ,

they will always satisfy $|A_t \cap B_t| = n-1$. So, for $t < \tau$, write $A_t = C_t \cup \{X(t)\}$ and $B_t = C_t \cup \{Y(t)\}$, where $C_t = A_t \cap B_t$. Before τ , the Markov process $(C_t, X(t), Y(t))$ has the following transitions from state (C, x, y) . In each case, $u \in C$ and $v \notin C \cup \{x, y\}$. The transitions are then to

$$\begin{array}{lll}
 (C_{u,v}, x, y) & \text{at rate} & p(u, v), \\
 (C, v, y) & \text{at rate} & p(x, v), \\
 (C, x, v) & \text{at rate} & p(y, v), \\
 (C, x, x) & \text{at rate} & p(y, x), \\
 (C, y, y) & \text{at rate} & p(x, y), \\
 (C_{u,x}, u, y) & \text{at rate} & p(u, x), \\
 (C_{u,y}, x, u) & \text{at rate} & p(u, y).
 \end{array}$$

Note that the marginal processes have the following rates: The pair $(X(t), Y(t))$ has the transition rates for two independent random walks with transition probabilities $p(\cdot, \cdot)$ and C_t evolves like an exclusion process, until the first time τ that $X(t) = Y(t)$. This again uses the symmetry of $p(\cdot, \cdot)$. To see this, write a configuration at some time as

$$\begin{array}{cccccc}
 A_t : & * & * & * & * & * \\
 B_t : & * & * & * & * & * \\
 & & C_t & & X(t) & Y(t)
 \end{array}$$

If $u \in C_t$ and $x = X(t)$, then $X(t)$ moves from x to u at rate $p(u, x)$, when it should make this transition at rate $p(x, u)$. But by symmetry, these are the same. Now by the recurrence assumption, $X(t)$ and $Y(t)$ will meet eventually, and hence $P(\tau < \infty) = 1$, as required.

3.2 The transient case

The proof in case $n = 1$ is immediate, since irreducible random walks have no nonconstant bounded harmonic functions. So, fix $n > 1$, and let

$$T = \{x = (x_1, \dots, x_n) \in S^n : x_i \neq x_j \text{ for all } i \neq j\}.$$

Then we can regard the exclusion process with n particles as a Markov chain on T . Let V_t be its semigroup, and let U_t be the semigroup for the system of n independent random walks $X(t) = (X_1(t), \dots, X_n(t))$. We need to prove that $V_t f = f$ implies f is constant if f is a bounded symmetric function

on T . Since U_t is the semigroup for a random walk on S^n , we know that $U_t f = f$ implies f is constant if f is a bounded function on S^n . Thus we need to compare U_t and V_t in some way. To do so, define g on S^n by

$$g(x) = P^x(X(t) \notin T \text{ for some } t \geq 0).$$

This function will play a role similar to that of the g in Chapter 2.

By coupling the exclusion process and the independent random walk system so that they agree until the first time the latter process hits T , we see that if $0 \leq f \leq 1$, then

$$|V_t f(x) - U_t f(x)| \leq g(x), \quad x \in T. \quad (3)$$

By Lemma 4(b) of Chapter 2 and the fact that $P^x(X(t) \in T) \rightarrow 1$,

$$\lim_{t \rightarrow \infty} U_t g(x) = 0, \quad x \in S^n. \quad (4)$$

($P^x(X(t) \in T) \rightarrow 1$ is needed here because Lemma 4(b) refers to the set

$$\{X_1(t), \dots, X_n(t)\}$$

rather than the vector $(X_1(t), \dots, X_n(t))$. Thus, for example, $g(\{u, u\}) = 0$, but $g(u, u) = 1$.)

We will argue in the following way. Suppose f is a symmetric function on T satisfying $0 \leq f \leq 1$ and $V_t f = f$ for all $t \geq 0$. Extend f to S^n by setting $f = 0$ on $S^n \setminus T$. By (3) and the fact that $g = 1$ on $S^n \setminus T$,

$$|f(x) - U_t f(x)| \leq g(x), \quad x \in S^n. \quad (5)$$

Applying U_s to this inequality gives

$$|U_s f(x) - U_{s+t} f(x)| \leq U_s g(x), \quad x \in S^n.$$

By (4), $U_s f$ has a limit as $s \rightarrow \infty$. Since this limit is harmonic for U_t , it is a constant; call it C . We conclude that

$$\lim_{t \rightarrow \infty} U_t f(x) = C, \quad x \in S^n.$$

Passing to the limit in (5) gives

$$|f(x) - C| \leq g(x), \quad x \in S^n.$$

Applying V_t to this and using $V_t f = f$ again leads to

$$|f(x) - C| = |V_t f(x) - C| \leq V_t g(x), \quad x \in T. \quad (6)$$

To complete the proof that f is constant on T , it then suffices to show that the right side of (6) tends to zero as $t \rightarrow \infty$.

In view of (4), it would be enough to show that

$$V_t g(x) \leq U_t g(x), \quad x \in T. \quad (7)$$

The intuition behind this inequality is the following: $g(x)$ is small if the coordinates of x are widely separated. Since the difference between the independent system and the interacting system is that the latter evolution keeps the coordinates of the process apart, one might expect that the coordinates are generally further apart in the interacting system than in the independent system. This relation of being “further apart” cannot be maintained in a pathwise sense, but (7) would say that the intuition is correct in an expected value sense.

It is easier to consider this issue for another function on S^n defined by

$$g^*(x) = \sum_{1 \leq i < j \leq n} \Delta(x_i - x_j),$$

where

$$\Delta(u) = P^u(Z(t) = 0 \text{ for some } t > 0).$$

Here $Z(t)$ is the underlying random walk on S . Since

$$g(x) \leq g^*(x) \leq \binom{n}{2} g(x), \quad x \in S^n,$$

(4) implies that

$$\lim_{t \rightarrow \infty} U_t g^*(x) = 0, \quad x \in S^n, \quad (8)$$

and we need only prove

$$V_t g^*(x) \leq U_t g^*(x), \quad x \in T. \quad (9)$$

We begin the proof of (9) by writing the integration by parts formula

$$U_t - V_t = \int_0^t V_{t-s}(U - V)U_s ds, \quad (10)$$

where U and V are the generators of U_t and V_t respectively. (To check this, integrate the following identity from 0 to t :

$$\frac{d}{ds} V_{t-s} U_s = V_{t-s} U U_s - V_{t-s} V U_s,$$

recalling that, while the U 's and V 's do not commute with each other, each semigroup commutes with its generator.) It follows from (10) that it will be enough to show that $(U - V)U_t g^* \geq 0$ on T , since V_t maps functions that are nonnegative on T to functions that are nonnegative on T . Since the two generators agree except for transitions that lead out of T ,

$$\begin{aligned} (U - V)U_t g^*(x) &= \sum_{i,j=1}^n p(x_i, x_j) [U_t g^*(x_1, \dots, x_{i-1}, x_j, x_{i+1}, \dots, x_n) \\ &\quad - U_t g^*(x)] \\ &= \frac{1}{2} \sum_{i,j=1}^n p(x_i, x_j) [U_t g^*(x_1, \dots, x_{i-1}, x_j, x_{i+1}, \dots, x_n) \\ &\quad + U_t g^*(x_1, \dots, x_{j-1}, x_i, x_{j+1}, \dots, x_n) - 2U_t g^*(x)], \end{aligned}$$

where we have used the symmetry of $p(\cdot, \cdot)$ again. To check that each summand on the right above is nonnegative, it is therefore enough to check that

$$\sum_{u,v \in S} U_t g^*(u, v, x_3, \dots, x_n) \beta(u) \beta(v) \geq 0 \quad (11)$$

for all choices of $x_3, \dots, x_n \in S$ if

$$\sum_{u \in S} |\beta(u)| < \infty, \quad \sum_{u \in S} \beta(u) = 0.$$

The $\beta(\cdot)$ that we are using in this application is

$$\beta(u) = \begin{cases} +1 & \text{if } u = x_1, \\ -1 & \text{if } u = x_2, \\ 0 & \text{otherwise.} \end{cases}$$

To compute the left side of (11), let

$$\Delta_t(u) = E^{(0,u)} \Delta(X_1(t) - X_2(t)) = P^u(Z(s) = 0 \text{ for some } s > 2t).$$

Then

$$U_t g^*(x) = \sum_{1 \leq i < j \leq n} \Delta_t(x_i - x_j),$$

so the left side of (11) is

$$\sum_{u, v \in S} \beta(u) \beta(v) \Delta_t(u - v).$$

The nonnegativity of this expression follows from the following computation, for which we take $0 < t_1 < \cdots < t_n$:

$$\begin{aligned} & \sum_{u, v} \beta(u) \beta(v) P^{(u, v)}(X_1(t_i) = X_2(t_i) \text{ for some } 1 \leq i \leq n) \\ &= \sum_{u, v} \beta(u) \beta(v) \sum_{i=1}^n P^{(u, v)}(X_1(t_i) = X_2(t_i), X_1(t_j) \neq X_2(t_j) \forall j > i) \\ &= \sum_{u, v} \beta(u) \beta(v) \sum_{i=1}^n \sum_w \left[P^u(X_1(t_i) = w) P^v(X_2(t_i) = w) \right. \\ & \quad \left. \times P^{(w, w)}(X_1(t_j - t_i) \neq X_2(t_j - t_i) \forall j > i) \right] \\ &= \sum_{i=1}^n \sum_w \left[\sum_u \beta(u) P^u(X_1(t_i) = w) \right]^2 \\ & \quad \times P^{(w, w)}(X_1(t_j - t_i) \neq X_2(t_j - t_i) \forall j > i) \geq 0. \end{aligned}$$

3.3 The general (non-translation invariant) case

We conclude this chapter by stating the results corresponding to Theorems 1 and 2 when S is a general countable set, and $p(x, y)$ are the transition probabilities for a general symmetric irreducible Markov chain on S . The proofs can be found in Section VIII.1 of Liggett (1985).

Define

$$\mathcal{H} = \left\{ \alpha : S \rightarrow [0, 1] : \sum_y p(x, y) \alpha(y) = \alpha(x) \right\}.$$

For $\alpha \in \mathcal{H}$, let ν_α be the product measure on $\{0, 1\}^S$ with marginals given by α :

$$\nu_\alpha\{\eta : \eta(x) = 1\} = \alpha(x).$$

Finally, let $p_t(x, y)$ be the transition probabilities for the continuous time Markov chain with jump probabilities $p(x, y)$ and exponential holding times with mean 1.

Theorem 3. (a) For every $\alpha \in \mathcal{H}$,

$$\mu_\alpha = \lim_{t \rightarrow \infty} \nu_\alpha S(t)$$

exists.

(b) $\mu_\alpha\{\eta : \eta(x) = 1\} = \alpha(x)$ for all $x \in S$, and

$$\mu_\alpha\{\eta : \eta(x) = 1, \eta(y) = 1\} \leq \alpha(x)\alpha(y)$$

for all $x \neq y \in S$.

(c) μ_α is a product measure if and only if α is a constant.

(d) $\mathcal{I}_e = \{\mu_\alpha : \alpha \in \mathcal{H}\}$.

(e) If the probability measure μ on $\{0, 1\}^S$ satisfies

$$\lim_{t \rightarrow \infty} \sum_y p_t(x, y) \mu\{\eta : \eta(y) = 1\} = \alpha(x)$$

for every $x \in S$, and

$$\lim_{t \rightarrow \infty} \sum_{y_1, y_2} p_t(x_1, y_1) p_t(x_2, y_2) \mu\{\eta : \eta(y_1) = 1, \eta(y_2) = 1\} = \alpha(x_1)\alpha(x_2)$$

for every $x_1, x_2 \in S$, then $\alpha \in \mathcal{H}$ and

$$\lim_{t \rightarrow \infty} \mu S(t) = \mu_\alpha.$$

Note that Theorem 1 is a special case of Theorem 3. For part (a), this uses the fact that random walks on Z^d have only constant bounded harmonic functions, so \mathcal{H} consists only of the constants in $[0, 1]$ in this case. To deduce part (b) of Theorem 1 from Theorem 3, use the argument at the end of the proof of Theorem 7 of Chapter 2.

Here is an example in which \mathcal{H} is very large. Let $S = T_2$, the tree in which each vertex has 3 neighbors. Let $p(x, y) = 1/3$ if x and y are neighbors, and $p(x, y) = 0$ otherwise. To construct a nonconstant $\alpha \in \mathcal{H}$, proceed as follows: Fix two adjacent vertices a, b , and let $\alpha(a) = 1/3$ and $\alpha(b) = 2/3$. If x is a distance n from a and $n+1$ from b , let $\alpha(x) = (1/3)2^{-n}$, while if x is a

distance n from b and $n+1$ from a , let $\alpha(x) = 1 - (1/3)2^{-n}$. Let T^* be the half of the tree made up of sites that are closer to b than to a . Then the harmonic function we have defined is $\alpha(x) = P^x(X(t) \in T^* \text{ from some time on})$. By Theorem 3, there is an extremal stationary distribution for the corresponding exclusion process with these marginals. It is not known explicitly what this measure looks like. There are of course many other elements of \mathcal{H} .

4 Translation Invariant Exclusion Processes

In the last chapter, we saw how duality can be used to analyze symmetric exclusion processes. There is no useful duality for asymmetric systems, so other techniques must be used. In fact, the behavior of asymmetric systems is more varied and interesting, so it is not just the tools that are different – the processes themselves are fundamentally different.

In the first two sections, we will let S and $p(x, y)$ be general. However, the main objective here is to treat translation invariant systems, so that in the rest of the chapter, we will consider the translation invariant case: $S = Z^d$ and $p(x, y) = p(0, y - x)$.

4.1 Stationary product measures

In this section, we will see when (possibly inhomogeneous) product measures are stationary for an exclusion process. If $\alpha : S \rightarrow [0, 1]$, let ν_α be the product measure on $\{0, 1\}^S$ with marginals given by $\nu_\alpha\{\eta : \eta(x) = 1\} = \alpha(x)$ for each $x \in S$.

Theorem 1. (a) *If $p(x, y)$ is doubly stochastic, then $\nu_\alpha \in \mathcal{I}$ for every constant $\alpha \in [0, 1]$. In particular, this is true in the translation invariant case.*

(b) *If $\pi(x) \geq 0$ and $\pi(x)p(x, y) = \pi(y)p(y, x)$ for all $x, y \in S$, (i.e., π is reversible with respect to $p(x, y)$) then $\nu_\alpha \in \mathcal{I}$, where*

$$\alpha(x) = \frac{\pi(x)}{1 + \pi(x)}. \quad (1)$$

Remark. If $p(x, y)$ is symmetric, then it satisfies both (a) and (b) (with $\pi = \text{constant}$). In this case, we checked that the homogeneous product measures are stationary using duality in Chapter 3.

Proof. By Theorem 1 of Chapter 1, it suffices to check that

$$\int \Omega f d\nu_\alpha = 0$$

for all cylinder functions f . In the interest of simplicity, we will carry out the computation in the case in which S is finite, so that all the sums below are finite. The general case follows by approximation, or by a more careful handling of the sums involved. (See Theorem 2.1 of Chapter VIII of Liggett (1985).) Then

$$\begin{aligned} \int \Omega f d\nu_\alpha &= \int \sum_{x,y:\eta(x)=1,\eta(y)=0} p(x,y) [f(\eta_{x,y}) - f(\eta)] d\nu_\alpha \\ &= \sum_{x,y} p(x,y) \int_{\eta:\eta(x)=1,\eta(y)=0} [f(\eta_{x,y}) - f(\eta)] d\nu_\alpha. \end{aligned}$$

It suffices to assume that $0 < \alpha(x) < 1$ for all x . Making the change of variables $\eta \rightarrow \eta_{x,y}$ in the integral, write

$$\begin{aligned} &\sum_{x,y} p(x,y) \int_{\eta:\eta(x)=1,\eta(y)=0} f(\eta_{x,y}) d\nu_\alpha \\ &= \sum_{x,y} p(x,y) \int_{\eta:\eta(x)=0,\eta(y)=1} f(\eta) \frac{\alpha(x)[1-\alpha(y)]}{[1-\alpha(x)]\alpha(y)} d\nu_\alpha. \end{aligned} \quad (2)$$

In case (a), $\alpha(x)$ is constant, so we get

$$\int \Omega f d\nu_\alpha = \int f(\eta) \sum_{x,y} p(x,y) \{ \eta(y)[1-\eta(x)] - \eta(x)[1-\eta(y)] \} d\nu_\alpha.$$

But in the doubly stochastic case,

$$\sum_{x,y} p(x,y) \{ \eta(y)[1-\eta(x)] - \eta(x)[1-\eta(y)] \} = \sum_{x,y} p(x,y) \{ \eta(y) - \eta(x) \} = 0.$$

(Recall that these sums are finite, since we are taking S finite in this computation.)

Turning to case (b) and interchanging the roles of x and y , (2) becomes

$$\sum_{x,y} p(y,x) \int_{\eta:\eta(y)=0,\eta(x)=1} f(\eta) \frac{\pi(y)}{\pi(x)} d\nu_\alpha.$$

Using the reversibility assumption, this becomes

$$\sum_{x,y} p(x,y) \int_{\eta:\eta(y)=0,\eta(x)=1} f(\eta) d\nu_\alpha,$$

so that $\int \Omega f d\nu_\alpha = 0$.

Remark. Note that in case (b), ν_α is not only stationary, but is reversible as well. This means that the stationary process η_t obtained by using ν_α as the initial distribution (when extended to $t \in (-\infty, \infty)$) has the property that $\{\eta_t, t \in (-\infty, \infty)\}$ and $\{\eta_{-t}, t \in (-\infty, \infty)\}$ have the same joint distributions. Analytically, it means that the semigroup of the process is symmetric on $L_2(\nu_\alpha)$:

$$\int f S(t) g d\nu_\alpha = \int g S(t) f d\nu_\alpha$$

for $f, g \in L_2(\nu_\alpha)$.

4.2 Coupling

The basic coupling of two copies of the exclusion process is the process (η_t, ζ_t) with the following transitions at rate $p(x, y)$:

$$\begin{aligned} (\eta, \zeta) &\rightarrow (\eta_{x,y}, \zeta_{x,y}) \text{ if } \eta(x) = \zeta(x) = 1 \text{ and } \eta(y) = \zeta(y) = 0, \\ (\eta, \zeta) &\rightarrow (\eta_{x,y}, \zeta) \quad \text{if } \eta(x) = 1, \eta(y) = 0 \text{ and } (\zeta(x) = 0 \text{ or } \zeta(y) = 1) \\ (\eta, \zeta) &\rightarrow (\eta, \zeta_{x,y}) \quad \text{if } \zeta(x) = 1, \zeta(y) = 0 \text{ and } (\eta(x) = 0 \text{ or } \eta(y) = 1). \end{aligned}$$

In other words, particles move together whenever they can. A discrepancy is a site x at which $\eta(x) \neq \zeta(x)$. An important property of the coupling is that, while discrepancies can move and disappear, they cannot be created. Another important property is that $\eta_0 \leq \zeta_0$ implies $\eta_t \leq \zeta_t$ for all $t > 0$, which shows that the process is attractive. An immediate consequence of this is that every extremal stationary distribution for the coupled process assigns probability zero or one to the sets $\{\eta \leq \zeta\}$, $\{\eta = \zeta\}$ and $\{\zeta \leq \eta\}$.

Let \mathcal{I}^* be the set of stationary distributions for the coupled process. More generally, $*$ will denote objects related to the coupled process. The next result provides some connections between \mathcal{I}^* and \mathcal{I} . By the translation invariant case, we will mean the case in which $S = Z^d$ and $p(x, y) = p(y - x)$ for some probability density $p(x)$ on Z^d . In this case, we will let \mathcal{S} and \mathcal{S}^* denote the class of shift invariant probability measure on $\{0, 1\}^S$ and $\{0, 1\}^S \times \{0, 1\}^S$ respectively.

Theorem 2. (a) If $\nu^* \in \mathcal{I}^*$ then its marginals are in \mathcal{I} .

(b) If $\nu_1, \nu_2 \in \mathcal{I}$, then there is a $\nu^* \in \mathcal{I}^*$ with marginals ν_1 and ν_2 .

(c) If $\nu_1, \nu_2 \in \mathcal{I}_e$, then the ν^* in part (b) can be taken to be in \mathcal{I}_e^* .

(d) In parts (b) and (c), if $\nu_1 \leq \nu_2$, then ν^* can be taken to concentrate on $\{\eta \leq \zeta\}$.

(e) In the translation invariant case, parts (a)-(d) hold if \mathcal{I} and \mathcal{I}^* are replaced by $\mathcal{I} \cap \mathcal{S}$ and $\mathcal{I}^* \cap \mathcal{S}^*$ respectively.

Proof. Part (a) follows from the fact that the marginal processes η_t and ζ_t are versions of the exclusion process. For part (b), let μ^* be the product measure $\nu_1 \times \nu_2$. Then ν^* can be taken to be the limit of a weakly convergent sequence

$$\frac{1}{t_n} \int_0^{t_n} \mu^* S(s) ds$$

of Cesaro averages, with $t_n \uparrow \infty$. Such a limit exists by compactness of the state space, and is in \mathcal{I}^* . Turning to part (c), take the ν^* obtained in part (b), and write it as a mixture of elements of \mathcal{I}_e^* . The marginals ν_1 and ν_2 are the corresponding mixtures of the marginals of these elements of \mathcal{I}_e^* , so by the extremality assumption on ν_1 and ν_2 , almost every such element has marginals ν_1 and ν_2 . Take any one of them to be the required element of \mathcal{I}_e^* . For part (d), instead of taking μ^* to be the product measure $\nu_1 \times \nu_2$, take it to be any measure with marginals ν_1 and ν_2 that concentrates on $\{\eta \leq \zeta\}$. For the last statement, carry out the above proofs with the replacements given.

Remark. The extension of (d) to the translation invariant setting is somewhat more subtle than the others. If $\nu_1 \leq \nu_2$ and both ν_i are shift invariant, one needs to find a coupling measure for them that concentrates on $\{(\eta, \zeta) : \eta \leq \zeta\}$, and is also shift invariant. To do so, take any coupling measure ν for ν_1 and ν_2 that concentrates on $\{(\eta, \zeta) : \eta \leq \zeta\}$ – it need not be shift invariant. Then take a limit of Cesaro averages of shifts of ν as the coupling measure to be used in the proof. To check that this limit is shift invariant, one needs to use the fact that the boundary of a cube of side length N in Z^d has a size that is of smaller order than the volume of the cube. Thus this proof would not work if Z^d is replaced by the homogeneous tree T_d , for example.

4.3 Shift invariant stationary measures

In this section, we consider the irreducible translation invariant case, and explain how to use coupling to show that the stationary distributions that are

shift invariant are exactly the exchangeable measures. Since all exchangeable measures are mixtures of homogeneous product measures by de Finetti's Theorem (see the exercise preceding Theorem 7 in Chapter 2), the fact that every exchangeable measure is stationary follows from Theorem 1(a). We begin with a basic fact about the coupled process – in a shift invariant equilibrium situation, discrepancies of opposite type cannot coexist. This is a consequence of the fact that discrepancies cannot be created in the coupled process.

Theorem 3. *If $\nu^* \in \mathcal{I}^* \cap \mathcal{S}^*$, then*

$$\nu^* \{(\eta, \zeta) : \eta(x) = \zeta(y) = 0, \eta(y) = \zeta(x) = 1\} = 0$$

for every $x, y \in S$

Proof. For any probability measure ν^* on $\{0, 1\}^S \times \{0, 1\}^S$,

$$\begin{aligned} \left. \frac{d}{dt} \nu^* S^*(t) \{(\eta, \zeta) : \eta(x) \neq \zeta(x)\} \right|_{t=0} = & \\ & \sum_y p(x, y) \nu^* \{(\eta, \zeta) : \eta(x) = \zeta(x) = 1, \eta(y) \neq \zeta(y)\} \\ & + \sum_y p(y, x) \nu^* \{(\eta, \zeta) : \eta(x) = \zeta(x) = 0, \eta(y) \neq \zeta(y)\} \\ & - \sum_y p(x, y) \nu^* \{(\eta, \zeta) : \eta(x) \neq \zeta(x), \eta(y) = \zeta(y) = 0\} \\ & - \sum_y p(y, x) \nu^* \{(\eta, \zeta) : \eta(x) \neq \zeta(x), \eta(y) = \zeta(y) = 1\} \\ & - \sum_y [p(x, y) + p(y, x)] \nu^* \{(\eta, \zeta) : \eta(x) = \zeta(y) \neq \eta(y) = \zeta(x)\}. \quad (3) \end{aligned}$$

The first two terms on the right of (3) correspond to a discrepancy moving from y to x , the third and fourth terms correspond to a discrepancy moving from x to y , and the final term corresponds to the destruction of the discrepancy at x . If $\nu^* \in \mathcal{I}^*$, then the left side of (3) is zero. If $\nu^* \in \mathcal{S}^*$, then the ν^* -probabilities appearing on the right are functions of $y - x$. Therefore in this case, the first and fourth terms cancel, and the second and third terms cancel. This implies that

$$\nu^* \{(\eta, \zeta) : \eta(x) = \zeta(y) \neq \eta(y) = \zeta(x)\} = 0 \quad (4)$$

for all x, y such that $p(x, y) + p(y, x) > 0$. Using the irreducibility of $p(x, y)$ and the stationarity of ν^* again, it then follows that (4) holds for all x, y as required. To check this, one shows by induction on n that if

$$x = x_0, x_1, \dots, x_n = y \text{ satisfy } p(x_i, x_{i+1}) > 0 \text{ for each } i, \quad (5)$$

then (4) holds for that pair x, y . We have proved the basis step $n = 1$ already. For the induction step, assume this is true for $n - 1$ (and all x, y which can be joined by a path of length $n - 1$ as in (5)), and take x, y joined by a path of length n as in (5). Use the notation:

$$\nu^* \begin{Bmatrix} 1 & 0 \\ 0 & 1 \\ u & v \end{Bmatrix} = \nu^* \{(\eta, \zeta) : \eta(u) = \zeta(v) = 0, \eta(v) = \zeta(u) = 1\},$$

for example. Then write

$$\begin{aligned} \nu^* \begin{Bmatrix} 1 & 0 \\ 0 & 1 \\ x & y \end{Bmatrix} &= \nu^* \begin{Bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ x & x_1 & y \end{Bmatrix} + \nu^* \begin{Bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ x & x_1 & y \end{Bmatrix} \\ &\quad + \nu^* \begin{Bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ x & x_1 & y \end{Bmatrix} + \nu^* \begin{Bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ x & x_1 & y \end{Bmatrix}. \end{aligned}$$

The last two terms on the right are zero by the induction hypothesis.

We will check that the second term is zero; the first term is treated in a similar manner. Suppose that at time zero, the process is in the situation

$$\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 0 & 1 \\ x & x_1 & y \end{array}$$

and that by time t , there has been an attempted transition from $x = x_0$ to x_1 , but no other attempted transitions involving x_0, \dots, x_n . Then at time t , the situation will be

$$\begin{array}{ccc} 0 & 1 & 0 \\ 0 & 0 & 1 \\ x & x_1 & y \end{array}$$

Therefore,

$$\nu^* S^*(t) \begin{Bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ x & x_1 & y \end{Bmatrix} \geq \nu^* \begin{Bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ x & x_1 & y \end{Bmatrix} te^{-(2n+2)t} p(x, x_1).$$

Since ν^* is stationary for the coupled process, the left side above is zero by the induction hypothesis, and therefore the right side is also zero as required.

Theorem 4. $(\mathcal{I} \cap \mathcal{S})_e = \{\nu_\rho, 0 \leq \rho \leq 1\}$.

Proof. By Theorem 1(a), $\nu_\rho \in \mathcal{I}$, and of course $\nu_\rho \in \mathcal{S}$. Furthermore, $\nu_\rho \in \mathcal{S}_e$, since it is spatially ergodic. Therefore, $\nu_\rho \in (\mathcal{I} \cap \mathcal{S})_e$.

For the converse, take $\nu \in (\mathcal{I} \cap \mathcal{S})_e$. By Theorem 2(e), for any $0 \leq \rho \leq 1$, there is a $\nu^* \in (\mathcal{I}^* \cap \mathcal{S}^*)_e$ with marginals ν_ρ and ν . By Theorem 3,

$$\nu^* \{(\eta, \zeta) : \eta \leq \zeta, \eta \neq \zeta\} + \nu^* \{(\eta, \zeta) : \zeta \leq \eta, \eta \neq \zeta\} + \nu^* \{(\eta, \zeta) : \eta = \zeta\} = 1.$$

Since the three sets above are closed for the evolution and translation invariant, and since ν^* is extremal, it follows that one of the three sets has full measure. Therefore, for every $0 \leq \rho \leq 1$, either $\nu \leq \nu_\rho$ or $\nu_\rho \leq \nu$. It follows that $\nu = \nu_{\rho_0}$ where ρ_0 is determined by

$$\begin{aligned} \nu &\leq \nu_\rho && \text{for } \rho > \rho_0, \\ \nu &\geq \nu_\rho && \text{for } \rho < \rho_0. \end{aligned}$$

4.4 Some extensions

The technique used in the proof of Theorem 4 can be used in some other situations. Here is an example.

Theorem 5. *If $p(x, y) = p(0, y - x)$ is irreducible on Z^1 ,*

$$\sum_x |x| p(0, x) < \infty, \text{ and } \sum_x x p(0, x) = 0,$$

then all stationary distributions are exchangeable.

The proof of this result can be found in Chapter VIII of Liggett (1985). Here we describe only the idea of the proof. The main point is again to prove an analogue of Theorem 3, without assuming that ν^* is shift invariant. Once that is done, one can proceed as in the proof of Theorem 4. Instead of considering the rate of change of the probability of having a discrepancy at x in (3) (which is negative only because of the cancellation that occurs in the translation invariant case), we now need to look at the rate of change of the expected number of discrepancies in a large interval $[-N, N]$. This quantity can change in both directions – discrepancies can enter $[-N, N]$ from

outside the interval, leading to an increase in discrepancies in the interval. This problem is handled by showing that for large N , the rate at which discrepancies cross N or $-N$ is small. When looked at near $\pm\infty$, the coupling measure ν^* is almost translation invariant (this requires a Cesaro argument). But Theorem 4 tells us that the translation invariant stationary distributions for η_t are exchangeable. Suppose, for example, that ν^* were shift invariant and had marginals ν_λ and ν_ρ respectively, with $\lambda < \rho$. Then by Theorem 3, one might expect that ν^* is the product measure with marginals

$$\begin{aligned}\nu^*\{(\eta, \zeta) : \eta(x) = \zeta(x) = 1\} &= \lambda, \\ \nu^*\{(\eta, \zeta) : \eta(x) = 0, \zeta(x) = 1\} &= \rho - \lambda, \\ \nu^*\{(\eta, \zeta) : \eta(x) = \zeta(x) = 0\} &= 1 - \rho.\end{aligned}$$

It would follow that the net rate of flow of discrepancies to the right across N , for example, is

$$\begin{aligned}& \sum_{x < N \leq y} p(x, y) \nu^*\{(\eta, \zeta) : \eta(x) = 0, \zeta(x) = 1, \eta(y) = \zeta(y) = 0\} \\ & + \sum_{x < N \leq y} p(y, x) \nu^*\{(\eta, \zeta) : \eta(x) = 0, \zeta(x) = 1, \eta(y) = \zeta(y) = 1\} \\ & - \sum_{x < N \leq y} p(y, x) \nu^*\{(\eta, \zeta) : \eta(x) = \zeta(x) = 0, \eta(y) = 0, \zeta(y) = 1\} \\ & - \sum_{x < N \leq y} p(x, y) \nu^*\{(\eta, \zeta) : \eta(x) = \zeta(x) = 1, \eta(y) = 0, \zeta(y) = 1\} \\ & = (\rho - \lambda)(1 - \rho - \lambda) \sum_u u p(0, u),\end{aligned}$$

which vanishes by the mean zero assumption. This is not quite the way the proof goes, but this computation does show how the mean zero assumption arises.

Remark. We are very far from understanding the structure of \mathcal{I}_e in the general translation invariant case on Z^d . In fact, for $d \geq 2$, \mathcal{I}_e has been completely described only in the symmetric case covered in Chapter 3. Certainly it is often not the case that

$$\mathcal{I}_e = \{\nu_\rho, 0 \leq \rho \leq 1\}.$$

For example, suppose there is a nonzero $v \in Z^d$ so that

$$p(0, x) = e^{\langle x, v \rangle} p(0, -x) \tag{6}$$

for all x . Then $\pi(x) = ce^{\langle x, v \rangle}$ is reversible for $p(x, y)$, so Theorem 1(b) provides examples on stationary distributions that are not shift invariant.

An instructive example is the following. Let $\{e_1, \dots, e_d\}$ be the standard basis vectors in Z^d , and take

$$p(x, x + e_i) = \alpha_i, \quad p(x, x - e_i) = \beta_i$$

with all other $p(x, y)$'s zero. Assume that all the α_i 's and β_i 's are strictly positive, and set

$$v_i = \log \frac{\alpha_i}{\beta_i}.$$

Then (6) holds with $v = (v_1, \dots, v_d)$. The corresponding stationary distribution ν satisfies

$$\nu\{\eta : \eta(x) = 1\} = \frac{ce^{\langle x, v \rangle}}{1 + ce^{\langle x, v \rangle}}. \quad (7)$$

Note that this is constant on hyperplanes that are orthogonal to v , while one might have expected it to be constant on hyperplanes that are orthogonal to the the mean vector

$$\sum_x xp(0, x) = (\alpha_1 - \beta_1, \dots, \alpha_d - \beta_d).$$

It would be very interesting to know that all the extremal stationary distributions in this case are the homogeneous product measures and the product measures with marginals given by (7).

The picture is much more complete in the one dimensional case, however. That is the topic of the next chapter.

5 One Dimensional Exclusion Processes

In this chapter, we will focus on translation invariant exclusion processes on $S = Z^1$. The first section deals with the nearest neighbor case. The more general case is described in Section 2. As we will see, the behavior of the process is much more interesting when the particles have a drift.

5.1 The nearest neighbor case

In this section, we take $p(x, x + 1) = p$ and $p(x, x - 1) = q$ where $p + q = 1, 0 < p < 1$. Letting

$$\pi(x) = \left(\frac{p}{q}\right)^x \quad \text{and} \quad \alpha(x) = \frac{p^x}{p^x + q^x},$$

we see from Theorem 1(b) of Chapter 4 that $\nu_\alpha \in \mathcal{I}$. Note that by the Borel-Cantelli Lemma, if $p > \frac{1}{2}$, then ν_α assigns full measure to

$$\Xi = \left\{ \eta \in \{0, 1\}^S : \sum_{x < 0} \eta(x) < \infty, \sum_{x > 0} [1 - \eta(x)] < \infty \right\}.$$

Decompose Ξ in the following way:

$$\Xi = \cup_{n=-\infty}^{\infty} \Xi_n, \quad \text{where} \quad \Xi_n = \left\{ \eta \in \Xi : \sum_{x < n} \eta(x) = \sum_{x \geq n} [1 - \eta(x)] \right\}.$$

Then the process η_t remains in Ξ_n if it begins in it, so η_t is simply an irreducible, countable state Markov chain on Ξ_n . But the conditional distribution

$$\mu_n(\cdot) = \nu_\alpha(\cdot \mid \Xi_n)$$

is stationary for this Markov chain, so the chain is positive recurrent.

Using coupling, much as it was used in Chapter 4, one can prove that we have now found all the stationary distributions for the process. The main part of the proof is an analogue of Theorem 3 of Chapter 4. That theorem says that discrepancies of opposite type cannot coexist in a translation invariant equilibrium for the coupled process. This is not the case in the present context, since if one couples the stationary measures $\nu_{\frac{1}{2}}$ and μ_0 , for example, there must be discrepancies of one type on the negative half line, and of the other type on the positive half line. This is the worst that can happen, though. It turns out that every element of \mathcal{I}^* concentrates on configurations that have all discrepancies of one type to the left of all discrepancies of the other type. (See Liggett (1976) for details.) Once this is established, the proof of the following result follows the lines of the proof of Theorem 4 of Chapter 4.

Theorem 1. *If $p > \frac{1}{2}$, then $\mathcal{I}_e = \{\nu_\rho, 0 \leq \rho \leq 1\} \cup \{\mu_n, -\infty < n < \infty\}$.*

Thus we see that the asymmetric case allows for the existence of more stationary distributions than does the symmetric case.

Next, we look at the limiting behavior of the process when the initial distribution is highly non-translation invariant. For $0 \leq \lambda, \rho \leq 1$, let $\nu_{\lambda, \rho}$ be the product measure with

$$\nu_{\lambda, \rho}\{\eta : \eta(x) = 1\} = \begin{cases} \lambda & \text{if } x < 0 \\ \rho & \text{if } x \geq 0. \end{cases}$$

By Theorem 3(e) of Chapter 3, if $p = \frac{1}{2}$ then

$$\lim_{t \rightarrow \infty} \nu_{\lambda, \rho} S(t) = \nu_{(\lambda + \rho)/2}.$$

By contrast, in the asymmetric case, we have the following quite different statement. (See Liggett (1977) and Andjel, Bramson and Liggett (1988) for the proof.)

Theorem 2. *If $p > \frac{1}{2}$, then*

$$\lim_{t \rightarrow \infty} \nu_{\lambda, \rho} S(t) = \begin{cases} \nu_{1/2} & \text{if } \lambda \geq 1/2 \text{ and } \rho \leq 1/2, \\ \nu_{\rho} & \text{if } \rho \geq 1/2 \text{ and } \lambda + \rho > 1, \\ \nu_{\lambda} & \text{if } \lambda \leq 1/2 \text{ and } \lambda + \rho < 1, \\ \frac{1}{2}\nu_{\lambda} + \frac{1}{2}\nu_{\rho} & \text{if } 0 < \lambda \leq \rho \text{ and } \lambda + \rho = 1, \\ \mu_0 & \text{if } \lambda = 0 \text{ and } \rho = 1. \end{cases}$$

To help understand why this might be limit in the various cases, we describe a heuristic argument. Take as an example the case $p = 1$, and write μ_t for $\mu S(t)$. Then

$$\frac{d}{dt} \mu_t \{ \eta : \eta(x) = 1 \} = \mu_t \{ \eta : \eta(x-1) = 1, \eta(x) = 0 \} - \mu_t \{ \eta : \eta(x) = 1, \eta(x+1) = 0 \}.$$

Letting $u(x, t) = \mu_t \{ \eta : \eta(x) = 1 \}$ and pretending that μ_t is a product measure (which it is not), this can be seen as a discrete approximation to Burgers' partial differential equation

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} [u(1 - u)] = 0.$$

Consider this differential equation with initial condition

$$u(x, 0) = \begin{cases} \lambda & \text{if } x < 0, \\ \rho & \text{if } x \geq 0. \end{cases}$$

Then the (entropic) solution is given by

$$u(x, t) = \begin{cases} \lambda & \text{if } x \leq (1 - 2\lambda)t, \\ \frac{t-x}{2t} & \text{if } (1 - 2\lambda)t \leq x \leq (1 - 2\rho)t, \\ \rho & \text{if } x \geq (1 - 2\rho)t, \end{cases}$$

provided that $\lambda \geq \rho$, and

$$u(x, t) = u(x - vt, 0), \quad \text{where } v = 1 - \lambda - \rho$$

if $\lambda \leq \rho$. Therefore, except in the case $\lambda + \rho = 1$, $\lambda < \rho$, the limit in Theorem 2 is the product measure with density given by

$$\lim_{t \rightarrow \infty} u(x, t)$$

(which is independent of x).

5.2 The general case

In this section, we assume that $S = \mathbb{Z}^1$, $p(x, y) = p(0, y - x)$, $\sum_x |x|p(0, x) < \infty$ and the corresponding random walk is irreducible. In this case, we know that $\mathcal{I}_e = \{\nu_\rho, 0 \leq \rho \leq 1\}$ if $\sum_x xp(0, x) = 0$ (Theorem 5 of Chapter 4), and that the analogues of the first three cases of Theorem 2 above hold if $\sum_x xp(0, x) > 0$ (Theorem 1.1 of Liggett (1977)). For the rest of this section, we focus on analogues of Theorem 1 above. Without loss of generality, we may assume from here on that $\sum_x xp(0, x) \geq 0$.

A probability measure on Ξ will be called a blocking measure, while a probability measure μ on $\{0, 1\}^S$ that satisfies

$$\lim_{x \rightarrow -\infty} \mu\{\eta : \eta(x) = 1\} = 0 \quad \text{and} \quad \lim_{x \rightarrow \infty} \mu\{\eta : \eta(x) = 1\} = 1$$

will be called a profile measure. Every blocking measure is a profile measure, but not conversely. The following results were proved recently in Bramson, Liggett and Mountford (2002). That paper was an outgrowth of two other recent papers: Bramson and Mountford (2002) and Ferrari, Lebowitz and Speer (2001).

Theorem 3. *Either (a) $\mathcal{I}_e = \{\nu_\rho, 0 \leq \rho \leq 1\}$, or*

(b) there is a profile invariant measure μ so that

$$\mathcal{I}_e = \{\nu_\rho, 0 \leq \rho \leq 1\} \cup \{\mu_n, -\infty < n < \infty\},$$

where μ_n is the shift of μ by a distance n .

We know that both situations can occur: (a) happens when $p(0, \cdot)$ has mean zero (Theorem 5 of Chapter 4), and (b) happens in the nearest neighbor case with positive drift (Theorem 1 in this chapter). Thus we would like to know exactly when each of (a) and (b) occurs. The next two theorems provide some answers to this question.

Theorem 4. *If $\sum_{x<0} x^2 p(0, x) = \infty$, then no stationary blocking measures exist.*

The idea behind this result is that if the negative tails of $p(0, \cdot)$ are very large, and if there were a blocking equilibrium (in which case all sites to the right of some point would be occupied), then there would have to be infinitely many particles to the left of the origin. Note that Theorem 4 does not rule out the existence of a stationary profile measure. It is not known whether nonblocking stationary profile measures exist. The following complements Theorem 4.

Theorem 5. *Suppose (a) $p(0, x)$ and $p(0, -x)$ are decreasing in x for $x \geq 1$, (b) $p(0, x) \geq p(0, -x)$ for all $x \geq 1$, with a strict inequality for some $x \geq 1$, and*

$$(c) \sum_{x<0} x^2 p(0, x) < \infty.$$

Then a stationary blocking measure exists.

The proof of this theorem is too long to explain here. However, we will describe the three main steps in the proof, including a computation that shows explicitly how the second moment assumption enters into the proof.

Outline of part 1 of the proof. Here the objective is to show that stationary blocking measures exist when the negative tails of p are small. Suppose, for example, that

$$\sum_{x<0} |x|^5 p(0, x) < \infty,$$

and choose $\epsilon > 0$ so small that

$$\sum_{x<0} |x|(1 + \epsilon|x|)^4 p(0, x) \leq \sum_{x>0} x(1 + \epsilon x)^{-4} p(0, x).$$

This is possible by the fifth moment assumption on the negative tails, and the fact the drift is strictly positive. Define the product measure ν_α as in Section 3.3, with $\alpha(x)$ chosen to be

$$\alpha(x) = \begin{cases} \frac{1}{1+(1+\epsilon|x|)^2} & \text{if } x < 0 \\ \frac{(1+\epsilon x)^2}{1+(1+\epsilon x)^2} & \text{if } x > 0. \end{cases}$$

Note that ν_α is a blocking measure. However, there is no reason for it to be stationary. However, because of assumption (a) and the choice of ϵ , one can

use a nonstandard coupling argument to show that

$$\sum_{x < u} \nu_\alpha S(t) \{\eta : \eta(x) = 1\}$$

is nonincreasing in t for every u . The limiting measure as $t \rightarrow \infty$ is then a stationary blocking measure.

An important step in checking this is the following symmetry: For $\eta \in \{0, 1\}^{\mathbb{Z}^1}$, define $\eta^* \in \{0, 1\}^{\mathbb{Z}^1}$ by

$$\eta^*(x) = 1 - \eta(-x).$$

Then η_t^* also evolves like the exclusion process. To see this, note that if a particle in η_t moves from u to v , then a nonparticle (a site at which $\eta_t(x) = 0$) moves from v to u . Since ν_α is invariant under the transformation $\eta \rightarrow \eta^*$, so is $\nu_\alpha S(t)$.

Outline of part 2 of the proof. To remove the moment assumption that we imposed in part 1, we will need to take the p that we are interested in, whose negative tails have only a second moment, and truncate it to obtain an approximating sequence p_k whose negative tails have a finite fifth moment. We will need to have some sort of *a priori* bound that will allow us to maintain control of the corresponding sequence of stationary blocking measures for p_k as $k \rightarrow \infty$. To do so, suppose ν is a stationary measure that satisfies

$$\sum_{x < 0} \nu\{\eta : \eta(x) = 1\} < \infty, \quad \sum_{x > 0} \nu\{\eta : \eta(x) = 0\} < \infty, \quad (1)$$

(which implies it is a blocking measure) and $\sum_{x < 0} x^2 p(0, x) < \infty$. While (1) is stronger than being blocking, note that the stationary measures constructed in part 1 have this stronger property. The result of this part will be applied to p_k for each k .

Let

$$M(n) = \sum_{x=-\infty}^{\infty} \nu\{\eta : \eta(x) = 1, \eta(x+n) = 0\}.$$

for $n \in \mathbb{Z}^1$. For $n = 1$, this is the expected number of 10's in the configuration η . Note that $M(n) < \infty$ for all n by (1). For $n \geq 1$,

$$\begin{aligned} M(-n) - M(n) &= \sum_x [\nu\{\eta(x) = 0, \eta(x+n) = 1\} - \nu\{\eta(x) = 1, \eta(x+n) = 0\}] \\ &= \sum_x [\nu\{\eta(x+n) = 1\} - \nu\{\eta(x) = 1\}] = n, \end{aligned}$$

since the latter series telescopes. Therefore,

$$M(-n) = M(n) + n. \quad (2)$$

Next, note that for a stationary measure, the net rate at which particles go from the left of x to the right of x is independent of x . If the stationary measure is also a profile measure, this rate must be zero (since it tends to zero at $\pm\infty$). Therefore, for fixed x ,

$$\sum_{u \leq x < v} p(u, v) \nu\{\eta(u) = 1, \eta(v) = 0\} = \sum_{u \leq x < v} p(v, u) \nu\{\eta(u) = 0, \eta(v) = 1\}.$$

Summing over all x yields

$$\sum_{n=1}^{\infty} np(0, n)M(n) = \sum_{n=1}^{\infty} np(0, -n)M(-n).$$

Using (2) in this expression gives

$$\sum_{n=1}^{\infty} n^2 p(0, -n) = \sum_{n=1}^{\infty} nM(n)[p(0, n) - p(0, -n)]. \quad (3)$$

By assumption (b) of Theorem 5, the coefficients of $M(n)$ on the right are nonnegative, and one of them is strictly positive. Therefore, the second moment of the negative tail of p controls the size of some $M(n)$. For simplicity, suppose that it is $M(1)$.

Outline of part 3 of the proof. Now take p satisfying the assumptions of the theorem, and approximate it by a sequence p_k , each of which satisfies those assumptions, and in addition has negative tails with a finite fifth moment. Let μ_k be the corresponding stationary blocking measures that concentrate on Ξ_0 . Their existence is guaranteed by part 1 of the proof. By part 2 of the proof, we may assume that

$$\sup_k \sum_x \mu_k\{\eta : \eta(x) = 1, \eta(x+1) = 0\} < \infty. \quad (4)$$

Informally speaking, there cannot be many 10's in μ_k . The stationary blocking measure μ we are trying to produce will be a limit point of μ_k as $k \rightarrow \infty$. It is easy to check that such a limit point is stationary (Proposition 2.14 of Chapter I of Liggett (1985).) The problem is to show that it is a blocking measure. In particular, we need to rule out the possibility that $\mu = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$.

So, suppose that μ_k converges weakly to $\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$. Then for large k , relative to μ_k , there will be a large number of 1's to the left of the origin with probability close to $\frac{1}{2}$. Since μ_k concentrates on Ξ_0 , there must also be a large number of 0's to the right of the origin. These 0's may occur in a large clump, or in many small clumps. The latter possibility is ruled out by (4). So, there must be something close to a long interval of 1's followed by a long interval of 0's in the configuration. By Theorem 2, the limiting distribution of the exclusion process beginning with the configuration

$$\dots 1 1 1 1 0 0 0 0 \dots$$

is $\nu_{1/2}$. (Recall that the first three parts of Theorem 2 have been proved in the nonnearest neighbor context.) But μ_k is stationary, and $\nu_{1/2}$ has many 10's, so this again contradicts (4). Thus $\mu \neq \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$.

Acknowledgments

(a) The author would like to thank J. Steif for his very helpful corrections and comments on the first draft of these notes.

(b) Preparation of these notes was supported in part by NSF Grant DMS-00-70465.

References

- [1] E.D. Andjel, M. D. Bramson, and T. M. Liggett, *Shocks in the asymmetric exclusion process*, Probab. Th. Rel. Fields **78** (1988), 231–247.
- [2] M.D. Bramson and T. Mountford, *Blocking measures for one dimensional nonzero mean exclusion processes*, Ann. Probab. (2002).
- [3] M.D. Bramson, T.M. Liggett, and T. Mountford, *Characterization of stationary measures for one dimensional exclusion processes*, Ann. Probab. (2002).
- [4] W. Feller, *An Introduction to Probability and its Applications*, Vol. II, Wiley, New York, 1966.
- [5] P.A. Ferrari, J.L. Lebowitz, and E. Speer, *Blocking measures for asymmetric exclusion processes via coupling*, Bernoulli (2001).
- [6] R. Holley and T.M. Liggett, *Ergodic theorems for weakly interacting infinite systems and the voter model*, Ann. Probab. **3** (1975), 643–663.
- [7] T.M. Liggett, *Coupling the simple exclusion process*, Ann. Probab. **4** (1976), 339–356.
- [8] T.M. Liggett, *Ergodic theorems for the asymmetric simple exclusion process II*, Ann. Probab. **5** (1977), 795–801.
- [9] T.M. Liggett, *Interacting Particle Systems*, Springer, New York, 1985.
- [10] T.M. Liggett, *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*, Springer, Berlin, 1999.