AN INTRODUCTION TO INFINITE PARTICLE SYSTEMS*

Richard DURRETT

Department of Mathematics, University of California, Los Angeles, CA 90024, U.S.A.

In 1970, Spitzer wrote a paper called "Interaction of Markov processes" in which he introduced several classes of interacting particle systems. These processes and other related models, collectively referred to as infinite particle systems, have been the object of much research in the last ten years. In this paper we will survey some of the results which have been obtained and some of the open problems, concentrating on six overlapping classes of processes: the voter model, additive processes, the exponential family, one dimensional systems, attractive systems, and the Ising model.

AMS 1970 Subj. Class.: Primary 60K35

Infinite particle system	attractive system
voter model	Ising model
contact process	Gibbs state
additive process	

1. Introduction

Since this survey is based on the lecture I gave in Evanston, I would like to begin my paper as I should have begun my talk – by describing in an informal fashion some of the processes we will consider and some of the questions we would like to answer. You should notice that I said "would like to answer" in the last sentence – there are many basic questions which have not been answered. This is, for me, one of the exciting aspects of infinite particle systems. There are many conjectures which are obviously true but which are very difficult to prove.

I will state some of these conjectures before long, but before I can do this I need to introduce some notation and terminology:

Let S be a countable set. S, which will usually be $Z^d = \{(n_1, \ldots, n_d): n_i$ are integers}, is a set of sites (locations in space) where the events of interest occur.

Let F be a two element set (or if the reader wants some finite set). F, which will always be $\{0, 1\}$ or $\{-1, 1\}$ in this paper, is the set of states in which we can find the various sites (occupied or not, infected or healthy, spin up or down).

Let $\chi = F^S =$ the set of all functions from S to F. If $\eta \in \chi$, then $\eta(x)$ gives the state at x, so η describe the configuration of the system.

Let $\Lambda =$ set of all subsets of S. If $a \in F$, the mapping $\eta \rightarrow \{x: \eta(x) = a\}$ gives a 1–1 correspondence of Λ and χ. In some examples Λ is a more convenient state space than χ.

In the abstract an infinite particle system is a strong Markov process with state space χ or Λ. To prove interesting results we will have to assume somewhat more than this and at a bare minimum we will always assume that the process is described by giving a function $c: \Lambda \times \chi \rightarrow [0, \infty)$ which determined the evolution through the equation:

$$\mathbf{P}(\eta_{t+\delta}(x) \neq \eta_t(x) \text{ for all } x \in A \mid \eta_t = \eta) =$$

$$= \left(\sum_{B \supset A} c(B, \eta) \right) \delta + o(\delta) \quad \text{as } \delta \rightarrow 0. \tag{1}$$

In words $c(A, \eta)$ the rate at which we flip the values at all the sites in A when the configuration is η. In most models that have been studied $c(A, \eta) = 0$ unless $|A|$, the cardinality of A, is 1 or 2, and in nearly all these cases the process is one of the following two types:

(a) *Spin flip systems.* $c(A, \eta) = 0$ unless $|A| = 1$. In these models only one site flips at a time so we write $c(\{x\}, \eta)$ as $c(x, \eta)$.

(b) *Particle motion systems.* $F = \{0, 1\}$ and $c(A, \eta) = 0$ unless $A = \{x, y\}$ and $\eta(x) = 1$, $\eta(y) = 0$ or $|A| = 1$. In these models, ones mark the locations of particles. In the first case a particle jumps from x to y. In the second a particle appears or disappears.

Taking (1) as our way of defining η_t there is immediately the question of whether the transition rates specify a unique Markov process. As the reader might expect the answer is "sometimes yes and sometimes no". Later in the paper we will state conditions on the $c(A, \eta)$ which guarantee that the flip rates specify a unique Markov process but for the moment our purpose is to discuss some examples so we will ask the reader to believe that the examples we will describe below have unambiguous definitions.

We start our consideration of examples with a trivial one:

Example 1. *Independent flips.* $S =$ any set, $F = \{0, 1\}$,

$$c(x, \eta) = \begin{cases} a, & \text{if } \eta(x) = 0, \\ b, & \text{if } \eta(x) = 1. \end{cases}$$

Since the value of the flip rate at x does not depend upon the values of the $\eta(y)$ for $y \neq x$, the coordinate processes $\{\eta_t(x), t \geq 0\}$ $x \in S$ are independent two state Markov chains which flip from 0 to 1 at rate a from 1 to 0 at rate b. The asymptotic behavior of this system as $t \rightarrow \infty$ is trivial to determine: η_t converges weakly (precise definition given below) to the probability measure on $\{0, 1\}^S$ which makes the coordinates independent and have $\mathbf{P}(\eta(x) = 1) = a/a + b$ (we call this distribution the product measure with density $a/a + b$ and denote it $\nu_{a/a+b}$).

Even though the last result is trivial it already indicates some of the difficulties in the subject. If we let η_t and η_t' be two realizations of this process which start from configurations which have $\eta_t(x) \neq \eta_t'(x)$ for infinitely many x, then the distributions of η_t and η_t' are mutually singular for all times and so the theory of Markov chains on a general state space cannot be used even in this trivial example.

Example 1 was trivial because the sites did not interact and hence the fact that S was infinite was irrelevant. Things become considerably more interesting when the sites interact even a little bit:

Example 2. *One sided nearest neighbor systems.* $S = Z^1$, $F = \{0, 1\}$,

$$c(x, \eta) = f(\eta(x), \eta(x + 1)) > 0 \quad \text{for all } x, \eta.$$

Even though these systems are simple it is not yet known whether the processes in this class always have a unique stationary distribution. The answer is known to be yes in three cases:

(a) $f(\eta(x), \eta(x + 1)) = g(\eta(x + 1))$ (or $g(\eta(x + y_1), \ldots, \eta(x + y_n))$, where all $y_i \neq 0$), then $\nu_{1/2}$, product measure with density $\frac{1}{2}$ is the unique stationary distribution (for a proof of this see Section 6).

(b) If $f(0, 1) \geq f(0, 0)$ and $f(1, 0) \geq f(1, 1)$, then the system is an additive process (see Section 4) and the duality theory associated with these processes allows us to show the stationary distribution is unique. The stationary distribution for this process cannot be described in 25 words or less but duality gives a (not very practical) procedure for calculating its finite dimensional distributions.

(c) $f(\eta_1, \eta_2) = 1 + a\eta_1 + b\eta_2 + c\eta_1\eta_2$. where $|a| + |b| + |c| < 1$. Any nonnegative function of η_1 and η_2 is a constant multiple of a function of this type. When a, b, c have special properties the process has a dual in a sense more general than Section 4. Holley and Stroock [79] have shown that the stationary distribution is unique if one of the following conditions are satisfied: (i) $abc < 0$, (ii) $a > 0$, $b, c < 0$, or (iii) $b > 0$, $a, c < 0$. This shows the stationary distribution is unique in at least $\frac{6}{8}$ of the cases.

On the basis of the last two examples and by analogy with what happens when S is finite, the reader might think that if $c(x, \eta) > 0$ for all x, η, then the process is 'irreducible' and will always have a unique stationary distribution. The next example shows that this is not true.

Example 3. *The Ising model.* $S = Z^d$, $F = \{-1, 1\}$,

$$c(x, \eta) = \exp\left(-\beta \sum_{u, \|u\| = 1} \eta(x)\eta(x + u)\right) \quad \beta \geq 0.$$

In the Ising model the sites are thought of as iron atoms whose individual magnetic north pole may be pointed up (1) or down (−1). Since $\eta(x)\eta(x + u) = 1$ if and only if $\eta(x) = \eta(x + u)$, the sum in the exponent measures the extent to which $\eta(x)$ is aligned with its neighbors. If the alignment is bad the flip rate is large and conversely.

The Ising model is not hard to analyze in one dimension. A simple calculation (done in Sections 5 and 6) shows that if μ_β is *the* distribution on $\{0, 1\}^Z$ which makes the coordinates $\eta(n) - \infty < n < \infty$ a Markov chain with transition matrix

$$(e^\beta + e^{-\beta})^{-1} \begin{pmatrix} e^\beta & e^{-\beta} \\ e^{-\beta} & e^\beta \end{pmatrix},$$

then μ_β is a stationary distribution. Holley and Stroock [75] have shown there are no other stationary distributions so in one dimension the stationary distribution is unique for all β.

Things get much more interesting in dimensions $d \geqslant 2$ (Section 8). In these cases there can be more than one stationary distribution. There is a *critical value* β_d so that if $\beta < \beta_d$, the stationary distribution is unique and if $\beta > \beta_d$, it is not. When $d = 2$ it is known that $\beta_2 = 2^{-1} \text{arcsinh}(1) = 0.44$ and if $\beta > \beta_2$, then the set of stationary distributions $\mathcal{I}$ is a convex set with two extreme points μ^+ and μ^- which have $\int f(\eta) \, d\mu^-(\eta) = \int f(-\eta) \, d\mu^+(\eta)$. The last sentence gives a complete description of the size of $\mathcal{I}$ in dimension 2. In contrast very little is known when $d = 3$: it is easy to show $\beta_3 \leqslant \beta_2$ and with some help from Dobrushin [6, 7] or van Bijeren [48] that if $\beta > \beta_2$, $\mathcal{I}$ has an infinite number of extreme points, but we cannot compute the value of β_3 (or to my knowledge even show rigorously that $\beta_3 < \beta_2$) and we are far from determining the structure of $\mathcal{I}$ for $\beta > \beta_3$. There are however precise conjectures about what happens. Based on numerical results it seems that $\beta_3 \approx \frac{1}{2}\beta_2$. As for $\mathcal{I}$, today I think it is reasonable to conjecture that $|\mathcal{I}_e|$, the number of extreme points of $\mathcal{I}$, is given by

$$|\mathcal{I}_e| = \begin{cases} 1 & \beta \in (-\infty, \beta_3], \\ 2 & \beta \in (\beta_3, \beta_2], \\ \infty & \beta \in (\beta_2, \infty) \end{cases}$$

but on other days I have thought that $|\mathcal{I}_e| = \infty$ for all $\beta_3 < \beta$. I will leave it to you to decide which side you want to bet on.

The Ising model (or to be precise its stationary distributions, the Gibbs states) arose in statistical mechanics in 1925 and has been much studied by mathematical physicists. Physics however has not been the only source of models. The last two examples I will mention have socio-political and biological interpretations.

Example 4. *Voter model.* $S = Z^d$, $F = \{0, 1\}$,

$$c(x, \eta) = \sum_{y \in S} p(y) 1_{(\eta(x+y) \neq \eta(x))}$$

where

$$p(y) \geqslant 0 \quad \text{and} \quad \sum_{y \in S} p(y) = 1.$$

In the voter model the sites can be considered to be the homes of individuals who live in a (large) idealized city. The states 1 and 0 represent being for and being against a particular issue or proposition. The individual at x assigns weight $p(y)$ to the opinion of the person at $x + y$ and changes his opinion at a rate which is equal to the sum of the weights for the opposite opinion.

In this system $\eta \equiv 0$ and $\eta \equiv 1$ are absorbing states so it is natural to ask if there are any other stationary distributions? The answer (due to Holley and Liggett [72]) is no if the random walk generated by $\bar{p}(x) = \frac{1}{2}(p(x) + p(-x))$ is recurrent and yes if it is transient. In the transient case they showed that $\mathcal{I}_e$, the extreme points of $\mathcal{I}$ is a one parameter family and in both cases they obtained convergence theorems which allow us to determine the limiting distribution for a large class of initial distribution. (The reader will find a precise statement of these results in Section 3.)

The last example we want to mention is:

Example 5. *Contact processes.* $S = Z^d$, $F = \{0, 1\}$,

$$
c(x, \eta) = \begin{cases} 1, & \text{if } n(x) = 1, \\ k\lambda, & \text{if } \eta(x) = 0 \text{ and } \sum_{y \in D} \eta(x + y) = k, \end{cases}
$$

where D is a finite subset of Z^d.

In a contact process, the sites should be thought of as plants arranged in regular rows or cells in the human body. The states 0 and 1 correspond to the site being 'healthy' or 'infected'. In these terms the flip rates say that infected sites recover at rate 1 while healthy individuals get infected at a rate proportional to the number of sites in $x + D$ which are infected.

Visualizing the contact process as the spread of an infection there are a number of natural questions to ask:

(a) For what parameter values is the process supercritical, i.e. starting from one particle at 0 (i.e. $\eta_0(x) = 1_{\{0\}}(x)$) is there positive probability that $\eta_t \neq 0$ for all t?

(b) If the infection has positive probability of persisting for all time then at what rate does the number of particles grow? Does the infected region have an asymptotic shape?

(c) $\eta \equiv 0$ is an absorbing state. Under what conditions is there a nontrivial stationary distribution?

(d) How are questions (a) and (c) related?

The list of questions could go on and on but I won't let it, since I'm not going to talk about contact processes — there just happens to be a survey paper on contact processes by David Griffeath in this issue (see pp. 151–185).

The list of examples above is just meant to whet your appetite. It does not exhaust the processes I will consider in this paper much less the list of processes which have been studied so far. In this paper I have concentrated on spin flip systems and have

not discussed, except briefly in Section 4, any of the particle motion systems. These processes are discussed extensively in Liggett [86] and Griffeath [64], sources which I would like to enthusiastically encourage the reader to consult for information on this and other topics.

Another obvious omission from the survey is any consideration of work in progress. I have tried to remedy this defect by listing in the bibliography the most recent papers and where they are to appear. If the reader tracks down these papers, he will see what was happening in infinite particle systems now and, hopefully, will enlarge the class of known results.

2. Definitions and preliminaries

In the introduction the reader got a preview of the processes to be considered in Sections 3-8. As is the curse of most aspects of mathematics, before we can have a serious discussion of the subject it is necessary to introduce a fair amount of notation and definitions. This section is dedicated to this task. Hopefully the reader will persevere. As David Williams might say, to visit some of the most beautiful Mayan ruins in Mexico, one must hike for several hours through the jungle.

The first thing to deal with is the state space $\chi = F^S$ or, equivalently, Λ the set of all subsets of S. Since F is a two element set, F is compact (in the obvious topology) and hence χ is compact in the product topology, i.e. the topology of coordinate-wise convergence. To get a mental picture of χ consider (w log) the case $F = \{0, 2\}$. $S = \{1, 2, \ldots\}$ and observe that $\eta \to \sum_{n=1}^{\infty} 3^{-n} \eta(n)$ is a continuous map which sends χ to the usual Cantor set.

Once we have a topology on χ we can define the Borel sets X to be the σ-algebra they generate and $\mathcal{P}(X)$ to be the set of probability measures on X. To define a topology on $\mathcal{P}(X)$ we say that probability measures μ_n converge weakly to a limit μ if $\int f \, d\mu_n \to \int f \, d\mu$ for all $f \in C(\chi)$, the continuous functions on χ. It is well known that $\mathcal{P}(X)$ is compact in this topology, a fact which will be useful below.

Our next concern is the flip rates $c(A, \eta)$ which define the evolution. The flip rates were defined above as functions from $\Lambda \times \chi \to [0, \infty)$ but we will consider only spin flip systems on Z^d so we will reduce ourselves to that level of generality now and write the flip rates as $c(x, \eta)$ instead of $c(\{x\}, \eta)$. The first step in constructing η_t from c is to define the infinitesimal generator

$$Lf(\eta) = \sum_x c(x, \eta)(f(\eta^x) - f(\eta)), \tag{1}$$

where $\eta^x \in \chi$ is the configuration with the spin at x flipped

$$\eta^x(y): \begin{cases} \neq \eta(x), & \text{if } y = x, \\ = \eta(y), & \text{otherwise} \end{cases}$$

and, to make sure that the sum in the definition of $Lf(\eta)$ converges, we assume $f \in C_0(\chi)$ the functions which only depend upon finitely many coordinates.

Having defined L on a subset of $C(\chi)$ which is dense (using the sup norm on $C(\chi)$) the next step is to apply the Hille–Yosida theorem to show that L (or, to be precise, its closure) is the generator of a semigroup S_t, $t \ge 0$ on $C(\chi)$ which defines a unique Markov process through the formula $S_t f(\eta) = \mathbf{E}(f(\eta_t) \mid \eta_0 = \eta)$. Needless to say this cannot be done without some assumptions on c. Liggett [85] has shown (see Liggett [86] for a more recent treatment) that if we assume c satisfies

(i) *translation invariance*: if $y \in Z^d$ and $\theta_y \eta$ has $\theta_y \eta(x) = \eta(x + y)$, then $c(x + y, \theta_y \eta) = c(x, \eta)$ and

(ii) *Lipschitz continuity*: if we let $\|g\|_\infty = \sup_\eta |g(\eta)|$ and $\|f\| = \sum_x \|f(\eta^x) - f(\eta)\|_\infty$, then $\|c(0, \cdot)\| < \infty$, then L generates a unique Markov semigroup S_t and we have

$$\frac{\mathrm{d}}{\mathrm{d}t} S_t f(\eta) = L S_t f(\eta) = S_t L f(\eta), \tag{2}$$

$$\|S_t f\| \le \mathrm{e}^{(\alpha - \beta)t} \|f\|, \tag{3}$$

where $\alpha = \|c(0, \cdot)\|$ and $\beta = \inf_\eta |c(0, \eta) + c(0, \eta^x)|$.

This result requires some explanations:

Remark 1. The assumption that $c(0, \cdot)$ is continuous is insufficient to obtain the desired conclusion. There are strictly positive continuous attractive translation invariant flip rates for which the conclusion is false, see Gray [62].

Remark 2. Assumptions (i) and (ii) hold whenever $c(x, \eta)$ has finite range, that is, $c(x, \eta) = g(\eta(x + x_1), \ldots, \eta(x + y_n))$, where $\{x_1, \ldots, x_n\}$ is some finite subset of Z^d.

Remark 3. It is only fair to mention that there are other approaches to constructing finite range and more general infinite particle systems which can be found in Harris [65], and Holley and Stroock [74]. The first is a direct construction of η_t, $0 \le t \le t_0$ using random islands, the second is a martingale approach.

Liggett's result guarantees the existence of all the infinite particle systems we will consider so we turn our attention now to the terminology required to discuss the two basic problems of infinite particle systems

(i) determining $\mathcal{I}$, the class of stationary distributions, and

(ii) computing the limiting distribution starting from any initial distribution.

If $\mu \in \mathcal{P}(X)$, let η_t^μ be a version of the process with initial distribution μ and let μS_t be the distribution of η_t^μ. In order for $\mu \in \mathcal{I}$ it is necessary and sufficient that for $f \in C_0(\chi)$,

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} \mu S_t f(\eta)\big|_{t=0} = \int L f(\eta) \, \mathrm{d}\mu(\eta) \tag{4}$$

(see Liggett [86] for details). If one writes out the right-hand side one sees that the finite dimensional distributions of μ satisfy an infinite system of linear equations.

(Exercise: do this for the basic contact process.) These equations give some information about μ but usually not enough to determine it so we have to resort to other tricks to determine $\mathscr{I}$.

The reader will see in Sections 3–8 that these tricks vary considerably from example to example. The situation is not total chaos, however there are some common themes (coupling, duality, monotonicity) and even several general results. The most basic is that $\mathscr{I}$ is a non-empty convex set. $\mathscr{I}$ is clearly convex. To show $\mathscr{I} \neq \emptyset$ let $\mu_n = (1/n) \int_0^n \mu S_t \, dt$, pick a convergent subsequence μ_{n_k}, and observe that if μ is the weak limit, then $\mu S_t = \mu$ for all t (computation left to the reader). The second general result I want to mention is a consequence of (3). If $\alpha < \beta$ (i.e. $c(0, \eta) \equiv c$ or almost constant), then $\alpha - \beta < 0$ so $\|S_t f\| \to 0$. By the previous result $\mathscr{I} \neq \emptyset$. If $\mu \in \mathscr{I}$, then $\mu S_t f$ is constant and since $\|S_t f\| \to 0$ we have $\|S_t f - \mu S_0 f\|_\infty \to 0$ showing that $\mathscr{I} = \{\mu\}$ and μ is the limit starting from any initial distribution. Other sufficient conditions for $|\mathscr{I}| = 1$ and for $|\mathscr{I}| > 1$ can be found in Sections 4 and 8.

3. The voter model

In this section we will study the voter model:

$$c(x, \eta) = \sum_y p(y) 1_{\{\eta(x+y) \neq \eta(x)\}},$$

where

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

The first step is to give a special construction of the process:

- Let $\{N_z(t): t \geq 0\}$, $z \in Z^d$ be independent Poisson processes with rate 1.
- Let $T_{z,n} = \inf\{t \geq 0: N_z(t) = n\}$ be the time of the nth event in N_z.
- Let $\{Y_{z,n}: n \geq 1\}$, $z \in Z^d$ be independent i.i.d. sequences with the property that $P(Y_{z,n} = y) = p(y)$ for all y, z, n.

In making these definitions we have in mind that at time $T_{z,n}$ the voter at z decides for the nth time to change his mind, he picks a neighbor $z + Y_{z,n}$ at random and adopts the opinion of that neighbor (which may be the same as his own). Since the voter only changes opinion when he picks a y with $\eta(z + y) \neq \eta(z)$ it is easy to see that we get a process having the flip rates given above.

To use the recipe given above to compute the state of the process at time t is not completely trivial because there are infinitely many flips in any positive amount of time. To carry out the construction we will use the graphical representation invented by Harris [69] and developed by Griffeath [64]. Draw the family of line segments $\{z\} \times [0, t]$, $z \in Z^d$. Mark the points $(z, T_{z,n})$, $z \in Z^d$, $n \geq 1$, with δ's and draw an arrow from $(z + Y_{z,n}, T_{z,n})$ to $(z, T_{z,n})$ (see Fig. 1 for a picture). The δ indicates that the voter at z has decided to change his mind and the arrow indicates the neighbor he chooses to imitate.

To construct the process from this 'percolation structure' we imagine fluid entering the bottom at the points, where $\eta_0(x) = 1$ and flowing up the structure – the δ's being dams and the arrows being pipes which allow the fluid to flow in the indicated direction. With this interpretation the $\{x: \eta_t(x) = 1\}$ is the set of wet sites at height t. To make this definition mathematical we say that there is a path from $(x, 0)$ to (y, t) if there is a sequence of times $0 < s_1 < s_2 \cdots < s_n < t$ and spatial locations $x_0 = x, x_1, x_2, \ldots, x_n = y$ so that

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

(ii) the vertical segments $\{x_i\} \times (s_i, s_{i+1})$, $i = 0, 1, \ldots, n$ $(s_0 = 0, s_{n+1} = t)$ do not contain any δ's.

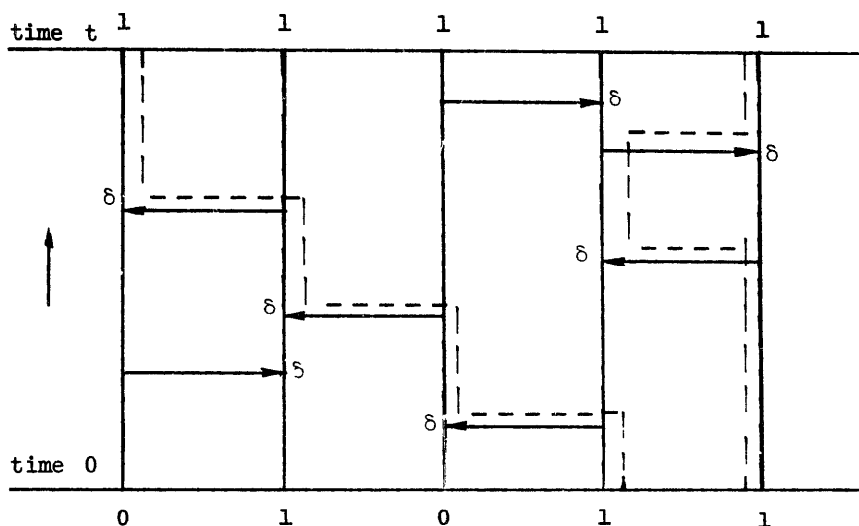(On Fig. 1 we have indicated some sample paths.)

Fig. 1. One realization of the graphical representation.

When there is a path from $(x, 0)$ to (y, t) it follows that the individual at time t has the same opinion as individual x at time 0. Since every individual at time t has the same opinion as some individual at time 0, it follows that if A is the set of individuals at time 0 who have opinion 1, then the set of individuals at time t which have that opinion is given by

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

The last definition gives the 'graphical representation' of the voter model. One of the nice aspects of this construction is that it defines all the ξ_t^A on the same probability space so that we have

monotonicity: if $A \subset B$, then $\xi_t^A \subset \xi_t^B$ for all t.   (1)

additivity: for any A, B, and t, $\xi_t^{A \cup B} = \xi_t^A \cup \xi_t^B$.   (2)

These two properties are very useful and we will consider them later. For computing the asymptotic behavior of the voter model the most important fact about the graphical representation is that it allows us to construct a dual process $\hat{\xi}_t^B$ which has the property that for any A, B

$$\mathbf{P}(\xi_t^A \cap B \neq \emptyset) = \mathbf{P}(\hat{\xi}_t^B \cap A \neq \emptyset). \tag{3}$$

One process which satisfies (3) is

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

for then

$$\{\xi_t^A \cap B \neq \emptyset\} = \left\{ \begin{array}{c} \text{for some } x \in A, y \in B \text{ there is} \\ \text{a path from } (x, 0) \text{ to } (y, t) \end{array} \right\} = \{\zeta_t^B \cap A \neq \emptyset\}.$$

The sample paths of ζ_t^B have huge jumps (consider a contact process) and ζ_t^B is not Markov so we will replace ζ_t^B by a tamer process $\hat{\xi}_t^B$ which has the same one dimensional distributions (and hence satisfies (3)).

To construct $\hat{\xi}_t^B$ $t \geq 0$ it suffices to give the distribution of $\hat{\xi}_s^B$, $0 \leq s \leq t$ for all t. To do this we define a dual percolation structure $\hat{\mathcal{P}}_t$ by reversing the arrows in the original structure $\mathcal{P}$ and changing time by the mapping $\hat{s} = t - s$ (see Fig. 2). Since the distribution of a Poisson process is unchanged by time reversal it is clear that
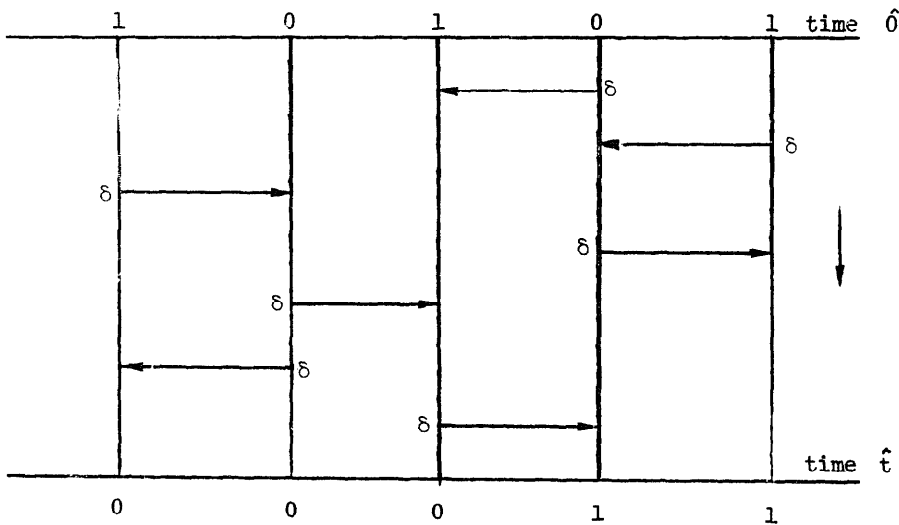

Fig. 2. The dual percolation structure of the example in Fig. 1.

(i) the finite dimensional distributions we have defined are consistent,

(ii) $\xi_t^B \overset{d}{=} \zeta_t^B$ for any t and

(iii) the dual process could be constructed from a single percolation structure $\hat{\mathcal{P}}$ with gadgets which are obtained from the original ones by reversing the arrows.

In order for (3) to be useful we need to know how $\hat{\xi}_t^B$ evolves (and if we want to get results, the evolution has to be simple). When ξ_t is the voter model the dual

percolation structure has 'δ-arrows' from x to $x + y$ at rate $p(y)$. By considering the four possible cases one sees that the effect of δ-arrow is given by the following table:

	before		after	
	$\xi(x)$	$\xi(x+y)$	$\xi'(x)$	$\xi'(x+y)$
	1	1	0	1
	1	0	0	1
	0	1	0	1
	0	0	0	0

$\xi(x) \quad \xi(x+y)$

$\delta \vdash\!\!\longrightarrow \quad \downarrow$ time

$\xi'(x) \quad \xi'(x+y)$

If we think of the 1's as particles and 0's as empty sites, then $\hat{\xi}_t$ may be described as a coalescing random walk. A δ-arrow from x to $x + y$ causes a particle at x to jump to $x + y$ and if $x + y$ is occupied the two particles coalesce to 1.

When the dual process starts from $B = \{0\}$ its behavior is trivial to compute. For any $t \geq 0$, $\hat{\xi}_t^0$ (our abbreviation for $\hat{\xi}_t^{\{0\}}$) has only one element and its position X_t is a random walk which takes steps with distribution $p(y)$ at rate 1. Combining this observation with the duality equation (3) shows

$$\mathbf{P}(0 \in \xi_t^A) = \mathbf{P}(X_t \in A). \tag{4}$$

The last formula allows us to compute the one-dimensional distributions of ξ_t^A and implies in particular that if the initial distribution is ν_ρ product measure with density ρ, then $\mathbf{P}(0 \in \xi_t) = \rho$, for all t. To determine the limiting behavior of ξ_t in this case we compute the two dimensional distributions. Let $\xi_t^A(y) = 1$, if $y \in \xi_t^A$ and 0 otherwise. Now

$$\mathbf{P}(\xi_t^A(0) \neq \xi_t^A(x)) \leq \mathbf{P}(\hat{\xi}_t^0 \neq \hat{\xi}_t^x)$$

and from the graphical representation of $\hat{\xi}_t$ we have $\hat{\xi}_t^0 \cup \hat{\xi}_t^x = \hat{\xi}_t^{\{0,x\}}$ so

$$\mathbf{P}(\xi_t^A(0) \neq \xi_t^A(x)) \leq \mathbf{P}(|\hat{\xi}_t^{\{0,x\}}| = 2).$$

The particles which make up $\hat{\xi}_t^0$ and $\hat{\xi}_t^x$ perform independent random walks until they hit so if X_t and X_t' are independent random walks with distribution p and $X_0 = 0$, $X_0' = x$, then

$$\mathbf{P}(|\hat{\xi}_t^{\{0,x\}}| = 2) = \mathbf{P}(X_s - X_s' \neq 0 \text{ for all } 0 \leq s \leq t).$$

$\bar{X}_s = X_s - X_s'$ is itself a random walk which takes steps with distribution $\bar{p}(x) = \frac{1}{2}(p(x) + p(-x))$ at rate 2. If $\bar{X}_s$ is recurrent, then

$$\mathbf{P}(\xi_t^A(0) \neq \xi_t^A(x)) \to 0 \quad \text{as } t \to \infty. \tag{5}$$

So the system approaches total consensus: that is, if A is a any initial configuration and B a finite set

$$\mathbf{P}(\xi_t^A(0) = \xi_t^A(x) \text{ for all } x \in B) \to 1. \tag{6}$$

If $\bar{X}$ is transient $\mathbf{P}(\hat{\xi}_t^0 \neq \hat{\xi}_t^x)$ does not converge to 0 so differences of opinion may

persist. If ξ_t^ρ is a voter model with initial distribution ν_ρ, then

$$\mathbf{P}(\xi_t^\rho(0) \neq \xi_t^\rho(x)) = 2\rho(1-\rho)\mathbf{P}(\hat{\xi}_t^0 \neq \hat{\xi}_t^x)$$

which converges to a nonzero limit as $t \to \infty$. With a little more work we can show that the finite dimensional distributions converge to those of a limit μ_ρ which is a stationary distribution for the voter model;

$$\mathbf{P}(\xi_t^\rho \cap B = \emptyset) = \int \mathbf{P}(\xi_t^A \cap B = \emptyset) \, d\nu_\rho(A)$$

$$= \int \mathbf{P}(\hat{\xi}_t^B \cap A = \emptyset) \, d\nu_\rho(A) = \mathbf{E}((1-\rho)^{|\hat{\xi}_t^B|})$$

Since $|\hat{\xi}_t^B|$ can only decrease, it follows that if B is finite, then as $t \uparrow \infty$ $|\hat{\xi}_t^B|$ decreases to a limit $|\hat{\xi}_\infty^B|$ and

$$\mathbf{P}(\xi_t^\rho \cap B = \emptyset) \to \mathbf{E}((1-\rho)^{|\hat{\xi}_\infty^B|}) \tag{7}$$

proving the asserted convergence of finite dimensional distributions.

At this point we have constructed a one parameter family $\{\mu_\rho : 0 \leqslant \rho \leqslant 1\}$ of stationary distributions so it is natural to ask if $\mathcal{I}$ is the closure (in the weak topology) of the convex hull of the μ_ρ ($\mathcal{I}$ certainly must contain this set of measures). Holley and Liggett [72, pp. 659–660] have shown that the anser to this question is yes. The idea behind their proof is to show that if $\mu \in \mathcal{I}$ and $B_{k,n}$ is a sequence of k element sets with $\inf\{\|x-y\| : x, y \in B_{k,n}, x \neq y\} \to \infty$, then

$$\lim_{n \to \infty} \mu(\eta : \eta \cap B_{k,n} = \emptyset) = \rho_k \tag{8}$$

and the numbers ρ_k satisfy

$$\sum_{r=0}^{m} \binom{m}{r} (-1)^r \rho_{k+r} \geqslant 0 \quad \text{for all } k, m \geqslant 0. \tag{9}$$

Looking at Theorem 2 in [99, Vol. II, p. 223] we see that these inequalities imply that there is a probability measure γ on $[0, 1]$ so that

$$\rho_k = \int_0^1 (1-\rho)^k \gamma(dp). \tag{10}$$

Since $\lim_{n\to\infty} \mu_\rho\{\eta : \eta \cap B_{k,n} = \emptyset\} = (1-\rho)^n$ the last equality suggests that $\mu(S) = \int \mu_\rho(S)\gamma(dp)$ for all measurable subsets S (and proves that if μ is some convex combination of the μ_ρ it must be this one). By investigating the function

$$h(F) = \mu(\eta : \eta \cap F = \emptyset)$$

which is harmonic for the process $\hat{\xi}_t^B$ (i.e. $h(B) = \mathbf{E}h(\hat{\xi}_t^B)$) Holley and Liggett prove this and thus conclude that all $\mu \in \mathcal{I}$ can be written as

$$\mu(\cdot) = \int \mu_\rho(\cdot)\gamma(dp). \tag{11}$$

It is easy to show using (10) that such a decomposition is unique so we have shown that $\mathcal{I}$ is a convex set with extreme points $\{\mu_p: 0 \leqslant p \leqslant 1\}$.

Having determined the set of stationary distributions the next step is to investigate the convergence of ξ_t^ν (which denotes the process with initial distribution ν). On this subject Holley and Liggett have the following remarkable result [72, Theorem 5.16, p. 660]:

(12) Theorem. Let $p_t(i, j)$ be the transition density of the random walk X_t. $\xi_t^\nu \Rightarrow \mu_p$ as $t \to \infty$ if and only if

(a) $\lim\limits_{t\to\infty} \sum\limits_j p_t(i, j)\nu\{\eta: \eta(j) = 0\} = 1 - \rho$

and

(b) $\lim\limits_{t\to\infty} \sum\limits_{j,k} p_t(i, j)p_t(i, k)\nu\{\eta: \eta(j) = 0, \eta(k) = 0\} = (1-\rho)^2$.

If ν is translation invariant and $\nu\{\eta: \eta(0) = 1\} = \rho$, then (a) holds. If in addition the coordinates $\eta(x)$ form an ergodic stationary sequence under ν, then (b) holds, so $\xi_t^\nu \Rightarrow \mu_p$ as $t \to \infty$.

The results above describe the stationary distributions for the voter model and the limiting behavior of the finite dimensional distributions. In the recurrent case the limit is total consensus so it is natural to inquire about the rate at which this is approached. In one dimension when $p(1) = p(-1) = \frac{1}{2}$ this is an easy problem. When $\xi_t^0 \neq \emptyset$ it is always an interval $[a_t, b_t]$ and the end points perform independent simple random walks until $b_t < a_t$ at which time ξ_t^0 becomes $\emptyset$. From this it follows easily that if $p_t = \mathbf{P}(\xi_t^0 \neq \emptyset)$, then $p_t \sim 1/\sqrt{\pi t}$ and

$$\mathbf{P}(p_t|b_t - a_t| > x|\xi_t^0 \neq \emptyset) \to e^{-x^2/2}$$

the Rayleigh distribution (see Durrett [56] for the relvant conditional limit theorems).

In dimensions $d \geqslant 2$ the situation is much harder. By exploiting the connection with a model studied by Sawyer [93] and others [57, 61, 78, 83, 89], Bramson and Griffeath [55] have shown that if $p(x) = 1/2d$ when $\|x\| = 1$, then as $t \to \infty$

$$p_t \sim \begin{cases} (\log t)/\pi t, & d = 2, \\ C_d/t, & d \geqslant 3, \end{cases}$$

and $(p_t|\xi_t^0\|\ |\xi_t^0 \neq \emptyset)$ converges to an exponential distribution with mean 1.

The results above describe how one person's opinion dies out and are designed to complement the trivial convergence theorems in the recurrent case. In the transient case we have a one-parameter family of stationary distributions which have complicated dependencies and are given by a horrible formula, so there is also the problem of describing the stationary distribution. One approach to this (which was motivated by similar activity in statistical mechanics) is called renormalization. In this approach one studies the asymptotic behavior of $S_n(\xi)$, the number of points in $\xi \cap [-n, n]^d$,

and other related weighted sums and then after an appropriate amount of technical intercourse renormalizes to obtain a limiting random field.

Bramson and Griffeath [53] have done this for the three dimensional voter model with the result that $S_n(\xi)/n^{5/2}$ converges to a normal distribution and the equilibrium state under appropriate renormalization approaches the 0 mass free field. Although the limits in these results are what we get from independent variables the presence of the norming constant $n^{5/2}$ instead of $n^{3/2}$ indicates that the equilibrium states have long range dependencies and gives an estimate of the strength.

For more information on these points see Griffeath [64, pp. 53–54], Bramson and Griffeath [53] or Major [88] who has given another treatment of the voter model. A number of infinite particle systems have been renormalized. The reader who is interested in these results should start with Holley and Stroock [76] and then progress to [77] and [78].

4. Additive processes

In this section we will consider a collection of models called additive processes which can be constructed using graphical representations similar to the one used in the last section. The motivation for considering this class is that all additive processes have an associated dual process $\hat{\xi}_t$ which is related to the original process ξ_t by the duality equation

$$P(\xi_t^A \cap B = \emptyset) = P(\hat{\xi}_t^B \cap A = \emptyset). \tag{1}$$

In the last section we saw how this relationship could be used to study the stationary distributions and limiting behavior of the voter model. In this section we will describe some other cases in which the duality equation (1) can be used to study the evolution. We will not dwell too long on these examples, since our main aim is to show that any additive process with positive translation invariant flip rates has a unique stationary distribution which is the limit as $t \to \infty$ starting from any initial distribution.

In the abstract, an additive process is a process for which it is possible to construct a family of realizations ξ_t^A, $A \subset S$ on the same probability space in such a way that we have

$$\xi_t^A \cup \xi_t^B = \xi_t^{A \cup B}. \tag{2}$$

Now when Harris introduced additive processes in [69] he showed that (2) implied ξ_t could be constructed in a special way (the graphical representation described below) which made it easy to define a process $\hat{\xi}_t$ satisfying (1). Rather than adopting this viewpoint and attempting an exhaustive description of the set of additive processes, we will introduce the collection, through a sequence of examples which can be combined to obtain the most general additive spin flip system (see Section 2 of Harris [69] for details).

All the examples we will consider (and in fact all additive processes) can be constructed by using a percolation structure constructed from a set of independent Poisson processes $\{N_{z,i}(t), t \geq 0\}$, $z \in S$, $i \in I = \{0, 1, 2, \ldots\}$ with $EN_{z,i}(t) = r_i t$ and $\sum_i r_i < \infty$. What we do at the Poisson arrival times $T_{z,i,n} = \inf\{t \geq 0: N_{z,i}(t) = n\}$ will depend upon the particular example. In the last section we studied

Example 1. *The voter model* (here and below $S = Z^d$, $F = \{0, 1\}$).

$$c(x, \eta) = \sum_y p(y) 1_{(\eta(x+y) \neq \eta(x))}.$$

To write the definitions of the last section in the present notation let $y_1, y_2, \ldots$ be an enumeration of Z^d and at time $T_{z,i,n}$ draw an arrow from $z + y_i$ to z and label z with a δ.

Example 2. *The basic contact process.*

$$c(x, \eta) = \begin{cases} 1, & \text{if } \eta(x) = 1, \\ k\lambda, & \text{if } \eta(x) = 0 \text{ and } \sum_{y \in G} \eta(x+y) = k, \end{cases}$$

where $G = \{g_1, \ldots, g_n\}$ is some finite set. At times $T_{z,0,n}$ we put a δ at z and at times $T_{z,i,n}$ $i \geq 1$ we draw an arrow from $z + g_i$ to z. If we define path and ξ_t^A as we did in Section 3, then the effect of a δ is to kill any particle at z and an arrow from $z + g_i$ to z spreads the infection to z if it is present at $z + g_i$, so we have defined a process with the flip rates given above.

Example 3. *Contact processes with more general birth rates.* Suppose we modify the flip rates of the contact process so that

$$c(x, \eta) = \begin{cases} 1, & \text{if } \eta(x) = 1, \\ \lambda_k, & \text{if } \eta(x) = 0 \text{ and } \sum_{y \in G} \eta(x+y) = k, \end{cases}$$

where the λ_k are constants ≥ 0 with $\lambda_0 = 0$. When $|G| = n$, the new system is additive if (and only if)

$$\sum_{r=0}^{k} (-1)^{1+r} \binom{k}{r} \lambda_{n-k+r} \geq 0 \quad \text{for } 1 \leq k \leq n$$

(when $n = 2$ this says $\lambda_1 \leq \lambda_2$ and $2\lambda_1 \geq \lambda_2$). To construct these examples we let $N = 2^n - 1$ and $G_1, \ldots, G_N$ be the N nonempty subsets of G and for $1 \leq i \leq N$ draw arrows at time $T_{z,i,n}$ from each of the points in $z + G_i = \{z + g: g \in G_i\}$ to z. If we use the previous definition of path, then this cluster of arrows has the effect of making $\xi(x) = 1$ if $\xi(y) = 1$ for some $y \in z + G_i$. If G is a two element set, then the G_i we have at our disposal are $\{g_1\}, \{g_2\}, \{g_1, g_2\}$. By symmetry $\{g_1\}$, and $\{g_2\}$ must have the same rate (say a) so we can construct any contact process with $\lambda_1 = a + b$, $\lambda_2 = 2a + b$, where $a, b \geq 0$, i.e. $\lambda_1 \leq \lambda_2 \leq 2\lambda_1$. When $|G| > 2$ similar reasoning and more complicated calculations lead to the general criteria given above.

In the last example we generalized the birth rates. As the reader can imagine we can also generalize the class of death rates by introducing gadgets which combine a δ at z with a collection of arrows from points in $z + G$ to z. (The effect of this gadget is to kill a particle at z if and only if there is one present at z and all the sites in $z + G$ are empty.) Once we have done this we have all the ingredients for constructing just about the most general additive process.

Our quest for the most general additive process is not an idle quasialgebraic curiosity. Additive processes have special properties which make them easier to study. The most important of these is the fact that the graphical representation of additive processes given above allows us to associate with each additive process a dual process, $\hat{\xi}_s$ which satisfies (1). This process is defined in exactly the same way as the dual of the voter model (the discussion following (3) in Section 3 gives the general definition).

Now that we have constructed a process $\hat{\xi}$ so that ξ and $\hat{\xi}$ satisfy (1), the question is "What can we do with $\hat{\xi}$?" As I have already mentioned several times we can analyze the voter model, but this is a special situation. In the voter model the number of particles in the dual process does not increase and this special property was important for our solution. The basic contact process (Example 2) illustrates another special case: $\hat{\xi}$ has the same distribution as ξ. In this special case the analysis is not so simple and after six years of work by various people we finally know what $|\mathcal{I}|$ is for all values of λ (see Griffeath's survey). When we generalize the contact process to get Example 3 we lose the self duality and things become more complicated. It has been possible to prove a few things about these processes but there are many basic questions which remain unanswered. The reader should see Chapter 2 of Griffeath [64] for information on this point and for some more examples which can be solved by duality.

Up to this point we have considered only systems in which $c(x, \eta) \equiv 0$ when $\eta \equiv 0$ (there are no paths if there are no starting points). To allow for 'creation from nothing' we have to introduce a new collection of independent Poisson processes $\{N_{z,\beta}(t), t \geq 0\}$, $z \in S$, with rate r_β. At the time $T_{z,\beta,n} = \inf\{t \geq 0: N_{z,\beta}(t) = n\}$ we write a β at z and view the point β as a possible starting point for paths. With this in mind we let $\mathcal{B}_t = \{(z, T_{z,\beta,n}): T_{z,\beta,n} \leq t\}$ and change the definition of the process to

$$\xi_t^A = \{y: \text{ there is a path to } (y, t) \text{ from some } (x, 0) \text{ with } x \in A \text{ or from some}$$
$$(z, s) \in \mathcal{B}_t\}.$$

With this new definition of ξ_t^A the effect of a β at x is to make $\xi(x) = 1$. With the introduction of β we can construct some contact processes with $\lambda_0 > 0$ (the condition above is still necessary and sufficient) as well as some other models.

Example 4. *The voter model with defections.*

$$c(x, \eta) = (a + b\eta(x)) + \sum_y p(y) 1_{(\eta(y+x) \neq \eta(x))},$$

where a, $a + b$, $p(y) \geq 0$ and $\sum_y p(y) = 1$. In this version the individuals change from opinion $0 \to 1$ at rate a and from $1 \to 0$ at rate $a + b$ even when they agree with all their

neighbors. To construct this process using the graphical representation, let y_i be an enumeration of Z^d, let $r_\beta = a$, $r_0 = a + b$, and $r_i = p(y_i)$ for $i \geq 1$. At times $T_{z,\beta,n}$ draw a β at z, at times $T_{z,0,n}$ draw a δ at z and at times $T_{z,i,n}$ for $i \geq 1$ follow the rules for the ordinary voter model (Example 1).

When there are β's in the graphical representation (1) is no longer correct since we are ignoring the paths which start in $\mathcal{B}_t$. Let $\Omega_{B,t} = \{$there is a path from $\mathcal{B}_t$ to $B \times \{t\}$ in $\mathcal{P}\}$. Since $\xi_t^A \cap B \neq \emptyset$ on $\Omega_{B,t}$ and the argument for the case with no β's holds on $\Omega_{B,t}^c$ we have

$$\mathbf{P}(\xi_t^A \cap B = \emptyset) = \mathbf{P}(\xi_t^A \cap B = \emptyset, \Omega_{B,t}^c) = \mathbf{P}(\hat{\xi}_t^B \cap A = \emptyset, \Omega_{B,t}^c).$$

If we let $\hat{\mathcal{B}}_t$ be the birth set in $\hat{\mathcal{P}}$ and $\gamma^B = \inf\{s$: there is a path from $B \times \{0\}$ to $\hat{\mathcal{B}}_t$ in $\hat{\mathcal{P}}\}$, then it follows from the definitions that

$$\mathbf{P}(\hat{\xi}_t^B \cap A, \Omega_{B,t}^c) = \mathbf{P}(\hat{\xi}_t^B \cap A, \gamma^B > t)$$

so we have

$$\mathbf{P}(\xi_t^A \cap B = \emptyset) = \mathbf{P}(\hat{\xi}_t^B \cap A, \gamma^B > t). \tag{1'}$$

The duality relationship (1') is not as aesthetically pleasing as (1), but it is much more useful, because it allows us to prove the result we mentioned earlier:

> If ξ_t is an additive process in which β's occur at a rate $r > 0$, then there is a unique stationary distribution π and for any A, B (3)

$$|\mathbf{P}(\xi_t^A \cap B = \emptyset) - \pi(\eta: \eta \cap B = \emptyset)| \leq e^{-rt}.$$

Proof. Let $\tau^B = \inf\{t: \hat{\xi}_t^B = \emptyset\}$. From (1') we have

$$\mathbf{P}(\xi_t^A \cap B = \emptyset) = \mathbf{P}(\hat{\xi}_t^B \cap A = \emptyset, \gamma^B > t)$$

$$= \mathbf{P}(\tau^B \leq t, \gamma^B > \tau^B) + \mathbf{P}(\hat{\xi}_t^B \cap A = \emptyset, \tau^B > t, \gamma^B > t).$$

When $\hat{\xi}_t^B \neq \emptyset$ there is always at least one particle which is being showered with β's at rate r so

$$\mathbf{P}(\tau^B > t, \gamma^B > t) \leq P(\gamma^B > t | \tau^B > t) \leq e^{-rt}.$$

If we let $\pi(\eta: \eta \cap B = \emptyset) = \mathbf{P}(\tau^B < \gamma^B)$, then it follows from the last equality above that

$$\mathbf{P}(\xi_t^A \cap B = \emptyset) - \pi(\eta: \eta \cap B = \emptyset) =$$

$$= -\mathbf{P}(t < \tau^B < \gamma^B) + \mathbf{P}(\hat{\xi}_t^B \cap A = \emptyset, \tau^B > t, \gamma^B > t).$$

Since the terms on the right-hand side are of opposite sign and each is smaller in absolute value than $\mathbf{P}(\tau^B > t, \gamma^B > t)$, we have proved (3).

As an immediate consequence of (3) we get that the voter model with defections has a unique stationary distribution (in any dimension!) or more generally that this is the case whenever we add spontaneous births to an additive process.

Given the last result the reader can perhaps understand our desire to find the class of all additive processes and to generalize duality. Holley and Stroock [79] were the first to do the latter by introducing 'parity' into the framework above (this idea incidentally was due to Matloff [89]). A portion of this duality is described in Griffeath [64] using the graphical representation described above and counting paths mod 2 to define the process.

$$\eta_t^A = \{y: \text{ there are an odd number of paths up to } (y, t) \text{ from } (x, 0) \text{ with}$$
$$x \in A\}.$$

The duality theory for these 'cancellative systems' is much different from the additive variety and is not very well developed. This is, I think, a fertile area for research which will attract attention in the near future.

In closing this section I would like to emphasize a point I made in passing earlier: duality is not tied to a graphical representation. It can be expressed as a relation between expectations

$$E_{\xi_0} f(\xi_t, \hat{\xi}_0) = E_{\hat{\xi}_0} f(\xi_0, \hat{\xi}_t)$$

which holds for *one* particular choice of f: ($f(A, B) = 1_{(A \cap B = \emptyset)}$ in the additive case.) This viewpoint has been generalized recently and used by various one and two element subsets of {Holley Liggett and Spitzer} (see [97, 87, 73]).

5. The exponential family

In this section we will consider a collection of infinite particle systems with $S = Z^d$ and $F = \{-1, 1\}$ which generalize the stochastic Ising model (Example 3 of the introduction). In these systems the flip rates are described by giving a potential J which is a real valued function defined on the finite subsets of Z^d and has the following properties:

(i) *translation invariance*: if $x \in Z^d$, $A \subset Z^d$ and $x + A = \{x + y: y \in A\}$, then $J(A) = J(x + A)$;

(ii) *finite range*: there is an $L < \infty$ so that if $0 \in A$ and $J(A) \neq 0$, then $A \subset \{x: \|x\| \leq L\}$ (the smallest such L in the range of J).

The flip rates are defined from the potential by

$$c(x, \eta) = \exp\left(-\sum_{A \ni x} J(A)\eta(A)\right), \quad \text{where } \eta(A) = \prod_{y \in A} \eta(y).$$

The flip rates in (1) may look mysterious so some interpretation is called for. When these flip rates are used in statistical mechanics the exponent (including the minus sign) represents the energy in the configuration which is due to the interaction of x

with the rest of the system. The term in the sum with $A = \{x\}$ represents the interaction with an external magnetic field of strength $|J(\{x\})|$. The terms with $A = \{x, y\}$ take into account the interactions between pairs of spins. The quantity $\eta(x)\eta(y) = 1$ if and only if $\eta(x) = \eta(y)$ so if there are only two particle interactions present (e.g. the Ising model) the energy is a measure of the relative alignment of sites. If $J(\{x, y\}) > 0$ (the attractive case discussed in Section 7) the flip rate is smaller when $\eta(x)$ and $\eta(y)$ are aligned. If $J(\{x, y\}) < 0$ the flip rate is smaller when they are opposite.

The flip rates in (1) though curious looking have some very special properties. The most important of these is the fact that if we consider approximating processes with flip rates

$$c_n(x, \eta) = c(x, \eta)1_{\{\|x\| \leqslant n\}},\tag{2}$$

then the stationary distributions of this system are easy to write down (after enough notation is introduced!).

Let $\Lambda_n = \{x : \|x\| \leqslant n\}$ and let $\Lambda_n^c = Z^d - \Lambda_n$. In the process with flip rates c_n the sites $x \in \Lambda_n^c$ do not flip, so for a fixed $\zeta \in \{-1, 1\}^{\Lambda_n^c}$ the system reduces to a Markov chain with state space $\{-1, 1\}^{\Lambda_n}$. If for $\xi \in \{-1, 1\}^{\Lambda_n}$, we let η be the configuration obtained by combining ξ and ζ,

$$\eta(y) = \begin{cases} \xi(y), & y \in \Lambda_n, \\ \zeta(y), & y \in \Lambda_n^c, \end{cases}\tag{3}$$

then I claim that

$$w_{n,\zeta}(\xi) = \exp\left(\sum_{A \cap \Lambda_n \neq \emptyset} J(A)\eta(A)\right)\tag{4}$$

defines a stationary measure for the Markov chain with state space $\{-1, 1\}^{\Lambda_n}$.

To check this claim let η^x denote the configuration in which the site at x is flipped:

$$\eta^x(y) = \begin{cases} \eta(y), & y \neq x, \\ -\eta(x), & y = x. \end{cases}\tag{5}$$

Simple arithmetic shows that if $x \in \Lambda_n$

$$\frac{c_n(x, \eta^x)}{c_n(x, \eta)} = \exp\left(2 \sum_{A \ni x} J(A)\eta(A)\right) = \frac{w_{n,\zeta}(\xi)}{w_{n,\zeta}(\xi^x)},\tag{6}$$

where ξ^x has the obvious definition, so

$$c_n(x, \eta^x)w_{n,\zeta}(\xi^x) = c_n(x, \eta)w_{n,\zeta}(\xi).\tag{7}$$

The last equation implies that $w_{n,\zeta}$ is a stationary measure for the Markov chain with state space $\{-1, 1\}^{\Lambda_n}$ (and in fact a special type of stationary measure called a

reversible measure). To prove this we define the 'Q matrix' for the chain on $\{-1, 1\}^{\Lambda_n}$

$$
Q^n_{\xi,\xi'} = \begin{cases} c_n(x, \eta), & \text{if } \xi' = \xi^x,\ x \in \Lambda_n, \\ -\sum_{x \in \Lambda_n} c_n(x, \eta), & \text{if } \xi' = \xi, \\ 0, & \text{otherwise} \end{cases} \tag{8}
$$

(here and in what follows $\xi \in \{-1, 1\}^{\Lambda_n}$, $\zeta \in \{-1, 1\}^{\Lambda_n^c}$ and η is their combination defined in (3)). Now a necessary and sufficient condition for ν to be a stationary measure is that

$$
\sum_{\xi} \nu(\xi') Q_{\xi',\xi} = 0 \quad \text{for all } \xi \in \{-1, 1\}^{\Lambda_n} \tag{9}
$$

and we have from (8) and (7) that

$$
\sum_{\xi'} w_{n,\zeta}(\xi') Q^n_{\xi',\xi} = -w_{n,\zeta}(\xi)\left(\sum_{x \in \Lambda_n} c(x, \eta)\right) + \sum_{x \in \Lambda_n} w_{n,\zeta}(\xi^x) c_n(x, \eta^x) = 0. \tag{10}
$$

Remark. The last computation shows that $w_{n,\zeta}$ is a stationary measure for the system in Λ_n with boundary condition ζ. By looking at formulas (6)–(10) we see that we have the same conclusion if we modify the flip rates to have the form

$$
c'(x, \eta) = c(x, \eta) b(x, \eta), \qquad \text{where } c(x, \eta) \text{ has the form given in (1) and}
$$
$$
b(x, \eta) = b(x, \eta^x) \in (0, \infty) \quad \text{for all } x \in \eta, \tag{11}
$$

so we will adopt this level of generality for the rest of the paper and refer to this class of flip rates as the exponential family.

Formula (10) shows that $w_{n,\zeta}$ are stationary measures. To obtain stationary distributions we have to let

$$
\mu_{n,\zeta}(\xi) = w_{n,\zeta}(\xi)/Z_{n,\zeta}, \quad \text{where } Z_{n,\zeta} = \sum_{\xi} w_{n,\zeta}(\xi). \tag{12}
$$

The normalizing constant $Z_{n,\zeta}$ is called the partition function (for the system in Λ_n with boundary condition ζ). We will return to it later. For the moment our purpose is to construct a collection of stationary distributions for the infinite particle system, so we will concentrate on the $\mu_{n,\zeta}$.

Let $\mathcal{G}_n$ be the closed convex hull of $\{\mu_{n,\zeta}: \zeta \in \{-1, 1\}^{\Lambda_n^c}\}$, which are now considered to be measures on $\{-1, 1\}^{Z^d}$. $\mathcal{G}_n$ is the class of equilibrium distributions for the system in Λ_n if we allow random boundary conditions. It is a consequence of a theorem which guarantees the existence of the processes under consideration (e.g. Liggett's theorem [86, p. 192]) that if $n_k \to \infty$, $\nu_k \in \mathcal{G}_{n_k}$ for all k, and $\nu_k \to \nu$, then ν is a stationary distribution for the infinite particle system. The collection of all stationary distributions which can be generated in this way are called the Gibbs states for the infinite particle system and denoted $\mathcal{G}$.

This terminology was invented before infinite particle systems. When the flip rates have the form given in (1) this reduces to Dobrushin's definition of Gibbs states for

infinite systems in terms of their conditional probabilities. In Dobrushin's approach the measures $\mu_{\Lambda,\zeta}$ are viewed as specifying the conditional distribution of the configuration in Λ given that the configuration in Λ^c is ζ and the Gibbs states are defined to be the probability measures ν which satisfy

$$\nu_\Lambda(\xi) = \int \nu_{\Lambda^c}(\mathrm{d}\zeta)\mu_{\Lambda,\zeta}(\xi) \quad \text{for all } \xi \in \{-1, 1\}^\Lambda, \tag{13}$$

where for $S \subset Z^d$, ν_S denotes the distribution ν induces on $\{-1, 1\}^S$.

The definition of Gibbs state given in (13) (called the DLR equations because of the contributions of Lanford and Ruelle) is equivalent to our previous definition. To prove this we observe that (13) says that in Λ, ν looks like $\mu_{\Lambda,\zeta(\nu)}$, where $\zeta(\nu)$ is a random boundary condition with distribution ν_{Λ^c}. Letting $\Lambda \uparrow \infty$ we see that anything which satisfies (13) is a Gibbs state with our previous definition. To prove the converse we need some notation:

if $S \subset Z^d$ and $\xi \in \{-1, 1\}^S$, then we let

$$A_{\xi,S} = \{\eta : \eta(x) = \xi(x) \text{ for all } x \in S\}. \tag{14}$$

(The configurations which agree with ξ in S.) Let $\Lambda_1 \subset \Lambda_2 \subset Z^d$ and $\Gamma = \Lambda_2 - \Lambda_1$. A little computation shows

$$\mu_{\Lambda_2,\zeta}(A_{\xi,\Lambda_1}) = \sum_\gamma \mu_{\Lambda_2,\zeta}(A_{\gamma,\Gamma})\mu_{\Lambda_1,(\zeta,\gamma)}(A_{\xi,\Lambda_1}), \tag{15}$$

where (ζ, γ) denotes the configuration in $\{-1, 1\}^{\Lambda_2^c}$ obtained by combining ζ and γ (see Fig. 3 to get oriented). Although (15) takes complicated notation to state
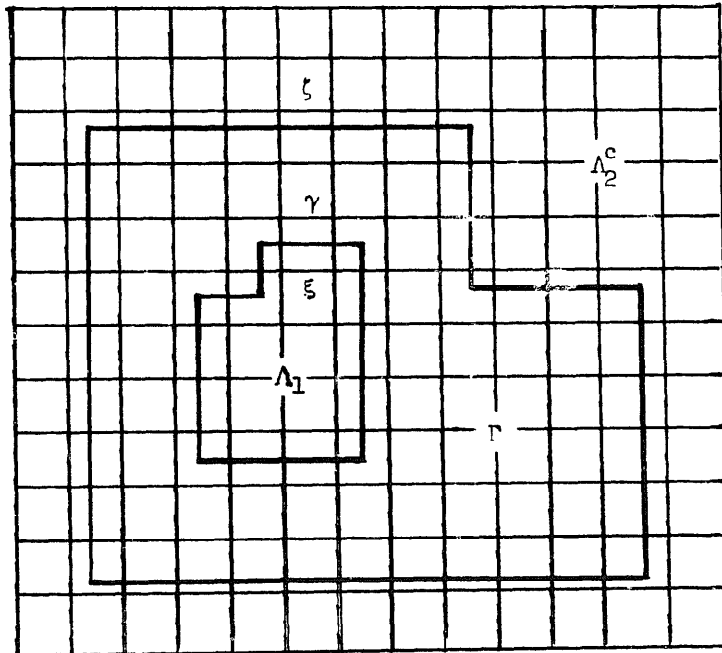

Fig. 3.

formally it is easy to say in words: if we want to compute the probability under $\mu_{A_2, \zeta}$ of seeing ξ in A_1 we first compute the distribution in $\Gamma = A_2 - A_1$ and then take the average of the probabilities under the $\mu_{A_1(\zeta, \gamma)}$. (This is a spatial Markov property.) Letting $A_2 \uparrow Z^d$ in (15) shows that any Gibbs states has the form given in (13).

At this point we have shown that our previous definition of Gibbs state is equivalent to (13). We have taken the time to do this because most people who study Gibbs states (including all the physicists) take (13) as the definition and do not consider the process η_t. For a good introduction to the subject see Preston [39] or Ruelle [41] or [43]. After being introduced, the reader can find out the current state of the art by looking at the Communications in Mathematical Physics.

6. One dimensional systems

In this section we will consider infinite particle systems with $S = Z$ concentrating primarily on those with flip rates in the exponential family. In one dimension the Gibbs states defined in Section 4 are easy to compute. Before introducing all the notation needed to treat the general case, let's consider the one dimensional Ising model which has flip rates

$$c(x, \eta) = \exp(-\beta(\eta(x)\eta(x + 1) + \eta(x)\eta(x - 1))). \tag{1}$$

In this case examining (4) and (12) of Section 5 shows that if $x \in [-n, n]$, then under $\mu_{[-n,n],\zeta}$ the coordinate vectors $(\eta(-n), \ldots, \eta(x-1))$ and $(\eta(x+1), \ldots, \eta(n))$ are conditionally independent given $\eta(x)$, so under $\mu_{[-n,n],\zeta}$, $\{\eta(x), -n \leq x \leq n\}$ is a Markov chain. A little computation shows that the transition matrix of this chain is

$$(2 \cosh \beta)^{-1}\begin{pmatrix} e^\beta & e^{-\beta} \\ e^{-\beta} & e^\beta \end{pmatrix} \tag{2}$$

and that $\mu_{[-n,n],\zeta}$ corresponds to conditioning this chain to have $\eta(-n-1) = \zeta(-n-1)$ and $\eta(n+1) = \zeta(n+1)$. Since the transition matrix in (2) has all positive entries $(\frac{1}{2}, \frac{1}{2})$ is unique stationary distribution and it follows from a standard Markov chain convergence theorem that the one dimensional Ising model has only one Gibbs state – the unique distribution on $\{-1, 1\}^Z$ which makes the coordinates $\eta(n)$, $-\infty < n < \infty$, a Markov chain with transition probability given in (2).

It is easy to generalize the argument above to systems defined by finite range potentials. If the range is L and $x \in [-n, n]$, then under $\mu_{[-n,n],\zeta}$ the coordinate vectors $(\eta(-n), \ldots, \eta(x-L-1))$ and $(\eta(x+L+1), \ldots, \eta(n))$ are conditionally independent given $(\eta(x-L), \ldots, \eta(x+L))$, so under $\mu_{[-n,n],\zeta}$, $\{(\eta(x-L), \ldots, \eta(x+L)): -n+L \leq x \leq n-L\}$ is a Markov chain whose transition matrix has the property that the $(2L + 1)$th power has all positive entries so again applying a standard Markov chain result we see that there is only one Gibbs state.

Having shown that any finite range one dimensional system in the exponential family has a unique Gibbs state it is natural (but somewhat optimistic) to ask if $\mathcal{G} = \mathcal{I}$.

There is probably a trivial example with long range interactions with $\mathcal{G} \neq \mathcal{I}$ but I think it is reasonable to expect that $\mathcal{G} = \mathcal{I}$ when the flip rates are finite range and translation invariant. Positive results on this question are almost nonexistent. Holley and Stroock have proved $\mathcal{G} = \mathcal{I}$ for finite range exponential systems in 1 and 2 dimensions. On p. 39 of [75], they prove (Theorem 1.7)

If $S = Z^d$, $d = 1$ or 2 and the flip rates $c(x, \eta)$
have the form given in (11) of Section 5 with

$$0 < \inf_{x, \eta} b(x, \eta) \leq \sup_{x, \eta} b(x, \eta) < \infty, \tag{3}$$

then $\mathcal{G} = \mathcal{I}$.

No one knows if this is true in dimensions $d \geq 3$.

(3) allows us to conclude that for one dimensional exponential systems with finite range potentials (and in particular Example 2(a) of the introduction) $|\mathcal{I}| = 1$. Since the flip rates in (11) of Section 5 encompass a wide variety of examples it is natural to wonder if we will have $|\mathcal{I}| = 1$ for all translation invariant flip rates with finite range. Five seconds thought reveals this is false since the process may have more than one absorbing state. If we assume in addition that $c(x, \eta) > 0$ for all x, η we eliminate these trivial examples and we have

(4) The positive rates conjecture. If $S = Z$ and the flip rates $c(x, \eta)$ are positive, translation invariant, and have finite range, then $|\mathcal{I}| = 1$.

I think this conjecture is based on the belief that, although the flip rates in the exponential class have a special form which allows us to compute the $\mu_{[-n,n],\zeta_n}$ the fact that the limit of the $\mu_{[-n,n],\zeta_n}$ is independent of the sequence of boundary conditions is due to the fact the correlation between two sites in $[-n, n]$ under $\mu_{[-n,n],\zeta_n}$ decays exponentially as a function of their distance, and this in turn is due to the positivity of the flip rates and the 'geometry' of Z. The proof of the uniqueness of the Gibbs state for systems in the exponential family supports this belief since all we needed there was that the Markov chains were irreducible. We should not however get too carried away with the reasoning – the last sentence does not mention Z explicitly and hence would lead us eventually to the false conclusion that in two dimensional Ising model, which has positive flip rates, we have $|\mathcal{I}| = 1$.

The positive rates conjecture represents just one of the ways the results for finite range exponential systems can be generalized. The rest of this section will be devoted to stating some results and an intriguing conjecture for exponential systems with potentials which have infinite range.

The first results for potentials with infinite range were due to Dobrushin [4] and Ruelle [40] who showed independently that if $J(\{x, y\}) = K(y - x)$ and $J(A) = 0$ when $|A| \neq 2$, then

$$|\mathcal{I}| = 1, \text{ whenever } \sum_{x=0}^{\infty} \sum_{y=-\infty}^{-1} K(x - y) = \sum_{n=1}^{\infty} nK(n) < \infty. \tag{5}$$

If one does not bother with the details it is easy to prove this result and a lot more. Thouless [47] has given a convincing heuristic argument which he attributes to Landau and Lifschitz [26]. In the next two paragraphs we give a moderately revised version of his argument.

If we have a configuration for the Ising model in $\{1, 2, \ldots, N\}$ which has average magnetization $\mu \neq 0$, then we can construct a distribution with average magnetization 0 by picking an integer L at random and flipping the spins at $1, 2, \ldots, L$. If we flip at L, the energy change is

$$\Delta E = 2 \sum_{n=1}^{L} \sum_{m=L+1}^{N} K(m-n)\eta(m)\eta(n). \tag{6}$$

Since there are N possible choices of L the entropy change $\Delta S = 0(\log n)$. For a short range interaction (.e. the case considered by Dobrushin and Ruelle) the sum on the right-hand side of (6) is bounded independent of L and N, so if N is sufficient large $\Delta E - T\Delta S$ is always negative and the ordered state cannot be the equilibrium state. If we consider interactions $K(n)$ that decrease monotonically as n goes to infinity, it is clear that no interaction falling off faster than $n^{-\alpha}$, $\alpha > 2$, can lead to an ordered state, because ΔE then increases slower than $\ln N$ as N increases.

The case of an interaction falling off as θn^{-2} is particularly interesting. In this case the change in the expectation value of the energy is

$$\Delta E = 2\mu^2 \theta \ln N + 0(1) \tag{7}$$

if the spin–spin correlation function is equal to μ^2 for large separation of the spins. If the width of the distribution of number of spins up is $N^{1/2}$, as it is for a system with the usual type of long range order, $\Delta S = (2CT)^{-1} \ln N$ and so if there is long range order in this case the inequality $2\mu^2 \theta \geq 1/2C$ must be satisfied. Writing μ as a function of θ this implies that if $\mu > 0$,

$$\mu(\theta) \geq (4C\theta)^{-1/2}. \tag{8}$$

Since $\mu(\theta)$ is increasing function of θ it follows that μ must be discontinuous at $\sup\{\theta: \mu(\theta) = 0\}$.

While the argument above may be convincing, it is certainly not a rigorous proof and, as I can testify, is difficult to fill the necessary details to make it one (the entropy computation was my stumbling block). Nonetheless there are rigorous results which suggest that the conclusions above are true. Dyson [8] has shown

$$\text{If } (\log \log n)^{-1} \sum_{m=1}^{n} mK(m) \to 0, \quad \text{then } |\mathcal{I}| = 1 \tag{9}$$

and also shown [10]

There is a number $a > 0$ so that if $K(n) \geq an^{-2} \log \log n$ for all n sufficiently large $|\mathcal{I}| > 1$. \tag{10}

These results are obtained by comparing the Ising system with a 'Hierarchical Model' which is easier to analyze. In [10] Dyson showed that the Hierarchical Models which corresponds to $K(n) = \theta n^{-2}$ have a jump discontinuity at the critical value but no one has succeeded in proving this for the Ising model although numerical results of Anderson and Yuvall reported in Dyson's paper seem to support this conclusion.

7. Attractive systems

In this section we will study a class of infinite systems with state space $\{-1, 1\}^S$ which have a special monotonicity property. These processes were first studied by Holley [70] who named them attractive. Later workers have confused the issue by calling them monotone so it is tempting to clarify the issue by calling them Holley processes. One look at the list of references however, reveals that this would be a highly ambiguous designation so I'll stick with the term attractive.

Define a partial ordering $\leq$ on $\{-1, 1\}^S$ by

$$\eta \leq \zeta \quad \text{if and only if} \quad \eta(x) \leq \zeta(x) \quad \text{for all } x \in S. \tag{1}$$

(Here and in what follows we will leave it to the reader to determine the meaning of $\leq$ from the context). We say that a real valued function f defined on $\{-1, 1\}^S$ is increasing if

$$\eta \leq \zeta \quad \text{implies} \quad f(\eta) \leq f(\zeta) \tag{2}$$

and decreasing if $-f$ is increasing. A collection of flip rates $c(x, \eta)$ is said to be attractive if

(a) $\eta \to c(x, \eta)$ is increasing on $\{\eta : \eta(x) = -1\}$,

(b) $\eta \to c(x, \eta)$ is decreasing on $\{\eta : \eta(x) = 1\}$.

$$\tag{3}$$

In attractive models the rate of flipping from -1 to 1 is an increasing function of the configuration and the rate of flipping back from 1 to -1 is a decreasing function so there is a tendency for blocks of 1's to form (and since the definition is symmetric the same can be said for -1's). In the last three sections we have encountered many examples of attractive systems – the voter model and contact processes have this property and, with a little patience, the reader can check that all the additive processes have this property. (There is no reason to do this however, since this will be obvious by the time we reach the next remark.)

The monotonicity property in (3) has a very useful consequence: if we are given initial configurations $\eta^1 \leq \eta^2$, then we can construct a pair of processes $\eta_t^1, \eta_t^2, t \geq 0$, on the same probability space in such a way that $\eta_t^1 \leq \eta_t^2$ for all $t \geq 0$ and for $i = 1, 2, \eta_t^i, t > 0$, has $\eta_0^i = \eta^i$ and is a Markov process with flip rates c. To do this we write down the following flip rates for a process with state space $\{(-1, -1), (-1, 1), (1, 1)\}^S$:

transition at x	flip rate in (η^1, η^2)
$(1, 1) \to (-1, -1)$	$c(x, \eta^2)$
$(1, 1) \to (-1, 1)$	$c(x, \eta^1) - c(x, \eta^2)$
$(-1, 1) \to (-1, -1)$	$c(x, \eta^2)$
$(-1, 1) \to (1, 1)$	$c(x, \eta^1)$
$(-1, -1) \to (1, 1)$	$c(x, \eta^1)$
$(-1, -1) \to (-1, 1)$	$c(x, \eta^2) - c(x, \eta^1)$.

$$(4)$$

In words when $\eta^1(x) = \eta^2(x)$ the processes flip together as much as they can and when $\eta^1(x) \neq \eta^2(x)$ they flip independently. This is called the basic coupling.

The reader should observe that each flip in (4) preserves the inequality $\eta^1(x) \leqslant \eta^2(x)$ and that since we have assumed c is attractive all the flip rates we have written down are $\geqslant 0$. If we assume (as we will throughout the rest of the paper) that c has finite range and is translation invariant, then it follows from a result of Liggett [85] that there is a unique Markov process (η^1_t, η^2_t), $t \geqslant 0$, with these flip rates. The reader can check by adding various rates in the table that the coordinates $\eta^1(x)$ and $\eta^2(x)$ individually flip at rates $c(x, \eta^1)$ and $c(x, \eta^2)$ so we have achieved the desired coupling.

Remark. In the additive case we do not need the construction above because the graphical representation gives the desired coupling simultaneously for all initial configurations.

Using the basic coupling we can immediately deduce several facts about the limiting behavior of attractive systems:

> Let η^+_t be an attractive process with $\eta^+_0 \equiv 1$. If f
> is an increasing function on $\{-1, 1\}^S$, then $\mathbf{E}f(\eta^+_t)$ is
> a decreasing function of t.   (5)

Proof. Let $s < t$, let $\eta^1 \equiv 1$ and let η^2 have distribution η^+_{t-s}. Since we have $\eta^1 \geqslant \eta^2$ a.s. it follows that we can construct η^1_u, η^2_u, $u \geqslant 0$, so that $\eta^1_s \geqslant \eta^2_s$ ($\overset{d}{=} \eta^+_t$) and hence $f(\eta^1_s) \geqslant f(\eta^2_s)$. Taking expectations gives $\mathbf{E}f(\eta^+_s) \geqslant \mathbf{E}f(\eta^+_t)$ the desired result.

(5) implies that the distribution   is a decreasing function of t in the partial ordering defined by

$$\mu \leqslant \nu \text{ if and only if } \int f \, d\mu \leqslant \int f \, d\nu \text{ for all increasing functions } f.   \quad (6)$$

Using this observation it is easy to show that as $t \to \infty$, η^+_t converges weakly to a limit. (5) implies that if g is a finite linear combination of increasing functions, then

$$\lim_{t \to \infty} \mathbf{E}g(\eta^+_t) \quad \text{xists.}$$

Since every function which depends upon only finitely many coordinates is in this class, it follows that

$$\text{as } t \to \infty, \ \eta_t^+ \text{ converges weakly to a limit } \mu^+. \tag{7}$$

It is a standard fact that (7) implies μ^+ is a stationary distribution. (Recall from the introduction that the associated semigroup is Feller.) It is called the upper invariant measure because it is clear from the basic coupling that if $\nu \in \mathcal{I}$, then $\nu \leq \mu^+$ (in the sense of (6)). All the reasoning above can be repeated for a system η_t^-, $t \geq 0$ with $\eta_0^- \equiv -1$, with the result that as $t \to \infty$, η_t^- converges to a lower invariant measure μ^- which has the property that if $\nu \in \mathcal{I}$, then $\mu^- \leq \nu \leq \mu^+$. From the last inequality we see there is a dichotomy for attractive systems: if $\mu^- \neq \mu^+$ the stationary distribution is not unique, if $\mu^- = \mu^+ = \mu$, then μ is the only stationary distribution and it follows from the basic coupling that starting from any initial distribution $\eta_t^+ \Rightarrow \mu$ as $t \to \infty$.

Although the dichotomy above is easy to establish, it provides useful information. If we apply this result to the Ising model, we see that starting from $\eta^+ \equiv 1$ and $\eta^- \equiv -1$ the process converges to limits μ^+ and μ^- which satisfy $\mu^+ \geq \mu^-$. In this situation it is easy to show that $\mu^+ = \mu^-$ if and only if $\mu^+\{\eta: \eta(x) = 1\} = \mu^-\{\eta: \eta(x) = 1\}$ for all x. Since the Ising model is symmetric under the interchange of 1 and -1 we have

$$\mu^-\{\eta: \eta(x) = 1\} = \mu^+\{\eta: \eta(x) = -1\} = 1 - \mu^+\{\eta: \eta(x) = 1\}$$

so $\mu^+ = \mu^-$ if and only if $\mu^+\{\eta: \eta(x) = 1\} = \frac{1}{2}$, i.e.

$$\int \eta(x) \, du^+(\eta) = 0.$$

Similar conclusions can of course be obtained for other systems with state space $\{-1, 1\}^S$. By a simple change of notation the results above can be extended to systems with state space $\{0, 1\}^S$. If they are applied to the contact process, we get the following useful result: as $t \to \infty$, ξ_t^Z (the contact process with $\xi_0^Z = Z$) converges weakly to a limit ξ_∞^Z and the contact process has $|\mathcal{I}| > 1$ if and only if ξ_∞^Z is not δ_0, the point mass on the configuration $\eta \equiv 0$.

The basic coupling can (sometimes) be used to compare 2 attractive processes. By writing out a new table of flip rates with $c_i(x, \eta^i)$ replacing $c(x, \eta^i)$ in (4) we see that if c_1 and c_2 are flip rates which satisfy

$$\begin{aligned} \text{(a)} \quad & c_1(x, \eta) \geq c_2(x, \zeta) \text{ if } \eta \leq \zeta \text{ and } \eta(x) = \zeta(x) = 1, \\ \text{(b)} \quad & c_1(x, \eta) \leq c_2(x, \zeta) \text{ if } \eta \leq \zeta \text{ and } \eta(x) = \zeta(x) = -1 \end{aligned} \tag{8}$$

and if $\eta^1 \leq \eta^2$, then we can construct a pair of processes $\eta^1 \leq \eta^2$, $t \geq 0$, on the same probability space in such a way that η_t^i has initial configuration η^i and flip rates c_i and $\eta_t^1 \leq \eta_t^2$ for all $t \geq 0$.

Applying the last result to the contact processes we see that if we let ξ_t^λ be a contact process with parameter λ and $\xi_0^\lambda = Z^d$ and if we have $\xi_t^\lambda \Rightarrow \delta_0$, then $\xi_t^\theta \Rightarrow \delta_0$ for all

$\theta < \lambda$. The last observation implies that if we let $\lambda_{cr} = \sup\{\lambda : \xi_t^\lambda \Rightarrow \delta_0\}$, then $\xi_t^\lambda \Rightarrow \delta_0$ for $\lambda < \lambda_{cr}$ and does not $\Rightarrow \delta_0$ for $\lambda > \lambda_{cr}$. This argument of course tells us nothing about λ_{cr} or about the behavior of the system at $\lambda = \lambda_{cr}$. These are both difficult problems.

While the result in (8) is useful for contact processes it does not apply to the Ising model

$$c(x, \eta) = \exp\left(-\beta \sum_{\nu, \|u\| = 1} \eta(x)\eta(x + u)\right).$$

If c_1 and c_2 are the flip rates for parameters $\beta_1 < \beta_2$, then (8) does not hold: if $\eta(x) = 1$ and $\sum_{u, \|u\| = 1} \eta(x + u) < 0$, $c_1(x, \eta) < c_2(x, \eta)$ so (a) does not hold.

In order to prove the existence of a critical value for the Ising model we have to use other means. The method in this case is a powerful correlation inequality which was discovered by Griffiths [18] and which is (due to the subsequent generalization of Kelley and Sherman [24]) called the GKS inequality. To state this result we need some definitions: Let Λ be a finite set. If $x \in \Lambda$ and $\xi \in \{-1, 1\}^\Lambda$, let

$$c(x, \xi) = \exp\left(-\sum_{A, x \in A \subset \Lambda} K_A \xi(A)\right), \quad \text{where } \xi(A) = \prod_{y \in A} \xi(y) \tag{9}$$

The $c(x, \xi)$ are flip rates for a Markov chain with state space $\{-1, 1\}^\Lambda$. This chain has a unique stationary distribution μ. If f is a function on $\{-1, 1\}^\Lambda$, let $\langle f \rangle$ denote $\int f \, d\mu$. With this notation the GKS inequality (which is really two inequalities) may be stated as

$$\text{If } A, B \subset \Lambda, \quad \text{then } \frac{\partial \langle \xi(A) \rangle}{\partial K_B} = \langle \xi(A)\xi(B) \rangle - \langle \xi(A) \rangle \langle \xi(B) \rangle \geqslant 0. \tag{10}$$

In words, $\xi(A)$ and $\xi(B)$ are positively correlated and $\langle \xi(A) \rangle$ is increased by increasing any of the K_B.

To apply the GKS inequality to flip rates in the exponential family we let Λ be a subset of Z^d, $\zeta \in \{-1, 1\}^{\Lambda^c}$ be a boundary condition and define

$$K_A = \sum_{B \subset \Lambda^c} J(A \cup B)\zeta(B), \quad \text{where } \zeta(B) = \prod_{y \in B} \zeta(y), \zeta(\emptyset) = 1. \tag{11}$$

With this definition of the K_A's it is easy to check that the stationary distribution for the flip rates in (10) is what we called $\mu_{\Lambda, \zeta}$ in Section 5. If we call this $\mu_{J, \Lambda, \zeta}$ now (to record the dependence on the potential), it follows from the GKS inequality that if $0 \leqslant J(S) \leqslant J'(S)$ for all $S \subset Z^d$, then we have

$$\int \xi(A) \, d\mu_{J, \Lambda, 1} \leqslant \int \xi(A) \, d\mu_{J', \Lambda, 1} \tag{12}$$

for all $A \subset \Lambda$.

Remark 1. The reader should note that the partial ordering of distribution defined in (12) is much different than the one defined in (6) (which was used for comparing attractive processes).

Remark 2. An inequality closely related to one part of the GKS inequality has been proved by Harris [68]. His results imply that in an attractive system in which $|F| = 2$ and only one site flips at a time the upper invariant measure has

$$\int fg \, d\mu^+ \geq \int f \, d\mu^+ \int g \, d\mu^+ \tag{13}$$

for all increasing f, g. This inequality says μ^+ has positive correlations and can be compared to the second inequality in (10) even though it is in substance much different. To my knowledge no one has proved an analogue of the first part of the GKS inequality (the monotonicity of $\eta(A)$ as a function of K_B) for a more general class of spin systems, even though, I think it is a common opinion that such a result would be very useful

Remark 3. Inequality (13) is called the FKG inequality which it is applied to a system in the exponential family. For a direct proof of this inequality see Fortuin, Ginibre and Kasteleyn [12], Ginibre [15] or Holley [71] (the last is my favorite).

Remark 4. The FKG and GKS are just the first two of an infinite sequence of similar inequalities which curiously all have three letter designations. Needless to say not all these inequalities are useful but the next one in the sequence, the GHS inequality (see Griffiths, Hurst and Sherman [20] or for an easier proof Lebowitz [29]) is very useful. Preston [38] has used this to show $|\mathcal{G}| = 1$ when the potential satisfies $J(A) \neq 0$ if $|A| = 1$, $J(A) \geq 0$ if $|A| = 2$ and $J(A) = 0$ if $|A| \geq 3$ (an attractive pair potential with a nonzero magnetic field). The analogue of this result for the partition function was first proved by Lee and Yang [31] using purely analytical methods, for a proof of the fact about Gibbs states see Ruelle [42] and Lebowitz and Martin Löf [30].

8. The Ising model

In this section we will study the d-dimensional Ising model: $S = Z^d$, $F = \{-1, 1\}$ and

$$c(x, \eta) = \exp\left(-\beta \sum_{u, \|u\| = 1} \eta(x)\eta(x + u)\right).$$

From results in the last three sections we know that

(1) $\mathcal{G} = 1$ if and only if $\mu^+ = \mu^-$, which occurs if and only if $\langle \eta(0)\rangle_+ = -\langle \eta(0)\rangle_- = 0$.

(2) $\langle \eta(0)\rangle_+$ is an increasing function of β so if we let $\beta_{cr} =$

$\sup\{\beta : \beta \langle \beta : \langle \eta(0) \rangle \rangle_+ = 0\}$, then

$$\text{if } \beta < \beta_{cr} \quad |\mathscr{I}| = 1,$$
$$\text{if } \beta > \beta_{cr} \quad |\mathscr{I}| > 1.$$

(3) in one dimension $\beta_{cr} = \infty$.

The results above describe $|\mathscr{I}|$ for the one dimensional Ising model so we now turn to the two dimensional Ising model. The first step is to show $|\mathscr{I}| > 1$ when β is sufficiently large. The main idea of the argument was given by Peierls [36], and rigorous versions developed independently by Dobruchin [3] and Griffiths [16]. We will prove this using the approach in Griffiths [19].

The first step in the proof is to introduce a geometric description of the spin configuration. Consider Z^2 as a graph with edges connecting all x, y with $\|x - y\| = 1$. If $\Lambda_n = [-n, n]^2$, $\xi \in \{-1, 1\}^{\Lambda_n}$ and $\zeta \in \{-1, 1\}^{\Lambda_n^c}$ has $\zeta \equiv 1$, then we create contour lines by drawing a unit segment perpendicular to center of each edge e which connects two sites x, y with opposite spins.

A glance at Fig. 4 shows that the collection of lines we generate is always a finite union of (non-self-intersecting) polygons and is hence called a multipolygon. Since the boundary condition is fixed to be $\zeta \equiv 1$, it is easy to see that there is a 1-1 correspondence between configurations in $\{-1, 1\}^{\Lambda_n}$ and multipolygons with vertices in the shifted lattice $(\frac{1}{2}, \frac{1}{2}) + Z^2$ and which lie in the box $\Lambda_n^* = [-n - \frac{1}{2}, n + \frac{1}{2}]^2$.
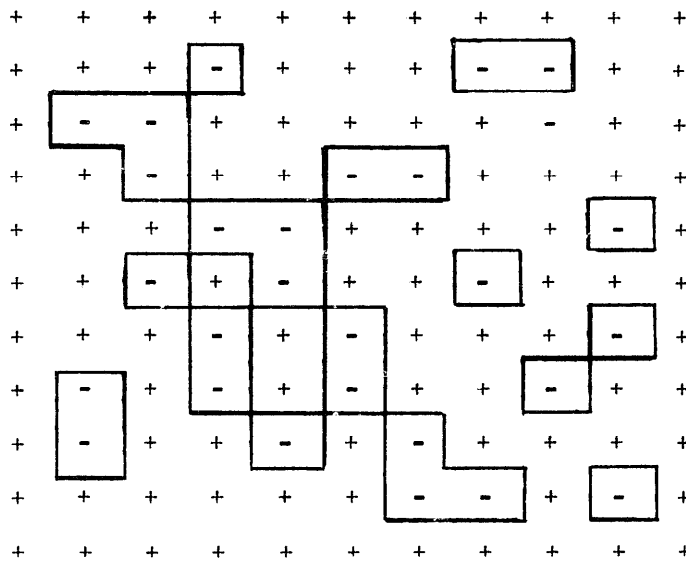

Fig. 4. Contours for a typical configuration in $\mu_{\Lambda_4, 1}$.

Although this is just a change of notation it is a useful one. if we let $\xi \in \{-1, 1\}^{\Lambda_n}$, let $K(\xi)$ the associated multipolygon, and let $|K(\xi)|$ be the number of edges it contains, then the exponent in the definition of $\omega_{\Lambda_{n,1}}(\xi)$ in (4) of Section 5 is equal to

$$-2\beta(\text{no. of times } \eta(A) = 1) + \beta(\text{no. of } A \text{ with } J(A) \neq 0) =$$
$$= -2\beta|K(\xi)| + \beta(2(2n + 1)(2n + 2)), \tag{5}$$

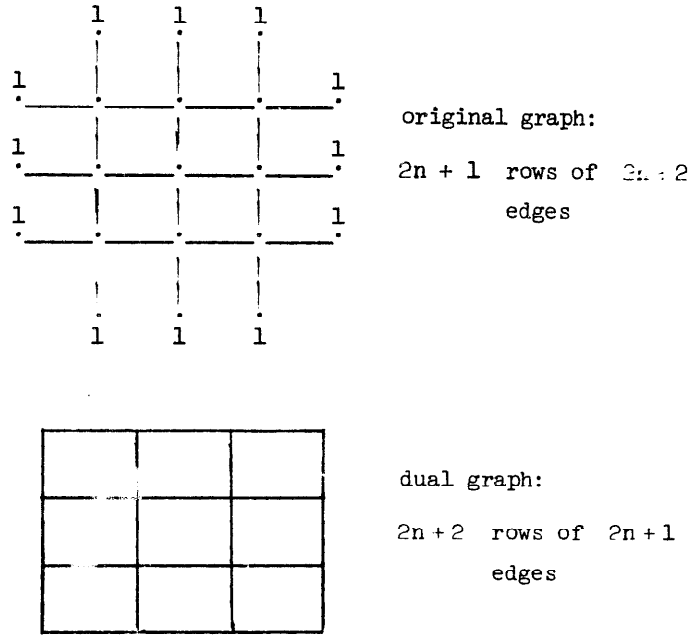where in the second equality we have used Fig. 5 to compute the value of the constant.

original graph:

2n + 1 rows of 2n + 2 edges

dual graph:

2n + 2 rows of 2n + 1 edges

Fig. 5. Λ_1 and Λ_1^*.

The first step in showing that $|\mathscr{P}| > 1$ when β is sufficiently large is to get a bound on the probability a (non-self-intersecting) polygon L with length $|L|$ is a subset of $K(\xi)$

$$\mu_{\Lambda_n,1}(L \subset K(\xi)) \leq e^{-2\beta|L|}/(1 + e^{-2\beta|L|}). \tag{6}$$

Proof. From (4) of Section 5 and (5) above it follows that

$$\mu_{\Lambda_n,1}(L \subset K(\xi)) = \sum_{\xi,K(\xi)\supset L} w_{\Lambda_n,1}(\xi) / \sum_{\xi} w_{\Lambda_n,1}(\xi)$$

$$= \sum_{\xi,K(\xi)\supset L} e^{-2\beta|K(\xi)|} \Big/ \sum_{\xi} e^{-2\beta|K(\xi)|}.$$

To estimate the last quantity we will show that with each configuration ξ with $K(\xi) \supset L$ we can associate a configuration ξ^* with $K(\xi^*) = K(\xi) - L$. We have supposed the polygon L is non-self-intersecting so it divides the plane into two regions – one unbounded (the outside) and one bounded (the inside, denoted L^0). Let

$$\xi^*(y) = \begin{cases} -\xi(y), & y \in L^0, \\ \xi(y), & \text{otherwise.} \end{cases}$$

It is easy to check that ξ^* has the desired property and that different ξ's give rise to different ξ^*'s. Combining this with the equalities above shows .

$$\mu_{\Lambda_n,1}(L \subset K(\xi)) \leq \frac{\sum\limits_{\xi,K(\xi) \supset L} e^{-2\beta|K(\xi)|}}{\sum\limits_{\xi,K(\xi) \supset L} e^{-2\beta|K(\xi)|}(1-e^{2\beta|L|})}$$

$$= \frac{1}{1+e^{2\beta|L|}} = e^{-2\beta|L|}/(1+e^{-2\beta|L|})$$

proving the desired result.

(6) allows us to estimate the probability a given polygon will occur. Inside a polygon of length l lying on the shifted lattice $(\frac{1}{2},\frac{1}{2}) + Z^d$ there are at most $(\frac{1}{4}l)^2$ sites in Z^d. Any -1 must be inside some polygon so if we let

$$M_n(\xi) = |\{x : \xi(x) = -1\}|,$$

then

$$M_n(\xi) \leq \sum_L \frac{|L|^2}{16} 1_{(L \subset K(\xi))},$$

where the sum is taken over all polygons lying in Λ_n^*. Taking expectations and using Fubini's theorem we get

$$\int M_n(\xi)\mu_{\Lambda_n,1}(d\xi) = 16^{-1} \sum_L |L|^2 \mu_{\Lambda_n,1}(L \subset K(\xi))$$

$$\leq 16^{-1} \sum_L |L|^2 e^{-2\beta|L|} = 16^{-1} \sum_{l=1}^{\infty} a_n(l)l^2 e^{-2\beta l}, \tag{7}$$

where $a_n(l)$ is the number of polygons of length l contained in Λ_n.

It is easy to get a rough estimate for the $a_n(l)$. A polygon may be constructed by picking an initial segment, which can be done in $2(2n+1)(2n+2) \leq 3(2n+1)^2$ ways, and then successively adding adjacent segments. Since we cannot repeat any edge at stages $2, \ldots, l$ there are at most 3 choices and it follows that

$$a_n(l) \leq 3(2n+1)^2 3^{l-1}/l. \tag{8}$$

Here we have used the fact that this procedure generates every polygon of length l l times to divide our estimate by l. Combining the last two inequalities above and observing that $a_n(l) = 0$ if l is odd

$$\int M_n(\xi)\mu_{\Lambda_n,1}(d\xi) \leq 16^{-1}(2n+1)^2 \sum_{m=1}^{\infty} 2m \, e^{(\log 3 - 2\beta)2m}$$

or letting $p = \exp(2(\log 3 - 2\beta))$

$$\frac{1}{(2n+1)^2} \int M_n(\xi)\mu_{\Lambda_n,1}(d\xi) \leq 8^{-1} \sum_{m=1}^{\infty} mp^m = 8^{-1}(p/(1-p)^2)$$

the right-hand side is $< \frac{1}{2}$ when $p/(1-p)^2 < 4$, i.e. $p < (9 - \sqrt{17})/8$. Combining the last observation with the definition of p we see that if β is sufficiently large

$$\frac{1}{(2n+1)^2} \int M_n(\xi) \mu_{A_n,1}(d\xi) \leq 0.499$$

for all n and consequently $\int \eta(0) d\mu^+(\eta) \geq 0.002$ (a little work is required to prove this rigorously).

In the last argument 0.499 could be replaced by any number <0.5 so we have $\int \eta(0) d\mu^+(\eta) > 0$ and hence $|\mathcal{I}| > 1$ whenever

$$2 \log 3 - 4\beta < \log\left(\frac{9 - \sqrt{17}}{8}\right),$$

i.e.

$$\beta > \tfrac{1}{2} \log 3 - \tfrac{1}{4} \log\left(\frac{9 - \sqrt{17}}{8}\right).$$

Remark 1. By fiddling with the computations above (e.g. noticing $a_n(2) = 0$, $a_n(4) = (2n+1)^2$ instead of $(9/2)(2n+1)^2, \ldots$) the value of the constant above can be improved but without a substitute for the estimate (8) we cannot do better than $\beta > (\log 3)/2$ (and we cannot even do this well unless someone can show that the finiteness of the sum in (7) implies $|\mathcal{I}| > 1$).

Remark 2. Even if we lived in the best of all possible worlds and used the best possible estimate for $a_n(l)$ the argument above is not sharp. It is known (see Smythe and Wierman [95], and Kesten [84]) that as $n \to \infty$

$$n^{-1} \log a_n(l) \to \kappa = 2.639 \cdots$$

where κ is the connectivity constant of the two dimensional lattice so by the argument above we could never do better than $\beta > (\log \kappa)/2 = 0.485$.

The arguments above show that if β is sufficiently large, $|\mathcal{I}| > 1$ or using the notation of (2) above that $\beta_{cr} < \infty$. Identifying β_{cr} is quite another matter. The solution of this problem developed slowly in a sequence of papers starting in 1925 when the Ising model (or more precisely its Gibbs states) was introduced and ending in 1972 when Bennetin, Gallavotti, Jono-Lasinio, and Stella put the finishing touches on the solution.

The first step in identifying the critical value for the two dimensional Ising model was taken in 1941 when Kramers and Wannier observed there was a special relationship between the $2d$ Ising models with parameters β and β^* where β^* is the solution of $\exp(-2\beta^*) = \tanh(\beta) = (e^\beta - e^{-\beta})/(e^\beta + e^{-\beta})$. To describe this relationship we have to introduce the physicist's notion of phase transition. In (12) of Section 5 we defined the partition function. In the case of the two dimensional Ising model

this definition says

$$Z_{n,\zeta} = \sum_{\xi} w_{n,\zeta}(\xi), \quad \text{where } w_{n,\zeta}(\xi) = \exp\left(\sum_{x,y} \beta\eta(x)\eta(y)\right) \tag{9}$$

and the sum is taken over pairs with $\|x - y\| = 1$ and $\{x, y\} \cap \Lambda_n \neq \emptyset$.

If we write the partition function as $Z_{n,\zeta}(\beta)$ to record the dependence on β, then it is known that for any sequence of boundary conditions ζ_n (see Ruelle [41, Section 2.4])

$$\lim_{n \to \infty} |\Lambda_n|^{-1} \log Z_{n,\zeta}(\beta) \text{ exists} \tag{10}$$

and the value of the limit $\varphi(\beta)$ is independent of the sequence of boundary conditions chosen. For a physicist the function $\varphi(\beta)$ gives a complete description of the infinite system. $f(\beta) = -\beta\varphi(\beta)$ gives the free energy per site while $U(\beta) = f'(\beta)$ gives the internal energy per site. The parameter β which gives the strength of the interaction $= 1/kT$, where k is Boltzman's constant so we can think of f and U as functions of T. If we differentiate U with respect to T, we get the specific heat. This quantity measures the system's susceptibility to change and if this is ∞ we have a highly unstable situation which suggests the presence of a phase transition. This consideration, which arose in the two dimensional Ising model, has been generalized and become the physicists definition of phase transition: the failure of $f(\beta)$ to be infinitely differentiable.

The partition function was first studied by two series expansions – one valid for small β (large T), and the other for large β (small T).

I. *High temperature expansion*: For each edge b which connects two sites x and y we let

$$\tilde{\eta}(b) = \eta(x)\eta(y) = \begin{cases} -1, & \text{if } K(\eta) \supset b, \\ 1, & \text{otherwise.} \end{cases}$$

With this notation it is immediate that

$$Z_{\Lambda_n,\zeta}(\beta) = \sum_{\xi} \exp\left(\beta \sum_{b \in \Gamma_n} \tilde{\eta}(b)\right) = \sum_{\xi} \prod_{b \in \Gamma_n} \exp(\beta\tilde{\eta}(b)),$$

where $\Gamma_n = $ set of edges contained in $(-n - 1, n + 1)^2$. Now if $u = 1$ or -1, we have

$$e^{\beta u} = \tfrac{1}{2}((e^{\beta} + e^{-\beta}) + u(e^{\beta} - e^{-\beta})) = \cosh\beta(1 + u(\tanh\beta)),$$

so we can write

$$Z_{\Lambda_n,\zeta}(\beta) = (\cosh\beta)^{|\Gamma_n|} \sum_{\xi} \prod_{b \in \Gamma_n} (1 + \tilde{\eta}(b)\tanh\beta)$$

$$= (\cosh\beta)^{|\Gamma_n|} \sum_{\xi} \sum_{B \subset \Gamma_n} (\tanh\beta)^{|B|} \left(\prod_{b \in B} \tilde{\eta}(b)\right). \tag{11}$$

Since $\tanh\beta \to 0$ as $\beta \to 0$ this expansion is good for small β.

II. *Low temperature expansion*. Using the notation from I we can write

$$Z_{\Lambda_n,\zeta}(\beta) = \sum_{\xi} \prod_{b \in \Gamma_n} \exp(\beta \tilde{\eta}(b))$$

$$= e^{\beta|\Gamma_n|} \sum_{\xi} \prod_{\substack{b \in \Gamma_n \\ \tilde{\eta}(b)=-1}} e^{-2\beta} = e^{\beta|\Gamma_n|} \sum_{\xi} e^{-2\beta L(\xi)}, \tag{12}$$

where $L(\xi) = \{b : \tilde{\eta}(b) = -1\}$.

If we consider the high temperature expansion with $\zeta_n \equiv 0$ (so called open boundary conditions, corresponding to an isolated system), then there are 2 types of terms in the sum in (11):

(i) Some $x \in \Lambda_n$ occurs an odd number of times. In this case pairing ξ with ξ^x shows

$$\sum_{\xi} \prod_{b \in B} \tilde{\eta}(b) = 0.$$

(ii) All $x \in \Lambda_n$ occur an even number of times in B. In case $\prod_{b \in B} \tilde{\eta}(b) = 1$ for all ξ so if we let K_n be the set of such B, we have

$$Z_{\Lambda_n,0}(\beta) = (\cosh \beta)^{|\Gamma_n|} \sum_{B \in K_n} (\tanh \beta)^{|B|}. \tag{13}$$

Kramers and Wannier's observation was that we could get a similar expression from the lower temperature expansion evaluated at β^* if we replaced Λ_n by $\Lambda_n^* = ((\frac{1}{2}, \frac{1}{2}) + Z^\beta) \cap [-n, n]^2$ and consider an Ising model on this lattice with boundary condition $\zeta_n \equiv 1$. In this case (12) becomes

$$Z_{\Lambda_n^*,1}(\beta^*) = e^{\beta^*|\Gamma_n^*|} \sum_{\xi} (\tanh \beta)^{|L(\xi)|}.$$

A glance at Fig. 6 shows that if $\xi \in \{-1, 1\}^{\Lambda_n^*}$ and we associate with each $b \in L(\xi)$ the
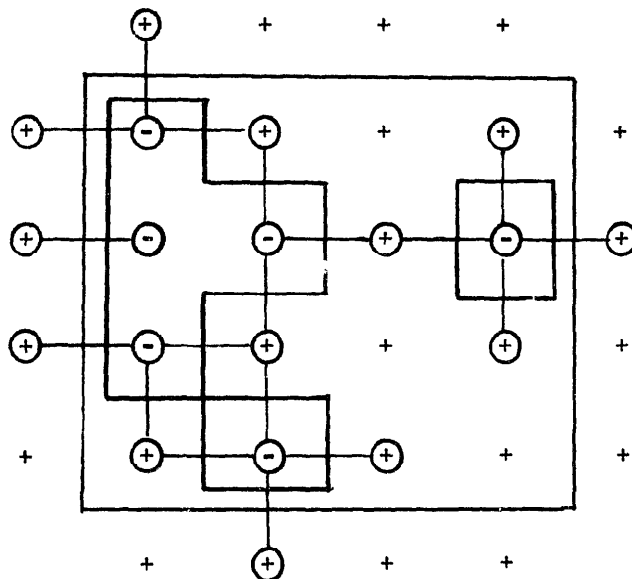

Fig. 6. A typical configuration $\xi \in \{-1, 1\}^{\Lambda_n^*}$.

arc $b^* \in \Gamma_n$ which intersects it, then $L^*(\xi) = \{b^* : b \in L(\xi)\} \in K_n$ and furthermore each $L \in K_n$ is associated with exactly one $\xi \in \{-1, 1\}^{\Lambda_n^*}$. This observation allows us to rewrite the last expression as

$$Z_{\Lambda_n^*,1}(\beta^*) = e^{\beta^* |\Gamma_n^*|} \sum_{B \in K_n} (\tanh \beta)^{|B|}, \tag{14}$$

so we have

$$\frac{Z_{\Lambda_n,0}(\beta)}{(\cosh \beta)^{|\Gamma_n|}} = \frac{Z_{\Lambda_n^*,1}(\beta^*)}{e^{\beta^* |\Gamma_n^*|}}. \tag{15}$$

If we assume that the free energy has only one singularity as a function of β, then (17) implies that this must occur at a point where $\beta = \beta^*$. Setting $\beta = \beta^*$ leads to the equation

$$\frac{e^{\beta} - e^{-\beta}}{e^{\beta} + e^{-\beta}} = e^{-2\beta}$$

which reduces to $\sinh(2\beta) = 1$.

The argument above suggests (but of course does not prove) that the phase transition point for the Ising model should be $\beta_0 = \frac{1}{2} \operatorname{arcsinh}(1)$. This fact was proved several years later by Onsager [34] who showed by a very ingenious calculation that

$$f(\beta) = \log 2 + \frac{1}{2\pi^2} \int_0^\pi d\theta_1 \int_0^\pi d\theta_2 \log(\cosh^2 2\beta - \sinh^2 \beta(\cos \theta_1 + \cos \theta_2)). \tag{16}$$

Differentiating gives

$$U(\beta) = \coth 2\beta \left(1 + \frac{(\sinh^2 2\beta - 1)}{\pi^2} \int_0^\pi \int_0^\pi \frac{d\theta_1 \, d\theta_2}{\cosh^2 2\beta - \sinh 2\beta(\cos \theta_1 + \cos \theta_2)}\right). \tag{17}$$

The integral in (17) diverges when $\cosh^2 2\beta - 2 \sinh 2\beta = 0$. This disaster does not make $U = \infty$, however. $\cosh^2 2\beta - \sinh^2 2\beta = 1$ so $\cosh^2 2\beta - 2 \sinh 2\beta = 0$ is equivalent to $\sinh 2\beta = 1$ and we have $(\sinh^2 2\beta - 1) = 0$. At this point we have seen that the integral is ∞ and the multiplier is 0. A little calculation shows that in the above $0 \cdot \infty = 0$ but taking one more derivative reveals the singularity

$$\partial U/\partial T \sim C \log|\beta - \beta_0| \quad \text{as } \beta \to \beta_0 \tag{18}$$

After the free energy was calculated the next problem was to compute the spontaneous magnetization, i.e. the extent to which the spins are aligned when $\beta > \beta_{cr}$. We have been deliberately vague in introducing the term spontaneous magnetization since it has been given several (ultimately equivalent) definitions. The original definition was

$$m(\beta) = \frac{\partial f(\beta, h)}{\beta \partial h}\bigg|_{h=0+}, \tag{19}$$

where $f(\beta, h)$ is the free energy for the system in the exponential family with $J(\{x\}) = -h$, $J(\{x, y\}) = \beta$ when $\|x - y\| = 1$, and $J(A) = 0$ otherwise. Although the definition in (19) may look mysterious it is the right one from the viewpoint of infinite particle systems. Lebowitz [28] has shown that $m(\beta) = \langle \eta(0) \rangle_+$ so $m(\beta)$ gives the net magnetic field in μ_+ and the value of $m(\beta)$ being >0 or $=0$ tells us whether $|\mathscr{I}| > 1$ or $=1$.

Although (19) appears to give an explicit formula for $m(\beta)$ it is useless since no one has been able to compute $f(\beta, h)$ for $h \neq 0$. Without a formula for $f(\beta, h)$ the determination of $m(\beta)$ was a slow process. The first step was taken by Onsager [35] who claimed that

$$m(\beta) = (1 - (\sinh 2\beta)^{-4})^{1/8} \tag{20}$$

but never published a proof of this result. The first rigorous result was due to Yang [49] who showed that if $m_0(\beta)$ is the right-hand side of (20), then

$$m_0(\beta) = \lim_{h \downarrow 0} \lim_{n \to \infty} \lim_{m \to \infty} \left(\frac{n}{h}\right) (f_{m,n}(\beta, h/n) - f_{m,n}(\beta, 0)), \tag{21}$$

where $f_{m,n}$ is the free energy for the Ising model in $\{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\}$ with periodic boundary conditions: $J(x - y) = \beta$ if $|(x_1 - y_1) \bmod m| + |(x_2 - y_2) \bmod m| = 1$.

The first connection with correlation functions was made by Montroll, Potts and Ward [33] who showed that if μ_p is the Gibbs state which is the limit of the equilibrium states in $[-n, n]^2$ with periodic boundary conditions, then

$$m_0(\beta) = \lim_{(x-y) \to \infty} \langle \eta(x) \eta(y) \rangle_p^{1/2}.$$

Having made the connection between $m_0(\beta)$ and $\langle \eta(x) \eta(y) \rangle_p$ the last step is to relate $\langle \eta(x) \eta(y) \rangle_p$ to $\langle \eta(x) \eta(y) \rangle_+$. This was done by Benettin, Gallavotti, Jona-Lasinio and Stella [2] who used duality to show that if $A \subset Z^d$ has an even number of elements, there are constants $\gamma_A(B)$ (depending also on β) so that

$$\langle \eta(A) \rangle_{0,\beta} = \sum_B \gamma_A(B) \langle \eta(B) \rangle_{+,B^*},$$

$$\langle \eta(A) \rangle_{+,\beta} \neq \sum_B \gamma_A(B) \langle \eta(B) \rangle_{0,\beta^*}$$

(see (2.3) and (2.4) in [2]). Since a $\beta > \beta_{cr}$ maps to a $\beta^* < \beta_{cr}$ it follows that for all $\beta \in (0, \infty)$,

$$\langle \eta(A) \rangle_{0,\beta} = \langle \eta(A) \rangle_{+,\beta^*}$$

for all $A \subset Z^d$ which have an even number of elements. The GKS inequalities show that

$$\langle \eta(x) \eta(y) \rangle_0 \leqslant \langle \eta(x) \eta(y) \rangle_p \leqslant \langle \eta(x) \eta(y) \rangle_+,$$

so we have

$$\langle \eta(x)\eta(y)\rangle_p = \langle \eta(x)\eta(y)\rangle_+.$$

It only remains to assemble the pieces. It is known that μ^+ is ergodic (see Holley [70]), so

$$\langle \eta(0)\rangle_+^2 = \lim_{n\to\infty}\left\langle\left(|\Lambda_n|^{-1}\sum_{x\in\Lambda_n}\eta(x)\right)^2\right\rangle_+$$

$$= \lim_{n\to\infty}|\Lambda_n|^{-2}\sum_{x,y\in\Lambda_n}\langle \eta(x)\eta(y)\rangle_+$$

$$= \lim_{n\to\infty}|\Lambda_n|^{-2}\sum_{x,y\in\Lambda_n}\langle \eta(x)\eta(y)\rangle_p$$

$$= \lim_{|x-y|\to\infty}\langle \eta(x)\eta(y)\rangle_p = m_0(\beta)^2$$

and we have

$$\langle \eta(0)\rangle_+ = m_0(\beta) = (1-(\sinh 2\beta)^{-4})^{1/8}. \tag{22}$$

The last result shows that $|\mathcal{J}| = 1$ if $\beta \leqslant \beta_0 = \frac{1}{2}\mathrm{arcsinh}(1)$ and $|\mathcal{J}| > 1$ if $\beta > \beta_0$, so we have identified the critical value. The next problem then is to study $\mathcal{J}$ when $|\mathcal{J}| > 1$. $\mathcal{J}$ has two obvious extreme points μ^- and μ^+ which are related by interchanging 1 and -1 so it is natural to conjecture that $\mathcal{J}_e = \{\mu^-, \mu^+\}$. The first step in proving this was taken in [32] by Messager and Miracle Sole who used duality ideas and some results of Ruelle to show that any translation invariant Gibbs state is a convex combination of μ_- and μ_+. This was the best known result until 1979 when Aizenmann [1] proved the result without the assumption of translation invariance.

The results above give a complete description of the invariant measures for the two dimensional Ising model so now we turn our attention to dimensions $d \geqslant 3$. Unfortunately there is not much to say. It follows from the GKS inequality that the critical value is a decreasing function of the dimension so the three dimensional critical value $\leqslant \beta_0$. With a little cleverness this result can be improved dramatically. By introducing the boundary conditions $\zeta_n(x) = \mathrm{sgn}(x_3 - \frac{1}{2})$ and using the GKS inequality van Beijeren [48] has shown that the three dimensional Ising model has nontranslation invariant Gibbs states for $\beta > \beta_0$.

The last result is obviously just the beginning of the study of the three dimensional Ising model but unfortunately is almost the end of the known results. We know that

(a) if β is sufficiently small $|\mathcal{J}| = 1$ (this is a consequence of the uniqueness result in Section 2);

(b) if β is sufficiently large every translation invariant stationary but is a convex combination of μ^+ and μ^-.

However, there are still many open questions

(1) Find β_{cr}^3 or at least show that $\beta_{cr}^3 < \beta_{cr}^2$. (The latter is not as easy as it sounds, although it is likely that someone knows how to do this.)

(2) Show that for any β every translation invariant stationary distribution is a convex combination of μ^+ and μ^-.

(3) Let $\gamma^3 = \sup\{\beta:$ all elements of $\mathcal{I}$ are translation invariant$\}$. Is $\gamma^3 > \beta_{cr}^3$? Or to be even bolder is $\gamma^3 = \beta_{cr}^2$?

Acknowledgment

References

The only papers listed below are the ones which were referred to above so the list does not come close to exhausting the literature. For more references on the Ising model see Gallavotti [13], or Ruelle [42, 43] and for more on other infinite particle systems see Liggett [28] or Griffeath [64].

I. The Ising model, etc.

[1] M. Aizenmann, Instability of phase coexistence and translation invariance in two dimensions. Phys. Rev. Lett. 43 (1979) 407–409.

[2] G. Bennetin, G. Gallavotti, G. Jona-Lasinio and A.L. Stella, On the Onsager–Yang value of the spontaneous magnetization, Comm. Math. Phys. 30 (1972) 45–54.

[3] R.L. Dobrushin, The existence of phase transition in two and three dimensional ferromagnets, Theoret. Probability Appl. 10 (1965) 209–230.

[4] R.L. Dobrushin, Description of a random field by means of conditional probabilities and conditions for its regularity, Theoret. Probability Appl. 13 (1968) 197–224.

[5] R.L. Dobrushin, The problem of uniqueness of a Gibbsian random field and the problem of phase transition, Funct. Anal. Appl. 2 (1969) 302–312.

[6] R.L. Dobrushin, Gibbsian states describing the coexistence of phases in the three dimensional Ising model, Theoret. Probability Appl. 17 (1972) 582–600.

[7] R.L. Dobrushin, Investigation of Gibbsian states for three dimensional lattice systems, Theoret. Probability Appl. 18 (1973) 253–271.

[8] F.J. Dyson, Existence of phase transition in a one dimensional ferromagnet, Comm. Math. Phys. 12 (1969) 91–107.

[9] F.J. Dyson, Nonexistence of spontaneous magnetization in a one-dimensional Ising ferromagnet, Comm. Math. Phys. 12 (1969) 212–215.

[10] F.J. Dyson, An Ising ferromagnet with discontinuous long range order, Comm. Math. Phys. 21 (1971) 269–283.

[11] F.J. Dyson, The existence and nature of phase transitions in one-dimensional ferromagnets. in: [37].

[12] C.M. Fortuin, J. Ginibre and P.W. Kasteleyn, Correlation inequalities on partially ordered sets, Comm. Math. Phys. 22 (1971) 89–103.

[13] G. Gallavotti and S. Miracle Sole, Equilibrium states of the Ising model in the two phase region, Phys. Rev. 5B (1972) 2555–2559.

[14] G. Gallavotti, Instabilities and phase transitions in the Ising model, a review. Riv. Nuovo Cimento 2 (1972) 133–169.

[15] J. Ginibre, Correlation inequalities in statistical mechanics, in: [37].

[16] R.B. Griffiths, Peierls proof of spontaneous magnetization in the two dimensional Ising model, Phys. Rev. 136A (1964) 437–439.

[17] R.B. Griffiths, Spontaneous magnetization in idealized ferromagnets, Phys. Rev. 152 (1966) 240–246.

[18] R.B. Griffiths, Correlations in Ising ferromagnets, I, J. Math. Phys. 8 (1967) 478–483; II, J. Math. Phys. 8 (1967) 484–489; III, Comm. Math. Phys. 6 (1967) 121–127.

[19] R.B. Griffiths, The Peierls argument for the existence of phase transitions, in: [37].

[20] R.B. Griffiths, C.A. Hurst and S.J. Sherman, Concavity of magnetization of an Ising Ferromagnet in a positive external field, J. Math. Phys. 11 (1970) 790–795.

[21] C. Gruber, A. Hinterman and D. Merlini, Analyticity and uniqueness of the invariant equilibrium state for general spin 1/2 classical lattice systems, Comm. Math. Phys. 40 (1975) 83–95.

[22] C. Gruber, A. Hinterman and D. Merlini, Group Analysis of Classical Lattice Systems, Lecture Notes in Physics 60 (Springer, Berlin, 1977).

[23] Ising, Beitrag zur theorie des ferromagnetismus, Z. Physik 31 (1925) 253–258.

[24] D.G. Kelley and S. Sherman, General Griffiths inequalities for correlations in Ising ferromagnets, J. Math. Phys. 9 (1968) 466–471.

[25] H.A. Kramers and G.K. Wannier, Statistics of the two dimensional ferromagnet I, II, Phys. Rev. 60 (1941) 252–276.

[26] Landau and Lifschitz, Statistical Physics (Pergammon Press, Oxford, 1958).

[27] O. Lanford and D. Ruelle, Observables at infinity and states with short range correlations in statistical mechanics, Comm. Math. Phys. 13 (1969) 194–215.

[28] J. Lebowitz, Bounds on the correlations and analyticity properties of ferromagnetic Ising spin systems, Comm. Math. Phys. 28 (1972) 313–321.

[29] J. Lebowitz, GHS and other inequalities, Comm. Math. Phys. 35 (1974) 87–92.

[30] J. Lebowitz and A. Martin Löf, On the uniqueness of the equilibrium state for Ising spin systems, Comm. Math. Phys. 25 (1972) 276–282.

[31] T.D. Lee and C.N. Yang, Statistical theory of equations of state and phase transitions I, Phys. Rev. 87 (1952) 404–409; II, Phys. Rev. 87 (1952) 410–419.

[32] Messager and Miracle Sole, Equilibrium states of the two dimensional Ising model in the two phase region, Comm. Math. Phys. 40 (1972) 187–196.

[33] E.W. Montroll, R.B. Potts and J.C. Ward, Correlations and spontaneous magnetization of the two dimensional Ising model, J. Math. Phys. 4 (1963) 308–322.

[34] L. Onsager, Crystal Statistics I, A two dimensional model with an order-disorder transition, Phys. Rev. 65 (1944) 117–149.

[35] L. Onsager, The spontaneous magnetization of the Ising model, Suppl. Nuovo. Cimento 6 (1949) 241–243.

[36] R. Peierls, On Ising's model of ferromagnetism, Proc. Cambridge Philos. Soc. 32 (1936) 477–48⌐.

[37] J.C.T. Pool, Editor, Mathematical Aspects of Statistical Mechanics, SIAM–AMS Proc. Vol. 5 (American Math. Society, Providence, RI, 1972).

[38] C.J. Preston, An application of the GHS inequalities to show the absence of phase transition for Ising spin systems, Comm. Math. Phys. 35 (1974) 253–255.

[39] C.J. Preston, Random Fields, Lecture Notes in Mathematics 534 (Springer, Berlin, 1976).

[40] D. Ruelle, Statistical mechanics of a one dimensional lattice gas, Comm. Math. Phys. 9 (1968) 267–278.

[41] D. Ruelle, Statistical Mechanics: Rigorous Results (Benjamin, New York, 1969).

[42] D. Ruelle, On the use of 'small external fields' in the problem of symmetry breakdown in statistical mechanics, Ann. Phys. 69 (1972) 364–374.

[43] D. Ruelle, Thermodynamic Formalism Encyclopedia of Mathematics Vol. 5 (Addison Wesley, New York, 1979).

[44] T.D. Schultz, D.C. Mattis and E.H. Lieb, The two dimensional Ising model as a soluble problem of many fermions, Rev. Mod. Phys. 36 (1964) 856–871.

[45] J. Slawny, Analyticity and uniqueness for spin 1/2 classical ferromagnetic systems at low temperatures, Comm. Math. Phys. 34 (1973) 271–296.

[46] J. Slawny, A family of equilibrium states relevant to low temperature behavior of spin 1/2 classical ferromagnets and breaking of translational symmetry, Comm. Math. Phys. 37 (1974) 297–305.

[47] D.J. Thouless, Long range order in one dimensional systems, Phys. Rev. 187 (1969) 732–733.

[48] H. van Beijeren, Interface sharpness in the Ising system, Comm. Math. Phys. 40 (1975) 1–6.

[49] C.N. Yang, The spontaneous magnetization of a two dimensional Ising model, Phys. Rev. 85 (1952) 806–816.

II. Other infinite particle systems

[50] R. Arratia, Coalescing Brownian motions on the line, PH.D. thesis, U. of Wisconsin (1979).

[51] R. Arratia, Limiting point processes for coalescing and annhilating random walks, preprint (1980).

[52] M. Bramson and L. Gray, A note on the survival of the long range contact process, preprint (1980).

[53] M. Bramson and D. Griffeath, Renormalizing the three dimensional voter model, Ann. Probability 7 (1979) 418–432.

[54] M. Bramson and D. Griffeath, Clustering and dispersion rates for some interacting particle systems on Z, Ann. Probability 8 (1980) 183–213.

[55] M. Bramson and D. Griffeath, Asymptotics for interacting particle systems on Z^d, preprint (1980).

[56] R. Durrett, Conditioned limit theorems for some null recurrent Markov processes, Ann. Probability 6 (1978) 798–828.

[57] R. Durrett, An infinite particle system with additive interactions, Advances Appl. Probability 11 (1979) 355–383.

[58] R. Durrett, On the growth of one dimensional contact processes, Ann. Probability 8 (1980) 890–907.

[59] R. Durrett and T. Liggett, The shape of the limit set in Richardson's growth model, Ann. Probability, to appear.

[60] R. Durrett and D. Griffeath, Contact processes in several dimensions, preprint (1980).

[61] J. Fleischman, Limiting distributions for branching random fields, Trans. Am. Math. Soc. 239 (1970) 353–389.

[62] L. Gray, Translation invariant spin flip processes and nonuniqueness, Z. Wahr scheinlichkeitstheorie 51 (1980) 171–184.

[63] L. Gray and D. Griffeath, A stability criterion for attractive nearest neighbor spin systems on Z, preprint (1980).

[64] D. Griffeath, Additive and Cancellative Interacting Particle Systems, Lecture Notes in Mathematics 724 (Springer, Berlin, 1979).

[65] T. Harris, Nearest neighbor Markov interaction processes on multidimensional lattices, Advances in Math. 9 (1972) 66–89.

[66] T. Harris, Contact interactions on a lattice, Ann. Probability 2 (1974) 969–988.

[67] T. Harris, On a class of set valued Markov processes, Ann. Probability 4 (1976) 175–194.

[68] T. Harris, A correlation inequality for Markov processes in partially ordered spaces, Ann. Probability 5 (1977) 451–454.

[69] T. Harris, Additive set valued Markov processes and percolation methods, Ann. Probability 6 (1978) 355–378.

[70] R. Holley, An ergodic theorem for interacting systems with attractive interactions, Z. Wahrscheinlichkeitstheorie 24 (1972) 325–334.

[71] R. Holley, Remarks on the FKG inequalities, Comm. Math. Phys. 36 (1974) 227–231.

[72] R. Holley and T. Liggett, Ergodic theorems for weakly interacting systems and the voter model, Ann. Probability 3 (1975) 643–663.

[73] R. Holley and T. Liggett, Generalized potlatch and smoothing processes, preprint (1980).

[74] R. Holley and D. Stroock, A martingale approach to infinite systems of interacting processes, Ann. Probability 4 (1976) 195–228.

[75] R. Holley and D. Stroock, In one and two dimensions every stationary distribution for the stochastic Ising model is a Gibbs state, Comm. Math. Phys. 55 (1977) 37–45.

[76] R. Holley and D. Stroock, Invariance principles for some infinite particle systems, Proc. of a conference at Northwestern, to appear.

[77] R. Holley and D. Stroock, Central limit phenomena of various interacting systems, preprint (1978).

[78] R. Holley and D. Stroock, Generalized Ornstein–Uhlenbeck processes and infinite particle branching Brownian motions, Publ. R.I.M.S. Kyoto 14 (1978).

[79] R. Holley and D. Stroock, Dual processes and tneir application to infinite interacting systems, Advances in Math. 32 (1979) 149–174.

[80] C.T. Hsaio, The stochastic time evolution of Gaussian interacting systems, Ph.D. thesis, Cornell University (1979).

[81] C.T. Hsaio, Infinite systems with locally additive interaction of components, preprint (1980).

[82] C.T. Hsaio, The stochastic processes with Gaussian interaction of components, preprint (1980).

[83] F.P. Kelley, The asymptotic behavior of an invasion process, J. Appl. Probability 14 (1977) 584–590.

[84] H. Kesten, On the number of self-avoiding random walks, J. Math. Phys. 4 (1963) 960–969.

[85] T. Liggett, Existence theorems for infinite particle systems, Trans. AMS. 165 (1972) 471–481.

[86] T. Liggett, The stochastic evolution of infinite systems of interacting particles, in: Ecole d'Été de Probabilités Saint-Flour VI, Lecture Notes in Mathematics 598 (Springer, Berlin, 1976).

[87] T. Liggett and F. Spitzer, Ergodic theorems for coupled random walks and other systems with locally interacting components, preprint (1980).

[88] P. Major, Renormalizing the voter model (1980).

[89] N. Matloff, Ergodicity conditions for a dissonant voter model, Ann. Probability 5 (1977) 371–386.

[90] S. Sawyer, Branching diffusion processes in population genetics, Advances Appl. Probability (1976) 659–689.

[91] S. Sawyer, Results for the stepping stone model for migration in population genetics, Ann. Probability 4 (1976) 699–728.

[92] S. Sawyer, Rates of consolidation in a selectively neutral migration model, Ann. Probability 5 (1977) 486–493.

[93] S. Sawyer, A limit theorem for patch sizes in a selectively neutral migration model, J. Appl. Probability 16 (1979) 482–495.

[94] S. Sawyer and J. Felsenstein, A continuous migration model with stable demography, preprint (1980).

[95] R. Smythe and J. Wierman, First Passage Percolation on the Square Lattice, Lecture Notes in Mathematics 671 (Springer, Berlin, 1978).

[96] F. Spitzer, Interaction of Markov processes, Advances in Math. 5 (1970) 246–290.

[97] F. Spitzer, Infinite systems with locally interacting components, 1979 Wald Lectures, Ann. Probability, to appear.

[98] W. Sullivan, A unified existence and ergodic theorem for Markov evolution of random fields, Z. Wahrscheinlichkeitstheorie 31 (1974) 47–56.

[99] W. Feller, Probability Theory and its Applications, Vol. II (Wiley, New York, 1970).