Interacting Particle Systems

J.M. Swart

July 20, 2011

Contents

Preface

Interacting particle systems, in the sense we will be using the word in these lecture notes, are countable systems of locally interacting Markov processes. Each interacting particle system is define on a lattice: a countable set with (usually) some concept of distance defined on it; the canonical choice is the d-dimensional integer lattice $\mathbb{Z}^d$. On each point in this lattice, there is situated a continuous-time Markov process with a finite state space (often even of cardinality two) whose jump rates depend on the states of the Markov processes on near-by sites. Interacting particle systems are often used as extremely simplified 'toy models' for stochastic phenomena that involve a spatial structure.

Although the definition of an interacting particle system often looks very simple, and problems of existence and uniqueness have long been settled, it is often surprisingly difficult to prove anything nontrivial about its behavior. With a few exceptions, explicit calculations tend not to be feasible, so one has to be satisfied with qualitative statements and some explicit bounds. Despite intensive research for over more than thirty years, some easy-to-formulate problems still remain open while the solution of others has required the development of nontrivial and complicated techniques.

Luckily, as a reward for all this, it turns out that despite their simple rules, interacting particle systems are often remarkably subtle models that capture the sort of phenomena one is interested in much better than might initially be expected. Thus, while it may seem outrageous to assume that "Plants of a certain type occupy points in the square lattice $\mathbb{Z}^2$, live for an exponential time with mean one, and place seeds on unoccupied neighboring sites with rate λ" it turns out that making the model more realistic often does not change much in its overall behavior. Indeed, there is a general philosophy in the field, that is still unsufficiently understood, which says that interacting particle systems come in 'universality classes' with the property that all models in one class have roughly the same behavior.

As a mathematical discipline, the subject of interacting particle systems is still relatively young. It started around 1970 with the work of R.L. Dobrushin and F. Spitzer,, with many other authors joining in during the next few years. By 1975, general existence and uniqueness questions had been settled, four classic models had been introduced (the exclusion process, the stochastic Ising model, the voter model and the contact process), and elementary (and less elementary) properties of these models had been proved. In 1985, when Liggett's published his famous book [Lig85], the subject had established itself as a mature field of study. Since then, it has continued to grow rapidly, to the point where it is impossible to accurately capture the state of the art in a single book. Indeed, it would be

possible to write a book on each of the four classic models mentioned above, while many new models have been introduced and studied.
While interacting particle systems, in the narrow sense indicated above, have apparently not been the subject of mathematical study before 1970, the subject has close links to some problems that are considerably older. In particular, the Ising model (without time evolution) has been studied since 1925 while both the Ising model and the contact process have close connections to percolation, which has been studied since the late 1950-ies. In recent years, more links between interacting particle systems and other, older subjects of mathematical research have been established, and the field continues to recieve new impulses not only from the applied, but also from the more theoretical side.

Chapter 1

Construction of interacting particle systems

In this section, we prove a general result on the construction, existence and uniqueness of interacting particle systems. As a preparation, we first review some necessary background theory about Markov processes and Poisson point sets. Proofs of these preliminary facts will mostly be omitted, although we sometimes give rough sketches of proofs when this is useful for developing intuition.

1.1 Probability on Polish spaces

By definition, a *Polish space* is a separable topological space E on which there exists a complete metric generating the topology. Polish spaces are particularly nice for doing probability theory on. We equip a Polish space E standardly with the *Borel-σ-field* $\mathcal{B}(E)$ generated by the open subsets of E. We let $B(E)$ denote space of bounded, real, $\mathcal{B}(E)$-measurable functions on E. Polish spaces have nice reproducing properties; for example, if E is a Polish space and F is a closed or an open subset of E, then the space F is also Polish (in the embedded topology). Also, if $E_1, E_2, \ldots$ is a finite or countably infinite sequence of Polish spaces, then the product space $E_1 \times E_2 \times \cdots$ equipped with the product topology is again Polish, and the Borel-σ-field on the product space coincides with the product-σ-field of the Borel-σ-fields on the individual spaces.

Let $\mathbb{E}$ be a Polish space and let $\mathcal{M}_1(E)$ be the space of probability measures on E, equipped with the topology of weak convergence. By definition, a set $\mathcal{R} \subset \mathcal{M}_1(E)$ is *tight* if

$$\forall \varepsilon > 0 \ \exists K \subset E \text{ s.t. } K \text{ is compact and } \mu(E \backslash K) \leq \varepsilon \ \forall \mu \in \mathcal{R}.$$

A well-known result says that the closure of $\mathcal{R}$ is compact (i.e., $\mathcal{R}$ is 'precompact') as a subset of $\mathcal{M}_1(E)$ if and only if $\mathcal{R}$ is tight. In particular, if $(\mu_n)_{n\geq 0}$ is a sequence of probability measures on E then we say that such a sequence is tight if the set $\{\mu_n : n \geq 0\} \subset \mathcal{M}_1(E)$ is tight. Note that each tight sequence of probability measures has a weakly convergent subsequence. Recall that a *cluster point* of a sequence is a limit of some subsequence of the sequence. We sometimes say 'weak cluster point' when we mean a 'cluster point in the topology of weak convergence'. One often needs tightness because of the following simple fact.

Lemma 1.1 (Tightness and weak convergence) *Let $(\mu_n)_{n\geq 0}$ be a tight sequence of probability measures on a Polish space E and assume that $(\mu_n)_{n\geq 0}$ has only one weak cluster point μ. Then μ_n converges weakly to μ.*

Note that if E is compact, then tightness comes for free, i.e., every sequence of probability measures on E is tight and $\mathcal{M}_1(E)$ is itself a compact space.

Let E, F be Polish spaces. By definition, a *probability kernel* from E to F is a function $K : E \times \mathcal{B}(F) \to \mathbb{R}$ such that

(i) $K(x,\,\cdot\,)$ is a probability measure on F for each $x \in E$,

(ii) $K(\,\cdot\,, A)$ is a real measurable function on E for each $A \in \mathcal{B}(F)$.

If $K(x, \mathrm{d}y)$ is a probability kernel on a Polish space E, then setting

$$Kf(x) := \int_E K(x, \mathrm{d}y) f(y) \qquad \big(x \in E \;\; f \in B(E)\big)$$

defines a linear operator $K : B(E) \to B(E)$. We sometimes use this notation also if f is not a bounded function, as long as the integral is well-defined for every x. If K, L are probability kernels on E, then we define the *composition* of K and L as

$$(KL)(x, A) := \int_E K(x, \mathrm{d}y) L(y, A) \qquad \big(x \in E \;\; f \in B(E)\big).$$

It is straightforward to check that this formula defines a probability kernel on E. If $K : B(E) \to B(E)$ and $L : B(E) \to B(E)$ are the linear operators associated with the probability kernels $K(x, \mathrm{d}y)$ and $L(x, \mathrm{d}y)$, then the linear operator associated with the composed kernel $(KL)(x, \mathrm{d}y)$ is just KL, the composition of the linear operators K and L.

Proposition 1.2 (Decomposition of probability measures) *Let E, F be Polish spaces and let μ be a probability measure on $E \times F$. Then there exist a (unique)* *probability measure ν on E and a (in general not unique) probability kernel K from E to F such that*

$$\int f \mathrm{d}\mu = \int_E \nu(\mathrm{d}x) \int_F K(x, \mathrm{d}y) f(x, y) \qquad \big(f \in B(E \times F)\big). \tag{1.1}$$

If K, K' are probability kernels from E to F such that (1.1) holds, then there exists a set $N \in \mathcal{B}(E)$ with $\nu(N) = 0$ such that $K(x, \cdot\,) = K(x, \cdot\,)$ for all $x \in E \backslash N$. Conversely, if ν is a probability measure on E and K is a probability kernel from E to F, then formula (1.1) defines a unique probability measure on $E \times F$.

Note that it follows obviously from (1.1) that

$$\nu(A) = \mu(A \times F) \qquad \big(A \in \mathcal{B}(E)\big),$$

i.e., ν is the first *marginal* of the probability measure μ.

If X and Y are random variables, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and taking values in E and F, respectively, then setting

$$\mu(A) := \mathbb{P}[(X, Y) \in A] \qquad \big(A \in \mathcal{B}(E \times F)\big)$$

defines a probability law on $E \times F$ which is called the joint law of X and Y. By Proposition 1.2, we may write μ in the form (1.1) for some probability law ν on E and probability kernel K from E to F. We observe that

$$\nu(A) = \mathbb{P}[X \in A] \qquad \big(A \in \mathcal{B}(E)\big),$$

i.e., ν is the law of X. We will often denote the law of X by $\mathbb{P}[X \in \cdot\,]$. Moreover, we introduce the notation

$$\mathbb{P}\big[Y \in A \,\big|\, X = x\big] := K(x, A) \qquad \big(x \in E,\ A \in \mathcal{B}(F)\big),$$

where $K(x, A)$ is the probability kernel from E to F defined in terms of μ as in (1.1). Note that $K(x, A)$ is defined uniquely for a.e. x with respect to the law of X. We call $\mathbb{P}[Y \in \cdot \,|\, X = x\big]$ the conditional law of Y given X. Note that with the notation we have just introduced, formula (1.1) takes the form

$$\mathbb{E}[f(X, Y)] = \int_E \mathbb{P}[X \in \mathrm{d}x] \int_F \mathbb{P}\big[Y \in \mathrm{d}y \,\big|\, X = x\big] f(x, y). \tag{1.2}$$

Closely related to this, one also defines

$$\mathbb{P}\big[Y \in A \,\big|\, X\big] := K(X, A) \qquad \big(A \in \mathcal{B}(F)\big).$$

Note that this is the random variable (defined on the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$) obtained by plugging X into the function $x \mapsto K(x, A)$.

If $f : F \to \mathbb{R}$ is a measurable function such that $E[|f(Y)|] < \infty$, then we let

$$\mathbb{E}\big[f(Y)\,\big|\,X = x\big] := \int_F \mathbb{P}\big[Y \in \mathrm{d}y\,\big|\,X = x\big] f(y)$$

denote the conditional expectation of $f(Y)$ given X. Note that for fixed f and Y, the map $x \mapsto \mathbb{E}\big[f(Y)\,\big|\,X = x\big]$ is a measurable real function on E. Plugging X into this function yields a random variable which we denote by $\mathbb{E}[f(Y)|X]$. We observe that for each $g \in B(E)$, one has

$$\begin{aligned}
\mathbb{E}\big[g(X)\mathbb{E}[f(Y)|X]\big] &= \int_E \mathbb{P}[X \in \mathrm{d}x] g(x) \mathbb{E}[f(Y)|X = x] \\
&= \int_E \mathbb{P}[X \in \mathrm{d}x] g(x) \int_F \mathbb{P}[Y \in \mathrm{d}y|X = x] f(y) \\
&= \int_E \mathbb{P}[X \in \mathrm{d}x] \int_F \mathbb{P}[Y \in \mathrm{d}y|X = x] g(x) f(y) \\
&= \int_{E\times F} \mathbb{P}[(X, Y) \in \mathrm{d}(x, y)] g(x) f(y) = \mathbb{E}[g(X) f(Y)].
\end{aligned}$$

Moreover, since $E[f(Y)|X]$ can be written as a function of X, it is easy to check that $E[f(Y)|X]$ is measurable with respect to the σ-field generated by X. One may take these properties as an alternative definition of $\mathbb{E}[f(Y)|X]$. More generally, if R is a real-valued random variable with $\mathbb{E}[|R|] < \infty$, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $\mathcal{G} \subset \mathcal{F}$ is a sub-σ-field, then there exists an a.s. (with respect to the underlying probability measure $\mathbb{P}$) unique random variable $\mathbb{E}[R|\mathcal{G}]$ such that $E[R|\mathcal{G}]$ is $\mathcal{G}$-measurable and

$$\mathbb{E}\big[G\,\mathbb{E}[R|\mathcal{G}]\big] = \mathbb{E}[GR] \qquad \forall \text{ bounded } \mathcal{G}\text{-measurable } G.$$

In the special case that $R = f(Y)$ and $\mathcal{G}$ is the σ-field generated by X one recovers $\mathbb{E}[f(Y)|X] = \mathbb{E}[R|\mathcal{G}]$.

1.2 Markov chains

Let E be a Polish space. By definition, a *Markov chain* with *state space* E is a discrete-time stochastic process $(X_k)_{k\geq 0}$ such that for all $0 \leq l \leq m \leq n$

$$\begin{aligned}
&\mathbb{P}\big[(X_l, \ldots, X_m) \in A,\ (X_m, \ldots, X_n) \in B\,\big|\,X_m\big] \\
&\quad = \mathbb{P}\big[(X_l, \ldots, X_m) \in A|\,X_m\big]\,\mathbb{P}(X_m, \ldots, X_n) \in B\,\big|\,X_m\big] \quad \text{a.s.}
\end{aligned} \tag{1.3}$$

for each $A \in \mathcal{B}(E^{m-l+1})$ and $B \in \mathcal{B}(E^{n-m+1})$. In words, formula (1.3) says that the past and the future are conditionally independent given the present. A similar definition applies to Markov chains $(X_k)_{k\in I}$ where $I \subset \mathbb{Z}$ is some interval (possibly unbounded on either side). It can be shown that (1.3) is equivalent to the statement that

$$\mathbb{P}\big[X_k \in A \,\big|\, (X_0,\ldots,X_{k-1})\big] = \mathbb{P}\big[X_k \in A \,\big|\, X_{k-1}\big] \quad \text{a.s.} \tag{1.4}$$

for each $k \geq 1$ and $A \in \mathcal{B}(E)$. For any sequence $(X_k)_{k\geq 0}$ of E-valued random variables, repeated application of (1.2) gives

$$\begin{aligned}
&\mathbb{E}\big[f(X_0,\ldots,X_n)\big] = \int \mathbb{P}\big[(X_0,\ldots,X_{n-1}) \in \mathrm{d}(x_0,\ldots,x_{n-1})\big] \\
&\quad \times \int \mathbb{P}\big[X_n \in \mathrm{d}x_n \,\big|\, (X_0,\ldots,X_{n-1}) = (x_0,\ldots,x_{n-1})\big] f(x_0,\ldots,x_n) \\
&\quad = \int \mathbb{P}\big[(X_0,\ldots,X_{n-2}) \in \mathrm{d}(x_0,\ldots,x_{n-2})\big] \\
&\qquad \times \int \mathbb{P}\big[X_{n-1} \in \mathrm{d}x_{n-1} \,\big|\, (X_0,\ldots,X_{n-2}) = (x_0,\ldots,x_{n-2})\big] \\
&\qquad \times \int \mathbb{P}\big[X_n \in \mathrm{d}x_n \,\big|\, (X_0,\ldots,X_{n-1}) = (x_0,\ldots,x_{n-1})\big] f(x_0,\ldots,x_n) \\
&= \int \mathbb{P}\big[X_0 \in \mathrm{d}x_0\big] \int \mathbb{P}\big[X_1 \in \mathrm{d}x_1 \,\big|\, X_0 = x_0\big] \int \mathbb{P}\big[X_2 \in \mathrm{d}x_2 \,\big|\, (X_0,X_1) = (x_0,x_1)\big] \\
&\qquad \times \cdots \times \int \mathbb{P}\big[X_n \in \mathrm{d}x_n \,\big|\, (X_0,\ldots,X_{n-1}) = (x_0,\ldots,x_{n-1})\big] f(x_0,\ldots,x_n).
\end{aligned}$$

If $(X_k)_{k\geq 0}$ is a Markov chain, then by (1.4) this simplifies to

$$\begin{aligned}
&\mathbb{E}\big[f(X_0,\ldots,X_n)\big] \\
&\quad = \int \mathbb{P}\big[X_0 \in \mathrm{d}x_0\big] \int \mathbb{P}\big[X_1 \in \mathrm{d}x_1 \,\big|\, X_0 = x_0\big] \\
&\qquad \times \cdots \times \int \mathbb{P}\big[X_n \in \mathrm{d}x_n \,\big|\, X_{n-1} = x_{n-1}\big] f(x_0,\ldots,x_n).
\end{aligned}$$

As this formula shows, the law of a Markov chain $(X_k)_{k\geq 0}$ is uniquely determined by its *initial law* $\mathbb{P}[X_0 \in \cdot\,]$ and its *transition probabilities* $\mathbb{P}\big[X_n \in \mathrm{d}x_n \,\big|\, X_{n-1} = x_{n-1}\big]$ $(k \geq 1)$. By definition, a Markov chain is *time-homogeneous* if its transitition probabilities are the same in each time step, more precisely, if there exists a probability kernel $P(x, \mathrm{d}y)$ on E such that

$$\mathbb{P}\big[X_n \in \cdot \,\big|\, X_{n-1} = x\big] = P(x, \cdot\,) \quad \text{for a.e. } x \text{ w.r.t. } \mathbb{P}[X_{n-1} \in \cdot\,],$$

which is equivalent to

$$\mathbb{P}\big[X_n \in \cdot \,\big|\, X_{n-1}\big] = P(X_{n-1}, \cdot\,) \quad \text{a.s.} \tag{1.5}$$

We will usually be interested in time-homogeneous Markov chains only. In fact, we will often fix a probability kernel $P(x, \mathrm{d}y)$ on E and then be interested in all possible Markov chains with this *transition kernel* (and arbitrary initial law). Note that we can combine (1.4) and (1.5) in a single condition: a sequence $(X_k)_{k\geq 0}$ of E-valued random variables is a Markov chain with transition kernel $P(x, \mathrm{d}y)$ (and arbitrary initial law) if and only if

$$\mathbb{P}\big[X_k \in \cdot \,\big|\, (X_0, \ldots, X_{k-1})\big] = P(X_{k-1}, \,\cdot\,) \quad \text{a.s.} \quad (k \geq 1), \tag{1.6}$$

which is equivalent to

$$\mathbb{E}\big[f(X_k) \big|\, (X_0, \ldots, X_{k-1})\big] = Pf(X_{k-1}) \quad \text{a.s.} \quad \big(k \geq 1,\ f \in B(E)\big), \tag{1.7}$$

where P denotes the linear operator from $B(E)$ to $B(E)$ associated with the kernel $P(x, \mathrm{d}y)$

If $(X_k)_{k\geq 0}$ is a Markov chain with transition kernel $P(x, \mathrm{d}y)$, and we let P^n denote the n-fold composition of the kernel / linear operator P with itself, where $P^0(x, \mathrm{d}y) := \delta_x(\mathrm{d}y)$ (the delta measure in x), then we may generalize (1.6) to

$$\mathbb{P}\big[X_{k+n} \in \cdot \,\big|\, (X_0, \ldots, X_k)\big] = P^n(X_k, \,\cdot\,) \quad \text{a.s.} \quad (k, n \geq 0), \tag{1.8}$$

which is equivalent to

$$\mathbb{E}\big[f(X_{k+n}) \,\big|\, (X_0, \ldots, X_k)\big] = P^n f(X_k) \quad \text{a.s.} \quad \big(k, n \geq 0,\ f \in B(E)\big). \tag{1.9}$$

1.3 Feller processes

Let E be a compact metrizable space. Such spaces are always separable and complete in any metric that generates the topology; in particular, they are therefore Polish. Let $\mathcal{C}(E)$ denote the space of continuous real functions on E, equipped with the supremumnorm

$$\|f\| := \sup_{x\in E} |f(x)| \qquad (f \in \mathcal{C}(E)).$$

We let $\mathcal{M}_1(E)$ denote the space of probability measures on E (equipped with the topology of weak convergence). We note that $\mathcal{C}(E)$ is a separable Banach space and that $\mathcal{M}_1(E)$ is a compact metrizable space.

By definition, a *continuous transition probability* on E is a collection $(P_t(x, \mathrm{d}y))_{t\geq 0}$ of probability kernels on E such that

(i) $(x, t) \mapsto P_t(x, \,\cdot\,)$ is a continuous map from $E \times [0, \infty)$ into $\mathcal{M}_1(E)$,

(ii) $\displaystyle\int_E P_s(x, \mathrm{d}y) P_t(y, \mathrm{d}z) = P_{s+t}(x, \mathrm{d}z) \quad \text{and} \quad P_0(x, \cdot\,) = \delta_x \quad (x \in E,\ s, t \geq 0).$

Each continuous transition probability defines a semigroup $(P_t)_{t\geq 0}$ by

$$P_t f(x) := \int_E P_t(x, \mathrm{d}y) f(y) \qquad \big(f \in B(E)\big). \tag{1.10}$$

It follows from the continuity of the transition probability that the operators P_t map the space $\mathcal{C}(E)$ into itself. Moreover, the collection of linear operators $(P_t)_{t\geq 0}$ associated with a continuous transition probability satisfies

(i) $\lim_{t\to 0} \|P_t f - f\| = 0 \qquad (f \in \mathcal{C}(E))$,

(ii) $P_s P_t f = P_{s+t} f \quad$ and $\quad P_0 f = f$,

(iii) $f \geq 0$ implies $P_t f \geq 0$,

(iv) $P_t 1 = 1$,

and conversely, each collection of linear operators $P_t : \mathcal{C}(E) \to \mathcal{C}(E)$ with these properties corresponds to a unique continuous transition probability on E. Such a collection of linear operators $P_t : \mathcal{C}(E) \to \mathcal{C}(E)$ is called a *Feller semigroup*.

By definition, the *generator* of a Feller semigroup is the operator

$$Gf := \lim_{t\to 0} t^{-1}\big(P_t f - f\big),$$

which is defined only for functions $f \in \mathcal{D}(G)$, where

$$\mathcal{D}(G) := \big\{f \in \mathcal{C}(E) : \text{the limit } \lim_{t\to 0} t^{-1}\big(P_t f - f\big) \text{ exists}\big\}.$$

Here, when we say that the limit exists, we mean the limit in the topology on $\mathcal{C}(E)$, which is defined by the supremumnorm $\|\cdot\|$.

We say that an operator A on $\mathcal{C}(E)$ with domain $\mathcal{D}(A)$ satisfies the *maximum principle* if, whenever a function $f \in \mathcal{D}(A)$ assumes its maximum over E in a point $x \in E$, we have $Af(x) \leq 0$. We say that a linear operator A with domain $\mathcal{D}(A)$ acting on a Banach space $\mathcal{V}$ (in our example the space $\mathcal{C}(E)$ equipped with the supremunorm) is *closed* if and only if its graph $\{(f, Af) : f \in \mathcal{D}(A)\}$ is a closed subset of $\mathcal{V} \times \mathcal{V}$. If the domain $\mathcal{D}(A)$ of a linear operator A is the whole Banach space $\mathcal{V}$, then A is closed if and only if A is *bounded* i.e., there exists a constant $C < \infty$ such that $\|Af\| \leq C\|f\|$. Note that as a consequence, the domain of a closed unbounded operator can never be the whole space $\mathcal{V}$. By definition, a linear operator A with domain $\mathcal{D}(A)$ on a Banach space $\mathcal{V}$ is *closable* if the closure of its graph (as a subset of $\mathcal{V}\times\mathcal{V}$) is the graph of a linear operator $\overline{A}$ with domain $\mathcal{D}(\overline{A})$, called the *closure* of A. The following proposition collects some important facts about Feller semigroups. Proofs of these facts can be found in [EK86, Sections 1.1, 1.2 and 4.2].

Proposition 1.3 (Feller semigroups) *A linear operator G on $\mathcal{C}(E)$ is the generator of a Feller semigroup $(P_t)_{t\geq 0}$ if and only if*

(i) $1 \in \mathcal{D}(G)$ *and* $G1 = 0$.

(ii) *G satisfies the maximum principle.*

(iii) *$\mathcal{D}(G)$ is dense in $\mathcal{C}(E)$.*

(iv) *For every $f \in \mathcal{D}(G)$ there exists a continuously differentiable function $t \mapsto u_t$ from $[0,\infty)$ into $\mathcal{C}(E)$ such that $u_0 = f$, $u_t \in \mathcal{D}(G)$, and $\frac{\partial}{\partial t}u_t = Gu_t$ for each $t \geq 0$.*

(v) *G is closed.*

Here, in point (iv), the differentiation with respect to t is in the Banach space $\mathcal{C}(E)$. If G is the generator of a Feller semigroup $(P_t)_{t\geq 0}$, then for each $f \in \mathcal{D}(G)$, the solution u to the equation $u_0 = f$, $u_t \in \mathcal{D}(G)$, and $\frac{\partial}{\partial t}u_t = Gu_t$ $(t \geq 0)$ is in fact unique and given by $u_t = P_t f$. Moreover, for each $t \geq 0$, the operator P_t is the closure of $\{(f, P_t f) : f \in \mathcal{D}(G)\}$.

If in addition to the properties from Proposition 1.3, the operator G is bounded (or equivalently, $\mathcal{D}(G) = \mathcal{C}(E)$), then one has

$$P_t = e^{Gt} := \sum_{n=0}^{\infty} \frac{1}{n!} G^n t^n \qquad (t \geq 0),$$

where the infinite sum converges absolutely in the operator norm, defined as $\|A\| := \sup\{\|Af\| : \|f\| \leq 1\}$. In many interesting cases, however, G will not be bounded and hence not everywhere defined. In these cases, it us usually not feasible to explicitly write down the full domain of the generator of a Feller semigroup. Instead, one often first defines a 'pregenerator' which is defined for a smaller class of functions, and then constructs the 'full generator' by taking the closure of the pregenerator.

The next result, which is a version of the Hille-Yosida theorem, is often useful. For a proof, we refer to Sections 1.1, 1.2 and 4.2, and in particular Theorem 4.2.2 of [EK86].

Theorem 1.4 (Hille-Yosida) *A linear operator G on $\mathcal{C}(E)$ with domain $\mathcal{D}(G)$ is closable and its closure $\overline{G}$ is the generator of a Feller semigroup if and only if*

(i) $(1,0) \in \overline{\{(f, Gf) : f \in \mathcal{D}(G)\}}$ *(i.e., $(1,0)$ is in the closure of the graph of G).*

(ii) *G satisfies the maximum principle.*

(iii) *$\mathcal{D}(G)$ is dense in $\mathcal{C}(E)$.*

(iv) *There exists an $r \in (0,\infty)$ and a dense subspace $\mathcal{D} \subset \mathcal{C}(E)$ with the property that for every $f \in \mathcal{D}$ there exists a $p_r \in \mathcal{D}(G)$ such that $(r - G)p_r = f$.*

If $\overline{G}$ is the generator of a Feller semigroup $(P_t)_{t\geq 0}$, then for each $r \in (0,\infty)$, the space $\{(r-G)p : p \in \mathcal{D}(G)\}$ is a dense linear subspace of $\mathcal{C}(E)$. For each $f \in \mathcal{C}(E)$, there is a unique $p_r \in \mathcal{D}(\overline{G})$ such that $(r - \overline{G})p_r = f$ and this function is given by $p_r = \int_0^\infty e^{-rt} P_t f \,\mathrm{d}t$.

By definition, we let $\mathcal{D}_E[0,\infty)$ denote the space of all functions from $[0,\infty)$ to E that are right-continuous with left limits, i.e., $\mathcal{D}_E[0,\infty)$ is the space of functions $w : [0,\infty) \to S$ such that

$$\begin{array}{lll} \text{(i)} & \lim_{t\downarrow s} w_t = w_s & (s \geq 0), \\ \text{(ii)} & \lim_{t\uparrow s} w_t =: w_{s-} \text{ exists} & (s > 0). \end{array}$$

We call $\mathcal{D}_E[0,\infty)$ the space of *cadlag* functions from $[0,\infty)$ to E. (After the French **c**ontinue **à d**roit, **l**imite **à g**auche.) It is possible to equip this space with a (rather natural) topology, called the *Skorohod topology*, such that $\mathcal{D}_E[0,\infty)$ is a Polish space and the Borel-σ-field on $\mathcal{D}_E[0,\infty)$ is generated by the coordinate projections $w \mapsto w_t$ $(t \geq 0)$; we will skip the details.
By definition, we say that an E-valued stochastic process $(X_t)_{t\geq 0}$ defined on some underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ has *cadlag sample paths* if for every $\omega \in \Omega$, the function $t \mapsto X_t(\omega)$ is cadlag. We may view such a stochastic process as a single random variable, taking values in the Polish space $\mathcal{D}_E[0,\infty)$. Now

$$\mathbb{P}\big[(X_t)_{t\geq 0} \in A\big] \qquad \big(A \in \mathcal{B}(\mathcal{D}_S[0,\infty))\big)$$

is a probability law on $\mathcal{D}_E[0,\infty)$ called the *law* of the process $(X_t)_{t\geq 0}$. Since the Borel-σ-field on $\mathcal{D}_E[0,\infty)$ is generated by the coordinate projections, this law is uniquely determined by the *finite dimensional distributions*

$$\mathbb{P}\big[(X_{t_1}, \ldots, X_{t_n}) \in A\big] \qquad (A \subset E^n).$$

We recall that a *filtration* is a collection $(\mathcal{F}_t)_{t\geq 0}$ of σ-fields such that $s \leq t$ implies $\mathcal{F}_s \subset \mathcal{F}_t$. If $(X_t)_{t\geq 0}$ is a stochastic process, then the *filtration generated by* $(X_t)_{t\geq 0}$ is defined as

$$\mathcal{F}_t := \sigma\big(X_s : 0 \leq s \leq t\big) \qquad (t \geq 0),$$

i.e., $\mathcal{F}_t$ is the σ-field generated by the random variables $(X_s)_{0\leq s\leq t}$.
By definition, a *Feller process* associated to a given Feller semigroup $(P_t)_{t\geq 0}$ is a stochastic process $(X_t)_{t\geq 0}$ with values in E and cadlag sample paths, such that (compare (1.9))

$$\mathbb{E}\big[f(X_t)\,\big|\,\mathcal{F}_s\big] = P_{t-s}f(X_s) \quad \text{a.s.} \quad \big(s\leq t,\ f\in\mathcal{C}(E)\big), \tag{1.11}$$

where $(\mathcal{F}_t)_{t\geq 0}$ is the filtration generated by $(X_t)_{t\geq 0}$. It can be shown that if $(P_t)_{t\geq 0}$ is a Feller semigroup, then for each probability law μ on E there exists a unique (in law) Feller process associated to $(P_t)_{t\geq 0}$ with *initial law* $\mathbb{P}[X_0\in\cdot\,]=\mu$. Feller processes have many nice properties, such as the strong Markov property.

1.4 Poisson point processes

Let E be a Polish space. Recall that a sequence of finite measures μ_n *converges weakly* to a limit μ, denoted as $\mu_n \Rightarrow \mu$, if and only if

$$\int f\mathrm{d}\mu_n \underset{n\to\infty}{\longrightarrow} \int f\mathrm{d}\mu \qquad \big(f\in\mathcal{C}_\mathrm{b}(E)\big),$$

where $\mathcal{C}_\mathrm{b}(E)$ denotes the space of bounded continuous real functions on E. We let $\mathcal{M}(E)$ denote the space of finite measures on E, equipped with the topology of weak convergence. It can be shown that $\mathcal{M}(E)$ is Polish and the Borel-σ-field $\mathcal{B}(\mathcal{M}(E))$ on $\mathcal{M}(E)$ coincides with the σ-field generated by the random variables $\mu\mapsto\mu(A)$ with $A\in\mathcal{B}(E)$. We let

$$\mathcal{N}(E) := \big\{\nu\in\mathcal{M}(E) : \exists n\geq 0,\ x_1,\ldots,x_n\in E \text{ s.t. } \nu=\sum_{i=1}^n \delta_{x_i}\big\}$$

denote the space of all *counting measures* on E, i.e., all measures that can be written as a finite sum of delta-measures. Being a closed subset of $\mathcal{M}(E)$, the space $\mathcal{N}(E)$ is again Polish.

For any counting measure $\nu\in\mathcal{N}(E)$ and $f\in B(E)$ we introduce the notation

$$f^\nu := \prod_{i=1}^n f(x_i) \qquad \text{where} \quad \nu=\sum_{i=1}^n \delta_{x_i},$$

with $f^0 := 1$ (where 0 denotes the counting measure that is identically zero). It is easy to see that $f^\nu f^{\nu'} = f^{\nu+\nu'}$. Let $\nu=\sum_{i=1}^n\delta_{x_i}$ be a counting measure, let $\phi\in B(E)$ satisfy $0\leq\phi\leq 1$, and let $\chi_1,\ldots,\chi_n$ be independent *Bernoulli random* *variables* (i.e., random variables with values in $\{0,1\}$) with $\mathbb{P}[\chi_i = 1] = \phi(x_i)$. Then the random counting measure

$$\nu' := \sum_{i=1}^{n} \chi_i \delta_{x_i}$$

is called a *ϕ-thinning* of the counting measure ν. Note that

$$\mathbb{P}[\nu' = 0] = \prod_{i=1}^{n} \mathbb{P}[\chi_i = 0] = (1-\phi)^\nu.$$

More generally, one has

$$\mathbb{E}\big[(1-f)^{\nu'}\big] = (1-f\phi)^\nu \qquad \big(f \in B(E),\ 0 \leq f \leq 1\big). \tag{1.12}$$

(Setting $f = 1$ here yields the previous formula.) To see this, note that if $\chi'_1, \ldots, \chi'_n$ are Bernoulli random variables with $\mathbb{P}[\chi'_i = 1] = f(x_i)$, independent of each other and of the χ_i's, and

$$\nu'' := \sum_{i=1}^{n} \chi'_i \chi_i \delta_{x_i},$$

then, since conditional on ν', the measure ν'' is distributed as an f-thinning of ν', one has

$$\mathbb{P}\big[\nu'' = 0\big] = \mathbb{E}\big[(1-f)^{\nu'}\big],$$

while on the other hand, since ν'' is an $f\phi$-thinning of ν, one has $\mathbb{P}[\nu'' = 0] = (1-f\phi)^\nu$. One can prove that (1.12) characterizes the law of the random counting measure ν' uniquely, and in fact suffices to check (1.12) for continuous $f : E \to [0,1]$.

Proposition 1.5 (Poisson counting measure) *Let E be a Polish space and let μ be a finite measure on E. Then there exists a random counting measure ν on E whose law is uniquely characterized by*

$$\mathbb{E}\big[(1-f)^\nu\big] = e^{-\int f \mathrm{d}\mu} \qquad \big(f \in B(E),\ 0 \leq f \leq 1\big). \tag{1.13}$$

If $A_1, \ldots, A_n$ are disjoint measurable subsets of E, then $\nu(A_1), \ldots, \nu(A_n)$ are independent Poisson distributed random variables with mean $\mathbb{E}[\nu(A_i)] = \mu(A_i)$ $(i = 1, \ldots, n)$.

Proof (sketch) Let N be a Poisson distributed random variable with mean $\mu(E)$ and let $X_1, X_2, \ldots$ be i.i.d. random variables with law $\mathbb{P}[X_i \in \cdot\,] = \mu(E)^{-1}\mu(\,\cdot\,)$, independent of N. Then one can check that the random counting measure

$$\nu := \sum_{i=1}^{N} \delta_{X_i}$$

has all the desired properties. ∎

The random measure ν whose law is defined in Proposition 1.5 is called a *Poisson counting measure* with *intensity* μ. In fact, to prove that a given random counting measure ν is a Poisson point measure with intensity μ, it suffices to check (1.13) for continuous $f : E \to [0,1]$.

Proposition 1.6 (Poisson as limit of thinning) *For $n \geq 1$, let ε_n be a non-negative constants and let $\nu_n := \sum_{i=1}^{N_n} \delta_{x_{n,i}}$ be counting measures on some Polish space E. Assume that $\varepsilon_n \to 0$ and*

$$\mu_n := \varepsilon_n \sum_{i=1}^{N_n} \delta_{x_{n,i}} \underset{n\to\infty}{\Longrightarrow} \mu$$

for some finite measure μ. Let ν'_n be a thinning of ν_n with the constant function ε_n. Then the $\mathcal{N}(E)$-valued random variables ν'_n converge weakly in law to a Poisson point measure with intensity μ.

Proof (sketch) For any $f \in \mathcal{C}(E)$ satisfying $c \leq f \leq 1$ for some $c > 0$, by (1.12), one has

$$\mathbb{E}\big[(1-f)^{\nu'_n}\big] = (1-\varepsilon_n f)^{\nu_n} = e^{\int \log(1-\varepsilon_n f)\mathrm{d}\nu_n} = e^{\int \varepsilon_n^{-1}\log(1-\varepsilon_n f)\mathrm{d}\mu_n} \underset{n\to\infty}{\longrightarrow} e^{-\int f\mathrm{d}\mu},$$

which (with some care) follows from the facts that $\varepsilon_n^{-1}\log(1-\varepsilon_n f) \to -f$ and $\mu_n \Rightarrow \mu$. By approximation, one obtains (1.13) for all continuous funtions $f : E \to [0,1]$, which suffices to prove that the ν'_n converge weakly in law to a Poisson point measure with intensity μ. ∎

Lemma 1.7 (Sum of independent Poisson counting measures) *Let E be a Polish space and let ν_1, ν_2 be independent Poisson counting measures on E with intensities μ_1, μ_2, respectively. Then $\nu_1 + \nu_2$ is a Poisson counting measure with intensity $\mu_1 + \mu_2$.*

Proof (sketch) One can straightforwardly check this from (1.13). Note that thinnings have a similar property, so the statement is also rather obvious from our approximation of Poisson counting measures with thinnings. ∎

Set

$$\mathcal{N}_1(E) := \big\{\nu \in \mathcal{N}(E) : \nu(\{x\}) \in \{0,1\} \ \forall x \in E\big\}.$$

Since $\mathcal{N}_1(E)$ is an open subset of $\mathcal{N}(E)$, it is a Polish space. We can identify elements of $\mathcal{N}_1(E)$ with finite subsets of E; indeed, $\nu \in \mathcal{N}_1(E)$ if and only if $\nu = \sum_{x\in\Delta} \delta_x$ for some finite $\Delta \subset E$. We skip the proof of the following lemma.

Lemma 1.8 (Poisson point set) *Let μ be a finite measure on a Polish space E and let ν be a Poisson counting measure with intensity μ. Then $\mathbb{P}[\nu \in \mathcal{N}_1(E)] = 1$ if and only if μ is* nonatomic, *i.e., $\mu(\{x\}) = 0$ for all $x \in E$.*

If μ is a nonatomic measure on some Polish space, ν is a Poisson counting measure with intensity μ, and Δ is the random finite set associated with ν, then we call Δ a *Poisson point set* with intensity μ.

If E is a locally compact space and μ is a locally finite measure on E (i.e., a measure such that $\mu(K) < \infty$ for each compact $K \subset E$), then Poisson counting measures and Poisson point sets with intensity μ are defined analogously to the finite measure case, where in (1.13), we now only allow functions f with compact support. We will in particular be interested in the case that $E = [0,\infty)$ and μ is a multiple of Lebesgue measure.

Lemma 1.9 (Exponential times) *Let $r > 0$ be a constant and let $(\sigma_k)_{k\geq 1}$ be i.i.d. exponentially distributed random variables with mean $\mathbb{E}[\sigma_k] = 1/r$ $(k \geq 1)$. Set $\tau_n := \sum_{k=1}^n \sigma_k$ $(n \geq 1)$. Then $\{\tau_n : n \geq 1\}$ is a Poisson point set on $[0,\infty)$ with density $r\,\mathrm{d}t$, where $\mathrm{d}t$ denotes Lebesgue measure.*

Proof (sketch) Fix $\varepsilon > 0$ and set $\nu := \sum_{i=1}^\infty \delta_{\varepsilon i}$. Let ν' be a thinning of ν with the constant function $r\varepsilon$. Then it is not hard to see that the distances between consecutive points in ν' are independent and geometrically distributed. Letting $\varepsilon \to 0$, we observe that ν' converges to a Poisson point set with density $r\,\mathrm{d}t$ and that the distances between consecutive points become exponentially distributed. ∎

Exercise 1.10 Let μ be a finite measure on a Polish space E and let ν be a Poisson counting measure with intensity μ. Let $A_1,\ldots,A_n$ be disjoint, measurable subsets of E and define $\nu_i(B) := \nu(B \cap A_i)$ $(B \in \mathcal{B}(E))$. Show that $\nu_1,\ldots\nu_n$ are independent.

1.5 Poisson construction of Markov processes

The interacting particle systems that we will consider in these lecture notes will be Feller processes with state space $\{0,1\}^{\mathbb{Z}^d}$, which in the product topology is a compact metrizable space. We will construct these Feller processes using a graphical representation based on Poisson point sets. To prepare for this, in the present section, we will show how Feller processes with finite state spaces can be constructed using Poisson point sets.

Let S be a finite set. If μ is a probability measure on S and $x \in S$, then we write $\mu(x) := \mu(\{x\})$ and likewise, if $K(x,A)$ is a probability kernel on S, then we write $K(x,y) := K(x,\{y\})$. Now probability kernels correspond to matrices and the composition of two kernels corresponds to the usual matrix product.

Let $(P_t)_{t\geq 0}$ be a continuous transition probability on S, or equivalently, a Feller semigroup on $\mathcal{C}(S)$. Since S is finite, any function $f : S \to \mathbb{R}$ is in fact continuous and there is no difference between $B(S)$ and $\mathcal{C}(S)$. Likewise, any probability kernel on S is continuous. Thus, in this more simple context, a continuous transition probability on S is just a collection $(P_t(x,y))_{t\geq 0}$ of probability kernels on S such that

$$\begin{array}{lll}\text{(i)} & \lim_{t\downarrow 0} P_t(x,y) = P_0(x,y) = \delta_x(y) & (x,y \in S), \\ \text{(ii)} & \sum_y P_s(x,y)P_t(y,z) = P_{s+t}(x,z) & (s,t \geq 0,\ x,z \in S).\end{array}$$

We will also call the associated Feller semigroup of linear operators $(P_t)_{t\geq 0}$ a *Markov semigroup*. (Usually, this is a more general, and less precisely defined term than Feller semigroup, since it does not entail any continuity assumptions, but in the present set-up of finite state spaces, continuity (in space) is not an issue.)

More generally, any collection $(A_t)_{t\geq 0}$ of linear operators on some finite-dimensional linear space such that $A_sA_t = A_{s+t}$ and $\lim_{t\downarrow 0} A_t = A_0 = 1$ (where 1 denotes the identity operator) is called a *linear semigroup*. One can show that each such linear semigroup is of the form

$$A_t = e^{Gt} := \sum_{n=0}^{\infty} \frac{1}{n!} G^n t^n \qquad (t \geq 0),$$

where

$$Gf := \lim_{t\downarrow 0} t^{-1}\big(A_t f - f\big)$$

is the *generator* of $(A_t)_{t\geq 0}$.

Proposition 1.11 (Markov generators) *Let S be a finite set and let G be a linear operator on $B(S)$. Then G is the generator of a Markov semigroup if and only if there exist nonnegative constants $r(x,y)$ $(x,y \in S,\ x \neq y)$ such that*

$$Gf(x) = \sum_{y \in S} r(x,y)\big(f(y) - f(x)\big) \qquad \big(x \in S,\ f \in B(S)\big). \tag{1.14}$$

Proof (sketch) Let $G(x,y)$ be the matrix associated with G. Then

$$P_t f(x) = f(x) + t \sum_y G(x,y) f(y) + O(t^2) \quad \text{as } t \to 0 \qquad \big(x \in S,\ f \in B(S)\big).$$

Now the condition that $P_t f \geq 0$ for all $f \geq 0$ implies that $Gf(x) \geq 0$ whenever $f(x) = 0$, hence $G(x,y) \geq 0$ for each $x \neq y$. Moreover, the condition that $P_t 1 = 1$ implies that

$$1 = 1 + t \sum_y G(x,y) + O(t^2) \quad \text{as } t \to 0 \qquad (x \in S),$$

which shows that $\sum_y G(x,y) = 0$ for each x. Setting $r(x,y) := G(x,y)$ for $x \neq y$ and using the fact that $G(x,x) = -\sum_{y \neq x} r(x,y)$, we see that G can be cast in the form (1.14). The fact that conversely, each generator of this form defines a Markov semigroup will follow from our explicit construction of the associated Markov process below. ∎

We call $r(x,y)$ the *rate* of *jumps* from x to y. By applying (1.14) and (1.15) to functions of the form $f(x) := 1_{\{x=y\}}$, we see that if $(X^x_t)_{t \geq 0}$ denotes the Markov process started in the initial law $X^x_0 := 1$, then

$$\mathbb{P}[X^x_t = y] = r(x,y)t + O(t^2) \quad \text{as } t \to 0 \quad (x,y \in S,\ x \neq y).$$

This says that if we start the process in the state x, then for small t, the probability that we jump from x to y somewhere in the interval $(0,t)$ is $tr(x,y)$ plus a term of order t^2.

Let S be a finite set, let $(X_t)_{t \geq 0}$ be a stochastic process with values in S and let $(P_t)_{t \geq 0}$ be a Markov semigroup on $B(S)$. Then, specializing from our definition of Feller processes, we say that $(X_t)_{t \geq 0}$ is a (time-homogeneous, continuous-time) Markov process with semigroup $(P_t)_{t \geq 0}$ if $(X_t)_{t \geq 0}$ has cadlag sample paths and

$$\mathbb{E}\big[f(X_t) \,\big|\, \mathcal{F}_s\big] = P_{t-s} f(X_s) \quad \text{a.s.} \quad \big(0 \leq s \leq t,\ f \in B(S)\big), \tag{1.15}$$

where $(\mathcal{F}_t)_{t \geq 0}$ is the filtration generated by $(X_t)_{t \geq 0}$. One can prove that for a given Markov semigroup $(P_t)_{t \geq 0}$ and probability law μ on S there exists a unique (in distribution) Markov process $(X_t)_{t\geq 0}$ with initial law $\mathbb{P}[X_0 \in \cdot\,] = \mu$ such that (1.15) holds.

We are now ready to state the first important theorem of this chapter, which tells us how to construct finite-state Markov processes based on a collection of Poisson point processes. Let S be a finite set and let $\mathcal{M}$ be a finite or countably infinite set whose elements are maps $m : S \to S$. Let $(r_m)_{m\in\mathcal{M}}$ be nonnegative constants and let Δ be a Poisson point set on $\mathcal{M} \times \mathbb{R} = \{(m,t) : m \in \mathcal{M},\ t \in \mathbb{R}\}$ with intensity $r_m \mathrm{d}t$, where $\mathrm{d}t$ denotes Lebesgue measure. Assume that

$$\sum_{m\in\mathcal{M}} r_m < \infty.$$

For $s \leq t$, set $\Delta_{s,t} := \Delta \cap (\mathcal{M} \times (s,t])$ and define random maps $\Psi_{\Delta,s,t} : S \to S$ by

$$\begin{aligned}&\Psi_{\Delta,s,t}(x) := m_n \circ \cdots \circ m_1(x)\\ &\quad\text{where}\quad \Delta_{s,t} := \{(m_1,t_1),\ldots,(m_n,t_n)\},\quad t_1 < \cdots < t_n,\end{aligned}$$

with the convention that $\Psi_{\Delta,s,t}(x) = x$ if $\Delta_{s,t} = \emptyset$. Note that since $\{t : (m,t) \in \Delta\}$ is a Poisson point set on $\mathbb{R}$ with finite intensity $\sum_{m\in\mathcal{M}} r_m$, the sets $\Delta_{s,t}$ are a.s. finite for each $s \leq t$. It is easy to check that $\Psi_{\Delta,t,u} \circ \Psi_{\Delta,s,t} = \Psi_{\Delta,s,u}$ $(s \leq t \leq u)$.

Theorem 1.12 (Poisson construction of Markov process) *Let X_0 be an S-valued random variable, independent of Δ. Then*

$$X_t := \Psi_{\Delta,0,t}(X_0) \qquad (t \geq 0) \tag{1.16}$$

defines a Markov process $(X_t)_{t\geq 0}$ with generator

$$Gf(x) = \sum_{m\in\mathcal{M}} r_m\big(f(m(x)) - f(x)\big) \qquad \big(x \in S,\ f \in B(S)\big). \tag{1.17}$$

It is not hard to see that each operator of the form (1.14) can be cast in the form (1.17) for some suitable finite collection $\mathcal{M}$ of maps $m : S \to S$ and nonnegative rates $(r_m)_{m\in\mathcal{M}}$. Thus, Theorem 1.12 can be used to prove that each collection of nonnegative rates $(r(x,y))_{x\neq y}$ defines a Markov semigroup and associated Markov process with generator given by (1.14). We note that while the rates $(r_m)_{m\in\mathcal{M}}$ determine the rates $(r(x,y))_{x\neq y}$ uniquely, the inverse problem is far from unique, i.e., there are usually many different ways of writing the generator G of a Markov process in the form (1.17). Once we have chosen a particular way of writing G in the form (1.17), Theorem 1.12 provides us with a natural way of *coupling* processes started in different initial states. Indeed, using the same Poisson point set Δ, setting $X^x_t := \Psi_{\Delta,0,t}(x)$ defines for each $x \in S$ a Markov process $(X^x_t)_{t\geq 0}$ with generator G started in the initial state $X^x_0 = x$, and all these processes (for different x) are in a natural way defined on one and the same underlying probability space (i.e., they are *coupled*). Such couplings are very important in the theory of interacting particle systems.

Proof of Theorem 1.12 Since $\{t : (m,t) \in \Delta\}$ is a Poisson point process on $\mathbb{R}$ with finite intensity $\sum_{m\in\mathcal{M}} r_m$, and since X is constant between times in this set and right-continuous at times in this set, it follows X has cadlag sample paths. (Note that not necessarily each time in $\{t : (m,t) \in \Delta\}$ is a time when X jumps, since it may happen that the associated map m maps X_{t-} onto itself.)
Next, we set

$$P_t(x,y) := \mathbb{P}[\Psi_{\Delta,0,t}(x) = y] \qquad (x,y \in S,\ t \geq 0).$$

Let $\mathcal{G}_t$ be the σ-field generated by the random variables X_0 and $\Delta_{0,t}$. Since $(X_s)_{0\leq s\leq t}$ is a function of X_0 and $\Delta_{0,t}$, it follows that $\mathcal{F}_t \subset \mathcal{G}_t$. Now fix $0 \leq s \leq t$ and look at the conditional law

$$\mathbb{P}[X_t \in \cdot \,|\, \mathcal{G}_s] = \mathbb{P}[X_t \in \cdot \,|\, X_0, \Delta_{0,s}].$$

Since X_0 is independent of Δ and since Δ is a Poisson point process, we see that X_0, $\Delta_{0,s}$ and $\Delta_{s,t}$ are independent. Since $\Delta_{s,t}$ is up to a time shift equally distributed with $\Delta_{0,t-s}$, it follows that

$$\mathbb{P}[X_t \in \cdot \,|\, \mathcal{G}_s] = \mathbb{P}\big[\Psi_{\Delta,s,t}(X_s) \in \cdot \,|\, X_0, \Delta_{0,s}] = P_{t-s}(X_s,\,\cdot\,).$$

Since $\mathcal{F}_s \subset \mathcal{G}_s$, it follows that for any $f \in B(S)$,

$$\mathbb{E}[f(X_t) \,|\, \mathcal{F}_s] = \mathbb{E}\big[\mathbb{E}[f(X_t) \,|\, \mathcal{G}_s] \,\big|\, \mathcal{F}_s\big] = \mathbb{E}[P_{t-s}f(X_s) \,|\, \mathcal{F}_s] = P_{t-s}f(X_s).$$

To finish the proof, we must show that $(P_t)_{t\geq 0}$ is a Markov semigroup with generator G given by (1.17). The fact that $\lim_{t\downarrow 0} P_t(x,y) = P_0(x,y) = \delta_x(y)$ follows from the fact that $\mathbb{P}[\Delta_{0,t} = \emptyset] \to 1$ as $t \downarrow 0$. To see that $P_sP_t = P_{s+t}$, let X^x be the process started in $X_0 = x$. By what we have already proved,

$$\begin{aligned} P_{s+t}f(x) &= \mathbb{E}[f(X^x_{s+t})] = \mathbb{E}\big[\mathbb{E}[f(X^x_{s+t}) \,|\, \mathcal{F}_s]\big] \\ &= \mathbb{E}[P_tf(X^x_s)] = P_sP_tf(X^x_0) = P_sP_tf(x). \end{aligned}$$

To see that the generator G of $(P_t)_{t\geq 0}$ is given by (1.17), we observe that

$$P_tf(x) = \mathbb{E}[f(X^x_t)] = f(x) + t\sum_{m\in\mathcal{M}} r_m\big(f(m(x)) - f(x)\big) + O(t^2) \qquad \text{as } t \downarrow 0,$$

which follows from the fact that $\mathbb{P}[|\Delta_{0,t}| \geq 2] = O(t^2)$ while

$$\mathbb{P}[\Delta_{0,t} = \{(m,s)\} \text{ for some } s \in (0,t)] = tr_m + O(t^2).$$

∎

1.6 Poisson construction of particle systems

Let S and Λ be a finite and countably infinite set, respectively. In what follows, we will mainly be interested in the case that $S = \{0,1\}$ and $\Lambda = \mathbb{Z}^d$, the d-dimensional integer lattice. We let S^Λ denote the space of all $x = (x(i))_{i\in\Lambda}$ with $x(i) \in S$ for all $i \in \Lambda$, i.e., S^Λ is the carthesian product of countably many copies of S, one for each point $i \in \Lambda$. Note that we can view an element $x \in S^\Lambda$ as a function that assigns to each lattice point $i \in \Lambda$ a value $x(i) \in S$. Recall that a sequence $x_n \in S^\Lambda$ converges to a limit x in the *product topology* on S^Λ if and only if $x_n \to x$ pointwise, i.e., $x_n(i) \to x(i)$ for all $i \in \Lambda$. Since S is finite and therefore compact, Tychonoff's theorem tells us that S^Λ, equipped with the product topology, is a compact metrizable space.

For any map $m : S^\Lambda \to S^\Lambda$ and $x, y \in \Lambda$, let

$$\mathcal{D}(m) := \big\{i \in \Lambda : \exists x \in S^\Lambda \text{ s.t. } m(x)(i) \neq x(i)\big\}$$

denote the set of lattice points whose values can possibly be changed by m. Let us say that a point $j \in \Lambda$ is *m-relevant* for some $i \in \Lambda$ if

$$\exists x, y \in S^\Lambda \text{ s.t. } m(x)(i) \neq m(y)(i) \text{ and } x(k) = y(k)\ \forall k \neq j,$$

i.e., changing the value of x in j may change the value of $m(x)$ in i. We say that a map $m : S^\Lambda \to S^\Lambda$ is *local* if both $\mathcal{D}(m)$ and the sets $(\mathcal{R}_i(m))_{i\in\mathcal{D}(m)}$ defined by

$$\mathcal{R}_i(m) := \big\{j \in \Lambda : j \text{ is } m\text{-relevant for } i\big\}$$

are all finite sets. Note that it is possible that $\mathcal{D}(m)$ is nonempty but $\mathcal{R}_i(m) = \emptyset$ for all $i \in \mathcal{D}(m)$.

Let $\mathcal{M}$ be a countable set whose elements are local maps $m : S^\Lambda \to S^\Lambda$, let $(r_m)_{m\in\mathcal{M}}$ be nonnegative constants, and let Δ be a Poisson point set on $\mathcal{M} \times \mathbb{R}$ with intensity $r_m \mathrm{d}t$. In analogy with Theorem 1.12, we wish to give a Poisson construction of the S^Λ-valued Markov process $(X_t)_{t\geq 0}$ with formal generator

$$Gf(x) := \sum_{m\in\mathcal{M}} r_m\big(f(m(x)) - f(x)\big). \tag{1.18}$$

The difficulty is that we will typically have that $\sum_{m\in\mathcal{M}} r_m = \infty$. As a result, $\{t : (t,m) \in \Delta\}$ will be a dense subset of $\mathbb{R}$, so it will no longer possible to order the elements of $\Delta_{0,t}$ according to their times. Nevertheless, since our maps m are local, we can hope that under suitable assumptions on the rates, only finitely many points of $\Delta_{0,t}$ are needed to determine the value $X_t(i)$ of our process at a given lattice point $i \in \Lambda$ and time $t \geq 0$.
To make this rigorous, we start by observing that for each $i \in \Lambda$, the set

$$\big\{t \in \mathbb{R} : \exists m \in \mathcal{M} \text{ s.t. } i \in \mathcal{D}(m),\ (m,t) \in \Delta\big\}$$

is a Poisson point set with intensity $\sum_{m\in\mathcal{M},\ \mathcal{D}(m)\ni i} r_m$. Therefore, provided that

$$K_0 := \sup_i \sum_{\substack{m\in\mathcal{M}\\ \mathcal{D}(m)\ni i}} r_m < \infty, \tag{1.19}$$

each finite time interval contains only finitely many events that have the potential to change the state of a given lattice point i. This does not automatically imply, however, that our process is well-defined, since events that happen at i might depend on events that happen at other sites at earlier times, and in this way a large and possibly infinite number of events and lattice points can potentially influence the state of a single lattice point at a given time.
With this in mind, we make the following definitions. By definition, by a *path* in Λ we will mean a pair of functions (γ_{t-}, γ_t) defined on some time interval $[s,u]$ with $s \leq u$ and taking values in Λ, such that

$$\begin{array}{ll} \lim_{t\downarrow t_0} \gamma_{t-} = \gamma_{t_0} & \big(t_0 \in [s,u)\big), \\ \lim_{t\uparrow t_0} \gamma_t = \gamma_{t_0-} & \big(t_0 \in (s,u]\big). \end{array} \tag{1.20}$$

Note that this definition allows for the case that $\gamma_{s-} \neq \gamma_s$; in this case, and only in this case, knowing only the function $t \mapsto \gamma_t$ is not enough to deduce the function $t \mapsto \gamma_{t-}$. We may identify a path, as we have just defined it, with the set $\gamma \subset \Lambda \times [0,\infty)$ defined by

$$\gamma := \big\{(\gamma_{t-}, t) : t \in [s,u]\big\} \cup \big\{(\gamma_t, t) : t \in [s,u]\big\}.$$

Note that both the functions γ_t and γ_{t-}, as well as the *starting time* s and *final time* u can be read off from the set γ.

For any $i, j \in \Lambda$ and $0 \leq s \leq u$, let us say that a path γ with starting time s and final time u is a *path of influence* from (i, s) to (j, u) if $\gamma_{s-} = i$, $\gamma_u = j$, and

$$\begin{array}{ll} \text{(i)} & \text{if } \gamma_{t-} \neq \gamma_t \text{ for some } t \in [s,u], \text{ then there exists some } m \in \mathcal{M} \\ & \text{such that } (m,t) \in \Delta,\ \gamma_t \in \mathcal{D}(m) \text{ and } \gamma_{t-} \in \mathcal{R}_{\gamma_t}(m), \\ \text{(ii)} & \text{for each } (m,t) \in \Delta \text{ with } t \in [s,u] \text{ and } \gamma_t \in \mathcal{D}(m), \\ & \text{one has } \gamma_{t-} \in \mathcal{R}_{\gamma_t}(m). \end{array} \tag{1.21}$$

For any finite set $A \subset \Lambda$ and $0 \leq s \leq u$, we set

$$\zeta_s^{A,u} := \big\{ i \in \Lambda : (i,s) \rightsquigarrow A \times \{u\} \big\}, \tag{1.22}$$

where $(i,s) \rightsquigarrow A \times \{u\}$ denotes the presence of a path of influence from (i,s) to some $(j,u) \in A \times \{u\}$. Note that $\zeta_t^{A,t}$ is the set of lattice points whose values at time zero are relevant for the state of the process in A at time t. The following lemma will be the cornerstone of our Poisson construction of interacting particle systems.

Lemma 1.13 (Exponential bound) *Assume that the rates $(r_m)_{m\in\mathcal{M}}$ satisfy (1.19) and that*

$$K := \sup_{i\in\Lambda} \sum_{\substack{m\in\mathcal{M} \\ \mathcal{D}(m)\ni i}} r_m\big(|\mathcal{R}_i(m)| - 1\big) < \infty. \tag{1.23}$$

Then, for each finite $A \subset \Lambda$, one has

$$\mathbb{E}\big[|\zeta_s^{A,u}|\big] \leq |A| e^{K(u-s)} \qquad (0 \leq s \leq u). \tag{1.24}$$

Proof To simplify notation, we fix A and u and write $\zeta_s := \zeta_s^{A,u}$. Let $\Lambda_n \subset \Lambda$ be finite sets such that $\Lambda_n \uparrow \Lambda$. For n large enough such that $A \subset \Lambda_n$, let us write

$$\zeta_s^n := \big\{ i \in \Lambda : (i,s) \rightsquigarrow_n A \times \{u\} \big\},$$

where $(i,s) \rightsquigarrow_n A \times \{u\}$ denotes the presence of a path of influence from (i,s) to $A \times \{u\}$ that stays in Λ_n. We observe that since $\Lambda_n \uparrow \Lambda$, we have

$$\zeta_s^n \uparrow \zeta_s \qquad (0 \leq s \leq u).$$

Let $\mathcal{M}_n := \{m \in \mathcal{M} : \mathcal{D}(m) \cap \Lambda_n \neq \emptyset\}$. For any $A \subset \Lambda_n$ and $m \in \mathcal{M}_n$, set

$$A^m := (A \backslash \mathcal{D}(m)) \ \cup \bigcup_{i\in A\cap\mathcal{D}(m)} \mathcal{R}_i(m).$$

It follows from (1.19) that $\sum_{m\in\mathcal{M}_n} r_m < \infty$, hence Theorem 1.12 implies that the process $(\zeta^n_{u-t})_{0\leq t\leq u}$ is a Markov process taking values in the (finite) space of all subsets of Λ_n, with generator

$$G_n f(A) := \sum_{m\in\mathcal{M}_n} r_m\big(f(A^m) - f(A)\big).$$

Let $(P^n_t)_{t\geq 0}$ be the associated semigroup and let f denote the function $f(A) := |A|$. Then

$$\begin{aligned}
G_n f(A) &= \sum_{m\in\mathcal{M}_n} r_m\big(f(A^m) - f(A)\big)\\
&\leq \sum_{m\in\mathcal{M}_n} r_m\Big(|A\backslash\mathcal{D}(m)| + \sum_{i\in A\cap\mathcal{D}(m)} |\mathcal{R}_i(m)| - |A|\Big)\\
&= \sum_{m\in\mathcal{M}_n} r_m\Big(\sum_{i\in A\cap\mathcal{D}(m)} \big(|\mathcal{R}_i(m)| - 1\big)\Big)\\
&= \sum_{i\in A} \sum_{\substack{m\in\mathcal{M}_n\\ \mathcal{D}(m)\ni i}} r_m\big(|\mathcal{R}_i(m)| - 1\big) \leq K|A|.
\end{aligned}$$

It follows that

$$\tfrac{\partial}{\partial t}\big(e^{-Kt}P^n_t f\big) = -Ke^{-Kt}P^n_t f + e^{-Kt}P^n_t Gf = e^{-Kt}P^n_t(Gf - Kf) \leq 0$$

and therefore $e^{-Kt}P^n_t f \leq e^{-K0}P^n_0 f = f$, which means that

$$\mathbb{E}\big[|\zeta^n_{u-t}|\big] \leq |A|e^{Kt} \qquad (0\leq t\leq u). \tag{1.25}$$

Letting $n\uparrow\infty$ we arrive at (2.10). ∎

Recall that $\Delta_{s,t} := \Delta\cap(\mathcal{M}\times(s,t])$. The next lemma shows that under suitable summability conditions on the rates, only finitely many Poisson events are relevant to determine the value of an interacting particle system at a given point in space and time.

Lemma 1.14 (Finitely many relevant events) *Assume that the rates $(r_m)_{m\in\mathcal{M}}$ satisfy (1.19) and that*

$$K_1 := \sup_{i\in\Lambda} \sum_{\substack{m\in\mathcal{M}\\ \mathcal{D}(m)\ni i}} r_m|\mathcal{R}_i(m)| < \infty. \tag{1.26}$$

Then, almost surely, for each $s\leq u$ and $i\in\Lambda$, the set

$$\big\{(m,t)\in\Delta_{s,u} : \mathcal{D}(m)\times\{t\} \rightsquigarrow (i,u)\big\}$$

is finite.

Proof Set

$$\begin{aligned}\xi_s^{A,u} &:= \bigcup_{t\in[s,u]} \zeta_t^{A,u}\\ &= \big\{i\in\Lambda : (i,s) \rightsquigarrow' A\times\{u\}\big\},\end{aligned}$$

where $\rightsquigarrow'$ is defined in a similar ways as $\rightsquigarrow$, except that we drop condition (ii) from the definition of a path of influence in (1.21). Lemma 1.13 does not automatically imply that $|\xi_s^{A,u}| < \infty$ for all $s \leq u$. However, applying the same method of proof to the Markov process $(\xi_{u-t}^{A,u})_{t\geq 0}$, replacing (1.23) by the slightly stronger condition (1.26), we can derive an exponential bound for $\mathbb{E}[|\xi_s^{A,u}|]$, proving that $\xi_s^{A,u}$ is a.s. finite for each finite A and $s \leq u$. Since by (1.19), there are only finitely many events $(m,t) \in \Delta_{s,u}$ such that $\mathcal{D}(m) \cap \xi_s^{A,u} \neq \emptyset$, our claim follows. ∎

Remark Conditions (1.19) and (1.26) can be combined in the condition

$$\sup_{i\in\Lambda} \sum_{\substack{m\in\mathcal{M}\\ \mathcal{D}(m)\ni i}} r_m\big(|\mathcal{R}_i(m)|+1\big) < \infty. \tag{1.27}$$

In view of Lemma 1.14, for any $0 \leq s \leq u$, we define maps $\Psi_{\Delta,s,u} : S^\Lambda \to S^\Lambda$ by

$$\begin{aligned}&\Psi_{\Delta,s,u}(x)(i) := m_n \circ \cdots \circ m_1(x)(i) \qquad (i\in\Lambda)\\ &\quad\text{where } \{(m_1,t_1),\ldots,(m_n,t_n)\} = \big\{(m,t)\in\Delta_{s,u} : \mathcal{D}(m)\times\{t\} \rightsquigarrow (i,u)\big\},\\ &\qquad t_1 < \cdots < t_n.\end{aligned}$$

We define probability kernels $P_t(x,\mathrm{d}y)$ on S^Λ by

$$P_t(x,\,\cdot\,) := \mathbb{P}\big[\Psi_{\Delta,0,t}(x) \in \cdot\,\big] \qquad (x\in S^\Lambda,\ t\geq 0). \tag{1.28}$$

Below is the main result of this chapter.

Theorem 1.15 (Poisson construction of particle systems) *Let $\mathcal{M}$ be a countable set whose elements are local maps $m : S^\Lambda \to S^\Lambda$, let $(r_m)_{m\in\mathcal{M}}$ be nonnegative constants satisfying (1.27), and let Δ be a Poisson point set on $\mathcal{M}\times[0,\infty)$ with intensity $r_m\mathrm{d}t$. Then (1.28) defines a Feller semigroup $(P_t)_{t\geq 0}$ on S^Λ. Moreover, if X_0 is an S^Λ-valued random variable, independent of Δ, then*

$$X_t := \Psi_{\Delta,0,t}(X_0) \qquad (t\geq 0) \tag{1.29}$$

defines a Feller process with semigroup $(P_t)_{t\geq 0}$.

Proof We start by observing that the process $(X_t)_{t\geq 0}$ defined in (1.29) has cadlag sample paths. Indeed, since we equip S^Λ with the product topology, this is equivalent to the statement that $t \mapsto X_t(i)$ is cadlag for each $i \in \Lambda$. But this follows directly from the way we have defined $\Psi_{\Delta,0,t}$ and the fact that the set of events that have the potential to change the state of a given lattice point i is a locally finite subset of $[0,\infty)$.
As in the proof of Theorem 1.12, let $\mathcal{F}_t$ and $\mathcal{G}_t$ be the σ-fields generated by $(X_s)_{0\leq s\leq t}$ and $X_0, \Delta_{0,t}$, respectively. Then, exactly in the same way as in the proof of Theorem 1.12, we see that

$$\mathbb{E}[f(X_t)\,|\,\mathcal{F}_s] = P_{t-s}f(X_s) \qquad \big(0 \leq s \leq t,\ f \in B(S^\Lambda)\big).$$

Also, the proof that $P_sP_t = P_{s+t}$ carries over without a change. Thus, to see that $(P_t)_{t\geq 0}$ is a Feller semigroup, it suffices to show that $(x,t) \mapsto P_t(x,\,\cdot\,)$ is a continuous map from $S^\Lambda \times [0,\infty)$ to $\mathcal{M}_1(S^\Lambda)$. In order to do this, it is convenient to ue negative times (Note that we have defined Δ to be a Poisson point process on $\mathcal{M} \times \mathbb{R}$, but so far we have only used points $(m,t) \in \Delta$ with $t > 0$.) Since the law of Δ is invariant under translations of time, we have (compare (1.28))

$$P_t(x,\,\cdot\,) := \mathbb{P}\big[\Psi_{\Delta,-t,0}(x) \in \cdot\,\big] \qquad (x \in S^\Lambda,\ t \geq 0).$$

Therefore, in order to prove that $P_{t_n}(x_n,\,\cdot\,)$ converges weakly to $P_t(x,\,\cdot\,)$ as we let $(x_n,t_n) \to (x,t)$, it suffices to prove that

$$\Psi_{\Delta,-t_n,0}(x_n) \underset{n\to\infty}{\longrightarrow} \Psi_{\Delta,-t,0}(x) \quad \text{a.s.}$$

as $(x_n,t_n) \to (x,t)$. Since we equip S^Λ with the product topology, we need to show that

$$\Psi_{\Delta,-t_n,0}(x_n)(i) \underset{n\to\infty}{\longrightarrow} \Psi_{\Delta,-t,0}(x)(i) \quad \text{a.s.}$$

for each $i \in \Lambda$. By Lemma 1.14, there exists some $\varepsilon > 0$ such that there are no points in $\Delta_{-t-\varepsilon,-t+\varepsilon}$ that are relevant for $(i,0)$, while by Lemma 1.13, $\zeta^{\{i\},0}_{-t}$ is a finite set. Therefore, for all n large enough such that $-t_n \in (-t-\varepsilon, -t+\varepsilon)$ and $x_n = x$ on $\zeta^{\{i\},0}_{-t}$, one has $\Psi_{\Delta,-t_n,0}(x_n)(i) = \Psi_{\Delta,-t,0}(x)(i)$, proving the desired a.s. convergence. ∎

1.7 Generator construction of particle systems

Although Theorem 1.15 gives us an explicit way how to construct the Feller semigroup associated with an interacting particle system, it does not tell us very much about its generator. To fill this gap, we need a bit more theory. For any continuous function $f : S^\Lambda \to \mathbb{R}$ and $i \in \Lambda$, we define

$$\delta f(i) := \sup\big\{|f(x) - f(y)| : x, y \in S^\Lambda,\ x(j) = y(j)\ \forall j \neq i\big\}.$$

Note that $\delta f(i)$ measures how much $f(x)$ can change if we change x only in the point i. We call δf the *variation* of f and we define a space of functions of 'summable variation' by

$$\begin{aligned}\mathcal{C}_{\rm sum} = \mathcal{C}_{\rm sum}(S^\Lambda) &:= \big\{f \in \mathcal{C}(S^\Lambda) : \sum_i \delta f(i) < \infty\big\},\\ \mathcal{C}_{\rm fin} = \mathcal{C}_{\rm fin}(S^\Lambda) &:= \big\{f \in \mathcal{C}(S^\Lambda) : \delta f(i) = 0 \text{ for all but finitely many } i\big\}.\end{aligned}$$

Exercise 1.16 Let us say that a function $f : S^\Lambda \to \mathbb{R}$ *depends on finitely many coordinates* if there exists a finite set $A \subset \Lambda$ and a function $f' : S^A \to \mathbb{R}$ such that

$$f\big((x(i))_{i\in\Lambda}\big) = f'\big((x(i))_{i\in F}\big) \qquad \big(x \in S^\Lambda\big).$$

Show that each function that depends on finitely many coordinates is continuous, that

$$\mathcal{C}_{\rm fin}(S^\Lambda) = \big\{f \in \mathcal{C}(S^\Lambda) : f \text{ depends on finitely many coordinates}\big\},$$

and that $\mathcal{C}_{\rm fin}(S^\Lambda)$ is a dense linear subspace of the Banach space $\mathcal{C}(S^\Lambda)$ of all continuous real functions on S^Λ, equipped with the supremumnorm.

Lemma 1.17 (Domain of pregenerator) *Assume that the rates $(r_m)_{m\in\mathcal{M}}$ satisfy (1.19). Then, for each $f \in \mathcal{C}_{\rm sum}(S^\Lambda)$,*

$$\sum_{m\in\mathcal{M}} r_m\big|f(m(x)) - f(x)\big| \leq K_0 \sum_{i\in\Lambda} \delta f(i),$$

where K_0 is the constant from (1.19). In particular, for each $f \in \mathcal{C}_{\rm sum}(S^\Lambda)$, the right-hand side of (1.18) is absolutely summable and Gf is well-defined.

Proof This follows by writing

$$\begin{aligned}\sum_{m\in\mathcal{M}} r_m\big|f(m(x)) - f(x)\big| &\leq \sum_{m\in\mathcal{M}} r_m \sum_{i\in\mathcal{D}(m)} \delta f(i)\\ &= \sum_{i\in\Lambda} \delta f(i) \sum_{\substack{m\in\mathcal{M}\\ \mathcal{D}(m)\ni i}} r_m \leq K_0 \sum_{i\in\Lambda} \delta f(i).\end{aligned}$$

∎

The following theorem is the main result of the present section.

Theorem 1.18 (Generator construction of particle systems) *Assume that the rates $(r_m)_{m\in\mathcal{M}}$ satisfy (1.27), let $(P_t)_{t\geq 0}$ be the Feller semigroup defined in (1.28) and let G be the linear operator with domain $\mathcal{D}(G) := \mathcal{C}_{\rm sum}$ defined by (1.18). Then G is are closeable and its closure $\overline{G}$ is the generator of $(P_t)_{t\geq 0}$. Moreover, if $G|_{\mathcal{C}_{\rm fin}}$ denotes the restriction of G to the smaller domain $\mathcal{D}(G|_{\mathcal{C}_{\rm fin}}) := \mathcal{C}_{\rm fin}$, then $G|_{\mathcal{C}_{\rm fin}}$ is also closeable and $\overline{G|_{\mathcal{C}_{\rm fin}}} = \overline{G}$.*

Remark Since $\mathcal{D}(G|_{\mathcal{C}_{\rm fin}}) \subset \mathcal{D}(G)$ and G is closeable, it is easy to see that $G|_{\mathcal{C}_{\rm fin}}$ is also closeable, $\mathcal{D}(\overline{G|_{\mathcal{C}_{\rm fin}}}) \subset \mathcal{D}(\overline{G})$, and $\overline{G|_{\mathcal{C}_{\rm fin}}}f = \overline{G}f$ for all $f \in \mathcal{D}(\overline{G|_{\mathcal{C}_{\rm fin}}})$. It is not immediately obvious, however, that $\mathcal{D}(\overline{G|_{\mathcal{C}_{\rm fin}}}) = \mathcal{D}(\overline{G})$. In general, if A is a closed linear operator and $\mathcal{D}' \subset \mathcal{D}(A)$, then we say that $\mathcal{D}'$ is a *core* for A if $\overline{A|_{\mathcal{D}'}} = A$. Then Theorem 1.18 says that $\mathcal{C}_{\rm fin}$ is a core for $\overline{G}$.

To prepare for the proof of Theorem 1.18 we need a few lemmas.

Lemma 1.19 (Generator on local functions) *Under the asumptions of Theorem 1.18, one has $\lim_{t\downarrow 0} t^{-1}(P_t f - f) = Gf$ for all $f \in \mathcal{C}_{\rm fin}$, where the limit exists in the topology on $\mathcal{C}(S^\Lambda)$.*

Proof Since $f \in \mathcal{C}_{\rm fin}$, there exists some finite $A \subset \Lambda$ such that f depends only on the coordinates in A. Let $\Delta_{0,t}(A) := \{(m,s) \in \Delta_{0,t} : \mathcal{D}(m) \cap A \neq \emptyset\}$. Then $|\Delta_{0,t}(A)|$ is Poisson distributed with mean $t\sum_{m\in\mathcal{M},\ \mathcal{D}(m)\cap A\neq\emptyset}$, which is finite by (1.19). Write

$$
\begin{aligned}
P_t f(x) = {} & f(x)\mathbb{P}\big[\Delta_{0,t}(A) = \emptyset\big] \\
& + \sum_{\substack{m\in\mathcal{M} \\ \mathcal{D}(m)\cap A\neq\emptyset}} f\big(m(x)\big)\mathbb{P}\big[\Delta_{0,t}(A) = \{(m,s)\} \text{ for some } 0 < s \leq t\big] \\
& + \mathbb{E}\big[f\big(\Psi_{\Delta,0,t}(x)\big)\,\big|\,|\Delta_{0,t}(A)| \geq 2\big]\,\mathbb{P}\big[|\Delta_{0,t}(A)| \geq 2\big].
\end{aligned}
$$

Since

$$\big|\mathbb{E}[f\big(\Psi_{\Delta,0,t}(x)\big)\,\big|\,|\Delta_{0,t}(A)| \geq 2]\big| \leq \|f\|,$$

and since $m(x) = x$ for all $m \in \mathcal{M}$ with $\mathcal{D}(m) \cap A = \emptyset$, we can write

$$P_t f(x) = f(x) + t\sum_{m\in\mathcal{M}} r_m\big(f(m(x)) - f(x)\big) + R_t(x),$$

where $\lim_{t\downarrow 0} t^{-1}\|R_t\| = 0$. ∎

Lemma 1.20 (Approximation by local functions) *Assume that the rates* $(r_m)_{m\in\mathcal{M}}$ *satisfy (1.19). Then for all* $f \in \mathcal{C}_{\rm sum}$ *there exist* $f_n \in \mathcal{C}_{\rm fin}$ *such that* $\|f_n - f\| \to 0$ *and* $\|Gf_n - Gf\| \to 0$.

Proof Choose finite $\Lambda_n \uparrow \Lambda$, set $\Gamma_n := \Lambda\backslash\Lambda_n$, fix $z \in S^\Lambda$, and for each $x \in S^\Lambda$ define $x_n \to x$ by

$$x_n(i) := \begin{cases} x(i) & \text{if } i \in \Lambda_n, \\ z(i) & \text{if } i \in \Gamma_n. \end{cases}$$

Fix $f \in \mathcal{C}_{\rm sum}$ and define $f_n(x) := f(x_n)$ $(x \in S^\Lambda)$. Then f_n depends only on the coordinates in Λ_n, hence $f_n \in \mathcal{C}_{\rm fin}$. We claim that for any $x \in S^\Lambda$,

$$|f(x_n) - f(x)| \leq \sum_{i\in\Gamma_n} \delta f(i) \qquad (x \in S^\Lambda,\ n \geq 1)$$

To see this, let $\Gamma_n := \{i_1, i_2, \ldots\}$ and define $(x_n^k)^{k=0,1,2,\ldots}$ with $x_n^0 = x_n$ and $x_n^k \to x$ as $k \to \infty$ by

$$x_n^k(i) := \begin{cases} x(i) & \text{if } i \in \Lambda_n \cup \{i_1, \ldots, i_k\}, \\ z(i) & \text{if } i \in \Gamma_n\backslash\{i_1, \ldots, i_k\}. \end{cases}$$

Then

$$|f(x_n) - f(x_n^k)| \leq \sum_{l=1}^{k} |f(x_n^{l-1}) - f(x_n^l)| \leq \sum_{l=1}^{k} \delta f(i_l),$$

from which our claim follows by letting $k \to \infty$, using the continuity of f. Since $f \in \mathcal{C}_{\rm sum}$, it follows that

$$\|f_n - f\| \leq \sum_{i\in\Gamma_n} \delta f(i) \underset{n\to\infty}{\longrightarrow} 0.$$

Moreover, we observe that

$$\begin{aligned} |Gf_n(x) - Gf(x)| &= \Big| \sum_{m\in\mathcal{M}} r_m\big(f_n(m(x)) - f_n(x)\big) - \sum_{m\in\mathcal{M}} r_m\big(f(m(x)) - f(x)\big)\Big| \\ &\leq \sum_{m\in\mathcal{M}} r_m \big|f(m(x)_n) - f(x_n) - f(m(x)) + f(x)\big|. \end{aligned} \tag{1.30}$$

On the one hand, we have

$$\begin{aligned} &\big|f(m(x)_n) - f(x_n) - f(m(x)) + f(x)\big| \\ &\quad\leq \big|f(m(x)_n) - f(x_n)\big| + \big|f(m(x)) - f(x)\big| \leq 2 \sum_{i\in\mathcal{D}(m)} \delta f(i), \end{aligned}$$

while on the other hand, we can estimate the same quantity as

$$\leq \big|f(m(x)_n) - f(m(x))\big| + \big|f(x_n) - f(x)\big| \leq 2 \sum_{i\in\Gamma_n)} \delta f(i).$$

Let $\Delta \subset \Lambda$ be finite. Inserting either of our two estimates into (1.30), depending on whether $\mathcal{D}(m) \cap \Delta \neq \emptyset$ or not, we find that

$$\begin{aligned}\|Gf_n - Gf\| &\leq 2 \sum_{\substack{m\in\mathcal{M}\\ \mathcal{D}(m)\cap\Delta\neq\emptyset}} r_m \sum_{i\in\Gamma_n)} \delta f(i) + 2 \sum_{\substack{m\in\mathcal{M}\\ \mathcal{D}(m)\cap\Delta=\emptyset}} r_m \sum_{i\in\mathcal{D}(m)} \delta f(i)\\ &\leq 2K_0|\Delta| \sum_{i\in\Gamma_n)} \delta f(i) + 2\sum_{i\in\Lambda} \delta f(i) \sum_{\substack{m\in\mathcal{M}\\ \mathcal{D}(m)\cap\Delta=\emptyset\\ \mathcal{D}(m)\ni i}} r_m.\end{aligned}$$

It follows that

$$\limsup_{n\to\infty} \|Gf_n - Gf\| \leq 2 \sum_{i\in\Lambda\backslash\Delta} \delta f(i) \sum_{\substack{m\in\mathcal{M}\\ \mathcal{D}(m)\ni i}} r_m \leq 2K_0 \sum_{i\in\Lambda\backslash\Delta} \delta f(i).$$

Since Δ is arbitrary, letting $\Delta \uparrow \Lambda$, we see that $\limsup_n \|Gf_n - Gf\| = 0$. ∎

Lemma 1.21 (Functions of summable variation) *Under the asumptions of Theorem 1.18, one has*

$$\sum_{i\in\Lambda} \delta P_t f(i) \leq e^{Kt} \sum_{i\in\Lambda} \delta f(i) \qquad \big(t \geq 0,\ f \in \mathcal{C}_{\rm sum}(S^\Lambda)\big),$$

where K is the constant from (1.23). In particular, for each $t \geq 0$, P_t maps $\mathcal{C}_{\rm sum}(S^\Lambda)$ into itself.

Proof For each $i \in \Lambda$ and $x, y \in S^\Lambda$ such that $x(j) = y(j)$ for all $j \neq i$, we have

$$\begin{aligned}&|P_t f(x) - P_t f(y)| = \big|\mathbb{E}[f(\Psi_{\Delta,0,t}(x))] - \mathbb{E}[f(\Psi_{\Delta,0,t}(y))]\big|\\ &\quad\leq \mathbb{E}\big[|f(\Psi_{\Delta,0,t}(x)) - f(\Psi_{\Delta,0,t}(y))|\big]\\ &\quad\leq \mathbb{E}\big[\textstyle\sum_{j:\,\Psi_{\Delta,0,t}(x)(j)\neq\Psi_{\Delta,0,t}(y)(j)} \delta f(j)\big]\\ &\quad= \sum_j \mathbb{P}\big[\Psi_{\Delta,0,t}(x)(j) \neq \Psi_{\Delta,0,t}(y)(j)\big]\delta f(j)\\ &\quad\leq \sum_j \mathbb{P}\big[(i,0) \rightsquigarrow (j,t)\big]\delta f(j).\end{aligned}$$

By Lemma 1.13, it follows that

$$\sum_i \delta P_t f(i) \leq \sum_{ij} \mathbb{P}\big[(i,0) \rightsquigarrow (j,t)\big] \delta f(j) = \sum_j \mathbb{E}\big[|\zeta_0^{\{j\},t}|\big] \delta f(j) \leq e^{Kt} \sum_j \delta f(j).$$

∎

Proof of Theorem 1.18 Let H be the full generator of $(P_t)_{t\geq 0}$ and let $\mathcal{D}(H)$ denote it domain. Then Lemma 1.19 shows that $\mathcal{C}_{\rm fin} \subset \mathcal{D}(H)$ and $Gf = Hf$ for all $f \in \mathcal{C}_{\rm fin}$. By Lemma 1.20, it follows that $\mathcal{C}_{\rm sum} \subset \mathcal{D}(H)$ and $Gf = Hf$ for all $f \in \mathcal{C}_{\rm sum}$.
To see that G is closeable and its closure is the generator of a Feller semigroup, we check conditions (i)–(iv) of Theorem 1.4. It is easy to see that $1 \in \mathcal{C}_{\rm sum}(S^\Lambda)$ and $G1 = 0$. If f assumes its maximum in a point $x \in S^\Lambda$, then each term on the right-hand side of (1.18) is nonpositive, hence $Gf(x) \leq 0$. The fact that $\mathcal{C}_{\rm sum}(S^\Lambda)$ is dense follows from Exercise 1.16 and the fact that $\mathcal{C}_{\rm fin}(S^\Lambda) \subset \mathcal{C}_{\rm sum}(S^\Lambda)$. To check condition (iv), we will show that for each $r > K$, where K is the constant from (1.23), and for each $f \in \mathcal{C}_{\rm fin}(S^\Lambda)$, there exists a $p_r \in \mathcal{C}_{\rm sum}(S^\Lambda)$ such that $(r - G)p_r = f$. Indeed, we will show that such a function is given by

$$p_r := \int_0^\infty e^{-rt} P_t f \,\mathrm{d}t.$$

Indeed, it follows from Theorem 1.4 that $p_r \in \mathcal{D}(H)$ and $(r - H)p_r = f$. Thus, it suffices to show that $p_r \in \mathcal{C}_{\rm sum}$. To see this, note that if $x(j) = y(j)$ for all $j \neq i$, then

$$\begin{aligned} |p_r(x) - p_r(y)| &= \Big| \int_0^\infty e^{-rt} P_t f(x) \,\mathrm{d}t - \int_0^\infty e^{-rt} P_t f(y) \mathrm{d}t \Big| \\ &\leq \int_0^\infty e^{-rt} \big| P_t f(x) - P_t f(y) \big| \,\mathrm{d}t \leq \int_0^\infty e^{-rt} \delta P_t f(i) \,\mathrm{d}t, \end{aligned}$$

and therefore, by Lemma 1.21,

$$\sum_i \delta p(i) \leq \int_0^\infty e^{-rt} \sum_i \delta P_t f(i) \,\mathrm{d}t \leq \big(\sum_i \delta f(i)\big) \int_0^\infty e^{-rt} e^{Kt} \,\mathrm{d}t < \infty,$$

which proves that $p_r \in \mathcal{C}_{\rm sum}$. This completes the proof that $\overline{G} = H$. By Lemma 1.20, we see that $\mathcal{D}(\overline{G|_{\mathcal{C}_{\rm fin}}}) \supset \mathcal{C}_{\rm sum}$ and therefore also $\overline{G|_{\mathcal{C}_{\rm fin}}} = H$. ∎

We conclude this section with the following lemma, which is sometimes useful.

Lemma 1.22 (Differentiation of semigroup) *Assume that the rates $(r_m)_{m\in\mathcal{M}}$ satisfy (1.27), let $(P_t)_{t\geq 0}$ be the Feller semigroup defined in (1.28) and let G be the linear operator with domain $\mathcal{D}(G) := \mathcal{C}_{\rm sum}(S^\Lambda)$ defined by (1.18). Then, for each $f \in \mathcal{C}_{\rm sum}(S^\Lambda)$, $t \mapsto P_t f$ is a continuously differentiable function from $[0,\infty)$ to $\mathcal{C}(S^\Lambda)$ satisfying $P_0 f = f$, $P_t f \in \mathcal{C}_{\rm sum}(S^\Lambda)$, and $\frac{\partial}{\partial t} P_t f = G P_t f$ for each $t \geq 0$.*

Proof This is a direct consequence of Proposition 1.3, Lemma 1.21, and Theorem 1.18. A direct proof based on our definition of $(P_t)_{t\geq 0}$ (not using Hille-Yosida theory) is also possible, but quite long and technical. ∎

Remarks Theorem 1.18 is similar to Liggett's [Lig85, Theorem I.3.9], but there are also some differences. Liggett does not construct his interacting particle systems using Poisson point sets, but rather gives a direct proof that the closure of G generates a Feller semigroup $(P_t)_{t\geq 0}$, and then uses general theory to conclude that there exists a Feller process asociated with $(P_t)_{t\geq 0}$. Also, Liggett allows for the case that S is a (not necessarily finite) compact metrizable space and he does not write his generator in terms of local maps but in terms of 'local transition measures' which satisfy conditions similar to (1.27).

Chapter 2

The contact process

In this chapter, we study the contact process. The contact process is one of the most basic and most intensively studied interacting particle systems. It was introduced in the mathematical literature by Harris in 1974 [Har74] and a few years later independently in the high-energy physics literature as the ‘reggeon spin model’. Many important questions about the behavior of the nearest-neighbour contact process on $\mathbb{Z}^d$ were solved by Bezuidenhout and Grimmett in 1990–1991 (building, of course, on the work of many others) [BG90, BG91]. Physicists, not hindered by the burden of rigorous proof, proceeded much faster. In fact, the only statements about the contact process that physicists consider nontrivial -concerning its critical behavior in low dimensions- remain largely unproved by mathematicians up to date. The contact process continues to be the subject of intense study in the mathematical literature. Questions about its critical behavior in high dimensions were recently answered in [HS05]. In addition, all kind of variations on the original process such as contact processes in a random environment [Lig92, Rem08] or contact processes on more general lattices (see [Lig99] as a general reference) have recieved a lot of attention.

2.1 Definition of the model

Recall that

$$\mathbb{Z}^d := \big\{ i = (i_1, \ldots, i_d) : i_k \in \mathbb{Z}\ \forall k = 1, \ldots, d \big\}$$

is the *d-dimensional integer lattice*. Points $i \in \mathbb{Z}^d$ are often called *sites*. The (standard, nearest-neighbor) *contact process* on $\mathbb{Z}^d$ with *infection rate* $\lambda \geq 0$ is the Feller process in $\{0,1\}^{\mathbb{Z}^d}$ whose generator is the closure of the operator G with domain $\mathcal{D}(G) := \mathcal{C}_{\rm sum}(\{0,1\}^{\mathbb{Z}^d})$ defined by

$$\begin{aligned} Gf(x) := \lambda \sum_i 1_{\{x(i)=0\}} \sum_{j:\,|i-j|=1} x(j)\big(f(x^{\{i\}}) - f(x)\big) \\ + \sum_i 1_{\{x(i)=1\}} \big(f(x^{\{i\}}) - f(x)\big), \end{aligned} \tag{2.1}$$

where for any $x \in \{0,1\}^{\mathbb{Z}^d}$ and $i \in \mathbb{Z}^d$, we define $x^{\{i\}} \in \{0,1\}^{\mathbb{Z}^d}$ by

$$x^{\{i\}}(j) := \begin{cases} 1 - x(j) & \text{if } i = j, \\ x(j) & \text{otherwise.} \end{cases} \tag{2.2}$$

Note that (2.1) says that if at some time t the state of the process is $x = (x(i))_{i\in\mathbb{Z}^d} \in \{0,1\}^{\mathbb{Z}^d}$, then

$$\begin{array}{ll} x(i) \text{ jumps:} & \\ 0 \mapsto 1 & \text{with rate } \lambda \sum_{j:\,|i-j|=1} x(j), \\ 1 \mapsto 0 & \text{with rate } 1. \end{array}$$

Here, as in (2.1), the sum over j runs over all *nearest neighbours* of the site i, i.e., all $j \in \mathbb{Z}^d$ such that $|i-j| = 1$, where $|\cdot|$ denotes the usual euclidean distance.
A common way of interpreting a contact process $X = (X_t)_{t\geq 0}$ is to say that at each site $i \in \mathbb{Z}^d$ is situated an organism (for example, in $d = 1$ or $d = 2$ we can think of trees along an infinite road or in an infinite orchard) that can be in two states. If $X_t(i) = 0$, then we say that at time t the organism at site i is *healthy*, while if $X_t(i) = 1$, we say that it is *infected* (with some disease, or bug). Then the generator in (2.1) says that healthy organisms become infected with a rate that is lambda times their number of infected neighbours, and infected trees get healthy with constant *recovery rate* 1.
We wish to construct the process based on Poisson point sets as in Theorem 1.15. As a first step, we must write G in terms of local maps. Let

$$\mathcal{E} := \big\{(i,j) : i,j \in \mathbb{Z}^d,\ |i-j| = 1\big\}$$

denote the set of all ordered nearest-neighbour pairs, and for each $(i,j) \in \mathcal{E}$, let us define a map $m_{ij} : \{0,1\}^{\mathbb{Z}^d} \to \{0,1\}^{\mathbb{Z}^d}$ by

$$(m_{ij}x)(k) := \begin{cases} 1 & \text{if } k = j,\ x(i) = 1, \\ x(k) & \text{otherwise,} \end{cases} \qquad \big(i,j,k \in \mathbb{Z}^d,\ x \in \{0,1\}^{\mathbb{Z}^d}\big).$$

Note that m_{ij} describes a potential infection from i to j, i.e., if the site i is infected in the configuration x, then the site j will be infected in the configuration $m_{ij}(x)$, regardless of whether it was infected in x or not. Likewise, for each $i \in \mathbb{Z}^d$, let us define $p_i : \{0,1\}^{\mathbb{Z}^d} \to \{0,1\}^{\mathbb{Z}^d}$ by

$$(p_i x)(j) := \begin{cases} 0 & \text{if } j = i, \\ x(j) & \text{if } j \neq i, \end{cases} \qquad \big(i, j \in \mathbb{Z}^d,\ x \in \{0,1\}^{\mathbb{Z}^d}\big).$$

Then p_i describes a potential recovery of the site i, i.e., in the configuration $p_i(x)$, the site i is healthy, regardless of its state in the configuration x.

In terms of these local maps, we may rewrite the generator G in (2.1) in the form

$$Gf(x) = \lambda \sum_{(i,j)\in\mathcal{E}} \big(f(m_{ij}(x)) - f(x)\big) + \sum_{i\in\mathbb{Z}^d} \big(f(p_i(x)) - f(x)\big). \tag{2.3}$$

$(x \in \{0,1\}^{\mathbb{Z}^d})$. We can now apply Theorem 1.15 to give a Poisson construction of our contact process. To make this concrete, let Δ^{i} and Δ^{r} be independent Poisson point sets on $\mathcal{E} \times \mathbb{R}$ and $\mathbb{Z}^d \times \mathbb{R}$ with intensities $\lambda\,\mathrm{d}t$ and $1\,\mathrm{d}t$, respectively. We interpret a point $(i, j, t) \in \Delta^{\mathrm{i}}$ as a potential infection from i to j and a point $(i, t) \in \Delta^{\mathrm{r}}$ as a potential recovery of the site i.

In Figure 2.1 we have drawn a finite piece of the sets Δ^{i} and Δ^{r} for the process on $\mathbb{Z}$. We have drawn space horizontally and time vertically. Infections $(i, j, t) \in \Delta^{\mathrm{i}}$ have been indicated by drawing an arrow from (i, t) to (j, t) while recoveries $(i, t) \in \Delta^{\mathrm{r}}$ have been indicated with a black box.

Recall the definition of path γ from Section 1.6 and of a *path of influence* in (1.21). In our present set-up, we see that $\mathcal{D}(m_{ij}) = \{j\}$, $\mathcal{R}_j(m_{ij}) = \{i, j\}$, $\mathcal{D}(p_i) = \{i\}$ and $\mathcal{R}_i(p_i) = \emptyset$. Therefore, a path γ is a path of influence if and only if

$$\begin{aligned} &(\gamma_{t-}, \gamma_t, t) \in \Delta^{\mathrm{i}} \quad \text{for all } t \in [s, u] \text{ s.t. } \gamma_{t-} \neq \gamma_t, \\ &\gamma \cap \Delta^{\mathrm{r}} = \emptyset. \end{aligned}$$

In words, this says that an open path must walk upwards in time, may use infection arrows, but must avoid recovery symbols. In the context of the contact process, such paths are usually called *open*.

We write $(i, s) \rightsquigarrow (j, u)$, to denote the event that there exists an open path γ with starting time s and final time u such that $\gamma_{s-} = i$ and $\gamma_u = j$. Then, by Theorems 1.15 and 1.18, setting

$$X^x_t(i) := \begin{cases} 1 & \text{if } \exists j \text{ s.t. } x(j) = 1,\ (j, 0) \rightsquigarrow (i, t), \\ 0 & \text{otherwise}, \end{cases} \tag{2.4}$$

$(t \geq 0,\ i \in \mathbb{Z}^d,\ x \in \{0,1\}^{\mathbb{Z}^d})$ defines a Feller process in $\{0,1\}^{\mathbb{Z}^d}$ whose generator is the closure of the operator G in (2.1) (or, equivalently, (2.3)). This construction is known as the *graphical representation* of the contact process.

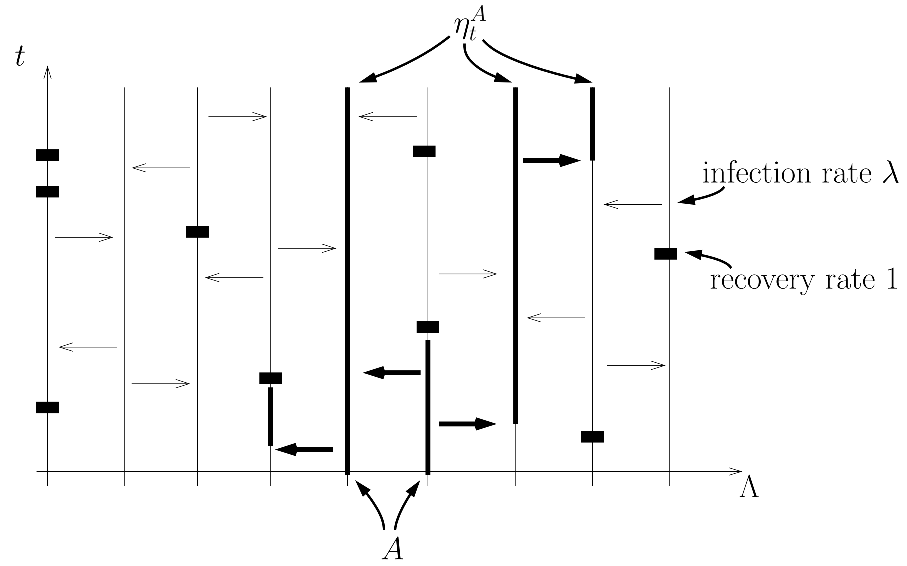

Figure 2.1: Graphical representation of the contact process.

In view of what follows, it will be useful to introduce some more notation. By identifying a set with its indicator function, we observe that the space $\{0,1\}^{\mathbb{Z}^d}$ is in a natural way isomorphic to the space $\mathcal{P}(\mathbb{Z}^d) := \{A : A \subset \mathbb{Z}^d\}$ of all subsets of $\mathbb{Z}^d$. With this in mind, for any $A \subset \mathbb{Z}^d$ and $s \in \mathbb{R}$, we define

$$\begin{aligned} \eta_t^{A,s} &:= \big\{i : A \times \{s\} \rightsquigarrow (i, s+t)\big\}, \\ \eta_t^{\dagger\, A,s} &:= \big\{i : (i, s-t) \rightsquigarrow A \times \{s\}\big\}, \end{aligned} \tag{2.5}$$

$(A \subset \mathbb{Z}^d,\ s \in \mathbb{R},\ t \geq 0)$, where $A \times \{0\} \rightsquigarrow (i,t)$ indicates the event that $(j,0) \rightsquigarrow (i,t)$ for some $j \in A$. We observe that if X^x is defined as in (2.4), then

$$x = 1_A \quad \text{implies} \quad X_t^x = 1_{\eta_t^{A,0}} \qquad (t \geq 0).$$

Moreover, we observe that for $A \subset \mathbb{Z}^d$ and $s \in \mathbb{R}$, the processes

$$\big(\eta_t^{A,0}\big)_{t\geq 0}, \quad \big(\eta_t^{A,s}\big)_{t\geq 0}, \quad \text{and} \quad \big(\eta_t^{\dagger\, A,s}\big)_{t\geq 0}$$

are all equal in law. The first equality (in law) follows from the fact that the law of our Poisson point processes is invariant under translations in time. To see that also $(\eta_t^{\dagger\, A,s})_{t\geq 0}$ is (in law) a contact process with initial state A, we turn the graphical representation in Figure 2.1 upside down and reverse the direction of all

arrows. The resulting picture is again Poisson, with the same rates as before, and what used to be an open path from (s, i) to (j, u) is now an open path from $(j, -u)$ to $(i, -s)$.
To simplify notation, we write

$$\eta_t^A := \eta_t^{A,0} \quad \text{and} \quad \eta_t^{\dagger\, A} := \eta_t^{\dagger\, A,0} \qquad (t \geq 0).$$

We observe that for any $t \geq 0$ and $A, B \subset \mathbb{Z}^d$,

$$\mathbb{P}\big[\eta_t^{A,0} \cap B \neq \emptyset\big] = \mathbb{P}\big[A \times \{0\} \rightsquigarrow B \times \{t\}\big] = \mathbb{P}\big[A \cap \eta_t^{\dagger\, B,t} \neq \emptyset\big].$$

This formula remains true if A and B are random sets, independent of the Poisson point processes of our graphical representation. This means that we have proved the following result.

Lemma 2.1 (Duality) *Let $(\eta_t)_{t\geq 0}$ and $(\eta_t^\dagger)_{t\geq 0}$ be independent contact processes on $\mathbb{Z}^d$ with infection rate λ. Then*

$$\mathbb{P}\big[\eta_t \cap \eta_0^\dagger \neq \emptyset\big] = \mathbb{P}\big[\eta_0 \cap \eta_t^\dagger \neq \emptyset\big] \qquad (t \geq 0).$$

This result is especially useful in view of the following fact.

Lemma 2.2 (Distribution determining functions) *Let μ, ν be probability laws on $\mathcal{P}(\mathbb{Z}^d)$ such that*

$$\int \mu(\mathrm{d}A) 1_{\{A\cap B\neq\emptyset\}} = \int \nu(\mathrm{d}A) 1_{\{A\cap B\neq\emptyset\}}$$

for all finite nonempty $B \subset \mathbb{Z}^d$. Then $\mu = \nu$.

Proof We start by recalling the *Stone-Weierstrass theorem.* Let E be a compact metrizable set. By definition, a subset $\mathcal{F}$ of $\mathcal{C}(E)$ is an *algebra* if $\mathcal{F}$ is a linear space, $\mathcal{F}$ contains the constant function 1, and $f, g \in \mathcal{F}$ implies $fg \in \mathcal{F}$. We say that $\mathcal{F}$ *separates points* if for every $x, y \in E$ with $x \neq y$ there exists an $f \in \mathcal{F}$ with $f(x) \neq f(y)$. The Stone-Weierstrass theorem says that if subset $\mathcal{F}$ of $\mathcal{C}(E)$ is an algebra that separates points, then $\mathcal{F}$ is dense in $\mathcal{C}(E)$.
Let $\mathcal{F}$ be the linear span of all functions of the form $A \mapsto 1_{\{A\cap B=\emptyset\}}$ with B a finite subset of $\mathbb{Z}^d$. Since $1_{\{A\cap\emptyset=\emptyset\}} = 1$ and $1_{\{A\cap B=\emptyset\}}1_{\{A\cap B'=\emptyset\}} = 1_{\{A\cap(B\cup B')=\emptyset\}}$ we see that $\mathcal{F}$ is an algebra. Since for all $A \neq A'$ there is a finite B such that $1_{\{A\cap B=\emptyset\}} \neq 1_{\{A'\cap B=\emptyset\}}$ we see that $\mathcal{F}$ separates points, hence by the Stone-Weierstrass theorem $\mathcal{F}$ is dense in $\mathcal{C}(\mathcal{P}(\mathbb{Z}^d))$.

It follows from our assumptions that

$$\int \mu(\mathrm{d}A)1_{\{A\cap B=\emptyset\}} = \int \nu(\mathrm{d}A)1_{\{A\cap B=\emptyset\}}$$

for each finite $B \subset \mathbb{Z}^d$, hence $\int \mu(\mathrm{d}A)f(A) = \int \nu(\mathrm{d}A)f(A)$ for all $f \in \mathcal{F}$ and therefore, since $\mathcal{F}$ is dense, $\int \mu(\mathrm{d}A)f(A) = \int \nu(\mathrm{d}A)f(A)$ for all $f \in \mathcal{C}(\mathcal{P}(\mathbb{Z}^d))$, which implies $\mu = \nu$. ∎

Exercise 2.3 For each unordered pair $\{i,j\}$ of nearest neighbors (i.e., $i,j \in \mathbb{Z}^d$ such that $|i-j|=1$), let us define a local map $\tilde{m}_{ij} : \{0,1\}^{\mathbb{Z}^d} \to \{0,1\}^{\mathbb{Z}^d}$ by

$$(\tilde{m}_{ij}x)(k) := \begin{cases} 1 & \text{if } k \in \{i,j\},\ \max\{x(i),x(j)\}=1, \\ x(k) & \text{otherwise,} \end{cases}$$

$(i,j,k \in \mathbb{Z}^d,\ x \in \{0,1\}^{\mathbb{Z}^d})$. Note that this says that if $(x(i),x(j)) = (0,1)$ or $(1,0)$, then $\tilde{m}_{ij}$ changes this into $(1,1)$; otherwise nothing happens. Show that the generator of the contact process can be written in the form (compare (2.3))

$$Gf(x) = \lambda \sum_{\{i,j\}} \big(f(\tilde{m}_{ij}(x)) - f(x)\big) + \sum_{i\in\mathbb{Z}^d} \big(f(p_i(x)) - f(x)\big), \tag{2.6}$$

$(x \in \{0,1\}^{\mathbb{Z}^d})$, where the sum now runs over all *unordered* nearest-neighbor pairs. What kind of graphical representation results from writing the generator in the form (2.6)?

Exercise 2.4 Invent graphical representations for the interacting particle systems on $\mathbb{Z}$ with generators (compare (2.1))

$$\begin{aligned} G'f(x) = {} & \lambda \sum_{i\in\mathbb{Z}} 1_{\{x(i)=0\}} 1_{\{x(i-1)+x(i+1)>0\}} \big(f(x^{\{i\}}) - f(x)\big) \\ & + \sum_{i\in\mathbb{Z}} 1_{\{x(i)=1\}} \big(f(x^{\{i\}}) - f(x)\big) \end{aligned}$$

and

$$\begin{aligned} G''f(x) = {} & \lambda \sum_{i\in\mathbb{Z}} 1_{\{x(i)=0\}} \big(x(i-1) + x(i+1)\big)^2 \big(f(x^{\{i\}}) - f(x)\big) \\ & + \sum_{i\in\mathbb{Z}} 1_{\{x(i)=1\}} \big(f(x^{\{i\}}) - f(x)\big). \end{aligned}$$

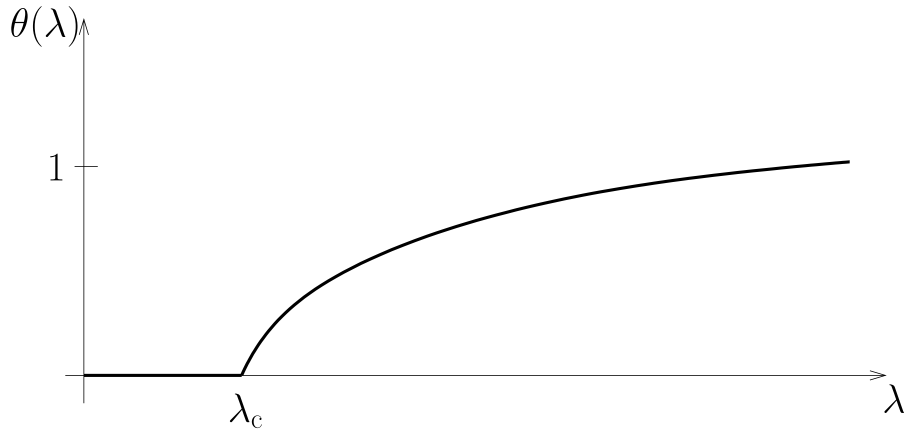

Figure 2.2: Survival probability.

2.2 The survival probability

By definition, we say that the nearest-neighbor contact process on $\mathbb{Z}^d$ with infection rate λ *survives* if

$$\theta(\lambda, d) = \theta(\lambda) := \mathbb{P}\big[\eta_t^{\{0\}} \neq \emptyset \ \forall t \geq 0\big] > 0.$$

If this probability is zero, then we say that the contact process *dies out* or *gets extinct*.
By a combination of rigorous mathematics, nonrigourous methods, and computer simulations, theoretical physicists have discovered the following properties of the function θ. There exists a *critical value* $\lambda_c = \lambda_c(d)$ with $0 < \lambda_c < \infty$ such that $\theta(\lambda) = 0$ for $\lambda \leq \lambda_c$ and $\theta(\lambda) > 0$ for $\lambda > \lambda_c$. The function θ is continuous, strictly increasing and concave on $[\lambda_c, \infty)$ and satisfies $\lim_{\lambda\to\infty} \theta(\lambda) = 1$. One has

$$\lambda_c(1) = 1.6489 \pm 0.0002.$$

Moreover, $\lambda_c(d)$ is decreasing in d and satisfies

$$\lambda_c(d) \approx \frac{1}{2d} \qquad \text{as } d \to \infty, \tag{2.7}$$

where the notation $f(z) \approx g(z)$ as $z \to z_0$ means that

$$\lim_{z\to z_0} \frac{f(z)}{g(z)} = 1 \qquad \text{as } z \to z_0.$$

The behavior of θ near the critical point is very interesting. One has

$$\theta(\lambda) \sim (\lambda - \lambda_{\rm c})^{\beta} \qquad \text{as } \lambda \downarrow \lambda_{\rm c}, \tag{2.8}$$

where we write $f(z) \sim g(z)$ as $z \to z_0$ if

$$\lim_{z \to z_0} \frac{\log(f(z))}{\log(g(z))} = 1.$$

The constant $\beta = \beta(d)$ is a *critical exponent*, approximately given by

$$\begin{aligned} \beta(1) &\cong 0.276487, \\ \beta(2) &\cong 0.584, \\ \beta(3) &\cong 0.81, \\ \beta(d) &= 1 \qquad\qquad (d \geq 4). \end{aligned}$$

In dimensions $d \neq 4$, it is believed that (2.8) can be strengthened to $\theta(\lambda) \approx c(\lambda - \lambda_{\rm c})^{\beta}$ for some $0 < c < \infty$.
Below, we will prove some of the easier properties of the function θ, such as monotonicity, the existence of a critical parameter $\lambda_{\rm c}$, and the fact that θ is right-continuous everywhere and left-continuous everywhere except possibly at the critical point $\lambda_{\rm c}$. Proving that θ is left-continuous at $\lambda_{\rm c}$, which by our previous remarks is equivalent to the statement that $\theta(\lambda_{\rm c}) = 0$, kept probabilists occupied for some 15 years, untill Bezuidenhout and Grimmett proved this in their celebrated paper [BG90]. Quite recently, it has been proved that (2.8) holds with $\beta = 1$ if the dimension d is sufficiently large. The critical behavior in dimensions $d = 1, 2, 3$ remains very much an unsolved problem. Physicists come to their prediction (2.8) using (nonrigorous) renormalization group arguments, where critical exponents can be related to eigenvectors of linearized renormalization transformations near a fixed point. Mathematically, there are big problems even *defining* these renormalization transformations rigorously, let alone studying them.
In dimension $d = 1$ it is known rigorously that $1.539 < \lambda_{\rm c} < 1.943$ [ZG88, Lig95]. For bounds in higher dimensions (including a proof of (2.7)), see [Lig85]. As far as I know, nobody has any idea how to prove that θ is concave on $[\lambda_{\rm c}, \infty)$.

Lemma 2.5 (Survival versus extinction) *If the contact process survives, then*

$$\mathbb{P}\big[\eta^A_t \neq \emptyset \ \forall t \geq 0\big] > 0 \tag{2.9}$$

for each finite nonempty $A \subset \mathbb{Z}^d$. If the contact process dies out, then this probability is zero for each finite nonempty $A \subset \mathbb{Z}^d$.

Proof Let A be finite and nonempty. For obvious reasons we also denote the probability in (2.9) by

$$\mathbb{P}\big[(A \times \{0\}) \rightsquigarrow \infty\big].$$

Now choose any $i \in A$. Then

$$\begin{aligned}\mathbb{P}\big[(0,0) \rightsquigarrow \infty\big] &= \mathbb{P}\big[(i,0) \rightsquigarrow \infty\big] \leq \mathbb{P}\big[(A \times \{0\}) \rightsquigarrow \infty\big] \\ &= \mathbb{P}\big[\exists j \in A \text{ s.t. } (j,0) \rightsquigarrow \infty\big] \leq \sum_{j \in A} \mathbb{P}\big[(j,0) \rightsquigarrow \infty\big] = |A| \mathbb{P}\big[(0,0) \rightsquigarrow \infty\big],\end{aligned}$$

where we have used translation invariance and $|A|$ denotes the number of elements in A. ∎

In this and the next section, we will prove the following result.

Theorem 2.6 (Critical infection rate) *For each $d \geq 1$ there exists a $\lambda_{\rm c} = \lambda_{\rm c}(d)$ with $0 < \lambda_{\rm c} < \infty$ such that the nearest-neighbour contact process on $\mathbb{Z}^d$ with infection rate λ survives for $\lambda > \lambda_{\rm c}$ and dies out for $\lambda < \lambda_{\rm c}$.*

Note that this theorem says nothing about survival or extinction if $\lambda = \lambda_{\rm c}(d)$.

As a first step towards Theorem 2.6, we prove the following fact.

Lemma 2.7 (Monotone coupling) *Let $(\eta_t)_{t\geq 0}$ and $(\eta'_t)_{t\geq 0}$ be contact processes on $\mathbb{Z}^d$ with infection rates $0 \leq \lambda \leq \lambda'$ and deterministic initial states $\eta_0 = A$ and $\eta'_0 = A'$ satisfying $A \subset A'$. Then $(\eta_t)_{t\geq 0}$ and $(\eta'_t)_{t\geq 0}$ can be coupled such that*

$$\eta_t \subset \eta'_t \qquad (t \geq 0).$$

In particular, survival of the contact process with infection rate λ implies survival of the contact process with infection rate λ'.

Proof Let $0 \leq \lambda \leq \lambda'$. Let $\Delta^{\rm i}$ and $\tilde{\Delta}^{\rm i}$ be independent Poisson point sets on $\mathcal{E} \times \mathbb{R}$ with intensities $\lambda\,\mathrm{d}t$ and $(\lambda' - \lambda)\,\mathrm{d}t$, respectively, and let $\tilde{\Delta}^{\rm r}$ be a Poisson point set on $\mathbb{Z}^d \times \mathbb{R}$ with intensity $1\,\mathrm{d}t$, independent of $\Delta^{\rm i}$ and $\tilde{\Delta}^{\rm i}$. Then $\Delta^{\rm i} \cup \tilde{\Delta}^{\rm i}$ is a Poisson point set on $\mathcal{E} \times \mathbb{R}$ with intensity $\lambda'\,\mathrm{d}t$. We interpret points in $\Delta^{\rm i}$ and $\tilde{\Delta}^{\rm i}$ as infection arrows and points in $\tilde{\Delta}^{\rm r}$ as recovery symbols. We let $\rightsquigarrow$ indicate the presence of an open path that may use infection arrows from $\Delta^{\rm i}$ only and we write $\rightsquigarrow'$ to indicate the presence of an open path that may use infection arrows from $\Delta^{\rm i} \cup \tilde{\Delta}^{\rm i}$. Then

$$\eta_t = \{i : A \times \{0\} \rightsquigarrow (i,t)\} \subset \{i : A' \times \{0\} \rightsquigarrow' (i,t)\} = \eta'_t \qquad (t \geq 0)$$

since $A \subset A'$ and the process $(\eta'_t)_{t\geq 0}$ has more arrows at its disposal. ∎

2.3 Extinction

It follows from Lemma 2.7 that the function $\lambda \mapsto \theta(\lambda)$ is nondecreasing and hence, for each $d \geq 1$, there exists a $0 \leq \lambda_{\rm c}(d) \leq \infty$ such that the nearest-neighbour contact process on $\mathbb{Z}^d$ with infection rate λ survives for $\lambda > \lambda_{\rm c}$ and dies out for $\lambda < \lambda_{\rm c}$. To prove Theorem 2.6, we must show that $0 < \lambda_{\rm c}(d) < \infty$. We start by proving a the lower bound on $\lambda_{\rm c}$, which is easiest.

Lemma 2.8 (Exponential bound) *For each finite $A \subset \mathbb{Z}^d$, one has*

$$\mathbb{E}\big[|\eta^A_t|\big] \leq |A|e^{(2d\lambda-1)t} \qquad (t \geq 0). \tag{2.10}$$

Proof For the contact process, the constant K from (1.23) is given by $K = 2d\lambda - 1$. Therefore, the statement is a direct consequence of Lemma 1.13. ∎

Lemma 2.8 has the following consequence.

Corollary 2.9 (Lower bound on critical infection rate) *The critical infection rate of the nearest-neighbour contact process on $\mathbb{Z}^d$ satisfies $\frac{1}{2d} \leq \lambda_{\rm c}$.*

Proof By (2.10), for each $\lambda < \frac{1}{2d}$,

$$\mathbb{P}\big[\eta^A_t \neq \emptyset\big] \leq \mathbb{E}\big[|\eta^A_t|\big] \underset{t\to\infty}{\longrightarrow} 0$$

for each finite $A \subset \mathbb{Z}^d$. ∎

In order to finish the proof of Theorem 2.6 we need to show that $\lambda_{\rm c} < \infty$. As a preparation for this, in the next section, we will start by studying a closely related problem. Before we do this, we apply the techniques developed so far to prove that the function $\theta(\lambda, d)$ is nondecreasing and right-continuity. Left-continuity, except (possibly) in the critical point $\lambda_{\rm c}$, will be proved in Proposition 2.22 below.

Proposition 2.10 (Monotonicty and right-continuity) *The survival probability $\theta(\lambda, d)$ is nondecreasing and right-continuous in λ, and nondecreasing in d.*

Proof The fact that $\theta(\lambda, d)$ is nondecreasing in λ follows from Lemma 2.7. The fact that $\theta(\lambda, d)$ is nondecreasing in d can be proved in a similar way, since if $d \leq d'$, then we may view $\mathbb{Z}^d$ as a subset of $\mathbb{Z}^{d'}$ and observe that if there is an open path that stays in $\mathbb{Z}^d$, then certainly there is an open path in $\mathbb{Z}^{d'}$.
To prove right continuity of $\theta(\lambda, d)$ in λ, we will improve the coupling used in the proof of Lemma 2.7 in such a way that we can define contact processes for any value of the infection rate on the same probability space. To this aim, consider

the space $\mathcal{E}\times\mathbb{R}\times[0,\infty)$ whose elements are triples $((i,j),t,\kappa)$ with $(i,j)\in\mathcal{E}$ and $t\in\mathbb{R},\kappa\geq 0$, and let $\overline{\Delta}^{\rm i}$ be a Poisson point set on this set with indensity $\mathrm{d}t\mathrm{d}\kappa$. Then, for each $\lambda\geq 0$,

$$\Delta^{\rm i}_\lambda := \big(\big((i,j),t\big) : \exists\big((i,j),t,\kappa\big)\in\overline{\Delta}^{\rm i} \text{ with } \kappa\leq\lambda\big\}.$$

is a Poisson point sets on $\mathcal{E}\times\mathbb{R}$ with intensity $\lambda\mathrm{d}t$. Let $\Delta^{\rm r}$ be an independent Poisson point set on $\mathbb{Z}^d\times\mathbb{R}$ with intensity $\mathrm{d}t$ and write $\rightsquigarrow_\lambda$ to indicate the presence of an open path in the graphical representations defined by $(\Delta^{\rm i}_\lambda,\Delta^{\rm r})$. Another way of saying this is that a point $((i,j),t,\kappa)\in\overline{\Delta}^{\rm i}$ indicates the presence of an arrow which has a value κ attached to it, and $\rightsquigarrow_\lambda$ indicates the presence of a path that may use only arrows with values $\kappa\leq\lambda$. Let $\lambda_n\downarrow\lambda_*$. Then we claim that

$$\begin{aligned}\lim_{\lambda_n\downarrow\lambda_*}\theta(\lambda_n) &= \lim_{\lambda\downarrow\lambda_*}\mathbb{P}[(0,0)\rightsquigarrow_\lambda\infty] = \mathbb{P}[(0,0)\rightsquigarrow_\lambda\infty\ \forall\lambda>\lambda_*]\\ &\overset{!}{=}\mathbb{P}[(0,0)\rightsquigarrow_{\lambda_*}\infty]=\theta(\lambda_*).\end{aligned}$$

The equality $\overset{!}{=}$ needs some explanation. It is obvious that $(0,0)\rightsquigarrow_{\lambda_*}\infty$ implies $(0,0)\rightsquigarrow_\lambda\infty\ \forall\lambda>\lambda_*$. On the other hand, if $(0,0)\not\rightsquigarrow_{\lambda_*}\infty$ then by Lemma 2.8

$$\mathcal{I} := \big\{(i,t)\in\mathbb{Z}^d\times\mathbb{R} : (0,0)\rightsquigarrow_{\lambda_*}(i,t)\big\}$$

is a compact subset of $\mathbb{Z}^d\times\mathbb{R}$, such that each each infection ends somewhere in a recovery sign, and all infection arrows starting in $\mathcal{I}$ and ending somewhere outside $\mathcal{I}$ have a value strictly larger than λ_*. Since there are only finitely many arrows with values $\kappa\in(\lambda_*,2\lambda_*)$ starting in $\mathcal{I}$ and ending somewhere outside $\mathcal{I}$, we know that there is some $\lambda'>\lambda_*$ such that all arrows starting in $\mathcal{I}$ and ending somewhere outside $\mathcal{I}$ have a value larger than λ', i.e., we know that $(0,0)\not\rightsquigarrow_{\lambda'}\infty$ for some $\lambda'>\lambda_*$. ∎

2.4 Oriented percolation

In order to prepare for the proof that the critical infection rate of the contact process is finite, in the present section, we will study *oriented* (or *directed*) *bond percolation* on $\mathbb{Z}^d$. For $i,j\in\mathbb{Z}^d$, we write $i\leq j$ if $i=(i_1,\ldots,i_d)$ and $j=(j_1,\ldots,j_d)$ satisfy $i_k\leq j_k$ for all $k=1,\ldots,d$. Let

$$\mathcal{A} := \big\{(i,j) : i,j\in\mathbb{Z}^d,\ i\leq j,\ |i-j|=1\big\}.$$

We view $\mathbb{Z}^d$ as an infinite directed graph, where elements $(i,j)\in\mathcal{A}$ represent arrows (or directed *bonds*) between neighbouring sites. Note that all arrows point 'upwards' in the sense of the natural order on $\mathbb{Z}^d$.

Now fix some *percolation parameter* $p \in [0,1]$ and let $(\omega_{(i,j)})_{(i,j)\in\mathcal{A}}$ be a collection of i.i.d. Bernoulli random variables with $\mathbb{P}[\omega_{(i,j)} = 1] = p$. We say that there is an *open path* from a site $i \in \mathbb{Z}^d$ to $j \in \mathbb{Z}^d$ if there exist $n \geq 0$ and a function $\gamma : \{0, \ldots, n\} \to \mathbb{Z}^d$ such that

$$(\gamma(k-1), \gamma(k)) \in \mathcal{A} \quad \text{and} \quad \omega_{(\gamma(k-1),\gamma(k))} = 1 \qquad (k = 1, \ldots, n).$$

We denote the presence of an open path by $\rightsquigarrow$. Note that open paths must walk upwards in the sense of the order on $\mathbb{Z}^d$. We write $0 \rightsquigarrow \infty$ to indicate the existence of an infinite open path starting at the origin $0 \in \mathbb{Z}^d$.

Theorem 2.11 (Critical percolation parameter) *For oriented percolation in dimensions $d \geq 2$ there exists a critical parameter $p_{\rm c} = p_{\rm c}(d)$ with $0 < p_{\rm c} < 1$ such that $\mathbb{P}[0 \rightsquigarrow \infty] = 0$ for $p < p_{\rm c}$ and $\mathbb{P}[0 \rightsquigarrow \infty] > 0$ for $p > p_{\rm c}$.*

Proof The existence of a critical parameter $p_{\rm c} \in [0,1]$ follows from a monotone coupling argument like the one we used in the proof of Lemma 2.7. To prove that $0 < p_{\rm c}$, let N_n denote the number of open paths of length n starting in 0. Since there are d^n different upward paths of length n starting at the origin, and each path has probability p^n to be open, we see that

$$\mathbb{E}\big[\sum_{n=1}^{\infty} N_n\big] = \sum_{n=1}^{\infty} d^n p^n < \infty \qquad (p < 1/d)$$

This shows that $\sum_{n=1}^{\infty} N_n < \infty$ a.s., hence $\mathbb{P}[0 \rightsquigarrow \infty] = 0$ if $p < 1/d$, and therefore

$$\frac{1}{d} \leq p_{\rm c}(d).$$

To prove that $p_{\rm c}(d) < 1$ for $d \geq 2$ it suffices to consider the case $d = 2$, for we may view $\mathbb{Z}^2$ as a subset of $\mathbb{Z}^d$ ($d \geq 3$) and then, if there is an open path that stays in $\mathbb{Z}^2$, then certainly there is an open path in $\mathbb{Z}^d$. (Note, by the way, that in $d = 1$ one has $\mathbb{P}[0 \rightsquigarrow \infty] = 0$ for all $p < 1$ hence $p_{\rm c}(1) = 1$.)

We claim that

$$p_{\rm c}(2) \leq \frac{8}{9}. \tag{2.11}$$

To prove this, we use a Peierls argument, named after R. Peierls who used a similar argument in 1936 for the Ising model [Pei36]. In Figure 2.3, we have drawn a piece of $\mathbb{Z}^2$. Open arrows are drawn in black; closed arrows are not drawn. Sites $i \in \mathbb{Z}^2$ such that $0 \rightsquigarrow i$ are indicated in black. These sites are called *wet*. Consider the *dual lattice*

$$\hat{\mathbb{Z}}^2 := \{(n + \tfrac{1}{2}, m + \tfrac{1}{2}) : (n,m) \in \mathbb{Z}^2\}.$$

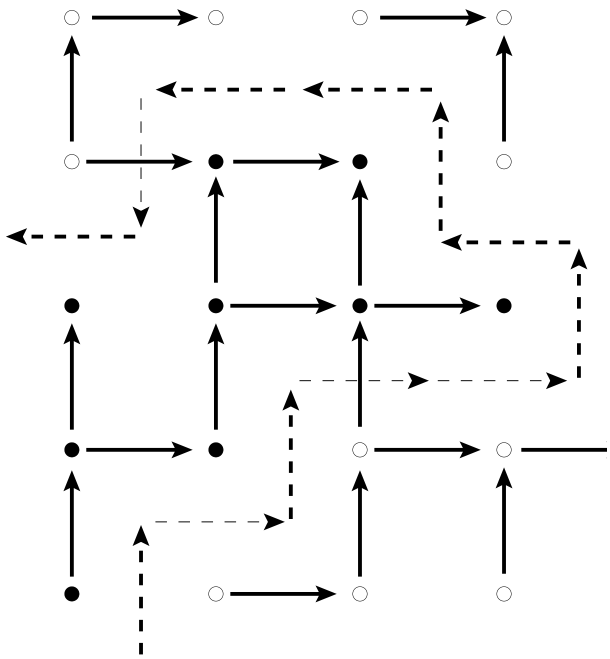

Figure 2.3: Peierls argument for oriented percolation.

If there are only finitely many wet sites, then the set of all non-wet sites contains one infinite connected component. (Here 'connected' is to be interpreted in terms of the unoriented graph $\mathbb{N}^2$ with nearest-neighbor edges.) Let γ be the boundary of this infinite component. Then γ is a nearest-neighbor path in $\hat{\mathbb{Z}}^2$, starting in some point $(n+\frac{1}{2},-\frac{1}{2})$ and ending in some point $(-\frac{1}{2},m+\frac{1}{2})$ with $n,m\geq 0$, such that all sites immediately to the left of γ are wet, and no open arrows starting at these sites cross γ. In Figure 2.3, we have indicated γ with dashed arrows.

From these considerations, we see that the following statement is true: one has $0\not\rightsquigarrow\infty$ if and only if there exists a path in $\hat{\mathbb{Z}}^2$, starting in some point $(n+\frac{1}{2},-\frac{1}{2})$ $(n\geq 0)$, ending in some point $(-\frac{1}{2},m+\frac{1}{2})$ $(m\geq 0)$, and passing to the northeast of the origin, such that all arrows of γ in the north and west directions (indicated in bold in the figure) are not be crossed by an open arrow. Let M_n be the number of paths of length n with these properties. Since there are n dual sites from where such a path of length n can start, and since in each step, there are three directions where it can go, there are at most $n3^n$ paths of length n with these properties. Since each path must make at least half of its steps in the north and east directions, the expected number of these paths satisfies

$$\mathbb{E}\big[\sum_{n=2}^{\infty}M_n\big]\leq\sum_{n=2}^{\infty}n3^n(1-p)^{n/2}<\infty\qquad(p>\tfrac{8}{9})$$

and therefore

$$\mathbb{P}[0\not\rightsquigarrow\infty]\leq\mathbb{P}\big[\sum_{n=2}^{\infty}M_n\geq 1\big]\leq\mathbb{E}\big[\sum_{n=2}^{\infty}M_n\big]<\infty.$$

This does not quite prove what we want yet, since we need the right-hand side of this equation to be less than one. To fix this, set $D_m:=\{0,\ldots,m\}^2$. Then, by the same arguments as before

$$\mathbb{P}[D_m\not\rightsquigarrow\infty]\leq\mathbb{P}\big[\sum_{n=2m}^{\infty}M_n\geq 1\big]\leq\mathbb{E}\big[\sum_{n=2m}^{\infty}M_n\big]\leq\sum_{n=2m}^{\infty}n3^n(1-p)^{n/2},$$

which in case $p>\frac{8}{9}$ can be made arbitrarily small by choosing m suffiently large. It follows that $\mathbb{P}[D_m\rightsquigarrow\infty]>0$ for some m, hence $\mathbb{P}[i\rightsquigarrow\infty]>0$ for some $i\in D_m$, and therefore, by translation invariance, also $\mathbb{P}[0\rightsquigarrow\infty]>0$. ∎

2.5 Survival

In the present section, we will complete the proof of Theorem 2.6 by showing that $\lambda_{\rm c}<\infty$. The method we will use is comparison with oriented percolation. This

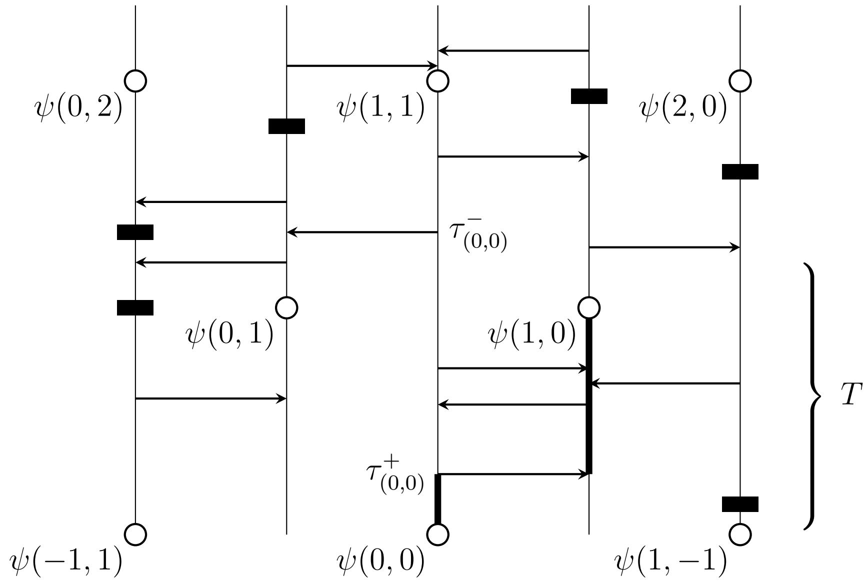

Figure 2.4: Comparison with oriented percolation. In this example, the good event $\mathcal{G}^+_{(0,0)}$ occurs because of the arrow at time $\tau^+_{(0,0)}$ and because of the fact that there are no recovery symbols on the thick line segments. On the other hand, the good event $\mathcal{G}^-_{(0,0)}$ does not occur since $\tau^-_{(0,0)} > \sigma_{(0,1)}$. Note that also $\mathcal{G}^-_{(0,1)}$ does not occur even though there is an open path $\psi(0,1) \rightsquigarrow \psi(0,2)$.

neither leads to a particularly short proof nor does it yield a very good upper bound on λ_c, but it has the advantage that it is a very robust method that can be applied to many other interacting particle systems. (Alternatively, it is also possible to adapt the proof of Theorem 2.11 to the continuous-time setting of the graphical representation of the one-dimensional contact process. This leads to a better bound on λ_c but the method is much less flexible.)

Let $\lambda_c(d)$ be the critical infection rate of the nearest-neighbour contact process on $\mathbb{Z}^d$. If $d \leq d'$, then we may view $\mathbb{Z}^d$ as a subset of $\mathbb{Z}^{d'}$, so by an obvious monotone coupling we see that

$$\lambda_c(d) \geq \lambda_c(d') \qquad (d \leq d').$$

In view of this, in order to finish the proof of Theorem 2.6, it suffices to show that $\lambda_c(1) < \infty$.

We use the graphical representation, i.e., we let Δ^{i} and Δ^{r} be Poisson point subsets of $\mathcal{E} \times \mathbb{R}$ and $\mathbb{Z}^d \times \mathbb{R}$, respectively. We fix $T > 0$ and define a map $\psi : \mathbb{Z}^2 \to \mathbb{Z} \times \mathbb{R}$

by

$$\psi(i) = \big(\kappa_i, \sigma_i\big) := \big(i_1 - i_2, T(i_1 + i_2)\big) \qquad \big(i = (i_1, i_2) \in \mathbb{Z}^2\big).$$

Recall from the previous section the definition of the set $\mathcal{A}$ of arrows on $\mathbb{Z}^d$. We wish to define a collection $(\omega_{(i,j)})_{(i,j)\in\mathcal{A}}$ of Bernoulli random variables such that

$$\omega_{(i,j)} = 1 \quad \text{implies} \quad (\kappa_i, \sigma_i) \rightsquigarrow (\kappa_j, \sigma_j) \qquad \big((i,j) \in \mathcal{A}\big).$$

For each $i \in \mathbb{Z}^2$ we let

$$\begin{aligned} \tau_i^- &:= \inf\big\{t \geq \sigma_i : (\kappa_i, \kappa_i - 1, t) \in \Delta^{\mathrm{i}}\big\}, \\ \tau_i^+ &:= \inf\big\{t \geq \sigma_i : (\kappa_i, \kappa_i + 1, t) \in \Delta^{\mathrm{i}}\big\}, \end{aligned}$$

denote the times of the first arrows out of i to the left and right, and we define the 'good events'

$$\begin{aligned} \mathcal{G}_i^- := \big\{&\tau_i^- < \sigma_i + T,\ \big(\{\kappa_i\} \times (\sigma_i, \tau_i^-)\big) \cap \Delta^{\mathrm{r}} = \emptyset, \\ &\big(\{\kappa_i - 1\} \times (\tau_i^-, \sigma_i + T)\big) \cap \Delta^{\mathrm{r}} = \emptyset\big\}, \\ \mathcal{G}_i^+ := \big\{&\tau_i^+ < \sigma_i + T,\ \big(\{\kappa_i\} \times (\sigma_i, \tau_i^+)\big) \cap \Delta^{\mathrm{r}} = \emptyset, \\ &\big(\{\kappa_i + 1\} \times (\tau_i^+, \sigma_i + T)\big) \cap \Delta^{\mathrm{r}} = \emptyset\big\}. \end{aligned}$$

We observe that the event $\mathcal{G}_i^\pm$ implies that $(\kappa_i, \sigma_i) \rightsquigarrow (\kappa_i \pm 1, \sigma_i + T)$ via an open path that stays in $\{\kappa_i, \kappa_i \pm 1\}$. In view of this, we set

$$\begin{aligned} \omega_{((i_1,i_2),(i_1+1,i_2))} &:= 1_{\mathcal{G}_i^+}, \\ \omega_{((i_1,i_2),(i_1,i_2+1))} &:= 1_{\mathcal{G}_i^-}. \end{aligned}$$

Then $\omega_{(i,j)} = 1$ implies the existence of an open path in the graphical representation for the contact process from (κ_i, σ_i) to (κ_j, σ_j) (with $(i,j) \in \mathcal{A}$), hence if we use the random variables $(\omega_{(i,j)})_{(i,j)\in\mathcal{A}}$ to define oriented percolation on $\mathbb{Z}^2$ in the usual way, then:

> $i \rightsquigarrow j$ in the oriented percolation on $\mathbb{Z}^2$ defined by the random variables $(\omega_{(i,j)})_{(i,j)\in\mathcal{A}}$ implies $(\kappa_i, \sigma_i) \rightsquigarrow (\kappa_i, \sigma_i)$ in the graphical representation for the contact process.

We observe that

$$p := \mathbb{P}[\omega_{(i,j)} = 1] = \mathbb{P}(\mathcal{G}_i^\pm) = (1 - e^{-\lambda T})e^{-T} \qquad \big((i,j) \in \mathcal{A}\big). \tag{2.12}$$

For λ sufficiently large, by a suitable choice of T, we can make p as close to one as we wish. We would like to conclude from this that $\mathbb{P}[(0,0) \rightsquigarrow \infty] > 0$ for the oriented percolation defined by the $\omega_{(i,j)}$'s, and therefore also $\mathbb{P}[(0,0) \rightsquigarrow \infty] > 0$ for the contact process. Unfortunately, the random variables $(\omega_{(i,j)})_{(i,j)\in\mathcal{A}}$ are not independent, and therefore Theorem 2.11 is not applicable. To fix this problem, we need a bit more theory. In the next section, we will introduce the concept of k-dependence. As will be clear from the definition there, the $(\omega_{(i,j)})_{(i,j)\in\mathcal{A}}$ are k-dependent for some suitable k, so by applying Theorem 2.12 from the next section we can estimate them from below by an independent collection of Bernoulli random variables $(\tilde\omega_{(i,j)})_{(i,j)\in\mathcal{A}}$ whose succes probability $\tilde p$ can be made arbitrarily close to one, so we are done. ∎

2.6 K-dependence

By definition, for $k \geq 0$, one says that a collection $(X_i)_{i\in\mathbb{Z}^d}$ of random variables, indexed by the integer square lattice, is *k-dependent* if for any $A, B \subset \mathbb{Z}^d$ with

$$\inf\{|i-j| : i \in A,\ j \in B\} > k,$$

the collections of random variables $(X_i)_{i\in A}$ and $(X_j)_{j\in B}$ are independent of each other. Note that in particular, 0-dependence means independence.

The most important property associated with k-dependence is that a collection of k-dependent Bernoulli random variables with success probability p can be stochastically estimated from below by a collection of independent Bernoulli random variables with a success probability $\tilde p$ that has the property that $\tilde p \to 1$ as $p \to 1$. It is a bit unfortunate that the term 'k-dependence' as it is standardly used explicity (and only) refers to random variables on $\mathbb{Z}^d$, while in fact, as the next theorem shows, for the property just mentioned the precise spatial structure is not very important. The next theorem is taken from [Lig99, Thm B26], who in turn cites [LSS97].

Theorem 2.12 (K-dependence) *Let Λ be a countable set and let $p \in (0,1)$, $K < \infty$. Assume that $(\chi_i)_{i\in\Lambda}$ are Bernoulli random variables with $P[\chi_i = 1] \geq p$ $(i \in \Lambda)$, such that for each $i \in \Lambda$ there exists a $\Delta_i \subset \Lambda$ with $i \in \Delta_i$ and $|\Delta_i| \leq K$, such that*

$$\chi_i \text{ is independent of } (\chi_j)_{j\in\Lambda\setminus\Delta_i}.$$

Then it is possible to couple $(\chi_i)_{i\in\Lambda}$ to a collection of independent Bernoulli random variables $(\tilde\chi_i)_{i\in\Lambda}$ with

$$P[\tilde\chi_i = 1] = \tilde p := \big(1 - (1-p)^{1/K}\big)^2, \tag{2.13}$$

in such a way that $\tilde\chi_i \leq \chi_i$ for all $i \in \Lambda$.

Proof of Theorem 2.12 Since we can always choose some arbitrary denumeration of Λ, we may assume that $\Lambda = \mathbb{N}$. Our strategy will be as follows. We will choose $\{0,1\}$-valued random variables $(\psi_i)_{i\in\Lambda}$ with $P[\psi_i = 1] = r$, independent of each other and of the $(\chi_i)_{i\in\mathbb{N}}$, and put

$$\chi'_i := \psi_i \chi_i \qquad (i \in \mathbb{N}). \tag{2.14}$$

Note that the $(\chi'_i)_{i\in\mathbb{N}}$ are a 'thinned out' version of the $(\chi_i)_{i\in\mathbb{N}}$. In particular, $\chi'_i \leq \chi_i$ $(i \in \mathbb{N})$. We will show that for an appropriate choice of r,

$$P[\chi'_n = 1 \,|\, \chi'_0, \ldots, \chi'_{n-1}] \geq \tilde{p} \tag{2.15}$$

for all $n \geq 0$, and we will show that this implies that the $(\chi'_i)_{i\in\mathbb{N}}$ can be coupled to independent $(\tilde{\chi}_i)_{i\in\Lambda}$ as in (2.13) in such a way that $\tilde{\chi}_i \leq \chi'_i \leq \chi_i$ $(i \in \mathbb{N})$.
We start with the latter claim. Imagine that (2.15) holds. Set

$$p'_n(\varepsilon_0, \ldots, \varepsilon_{n-1}) := P[\chi'_n = 1 \,|\, \chi'_0 = \varepsilon_0, \ldots, \chi'_{n-1} = \varepsilon_{n-1}] \tag{2.16}$$

whenever $P[\chi'_0 = \varepsilon_0, \ldots, \chi'_{n-1} = \varepsilon_{n-1}] > 0$. Let $(U_n)_{n\in\mathbb{N}}$ be independent, uniformly distributed $[0,1]$-valued random variables. Set

$$\tilde{\chi}_n := 1_{\{U_n < \tilde{p}\}} \qquad (n \in \mathbb{N}) \tag{2.17}$$

and define inductively

$$\chi'_n := 1_{\{U_n < p'_n(\chi'_0, \ldots, \chi'_{n-1})\}} \qquad (i \in \mathbb{N}). \tag{2.18}$$

Then

$$P[\chi'_n = \varepsilon_n, \ldots, \chi'_0 = \varepsilon_0] = p_n(\varepsilon_0, \ldots, \varepsilon_{n-1}) \cdots p_0. \tag{2.19}$$

This shows that these new χ'_n's have the same distribution as the old ones, and they are coupled to $\tilde{\chi}_i$'s as in (2.13) in such a way that $\tilde{\chi}_i \leq \chi'_i$.
What makes life complicated is that (2.15) does not always hold for the original $(\chi_i)_{i\in\mathbb{N}}$, which is why we have to work with the thinned variables $(\chi'_i)_{i\in\mathbb{N}}$.[1] We observe that

$$P[\chi'_n = 1 \,|\, \chi'_0 = \varepsilon_0, \ldots, \chi'_{n-1} = \varepsilon_{n-1}] = rP[\chi_n = 1 \,|\, \chi'_0 = \varepsilon_0, \ldots, \chi'_{n-1} = \varepsilon_{n-1}]. \tag{2.20}$$

[1] Indeed, let $(\phi_n)_{n\geq 0}$ be independent $\{0,1\}$-valued random variables with $P[\phi_n = 1] = \sqrt{p}$ for some $p < 1$, and put $\chi_n := \phi_n \phi_{n+1}$. Then the $(\chi_n)_{n\geq 0}$ are 1-dependent with $P[\chi_n = 1] = p$, but $P[\chi_n = 1 | \chi_{n-1} = 0, \chi_{n-2} = 1] = 0$.

We will prove by induction that for an appropriate choice of r,

$$P[\chi_n = 0 \,|\, \chi'_0 = \varepsilon_0, \ldots, \chi'_{n-1} = \varepsilon_{n-1}] \leq 1 - r. \tag{2.21}$$

Note that this is true for $n = 0$ provided that $r \leq p$. Let us put

$$\begin{aligned} E_0 &:= \{i \in \Delta_n : 0 \leq i \leq n-1,\ \varepsilon_i = 0\}, \\ E_1 &:= \{i \in \Delta_n : 0 \leq i \leq n-1,\ \varepsilon_i = 1\}, \\ F &:= \{i \notin \Delta_n : 0 \leq i \leq n-1\}. \end{aligned} \tag{2.22}$$

Then

$$\begin{aligned} & P[\chi_n = 0 \,|\, \chi'_0 = \varepsilon_0, \ldots, \chi'_{n-1} = \varepsilon_{n-1}] \\ &\quad = P\big[\chi_n = 0 \,\big|\, \chi'_i = 0\ \forall i \in E_0,\ \chi_i = 1 = \psi_i\ \forall i \in E_1,\ \chi'_i = \varepsilon_i\ \forall i \in F\big] \\ &\quad = P\big[\chi_n = 0 \,\big|\, \chi'_i = 0\ \forall i \in E_0,\ \chi_i = 1\ \forall i \in E_1,\ \chi'_i = \varepsilon_i\ \forall i \in F\big] \\ &\quad = \frac{P\big[\chi_n = 0,\ \chi'_i = 0\ \forall i \in E_0,\ \chi_i = 1\ \forall i \in E_1,\ \chi'_i = \varepsilon_i\ \forall i \in F\big]}{P\big[\chi'_i = 0\ \forall i \in E_0,\ \chi_i = 1\ \forall i \in E_1,\ \chi'_i = \varepsilon_i\ \forall i \in F\big]} \\ &\quad \leq \frac{P\big[\chi_n = 0,\ \chi'_i = \varepsilon_i\ \forall i \in F\big]}{P\big[\psi_i = 0\ \forall i \in E_0,\ \chi_i = 1\ \forall i \in E_1,\ \chi'_i = \varepsilon_i\ \forall i \in F\big]} \\ &\quad = \frac{P\big[\chi_n = 0 \,\big|\, \chi'_i = \varepsilon_i\ \forall i \in F\big]}{P\big[\psi_i = 0\ \forall i \in E_0,\ \chi_i = 1\ \forall i \in E_1 \,\big|\, \chi'_i = \varepsilon_i\ \forall i \in F\big]} \\ &\quad \leq \frac{1-p}{(1-r)^{|E_0|} P\big[\chi_i = 1\ \forall i \in E_1 \,\big|\, \chi'_i = \varepsilon_i\ \forall i \in F\big]} \leq \frac{1-p}{(1-r)^{|E_0|}\, r^{|E_1|}}, \end{aligned} \tag{2.23}$$

where in the last step we have used K-dependence and the (nontrivial) fact that

$$P\big[\chi_i = 1\ \forall i \in E_1 \,\big|\, \chi'_i = \varepsilon_i\ \forall i \in F\big] \geq r^{|E_1|}. \tag{2.24}$$

We claim that (2.24) is a consequence of the induction hypothesis (2.21). Indeed, we may assume that the induction hypothesis (2.21) holds regardless of the ordering of the first n elements, so without loss of generality we may assume that $E_1 = \{n-1, \ldots, m\}$ and $F = \{m-1, \ldots, 0\}$, for some m. Then the left-hand side of (2.24) may be written as

$$\begin{aligned} & \prod_{k=m}^{n-1} P\big[\chi_k = 1 \,\big|\, \chi_i = 1\ \forall m \leq i < k,\ \chi'_i = \varepsilon_i\ \forall 0 \leq i < m\big] \\ &\quad = \prod_{k=m}^{n-1} P\big[\chi_k = 1 \,\big|\, \chi'_i = 1\ \forall m \leq i < k,\ \chi'_i = \varepsilon_i\ \forall 0 \leq i < m\big] \geq r^{n-m}. \end{aligned} \tag{2.25}$$

If we assume moreover that $r \geq \frac{1}{2}$, then $r^{|E_1|} \geq (1-r)^{|E_1|}$ and therefore the right-hand side of (2.23) can be further estimated as

$$\frac{1-p}{(1-r)^{|E_0|}\, r^{|E_1|}} \leq \frac{1-p}{(1-r)^{|\Delta_n \cap \{0,\ldots,n-1\}|}} \leq \frac{1-p}{(1-r)^{K-1}}. \tag{2.26}$$

We see that in order for our proof to work, we need $\frac{1}{2} \leq r \leq p$ and

$$\frac{1-p}{(1-r)^{K-1}} \leq 1-r. \tag{2.27}$$

In particular, choosing $r = 1-(1-p)^{1/K}$ yields equality in (2.27). Having proved (2.21), we see by (2.20) that (2.15) holds provided that we put $\tilde{p} := r^2$. ∎

Exercise 2.13 Combine formulas (2.11), (2.12) and (2.13) to derive an explicit upper bound on the critical infection rate $\lambda_{\rm c}$ of the one-dimensional contact process.

2.7 The upper invariant law

For any Feller semigroup $(P_t)_{t\geq 0}$ on $\mathcal{C}(E)$, where E is some compact metrizable space, we say that a probability measure μ on E is an *invariant law* if

$$\int \mu(\mathrm{d}x)\mathbb{P}_t(x,\,\cdot\,) = \mu \qquad (t \geq 0).$$

Note that this says that for the associated Feller process, $\mathbb{P}[X_0 \in \cdot\,] = \mu$ implies $\mathbb{P}[X_t \in \cdot\,] = \mu$ for each $t \geq 0$. If μ is an invariant law, then it is possible to construct a stationary process $(X_t)_{t\in\mathbb{R}}$ that is also defined for negative times, such that $\mathbb{P}[X_t \in \cdot\,] = \mu$ for all $t \in \mathbb{R}$ and $(X_t)_{t\in\mathbb{R}}$ is a Feller process associated with $(P_t)_{t\geq 0}$, where we generalize (1.11) in the obvious way to negative times.

Using our graphical representation of the contact process, we define

$$\overline{\eta}_t := \big\{ i \in \mathbb{Z}^d : -\infty \rightsquigarrow (i,t) \big\} \qquad (t \in \mathbb{R}).$$

where $-\infty \rightsquigarrow (i,t)$ indicates the presence of an open path $\gamma : (-\infty, t] \to \mathbb{Z}^d$. (Note that so far, we have only defined paths with finite starting and ending times but the definition of an (open) path can easily be extended to allow for infinite paths.) Using the independence of restrictions of Poisson point processes to disjoint parts of space, we see that

$$\mathbb{P}\big[\overline{\eta}_u \in \cdot \,\big|\, (\overline{\eta}_s)_{s\leq t}\big] = P_{u-t}(\overline{\eta}_t,\,\cdot\,) \quad \text{a.s.} \quad (t \leq u),$$

hence $(\overline{\eta}_t)_{t\in\mathbb{R}}$ is a stationary contact process and its law at any given time

$$\overline{\nu} := \mathbb{P}\big[\overline{\eta}_t \in \cdot\,\big] \qquad (t \in \mathbb{R})$$

is an invariant law of the contact process. We call $\overline{\nu}$ the *upper invariant law* of the contact process (with given infection rate). As we will see in a moment, in a certain sense, it is the 'largest' invariant law of our process.

By definition, we say that a function $f : \{0,1\}^{\mathbb{Z}^d}$ is *monotone* if $f(x) \leq f(y)$ for all $x \leq y$.

Proposition 2.14 (Stochastic order) *Let μ, ν be probability laws on $\{0,1\}^{\mathbb{Z}^d}$. Then the following statements are equivalent:*

(i) $\int \mu(\mathrm{d}x) f(x) \leq \int \nu(\mathrm{d}x) f(x)\ \forall$ *monotone* $f \in \mathcal{C}(\{0,1\}^{\mathbb{Z}^d})$,

(ii) $\int \mu(\mathrm{d}x) f(x) \leq \int \nu(\mathrm{d}x) f(x)\ \forall$ *monotone* $f \in B(\{0,1\}^{\mathbb{Z}^d})$,

(iii) *It is possible to couple random variables X, Y with laws $\mu = P[X \in \cdot\,]$ and $\nu = P[Y \in \cdot\,]$ in such a way that $X \leq Y$.*

Proof The implications (iii)⇒(ii)⇒(i) are trivial. For the nontrivial implication (i)⇒(iii) (which we will never actually need to use), see [Lig85, Theorem II.2.4]. ∎

Remark The statements of Proposition 2.14 are not restricted to probability laws on $\{0,1\}^{\mathbb{Z}^d}$. Analogue statements hold for probability laws on $\mathbb{R}$ (equipped with the usual order) or more generally $\mathbb{R}^n$, as well as for other metric spaces on which an order is defined that is somehow 'compatible' with the topology; see [Lig85, Theorem II.2.4].

If probability laws μ, ν on $\{0,1\}^{\mathbb{Z}^d}$ satisfy the equivalent conditions (i)–(iii) from Proposition 2.14, then we say that μ and ν are *stochastically ordered* and we write[2] $\mu \leq \nu$.

Lemma 2.15 (Upper invariant law) *Let $\overline{\nu}$ be the upper invariant law of the contact process and let ν be any other invariant law. Then $\nu \leq \overline{\nu}$ in the stochastic order.*

[2]This notation may look a bit confusing at first sight, since, if μ, ν are probability measures on any measurable space $(\Omega, \mathcal{F})$, then one might interpret $\mu \leq \nu$ in a pointwise sense, i.e., in the sense that $\mu(A) \leq \nu(A)$ for all $A \in \mathcal{F}$. In practice, this does not lead to confusion, since pointwise inequality for probability measures is a very uninteresting property. Indeed, it is easy to check that probability measures μ, ν satisfy $\mu \leq \nu$ in a pointwise sense if and only if $\mu = \nu$.

Proof Let A be a random variable, taking values in $\mathcal{P}(\mathbb{Z}^d)$, with law $\mathbb{P}[A \in \cdot\,] = \nu$, and assume that A is independent of the Poisson point processes used in our graphical representation. Then, since ν is an invariant law, we have $\nu = \mathbb{P}[\eta^A_t \in \cdot\,]$ for all $t \geq 0$. Since the random variables

$$(\eta^A_t, \overline{\eta}_t),$$

take values in the compact space $\mathcal{P}(\mathbb{Z}^d)^2$, their laws are automatically tight, hence we can select a subsequence $t_n \to \infty$ such that the $(\eta^A_{t_n}, \overline{\eta}_{t_n})$ converge weakly in law to some limiting random variable (η^1, η^2), say, where η^1 has the law ν, η^2 has the law $\overline{\nu}$, and moreover

$$\begin{aligned}
\mathbb{P}[i \in \eta^1,\ i \notin \eta^2] &= \lim_{n\to\infty} \mathbb{P}[i \in \eta^A_{t_n},\ i \notin \overline{\eta}_{t_n}] = \lim_{n\to\infty} \mathbb{P}[i \in \eta^{A,-t_n}_0,\ i \notin \overline{\eta}_0] \\
&\leq \lim_{n\to\infty} \mathbb{P}[\mathbb{Z}^d \times \{-t_n\} \rightsquigarrow (i,0),\ -\infty \not\rightsquigarrow (i,0)] = 0 \qquad (i \in \mathbb{Z}^d),
\end{aligned}$$

where we have used that the events $\mathbb{Z}^d \times \{-t_n\} \rightsquigarrow (i,0)$ decrease monotonically to the event $-\infty \rightsquigarrow (i,0)$, hence the events $\mathbb{Z}^d \times \{-t_n\} \rightsquigarrow (i,0),\ -\infty \not\rightsquigarrow (i,0)$ decrease monotonically to the empty set. We conclude that $\eta^1 \subset \eta^2$ a.s., hence $\nu \leq \overline{\nu}$ in the stochastic order. ∎

By definition, we say that a probability law μ on $\mathcal{P}(\mathbb{Z}^d)$ is *nontrivial* if

$$\mu(\{\emptyset\}) = 0,$$

i.e., if μ gives zero probability to the configuration in which all sites are healthy.

Lemma 2.16 (Survival and the upper invariant law) *For the contact process on $\mathbb{Z}^d$ with infection rate $\lambda \geq 0$, the following statements are equivalent:*

(i) *The contact process survives, i.e., $\theta(\lambda, d) > 0$.*

(ii) *The upper invariant law $\overline{\nu}$ is nontrivial.*

(iii) *There exists a nontrivial invariant law.*

Moreover, if the contact process dies out, then $\overline{\nu} = \delta_\emptyset$.

Proof The implication (ii)⇒(iii) is trivial and (iii)⇒(ii) follows from Lemma 2.15. To see that (i) and (ii) are equivalent, we start by observing that by duality, for each finite $B \subset \mathbb{Z}^d$

$$\begin{aligned}
&\int \overline{\nu}(\mathrm{d}A) 1_{\{A\cap B\neq\emptyset\}} = \mathbb{P}[\overline{\eta}_0 \cap B \neq \emptyset] \\
&\quad = \mathbb{P}[\eta^{\dagger\,B,0}_t \neq \emptyset\ \forall t \geq 0] = \mathbb{P}[\eta^B_t \neq \emptyset\ \forall t \geq 0].
\end{aligned} \tag{2.28}$$

Note that by Lemma 2.2, this formula determines the law $\overline{\nu}$ uniquely. In particular, we see that $\overline{\nu} = \delta_\emptyset$ if the contact process dies out. On the other hand, if the contact process survives, then $\overline{\nu} \neq \delta_\emptyset$. This is not quite the same as saying that $\overline{\nu}$ is nontrivial, but at least it tells us that $\mathbb{P}[\overline{\eta}_0 \neq \emptyset] > 0$. We observe that

$$\mathbb{P}\big[\overline{\eta}_t \neq \emptyset\big] = \mathbb{P}\big[\overline{\eta}_t \neq \emptyset \,\big|\, \overline{\eta}_0 \neq \emptyset\big]\mathbb{P}\big[\overline{\eta}_0 \neq \emptyset\big],$$

which by the stationarity of $\overline{\eta}$ implies that

$$\mathbb{P}\big[\overline{\eta}_t \neq \emptyset \,\big|\, \overline{\eta}_0 \neq \emptyset\big] = 1 \qquad (t \geq 0).$$

It follows that the conditioned law

$$\overline{\nu}(\,\cdot\,|\{A : A \neq \emptyset\})$$

is a nontrivial invariant law for the contact process, hence by the equivalence of (ii) and (iii), we must have that $\overline{\nu}$ is nontrivial. ∎

2.8 Ergodic behavior

We define translation operators $T_i : \mathcal{P}(\mathbb{Z}^d) \to \mathcal{P}(\mathbb{Z}^d)$ by

$$T_i(A) := \{j + i : j \in A\} \qquad (i \in \mathbb{Z}^d).$$

Below, we will sometimes also use the notation

$$T_i(A) = i + A.$$

We say that a probability law μ on $\mathcal{P}(\mathbb{Z}^d)$ is *homogeneous* or *translation invariant* if $\mu \circ T_i^{-1} = \mu$ for all $i \in \mathbb{Z}^d$. A lot of work in the theory of interacting particle systems is concerned with classifying all invariant laws of a given system, and proving that the system started from certain initial laws converges in law to a certain invariant law. In the context of interacting particle systems, invariant laws are sometimes also called *equilibria* or *equilibrium laws*. If an interacting particle system has a unique invariant law, which is the limit law of the process started in any initial state, then it is often said that the system is *ergodic*.[3]

[3]This is not very good terminology since it may lead to confusion with another, more usual concept of ergodicity. If $(X_t)_{t\in\mathbb{R}}$ is a stationary process, for example an interacting particle system in equilibrium, then by definition $(X_t)_{t\in\mathbb{R}}$ is *ergodic* if the law of $(X_t)_{t\in\mathbb{R}}$ gives probability zero or one to all events that are invariant under time shifts. In fact, if a Markov process is ergodic as defined in the text above, then the corresponding stationary process is ergodic in the sense defined here, but the converse does not hold in general.

The main aim of the present section is to prove the following result. It seems this result is due to Harris [Har76], who in particular seems to have invented the use of Hölder's inequality in the proof of Lemma 2.19. I believe an earlier version of Theorem 2.17, for a somewhat more limited class of initial laws, was proved by a Russian mathematician, but I forgot who.

Theorem 2.17 (Convergence to upper invariant law) *Let $(\eta_t)_{t\geq 0}$ be a contact process started in a homogeneus nontrivial initial law $\mathbb{P}[\eta_0 \in \cdot\,]$. Then*

$$\mathbb{P}[\eta_t \in \cdot\,] \underset{t\to\infty}{\Longrightarrow} \overline{\nu},$$

where $\overline{\nu}$ is the upper invariant law.

We start with two preparatory lemmas.

Lemma 2.18 (Extinction versus unbounded growth) *For each finite $A \subset \mathbb{Z}^d$, one has*

$$\eta^A_t = \emptyset \textit{ for some } t \geq 0 \quad \textit{or} \quad |\eta^A_t| \underset{t\to\infty}{\longrightarrow} \infty \quad \text{a.s.} \tag{2.29}$$

Proof Define

$$\rho(A) := \mathbb{P}\big[\eta^A_t \neq \emptyset\ \forall t \geq 0\big] \qquad (A \subset \mathbb{Z}^2,\ |A| < \infty).$$

It is not hard to see that for each $N \geq 0$ there exists an $\varepsilon > 0$ such that

$$|A| \leq N \quad \text{implies} \quad \rho(A) \leq 1 - \varepsilon. \tag{2.30}$$

We first argue why it is plausible that this implies (2.29) and then give a rigorous proof. Imagine that $|\eta^A_t| \not\to \infty$. Then, in view of (2.30), the process infinitely often gets a chance of at least ε to die out, hence eventually it should die out.
To make this rigorous, let

$$\mathcal{A}_A := \{\eta^A_t \neq \emptyset\ \forall t \geq 0\} \qquad (A \subset \mathbb{Z}^2,\ |A| < \infty).$$

denote the event that the process $(\eta^A_t)_{t\geq 0}$ survives and let $\mathcal{F}_t$ be the σ-field generated by the Poisson point processes used in our graphical representation till time t. Then

$$\rho(\eta^A_t) = \mathbb{P}\big[\mathcal{A}_A \,\big|\, \mathcal{F}_t\big] \underset{t\to\infty}{\longrightarrow} 1_{\mathcal{A}_A} \quad \text{a.s.}, \tag{2.31}$$

where we have used an elementary result from probability theory which says that if $\mathcal{F}_n$ is an increasing sequence of σ-fields and $\mathcal{F}_\infty = \sigma(\bigcup_n \mathcal{F}_n)$, then $\lim_n \mathbb{P}[A|\mathcal{F}_n] = \mathbb{P}[A|\mathcal{F}_\infty]$ a.s. for each measurable event A. (See [Loe63, § 29, Complement 10 (b)].) In view of (2.30), formula (2.31) implies (2.29). ∎

Lemma 2.19 (Nonzero intersection) *Let $(\eta_t)_{t\geq 0}$ be a contact process started in a homogeneus nontrivial initial law $\mathbb{P}[\eta_0 \in \cdot\,]$. Then for each $s, \varepsilon > 0$ there exists an $N \geq 1$ such that for any subset $A \subset \mathbb{Z}^d$*

$$|A| \geq N \quad \textit{implies} \quad \mathbb{P}\big[A \cap \eta_s \neq \emptyset\big] \geq 1 - \varepsilon.$$

Proof By duality (Lemma 2.1)

$$\mathbb{P}\big[A \cap \eta_s \neq \emptyset\big] = \mathbb{P}\big[\eta^A_s \cap \eta_0 \neq \emptyset\big]$$

where η_0 is independent of the graphical representation used to define η^A_s. Set $\Lambda_M := \{-M, \ldots, M\}^d$. It is not hard to see that each set $A \subset \mathbb{Z}^d$ with $|A| \geq N$ contains a subset $A' \subset A$ with $|A'| \geq N/|\Lambda_M|$ such that the sets

$$\big\{i + \Lambda_M : i \in A'\big\}$$

are disjoint, where as before we define $i + \Lambda_M := \{i + j : j \in \Lambda_M\}$. Write $\rightsquigarrow_{i+\Lambda_M}$ to indicate the presence of an open path that stays in $i + \Lambda_M$ and set

$$\eta^{\{i\}\,(M)}_s := \big\{j \in \mathbb{Z}^d : (i, 0) \rightsquigarrow_{i+\Lambda_M} (j, s)\big\}.$$

Then, using Hölder's inequality[4] in the inequality marked with an exclamation mark, we have

$$\begin{aligned}
\mathbb{P}\big[\eta^A_s \cap \eta_0 = \emptyset\big] &= \int \mathbb{P}[\eta_0 \in \mathrm{d}B]\mathbb{P}\big[\eta^A_s \cap B = \emptyset\big] \\
&\leq \int \mathbb{P}[\eta_0 \in \mathrm{d}B]\mathbb{P}\big[\bigcup_{i\in A'} \eta^{\{i\}\,(M)}_s \cap B = \emptyset\big] \\
&= \int \mathbb{P}[\eta_0 \in \mathrm{d}B] \prod_{i\in A'} \mathbb{P}\big[\eta^{\{i\}\,(M)}_s \cap B = \emptyset\big] \\
&\overset{!}{\leq} \prod_{i\in A'} \Big(\int \mathbb{P}[\eta_0 \in \mathrm{d}B]\mathbb{P}\big[\eta^{\{i\}\,(M)}_s \cap B = \emptyset\big]^{|A'|}\Big)^{1/|A'|} \\
&= \prod_{i\in A'} \Big(\int \mathbb{P}[\eta_0 \in \mathrm{d}B]\mathbb{P}\big[\eta^{\{0\}\,(M)}_s \cap B = \emptyset\big]^{|A'|}\Big)^{1/|A'|} \\
&= \int \mathbb{P}[\eta_0 \in \mathrm{d}B]\mathbb{P}\big[\eta^{\{0\}\,(M)}_s \cap B = \emptyset\big]^{|A'|},
\end{aligned}$$

where we have used the homogeneity of $\mathbb{P}[\eta_0 \in \cdot\,]$ in the one but last equality. Our arguments so far show that $|A| \geq N$ implies that

$$\mathbb{P}\big[A \cap \eta_s = \emptyset\big] \leq \int \mathbb{P}[\eta_0 \in \mathrm{d}B]\mathbb{P}\big[\eta^{\{0\}\,(M)}_s \cap B = \emptyset\big]^{N/|\Lambda_M|} =: f(N, M).$$

[4] Recall that Hölder's inequality says that $1/p + 1/q = 1$ implies $\|fg\|_1 \leq \|f\|_p \|g\|_q$, where $\|f\|_p := (\int |f|^p \mathrm{d}\mu)^{1/p}$. By induction, this gives $\|\prod_{i=1}^n f_i\|_1 \leq \prod_{i=1}^n \|f_i\|_n$.

Here, using the fact that

$$\mathbb{P}\big[\eta_s^{\{0\}\,(M)} \cap B = \emptyset\big] < 1 \quad \text{if } B \cap \Lambda_M \neq \emptyset,$$

we see that

$$\lim_{N\uparrow\infty} f(N,M) = \int \mathbb{P}[\eta_0 \in \mathrm{d}B] 1_{\{B\cap\Lambda_M=\emptyset\}} = \mathbb{P}[\eta_0 \cap \Lambda_M = \emptyset].$$

Since $\mathbb{P}[\eta_0 \in \cdot\,]$ is nontrivial, we have moreover

$$\lim_{M\uparrow\infty} \mathbb{P}[\eta_0 \cap \Lambda_M = \emptyset] = \mathbb{P}[\eta_0 = \emptyset] = 0.$$

Thus, we have shown that

$$\lim_{M\to\infty} \lim_{N\to\infty} f(N,M) = 0.$$

By a diagonal argument, for each $\varepsilon > 0$ we can choose N and M_N such that $f(N, M_N) \leq \varepsilon$, proving our claim. ∎

Exercise 2.20 Show by counterexample that the statement of Lemma 2.19 is false for $s = 0$.

Proof of Theorem 2.17 Since the space $\mathcal{P}(\mathbb{Z}^d)$ is compact, the laws of the η_t with $t \geq 0$ are tight, hence by Lemma 1.1 it suffices to prove that $\overline{\nu}$ is the only weak cluster point. By Lemma 2.2 and formula (2.28), it suffices to show that

$$\lim_{t\to\infty} \mathbb{P}\big[A \cap \eta_t \neq \emptyset\big] = \mathbb{P}\big[A \cap \overline{\eta}_0 \neq \emptyset\big] = \mathbb{P}\big[\eta_u^A \neq \emptyset\ \forall u \geq 0\big] =: \rho(A)$$

for all finite $A \subset \mathbb{Z}^d$. By duality (Lemma 2.1), this is equivalent to showing that

$$\lim_{t\to\infty} \mathbb{P}\big[\eta_{t-s}^A \cap \eta_s \neq \emptyset\big] = \rho(A) \qquad \big(A \subset \mathbb{Z}^d,\ |A| < \infty\big),$$

where $(\eta_t^A)_{t\geq 0}$ and $(\eta_t)_{t\geq 0}$ are independent and $s > 0$ is some fixed constant. For each $\varepsilon > 0$, we can choose N as in Lemma 2.19, and write

$$\begin{aligned}\mathbb{P}\big[\eta_t^A \cap \eta_s \neq \emptyset\big] = {}& \mathbb{P}\big[\eta_t^A \cap \eta_s \neq \emptyset \,\big|\, |\eta_t^A| = 0\big]\, \mathbb{P}\big[|\eta_t^A| = 0\big] \\ &+ \mathbb{P}\big[\eta_t^A \cap \eta_s \neq \emptyset \,\big|\, 0 < |\eta_t^A| < N\big]\, \mathbb{P}\big[0 < |\eta_t^A| < N\big] \\ &+ \mathbb{P}\big[\eta_t^A \cap \eta_s \neq \emptyset \,\big|\, |\eta_t^A| \geq N\big]\, \mathbb{P}\big[|\eta_t^A| \geq N\big].\end{aligned}$$

Here, by Lemma 2.18 and our choice of N,

$$
\begin{aligned}
&\text{(i)} \quad \mathbb{P}\big[\eta_t^A \cap \eta_s \neq \emptyset \,\big|\, |\eta_t^A| = 0\big] = 0,\\
&\text{(ii)} \quad \lim_{t\to\infty} \mathbb{P}\big[0 < |\eta_t^A| < N\big] = 0,\\
&\text{(iii)} \quad \liminf_{t\to\infty} \mathbb{P}\big[\eta_t^A \cap \eta_s \neq \emptyset \,\big|\, |\eta_t^A| \geq N\big] \geq 1-\varepsilon,\\
&\text{(iv)} \quad \lim_{t\to\infty} \mathbb{P}\big[|\eta_t^A| \geq N\big] = \rho(A),
\end{aligned}
$$

from which we conclude that

$$
(1-\varepsilon)\rho(A) \leq \liminf_{t\to\infty} \mathbb{P}\big[\eta_t^A \cap \eta_s \neq \emptyset\big] \leq \limsup_{t\to\infty} \mathbb{P}\big[\eta_t^A \cap \eta_s \neq \emptyset\big] \leq \rho(A).
$$

Since $\varepsilon > 0$ is arbitrary, our proof is complete. ∎

Theorem 2.17 has a simple corollary.

Corollary 2.21 (Homogeneous invariant laws) *All homogeneous invariant laws of a contact process are convex combinations of $\delta_\emptyset$ and $\overline{\nu}$.*

Proof If the contact process dies out, then $\overline{\nu} = \delta_\emptyset$ is the largest invariant law with respect to the stochastic order, hence $\delta_\emptyset$ is the only invariant law and the statement is trivially true. On the other hand, if the contact process survives, then $\overline{\nu}$ is nontrivial (recall Lemma 2.16). Moreover, by Theorem 2.17, if μ is a nontrivial homogeneous invariant law and $(\eta_t)_{t\geq 0}$ is a contact process started in the initial law μ, then

$$
\mu = \mathbb{P}[\eta_t \in \cdot\,] \underset{t\to\infty}{\Longrightarrow} \overline{\nu}.
$$

In particular, this should hold if μ is an invariant law, hence $\overline{\nu}$ is the only nontrivial homogeneous invariant law. We recall from the proof of Lemma 2.16 that if μ is any homogeneous invariant law, then we may write

$$
\mu = \mu(\{\emptyset\})\delta_\emptyset + \big(1 - \mu(\{\emptyset\})\big)\mu(\,\cdot\,|\{A : A \neq \emptyset\})
$$

where $\mu(\,\cdot\,|\{A : A \neq \emptyset\})$ is a nontrivial homogeneous invariant law. From this we see that all homogeneous invariant laws are convex combinations of $\delta_\emptyset$ and $\overline{\nu}$. ∎

Recall from Proposition 2.10 that the function $\lambda \mapsto \theta(\lambda)$ is right-continuous everywhere. As an application of Theorem 2.17, we prove the following result.

Proposition 2.22 (Left-continuity) *The function $\lambda \mapsto \theta(\lambda)$ is left-continuous on $(\lambda_{\rm c}, \infty)$.*

We first prove two preparatory lemmas. Let $\mathcal{C}_{\text{fin}}(\mathcal{P}(\mathbb{Z}^d))$ denote the space of real functions on $\mathcal{P}(\mathbb{Z}^d)$ that depend on finitely many coordinates, as defined in Section 1.7.

Lemma 2.23 (Convergence of semigroups) *Let $(P^\lambda_t)_{t\geq 0}$ be the Markov semigroup of the contact process on $\mathbb{Z}^d$ with infection rate λ. Then*

$$\|P^{\lambda_n}_t f - P^\lambda_t f\| \underset{\lambda_n\to\lambda}{\longrightarrow} 0 \qquad \big(t,\lambda \geq 0,\ f\in\mathcal{C}_{\text{fin}}(\{0,1\}^{\mathbb{Z}^d})\big).$$

where $\|\cdot\|$ denotes the supremumnorm.

Proof We use the coupling from the proof of Proposition 2.10 and set

$$\eta^{A,\lambda}_t := \big\{i : A\times\{0\} \rightsquigarrow_\lambda (i,t)\big\}.$$

Let f be a function that depends only on the coordinates in a finite set $\Lambda\subset\mathbb{Z}^d$. Let $0\leq\lambda_n\to\lambda$ and choose some λ' such that $\lambda_n\leq\lambda'$ for all n. Let Γ be the collection of all infection arrows $((i,j),t,\kappa)\in\overline{\Delta}^{\rm i}$ that are used in some infection path along arrows with values $\kappa\leq\lambda'$ starting at time zero and ending somewhere in the finite set Λ. Then Γ contains all arrows that are relevant for deciding which points belong to the set $\Lambda\cap\eta^{A,\lambda}_t$. Let Γ_λ denote the set of arrows in Γ that have a value $\kappa\leq\lambda$. Since Γ is a.s. finite (by Lemma 2.8), there a.s. exists some random m such that $\Gamma_{\lambda_n}=\Gamma_\lambda$ for all $n\geq m$. It follows that

$$\begin{aligned}&|P^{\lambda_n}_t f(A) - P^\lambda_t f(A)| = \big|\mathbb{E}\big[f'(\Lambda\cap\eta^{A,\lambda_n}_t) - f'(\Lambda\cap\eta^{A,\lambda}_t)\big]\big|\\ &\quad\leq 2\|f\|\,\mathbb{P}[\Gamma_{\lambda_n}\neq\Gamma_\lambda] \underset{n\to\infty}{\longrightarrow} 0,\end{aligned}$$

and this convergence is *uniform* in A, as claimed. ∎

Lemma 2.24 (Convergence of invariant laws) *Let ν_n,ν be probability laws on $\mathcal{P}(\mathbb{Z}^d)$ such that $\nu_n\Rightarrow\nu$ and let $0\leq\lambda_n\to\lambda$. Assume that ν_n is an invariant law for the contact process with infection rate λ_n, for each n. Then ν is an invariant law for the contact process with infection rate λ.*

Proof We introduce the notation

$$\mu f := \int \mu(\mathrm{d}x) f(x).$$

With this notation, if $(X_t)_{t\geq 0}$ is a Markov process with Markov semigroup $(P_t)_{t\geq 0}$, started in the initial law $\mathbb{P}[X_0\in\cdot\,]=\mu$, then

$$\mu P_t f = \int \mathbb{P}[X_0\in\mathrm{d}x]\int P_t(x,\mathrm{d}y)f(y) = \mathbb{E}[f(X_t)].$$

We write

$$|\nu P_t^\lambda f - \nu f| \leq |\nu P_t^\lambda f - \nu_n P_t^\lambda f| + |\nu_n P_t^\lambda f - \nu_n P_t^{\lambda_n} f| + |\nu_n P_t^{\lambda_n} f - \nu_n f| + |\nu_n f - \nu f|$$

where of course $|\nu_n P_t^{\lambda_n} f - \nu_n f| = 0$ since ν_n is an invariant law for the process with infection rate λ_n. It follows from the Feller property of the contact process (which is a consequence of Theorem 1.15) that $\mathbb{P}_t^\lambda$ maps continuous functions into continuous functions, hence by our assumption that $\nu_n \Rightarrow \nu$ we have

$$|\nu_n P_t^\lambda f - \nu P_t^\lambda f| \underset{n\to\infty}{\longrightarrow} 0 \quad \text{and} \quad |\nu_n f - \nu f| \underset{n\to\infty}{\longrightarrow} 0$$

for each $f \in \mathcal{C}(\mathcal{P}(\mathbb{Z}^d))$ and $t \geq 0$. Assuming that moreover $f \in \mathcal{C}_{\rm fin}(\mathcal{P}(\mathbb{Z}^d))$, we have by Lemma 2.23 that

$$|\nu_n P_t^\lambda f - \nu_n P_t^{\lambda_n} f| \leq \|P_t^\lambda f - P_t^{\lambda_n} f\| \underset{n\to\infty}{\longrightarrow} 0.$$

It follows that

$$\nu P_t^\lambda f = \nu f \qquad \big(t \geq 0,\ f \in \mathcal{C}_{\rm fin}(\mathcal{P}(\mathbb{Z}^d))\big),$$

hence ν is an invariant law for the contact process with infection rate λ. ∎

Proof of Proposition 2.22 Let $\overline{\nu}_\lambda$ denote the upper invariant law of the contact process with infection rate λ. Choose $\lambda_{\rm c} < \lambda_n \uparrow \lambda$. Since the space $\{0,1\}^{\mathbb{Z}^d}$ is compact, the measures $\overline{\nu}_{\lambda_n}$ are tight. By Lemma 2.24, each weak cluster point of the $\overline{\nu}_{\lambda_n}$ is a homogeneous invariant law of the contact process with infection rate λ. Since each $\overline{\nu}_{\lambda_n}$ is the law of a process $\overline{\eta}_t$ as defined in Section 2.7, we see in the same way as in the proof of Lemma 2.7 that the laws $\overline{\nu}_{\lambda_n}$ are increasing in the stochastic order, hence each weak cluster point of the $\overline{\nu}_{\lambda_n}$ is nontrivial. By Corollary 2.21, it follows that each weak cluster point must equal $\overline{\nu}_\lambda$, hence we conclude that

$$\overline{\nu}_{\lambda_n} \underset{n\to\infty}{\Longrightarrow} \overline{\nu}_\lambda.$$

Since by (2.28),

$$\theta(\lambda) = \int \overline{\nu}_\lambda(\mathrm{d}A) 1_{\{0\in A\}},$$

this implies that $\theta(\lambda_n) \to \theta(\lambda)$. ∎

2.9 Other topics

Corollary 2.21 tells us that all homogeneous invariant laws of a contact process are convex combinations of $\delta_\emptyset$ and the upper invariant law. One may wonder if there exist inhomogeneous invariant laws. The answer to this question is known to be negative. This follows from the following theorem, that strengthens Theorem 2.17 quite a bit:

Theorem 2.25 (Complete convergence) *The contact process started in any initial state satisfies*

$$\mathbb{P}[\eta_t^A \in \cdot\,] \underset{t\to\infty}{\Longrightarrow} \rho(A)\overline{\nu} + (1-\rho(A))\delta_\emptyset,$$

where $\rho(A) := \mathbb{P}\big[\eta_t^A \neq \emptyset\ \forall t \geq 0\big]$.

Complete convergence was proved first only for λ sufficiently large. In [BG90], this was extended to arbitrary $\lambda \geq 0$. In fact, more is known: it is known that if the process survives, then the infected area grows approximately linear and has a deterministic limiting shape. This result is known as the *shape theorem*.
The proof of complete convergence is quite a bit more involved than the proof of Theorem 2.17. To understand why this is so, it is useful to generalize a bit and consider contact processes on more general lattices, e.g., infinite graphs. As long as the graph has some sort of translation invariant structure, Theorem 2.17 still holds (and the proof basically carries through without a change). However, complete convergence does not hold in this generality. In particular, for processes on trees, it is known that there exist two critical values $\lambda_c < \lambda_c'$ such that in the intermediate regime complete convergence does not hold and there exist inhomogeneous invariant laws. The study of contact processes on more general lattices is quite a lively modern subject with several nice open problems.

Chapter 3

The Ising model

3.1 Introduction

In this chapter, we study the Ising model. The Ising model model has been introduced by E. Ising in 1925 [Isi25] as a simple model for a ferromagnetic material, based on the theory of Gibbs measures, which dated from the late nineteenth century when people like Boltzmann tried to find a microscopic basis for the laws of thermodynamics that had been discovered earlier in that century. In his Phd thesis, Ising showed that the one-dimensional model that now bears his name does not exhibit a phase transition, and based on this he incorrectly concluded that the same is true in any dimension. In 1936, Peierls [Pei36] used his famous argument (a variation on which we have already seen in Chapter 2) to prove this conjecture wrong in dimensions two and more. In 1944, Onsager showed that the two-dimensional model can, in a certain sense, be solved explicitly [Ons44]. (No explicit solutions are known or believed to exist in dimensions three and more.)

The Ising model as such, it should be pointed out, is not an interating particle system. Rather, it is a certain probability law (Gibbs measure) on spin configurations, depending on a certain parameter related to the temperature of the system. It is possible, however, and physically meaningful, to construct interacting particle systems that have these Gibbs measures as invariant measures. Such interacting particle systems are called *stochastic Ising models*. The first one to contruct stochastic Ising models was Glauber [Gla63]. The subject was taken up again and studied more profoundly by Dobrushin in a series of papers starting with [Dob71]. Using the 'interacting particle systems approach', it is possible to give nice short proofs of certain properties of the Ising model. Conversely, the Ising model gives in a natural way rise to a number of interesting interacting particle systems which have sufficiently many pleasant properties to make it possible to prove things about them, while on the other hand they are sufficiently 'difficult' to be interesting.

3.2 Definition, construction, and ergodicity

For definiteness, we will introduce one stochastic Ising model, i.e., an interacting particle system that has the Gibbs measures of the Ising model as its invariant law(s), that we will mostly focus on. As we will see later, there exist several ways to invent a dynamics for the Ising model, and many things that we will prove for our specific model are valid more generally.
The model that we will mostly focus our attention on is the interacting particle system with the following description. At each site $i \in \mathbb{Z}^d$, there is an atom which has a property called *spin* which makes it act like a small magnet that can either point up, in which case we say the site i is in the state $+1$, or down, in which case we say the site i is in the state -1. Our stochastic Ising model is therefore a Markov process $(X_t)_{t\geq 0}$ with state space $\{-1,1\}^{\mathbb{Z}^d}$. We will consider the following dynamics: if the process is in a state $x = (x(i))_{i\in\mathbb{Z}^d} \in \{-1,1\}^{\mathbb{Z}^d}$, then

$$
\begin{array}{ll}
x(i) \text{ jumps:} & \\
-1 \mapsto 1 & \text{with rate } e^{-\beta \sum_{j:\,|i-j|=1} 1_{\{x(j)=-1\}}}, \\
1 \mapsto -1 & \text{with rate } e^{-\beta \sum_{j:\,|i-j|=1} 1_{\{x(j)=1\}}}.
\end{array}
$$

Here $\beta > 0$ is a parameter (loosely) called the *inverse temperature*. Indeed, in the physical interpretation of the model, $\beta = J/kT$ where T is the temperature, J is the energy difference between aligned and unaligned neighboring spins, and k is Boltzmann's constant. The motivation for our dynamics is roughly as follows: due to the constant motion of atoms, spins tend to *flip* in a random way between the $+1$ and -1 state. However, because of the magnetic interaction between neighboring atoms, neighboring spins like to be aligned (i.e., point in the same direction). This is expressed by making a spin less likely to flip when it has a lot of neighbors that point in the same direction. This effect is stronger when β is large (i.e., when the temperature is low). Note that (contrary to what we saw for the contact process) our dynamics treat the two values $-1, +1$ for the spins symmetrically.

In order to construct our process rigorously, we use a graphical representation. We first write down the formal generator of our process, which is

$$
Gf(x) := \sum_{i\in\mathbb{Z}^d} e^{-\beta \sum_{j\in\mathcal{N}_i} 1_{\{x(j)=x(i)\}}} \big(f(x^{\{i\}}) - f(x)\big) \qquad (x \in \{-1,1\}^{\mathbb{Z}^d}), \tag{3.1}
$$

where in analogy with (2.2), we define $x^{\{i\}}$ by

$$x^{\{i\}}(j) := \begin{cases} -x(j) & \text{if } i = j, \\ x(j) & \text{otherwise,} \end{cases}$$

and

$$\mathcal{N}_i := \{j \in \mathbb{Z}^d : |i - j| = 1\} \qquad (i \in \mathbb{Z}^d)$$

denotes the set of neighbors of a site i. To invent a graphical representation, we need to rewrite our generator in terms of local maps. For each $i \in \mathbb{Z}^d$ and subset $L \subset \mathcal{N}_i$, let us define the maps $m^-_{i,L}, m^+_{i,L}$ by

$$(m^-_{i,L}x)(k) := \begin{cases} -1 & \text{if } k = i,\ x(j) = -1\ \forall j \in L, \\ x(k) & \text{otherwise,} \end{cases}$$

$$(m^+_{i,L}x)(k) := \begin{cases} +1 & \text{if } k = i,\ x(j) = +1\ \forall j \in L, \\ x(k) & \text{otherwise,} \end{cases}$$

Then we may write our generator in the form

$$\begin{aligned} Gf(x) := &\sum_{i \in \mathbb{Z}^d} \sum_{L \subset \mathcal{N}_i} p^{|L|}(1-p)^{2d-|L|}\big(f(m^-_{i,L}x) - f(x)\big) \\ &+ \sum_{i \in \mathbb{Z}^d} \sum_{L \subset \mathcal{N}_i} p^{|L|}(1-p)^{2d-|L|}\big(f(m^+_{i,L}x) - f(x)\big), \end{aligned}$$

where

$$p := 1 - e^{-\beta}.$$

To see why this is correct, note that according to our new formulation of the generator, the spin at site i flips from -1 to $+1$ at rate

$$\sum_{L \subset \mathcal{N}_i} p^{|L|}(1-p)^{2d-|L|} 1_{\{x(j)=+1\ \forall j \in L\}}. \tag{3.2}$$

Let $\mathcal{L}$ be a random subset of $\mathcal{N}_i$ such that independently for each neighbor j of i, one has $\mathbb{P}[j \in \mathcal{L}] = p$. Then the rate in (3.2) may be rewritten as

$$\begin{aligned} &\mathbb{P}[x(j) = +1\ \forall j \in \mathcal{L}] = \textstyle\prod_{j \in \mathcal{N}_i :\, x(j)=-1} \mathbb{P}[j \notin \mathcal{L}] \\ &\quad = (1-p)^{\sum_{j \in \mathcal{N}_i} 1_{\{x(j)=-1\}}} = e^{-\beta \sum_{j \in \mathcal{N}_i} 1_{\{x(j)=-1\}}}, \end{aligned}$$

as required. By symmetry, a similar argument holds for flips from $+1$ to -1.

Using these observations, we can define a graphical representation for our process as follows. Let $\mathcal{H}$ be the space of all triples of the form

$$(\sigma, i, L) \quad \text{with } \sigma \in \{-,+\},\ i \in \mathbb{Z}^d,\ L \subset \mathcal{N}_i,$$

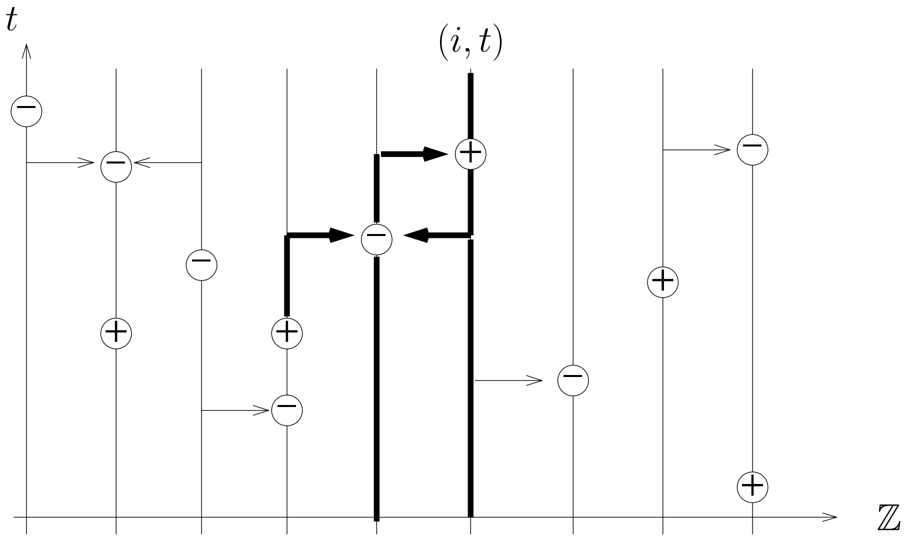

Figure 3.1: Graphical representation of our stochastic Ising model.

and let Δ be a Poisson point process on $\mathcal{H} \times \mathbb{R}$ with intensity $p^{|L|}(1-p)^{2d-|L|}\mathrm{d}t$. We interpret a point (σ, i, L, t) as saying that at time t, the state of of system changes according to the local map $m^{\sigma}_{i,L}$. To draw this in a picture, for each point $(\sigma, i, L, t) \in \Delta$, we draw a circle at the point $(i,t) \in \mathbb{Z}^d \times \mathbb{R}$ with the sign s in it, and we draw arrows starting at each point $j \in L$ and ending in i (see Figure 3.1). (Note that L may contain anything between zero and $2d$ elements.)

By Theorems 1.15 and 1.18, this graphical representation defines a Feller process with values in $\{-1,+1\}^{\mathbb{Z}^d}$ whose generator is the closure of the operator in (3.1). Recall the definition of path γ from Section 1.6 and of a *path of influence* in (1.21). We observe that $\mathcal{D}(m^{\pm}_{i,L}) = \{i\}$ and

$$\mathcal{R}_i(m^{\pm}_{i,L}) = \begin{cases} L \cup \{i\} & \text{if } L \neq \emptyset, \\ \emptyset & \text{if } L = \emptyset. \end{cases}$$

Therefore, a path γ is a path of influence if and only if

(i) $\forall t \in [s,u]$ with $\gamma_{t-} \neq \gamma_t$ $\exists (\sigma, i, L, t) \in \Delta$ s.t. $\gamma_{t-} \in L,\ \gamma_t = i$,

(ii) $\nexists (\sigma, i, L, t) \in \Delta$ s.t. $|L| = \emptyset,\ t \in [s,u],\ \gamma_t = i$.

In our picture, this says that a path may use arrows but must avoid points (i,t), marked with an $\ominus$ or $\oplus$ where no arrows come in. Note that at such points, the spin at site i flips to the state -1 or $+1$, regardless of the state of the system prior to time t. The constant K from (1.23) is given by

$$\begin{aligned} K &= 2\sum_{L\subset\mathcal{N}_i} p^{|L|}(1-p)^{2d-|L|}(|L|-1_{\{L=\emptyset\}}) \\ &= 2\big(\mathbb{E}\big[|\mathcal{L}|\big]-\mathbb{P}[\mathcal{L}=\emptyset]\big) = 2(2dp-(1-p)^{2d}), \end{aligned}$$

where as before $\mathcal{L}$ denotes a random subset of $\mathcal{N}_i$ such that $\mathbb{P}[j\in\mathcal{L}]=p$ independently for all $j\in\mathcal{N}_i$.

Since $K<0$ for β sufficiently small, we can draw an interesting conclusion.

Proposition 3.1 (Ergodicity for high temperature) *Let $\beta' := \sup\{\beta>0 : 2d(1-e^{-\beta})-e^{-2d\beta}<0\}$. Then, for each $\beta<\beta'$, our stochastic Ising model has a unique invariant measure ν and the process started from any initial law satisfies*

$$\mathbb{P}[X_t\in\cdot\,] \underset{t\to\infty}{\Longrightarrow} \nu. \tag{3.3}$$

Proof We observe that for each $\beta<\beta'$, the constant K from (1.23) is strictly negative. Write

$$\mathbb{P}[X_t\in\cdot\,] = \mathbb{P}\big[\Psi_{\Delta,-t,0}(X_0)\in\cdot\,\big],$$

where $\Psi_{\Delta,s,t}$ is defined as in Section 1.6. Define $\zeta^{\{i\},t}_s$ as in (1.22) and set

$$\sigma_{(i,t)} := \sup\{s\le t : \zeta^{\{i\},t}_s=\emptyset\}.$$

By Lemma 1.13 and the fact that $K<\infty$, we observe that $-\infty<\sigma_{(i,t)}$ for each $(i,t)\in\mathbb{Z}^d\times\mathbb{R}$. Fix some arbitrary $x\in\{-1,+1\}^{\mathbb{Z}^d}$ and define a process $(Y_t)_{t\in\mathbb{R}}$ with values in $\{-1,+1\}^{\mathbb{Z}^d}$ by

$$Y_t(i) := \Psi_{\Delta,\sigma_{(i,t)}-1,t}(x)(i) \qquad (t\in\mathbb{R},\ i\in\mathbb{Z}^d).$$

Note that this definition does not depend on the choice of x since no path of influence ending at (i,t) starts before time $\sigma_{(i,t)}$. Moreover, $(Y_t)_{t\in\mathbb{R}}$ is a stationary stochastic Ising model. Since

$$\Psi_{\Delta,-t,0}(X_0)(i) = Y_0(i) \qquad \forall -t<\sigma(i,0),$$

we see that $\Psi_{\Delta,-t,0}(X_0)$ converges pointwise to Y_0 as $t\to\infty$ and therefore the law of X_t converges weakly to the law of Y_0. ∎

Alternative proof of Proposition 3.1 (sketch) Since $K<0$ for all $\beta<\beta'$, it follows from Lemma 1.21 that

$$\sup_x P_tf(x)-\inf_x P_tf(x) = \sum_i \delta P_tf(i) \underset{t\to\infty}{\longrightarrow} 0 \qquad \big(f\in\mathcal{C}_{\rm sum}(\{-1,+1\}^{\mathbb{Z}^d})\big).$$

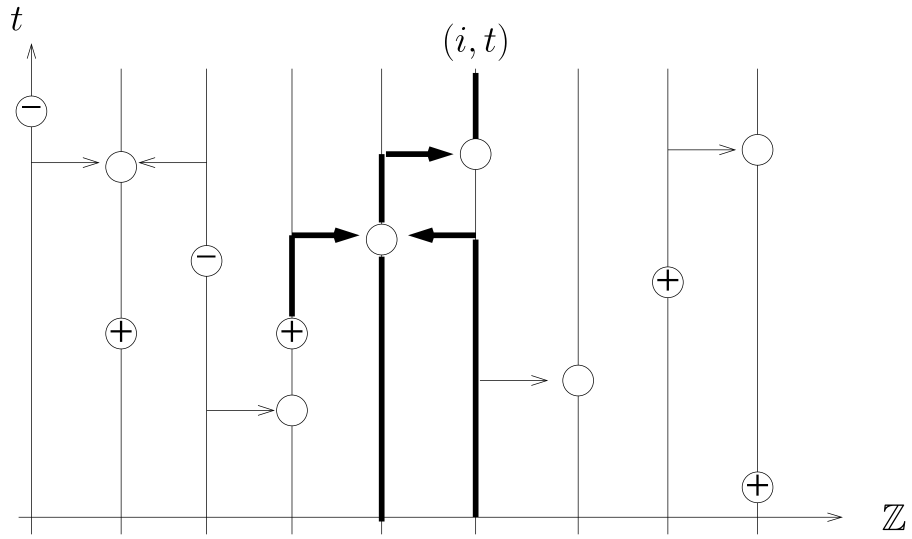

Figure 3.2: Alternative graphical representation of our stochastic Ising model.

Since $\sup_x P_t f(x)$ (resp. $\inf_x P_t f(x)$) is a nonincreasing (resp. nondecreasing) function of t, it follows that for each $f \in \mathcal{C}_{\rm sum}(\{-1,+1\}^{\mathbb{Z}^d})$, there exists a constant c_f such that

$$P_t f(x) \underset{t\to\infty}{\longrightarrow} c_f \qquad \big(x \in \{-1,+1\}^{\mathbb{Z}^d}\big).$$

Since the space $\{-1,+1\}^{\mathbb{Z}^d}$ is compact, for each $t_n \to \infty$, by going to a subsequence if necessary, we may assume that

$$\int \mathbb{P}[X_0 \in \mathrm{d}x] P_{t_n}(x,\,\cdot\,) \underset{n\to\infty}{\Longrightarrow} \nu$$

for some probability law ν. By what we have just proved,

$$\int \nu(\mathrm{d}x) f(x) = c_f \qquad \big(f \in \mathcal{C}_{\rm sum}(\{-1,+1\}^{\mathbb{Z}^d})\big).$$

Since $\mathcal{C}_{\rm sum}(\{-1,+1\}^{\mathbb{Z}^d})$ is dense in $\mathcal{C}(\{-1,+1\}^{\mathbb{Z}^d})$, it follows that all subsequential limits of the measures $\int \mathbb{P}[X_0 \in \mathrm{d}x] P_t(x,\,\cdot\,)$ are equal, and independent of the law of X_0. ∎

Exercise 3.2 (Alternative graphical representation) For each $i \in \mathbb{Z}^d$ let us define the maps

$$(m_i^- x)(k) := \begin{cases} -1 & \text{if } k = i, \\ x(k) & \text{otherwise}, \end{cases} \qquad (m_i^+ x)(k) := \begin{cases} +1 & \text{if } k = i, \\ x(k) & \text{otherwise}, \end{cases}$$

Moreover, for each $i \in \mathbb{Z}^d$ and nonempty subset $L \subset \mathcal{N}_i$, let us define

$$(m_{i,L}x)(k) := \begin{cases} -x(i) & \text{if } k = i,\ x(j) \neq x(i)\ \forall j \in L, \\ x(k) & \text{otherwise.} \end{cases}$$

Then we may rewrite our generator in the form

$$\begin{aligned} Gf(x) := {} & (1-p)^{2d} \sum_{i \in \mathbb{Z}^d} \big(f(m_i^- x) + f(m_i^+ x) - 2f(x) \big) \\ & + \sum_{i \in \mathbb{Z}^d} \sum_{\emptyset \neq L \subset \mathcal{N}_i} p^{|L|} (1-p)^{2d-|L|} \big(f(m_{i,L}x) - f(x) \big). \end{aligned}$$

Based on this, we may introduce an alternative graphical representation for our stochastic Ising model (see Figure 3.2). Use this to improve Proposition 3.1 by proving ergodicity for a larger range of the parameter.

3.3 Gibbs measures and finite systems

Let Λ be some finite set and let $H : \{-1,+1\}^\Lambda \to \mathbb{R}$ be some function. By definition, the *Gibbs measure* belonging to the *Hamiltonian* (or *energy function*) H and *inverse temperature* β is the probability measure on $\{-1,+1\}^\Lambda$ given by

$$\mu(\{x\}) = \frac{1}{Z} e^{-\beta H(x)} \qquad \big(x \in \{-1,+1\}^\Lambda \big), \tag{3.4}$$

where

$$Z := \sum_{x \in \{-1,+1\}^\Lambda} e^{-\beta H(x)} \tag{3.5}$$

is a normalization constant, also called the *partition sum*. Note that if H, H' are two energy functions that differ only by a constant, then the associated Gibbs measures are the same. Indeed, if $H(x) = H'(x) + c$ and μ, μ' are the associated Gibbs measures, then all probabilities in μ' get an extra factor $e^{-\beta c}$, but this disappears in the normalization. Indeed, we make the following simple observation.

Lemma 3.3 (Relative probabilities)
(a) *If Λ is a finite set and μ is the Gibbs measure on $\{-1,+1\}^\Lambda$ with Hamiltonian H and inverse temperature β, then*

$$\frac{\mu(\{x'\})}{\mu(\{x\})} = e^{-\beta(H(x') - H(x))} \qquad \big(x, x' \in \{-1,+1\}^\Lambda \big). \tag{3.6}$$

(b) *Conversely, if μ is a measure on $\{-1,+1\}^\Lambda$ and*

$$\frac{\mu(\{x^{\{i\}}\})}{\mu(\{x\})} = e^{-\beta(H(x^{\{i\}}) - H(x))} \qquad \big(i \in \Lambda,\ x \in \{-1,+1\}^\Lambda\big), \tag{3.7}$$

then μ must be the Gibbs measure on $\{-1,+1\}^\Lambda$ associated with H and β.

Proof Part (a) is trivial. To prove part (b), we note that for each $x, x' \in \{-1,+1\}^{\mathbb{Z}^d}$ we can find $x_0, \ldots, x_n$ such that $x = x_0$, $x' = x_n$, and x_k differs only in one point from x_{k-1} ($k = 1, \ldots, n$). In view of this, (3.7) implies (3.6). Choosing some arbitrary reference state x', we see that (3.6) determines all probabilities up to an overall multiplicative constant, which follows from the normalization. ∎

We need to introduce some notation. If S, R are disjoint sets, $x \in \{-1,+1\}^S$, and $y \in \{-1,+1\}^R$, then we define $x\&y \in \{-1,+1\}^{S\cup R}$ as $(x\&y)(i) := x(i)$ if $i \in S$ and $(x\&y)(i) := y(i)$ if $i \in R$. Now let Λ be a finite set, let $H : \{-1,+1\}^\Lambda \to \mathbb{R}$ be a function, and let $\mu^{\Lambda,\beta}$ be the Gibbs measure on $\{-1,+1\}^\Lambda$ with Hamiltonian H and inverse temperature β. For each $\Delta \subset \Lambda$ and $y \in \{-1,+1\}^{\Lambda\backslash\Delta}$, let $H^\Delta_y : \{-1,+1\}^\Delta \to \mathbb{R}$ be a function such that

$$H^\Delta_y(x) = H(x\&y) + c^\Delta_y \qquad \big(x \in \{-1,+1\}^\Delta\big),$$

where c^Δ_y is a constant that may depend on Δ and y but not on x. Let $\mu^{\Delta,\beta}_y$ be the Gibbs measure on $\{-1,+1\}^\Delta$ associated with H^Δ_y and β. (Note that this Gibbs measure is uniquely defined even though H^Δ_y is defined only up to a constant.) We make the following observations:

Lemma 3.4 (Conditional distributions)
(a) *If $(X(i))_{i\in\Lambda}$ is a random variable with law $\mu^{\Lambda,\beta}$, then for each $\Delta \subset \Lambda$, the conditional law of X inside Δ given its values outside Δ is given by*

$$\mathbb{P}\big[(X(i))_{i\in\Delta} \in \cdot \,\big|\, (X(i))_{i\in\Lambda\backslash\Delta} = y\big] = \mu^{\Delta,\beta}_y. \tag{3.8}$$

(b) *Conversely, if $(X(i))_{i\in\Lambda}$ is a random variable with values in $\{-1,+1\}^\Lambda$ and (3.8) holds for each $\Delta \subset \Lambda$ such that $|\Delta| = 1$, then the law of X must be equal to $\mu^{\Lambda,\beta}$.*

Proof We observe that

$$\begin{aligned}
&\frac{\mathbb{P}\big[(X(i))_{i\in\Delta} = x' \,\big|\, (X(i))_{i\in\Lambda\backslash\Delta} = y\big]}{\mathbb{P}\big[(X(i))_{i\in\Delta} = x \,\big|\, (X(i))_{i\in\Lambda\backslash\Delta} = y\big]} \\
&\quad = \frac{\mathbb{P}\big[(X(i))_{i\in\Delta} = x',\ (X(i))_{i\in\Lambda\backslash\Delta} = y\big]}{\mathbb{P}\big[(X(i))_{i\in\Delta} = x,\ (X(i))_{i\in\Lambda\backslash\Delta} = y\big]} = \frac{e^{-\beta H(x'\&y)}}{e^{-\beta H(x\&y)}} = e^{-\beta(H_y(x') - H_y(x))}.
\end{aligned}$$

In view of this, the statements follow from Lemma 3.3. ∎

The fact that we would like to prove is that Gibbs measures associated with the Hamiltonian [1]

$$H(x) := \sum_{\{i,j\}\in\mathcal{B}} 1_{\{x(i)\neq x(j)\}} \tag{3.9}$$

are reversible invariant measures for the stochastic Ising model constructed in the previous section. A "slight" problem with this statement is that the sum in this definition runs over the set

$$\mathcal{B} := \big\{\{i,j\} : i,j \in \mathbb{Z}^d,\ |i-j|\big\}$$

of all (unordered) nearest neighbor pairs in $\mathbb{Z}^d$. As a consequence, for most x, the sum in (3.9) is actually infinite. In addition, the set $\{-1,+1\}^{\mathbb{Z}^d}$ is uncountable, so it is clear that we cannot define Gibbs measures on $\{-1,+1\}^{\mathbb{Z}^d}$ in the same way as we have done for finite lattices.

The solution to this problem is suggested by Lemma 3.4. Instead of looking at the absolute probability of one particular configuration x (which will typically be zero), we will look at conditional probabilities of finding certain configurations inside a finite set $\Lambda\subset\mathbb{Z}^d$, given what is outside.

To this aim, for each $\Lambda\subset\mathbb{Z}^d$, we define

$$\partial\Lambda := \{i\in\mathbb{Z}^d\backslash\Lambda : \exists j\in\Lambda \text{ s.t. } |i-j|=1\}$$

and let

$$\mathcal{B}_\Lambda := \big\{\{i,j\} : i,j\in\Lambda,\ |i-j|\big\} \quad\text{and}\quad \partial\mathcal{B}_\Lambda := \big\{(i,j) : i\in\Lambda,\ j\in\partial\Lambda,\ |i-j|\big\}$$

denote the set of nearest-neighbor edges inside Λ and pointing out of Λ, respectively. For each finite $\Lambda\subset\mathbb{Z}^d$, $x\in\{-1,+1\}^\Lambda$ and $y\in\{-1,+1\}^{\mathbb{Z}^d\backslash\Lambda}$, we define

$$H^\Lambda_y(x) := \sum_{\{i,j\}\in\mathcal{B}_\Lambda} 1_{\{x(i)\neq x(j)\}} + \sum_{(i,j)\in\partial\mathcal{B}_\Lambda} 1_{\{x(i)\neq y(j)\}}.$$

We let $\mu^{\Lambda,\beta}_y$ denote the finite-volume Gibbs measure associated with H_y and β. We call this the finite-volume Gibbs measure with *boundary condition y*.

[1] Here I deviate from the usual definition of the Hamiltonian for the Ising model, which is

$$H'(x) := -\sum_{\{i,j\}\in\mathcal{B}} x(i)x(j) = \sum_{\{i,j\}\in\mathcal{B}} (21_{\{x(i)\neq x(j)\}}-1).$$

We observe that $H'(x) = 2H(x)+c$, where $c := |\mathcal{B}|$ is an irrelevant additive constant. In view of this, what is β in these lecture notes, is 2β in most of the literature on the Ising model.

Definition 3.5 (Infinite-volume Gibbs measures) *We say that the law μ of an $\{-1,+1\}^{\mathbb{Z}^d}$-valued random variable $(X(i))_{i\in\mathbb{Z}^d}$ is a Gibbs measure associated with the formal Hamiltonian (3.9) and inverse temperature β, if for each finite $\Lambda\subset\mathbb{Z}^d$, one has*

$$\mathbb{P}\big[(X(i))_{i\in\Lambda}\in\cdot\,\big|\,(X(i))_{i\in\mathbb{Z}^d\backslash\Lambda}=y\big]=\mu_y^{\Lambda,\beta}$$

for a.e. y w.r.t. μ.

We need to show that such infinite-volume Gibbs measures exist and are reversible invariant laws for the stochastic Ising model constructed in the previous section.

We first need some basic facts about Markov processes with finite state spaces. Let $(X_t)_{t\geq0}$ be a Markov process with finite state space S, Markov semigroup $(P_t)_{t\geq0}$, generator G and jump rates $\{r(x,y):x,y\in S,\ x\neq y\}$. By definition, a Markov process with given jump rates $\{r(x,y):x,y\in S,\ x\neq y\}$ is *irreducible* if

$$\forall S'\subset S \text{ with } S'\neq\emptyset,S\ \ \exists x\in S',\ y\notin S' \text{ such that } r(x,y)>0.$$

Proposition 3.6 (Ergodicity) *Consider a Markov process on a finite state space S with jump rates $\{r(x,y):x,y\in S,\ x\neq y\}$. If the jump rates are irreducible, then the Markov process has a unique invariant law μ and the process $(X_t)_{t\geq0}$ started in any initial law satisfies*

$$\mathbb{P}[X_t=x]\underset{t\to\infty}{\longrightarrow}\mu(x)\qquad(x\in S).\tag{3.10}$$

We recall from (1.3) that the Markov property is symmetric with respect to time reversal. Thus, if $(X_1,\ldots,X_n)$ is a (finite) Markov chain, then so is $(X_n,\ldots,X_1)$; similar statements hold for continuous-time processes. However, if a Markov process is time-homogeneous, then the same need not be true for the time-reversed process. An exception are stationary Markov processes: reversing the time in a stationary Markov process yields a stationary, hence time-homogeneous Markov process. The transition probabilities of this time-reversed process need not be the same as those of the original process, however. This leads to the following definition.

Definition 3.7 (Reversibility) *Let S be a finite set and let $(P_t)_{t\geq0}$ be a Markov semigroup on S. Then, by definition, we say that an invariant law μ of $(P_t)_{t\geq0}$ is* reversible *if the stationary process in with law $\mathbb{P}[X_t\in\cdot\,]=\mu$ satisfies*

$$\mathbb{P}\big[(X_{-t})_{t\in\mathbb{R}}\in\cdot\,\big]=\mathbb{P}\big[(X_{t-})_{t\in\mathbb{R}}\in\cdot\,\big].\tag{3.11}$$

Note that $(X_{-t})_{t\in\mathbb{R}}$ has *left*-continuous sample paths, which is why we compare this in (3.11) with $(X_{t-})_{t\in\mathbb{R}}$, the left-continuous modification of $(X_t)_{t\in\mathbb{R}}$. We state the following fact without proof.

Proposition 3.8 (Detailed balance) *A probability law μ on a finite set S is a reversible invariant law for a Markov process in S with jump rates $\{r(x,y) : x, y \in S,\ x \neq y\}$ if and only if*

$$\mu(x)r(x,y) = \mu(y)r(y,x) \qquad (x, y \in S). \tag{3.12}$$

Condition (3.12) is called *detailed balance*. Note that this says that in equilibrium, jumps from x to y happen with the same frequency as jumps from y to x.

We now apply these facts to study finite-volume Gibbs measures $\mu_y^{\Lambda,\beta}$ with fixed boundary conditions y. Our first result says that such finite-volume Gibbs measures are reversible invariant laws for a suitable finite-volume version of our stochastic Ising model.

Proposition 3.9 (Gibbs reversible law) *Let $\Lambda \subset \mathbb{Z}^d$ be a finite set, let $y \in \{-1,+1\}^{\mathbb{Z}^d\setminus\Lambda}$, and let $(X_t)_{t\geq 0}$ be the finite state Markov process in $\{-1,+1\}^\Lambda$ that jumps as*

$$x \mapsto x^{\{i\}} \quad \textit{with rate} \quad e^{-\beta\left(\sum_{j\in\mathcal{N}_i\cap\Lambda} 1_{\{x(i)=x(j)\}} + \sum_{j\in\mathcal{N}_i\cap\partial\Lambda} 1_{\{x(i)=y(j)\}}\right)}.$$

Then the Gibbs measure $\mu_y^{\Lambda,\beta}$ is a reversible invariant law for $(X_t)_{t\geq 0}$. Moreover, the process $(X_t)_{t\geq 0}$ started from any initial law satisfies $P[X_t \in \cdot\,] \underset{t\to\infty}{\Longrightarrow} \mu_y^{\Lambda,\beta}$.

Proof We must check detailed balance (3.12). Fix $i \in \Lambda$ and $x \in \{-1,+1\}^{\Lambda\setminus\{i\}}$, and define $x^-, x^+ \in \{-1,+1\}^\Lambda$ by

$$x^-(j) := \begin{cases} -1 & \text{if } j = i, \\ x(j) & \text{otherwise,} \end{cases} \qquad \text{and} \quad x^+(j) := \begin{cases} +1 & \text{if } j = i, \\ x(j) & \text{otherwise.} \end{cases}$$

We must check that

$$\mu_y^{\Lambda,\beta}(\{x^-\})r(x^-,x^+) = \mu_y^{\Lambda,\beta}(\{x^+\})r(x^+,x^-), \tag{3.13}$$

where $r(x^-,r^+)$ and $r(x^+,x^-)$ are the rates with which our process jumps from x^- to x^+ and back, respectively. Let

$$\begin{aligned} n_+ &:= \sum_{j\in\mathcal{N}_i\cap\Lambda} 1_{\{x(j)=+1\}} + \sum_{j\in\mathcal{N}_i\cap\partial\Lambda} 1_{\{y(j)=+1\}}, \\ n_- &:= \sum_{j\in\mathcal{N}_i\cap\Lambda} 1_{\{x(j)=-1\}} + \sum_{j\in\mathcal{N}_i\cap\partial\Lambda} 1_{\{y(j)=-1\}}. \end{aligned}$$

Then

$$\frac{\mu_y^{\Lambda,\beta}(\{x^+\})}{\mu_y^{\Lambda,\beta}(\{x^-\})} = \frac{e^{-\beta n_-}}{e^{-\beta n_+}} = \frac{r(x^-,x^+)}{r(x^+,x^-)},$$

which implies (3.13). To check that the process $(X_t)_{t\geq 0}$ is ergodic, it suffices to check irreducibility and apply Proposition 3.6. ∎

3.4 The upper and lower invariant laws

We still need to show the existence of infinite-volume Gibbs measures, as well as the fact that these are reversible invariant laws for our infinite-volume stochastic Ising model. We will concentrate on two special infinite-volume Gibbs measures, which are the upper and lower invariant laws of our stochastic Ising model.

Proposition 3.10 (Upper and lower invariant laws) *Let $(X_t)_{t\geq 0}$ be the stochastic Ising model from Section 3.2, started in the initial state $X_0(i) = +1$ for all $i \in \mathbb{Z}^d$. Then*

$$\mathbb{P}\big[X_t \in \cdot\,\big] \underset{t\to\infty}{\Longrightarrow} \overline{\nu},$$

where $\overline{\nu}$ is an invariant law of the process with the property that if ν is any other invariant law, then $\nu \leq \overline{\nu}$ in the stochastic order. Likewise, if $(X_t)_{t\geq 0}$ is started in $X_0(i) = -1$ for all $i \in \mathbb{Z}^d$, then

$$\mathbb{P}\big[X_t \in \cdot\,\big] \underset{t\to\infty}{\Longrightarrow} \underline{\nu},$$

where $\underline{\nu}$ is an invariant law of the process with the property that if ν is any other invariant law, then $\underline{\nu} \leq \nu$ in the stochastic order.

Proof Our graphical representation shows that the Ising model is *monotone*, i.e., if X^x and $X^{x'}$ are processes started in initial states such that $x \leq x'$, then we can couple X^x and $X^{x'}$ such that $X^x_t \leq X^{x'}_t$ for all $t \geq 0$. In terms of the semigroup $(P_t)_{t\geq 0}$ of our process, this says that if μ, μ' are laws on $\{-1,+1\}^{\mathbb{Z}^d}$ such that $\mu \leq \mu'$ in the stochastic order, then $\mu P_t \leq \mu' P_t$ for all $t \geq 0$. Applying this to $\mu = \delta_{\underline{+1}} P_{t-s}$ and $\mu' = \delta_{\underline{+1}}$, where $\underline{+1}$ denotes the all plus configuration, we see that $\delta_{\underline{+1}} P_s \geq \delta_{\underline{+1}} P_{t-s} P_s = \delta_{\underline{+1}} P_t$ for all $0 \leq s \leq t$. This means that for each sequence of times $t_n \uparrow \infty$ we can couple the random variables X_{t_n} such that the X_{t_n} decrease to some a.s. limit. It is not hard to see that this implies that $\mathbb{P}[X_t \in \cdot\,]$ converges weakly to some limit law $\overline{\nu}$ as $t \to \infty$, and using this we can prove that $\overline{\nu}$ is an invariant law. (We skip the details.) The fact that $\overline{\nu}$ is the largest invariant law in the stochastic order can be proved similar to the proof of Lemma 2.15. By symmetry, similar arguments apply to $\underline{\nu}$. ∎

For each finite $\Lambda \subset \mathbb{Z}^d$, we let $H_+^\Lambda(x)$ and $\mu_+^{\Lambda_n,\beta}$ denote the Hamiltonian $H_y^\Lambda(x)$ and finite-volume Gibbs measure, respectively, with boundary condition y given by $y(i) = +1$ for all $i \in \mathbb{Z}^d \backslash \Lambda$. We define $H_-^\Lambda(x)$ and $\mu_-^{\Lambda,\beta}$ similarly, with minus boundary conditions.

Proposition 3.11 (Limits of finite volume Gibbs measures) *Let $\Lambda_n \subset \mathbb{Z}^d$ be finite sets such that $\Lambda_n \uparrow \mathbb{Z}^d$. For each n, let $X^{\Lambda_n} = (X^{\Lambda_n}(i))_{i \in \mathbb{Z}^d}$ be a random variable such that $X(i) = +1$ for all $i \in \mathbb{Z}^d \backslash \Lambda_n$ and*

$$\mathbb{P}\big[(X^{\Lambda_n}(i))_{i \in \Lambda} \in \cdot\,\big] = \mu_+^{\Lambda_n,\beta}. \tag{3.14}$$

Then

$$\mathbb{P}\big[(X^{\Lambda_n}(i))_{i \in \mathbb{Z}^d} \in \cdot\,\big] \underset{n\to\infty}{\Longrightarrow} \overline{\nu}.$$

A similar statement holds for minus boundary conditions, in which case the limit is $\underline{\nu}$. Moreover, $\overline{\nu}$ and $\underline{\nu}$ are infinite-volume Gibbs measures in the sense of Definition 3.5.

Proof Let $(X_t)_{t\geq 0}$ be our infinite-volume stochastic Ising model started in $X_0 = \underline{+1}$ and for each n, let $(X_t^{\Lambda_n})_{t\geq 0}$ be a process such that $X_t^{\Lambda_n}(i) = +1$ for all $i \in \mathbb{Z}^d \backslash \Lambda_n$ and $t \geq 0$, while inside Λ, the process evolves as in Proposition 3.9, with plus boundary conditions and initial state $X_0^{\Lambda_n}(i) = +1$ for all i. Using the graphical representation, we see that we can couple our processes such that $X_t^{\Lambda_n} \geq X_t^{\Lambda_m} \geq X_t$ for all $t \geq 0$ and $n \leq m$. Taking the limit $t \to \infty$ we see that the random variables X^{Λ_n} from (3.14) can be coupled such that they decrease to an a.s. limit; in particular, this implies that their laws converge weakly to some limit ν. By using techniques similar to the proof of Lemma 2.24, we can prove that ν is an invariant law for the infinite-volume stochastic Ising model, while our coupling shows that $\nu \geq \overline{\nu}$. Since $\overline{\nu}$ is the largest invariant law, it follows that $\nu = \overline{\nu}$. The fact that $\overline{\nu}$ and $\underline{\nu}$ are infinite-volume Gibbs measures in the sense of Definition 3.5 is obvious, since the approximating finite-volume Gibbs measures have the right conditional distributions. ∎

3.5 The spontaneous magnetization

Let $\overline{\nu}$ be the upper invariant law of the Ising model. By definition, the quantity (which by translation invariance does not depend on $i \in \mathbb{Z}^d$)

$$m^*(\beta, d) = m^*(\beta) := \int \overline{\nu}(\mathrm{d}x)x(i) \qquad (\beta \geq 0)$$

is called the *spontaneous magnetization*. By symmetry, we have

$$\int \underline{\nu}(\mathrm{d}x)x(i) = -m^*(\beta).$$

Since moreover $\underline{\nu} \leq \overline{\nu}$, it follows that $\underline{\nu} \neq \overline{\nu}$ if and only if $m^*(\beta) > 0$. Since $\underline{\nu}$ and $\overline{\nu}$ are the lowest and highest invariant law in the stochastic order, this implies that our stochastic Ising model has a unique invariant law if and only if $m^*(\beta) = 0$. In this and the next section, we will prove the following theorem.

Theorem 3.12 (Phase transition of the Ising model) *The function $m^*(\beta, d)$ is nondecreasing and right-continuous in β and nondecreasing in d. In dimension $d = 1$ one has $m^*(\beta) = 0$ for all $\beta \geq 0$. On the other hand, for all dimensions $d \geq 2$, there exists a critical value $0 < \beta_{\rm c} < \infty$ such that $m^*(\beta) = 0$ for $\beta < \beta_{\rm c}$ and $m^*(\beta) > 0$ for $\beta > \beta_{\rm c}$.*

In the present section, we will prove that $\beta \mapsto m^*(\beta, d)$ is nondecreasing and right-continuous and $d \mapsto m^*(\beta, d)$ is nondecreasing. In the next section, we will prove that $\beta_{\rm c} = \infty$ in dimension $d = 1$ and $\beta_{\rm c} < \infty$ in dimensions $d \geq 2$.

At first, one might think that monotonicity of the spontaneous magnetization in β and d can be proved by the same sort of monotonicity arguments that we have used so far, by coupling Markov processes (in our case, stochastic Ising models) with different values of β in such a way that one process 'stays above' the other. It seems, however, that this idea does not work. Indeed, increasing β means that spins 'like more to be aligned'. Since our dynamics treat pluses and minuses in a symmetric way, this means that pluses are more favored near other pluses and minuses are more favored near minuses, an effect that can work both ways. In view of this, we have to take a different approach. Our proof will be based on Griffiths' inequalities. An alternative proof (not given here) uses a representation of our Gibbs measures in terms of the so-called random cluster model. It can be shown that the latter is monotone in β and d in the usual sense, leading to the desired monotonicities for $m^*(\beta, d)$.

Let Λ be a finite set, let $\mathcal{P}(\Lambda)$ denote the set of all subsets of Λ, and let $\mathcal{P}(\Lambda) \ni A \mapsto J_A \in \mathbb{R}$ be any function. For any $A \in \mathcal{P}(\Lambda)$ and $x \in \{-1, +1\}^\Lambda$, we write

$$x_A := \prod_{i \in A} x(i),$$

where $x_\emptyset := +1$. We will be interested in Gibbs measures on $\{-1, +1\}^\Lambda$ of the form

$$\mu_J(\{x\}) := \frac{1}{Z_J} e^{\sum_{A \subset \Lambda} J_A x_A},$$

where Z_J is the normalization constant (also known as partition sum)

$$Z_J := \sum_x e^{\sum_A J_A x_A}.$$

We start by observing that

$$\begin{aligned} &\text{(i)} && \tfrac{\partial}{\partial J_A} \log Z_J = \int \mu_J(\mathrm{d}x) x_A, \\ &\text{(ii)} && \tfrac{\partial^2}{\partial J_A \partial J_B} \log Z_J = \int \mu_J(\mathrm{d}x) x_A x_B - \int \mu_J(\mathrm{d}x) x_A \int \mu_J(\mathrm{d}x) x_B. \end{aligned} \tag{3.15}$$

To see this, just write

$$\tfrac{\partial}{\partial J_A} \log Z_J = \frac{\tfrac{\partial}{\partial J_A} Z_J}{Z_J}$$

and

$$\tfrac{\partial^2}{\partial J_A \partial J_B} \log Z_J = \tfrac{\partial}{\partial J_B} \frac{\tfrac{\partial}{\partial J_A} Z_J}{Z_J} = \frac{Z_J \tfrac{\partial^2}{\partial J_A \partial J_B} Z_J - (\tfrac{\partial}{\partial J_A} Z_J)(\tfrac{\partial}{\partial J_B} Z_J)}{Z_J^2},$$

where

$$\tfrac{\partial}{\partial J_A} Z_J = \tfrac{\partial}{\partial J_A} \sum_x e^{\sum_C J_C x_C} = \sum_x x_A e^{\sum_C J_C x_C}$$

and

$$\tfrac{\partial^2}{\partial J_A \partial J_B} Z_J = \tfrac{\partial}{\partial J_B} \sum_x x_A e^{\sum_C J_C x_C} = \sum_x x_A x_B e^{\sum_C J_C x_C}.$$

Proposition 3.13 (Griffiths' inequalities) *Assume that $J_A \geq 0$ for all $A \subset \Lambda$. Then*

$$\begin{aligned} &\text{(i)} && \tfrac{\partial}{\partial J_A} \log Z_J \geq 0, \\ &\text{(ii)} && \tfrac{\partial^2}{\partial J_A \partial J_B} \log Z_J \geq 0 \end{aligned}$$

for all $A, B \subset \Lambda$.

Proof We observe that

$$\begin{aligned} Z_J &= \sum_x e^{\sum_A J_A x_A} \\ &= \sum_x \sum_{n=0}^{\infty} \frac{1}{n!} \big(\sum_A J_A x_A \big)^n \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{A_1} \cdots \sum_{A_n} \big(\prod_{k=1}^n J_{A_k} \big) \sum_x \prod_{k=1}^n x_{A_k}. \end{aligned}$$

Since

$$x_A x_B = x_{A \triangle B},$$

where $A \triangle B$ denotes the symmetric difference of A and B, we see that

$$\sum_x \prod_{k=1}^n x_{A_k} = \sum_x x_{A_1 \triangle \cdots \triangle A_n} = \begin{cases} 2^{|\Lambda|} & \text{if } A_1 \triangle \cdots \triangle A_n = \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

Thus

$$Z_J = 2^{|\Lambda|} \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{A_1} \cdots \sum_{A_n} 1_{\{A_1 \triangle \cdots \triangle A_n = \emptyset\}} \prod_{k=1}^n J_{A_k}.$$

Likewise

$$\begin{aligned} \tfrac{\partial}{\partial J_A} \log Z_J &= \frac{1}{Z_J} \sum_x x_A e^{\sum_C J_C x_C} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{A_1} \cdots \sum_{A_n} \Big(\prod_{k=1}^n J_{A_k}\Big) x_A \sum_x \prod_{k=1}^n x_{A_k} \\ &= \frac{1}{Z_J} 2^{|\Lambda|} \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{A_1} \cdots \sum_{A_n} 1_{\{A \triangle A_1 \triangle \cdots \triangle A_n = \emptyset\}} \prod_{k=1}^n J_{A_k}, \end{aligned}$$

which is clearly nonnegative provided the $J_A \geq 0$ for all A. To prove also Griffiths' second inequality, we write, using (3.15),

$$\begin{aligned} \tfrac{\partial^2}{\partial J_A \partial J_B} \log Z_J &= \frac{1}{Z_J^2} \Big(\sum_x x_A x_B e^{\sum_C J_C x_C} \Big) \Big(\sum_y e^{\sum_C J_C y_C} \Big) \\ &\quad - \frac{1}{Z_J^2} \Big(\sum_x x_A e^{\sum_C J_C x_C} \Big) \Big(\sum_y y_A e^{\sum_C J_C y_C} \Big) \\ &= \frac{1}{Z_J^2} \sum_{x,y} \big(x_A x_B - x_A y_B \big) e^{\sum_C J_C (x_C + y_C)}. \end{aligned}$$

Using the facts that $y_B = (x_B)^2 y_B = x_B (xy)_B$ and $x_A x_B = x_{A \triangle B}$, we may rewrite our formula as

$$\begin{aligned} \tfrac{\partial^2}{\partial J_A \partial J_B} \log Z_J &= \frac{1}{Z_J^2} \sum_{x,y} x_{A \triangle B} \big(1 + (xy)_B\big) e^{\sum_C J_C x_C (1 + (xy)_C)} \\ &= \frac{1}{Z_J^2} \sum_{x,z} x_{A \triangle B} \big(1 + z_B\big) e^{\sum_C J_C x_C (1 + z_C)} \\ &= \frac{1}{Z_J^2} \sum_z \big(1 + z_B\big) \sum_x x_{A \triangle B} e^{\sum_C J_C^z x_C}, \end{aligned}$$

where we have defined $J^z_C := \big(1+z_C\big)J_C$. Since $|z_C|=1$, we have $J^z_C \geq 0$ for all z, hence by Griffiths' first inequality

$$\sum_x x_{A\triangle B}\, e^{\sum_C J^z_C x_C} \geq 0$$

for each $z \in \{-1,+1\}^\Lambda$. Summing over x we obtain Griffiths' second inequality. ∎

The monotonicity of the spontaneous magnetization in β follows from Proposition 3.11 and the following simple consequence of Proposition 3.13.

Lemma 3.14 (Monotonicity of magnetization) *For any finite set $\Lambda \subset \mathbb{Z}^d$ and $i \in \Lambda$, one has*

$$\tfrac{\partial}{\partial\beta}\int \mu^{\Lambda,\beta}_+(\mathrm{d}x)x(i) \geq 0.$$

Proof We claim that $\mu^{\Lambda,\beta}_+ = \mu_J$ for a suitable function J. Indeed, up to an irrelevant additive constant, we may rewrite our Hamiltonian as

$$H^\Lambda_+(x) = -\tfrac{1}{2}\sum_{\{i,j\}\in\mathcal{B}_\Lambda} x(i)x(j) - \tfrac{1}{2}\sum_{(i,j)\in\partial\mathcal{B}_\Lambda} x(i).$$

In view of this, our finite volume Gibbs measures are generated by the function J defined by

$$J_{\{i,j\}} := \tfrac{1}{2}\beta$$

if $i,j \in \Lambda$, $|i-j|=1$,

$$J_{\{i\}} := \tfrac{1}{2}\beta|\{j \in \mathbb{Z}^d\backslash\Lambda : |i-j|=1\}$$

and $J_A := 0$ in all other cases. It is now clear that increasing β means increasing the function J and hence, by Proposition 3.13, increasing $\int \mu_J(\mathrm{d}x)x(i)$. ∎

The monotonicity of $m^*(\beta,d)$ in d is proved in a similar way. Indeed, if $d \leq d'$, then we may view $\mathbb{Z}^d$ as a subset of $\mathbb{Z}^{d'}$. With positive boundary conditions, if we switch on the interaction between sites inside $\mathbb{Z}^d$ and sites in $\mathbb{Z}^{d'}\backslash\mathbb{Z}^d$, then by Proposition 3.13 this will lead to a higher magnetization in any point in $\mathbb{Z}^d$.
We conclude this section with the following result.

Lemma 3.15 (Right-continuity) *The spontaneous magnetization $m^*(\beta)$ is a right-continuous function of β.*

Proof Let $\overline{\nu}_\beta$ denote the upper invariant law at inverse temperature β and let $\beta_n \downarrow \beta$. Using the compactness of our state space, going to a subsequence if necessary, we may assume that $\overline{\nu}_{\beta_n} \Rightarrow \nu$ for some probability law ν. Just as in Lemma 2.24, we can show that ν is an invariant law for the stochastic Ising model with inverse temperature β. Moreover, since $\beta \mapsto m^*(\beta)$ is nondecreasing, we must have

$$\lim_{\beta_n \downarrow \beta} m^*(\beta_n) = \int \nu(\mathrm{d}x)x(0) \geq m^*(\beta).$$

Since $\overline{\nu}_\beta$ is the largest invariant law w.r.t. the stochastic order, we must have

$$\int \nu(\mathrm{d}x)x(0) \leq \int \overline{\nu}_\beta(\mathrm{d}x)x(0) = m^*(\beta),$$

proving our claim. ∎

3.6 Existence of a phase transition

We conclude this chapter with two of the oldest results in the field, namely, the result by Ising on the nonexistence of a phase transition for his model in dimension $d = 1$, and the result by Peierls on the existence of a phase transition in dimensions $d \geq 2$. We start with Ising's result.

Lemma 3.16 (No phase transition in one dimension) *In dimension $d = 1$, for each $\beta \geq 0$, there exists a unique infinite-volume Gibbs measure μ associated with the formal Hamiltonian (3.9) and inverse temperature β. If $X = (X(i))_{i\in\mathbb{Z}}$ is a random variable with law μ, then X is a stationary Markov chain with transition probabilities*

$$\mathbb{P}\big[X(i+1) \neq X(i) \,\big|\, X(i)\big] = \frac{e^{-\beta}}{e^{-\beta}+1}. \tag{3.16}$$

Proof Let $\overline{\nu}$ be the upper invariant law of the one-dimensional Ising model with inverse temperature β and let $X = (X(i))_{i\in\mathbb{Z}}$ is a random variable with law $\overline{\nu}$. We claim that X is a Markov chain. By Proposition 3.11 it suffices to prove that for any finite interval $\Lambda_n = \{-n, \ldots, n\}$, the finite-volume Gibbs measures $\mu_+^{\Lambda_n,\beta}$ are the laws of a finite Markov chain. Let X^{Λ_n} be a random variable with law $\mu_+^{\Lambda_n,\beta}$. We need to show that for any $-n \leq k \leq n$, the random variables

$$(X^{\Lambda_n}(i))_{-n\leq i<k} \quad \text{and} \quad (X^{\Lambda_n}(i))_{k<i\leq n}$$

are conditionally independent given $X^{\Lambda_n}(k)$. But this follows from Lemma 3.4 and the structure of the finite-volume Gibbs measures $\mu_y^{\Lambda_n\setminus\{k\},\beta}$ with $y(i) = +1$ for $i \notin \Lambda_n$ and $y(k) = -1$ or $+1$.

Since (by Proposition 3.10) the upper invariant law is invariant under translations and mirror images, the Markov chain $X = (X(i))_{i\in\mathbb{Z}}$ is stationary and reversible. Set

$$p := \mathbb{P}\big[X(i+1) = +1 \,\big|\, X(i) = -1\big] \quad \text{and} \quad q := \mathbb{P}\big[X(i+1) = -1 \,\big|\, X(i) = +1\big].$$

Then

$$\mathbb{P}[X(i) = +1] = \frac{p}{p+q}.$$

From the fact that X is an infinite volume Gibbs measure for the Ising model, by Lemma 3.3, we know that

$$\frac{\mathbb{P}[X(i) = +1 \,|\, X(i-1) = -1 = X(i+1)]}{\mathbb{P}[X(i) = -1 \,|\, X(i-1) = -1 = X(i+1)]} = \frac{e^{-2\beta}}{1}.$$

Since

$$\begin{aligned}\mathbb{P}[X(i-1) = -1, X(i) = +1, X(i+1) = -1] &= \frac{p}{p+q}pq,\\ \mathbb{P}[X(i-1) = -1, X(i) = -1, X(i+1) = -1] &= \frac{p}{p+q}(1-p)^2,\end{aligned}$$

this leads to the equation

$$\frac{pq}{(1-p)^2} = e^{-2\beta}.$$

Likewise, since

$$\frac{\mathbb{P}[X(i) = +1 \,|\, X(i-1) = -1, X(i+1) = +1]}{\mathbb{P}[X(i) = -1 \,|\, X(i-1) = -1, X(i+1) = +1]} = \frac{e^{-\beta}}{e^{-\beta}}$$

and

$$\begin{aligned}\mathbb{P}[X(i-1) = -1, X(i) = +1, X(i+1) = +1] &= \frac{p}{p+q}p(1-q),\\ \mathbb{P}[X(i-1) = -1, X(i) = -1, X(i+1) = +1] &= \frac{p}{p+q}(1-p)p,\end{aligned}$$

we see that

$$\frac{1-q}{1-p} = \frac{e^{-\beta}}{e^{-\beta}},$$

hence $p = q$. By our previous equation this implies

$$\Big(\frac{p}{1-p}\Big)^2 = e^{-2\beta},$$

which in turn implies (3.16). It follows that $\mathbb{E}[X(i)] = p/(p+q) = 1/2$, hence $m^*(\beta, 1) = 0$ for all $\beta \geq 0$. ∎

Since $m^*(\beta, d)$ is nondecreasing in d, in order to prove the existence of a phase transition in dimensions $d \geq 2$, it suffices to treat the case $d = 2$.

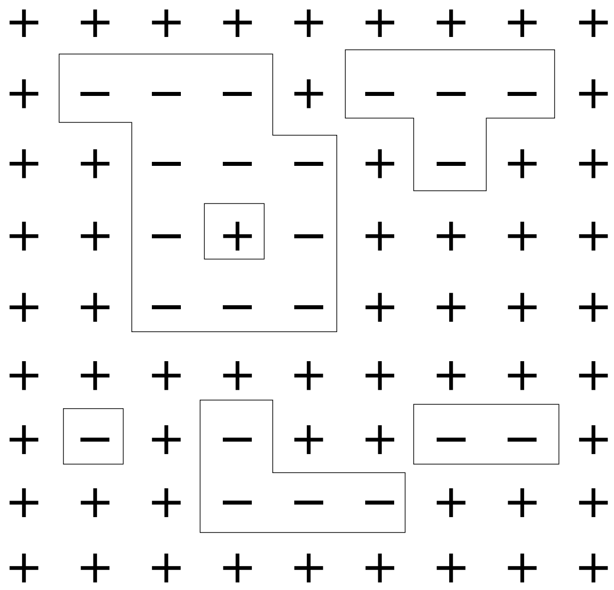

Figure 3.3: Peierls argument for Ising model.

Proposition 3.17 (Estimate on critical temperature) *One has* $m^*(\beta,2)>0$ *for all* $\beta>\log 3$.

Proof We will use the original Peierls argument from [Pei36]. Let

$$\Lambda_n := \{-n,\ldots,n\}^2.$$

We may view Λ_n as a graph with edges between nearest neighbors. In this picture, for a given spin configuration $x\in\{-1,=1\}^{\Lambda_n}$, we may group the -1 spins and $+1$ spins into connected components, each surounded by a closed curve (see Figure 3.3).
There is a one-to-one correspondence between configurations of curves and configurations of spins. In particular, the origin has a $+1$ spin if and only if it is surrounded by an even number of curves. More formally, for each $x\in\{-1,+1\}^{\Lambda_n}$, define $\overline{x}\in\{-1,+1\}^{\Lambda_{n+1}}$ by

$$\overline{x}(i) := \begin{cases} x(i) & \text{if } i\in\Lambda_n,\\ +1 & \text{if } i\in\partial\Lambda_n,\end{cases}$$

let $\mathcal{E}_n$ be the collection of all pairs $\{i,j\}$ with $|i-j|=1$, $i,j\in\Lambda_{n+1}$, and define $\Gamma(x)\subset\mathcal{E}_n$ by

$$\Gamma(x) := \big\{\{i,j\} : \overline{x}(i)\neq\overline{x}(j)\big\}.$$

Let

$$\mathcal{G}_n := \{\Gamma(x) : x \in \{-1,+1\}^{\Lambda_n}\}$$

be the configuration of all 'configurations of curves'. Then the probability of seeing a certain configuration of curves is given by

$$\rho(\{\Gamma\}) = \frac{1}{Z} e^{-\beta|\Gamma|},$$

where $|\Gamma|$ is the total length of the curves in the configuration Γ and

$$Z := \sum_{\Gamma \in \mathcal{G}_n} e^{-\beta|\Gamma|}$$

is a normalization constant. Now let $\gamma \subset \mathcal{E}$ be a collection of nearest-neighbor edges that form a closed curve (not a configuration of curves but just one single curve) surrounding the origin. We can ask what the probability is of seeing a configuration of curves in which this this particular curve is present. This probability is, of course,

$$\begin{aligned}
&\frac{1}{Z} \sum_{\Gamma \in \mathcal{G}_n:\, \gamma \subset \Gamma} e^{-\beta|\Gamma|} \\
&\quad = \frac{\sum_{\Gamma:\, \gamma\subset\Gamma} e^{-\beta|\Gamma|}}{\sum_{\Gamma} e^{-\beta|\Gamma|}} \\
&\quad \leq \frac{\sum_{\Gamma:\, \gamma\subset\Gamma} e^{-\beta|\Gamma|}}{\sum_{\Gamma:\, \gamma\subset\Gamma} e^{-\beta|\Gamma|} + \sum_{\Gamma:\, \gamma\cap\Gamma=\emptyset} e^{-\beta|\Gamma|}} \\
&\quad = \frac{\sum_{\Gamma:\, \gamma\subset\Gamma} e^{-\beta|\Gamma|}}{\sum_{\Gamma:\, \gamma\subset\Gamma} e^{-\beta|\Gamma|} + e^{\beta|\gamma|} \sum_{\Gamma:\, \gamma\subset\Gamma} e^{-\beta|\Gamma|}} = \frac{e^{-\beta|\gamma|}}{e^{-\beta|\gamma|}+1} \leq e^{-\beta|\gamma|}.
\end{aligned}$$

Here we use that for every configuration of curves in which γ is present, there is another configuration in which γ is completely removed, which is a factor $e^{\beta|\gamma|}$ more likely than the configuration in which γ is present. Since there are at most $k3^k$ different curves γ of length k surrounding the origin, we find that the expected number of curves surrounding the origin can be estimated from above by

$$\sum_{k=4}^{\infty} k3^k e^{-k\beta}.$$

By choosing β sufficiently small, we can make this number as close to zero as we wish; in particular, this proves that (uniformly (!) in n) $\int \mu_+^{\Lambda_n,\beta}(\mathrm{d}x)x(0) > \frac{1}{2}$ for β sufficiently large.

Unfortunately, this does not quite give the explicit bound we are after. If $\beta > \log 3$, then we see that the expected number of curves surrounding the origin is finite (where, again, our estimate is uniform in n), but this is not enough to conclude that $\int \mu_+^{\Lambda_n,\beta}(\mathrm{d}x)x(0) > \frac{1}{2}$, hence $m^*(\beta) > 0$.
To fix this problem, we use a trick. We fix some $m \leq n$ and look at the proportion of probabilities

$$\frac{\int \mu_+^{\Lambda_n,\beta}(\mathrm{d}x)1_{\{x(i)=-1\ \forall i\in\Lambda_m\}}}{\int \mu_+^{\Lambda_n,\beta}(\mathrm{d}x)1_{\{x(i)=+1\ \forall i\in\Lambda_m\}}}.$$

We note that the event $\{x(i) = -1\ \forall i \in \Lambda_m\}$ occurs if and only if there are no contours inside Λ_m and there is an odd number of contours surroundig Λ_m. Likewise, the event $\{x(i) = +1\ \forall i \in \Lambda_m\}$ occurs if and only if there are no contours inside Λ_m and there is an even number of contours surroundig Λ_m. We can estimate the proportion of the probabilities of these events by estimating the expected number of contours surrounding Λ_m, conditional on the event that there are no contours inside Λ_m. By the same arguments as above, this expectation can be estimated by

$$\sum_{k=4m}^{\infty} k3^k e^{-k\beta},$$

which in case $\beta > \log 3$ can be made arbitrarily small by choosing m sufficiently large. Now, letting $\Lambda_n \uparrow \infty$ while keeping m fixed, using Proposition 3.11, we see that the upper invariant measure $\overline{\nu}$ satisfies

$$\frac{\int \overline{\nu}(\mathrm{d}x)1_{\{x(i)=-1\ \forall i\in\Lambda_m\}}}{\int \overline{\nu}(\mathrm{d}x)1_{\{x(i)=+1\ \forall i\in\Lambda_m\}}} < 1$$

for some m. In particular, this shows that $\overline{\nu}$ is not symmetric with respect to a simultaneous flip of all spins, hence $\overline{\nu} \neq \underline{\nu}$. As we have already seen, this implies that $m^*(\beta) > 0$. ∎

3.7 Other topics

For the Ising model on $d = 2$, Onsager has shown [Ons44] that

$$\beta_\mathrm{c} = \log(1 + \sqrt{2})$$

and

$$m^*(\beta, 2) = \big(1 - \sinh(\beta)^{-4}\big)^{1/8} \qquad (\beta \geq \beta_\mathrm{c}),$$

where

$$\sinh(\beta) = \tfrac{1}{2}\big(e^{\beta} - e^{-\beta}\big)$$

is the sinus hyperbolicus. Note that in light of this, the estimate $\beta_{\rm c} \leq \log 3$ arising from Proposition 3.17 is not so bad! Onsager's solution also implies that

$$m^*(\beta, 2) \sim (\beta - \beta_{\rm c})^{1/8} \quad \text{as } \beta \downarrow \beta_{\rm c},$$

which shows that the critical exponent associated with the spontaneous magnetrization is 1/8 in dimension $d = 2$. It is supposed that

$$m^*(\beta, 3) \sim (\beta - \beta_{\rm c})^{0.308} \quad \text{as } \beta \downarrow \beta_{\rm c},$$

but there is no mathematical theory to explain this. (There is -nonrigorous- renormalization group theory that sort of 'explains' this and even allows one to calculate the critical exponent with some precision.) This critical exponent can actually be measured and has been experimentally observed for various magnetic systems and gasses near the critical point. Obviously, these physical systems are locally not very similar to the (nearest-neighbor) Ising model, but it is believed (and up to some level understood by renormalization group theory) that this critical exponent is universal and shared by a large number of different models.

Similar to what we know for the contact process, one can prove that for the Ising model, all spatially homogeneous infinite volume Gibbs measures are convex combinations of $\overline{\nu}$ and $\underline{\nu}$. In dimension 2, these are in fact all infinite volume Gibbs measures, but, contrary to what we saw for the contact process, in dimensions $d \geq 3$ there exist infinite volume Gibbs measures for the Ising model that are not translation invariant.

Generalizing from the Ising model, one may look at models where spins can take $q = 2, 3, \ldots$ values, described by Gibbs measures with a Hamiltonian of the form (3.9). These models are called *Potts models*. An interesting feature of these models is that while the spontaneous magnetization $m^*(\beta)$ is (supposed to be) a continuous of β for the Ising model, it is known that the same is not always true for Potts models. Ising and Potts models can be studied in a nice uniform framework using the random cluster model.

Chapter 4

Voter models

4.1 The basic voter model

The standard, nearest-neighbor voter model on $\mathbb{Z}^d$ is the Markov process with values in $\{0,1\}^{\mathbb{Z}}$ defined by the generator

$$Gf(x) := \sum_i \tfrac{1}{2d} \sum_{j:\,|i-j|=1} 1_{\{x(j)\neq x(i)\}} \big(f(x^{\{i\}}) - f(x)\big), \tag{4.1}$$

where $x^{\{i\}}$ is defined as in (2.2) and the generator is first defined for functions in $\mathcal{C}_{\rm sum}(\{0,1\}^{\mathbb{Z}^d})$ and then by closure for a larger class of functions, as explained in Chapter 1.

Note that there is no real parameter in (4.1) in which we could observe a phase transition, comparable to the infection rate of the contact process or the inverse temperature of the Ising model. The only free parameter is the dimension d. Indeed, we will see that the behavior of the voter model depends on whether $d \leq 2$ or $d > 2$, where $d = 2$ is the 'critical dimension'.

In the classical interpretation of the voter model, sites represent individuals and the *type* $X_t(i)$ of a site represents the political opinion held by individual i at time t. (The model was obviously invented in a country with a two-party system like, for example, the USA.) Then (4.1) says that people behave in a conformist way, i.e., they change their opinion at a rate proportional to the number of neighbors that hold the other opinion. An equivalent way of describing this is to say that at times of a rate one Poisson point process, an individual decides to update its opinion by choosing one of its $2d$ neigbors at randon and adopting its opinion (with the result that nothing happens if the newly adopted opinion is the same as the original opinion). This description suggests a natural way to write (4.1) in terms of local maps, and hence a graphical representation. For each $i, j \in \mathbb{Z}^d$, let

$$m_{ij}(x)(k) := \begin{cases} x(i) & \text{if } k = j, \\ x(k) & \text{otherwise.} \end{cases} \tag{4.2}$$

Then we may rewrite (4.1) as

$$Gf(x) := \sum_i \tfrac{1}{2d} \sum_{j:\,|i-j|=1} \big(f(m_{ij}(x)) - f(x)\big). \tag{4.3}$$

In a graphical representation, the local map m_{ij} is usually represented by an arrow from i to j. In a more serious interpretation, we may interpret the type $X_t(i)$ of a site at a given time as the genetic type of some organism living at this position. Then (4.3) says that organisms die at rate one and after their deaths are replaced by a descendant of one of their neighbors. Thus, the voter model can be used to model *neutral evolution* where the genetic type of an organism has no influence on its fitness. This is often a useful first step towards more complicated models where the fitness of types matters. Also, it is believed that large parts of the DNA of most organisms consist of 'junk' that has no influence on fitness but may still be interesting in the study descendancy relations.

4.2 Coalescing ancestries

Imagine that for a given voter model, constructed with a graphical representation as described in the previous section, we want to know the type $X_t(i)$ of a site i at some time $t > 0$. Then we may look at the last time $\sigma_1 < t$ when an arrow ends at i and hence the site i possibly changed its type. If $\sigma_1 < 0$, we are done; otherwise, we know that the organism at i was at time σ_1 replaced by a descendant of the organism at the neighboring site j where the arrow starts, so it suffices to follow this organism back in time till the last time $\sigma_2 < \sigma_1$ that an arrow ended there. Continuing this process, following arrows backwards whenever we meet one that ends at the site where we currently are, we see that for each $(i, t) \in \mathbb{Z}^d \times \mathbb{R}$ we can define a $\mathbb{Z}^d$-valued process $(\xi^{(i,t)}_s)_{s\geq 0}$ such that $\xi^{(i,t)}_s$ is the position where the (unique) *ancestor* of (i, t) lived at time $t - s$. In particular, following ancestors back till time zero, we find that

$$X_t(i) = X_0(\xi^{(i,t)}_t) \qquad (i \in \mathbb{Z}^d,\ t \geq 0). \tag{4.4}$$

It is easy to see (compare Theorem 1.12) that $(\xi^{(i,t)}_s)_{s\geq 0}$ is a random walk that jumps with rate one to a uniformly chosen neighboring site. Moreover, two 'ancestral' random walks started at two different space-time points (i,t) and (i',t') behave independently as long as they are apart and *coalesce* (i.e., go together as one) as soon as they meet.

The coalescing random walks $(\xi_s^{(i,t)})_{s\geq 0}$ are often called *lines of descent.* They are exactly the 'paths of influence' from Chapter 1. In view of this, it is easy to see that the constant K from (1.23) is zero. Therefore, unlike for the contact process and Ising model, the value of this constant tells us nothing about ergodicity of the process. Indeed, since the constant configurations 0 and 1 are traps for the model, the delta measures δ_0 and δ_1 are always invariant measures so the invariant law is never unique. In many ways, the voter model in any dimension is a 'critical' model, i.e., it behaves in many ways similar to other models at their critical point.

Proposition 4.1 (Invariant laws in low dimensions) *In dimensions $d = 1,2$, all invariant laws of the voter model are convex combinations of the delta measures δ_0 and δ_1 on the constant configurations 0 and 1.*

Proof The voter model stared in any initial law satisfies

$$\mathbb{P}[X_t(i) = X_t(j)] = \mathbb{P}[X_0(\xi_t^{(i,t)}) = X_0(\xi_t^{(j,t)})] \geq \mathbb{P}[\xi_t^{(i,t)} = \xi_t^{(j,t)}]. \tag{4.5}$$

We observe that $\big(\xi_s^{(i,t)} - \xi_s^{(j,t)}\big)_{s\geq 0}$ is a Markov process that everywhere outside the origin jumps with rate 2 to a uniformly chosen neighbor and when it reaches the origin is trapped there. Since one-dimensional nearest-neighbor random walk is recurrent in dimensions $d = 1,2$, it follows that the right-hand side of (4.5) tends to zero as $t \to \infty$. In particular, this means that any invariant law must be concentrated on constant configurations. Since, on the other hand, δ_0 and δ_1 are invariant laws for the voter model, this proves our claim. ∎

4.3 Duality

Our next aim is to prove that in dimensions $d > 2$, there are invariant laws of the voter model that are not convex combinations of δ_0 and δ_1. A convenient way to do this is to use duality.

We have already met the self-duality of the contact process (Lemma 2.1). More generally, two continuous-time Markov processes X and Y with generators G_X and G_Y and state spaces S and T are called *dual* to each other, with respect to a *duality function* $\Psi : S \times T \to \mathbb{R}$, if

$$\mathbb{E}[\Psi(X_0, Y_t)] = \mathbb{E}[\Psi(X_t, Y_0)] \qquad (t \geq 0), \tag{4.6}$$

whenever X and Y are independent (with arbitrary initial laws). Since we can always integrate over the initial laws, in order to check (4.6), it suffices to prove the statement for deterministic initial states, i.e., (4.6) is equivalent to

$$\mathbb{E}^y[\Psi(x,Y_t)] = \mathbb{E}^x[\Psi(X_t,y)] \qquad (t \geq 0,\ x \in S,\ y \in T), \tag{4.7}$$

where $\mathbb{E}^x$ (resp. $\mathbb{E}^y$) denotes expectation with respect to the law of the process X (resp. Y) started in $X_0 = x$ (resp. $Y_0 = y$). If the state spaces S, T are finite, then a necessary and sufficient condition for (4.6) is that

$$G_X \Psi(\,\cdot\,, y)(x) = G_Y \Psi(x, \,\cdot\,)(y) \qquad (x \in S,\ y \in T). \tag{4.8}$$

This condition, plus some technical assumptions, is also often sufficient for processes with infinite state spaces, see e.g. [AS09a]. Note that the self-duality of the contact process fits in this general scheme, where $\Psi(A,B) = 1_{\{A\cap B\neq\emptyset\}}$.

Below, for any $x, y \in \{0,1\}^{\mathbb{Z}^d}$, we write $|x| := \sum_i x(i)$ and $(xy)(i) := x(i)y(i)$. Note that in this notation, the self-duality function of the contact process can be rewritten as $\Psi(x,y) = 1_{\{|xy|\neq 0\}}$. It turns out that the same duality function also yields a useful dual for the voter model.

Lemma 4.2 (Coalescent dual of the voter model) *Let X be a voter model and let Y be the Markov process with values in $\{0,1\}^{\mathbb{Z}^d}$ and generator*

$$G_{\rm coal} f(y) := \sum_i \tfrac{1}{2d} \sum_{j:\,|i-j|=1} \big(f(c_{ij}(y)) - f(y)\big), \tag{4.9}$$

where

$$c_{ij}(y)(k) := \begin{cases} 0 & \text{if } k = i, \\ y(i) \vee y(i) & \text{if } k = j, \\ y(k) & \text{otherwise.} \end{cases} \tag{4.10}$$

Then

$$\mathbb{P}\big[|X_t Y_0| \neq 0\big] = \mathbb{P}\big[|X_0 Y_t| \neq 0\big] \qquad (t \geq 0) \tag{4.11}$$

whenever X and Y are independent.

Proof It suffices to prove the statement for deterministic initial states $X_0 = x$ and $Y_0 = y$. Set $A := \{i : y(i) = 1\}$. Then

$$Y_s(i) := 1_{\{\exists j \in A\,:\ \xi^{(j,t)}_s = i\}} \qquad (s \geq 0) \tag{4.12}$$

defines a Markov process with generator as in (4.9). It follows that

$$\mathbb{P}\big[|X_t y| \neq 0\big] = \mathbb{P}[x(\xi^{(i,t)}_t) = 1 \text{ for some } i \in A] = \mathbb{P}\big[|X_0 Y_t| \neq 0\big] \qquad (t \geq 0). \tag{4.13}$$

∎

The voter model also has an annihilating dual.

Lemma 4.3 (Annihilating dual of the voter model) *Let X be a voter model and let Y be the Markov process with values in $\{0,1\}^{\mathbb{Z}^d}$ and generator*

$$G_{\rm ann} f(y) := \sum_i \tfrac{1}{2d} \sum_{j:\,|i-j|=1} \big(f(a_{ij}(y)) - f(y)\big), \tag{4.14}$$

where

$$a_{ij}(y)(k) := \begin{cases} 0 & \textit{if } k = i, \\ y(i) + y(j) \mod(2) & \textit{if } k = j, \\ y(k) & \textit{otherwise.} \end{cases} \tag{4.15}$$

Then

$$\mathbb{P}\big[|X_t Y_0| \textit{ is odd}\big] = \mathbb{P}\big[|X_0 Y_t| \textit{ is odd}\big] \qquad (t \geq 0) \tag{4.16}$$

whenever X and Y are independent and either X_0 or Y_0 are a.s. finite.

Proof As in the previous proof, set $A := \{i : y(i) = 1\}$. Now

$$Y_s(i) := 1_{\{\xi^{(j,t)}_s = i \text{ for an odd number of } j \in A\}} \qquad (s \geq 0) \tag{4.17}$$

defines a Markov process with generator as in (4.14). It follows that

$$\begin{aligned} \mathbb{P}\big[|X_t y| \text{ is odd}\big] &= \mathbb{P}[x(\xi^{(j,t)}_t) = 1 \text{ for an odd number of } j \in A] \\ &= \mathbb{P}\big[|X_0 Y_t| \text{ is odd}\big] \qquad (t \geq 0). \end{aligned} \tag{4.18}$$

∎

Remark We may rewrite (4.11) as

$$\mathbb{E}[0^{|X_t Y_0|}] = \mathbb{E}[0^{|X_0 Y_t|}] \qquad (t \geq 0) \tag{4.19}$$

and (4.16) as

$$\mathbb{E}[(-1)^{|X_t Y_0|}] = \mathbb{E}[(-1)^{|X_0 Y_t|}] \qquad (t \geq 0). \tag{4.20}$$

More generally, for many nearest-neighbor interacting particle systems, there exist duals with respect to a duality function of the form $\Psi(x, y) = \eta^{|xy|}$, where η is a real parameter; see [SL95, SL97, Sud00].

4.4 Clustering versus stability

In this section, we study the voter model started in product initial laws. Thus, we will assume that the $(X_0(i))_{i\in\mathbb{Z}^d}$ are i.i.d. with intensity $\mathbb{P}[X_0(i)=1]=\theta\in[0,1]$.

Theorem 4.4 (Process started in product law) *Let X be a d-dimensional voter model started in product law with intensity $0\leq\theta\leq 1$. Then*

$$\mathbb{P}[X_t\in\cdot\,]\underset{t\to\infty}{\Longrightarrow}\nu_\theta, \tag{4.21}$$

where ν_θ is an invariant law of the process. If $d=1,2$, then

$$\nu_\theta=(1-\theta)\delta_0+\theta\delta_1. \tag{4.22}$$

On the other hand, in dimensions $d\geq 3$, for $0<\theta<1$ the measures ν_θ are concentrated on configurations that are not constant.

In dimensions $d=1,2$, this theorem says that in any finite environment of the origin, at sufficient large times, we see with high probability locally either the constant configuration 0 or 1. This implies that in the system, there must be large regions constant type, called *clusters*, of a size that grows in time. This behavior is called *clustering*.

On the other hand, in dimensions $d>2$, it can be shown that the measures ν_θ are concentrated on configurations in which the spatial intensity of ones (averaged over large blocks) is θ. This type of behavior is called *stable* behavior.

Proof of Theorem 4.4 We use duality with coalescing random walks. For each $y\in\{0,1\}^{\mathbb{Z}^d}$ with $|y|<\infty$, let Y^y denote the process with generator as in (4.9). Since $|Y^y_t|$ is a nonincreasing function of time, the limit

$$N^y:=\lim_{t\to\infty}|Y^y_t| \tag{4.23}$$

exists a.s. It follows that

$$\mathbb{P}\big[|X_ty|=0\big]=\mathbb{P}\big[|X_0Y^y_t|=0\big]=\mathbb{E}\big[(1-\theta)^{|Y^y_t|}\big]\underset{t\to\infty}{\longrightarrow}\mathbb{E}\big[(1-\theta)^{N^y}\big]. \tag{4.24}$$

Since the space $\{0,1\}^{\mathbb{Z}^d}$ is compact, the laws $\mathbb{P}[X_t\in\cdot\,]$ are automatically tight, so to prove convergence it suffices that all weak cluster points coincide. Therefore, by Lemma 2.2 we see that $\mathbb{P}[X_t\in\cdot\,]$ converges weakly to a probability law ν_θ which is uniquely characterized by

$$\int\nu_\theta(\mathrm{d}x)1_{\{|xy|=0\}}=\mathbb{E}\big[(1-\theta)^{N^y}\big]\qquad(|y|<\infty). \tag{4.25}$$

It follows from general arguments (which we now skip) that any long-time limit law of an interaction particle system must be an invariant law. In dimensions $d = 1, 2$, one has $N^y = 1$ a.s., which implies that $\nu_\theta = (1-\theta)\delta_0 + \theta\delta_1$. On the other hand, in dimensions $d > 2$, by the transience of random walk, $\mathbb{P}[N^y \geq 2] > 0$ for all $|y| \geq 2$. In fact, it is not hard to see (again we skip the details) that by choosing a configuration y in which all ones are sufficiently far from each other, for each $n \geq 1$ and $\varepsilon > 0$ we can find y with $|y| = n$ and $\mathbb{P}[N^y = n] \geq 1-\varepsilon$. It follows that $\nu_\theta(\{0\}) \leq \int \nu_\theta(\mathrm{d}x) 1_{\{|xy|=0\}} \leq (1-\varepsilon)(1-\theta)^n + \varepsilon$. Since n and ε are arbitrary, if $0 < \theta$ this proves that ν_θ gives zero probability to the constant configuration 0. By symmetry between the types, if $\theta < 1$, then ν_θ gives moreover zero probability to the constant configuration 1. ∎

Remark Let $\mathcal{I}$ denote the set of all invariant laws of the voter model. It is not hard to show that $\mathcal{I}$ is a compact, convex subset of the set of all probability measures on $\{0,1\}^{\mathbb{Z}^d}$. By definition, an element $\nu \in \mathcal{I}$ is called *extremal* if it cannot be written as a nontrivial convex combination of other elements of $\mathcal{I}$, i.e., there do not exist $\nu_1, \nu_2 \in \mathcal{I}$ with $\nu_1 \neq \nu_2$ and $0 < p < 1$ such that $\nu = p\nu_1 + (1-p)\nu_2$. We let $\mathcal{I}_\mathrm{e}$ denote the set of all extremal elements of $\mathcal{I}$. Then it is known that for voter models in dimensions $d > 2$, the measures ν_θ from (4.21) satisfy $\nu_\theta \in \mathcal{I}_\mathrm{e}$ for all $\theta \in [0, 1]$.

4.5 Selection and mutation

In this section, we consider a generalization of the voter model with generator

$$\begin{aligned} Gf(x) := (1+s)\sum_i \tfrac{1}{2d} \sum_{j:\,|i-j|=1} & 1_{\{x(i)=0,\ x(j)=1\}}\big(f(x^{\{i\}}) - f(x)\big) \\ + \sum_i \tfrac{1}{2d} \sum_{j:\,|i-j|=1} & 1_{\{x(i)=1,\ x(j)=0\}}\big(f(x^{\{i\}}) - f(x)\big) \\ + m\sum_i & 1_{\{x(i)=1\}}\big(f(x^{\{i\}}) - f(x)\big), \end{aligned} \tag{4.26}$$

where $s, m \geq 0$ are constants. Note that this says that sites of type 0 adopt the type 1 at a rate that is $\frac{1}{2d}(1+s)$ times the number of neighbors of type 1, but sites of type 1 adopt the type 0 at a rate that is only $\frac{1}{2d}$ times the number of neighbors of type 0. Thus, sites of type 1 invade other sites more easily than sites of type 0, which we can interpret as saying that type 1 has a higher fitness than the other type. In addition, sites of type 1 spontaneously change into sites of type 0, which models mutation. For example, we can imagine that organisms of type 1 own a complex gene that gives them a selective advantage, while organisms of type

0 own a damaged, nonfunctional version of the same gene. We ignore mutations that spontaneously restore the functionality of a damaged gene. We call s and m the *selection* and *mutation* rates, respectively.

In order to write the generator in (4.26) in terms of local maps, we define

$$s_{ij}(x)(k) := \begin{cases} 1 & \text{if } k=j,\ x(i)=1, \\ x(k) & \text{otherwise,} \end{cases} \tag{4.27}$$

and

$$p_i(x)(k) := \begin{cases} 0 & \text{if } k=i, \\ x(k) & \text{otherwise.} \end{cases} \tag{4.28}$$

Then

$$\begin{aligned} Gf(x) = &\sum_i \tfrac{1}{2d} \sum_{j:\,|i-j|=1} \big(f(m_{ij}(x)) - f(x)\big) \\ &+ s\sum_i \tfrac{1}{2d} \sum_{j:\,|i-j|=1} \big(f(s_{ij}(x)) - f(x)\big) \\ &+ m\sum_i \big(f(p_i(x)) - f(x)\big). \end{aligned} \tag{4.29}$$

In the graphical representation, we represent the maps m_{ij} as before by an arrow (black, let us say) from i to j. Likewise, we represent the maps s_{ij} by red arrows and the maps p_i by a black box. Then black arrows indicate 'invasion' events where the type at the beginning of the arrow takes over the site at the tip of the arrow. Red arrows are similar, except that they can only be used for the invasion of type 1. The black boxes indicate mutations where the type changes to 0 regardless of what it was before.

It turns out that the coalescent dual of the voter model can be generalized to the present more general set-up. Let c_{ij} be defined as in (4.10) and define local maps b_{ij} by

$$b_{ij}(y)(k) := \begin{cases} y(i) \vee y(j) & \text{if } k=j, \\ y(k) & \text{otherwise.} \end{cases} \tag{4.30}$$

Lemma 4.5 (Branching-coalescent dual) *Let X be a voter model with selection and mutation, with generator as in (4.26), and let Y be the Markov process with values in $\{0,1\}^{\mathbb{Z}^d}$ and generator*

$$\begin{aligned} G_{\rm coal}f(y) := &\sum_i \tfrac{1}{2d} \sum_{j:\,|i-j|=1} \big(f(c_{ij}(y)) - f(y)\big) \\ &+ s\sum_i \tfrac{1}{2d} \sum_{j:\,|i-j|=1} \big(f(b_{ij}(y)) - f(y)\big) \\ &+ m\sum_i \big(f(p_i(y)) - f(y)\big). \end{aligned} \tag{4.31}$$

Then

$$\mathbb{P}\big[|X_tY_0| \neq 0\big] = \mathbb{P}\big[|X_0Y_t| \neq 0\big] \qquad (t \geq 0) \tag{4.32}$$

whenever X and Y are independent.

Proof When we try to follow lines of descent backwards in time as for the usual voter model, we run into the difficulty that when we encounter the tip of a red arrow, we do not know which site is the true ancestor of our site. We do know, however, that if either the site at the beginning or at the tip of the arrow was of type 1 before, then the site at the tip will be of type 1 afterwards. Thus, the solution is to follow the ancestry of both sites backward in time. In this way, we arrive at a dual process consisting of coalescing random walks with additional branching events corresponding to the red arrows and deaths corresponding to mutations.

More formally, the proof goes as follows. It suffices to prove the statement for deterministic initial states $X_0 = x$ and $Y_0 = y$. Let us say that a path $\gamma : [s,u] \to \mathbb{Z}^d$ is *open* if it satisfies the following rules:

$$\begin{array}{rl} \text{(i)} & \text{If } \gamma_{t-1} \neq \gamma_t, \text{ then there must be an arrow from } (\gamma_{t-}, t) \text{ to } (\gamma_t, t), \\ \text{(ii)} & \text{if at time } t \text{ there is a black arrow from } i \text{ to } j \text{ and } \gamma_t = j, \text{ then } \gamma_{t-} = i, \\ \text{(iii)} & \text{there is no black box at } (\gamma_t, t) \text{ for any } t \in [s,u]. \end{array} \tag{4.33}$$

Write $(i,s) \rightsquigarrow (j,u)$ if there is an open path $\gamma : [s,t] \to \mathbb{Z}^d$ such that $\gamma_{s-} = i$ and $\gamma_u = j$. Then

$$Y_s(i) := 1_{\{\exists j \text{ s.t. } y(j)=1,\ (i,t-s) \rightsquigarrow (j,t)\}} \tag{4.34}$$

defines a Markov process with generator as in (4.31). Thus, (4.32) follows from the claim that

$$|X_tY_0| \neq 0 \quad \text{if and only if} \quad \exists i,j \text{ s.t. } x(i)=1,\ y(j)=1,\ (i,0) \rightsquigarrow (j,t). \tag{4.35}$$

Indeed, if there is an open path from $(i,0)$ to (j,t) with $x(i) = 1$, then it is easy to check that X must be one all along this path, hence in particular $X_t(j) = 1$. Conversely, if no such path exists, then all 'potential ancestors' of the site j at time t can be traced back to sites that were of type 0 at time zero or to mutation events. ∎

Both the voter model X with selection and mutation and its dual system Y of random walks with coalescence, branching and deaths are very similar in their behavior to the contact process. If the mutation rate is too high, then the processes started in any finite initial state die out and the delta measure on the constant zero configuration is the only invariant law. If the mutation rate is sufficiently low (but $s > 0$), then the processes started in a finite initial state survives with positive probability and there exists a nontrivial invariant law.

Exercise 4.6 Show that the standard voter model ($s = m = 0$) started in a finite initial state (i.e., $|X_0| < \infty$ a.s.) dies out a.s. (i.e., $\mathbb{P}[\exists t \geq 0 \text{ s.t. } X_t = 0] = 1$). Hint: martingale convergence.

4.6 Rebellious voter models

In the biological interpretation of the voter model, let us imagine that the types 0 and 1 represent two closely related species, which compete for space. We may imagine that the cause of death of organisms of these species is often competition with organisms on neighboring sites, and that organisms experience more competition from other organisms of their own species than from those of the other species. The explanation is that the species occupy slightly different ecological niches, hence each species has some resources not available to the other species, which reduces competition between different species. To model this effect, we define local maps

$$r_{ijk}(y)(l) := \begin{cases} y(i) & \text{if } l = j \text{ and } y(j) = y(k), \\ y(l) & \text{otherwise,} \end{cases} \tag{4.36}$$

which says that if the organisms living at j and k are of the same type, then the organism at j dies and is replaced by an organism of the type living at i. Then we are interested in the Markov process with generator

$$\begin{aligned} Gf(x) := &\sum_i \tfrac{1}{2d} \sum_{j:\,|i-j|=1} \big(f(m_{ij}(x)) - f(x)\big) \\ &+ 2\beta \sum_i \tfrac{1}{(2d)^2} \sum_{j:\,|i-j|=1} \sum_{k:\,|j-k|=1} \big(f(r_{ijk}(x)) - f(x)\big), \end{aligned} \tag{4.37}$$

where $2\beta \geq 0$ is the extra death rate due to competition with organisms of the same species. Up to a trivial rescaling of time and renaming of parameters, this model is a special case of a model introduced in [NP99]. This and similar models have also been studied in [SS08, SV10]. In the traditional interpretation of voter models, the model in (4.37) describes a population where individuals like to disagree with their neighbors. In view of this, a one-dimensional model similar to the one in (4.37) has been called the 'rebellious voter model' in [SS08].

It turns out that the annihilating dual of the voter model can be generalized to the present more general set-up. To see this, we start by noting that the generator in (4.37) can be rewritten as

$$\begin{aligned} Gf(x) :=& \sum_i \tfrac{1}{2d} \sum_{j:\,|i-j|=1} \big(f(m_{ij}(x)) - f(x)\big) \\ &+\beta \sum_i \tfrac{1}{(2d)^2} \sum_{j:\,|i-j|=1} \sum_{k:\,|j-k|=1} \big(f(q_{ijk}(x)) - f(x)\big), \end{aligned} \tag{4.38}$$

where q_{ijk} is the local map

$$q_{ijk}(x)(l) := \begin{cases} x(i) + x(j) + x(k) \mod(2) & \text{if } l = j, \\ x(l) & \text{otherwise,} \end{cases} \tag{4.39}$$

which combines the effects of r_{ijk} and r_{kji}. Let a_{ij} be defined as in (4.15) and define local maps t_{ijk} by

$$t_{ijk}(y)(l) := \begin{cases} y(l) + y(j) \mod(2) & \text{if } l \in \{i, k\} \\ y(l) & \text{otherwise.} \end{cases} \tag{4.40}$$

Then we can prove the following result.

Lemma 4.7 (Double branching-annihilating dual) *Let X be a 'rebellious' voter model with generator as in (4.38), and let Y be the Markov process with values in $\{0,1\}^{\mathbb{Z}^d}$ and generator*

$$\begin{aligned} G_{\rm dbran} f(y) :=& \sum_i \tfrac{1}{2d} \sum_{j:\,|i-j|=1} \big(f(a_{ij}(y)) - f(y)\big) \\ &+\beta \sum_i \tfrac{1}{(2d)^2} \sum_{j:\,|i-j|=1} \sum_{k:\,|j-k|=1} \big(f(t_{ijk}(y)) - f(y)\big). \end{aligned} \tag{4.41}$$

$$\mathbb{P}\big[|X_t Y_0| \text{ is odd}\big] = \mathbb{P}\big[|X_0 Y_t| \text{ is odd}\big] \qquad (t \geq 0) \tag{4.42}$$

whenever X and Y are independent and either X_0 or Y_0 are a.s. finite.

Proof In the graphical representation, we denote the local map q_{ijk} by drawing two red arrows, starting at i and k, respectively, and both ending at j. Drawing the usual arrows of the voter model dynamics in black, we call a path $\gamma : [s,u] \to \mathbb{Z}^d$ *open* if it satisfies the following rules:

$$\begin{array}{rl} \text{(i)} & \text{If } \gamma_{t-1} \neq \gamma_t, \text{ then there must be an arrow from } (\gamma_{t-}, t) \text{ to } (\gamma_t, t), \\ \text{(ii)} & \text{if at time } t \text{ there is a black arrow from } i \text{ to } j \text{ and } \gamma_t = j, \text{ then } \gamma_{t-} = i. \end{array} \tag{4.43}$$

It suffices to prove the statement of the lemma for deterministic initial states $X_0 = x$ and $Y_0 = y$. Set $A := \{i : x(i) = 1\}$ and $B := \{i : y(i) = 1\}$, fix $t > 0$ and define

$$Y_s(i) := 1_{\{\text{the number of open paths from } (i, t-s) \text{ to } B \times \{t\} \text{ is odd}\}} \tag{4.44}$$

$(s \geq 0)$. Then $(Y_s)_{s\geq 0}$ is a Markov process with generator as in (4.41). It follows that

$$\begin{aligned} &|X_t y| \text{ is odd} \\ &\quad\Leftrightarrow \quad \text{the number of open paths from } A \times \{0\} \text{ to } B \times \{t\} \text{ is odd} \\ &\quad\Leftrightarrow \quad |xY_t| \text{ is odd,} \end{aligned} \tag{4.45}$$

which proves (4.42). ∎

Note that the process Y with generator a in (4.41) describes a system of annihilating random walks, where in addition, with rate β, particles branch by producing two new particles on their neighboring sites, which immediately annihilate with any particles that may already be present on these sites. Note that since the number of particles always increases or decreases by a multiple of 2, the process Y is *parity preserving*, i.e., if $|y|$ is finite and odd (resp. even), then $|Y_t|$ is odd (resp. even) at all $t \geq 0$. In view of this, the process started with an odd number of particles cannot die out. The following facts are known:

- The process X exhibits *coexistence*, i.e., there is an invariant law concentrated on configurations that are not constant 0 or 1, if and only if the process Y, started with an even number of particles, survives with positive probability.

- In dimension $d = 1$, the process Y started in an even initial state dies out a.s. for any $\beta \geq 0$.

- In dimensions $d \geq 2$, and also for non-nearest neighbor processes in dimension $d = 1$, if β is sufficiently large, then the process Y started in an even initial state survives with positive probability.

- In dimensions $d \geq 2$, if β is sufficiently *small*, then the process Y started in an even initial state survives with positive probability.

The last statement, which has been proved in a series of rather long and technical papers by Cox and Perkins [CP05, CP06, CMP10], is supposed to be false for non-nearest neighbor processes in dimension $d = 1$, but here is no proof of this. Also, perhaps surprisingly, it is not known if survival of the even process Y is monotone in the branching rate β, hence in dimensions $d \geq 2$ is is also not known if the even process survives for intermediate values of β (although both statements are believed to be true). In general, the study of rebellious voter models and their duals is made difficult by the fact that they are not monotonous.

Bibliography

[AS09a] S.R. Athreya and J.M. Swart. Branching-coalescing particle systems. Erratum. *Prob. Theory Relat. Fields.* 145(3-4), 639–640, 2009.

[BFM78] R.C. Brower, M.A. Furman and M. Moshe. Critical exponents for the reggeon quantum spin model. *Phys. Lett. B* 76, 213–219, 1978.

[BG90] C. Bezuidenhout and G. Grimmett. The critical contact process dies out. *Ann. Probab.* 18(4), 1462–1482, 1990.

[BG91] C. Bezuidenhout and G. Grimmett. Exponential decay for subcritical contact and percolation processes. *Ann. Probab.* 19(3), 984–1009, 1991.

[CMP10] J.T. Cox, M. Merle and E.A. Perkins. Coexistence in a two-dimensional Lotka-Volterra model. *Electron. J. Probab.* 15, paper No. 38, 1190–1266, 2010.

[CP05] J.T. Cox and E.A. Perkins. Rescaled Lotka-Volterra models converge to super-Brownian motion. *Ann. Probab.* 33(3), 904–947, 2005.

[CP06] J.T. Cox and E.A. Perkins. Survival and coexistence in stochastic spatial Lotka-Volterra models. *Probab. Theory Relat. Fields* 139(1/2), 89–142, 2007.

[Dob71] R.L. Dobrushin. Markov processes with a large number of locally interacting components: existence of a limit process and its ergodicity. *Problems Inform. Transmission* 7, 149–164, 1971.

[Dur88] R. Durrett. *Lecture notes on particle systems and percolation.* Wadsworth & Brooks/Cole, Pacific Grove, 1988.

[EK86] S.N. Ethier and T.G. Kurtz. *Markov Processes: Characterization and Convergence.* John Wiley & Sons, New York, 1986.

[Gla63] R.J. Glauber. Time-dependent statistics of the Ising model. *J. Math. Phys.* 4, 294–307, 1963.

[Har74] T.E. Harris. Contact interactions on a lattice. *Ann. Probab.* 2, 969–988, 1974.

[Har76] T.E. Harris. On a class of set-valued Markov processes. *Ann. Probab.* 4, 175–194, 1976.

[HS05] R. van der Hofstad and A. Sakai. Critical points for spread-out self-avoiding walk, percolation and the contact process above the upper critical dimensions. *Probab. Theory Relat. Fields* 132(3), 438–470, 2005.

[Isi25] E. Ising. Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik* 31, 253–258, 1925.

[Lig85] T.M. Liggett. *Interacting Particle Systems.* Springer-Verlag, New York, 1985.

[Lig92] T.M. Liggett. The survival of one-dimensional contact processes in random environments. *Ann. Probab.* 20, 696-723, 1992.

[Lig95] T.M. Liggett. Improved upper bounds for the contact process critical value. *Ann. Probab.* 23, 697–723, 1995.

[Lig99] T.M. Liggett. *Stochastic interacting systems: contact, voter and exclusion processes.* Springer-Verlag, Berlin, 1999.

[LSS97] T.M. Liggett, R.H. Schonmann and A.M. Stacey. Domination by product measures. *Ann. Probab.* 25(1), 71–95, 1997.

[Loe63] M. Loève. *Probability Theory, 3rd edition.* Van Nostrand, Princeton, 1963.

[NP99] C. Neuhauser and S.W. Pacala. An explicitly spatial version of the Lotka-Volterra model with interspecific competition. *Ann. Appl. Probab.* 9(4): 1226–1259, 1999.

[Ons44] L. Onsager. Crystal statistics, I. A two-dimensional model with an order-disorder transition. *Phys. Rev.*65, 117–149, 1944.

[Pei36] R. Peierls. On Ising's model of ferromagnetism. *Proc. Camebridge Philos. Soc.* 36, 477-481, 1936.

[Rem08] D. Remenik. The contact process in a dynamic random environment. *Ann. Appl. Probab.* 18(6), 2392-2420, 2008.

[SL95] A. Sudbury and P. Lloyd. Quantum operators in classical probability theory. II: The concept of duality in interacting particle systems. *Ann. Probab.* 23(4), 1816–1830, 1995.

[SL97] A. Sudbury and P. Lloyd. Quantum operators in classical probability theory. IV: Quasi-duality and thinnings of interacting particle systems *Ann. Probab.* 25(1), 96–114, 1997.

[SS08] A. Sturm and J.M. Swart. Voter models with heterozygosity selection. *Ann. Appl. Probab.* 18(1), 59-99, 2008.

[Sud00] A. Sudbury. Dual families of interacting particle systems on graphs. *J. Theor. Probab.* 13(3), 695–716, 2000.

[SV10] J.M. Swart and K. Vrbenský. Numerical analysis of the rebellious voter model. *J. Stat. Phys.* 140(5), 873–899, 2010.

[ZG88] H. Ziezold and C. Grillenberger. On the critical infection rate of the one-dimensional basic contact process: numerical results. *J. Appl. Prob.* 25, 1–8, 1988.

Index