A Course in Interacting Particle Systems

J.M. Swart

October 23, 2015

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 forty years, some easy-to-formulate problems still remain open while the solutions of others have 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, that 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.

The present notes are loosely based on an older set of lecture notes for courses that I gave at Charles University in Prague in 2009 and 2011. Another imput came from slides for a course I gave at Verona University in 2014. Compared to the lecture notes of 2011, most of the text has been rewritten. Many figures have been added, as well as a chapter on the mean-field limit. The old lecture notes were organized around three classical models: the contact process, the Ising model, and the voter model. Instead, the present notes are organized around methods: the mean-field limit, graphical representations, monotone coupling, duality, and comparison with oriented percolation. Compared to the older notes, some results have been removed, in particular about the Ising model, whose study requires rather different techniques from the other models. Another omission are positive correlations. On the other hand, a wide range of interacting particle systems not (or barely) mentioned in the previous lecture notes are now used as examples throughout the notes, to give a better impression of the modern literature of the subject.

I am indebted to Tibor Mach for a careful reading of the lecture notes from 2011 that led to a large number of typoes being corrected. For all new errors introduced in the present text only I am responsible.

Chapter 1

Introduction

1.1 General set-up

Let S be a finite set, called the *local state space*, and let Λ be a countable set, called the *lattice*. We let S^Λ denote the Carthesian product space of Λ copies of S, i.e., elements x of S^Λ are of the form

$$x = \big(x(i)\big)_{i\in\Lambda} \quad \text{with} \quad x(i) \in S \ \forall\, i \in \Lambda.$$

Equivalently, S^Λ is nothing else than the set of all functions $x : \Lambda \to S$.

Interacting particle systems are continuous-time Markov processes $X = (X_t)_{t\geq 0}$ with a state space of the form S^Λ, that are defined in terms of *local maps*. Thus, $(X_t)_{t\geq 0}$ is a Markov process such that at each time $t \geq 0$, the state of X is of the form

$$X_t = \big(X_t(i)\big)_{i\in\Lambda} \quad \text{with} \quad X_t(i) \in S \ \forall\, i \in \Lambda.$$

We call $X_t(i)$ the *local state* of X at time t and at the *position* i. Positions $i \in \Lambda$ are also often called *sites*.

The time evolution of continuous-time Markov processes is usually characterized by their *generator* G, which is an operator acting on functions $f : \mathcal{S} \to \mathbb{R}$, where $\mathcal{S}$ is the state space. For example, in the case of Brownian motion, the state space is $\mathbb{R}$ and the generator is the differential operator $G = \frac{1}{2}\frac{\partial^2}{\partial x^2}$. In the case of an interacting particle system, the state space is of the form $\mathcal{S} = S^\Lambda$ and the generator takes the form

$$Gf(x) = \sum_{m\in\mathcal{G}} r_m \big\{ f\big(m(x)\big) - f\big(x\big) \big\} \qquad (x \in S^\Lambda). \tag{1.1}$$

Here $\mathcal{G}$ is a set whose elements are *local maps* $m : S^\Lambda \to S^\Lambda$ and $(r_m)_{m\in\mathcal{G}}$ is a collection of nonnegative constants called *rates*, that say with which

Poisson intensity the local map m should be applied to the configuration X_t. The precise definitions will be given in later chapters, but at the moment it suffices to say that if we approximate $(X_t)_{t\geq 0}$ by a discrete-time Markov chain where time is increased in steps of size $\mathrm{d}t$, then

$$r_m\,\mathrm{d}t \quad \text{is the probability that the map } m \text{ is applied during the time interval } (t, t+\mathrm{d}t].$$

Often, the lattice Λ has the structure of an (undirected) graph. In this case, we let E denote the corresponding *edge set*, i.e., a set of unordered pairs $\{i,j\}$ called *edges*, with $i,j \in \Lambda$, $i \neq j$, that in drawings of the graph are connected by a line segment. We let

$$\mathcal{E} := \big\{(i,j) : \{i,j\} \in E\big\}$$

denote the corresponding set of all *ordered* pairs (i,j) that correspond to an edge. We call

$$\mathcal{N}_i := \big\{j \in \Lambda : \{i,j\} \in E\big\} \tag{1.2}$$

the *neighborhood* of the site i.

Many well-known and well-studied interacting particle systems are defined on the *d-dimensional integer lattice* $\mathbb{Z}^d$. We denote the origin by $0 = (0,\ldots,0) \in \mathbb{Z}^d$. For any $i = (i_1,\ldots,i_d) \in \mathbb{Z}^d$, we let

$$\|i\|_1 := \sum_{k=1}^d |i_k| \quad \text{and} \quad \|i\|_\infty := \max_{k=1,\ldots,d} |i_k| \qquad (i \in \mathbb{Z}^d)$$

denote the ℓ_1-norm and supremumnorm, respectively. For $R \geq 1$, we set

$$E^d := \big\{\{i,j\} : \|i-j\|_1 = 1\big\} \quad \text{and} \quad E^d_R := \big\{\{i,j\} : 0 < \|i-j\|_\infty \leq R\big\}. \tag{1.3}$$

Then $(\mathbb{Z}^d, E^d)$ is the integer lattice equipped with the *nearest neighbor* graph structure and $(\mathbb{Z}^d, E^d_R)$ is the graph obtained by connecting all edges within $\|\cdot\|_\infty$-diestance R with an edge. We let $\mathcal{E}^d$ and $\mathcal{E}^d_R$ denote the corresponding sets of ordered pairs (i,j).

Before we turn to rigorous mathematical theory, it is good to see a number of examples. It is easy to simulate interacting particle systems on a computer. In simulations, the infinite graphs $(\mathbb{Z}^d, E^d)$ or $(\mathbb{Z}^d, E^d_R)$ are replaced by a finite piece of $\mathbb{Z}^d$, with some choice of the boundary conditions (e.g. periodic boundary conditions).

1.2 The voter model

For each $i, j \in \Lambda$, the *voter model map* $\mathrm{vot}_{ij} : S^\Lambda \to S^\Lambda$ is defined as

$$\texttt{vot}_{ij}(x)(k) := \begin{cases} x(i) & \text{if } k = j, \\ x(k) & \text{otherwise.} \end{cases} \tag{1.4}$$

Applying $\texttt{vot}_{ij}$ to a configuration x has the effect that local state of the site i is copied onto the site j. The *nearest neighbor voter model.* is the interacting particle system with generator

$$G_{\mathrm{vot}} f(x) = \frac{1}{|\mathcal{N}_0|} \sum_{(i,j) \in \mathcal{E}^d} \big\{ f\big(\texttt{vot}_{ij}(x)\big) - f\big(x\big) \big\} \qquad (x \in S^\Lambda). \tag{1.5}$$

Here $\mathcal{N}_0$ is the neighborhood of the origin and $|\mathcal{N}_0| = 2d$ denotes its cardinality. Similarly, replacing the set of oriented edges $\mathcal{E}^d$ by $\mathcal{E}^d_R$ and replacing $\mathcal{N}_0$ by the appropriate set of neighbors in this new graph, we obtain the *range* R voter model.

In the context of the voter model, the local state $x(i)$ at a site i is often called the *type* at i. The voter model is often used to model biological populations, where organisms with different genetic types occupy sites in space. Note that since each site j has $|\mathcal{N}_j| = |\mathcal{N}_0|$ neighbors, the total rate of all maps $\texttt{vot}_{ij}$ with $i \in \mathcal{N}_j$ is one. In view of this, an alternative way to describe the dynamics in (1.5) is to say that with rate 1, the organism living at a given site dies, and is replaced by a descendant chosen with equal probability from its neighbors.

An alternative interpretation, that has given the voter model its name, is that sites represent people and types represent political opinions. With rate one, an individual becomes unsure what political party to vote for, asks a randomly chosen neighbor, and copies his/her opinion.

In Figure 1.1, we see the four snapshots of the time evolution of a two-dimensional nearest-neighbor voter model. The initial state is constructed by assigning i.i.d. types to the sites. Due to the copying dynamics, we see patches appear where every site in a local neighborhood has the same type. As time proceeds, these patches, usually called *clusters*, grow in size, so that eventually, for any $N \geq 1$, the probability that all sites within distance N of the origin are of the same type tends to one.[1]

It turns out that this sort of behavior, called *clustering*, is dimension dependent. The voter model clusters in dimensions 1 and 2, but not in

[1] In spite of this, for the model on the infinite lattice, it is still true that the origin changes its type infinitely often.

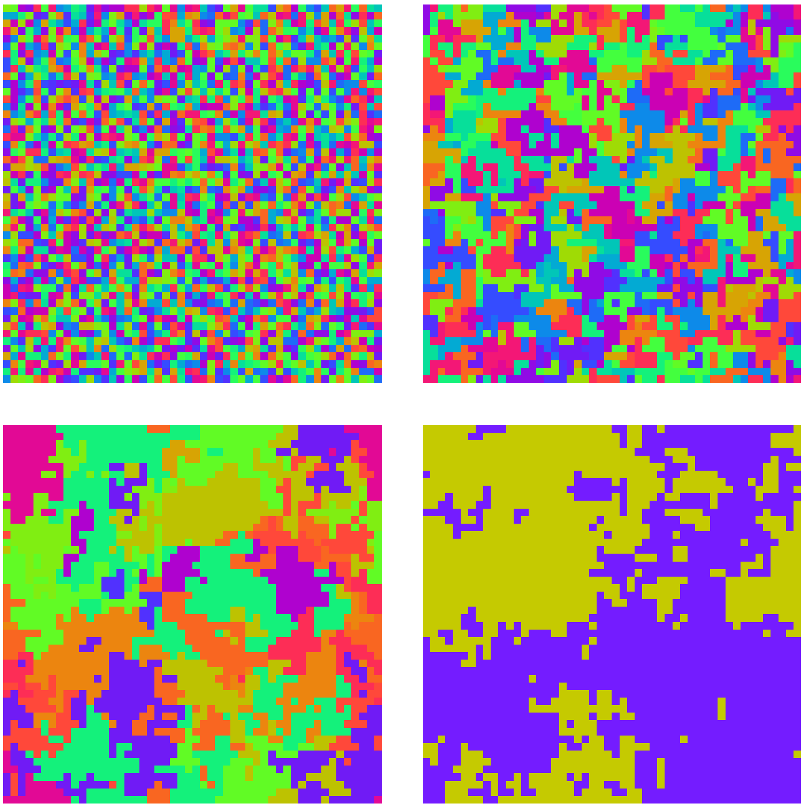


Figure 1.1: Four snapshots of a two-dimensional voter model with periodic boundary conditions. Initially, the types of sites are i.i.d. Time evolved in these pictures is 0, 1, 32, and 500.

dimensions 3 and more. In Figure 1.2, we see the four snapshots of the time evolution of a three-dimensional voter model. The model is simulated on a cube with periodic boundary conditions, and the types of the middle layer are shown in the pictures. In this case, we see that even after a long time, there are still many different types near the origin.[2]

[2]On a finite lattice, such as we use in our simulations, one would eventually see one type take over, but the time one has to wait for this is very long compared to dimensions 1 and 2. On the infinite lattice, the probability that the origin has a different type from its right neighbor tends to a positive limit as time tends to infinity.

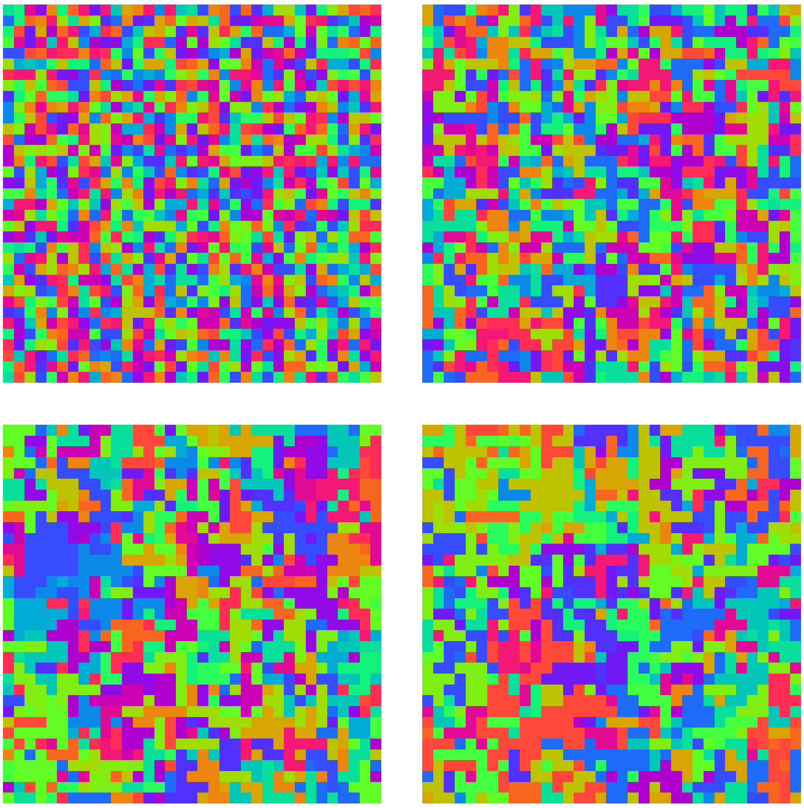


Figure 1.2: Four snapshots of the transsection of a three-dimensional voter model with periodic boundary conditions. Initially, the types of sites are i.i.d. Time evolved in these pictures is 0, 4, 32, and 250.

1.3 The contact process

The contact process is another interacting particle system with a biological interpretation. For this process, we choose the local state space $S = \{0, 1\}$. We interpret a site such that $X_t(i) = 1$ as *occupied* by an organism, and a site such that $X_t(i) = 0$ as *empty*. Alternatively, the contact process can be seen as a model for the spread of an infection. In this case, sites with $X_t(i) = 1$ are called *infected* and sites with $X_t(i) = 0$ are called *healthy*.

For each $i, j \in \Lambda$, we define a *branching map* $\mathtt{bra}_{ij} : \{0,1\}^\Lambda \to \{0,1\}^\Lambda$ as

$$\mathtt{bra}_{ij}(x)(k) := \begin{cases} x(i) \vee x(j) & \text{if } k = j, \\ x(k) & \text{otherwise.} \end{cases} \tag{1.6}$$

Note that this says that if prior to the application of $\mathtt{bra}_{ij}$, the site i is occupied, then after the application of $\mathtt{bra}_{ij}$, the site j will also be occupied, regardless of its previous state. If initially i is empty, then nothing happens. We interpret this as the organism at i giving *birth* to a new organism at j, or the infected site i *infecting* the site j. If j is already occupied/infected, then nothing happens.

For each $i \in \Lambda$, we also define a *death map* $\texttt{death}_i : \{0,1\}^\Lambda \to \{0,1\}^\Lambda$ as

$$\texttt{death}_i(x)(k) := \begin{cases} 0 & \text{if } k = i, \\ x(k) & \text{otherwise.} \end{cases} \tag{1.7}$$

If the map $\texttt{death}_i$ is applied, then an organism at i, if there is any, dies, respectively, the site i, if it is infected, *recovers* from the infection.

Figure 1.3: Four snapshots of a two-dimensional contact process. Initially, only a single site is infected. The infection rate is 2, the death rate is 1, and time evolved in these pictures is 1, 5, 10, and 20.

Recalling (1.3), the (nearest neighbor) contact process with *infection rate*

$\lambda \geq 0$ and *death rate* $\delta \geq 0$ is the interacting particle system with generator

$$\begin{aligned} G_{\rm cont} f(x) := \lambda \sum_{(i,j)\in\mathcal{E}^d} & \big\{ f\big(\big(\mathtt{bra}_{ij}(x)\big) - f\big(x\big) \big\} \\ +\delta \sum_{i\in\mathbb{Z}^d} & \big\{ f\big(\big(\mathtt{death}_i(x)\big) - f\big(x\big) \big\} \qquad (x \in \{0,1\}^{\mathbb{Z}^d}). \end{aligned} \tag{1.8}$$

This says that infected sites infect each healthy neighbor with rate λ, and infected sites recover with rate δ.

In Figure 1.3, we see the four snapshots of the time evolution of a two-dimensional contact process. Occupied sites are black and empty sites are white. Initially, only the origin is occupied. The infection rate is 2 and the death rate is 1. In this example, the infection spreads through the whole population, eventually reaching a steady state[3] where a positive fraction of the population is infected. Of course, starting from a single infected site, there is always a positive probability that the infection dies out in the initial stages of the epidemy.

Unlike the voter model, the behavior of the contact process is roughly similar in different dimensions. On the other hand, the proportion λ/δ of the infection rate to the death rate is important for the behavior. By changing the speed of time, we can without loss of generality choose one of the constants λ and δ to be one, and it is customary to set $\delta := 1$. In Figure 1.4, we have plotted the *survival probability*

$$\theta(\lambda) := \mathbb{P}^{1_{\{0\}}}[X_t \neq 0 \ \forall t \geq 0] \tag{1.9}$$

of the one-dimensional contact process, started in $X_0 = 1_{\{0\}}$, i.e., with a single infected site at the origin, as a function of the infection rate λ. For reasons that we cannot explain here, this is in fact the same as the probability that the origin is infected in equilibrium.

It turns out that for the nearest-neighbor contact process on $\mathbb{Z}^d$, there exists a *critical value* $\lambda_{\rm c} = \lambda_{\rm c}(d)$ with $0 < \lambda_{\rm c} < \infty$ such that $\theta(\lambda) = 0$ for $\lambda \leq \lambda_{\rm c}$ and $\theta(\lambda) > 0$ for $\lambda > \lambda_{\rm c}$. The function θ is continuous, strictly increasing and concave on $[\lambda_{\rm c}, \infty)$ and satisfies $\lim_{\lambda\to\infty} \theta(\lambda) = 1$. One has

$$\lambda_{\rm c}(1) = 1.6489 \pm 0.0002. \tag{1.10}$$

Proving these statements is not easy, however. For example, continuity of the function θ in the point $\lambda_{\rm c}$ was proved only in 1990 [BG90], seventeen years

[3] In fact, on the finite square used in our simulations, one can prove that the infection dies out a.s. However, the time one has to wait for this is exponentially large in the system size. For the size of system shown in Figure 1.3, this time is already too long to be numericaly observable.

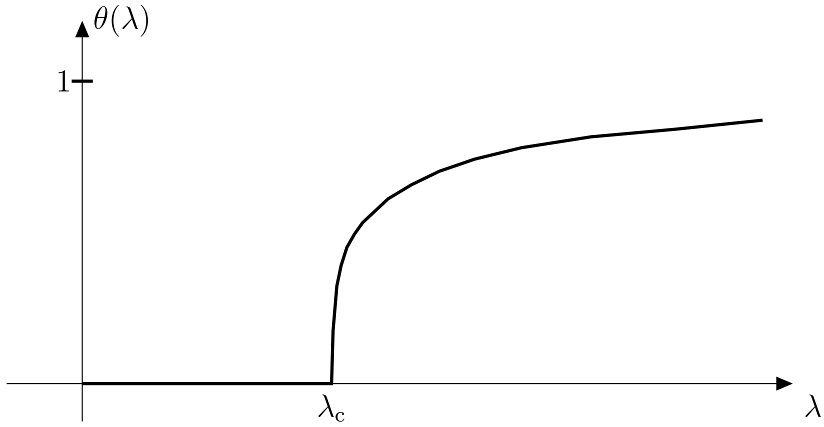


Figure 1.4: Survival probability of the one-dimensional contact process.

after the introduction of the model in [CS73, Har74]. The best[4] rigorous upper bound on the constant from (1.10) is $\lambda_c(1) \leq 1.942$ which is proved in [Lig95].

1.4 Ising and Potts models

In an *Ising model*, sites in the lattice $\mathbb{Z}^d$ are interpreted as atoms in a crystal, that can have two possible local states, usually denoted by -1 and $+1$. In the traditional interpretation, these states describe the direction of the magnetic field of the atom, and because of this, the local state $x(i)$ of a site i is usually called the *spin* at i. More generally, one can consider *Potts models* where each "spin" can have $q \geq 2$ possible values. In this case, the local state space is traditionally denoted as $S = \{1, \ldots, q\}$, the special case $q = 2$ corresponding to the Ising model (except for a small difference in notation between $S = \{-1, +1\}$ and $S = \{1, 2\}$).

Given a state x and site i, we let

$$N_{x,i}(\sigma) := \sum_{j \in \mathcal{N}_i} 1_{\{x(j) = \sigma\}} \qquad (\sigma \in S) \tag{1.11}$$

denote the number of neighbors of the site i that have the spin value $\sigma \in S$. In the Ising and Potts models, sites like or dislike to have the same spin

[4]There exists a sequence of rigorous upper bounds on the constant from (1.10) that is known to converge to the real value, but these bounds are so difficult to calculate that the best bound that has really been achieved by this method is much worse than the one in [Lig95].

value as their neighbors, depending on a parameter $\beta \in \mathbb{R}$. Adding a so-called *Glauber dynamics* to the model, sites update their spin values with rate one, and at such an event choose a new spin value with probabilities that depend on the values of their neighbors. More precisely,

$$\text{site } i \text{ flips to the value } \sigma \text{ with rate } \quad \frac{e^{\beta N_{x,i}(\sigma)}}{\sum_{\tau \in S} e^{\beta N_{x,i}(\tau)}}. \tag{1.12}$$

If $\beta > 0$, then this means that sites prefer to have spin values that agree with as many neighbors as possible, i.e., the model is *ferromagnetic*. For $\beta < 0$, the model is *antiferromagnetic*. These terms reflect the situation that in some materials, neighboring spins like to line up, which can lead to long-range order that has the effect that the material can be magnetized. Antiferromagnetic materials, on the other hand, lack this effect.

Alternatively, Potts models can also be interpreted as social or economic models, where sites represent people or firms and spin values represent opinions or the state (financially healthy or not) of a firm [BD01].

In Figure 1.5 we see four snapshots of a two-dimensional nearest-neighbor Potts model with four possible spin values. We have used periodic boundary conditions, and the value of the parameter β is 1.2. Superficially, the behavior is similar to that of a voter model, in the sense that the system forms clusters of growing size that in the end take over any finite neighborhood of the origin. Contrary to the voter model, however, even in the middle of large cluster that is predominantly of one color, sites can still flip to other values as is clear from (1.12), so in the simulations we see many small islands of different colors inside large clusters where one color dominates. Another difference is clustering actually happens only when the value of the parameter β is large enough. For small values of β, the behavior is roughly similar to the voter model in dimensions $d \geq 3$. There is a critical value $0 < \beta_{\rm c} < \infty$ where the model changes from one type of behavior to the other type of behavior. In this respect, the model is similar to the contact process.

To make this critical value visible, imagine that instead of periodic boundary conditions, we would use frozen boundary conditions where the sites at the boundary are kept fixed at one chosen color, say color 1. Then the system has a unique invariant law (equilibrium), in which for sufficiently large values of β the color 1 is (much) more frequent than the other colors, but for low values of β all colors occur with the same frequency. In particular, for the Ising model, where the set of possible spin values in $\{-1,+1\}$, we let

$$m_*(\beta) := \begin{array}{l}\text{the expectation of } x(0) \text{ with } +1 \text{ boundary} \\ \text{conditions, in the limit of large system size.}\end{array} \tag{1.13}$$

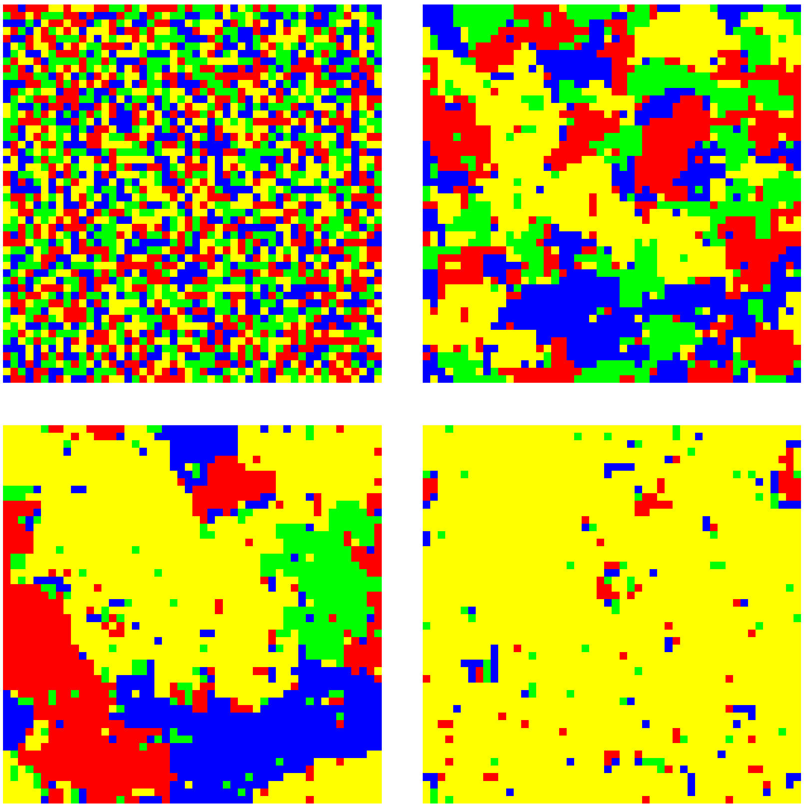


Figure 1.5: Four snapshots of a $q = 4$, $\beta = 1.2$ Potts model with Glauber dynamics and periodic boundary conditions. Initially, the types of sites are i.i.d. Time evolved in these pictures is 0, 4, 32, 500.

This function is called the *spontaneous magnetization*. For the Ising model in two dimensions, the spontaneous magnetization can be explicitly calculated, as was first done my Onsager [Ons44]. The formula is

$$m_*(\beta) = \begin{cases} \big(1 - \sinh(\beta)^{-4}\big)^{1/8} & \text{for } \beta \geq \beta_{\rm c} := \log(1+\sqrt{2}), \\ 0 & \text{for } \beta \leq \beta_{\rm c}. \end{cases} \tag{1.14}$$

This function is plotted in Figure 1.6. In this case, the critical point $\beta_{\rm c}$ is known explicitly.

For Ising models in dimensions $d \geq 3$, the graph of $m_*(\beta)$ looks roughly similar to Figure 1.6, but no explicit formulas are known.

In dimension one, one has $m^*(\beta) = 0$ for *all* $\beta \geq 0$. More generally, one-dimensional Potts models do not show long range order, even if β is very

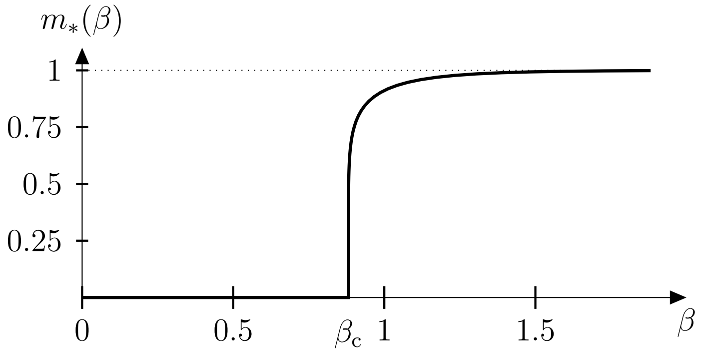


Figure 1.6: The spontaneous magnetization of the two-dimensional Ising model.

large.[5] By this we mean that in equilibrium, the correlation between the spin values at 0 and a point $i \in \mathbb{Z}$ tends to zero as $i \to \infty$ for any value of β (even though the decay is slow if β is large). In Figure 1.7, we compare the time evolution of a one-dimensional Potts model (with a large value of β) with the time evolution of a one-dimensional voter model. In the voter model, the cluster size keeps growing, but in the Potts model, the typical cluster size converges to a finite limit.

The careful reader will have noticed that so far, we have not given a formula for the generator of our Ising and Potts models, but instead described the Glauber dynamics by formula (1.12). It is possible to give a formula for the generator in terms of local maps that are defined with appropriate rates as in (1.1), but this is a bit complicated (at least notationally) and in fact there is more than one good way to do this.

1.5 Phase transitions

Figures 1.4 and 1.6 are examples of a phenomenon that is often observed in interacting particle systems. As a parameter governing the dynamics is crosses a particular value, the system goes through an abrupt change in behavior. This is called a *phase transition* and the value of the parameter is called the *point of the phase transition* or, in the mathematical literature, *critical point*. As we will see in a moment, in the physics literature, the term

[5]This was first noticed by Ising [Isi25], who introduced the model but noticed that it was uninteresting, incorrectly assuming that what he had proved in dimension 1 would probably hold in any dimension. Peierls [Pei36] realized that dimension matters and proved that the Ising model in higher dimensions does show long range order.

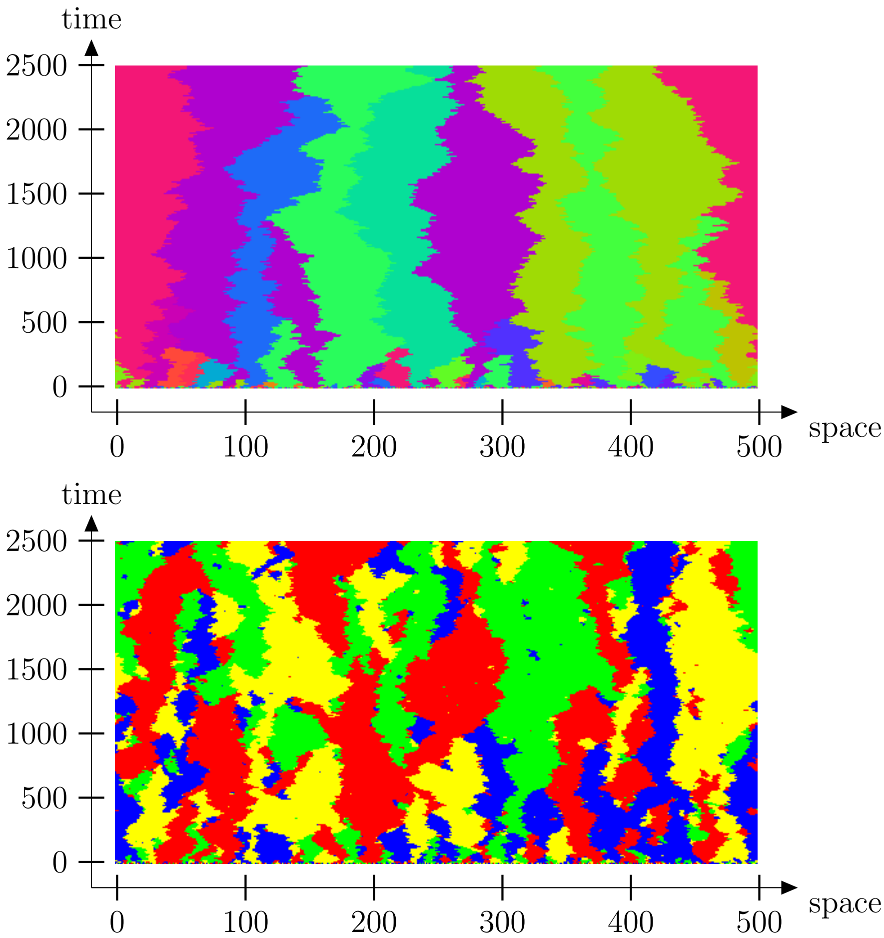


Figure 1.7: Time evolution of a one-dimensional voter model (above) and a one-dimensional Potts model (below) with a high value of β.

critical point has a more restricted meaning. The term "phase transition" of course also describes the behavior that certain materials change from a gas, fluid, or solid phase into another phase at a particular value of the temperature, pressure etc., and from the theoretical physicist's point of view, this is indeed the same phenomenon.

In both Figure 1.4 and 1.6, the point of the phase transition in fact separates two regimes, one where the interacting particle systems (on the infinite lattice) has a unique invariant law (below λ_c and β_c) and another regime where there are more invariant laws (above λ_c and β_c). Indeed, for the contact process, the delta measure on the empty configuration is always an invariant law, but above λ_c, a second, nontrivial invariant also appears. Potts models have q invariant laws (one corresponding to each color) above the critical point. Multiple invariant laws are a general phenomenon associated with phase transitions.

Phase transitions are classified into *first order* and *second order* phase transitions.[6] Second order phase transitions are also called *continuous* phase transitions. The phase transitions in Figures 1.4 and 1.6 are both second order, since the functions θ and m_* are continuous at the critical points $\lambda_{\rm c}$ and $\beta_{\rm c}$, respectively. Also, second order phase transitions are characterized by the fact that at the critical point, there is only one invariant law. By contrast, if we would draw the function $m_*(\beta)$ of a Potts model for sufficiently large values of q (in dimension two, for $q > 4$), then the plot of m_* would make a jump at $\beta_{\rm c}$ and the system would have multiple invariant laws at this point, which means that this phase transition is first order.

It can be difficult to prove whether a given phase transition is first or second order. While for the two-dimensional Ising model, continuity of the magnetization follows from Onsager's solution [Ons44], the analogue statement for the three-dimensional Ising model was only proved recently [ADS15] (70 years after Onsager!).

For the Ising model, it is known (but only partially proved) that

$$m_*(\beta) \propto (\beta - \beta_{\rm c})^c \quad \text{as } \beta \downarrow \beta_{\rm c},$$

where c is a *critical exponent*, which is given by

$$c = 1/8 \text{ in dim } 2, \quad c \approx 0.326 \text{ in dim } 3, \quad \text{and} \quad c = 1/2 \text{ in dim} \geq 4.$$

For the contact process, one observes that

$$\theta(\lambda) \propto (\lambda - \lambda_{\rm c})^c \qquad \text{as } \lambda \downarrow \lambda_{\rm c},$$

with a critical exponent

$$\begin{aligned} &c \approx 0.276 \text{ in dim } 1, \quad c \approx 0.583 \text{ in dim } 2, \\ &c \approx 0.813 \text{ in dim } 3, \quad \text{and} \quad c = 1 \text{ in dim} \geq 4. \end{aligned}$$

In theoretical physics, (nonrigorous) *renormalization group theory* is used to explain these critical exponents and calculate them. According to this theory, critical exponents are *universal*. For example, the nearest-neighbor

[6]This terminology was introduced by Paul Ehrenfest. The idea is that in first order phase transitions, the first derivative of the free energy has a discontinuity, while in a second order phase transitions, the first derivative of the free energy is continuous and only the second derivative makes a jump.

model and the range R models with different values of R all have different values of the critical point, but the critical exponent c has the same value for all these models. Also, changing from the square lattice to, e.g., the triangular lattice has no effect on c.

Critical exponents are associated only with second order phase transitions. At the critical point of a second order phase transition, one observes *critical behavior*, which involves, e.g., power-law decay of correlations. For this reason, physicists use the term "critical point" only for second order phase transitions.

So far, there is no mathematical theory that can explain critical behavior, except in high dimensions (where one uses a technique called the *lace expansion*) and in a few two-dimensional models.

1.6 Variations on the voter model

Apart from the models discussed so far, lots of other interacting particle systems have been introduced ans studied in the literature to model a phlectora of phenomena. Some of these behave very similarly to the models we have already seen (and even appear to have the same critical exponents), while others are completely different. In this and the next section, we take a brief look at some of these models to get an impression of the possibilities.

The *biased voter model* with *bias* $s \geq 0$ is the interacting particle system with state space $\{0,1\}^{\mathbb{Z}^d}$ and generator (compare (1.5))

$$\begin{aligned} G_{\rm bias} f(x) := & \frac{1}{2d} \sum_{(i,j)\in\mathcal{E}^d} \big\{ f\big(\big(\mathtt{vot}_{ij}(x)\big) - f\big(x\big)\big\} \\ & + \frac{s}{2d} \sum_{(i,j)\in\mathcal{E}^d} \big\{ f\big(\big(\mathtt{bra}_{ij}(x)\big) - f\big(x\big)\big\}, \end{aligned} \tag{1.15}$$

where $\mathtt{vot}_{ij}$ and $\mathtt{bra}_{ij}$ are the voter and branching maps defined in (1.4) and (1.6). The biased voter model describes a situation where one genetic type of an organism (in this case, type 1) is more fit than the other type, and hence reproduces at a larger rate. Alternatively, this type may represent a new idea of opinion that is more attractive than the current opinion. Contrary to the normal voter model, even if we start with just a single invidual of type 1, there is a positive probability that type 1 never dies out and indeed takes over the whole population, as can be seen in Figure 1.8.

Fix $i \in \mathbb{Z}^d$ and for any $x \in \{0,1\}^{\mathbb{Z}^d}$, let

$$f_\tau(x) := \frac{1}{|\mathcal{N}_i|} \sum_{j\in\mathcal{N}_i} 1_{\{x(j) = \tau\}} \qquad (\tau = 0, 1)$$

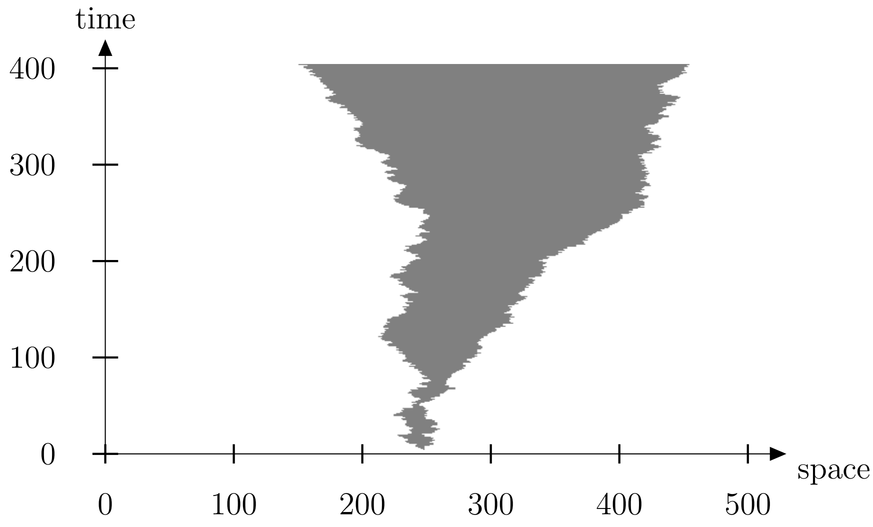


Figure 1.8: Time evolution of a one-dimensional biased voter model with bias $s = 0.2$.

be the frequency of type τ in the neighborhood $\mathcal{N}_i$. In the standard voter model, if the present state is x, then the site i changes its type with the following rates:

$$\begin{aligned} 0 \mapsto 1 &\quad \text{with rate } f_1(x), \\ 1 \mapsto 0 &\quad \text{with rate } f_0(x). \end{aligned}$$

In the biased voter model, this is changed to

$$\begin{aligned} 0 \mapsto 1 &\quad \text{with rate } (1+s)f_1(x), \\ 1 \mapsto 0 &\quad \text{with rate } f_0(x). \end{aligned}$$

Another generalization of the voter model, introduced in [NP99], is defined by the rates

$$\begin{aligned} 0 \mapsto 1 &\quad \text{with rate } f_1(x)\big(f_0(x) + \alpha f_1(x)\big), \\ 1 \mapsto 0 &\quad \text{with rate } f_0(x)\big(f_1(x) + \alpha f_0(x)\big), \end{aligned} \tag{1.16}$$

where $0 \leq \alpha \leq 1$ is a model parameter. Another way of expressing this is to say that if the individual at i is of type τ, then this individual dies with rate

$$f_\tau(x) + \alpha f_{1-\tau}(x), \tag{1.17}$$

and once an individual has died, just as in the normal contact process, it is replaced by a descendant of a uniformly chosen neighbor.

If $\alpha = 1$, then the rate of dying in (1.17) is one and we are back at the standard voter model, but for $\alpha < 1$, individuals die less often if they are surrounded by a lot of individuals of the other type. In biology, this models *balancing selection*. This is the effect that individuals that differ from their neighbors experience less competition, which results in a selective drive for high biodiversity.

In the social interpretation of the voter model, we may interpret (1.17) as saying that persons change their mind *less* often if they disagree with a lot of neighbors, i.e., the model in (1.16) has "rebellious" behavior.

Numerical simulations, shown in Figure 1.9, suggest that in one dimension and for ranges $R \geq 2$, the model in (1.16) exhibits a phase transition in α. For α sufficiently close to 1, the model behaves essentially as a voter model, with clusters growing in time, but for small values of α (which represent strong rebellious behavior), the cluster size tends to a finite limit.

1.7 Further models

For each $i, j \in \mathbb{Z}^d$, we define a *coalescing random walk map* $\mathrm{rw}_{ij} : \{0,1\}^{\mathbb{Z}^d} \to \{0,1\}^{\mathbb{Z}^d}$ by

$$\mathtt{rw}_{ij}(x)(k) := \begin{cases} 0 & \text{if } k = i, \\ x(i) \vee x(j) & \text{if } k = j, \\ x(k) & \text{otherwise.} \end{cases} \tag{1.18}$$

Applying $\mathtt{rw}_{ij}$ to a configuration x has the effect that if the site i is occupied by a particle, then this particle jumps to the site j. If there is already a particle at j, then the two particles coalesce.

The interacting particle system with generator

$$G_{\mathrm{rw}} f(x) = \frac{1}{|\mathcal{N}_0|} \sum_{(i,j) \in \mathcal{E}^d} \big\{ f\big((\mathtt{rw}_{ij}(x)\big) - f\big(x\big) \big\} \qquad (x \in \{0,1\}^{\mathbb{Z}^d}) \tag{1.19}$$

describes a system of coalescing random walks, where each particle jumps with rate 1 to a uniformly chosen neighboring site, and two particles on the same site coalesce; see Figure 1.10. Likewise, replacing the coalescing random walk map by the *annihilating random walk map* defined as

$$\mathtt{ann}_{ij}(x)(k) := \begin{cases} 0 & \text{if } k = i, \\ x(i) + x(j) \mod(2) & \text{if } k = j, \\ x(k) & \text{otherwise,} \end{cases} \tag{1.20}$$

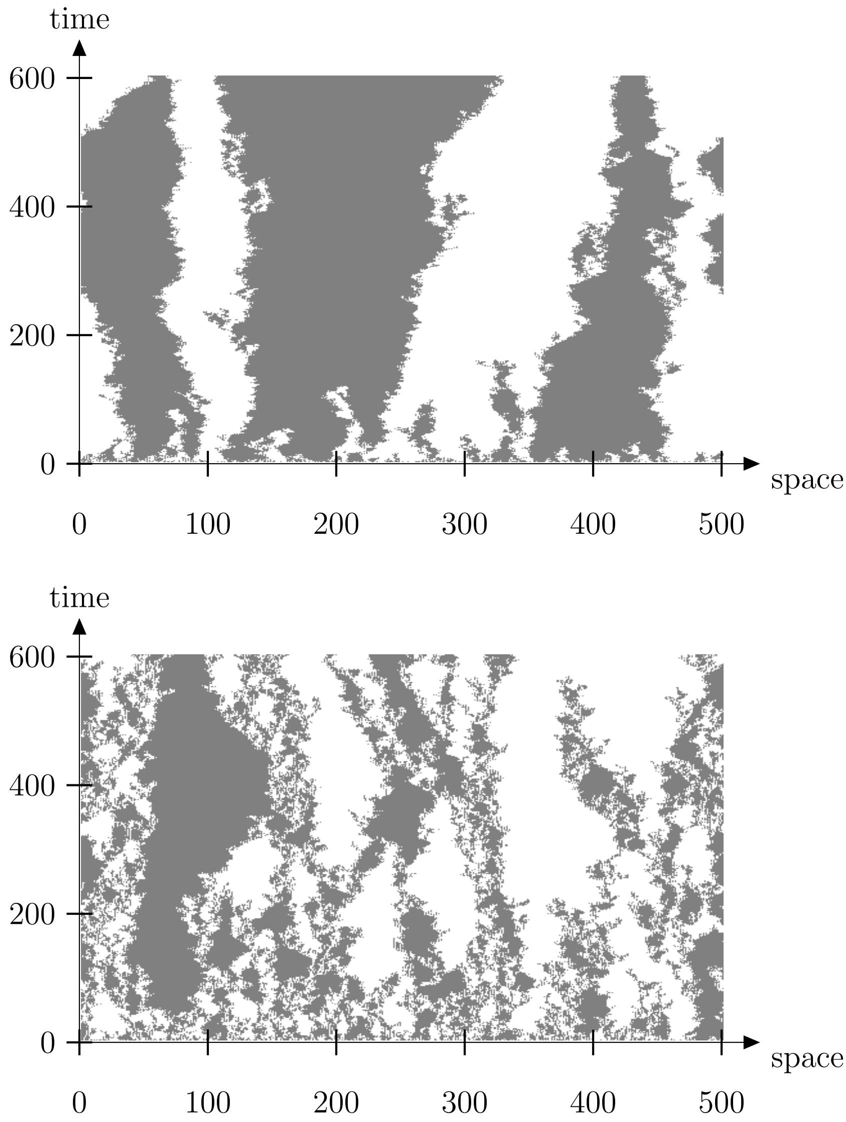


Figure 1.9: Evolution of "rebellious" voter models with $\alpha = 0.8$ and $\alpha = 0.3$, respectively.

yields a system of annihilating random walks, that kill each other as soon as two particles land on the same site; see Figure 1.10.

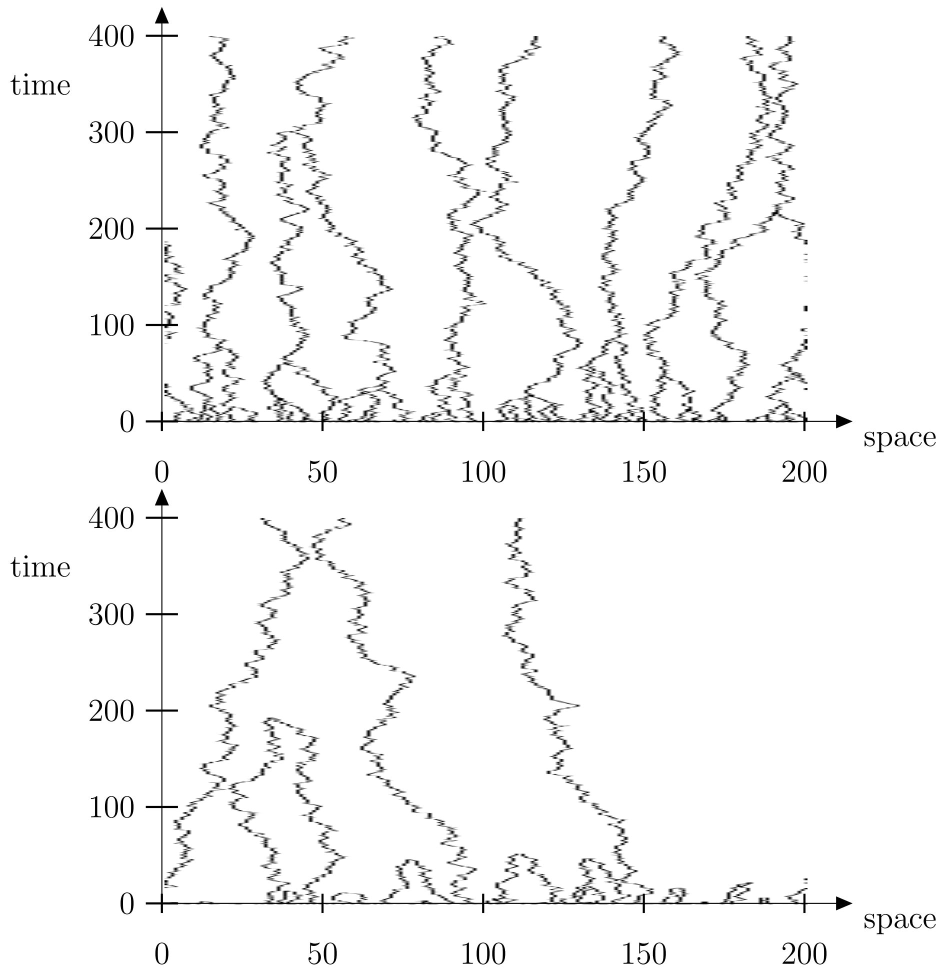


Figure 1.10: Systems of coalescing random walks (above) and annihilating random walks (below).

For each $i, j \in \mathbb{Z}^d$, we define an *exclusion map* $\mathrm{excl}_{ij} : S^{\mathbb{Z}^d} \to S^{\mathbb{Z}^d}$ by

$$\mathtt{excl}_{ij}(x)(k) := \begin{cases} x(j) & \text{if } k = i, \\ x(i) & \text{if } k = j, \\ x(k) & \text{otherwise.} \end{cases} \tag{1.21}$$

Applying $\mathtt{excl}_{ij}$ to a configuration x has the effect of interchanging the types of j and j. The interacting particle system with state space $\{0,1\}^{\mathbb{Z}^d}$ and generator

$$G_{\rm excl}f(x)=\frac{1}{|\mathcal{N}_0|}\sum_{(i,j)\in\mathcal{E}^d}\big\{f\big(\big(\texttt{excl}_{ij}(x)\big)-f\big(x\big)\big\}\qquad(x\in\{0,1\}^{\mathbb{Z}^d})\quad(1.22)$$

is called the (symmetric) *exclusion process*. In the exclusion process, individual particles move according to random walks, that are independent as long as the particles are sufficiently far apart. Particles never meet, and the total number of particles is preserved.

The previous three maps (coalescing random walk map, annihilating random walk map, and exclusion map) can be combined with, e.g., the branching map and death map from (1.6) and (1.7). In particular, adding coalescing random walk or exclusion dynamics to a contact process models displacement (migration) of organisms. Since in many organisms, you actually need two parents to produce offspring, several authors [Nob92, Dur92, Neu94, SS15a] have studied particle systems where the branching map is replaced by the *cooperative branching map*

$$\texttt{coop}_{ijk}(x)(l):=\begin{cases}1 & \text{if } l=k,\ x(i)=1,\ x(j)=1,\\ x(l) & \text{otherwise.}\end{cases}\qquad(1.23)$$

See Figure 1.11 for a one-dimensional interacting particle system involving cooperative branching and coalescing random walks.

We define a *killing map* by

$$\texttt{kill}_{ij}(x)(k):=\begin{cases}0 & \text{if } k=j,\ x(i)=1,\ x(j)=1,\\ x(k) & \text{otherwise.}\end{cases}\qquad(1.24)$$

In words, this says that if there are particles at i and j, then the particle at i kills the particle at j. Sudbury [Sud97, Sud99] has studied a "biased annihilating branching process" with generator of the form

$$\begin{aligned}G_{\rm babp}f(x)&:=\lambda\sum_{(i,j)\in\mathcal{E}^1}\big\{f\big(\big(\texttt{bra}_{ij}(x)\big)-f\big(x\big)\big\}\\ &:=\sum_{(i,j)\in\mathcal{E}^1}\big\{f\big(\big(\texttt{kill}_{ij}(x)\big)-f\big(x\big)\big\}\qquad(x\in\{0,1\}^{\mathbb{Z}}).\end{aligned}\qquad(1.25)$$

Figure 1.12 shows a simulation of such a system when $\lambda=0.2$.

Although many interacting particle systems studied in the literature have only two possible local states (usually denoted by 0 and 1), this is not always so. For example, in [Kro99], a *two-stage contact process* is introduced. Here, the local state space is $\{0,1,2\}$ where 0 represents an empty site, 1 a young organism, and 2 an adult organism. The behavior of this model is similar to that of the contact process.

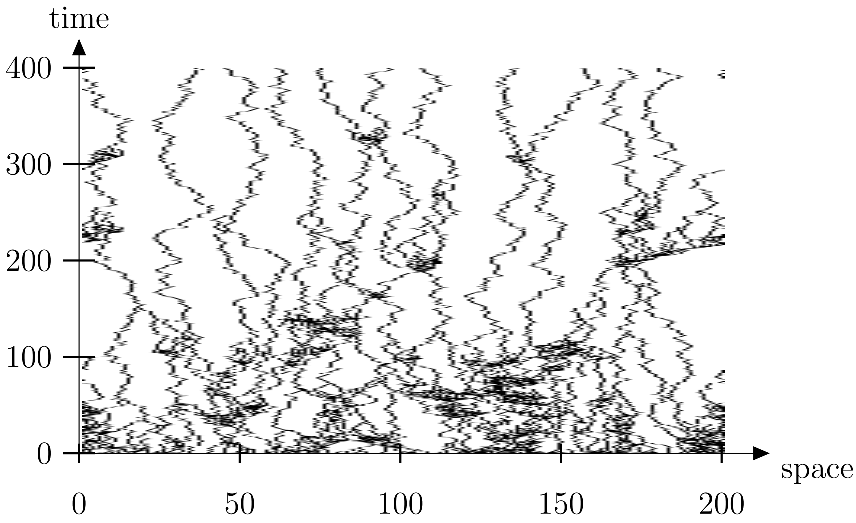


Figure 1.11: A one-dimensional interacting particle system with cooperative branching and coalescing random walk dynamics.

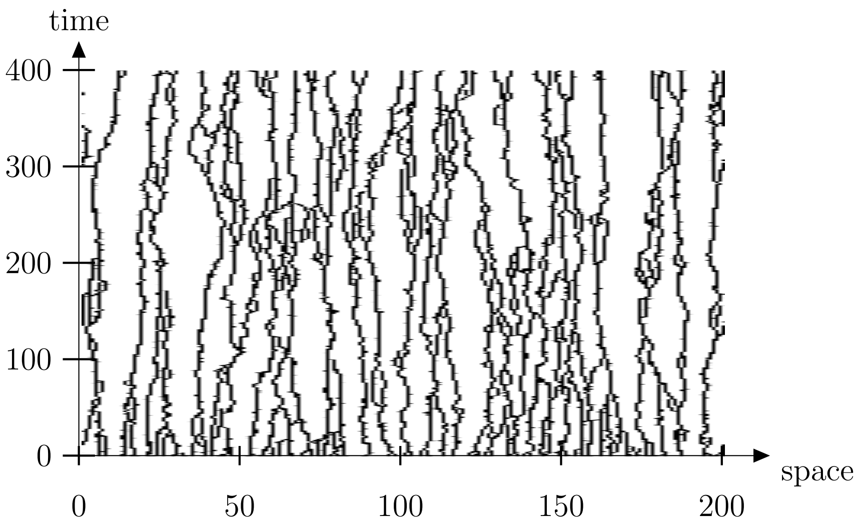


Figure 1.12: A system with branching and killing.

Chapter 2

Continuous-time Markov chains

2.1 Poisson point sets

Let S be a σ-compact[1] metrizable space. We will mainly be interested in the case that $S = \mathbb{R} \times \Lambda$ where Λ is a countable set. We let $\mathcal{S}$ denote the Borel-σ-field on S. A *locally finite measure* on $(S, \mathcal{S})$ is a measure μ such that $\mu(C) < \infty$ for all compact $C \subset S$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be our underlying probability space. A random measure on S is a function $\xi : \Omega \times \mathcal{S} \to [0, \infty]$ such that for fixed $\omega \in \Omega$, the function $\xi(\omega, \cdot\,)$ is a locally finite measure on $(S, \mathcal{S})$, and for fixed $A \in \mathcal{S}$, the function $\xi(\,\cdot\,, A)$ is measurable. By [Kal97, Lemma 1.37], we can think of ξ as a random variable with values in the space of locally finite measures on $(S, \mathcal{S})$, equipped with the σ-field generated by the maps $\mu \mapsto \mu(A)$ with $A \in \mathcal{S}$. Then the integral $\int f \mathrm{d}\xi$ defines a $[0, \infty]$-valued random variable for all measurable $f : S \to [0, \infty]$. There exists a unique measure, denoted by $\mathbb{E}[\xi]$, such that

$$\int f \,\mathrm{d}\mathbb{E}[\xi] = \mathbb{E}\Big[\int f \,\mathrm{d}\xi\Big]$$

for all measurable $f : S \to [0, \infty]$. The measure $\mathbb{E}[\xi]$ is called the *intensity* of ξ.

The following result follows from [Kal97, Lemma 10.1 and Prop. 10.4].[2] Below, $\hat{\mathcal{S}} := \{A \in \mathcal{S} : \overline{A} \text{ is compact}\}$ denotes the set of measurable subsets of S whose closure is compact.

[1] This means that there exists a countable collection of compact sets $S_i \subset S$ such that $\bigcup_i S_i = S$.

[2] In fact, [Kal97, Prop. 10.4] shows that it is possible to construct Poisson point measures on arbitrary measurable spaces, assuming only that the intensity measure is σ-finite, but we will not need this generality.

Proposition 2.1 (Poisson point measures) *Let μ be a locally finite measure on $(S, \mathcal{S})$. Then there exists a random measure ξ, unique in distribution, such that for any disjoint $A_1, \ldots, A_n \in \hat{\mathcal{S}}$, the random variables $\xi(A_1), \ldots, \xi(A_n)$ are independent and $\xi(A_i)$ is Poisson distributed with mean $\mu(A_i)$.*

We call a random measure ξ as in (2.1) a *Poisson point measure* with *intensity*μ. Indeed, one can check that $\mathbb{E}[\xi] = \mu$. We note that $\xi(A) \in \mathbb{N}$ for all $A \in \hat{\mathcal{S}}$. Such measures are called (locally finite) *counting measures*. Each locally finite counting measure ν on S is of the form

$$\nu = \sum_{x \in \mathrm{supp}(\nu)} n_x \delta_x,$$

where the supp(ν), the support of ν, is a locally finite subset of S, the n_x are positive integers, and δ_x denotes the delta-measure at x. We say that ν is *simple* if $n_x = 1$ for all $x \in \mathrm{supp}(\nu)$. Recall that a measure μ has an *atom* at x is $\mu(\{x\}) > 0$. A measure μ is called *atomless* if it has no atoms, i.e., $\mu(\{x\}) = 0$ for all $x \in S$. The already mentioned [Kal97, Prop. 10.4] tells us the following.

Lemma 2.2 (Simple Poisson point measures) *Let ξ be a Poisson point measure with locally finite intensity μ. Then ξ is a.s. simple if and only if μ is atomless.*

If μ is atomless, then a Poisson point measure ξ with intensity μ is characterized by its support $\omega := \mathrm{supp}(\xi)$. We call ω a *Poisson point set* with intensity μ. Intuitively, ω is a set such that $\mathbb{P}[\omega \cap \mathrm{d}x \neq \emptyset] = \mu(\mathrm{d}x)$, independently for each infinitesimal subset $\mathrm{d}x \subset S$.

For any counting measure ν on S and measurable function $f : S \to [0, 1]$ we introduce the notation

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

Here, by definition, $f^0 := 1$, where 0 denotes the counting measure that is identically zero. Alternatively, our definition says that

$$f^\nu = e^{\int (\log f) \mathrm{d}\nu},$$

where $\log 0 := -\infty$ and $e^{-\infty} := 0$. It is easy to see that $f^\nu f^{\nu'} = f^{\nu + \nu'}$.

Lemma 2.3 (Laplace functionals) *Let μ be a localy finite measure on $(S, \mathcal{S})$ and let ξ be a Poisson point measure with intensity μ. Then*

$$\mathbb{E}\big[(1-f)^{\nu}\big] = e^{-\int f \mathrm{d}\mu} \tag{2.1}$$

for each measurable $f : S \to [0,1]$. Conversely, if ξ is a random counting measure and (2.1) holds for all continuous, compactly supported f, then ξ is a Poisson point measure with intensity μ.

Proof The fact that Poisson point measures satisfy (2.1) is proved in [Kal97, Lemma 10.2], which is written in terms of $-\log f$, rather than f. The fact that (2.1) determines the law of ξ uniquely follows from [Kal97, Lemma 10.1]. ∎

Formula (2.1) can be interpreted in terms of thinning. Consider a counting measure $\nu = \sum_i \delta_{x_i}$, let $f : S \to [0,1]$ be measurable, and let χ_i be independent *Bernoulli random variables* (i.e., random variables with values in $\{0,1\}$) with $\mathbb{P}[\chi_i = 1] = f(x_i)$. Then the random counting measure

$$\nu' := \sum_i \chi_i \delta_{x_i}$$

is called an *f-thinning* of the counting measure ν. Note that

$$\mathbb{P}[\nu' = 0] = \prod_i \mathbb{P}[\chi_i = 0] = (1-f)^{\nu}.$$

In view of this, the left-hand side of (2.1) can be interpreted as the probability that after thinning the random counting measure ξ with f, no points remain. By [Kal97, Lemma 10.1], knowing this probability for each continuous, compactly supported f uniquely determines the law of a random counting measure.

Using Lemma 2.3, it is easy to prove that if ξ_1 and ξ_2 are independent Poisson point measures with intensities μ_1 and μ_2, then $\xi_1 + \xi_2$ is a Poisson point measure with intensity $\mu_1 + \mu_2$. We also mention [Kal97, Lemma 10.17], which says the following.

Lemma 2.4 (Poisson points on the halfline) *Let $(\tau_k)_{k\geq 0}$ be real random variables such that $\tau_0 = 0$ and $\sigma_k := \tau_k - \tau_{k-1} > 0$ $(k \geq 1)$. Then $\omega := \{\tau_k : k \geq 1\}$ is a Poisson point set on $[0,\infty)$ with intensity $c\ell$, where ℓ denotes the Lebesgue measure, if and only if the random variables $(\sigma_k)_{k\geq 1}$ are i.i.d. exponentially distributed with mean c^{-1}.*

2.2 Transition probabilities and generators

Let S be any finite set. A (real) matrix indexed by S is a collection of real constants $A = (A(x,y))_{x,y\in S}$. We calculate with such matrices in the same way as with normal finite matrices. Thus, the product AB of two matrices is defined as

$$(AB)(x,z) := \sum_{y\in S} A(x,y)B(y,z) \qquad (x,z\in S).$$

We let 1 denote the identity matrix $1(x,y) = 1_{\{x=y\}}$ and define A^n in the obvious way, with $A^0 := 1$. If $f : S \to \mathbb{R}$ is a function, then we also define

$$Af(x) := \sum_{y\in S} A(x,y)f(y) \quad \text{and} \quad fA(y) := \sum_{x\in S} f(x)A(x,y). \tag{2.2}$$

A *probability kernel* on S is a matrix $K = (K(x,y))_{x,y\in S}$ such that $K(x,y) \geq 0$ $(x,y \in S)$ and $\sum_{y\in S} K(x,y) = 1$ $(x \in S)$. Clearly, the composition of two probability kernels yields a third probability kernel. A *Markov semigroup* is a collection of probability kernels $(P_t)_{t\geq 0}$ such that

$$\lim_{t\downarrow 0} P_t = P_0 = 1 \quad \text{and} \quad P_sP_t = P_{s+t} \qquad (s,t\geq 0).$$

Each such Markov semigroup is of the form

$$P_t = e^{tG} := \sum_{n=0}^{\infty} \frac{1}{n!}(tG)^n,$$

where the *generator* G is a matrix of the form

$$G(x,y) \geq 0 \quad (x\neq y) \quad \text{and} \quad \sum_y G(x,y) = 0. \tag{2.3}$$

By definition, a *Markov process* with semigroup $(P_t)_{t\geq 0}$ is a stochastic process $X = (X_t)_{t\geq 0}$ with values in S and piecewise constant, right-continuous sample paths, such that

$$\mathbb{P}\big[X_u \in \cdot \,\big|\, (X_s)_{0\leq s\leq t}\big] = P_{u-t}(X_t,\,\cdot\,) \quad \text{a.s.} \qquad (0\leq t\leq u). \tag{2.4}$$

Here, in the left-hand side, we condition on the σ-field generated by the random variables $(X_s)_{0\leq s\leq t}$. Formula (2.4) is equivalent to the statement that

$$\begin{aligned}&\mathbb{P}\big[X_0 = x_0, \ldots, X_{t_n} = x_n\big]\\ &\quad = \mathbb{P}[X_0 = x_0]P_{t_1-t_0}(x_0,x_1)\cdots P_{t_n-t_{n-1}}(x_{n-1},x_n) \qquad (0 < t_1 < \cdots < t_n).\end{aligned} \tag{2.5}$$

From this last formula, we see that for each initial law $\mathbb{P}[X_0 = \cdot\,] = \mu$, there is a unique Markov process with semigroup $(P_t)_{t\geq 0}$ and this initial law. Moreover, recalling our notation (2.2), we see that

$$\mu P_t(x) = \mathbb{P}[X_t = x] \qquad (x \in S)$$

is the law of the process at time t. It is custom to let $\mathbb{P}^x$ denote the law of the Markov process with deterministic initial state $X_0 = x$ a.s. We note that

$$\mathbb{P}^x[X_t = y] = P_t(x, y) = 1_{\{x=y\}} + tG(x, y) + O(t^2) \quad \text{as } t \downarrow 0.$$

For $x \neq y$, we call $G(x, y)$ the *rate* of jumps from x to y. Intuitively, if the process is in x, then in the next infinitesimal time interval of length $\mathrm{d}t$ it has a probability $G(x, y)\mathrm{d}t$ to jump to y, independently for all $y \neq x$.

Let X be the process started in x and let $\tau := \inf\{t \geq 0 : X_t \neq x\}$. Then one can show that τ is exponentially distributed with mean r^{-1}, where $r := \sum_{y\neq x} G(x, y)$ is the total rate of all jumps from x. Moreover,

$$\mathbb{P}^x[X_\tau = y] = \frac{G(x, y)}{\sum_{z\neq x} G(x, z)} \qquad (y \in S,\ y \neq x). \tag{2.6}$$

Conditional on $X_\tau = y$, the time of the next jump is again exponentially distributed, and this leads to a construction of $(X_t)_{t\geq 0}$ based on an *embedded Markov chain* with transition kernel $K(x, y)$ given by the right-hand side of (2.6), and exponential holding times. For us, a somewhat different construction based on maps that are applied at Poissonian times will be more useful.

2.3 Poisson construction of Markov processes

Let S be a finite set. Let $\mathcal{G}$ be a set whose elements are maps $m : S \to S$, and let $(r_m)_{m\in\mathcal{G}}$ be nonnegative constants. We equip the space $\mathcal{G} \times \mathbb{R}$ with the measure

$$\rho\big(\{m\} \times A\big) := r_m\, \ell(A) \qquad \big(A \in \mathcal{B}(\mathbb{R})\big),$$

where $\mathcal{B}(\mathbb{R})$ denotes the Borel-σ-field on $\mathbb{R}$ and ℓ denotes the Lebesgue measure. Let ω be a Poisson point set with intensity ρ. Then $\nu := \sum_{(m,t)\in\omega} \delta_t$ is a Poisson point measure on $\mathbb{R}$ with intensity $r\ell$, where $r := \sum_{m\in\mathcal{G}} r_m$. Since the Lebesgue measure is atomless, by Lemma 2.2, this Poisson point measure is simple, i.e., for each $t \in \mathbb{R}$ there exists at most one m such that $(m, t) \in \omega$. Since $r < \infty$, the Poisson point measure ν is moreover locally finite, so, setting

$$\omega_{s,u} := \big\{(m, t) \in \omega : t \in (s, u]\big\} \qquad (s \leq u),$$

we can order the elements of $\omega_{s,u}$ as

$$\omega_{s,u} = \big\{(m_1,t_1),\ldots,(m_n,t_n)\big\} \quad \text{with} \quad t_1 < \cdots < t_n. \tag{2.7}$$

We use this to define a collection of random maps $(\mathbf{X}_{s,u})_{s\leq u}$ by

$$\mathbf{X}_{s,u} := m_n \circ \cdots \circ m_1,$$

where $m_1,\ldots,m_n$ are as in (2.7). Here, by definition, the composition of no maps is the identity map, i.e., $\mathbf{X}_{s,u}$ is the identity map if $\omega_{s,u} = \emptyset$. It is not hard to see that

$$\lim_{t\downarrow s} = \mathbf{X}_{s,t} = \mathbf{X}_{s,s} = 1 \quad \text{and} \quad \mathbf{X}_{t,u} \circ \mathbf{X}_{s,t} = \mathbf{X}_{s,u} \qquad (s \leq t \leq u), \tag{2.8}$$

i.e., the maps $(\mathbf{X}_{s,t})_{s\leq t}$ form a *stochastic flow.* Also, $\mathbf{X}_{s,t}$ is right-continuous in both s and t. Finally, $(\mathbf{X}_{s,t})_{s\leq t}$ has *independent increments* in the sense that

$$\mathbf{X}_{t_0,t_1},\ldots,\mathbf{X}_{t_{n-1},t_n} \quad \text{are independent} \quad \forall\, t_0 < \cdots < t_n.$$

Proposition 2.5 (Poisson construction of Markov processes) *Define a stochastic flow $(\mathbf{X}_{s,t})_{s\leq t}$ as above in terms of a Poisson point set ω. Let X_0 be an S-valued random variable, independent of ω. Then*

$$X_t := \mathbf{X}_{0,t}(X_0) \qquad (t \geq 0) \tag{2.9}$$

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

$$Gf(x) = \sum_{m\in\mathcal{G}} r_m \big\{f\big(m(x)\big) - f\big(x\big)\big\}. \tag{2.10}$$

Proof The process $X = (X_t)_{t\geq 0}$, defined in (2.9), has piecewise constant, right-continuous sample paths. Define

$$P_t(x,y) := \mathbb{P}[\mathbf{X}_{s,s+t}(x) = y] \qquad (t \geq 0), \tag{2.11}$$

where the definition does not depend on the choice of $s \in \mathbb{R}$ since the law of the Poisson process ω is invariant under translations in the time direction. Using the fact that $(\mathbf{X}_{s,t})_{s\leq t}$ has independent increments and X_0 is independent of ω, we see that the finite-dimensional distributions of X satisfy (2.5).

It follows from (2.8) that the probability kernels $(P_t)_{t\geq 0}$ defined in (2.11) form a Markov semigroup. To see that its generator G is given by (2.10), we observe that by the properties of Poisson processes,

$$\mathbb{P}\big[|\omega_{0,t}| \geq 2\big] = O(t^2) \quad \text{as } t \downarrow 0,$$

while

$$\mathbb{P}\big[\omega_{0,t} = \{(m,s)\} \text{ for some } s \in (0,t]\big] = r_m\, t + O(t^2) \quad \text{as } t \downarrow 0.$$

Using this, it follows that for any $f : S \to \mathbb{R}$, as $t \downarrow 0$,

$$P_t f(x) = \mathbb{E}\big[f\big(\mathbf{X}_{0,t}(x)\big)\big] = f(x) + t \sum_{m \in \mathcal{G}} r_m \big\{f\big(m(x)\big) - f\big(x\big)\big\} + O(t^2).$$

Since $P_t f = f + tGf + O(t^2)$, this proves that G is given by (2.10). ∎

2.4 Examples of Poisson representations

We call (2.10) a *random mapping representation* of the generator G. Such random mapping representations are generally not unique. Consider the following example. We choose the state space $S := \{0,1\}$ and the generator G defined by[3]

$$G(0,1) := 2 \quad \text{and} \quad G(1,0) := 1,$$

which corresponds to a Markov process that jumps

$$0 \mapsto 1 \quad \text{with rate } 2 \quad \text{and} \quad 1 \mapsto 0 \quad \text{with rate } 1.$$

We define maps `down`, `up`, and `swap`, mapping the state space $S = \{0,1\}$ into itself, by

$$\left.\begin{aligned} \texttt{down}(x) &:= 0, \\ \texttt{up}(x) &:= 1, \\ \texttt{swap}(x) &:= 1 - x \end{aligned}\right\} \quad (x \in S).$$

It is straightforward to check that the generator G can be represented in terms of the set of maps $\mathcal{G} := \{\texttt{down}, \texttt{up}\}$ as

$$Gf(x) = r_{\texttt{down}}\big\{f\big(\texttt{down}(x)\big) - f\big(x\big)\big\} + r_{\texttt{up}}\big\{f\big(\texttt{up}(x)\big) - f\big(x\big)\big\}, \tag{2.12}$$

where

$$r_{\texttt{down}} := 1 \quad \text{and} \quad r_{\texttt{up}} := 2.$$

But the same generator G can also be represented in terms of the set of maps $\mathcal{G}' := \{\texttt{swap}, \texttt{up}\}$ as

$$Gf(x) = r'_{\texttt{swap}}\big\{f\big(\texttt{down}(x)\big) - f\big(x\big)\big\} + r'_{\texttt{up}}\big\{f\big(\texttt{up}(x)\big) - f\big(x\big)\big\}, \tag{2.13}$$

[3] By (2.3), if G is a Markov generator, then $G(x,x) = -\sum_{y:\, y \neq x} G(x,y)$, so it order to specify a Markov generator, it suffices to give its off-diagonal elements.

where

$$r'_{\mathtt{swap}} := 1 \quad \text{and} \quad r'_{\mathtt{up}} := 1.$$

The random mapping representations (2.12) and (2.13) lead to different ways to construct the same Markov process. In the first construction, we start with a Poisson point set $\omega \subset \mathcal{G} \times \mathbb{R}$, which then defines a stochastic flow $(\mathbf{X}_{s,t})_{s \le t}$, while in the second construction, we start with a Poisson point set $\omega' \subset \mathcal{G}' \times \mathbb{R}$, which defines a different stochastic flow $(\mathbf{X}'_{s,t})_{s \le t}$, that nevertheless can be used to construct the same Markov process.

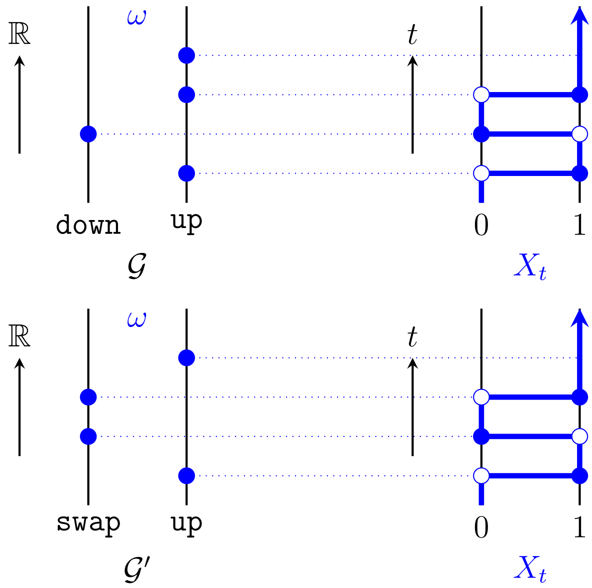


Figure 2.1: Two stochastic flows representing the same Markov process.

The situation is illustrated in Figure 2.1. Note that in the second representation, both the maps `swap` and `up` make the process jump $1 \mapsto 0$ if its previous state is 1. Therefore, the total rate of jumps $1 \mapsto 0$ is

$$r'_{\text{swap}} + r'_{\text{up}} = 2,$$

just as in the first representation.

Chapter 3

The mean-field limit

3.1 Processes on the complete graph

In Chapter 1, we have made acquaintances with a number of interacting particle systems. While some properties of these systems sometimes turn out easy to prove, other seemingly elementary questions can sometimes be remarkably difficult. A few examples of such hard problems have been mentioned in Chapter 1. In view of this, interacting particle systems are being studied by a range of different methods, from straightforward numerical simulations as we have seen in Chapter 1, to nonrigorous renormalization group techniques and rigorous mathematical methods. All these approaches complement each other. In addition, when a given problem appears too hard, one often looks for simpler models that (one hopes) still catch the essence, or at least some essential features of the behavior that one is interested in.

A standard way to turn a difficult model into an (often) much easier model is to take the *mean-field limit*, which we explain in the present chapter. Basically, this means that one replaces the graph structure of the underlying lattice that one is really interested in (in practice often $\mathbb{Z}^d$) by the structure of the complete graph with N vertices, and then takes the limit $N \to \infty$. As we will see, many properties of "real" interacting particle systems are already reflected in these mean-field models. In particular, phase transitions can often already be observed and even the values of critical exponents of high-dimensional models are correctly predicted by the mean-field model. In view of this, studying the mean-field limit is a wise first step in the study of any more complicated model that one may encounter.

Of course, not all phenomena can be captured by replacing the graph structure that one is really interested in by the complete graph. Comparing the real model with the mean-field model, one can learn which elements of the observed behavior are a consequence of the specific spatial structure of the lattice, and which are not. Also for this reason, studying the mean-field limit should be part of a complete study of any interacting particle system.

3.2 The mean-field limit of the Ising model

In this section we study the mean-field Ising model, also known as the , with Glauber dynamics.

We recall from formulas (1.11) and (1.12) in Chapter 1 that the Ising model is an interacting particle system with local state space $S = \{-1, +1\}$, where each site i updates its spin value $x(i) \in \{-1, +1\}$ at rate one. When a spin value is updated, the probability that the new value is $+1$ resp. -1 is proportional to $e^{\beta N_{x,i}}(+1)$ resp. $e^{\beta N_{x,i}}(-1)$, where $N_{x,i}(\sigma) := \sum_{j\in\mathcal{N}_i} 1_{\{x(j)=\sigma\}}$ denotes the number of neighboring sites that have the spin value σ.

For the aim of taking the mean-field model, it will be convenient to formulate the model slightly differently. We let

$$\overline{N}_{x,i} := \frac{1}{|\mathcal{N}_i|} \sum_{j\in\mathcal{N}_i} 1_{\{x(j)=\sigma\}}$$

denote the fraction of neighbors that have the spin value σ, and consider the model where (compare (1.12))

$$\text{site } i \text{ flips to the value } \sigma \text{ with rate} \quad \frac{e^{\beta \overline{N}_{x,i}(\sigma)}}{\sum_{\tau\in S} e^{\beta \overline{N}_{x,i}(\tau)}}. \tag{3.1}$$

Assuming that $|\mathcal{N}_i|$ is just a constant that does not depend on $i \in \Lambda$ (as is the case, e.g., for the model on $\mathbb{Z}^d$), this is just a reparametrization of the original model where the parameter β is replaced by $\beta/|\mathcal{N}_i|$.

We now wish to construct the mean-field model, i.e., the model on a complete graph Λ_N with $|\Lambda_N| = N$ vertices (sites), where each site is a neighbor of each other site. For mathematical simplicity, we even count a site as a neighbor of itself, i.e., we set

$$\mathcal{N}_i := \Lambda_N \quad \text{and} \quad |\mathcal{N}_i| = N.$$

A consequence of this choice is that the *average magnetization*

$$\overline{X}_t := \frac{1}{N} \sum_{i\in\Lambda_N} X_t(i) \qquad (t \geq 0)$$

forms a Markov process $\overline{X} = (\overline{X}_t)_{t\geq 0}$. Indeed, $\overline{X}_t$ takes values in the space

$$\big\{ -1, -1+\tfrac{2}{N}, \ldots, 1-\tfrac{2}{N}, 1\big\},$$

and jumps

$$\begin{aligned}
\overline{x} &\mapsto \overline{x} + \tfrac{2}{N} \qquad \text{with rate} \qquad N_x(-1)\frac{e^{\beta N_x(+1)/N}}{e^{\beta N_x(-1)/N} + e^{\beta N_x(+1)/N}},\\
\overline{x} &\mapsto \overline{x} - \tfrac{2}{N} \qquad \text{with rate} \qquad N_x(+1)\frac{e^{\beta N_x(-1)/N}}{e^{\beta N_x(-1)/N} + e^{\beta N_x(+1)/N}},
\end{aligned}$$

where $N_x(\sigma) := N_{x,i}(\sigma) = \sum_{j\in\Lambda_n} 1_{\{x(j)=\sigma\}}$ does not depend on $i \in \Lambda_N$. We observe that

$$N_x(+1)/N = (1+\overline{x})/2 \quad \text{and} \quad N_x(-1)/N = (1-\overline{x})/2.$$

In view of this, we can rewrite the jump rates of $\overline{X}$ as

$$\begin{aligned}
\overline{x} &\mapsto \overline{x} + \tfrac{2}{N} \qquad \text{with rate} \qquad N(1-\overline{x})/2\,\frac{e^{\beta(1+\overline{x})/2}}{e^{\beta(1-\overline{x})/2} + e^{\beta(1+\overline{x})/2}},\\
\overline{x} &\mapsto \overline{x} - \tfrac{2}{N} \qquad \text{with rate} \qquad N(1+\overline{x})/2\,\frac{e^{\beta(1-\overline{x})/2}}{e^{\beta(1-\overline{x})/2} + e^{\beta(1+\overline{x})/2}}.
\end{aligned}$$

In particular, since these rates are a function of $\overline{x}$ only (and do not depend on other functions of $x = (x(i))_{i\in\Lambda_N}$), we see that $\overline{X} = (\overline{X}_t)_{t\geq 0}$, on its own, is a Markov process. Cancelling a common factor $e^{\beta/2}$ in the nominator and denominator of the rates, we can simplify our formulas a bit to

$$\begin{aligned}
\overline{x} &\mapsto \overline{x} + \tfrac{2}{N} \qquad \text{with rate} \qquad r_+(\overline{x}) := N(1-\overline{x})/2\,\frac{e^{\beta\overline{x}/2}}{e^{-\beta\overline{x}/2} + e^{\beta\overline{x}/2}},\\
\overline{x} &\mapsto \overline{x} - \tfrac{2}{N} \qquad \text{with rate} \qquad r_-(\overline{x}) := N(1+\overline{x})/2\,\frac{e^{-\beta\overline{x}/2}}{e^{-\beta\overline{x}/2} + e^{\beta\overline{x}/2}}.
\end{aligned} \tag{3.2}$$

In Figure 3.1 we can see simulations of the Markov process in (3.2) on a lattice with $N = 10$, 100, 1000, and $10,000$ sites, respectively. It appears that in the limit $N \to \infty$, the process $\overline{X}_t$ is given by a smooth, deterministic function.

It is not hard to guess what this function is. Indeed, denoting the generator of the process in (3.2) by $\overline{G}_{N,\beta}$, we see that the *local drift* of the process $\overline{X}$ is given by

$$\mathbb{E}^{\overline{x}}[\overline{X}_t] = \overline{x} + t g_\beta(\overline{x}) + O(t^2) \quad \text{where} \quad g_\beta(\overline{x}) := \overline{G}_{N,\beta} f(\overline{x}) \quad \text{with} \quad f(\overline{x}) := \overline{x},$$

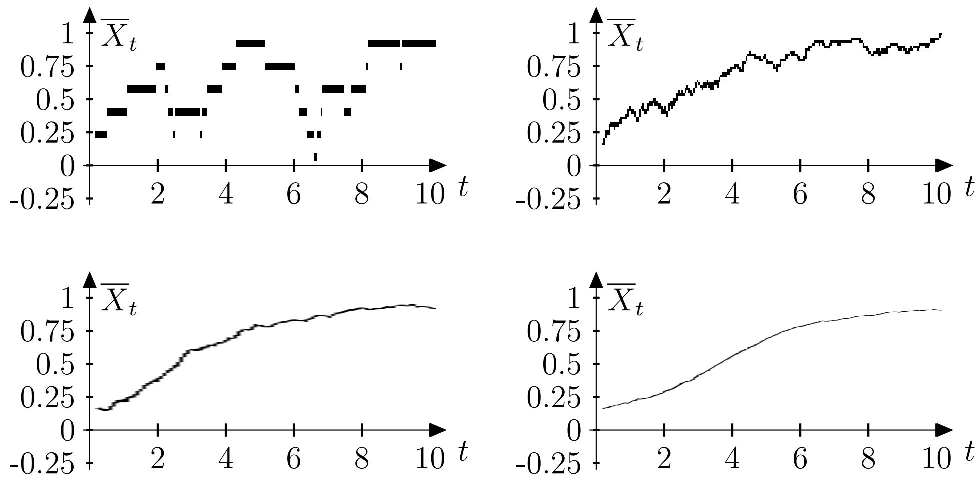


Figure 3.1: The mean-field Ising model on lattice with $N = 10$, 100, 1000, and $10,000$ sites, respectively. In these simulations, the parameter is $\beta = 3$, and the initial state is $\overline{X}_0 = 0.1$, except in the first picture, where $\overline{X}_0 = 0.2$.

where

$$\begin{aligned} g_\beta(\overline{x}) &= r_+(\overline{x}) \cdot \frac{2}{N} - r_-(\overline{x}) \cdot \frac{2}{N} = \frac{(1-\overline{x})e^{\beta\overline{x}/2} - (1+\overline{x})e^{-\beta\overline{x}/2}}{e^{\beta\overline{x}/2} + e^{-\beta\overline{x}/2}} \\ &= \frac{e^{\beta\overline{x}/2} - e^{-\beta\overline{x}/2}}{e^{\beta\overline{x}/2} + e^{-\beta\overline{x}/2}} - \overline{x} = \tanh(\tfrac{1}{2}\beta\overline{x}) - \overline{x}. \end{aligned} \tag{3.3}$$

Note that the constant N cancels out of this formula. In view of this, by some law of large numbers (that will be made rigorous in Theorem 3.2 below), we expect $(\overline{X}_t)_{t\geq 0}$ to converge in distribution, as $N \to \infty$, to a solution of the differential equation

$$\tfrac{\partial}{\partial t}\overline{X}_t = g_\beta(\overline{X}_t) \qquad (t \geq 0). \tag{3.4}$$

3.3 Analysis of the mean-field model

Assuming the correctness of (3.4) for the moment, we can study the behavior of the mean-field Ising model $\overline{X}$ in the limit that we first send $N \to \infty$, and then $t \to \infty$. A simple analysis of the function g_β (see Figure 3.2) reveals that the differential equation (3.4) has a single fixed point for $\beta \leq 2$, and three fixed points for $\beta > 2$.

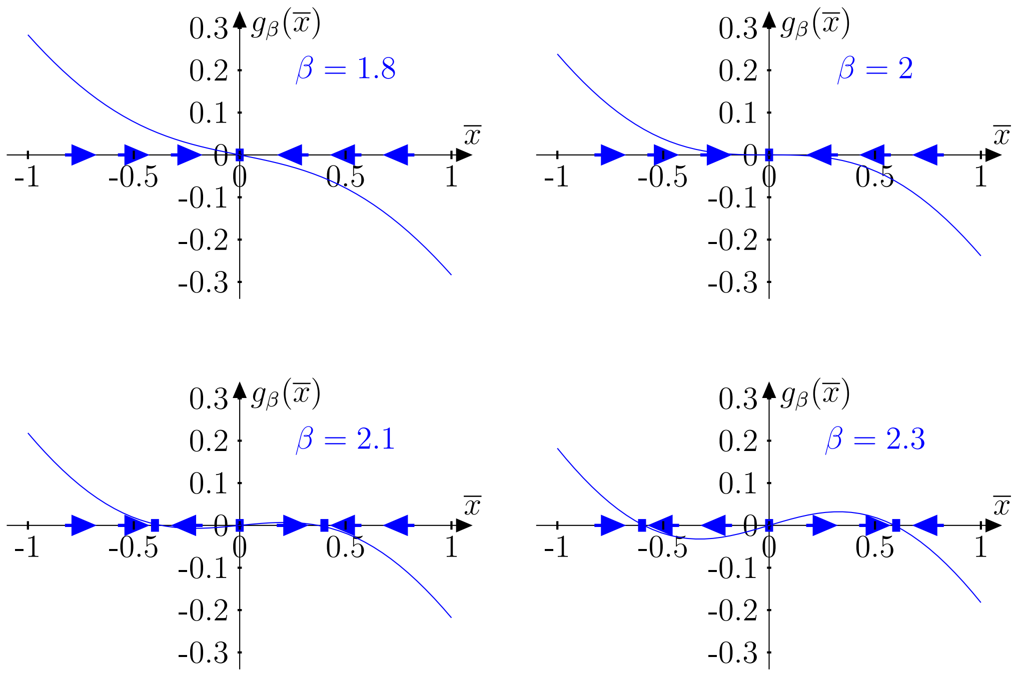


Figure 3.2: The drift function g_β for $\beta = 1.8$, 2, 2.1, and 2.3, respectively. For $\beta > 2$, the fixed point $\overline{x} = 0$ becomes unstable and two new fixed points appear.

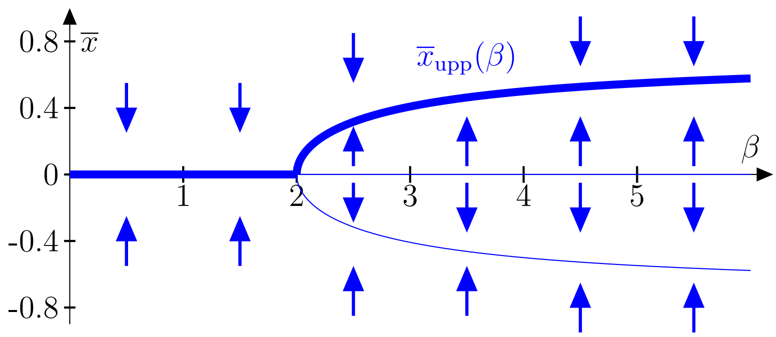


Figure 3.3: Fixed points of the mean-field Ising model as a function of β, with their domains of attraction. The upper fixed point as a function of β is indicated with a bold line.

Indeed, using the facts that tanh is an odd function that is concave on $[0, \infty)$ and satisfies $\frac{\partial}{\partial x} \tanh(x)|_{x=0} = 1$, we see that:

- For $\beta \leq 2$, the equation $g_\beta(x) = 0$ has the unique solution $x = 0$.

- For $\beta > 2$, the equation $g_\beta(x) = 0$ has three solutions $x_- < 0 < x_+$.

For $\beta \leq 2$, solutions to the differential equation (3.4) converge to the unique fixed point $x = 0$ as time tends to zero. On the other hand, for $\beta > 2$, the fixed point $x = 0$ becomes unstable. Solutions $\overline{X}$ to the differential equation (3.4) starting in $\overline{X}_0 > 0$ converge to x_+, while solutions starting in $\overline{X}_0 < 0$ converge to x_-.

In Figure 3.3, we have plotted the three fixed points $x_- < 0 < x_+$ as a function of β, and indicated their domains of attraction. The function

$$x_{\rm upp}(\beta) := \begin{cases} 0 & \text{if } \beta \leq 2, \\ \text{the unique pos. sol. of } \tanh(\tfrac{1}{2}\beta x) = x & \text{if } \beta > 2 \end{cases} \tag{3.5}$$

plays a similar role as the spontaneous magnetization $m_*(\beta)$ for the Ising model on $\mathbb{Z}^d$ (see formula (1.13)). More precisely, for mean-field processes started in initial states $\overline{X}_0 > 0$, the quantity $x_{\rm upp}$ describes the double limit

$$\lim_{t\to\infty} \lim_{N\to\infty} \overline{X}_t = x_{\rm upp}. \tag{3.6}$$

We see from (3.5) that the mean-field Ising model (as formulated in (3.1)) exhibits a second-order (i.e., continuous) phase transition at the critical point $\beta_{\rm c} = 1$. Since

$$x_{\rm upp}(\beta) \propto (\beta - \beta_{\rm c})^{1/2} \quad \text{as} \quad \beta \downarrow \beta_{\rm c},$$

the *mean-field critical exponent* associated with the magnetization[1] is $c = 1/2$, which is the same as for the Ising model on $\mathbb{Z}^d$ in dimensions $d \geq 4$ (see Section 1.5). Understanding why the mean-field model correctly predicts the critical exponent in sufficiently high dimensions goes beyond the scope of the present chapter.

To conclude the present section, we note that the two limits in (3.6) cannot be interchanged. Indeed, for each fixed N, the Markov process $\overline{X}$ is irreducible, and hence, by standard theory, has a unique equilibrium law that is the long-time of the law at time t, started from an arbitrary initial state. In view of the symmetry of the problem, the magnetization in equilibrium must be zero, so regardless of the initial state, we have, for each fixed N,

$$\lim_{t\to\infty} \mathbb{E}[\overline{X}_t] = 0.$$

The reason why this can be true while at the same time (3.6) also holds is that the speed of convergence to equilibrium of the Markov process $\overline{X}$ becomes very slow as $N \to \infty$.

[1] In general, for a given second-order phase transition, there are several quantities of interest that all show power-law behavior near the critical point, and hence there are also several critical exponents associated with a given phase transition.

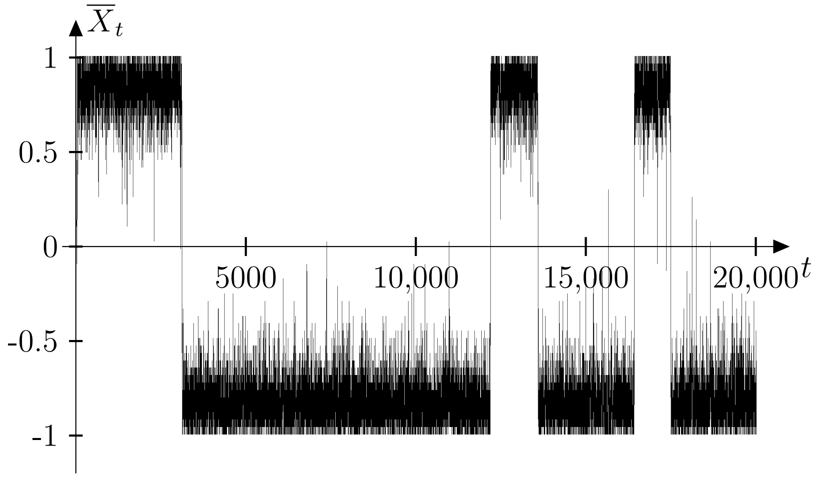


Figure 3.4: Metastable behavior of a mean-field Ising model with $N = 50$ and $\beta = 3$. Note the different time scale compared to Figure 3.1.

In Figure 3.4, we have plotted the time evolution of a mean-field Ising model $\overline{X}$ on a lattice with $N = 50$ sites, for a value of β above the critical point (concretely $\beta = 3$, which lies above $\beta_c = 2$). Although the average of $\overline{X}$ in the long run is 0, we see that the process spends most of its time around the values $x_{\rm upp}$ and $-x_{\rm upp}$, with rare transitions between the two. This sort of behavior is called *metastable behavior*.

The value $N = 50$ was near the highest possible value for which I could still numerically observe this sort of behavior. For $N = 100$ the transitions between the two metastable states $x_{\rm upp}$ and $-x_{\rm upp}$ become so rare that my program was no longer able to see them within a reasonable runtime. With the help of *large deviations theory*, one can show that the time that the system spends in one metastable state is approximately exponentially distributed (with a large mean), and calculate the asymptotics of the mean waiting time as $N \to \infty$. It turns out that the mean time one has to wait for a transition grows exponentially fast in N.

3.4 Functions of Markov processes

In the present section we formulate a proposition and a theorem that we have already implicitly used. Both are concerned with functions of Markov processes. Let $X = (X_t)_{t\geq 0}$ be a Markov process with finite state space S, generator G, and semigroup $(P_t)_{t\geq 0}$. Let T be another finite set and let

$f : S \to T$ be a function. For each $x \in S$ and $y' \in T$ such that $f(x) \neq y'$, let

$$\mathcal{H}(x, y') := \sum_{x' \in S:\, f(x') = y'} G(x, x') \tag{3.7}$$

be the total rate at which $f(X_t)$ jumps to the state y', when the present state is $X_t = x$. The next proposition says that if these rates are a function of $f(x)$ only, then the process $Y = (Y_t)_{t \geq 0}$ defined by

$$Y_t := f(X_t) \qquad (t \geq 0) \tag{3.8}$$

is itself a Markov process.

Proposition 3.1 (Autonomous Markov process) *Assume that the rates in (3.7) are of the form*

$$\mathcal{H}(x, y') = H\big(f(x), y'\big) \qquad (x \in S,\ y' \in T,\ f(x) \neq y') \tag{3.9}$$

where H is a Markov generator of some process in T. Then the process Y defined in (3.8) is a Markov process with generator H.

Proof Let $(Q_t)_{t \geq 0}$ be the semigroup generated by H. We claim that

$$Q_t\big(f(x), y'\big) = \sum_{x':\, f(x') = y'} P_t(x, x') \qquad \big(t \geq 0,\ x \in S,\ y' \in T\big). \tag{3.10}$$

By (2.5), it then follows that the finite dimensional distributions of Y are given by

$$\begin{aligned}
&\mathbb{P}\big[Y_0 = y_0, \ldots, Y_{t_n} = y_n\big] \\
&\quad = \sum_{x_0, \ldots, x_n} \mathbb{P}[X_0 = x_0] P_{t_1 - t_0}(x_0, x_1) \cdots P_{t_n - t_{n-1}}(x_{n-1}, x_n) \\
&\quad = \mathbb{P}[Y_0 = y_0] Q_{t_1 - t_0}(y_0, y_1) \cdots Q_{t_n - t_{n-1}}(y_{n-1}, y_n),
\end{aligned}$$

where we sum over all $x_0, \ldots, x_n$ such that $f(x_0) = y_0, \ldots, f(x_n) = y_n$. Again by (2.5), this implies that Y is a Markov process with generator H.

It therefore suffices to prove (3.10). We observe that for any $g : T \to \mathbb{R}$,

$$\begin{aligned}
P_t(g \circ f)(x) &= \sum_{x'} P_t(x, x') g(f(x')) = \sum_{y'} \sum_{x':\, f(x') = y'} P_t(x, x') g(y'), \\
(Q_t g) \circ f(x) &= \sum_{y'} Q_t(f(x), y') g(y').
\end{aligned}$$

The right-hand sides of these equations are equal for all $g : T \to \mathbb{R}$ if and only if (3.10) holds, so (3.10) is equivalent to the statement that

$$P_t(g \circ f) = (Q_t g) \circ f \qquad (t \geq 0)$$

for all $g : T \to \mathbb{R}$. Writing $P_t = e^{Gt} = \sum_{n=0}^{\infty} \frac{1}{n!}(Gt)^n$ and likewise $Q_t = e^{Ht}$, we see that our claim follows from (and is in fact equivalent to)

$$G(g \circ f) = (Hg) \circ f \qquad (t \geq 0,\ g : T \to \mathbb{R}).$$

By a similar calculation as above, this is equivalent to

$$H\big(f(x), y'\big) = \sum_{x' :\, f(x') = y'} G(x, x') \qquad \big(x \in S,\ y' \in T\big).$$

If $f(x) \neq y'$, then this is our assumption (3.9). To see that we also have equality if $f(x) = y'$, we note that

$$\begin{aligned} H\big(f(x), f(x)\big) &= -\sum_{y' :\, y' \neq f(x)} H(f(x), y') = -\sum_{y' :\, y' \neq f(x)} \sum_{x' :\, f(x') = y'} G(x, x') \\ &= -\sum_{x' :\, f(x') \neq f(x)} G(x, x') = \sum_{x' :\, f(x') = f(x)} G(x, x'), \end{aligned}$$

where we have used that since H and G are Markov generators, one has $\sum_{y' \in T} H(f(x), y') = 0$ and $\sum_{x' \in S} G(x, x') = 0$. ∎

Summarizing, Proposition 3.1 says that if $Y_t = f(X_t)$ is a function of a Markov process, and the jump rates of Y are a function of the present state of Y only (and do not otherwise depend on the state of X), then Y is itself a Markov process. In such a situation, we will say that Y is an *autonomous* Markov process. We have already implicitly used Proposition 3.1 in Section 3.2, when we claimed that the process $\overline{X}$ is a Markov process with jump rates as in (3.2).

Our next aim is to make the claim rigorous that for large N, the process $\overline{X}$ can be aproximated by solutions to the differential equation (3.4). We will apply a theorem from [DN08]. Although the proof is not very complicated, it is a bit lengthy and would detract from our main objects of interest here, so we only show how the theorem below can be deduced from a theorem in [DN08]. That paper also treats the multi-dimensional case and gives explicit estimates on probabilities of the form (3.14) below.

For each $N \geq 1$, let $X^N = (X^N_t)_{t \geq 0}$ be a Markov process with finite state space S_N, generator G_N, and semigroup $(P^N_t)_{t \geq 0}$, and let $f_N : S_N \to \mathbb{R}$ be functions. We will be interested in conditions under which the processes

$(f_N(X_t^N))_{t\geq 0}$ apprioximate the solution $(y_t)_{t\geq 0}$ of a differential equation, in the limit $N \to \infty$. Note that we do not require that $f_N(X_t^N)$ is an autonomous Markov process. To ease notation, we will sometimes drop the super- and subscripts N when no confusion arises.

We define two functions $\alpha = \alpha_N$ and $\beta = \beta_N$ that describe the quadratic variation and drift, respectively, of the process $f(X_t)$. More precisely, these functions are given by

$$\begin{aligned}\alpha(x) &:= \sum_{x'\in S} G(x,x')\big(f(x') - f(x)\big)^2,\\ \beta(x) &:= \sum_{x'\in S} G(x,x')\big(f(x') - f(x)\big).\end{aligned}$$

The idea is that if α tends to zero and β approximates a nice, Lipschitz contunuous function of $f(X_t)$, then $f(X_t)$ should in the limit be given by the solution of a differential equation.

We assume that the functions f_N all take values in a closed interval $I \subset \mathbb{R}$ with left and right boundaries $I_- := \inf I$ and $I_+ := \sup I$, which may be finite or infinite. We also assume that there exists a globally Lipschitz function $b : I \to \mathbb{R}$ such that

$$\sup_{x\in S_N} \big|\beta_N(x) - b\big(f_N(x)\big)\big| \underset{N\to\infty}{\longrightarrow} 0, \tag{3.11}$$

i.e., the drift function β is uniformly approximated by $b \circ f_N$. Assuming also that

$$b(I_-) \geq 0 \quad \text{if } I_- > -\infty \quad \text{and} \quad b(I_+) \leq 0 \quad \text{if } I_+ < -\infty, \tag{3.12}$$

the differential equation

$$\tfrac{\partial}{\partial t} y_t = b(y_t) \qquad (t \geq 0)$$

has a unique I-valued solution $(y_t)_{t\geq 0}$ for each initial state $y_0 \in I$. The following theorem gives sufficient conditions for the I-valued processes $(f_N(X_t^N))_{t\geq 0}$ to approximate a solution of the differential equation.

Theorem 3.2 (Limiting differential equation) *Assume that $f_N(X_0^N)$ converges in probability to y_0 and that apart from (3.11), one has*

$$\sup_{x\in S_N} \alpha_N(x) \underset{N\to\infty}{\longrightarrow} 0. \tag{3.13}$$

Then, for each $T < \infty$ and $\varepsilon > 0$,

$$\mathbb{P}\big[|f_N(X_t^N) - y_t| \leq \varepsilon \ \forall t \in [0,T]\big] \underset{N\to\infty}{\longrightarrow} 1. \tag{3.14}$$

Proof We apply [DN08, Thm 4.1]. Fix $T < \infty$ and $\varepsilon > 0$ and also fix $y_0 \in I$. Let L denote the Lipschitz constant of b. The assumptions of [DN08, Thm 4.1] allow for the case that f_N does not in general take values in I, but only under the additional condition that $f_N(x)$ not further than ε from a possible value the solution of the differential equation. In our case, these more general assumptions are autimatically satisfied. Set $\delta := \frac{1}{3}\varepsilon e^{-LT}$. We consider the events

$$\Omega_0 := \big\{|f(X_0) - y_0| \leq \delta\big\} \quad \text{and} \quad \Omega_1 := \Big\{\int_0^T |\beta(X_t) - b\big(f(X_t)\big)|\,\mathrm{d}t \leq \delta\Big\}.$$

For $K > 0$, we also define

$$\Omega_{K,2} := \Big\{\int_0^T \alpha(X_t)\,\mathrm{d}t \leq KT\Big\}.$$

Then [DN08, Thm 4.1] tells us that

$$\mathbb{P}\big[\sup_{t\in[0,T]} |f(X_t) - y_t| > \varepsilon\big] \leq 4KT\delta^{-2} + \mathbb{P}\big(\Omega_0^{\rm c} \cup \Omega_1^{\rm c} \cup \Omega_{K,2}^{\rm c}\big). \tag{3.15}$$

Our assumption that $f_N(X_0^N) \to y_0$ in probability implies that $\mathbb{P}(\Omega_0^{\rm c}) \to 0$ as $N \to \infty$. Set

$$A_N := \sup_{x\in S_N} \alpha_N(x) \quad \text{and} \quad B_N := \sup_{x\in S_N} \big|\beta_N(x) - b\big(f_N(x)\big)\big|$$

Then $A_N \to 0$ by (3.13) and $B_N \to 0$ by (3.11). Since

$$\int_0^T |\beta(X_t) - b\big(f(X_t)\big)|\,\mathrm{d}t \leq B_N T \leq \delta$$

for N sufficiently large, we see that $\mathbb{P}(\Omega_1^{\rm c}) = 0$ for N sufficiently large. Also, since

$$\int_0^T \alpha(X_t)\,\mathrm{d}t \leq A_N T,$$

we see that $\mathbb{P}(\Omega_{A_N,2}^{\rm c}) = 0$ for all N. Inserting $K = A_N$ in (3.15), we see that the right-hand side tends to zero as $N \to \infty$. ∎

Using Theorem 3.2, we can make the approximation of the mean-field Ising model by the differential equation (3.4) rigorous. Let $X^N = (X_t^N)_{t\geq 0}$ denote the Markov process with state space $\{-1,+1\}^{\Lambda_N}$, where Λ_N is a set containing N elements and the jump rates of X^N are given in (3.1). By Propositon 3.1, the process $\overline{X}_t^N := \frac{1}{N}\sum_{i\in\Lambda_N} X_t(i)$ is itself a Markov process with jump rates as in (3.2). We can either apply Theorem 3.2 directly to the Markov processes X^N and the functions $f_N(x) := \frac{1}{N}\sum_{i\in\Lambda_N} x(i)$, or we can apply Theorem 3.2 to the Markov processes $\overline{X}^N$ and choose for f_N the identity function $f_N(\overline{x}) = \overline{x}$. In either case, the assumption (3.11) is already verified in (3.3). To check also (3.13), we calculate

$$\alpha_N(x) = r_+(\overline{x})\Big(\frac{2}{N}\Big)^2 + r_-(\overline{x})\Big(\frac{2}{N}\Big)^2 = \frac{2}{N}\Big(1+\overline{x}\frac{e^{-\beta\overline{x}/2}-e^{\beta\overline{x}/2}}{e^{-\beta\overline{x}/2}+e^{\beta\overline{x}/2}}\Big),$$

which clearly tends uniformly to zero as $N\to\infty$.

3.5 The mean-field contact process

Recall the definition of the generator of the contact process from (1.8). We slightly reformulate this as

$$\begin{aligned} G_{\mathrm{cont}}f(x) := \lambda \sum_{(i,j)\in\mathcal{E}} &\big\{f\big(\big(\mathtt{bra}_{ij}(x)\big) - f\big(x\big)\big\} \\ + \sum_{i\in\mathbb{Z}^d} &\big\{f\big(\big(\mathtt{death}_i(x)\big) - f\big(x\big)\big\} \qquad (x\in\{0,1\}^\Lambda), \end{aligned} \tag{3.16}$$

where as customary we have set the death rate to $\delta = 1$, and we have also reparametrized the infection rate so that λ denotes the total rate of all outgoing infections from a given site, instead of the infection rate per neighbor.

We will be interested in the contact process on the complete graph, which means that we take for $\Lambda = \Lambda_N$ a set with N elements, which we equip with the structure of a complete graph with (undirected) edge set $E = E_N := \{\{i,j\} : i,j\in\Lambda_N,\ i\neq j\}$ and corresponding set of oriented edges $\mathcal{E} = \mathcal{E}_N$. We will be interested in the fraction of infected sites

$$\overline{X}_t = \overline{X}^N_t := \frac{1}{N}\sum_{i\in\Lambda_N} X_t(i) \qquad (t\geq 0),$$

which jumps with the following rates

$$\begin{aligned} &\overline{x}\mapsto\overline{x}+\tfrac{1}{N} \qquad \text{with rate} \qquad r_+(\overline{x}) := \lambda N\overline{x}(1-\overline{x}), \\ &\overline{x}\mapsto\overline{x}-\tfrac{1}{N} \qquad \text{with rate} \qquad r_-(\overline{x}) := N\overline{x}. \end{aligned} \tag{3.17}$$

Here $N(1-\overline{x})$ is the number of healthy sites, each of which gets infected with rate $\lambda\overline{x}$, and $N\overline{x}$ is the number of infected sites, each of which recovers with rate one. Note that since these rates ar a function of $\overline{x}$ only, by Proposition 3.1, the process $(\overline{X}_t)_{t\geq 0}$ is an autonomous Markov chain.

We wish to apply Theorem 3.2 to conclude that $\overline{X}$ can, for large N¡ be approximated by the solution of a differential equation. To this aim, we calculate the drift β and quadractic variation function α.

$$\alpha_N(x) = r_+(\overline{x})\tfrac{1}{N^2} + r_-(\overline{x})\tfrac{1}{N^2} = \frac{1}{N}\big(\lambda\overline{x}(1-\overline{x}) + \overline{x}\big),$$
$$\beta_N(x) = r_+(\overline{x})\tfrac{1}{N} - r_-(\overline{x})\tfrac{1}{N} = \lambda\overline{x}(1-\overline{x}) - \overline{x}.$$

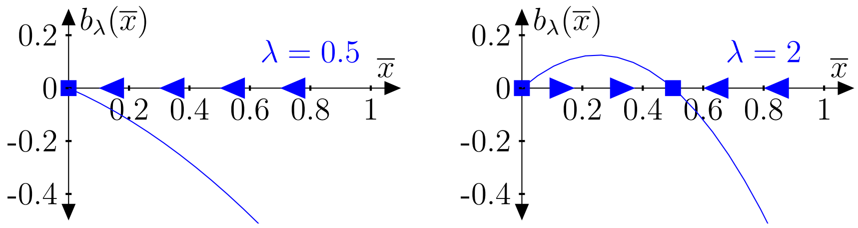
By Theorem 3.2, it follows that in the mean-field limit $N \to \infty$, the fraction of infected sites can be approximated by solutions of the differential equation

$$\tfrac{\partial}{\partial t}\overline{X}_t = b_\lambda(\overline{X}_t) \quad (t \geq 0), \quad \text{where} \quad b_\lambda(\overline{x}) := \lambda\overline{x}(1-\overline{x}) - \overline{x}. \tag{3.18}$$

The equation $b_\lambda(\overline{x}) = 0$ has the solutions

$$\overline{x} = 0 \quad \text{and} \quad \overline{x} = 1 - \lambda^{-1}.$$

The second solution lies inside the interval $[0,1]$ of possible values of $\overline{X}_t$ if and only if $\lambda \geq 1$. Plotting the function b_λ for $\lambda < 1$ and $\lambda > 1$ yields the following pictures.



We see from this that the fixed point $\overline{x} = 0$ is stable for $\lambda \leq 1$ but becomes unstable for $\lambda > 1$, in which case $\overline{x} = 1 - \lambda^{-1}$ is the only stable fixed point that attracts all solutions started in a nonzero initial state. The situation is summarized in Figure 3.5.

Letting $\overline{x}_{\rm upp}(\lambda) := 0 \vee (1 - \lambda^{-1})$ denote the stable fixed point, we see that the mean-field contact process exhibits a second-order phase transition at the critical point $\lambda_{\rm c} = 1$. Since

$$\overline{x}_{\rm upp}(\lambda) \propto (\lambda - \lambda_{\rm c}) \quad \text{as } \lambda \downarrow \lambda_{\rm c},$$

the associated critical point is $c = 1$, in line with what we know for contact processes in dimensions $d \geq 4$ (see the discussion in Section 1.5).

3.6 The mean-field voter model

Recall the definition of the generator of the voter model from (1.5). For simplicity, we will only consider the two-type model and as the local state

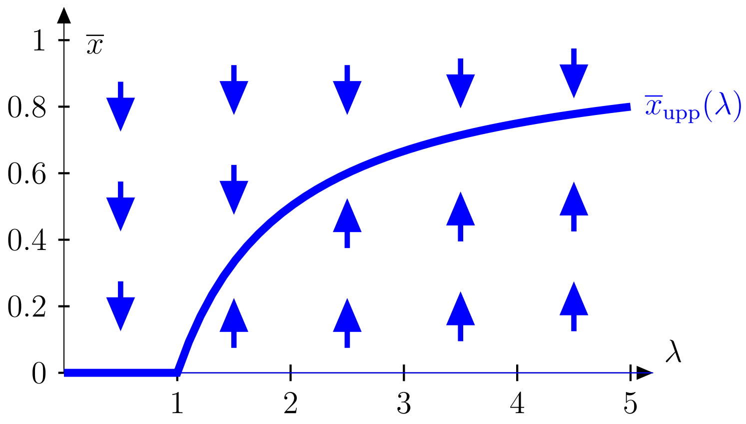


Figure 3.5: Mean-field analysis of the contact process.

space we will choose $S = \{0, 1\}$. Specializing to the complete graph $\Lambda = \Lambda_N$ with N vertices, the generator becomes

$$G_{\rm vot} f(x) = \frac{1}{|\Lambda|} \sum_{(i,j)\in\mathcal{E}} \big\{ f\big(\big(\mathtt{vot}_{ij}(x)\big) - f\big(x\big)\big\} \qquad (x \in \{0,1\}^\Lambda). \tag{3.19}$$

Note that the factor $|\Lambda|^{-1}$ says that each site i updates its type with rate one, and at such an event chooses a new type from a uniformly chosen site j (allowing for the case $i = j$, which has no effect).

We are interested in the fraction of sites of type 1,

$$\overline{X}_t = \overline{X}_t^N := \frac{1}{N} \sum_{i\in\Lambda_N} X_t(i) \qquad (t \geq 0),$$

which jumps as (compare (3.17))

$$\begin{aligned} \overline{x} &\mapsto \overline{x} + \tfrac{1}{N} \qquad \text{with rate} \qquad r_+(\overline{x}) := N\overline{x}(1 - \overline{x}), \\ \overline{x} &\mapsto \overline{x} - \tfrac{1}{N} \qquad \text{with rate} \qquad r_-(\overline{x}) := N\overline{x}(1 - \overline{x}). \end{aligned}$$

Note that $N(1 - \overline{x})$ is the number of sites of type 0, and that each such site adopts the type 1 with rate $\overline{x}$. The derivation of $r_-(\overline{x})$ is similar. We calculate the drift β and quadractic variation function α.

$$\begin{aligned} \alpha_N(x) &= r_+(\overline{x})\tfrac{1}{N^2} + r_-(\overline{x})\tfrac{1}{N^2} = \frac{2}{N}\overline{x}(1 - \overline{x}), \\ \beta_N(x) &= r_+(\overline{x})\tfrac{1}{N} - r_-(\overline{x})\tfrac{1}{N} = 0. \end{aligned}$$

Applying Theorem 3.2, we see that in the limit $N \to \infty$, the process $(\overline{X}_t)_{t\geq 0}$ is well approximated by solutions to the differential equation

$$\tfrac{\partial}{\partial t}\overline{X}_t = 0 \qquad (t \geq 0),$$

i.e., $\overline{X}_t$ is approximately constant as a function of t.

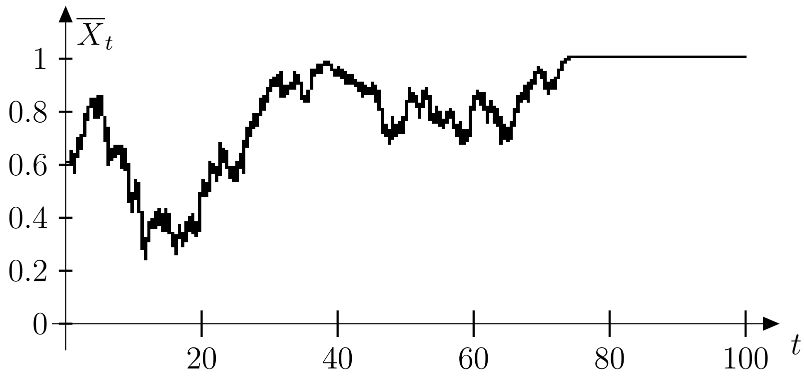


Figure 3.6: The fraction of type 1 individuals in the mean-field voter model from (3.20) on $N = 100$ sites. This process approximates the Wright-Fisher diffusion.

Of course, if we go to larger time scales, then $\overline{X}_t$ will no longer be constant; compare Figure 3.4. In fact, we can determine the time scale at which $\overline{X}_t$ fluctuates quite precisely. Scaling up time by a factor $|\Lambda| = N$ is the same as multiplying all rates by a factor $|\Lambda|$. If we repeat our previous calculations for the process with generator

$$G_{\rm vot} f(x) = \sum_{(i,j)\in\mathcal{E}} \big\{f\big(\big(\mathtt{vot}_{ij}(x)\big) - f\big(x\big)\big\} \qquad (x \in \{0,1\}^\Lambda), \tag{3.20}$$

then the drift and quadractic variation are given by

$$\begin{aligned} \alpha_N(x) &= 2\overline{x}(1-\overline{x}),\\ \beta_N(x) &= 0. \end{aligned}$$

In this case, the quadratic variation does not go to zero, so Theorem 3.2 is no longer applicable. One can show, however, that in the limit $N \to \infty$ the new, sped-up process is well approximated by solutions to the (Itô) stochastic differential equation (SDE)

$$\mathrm{d}\overline{X}_t = \sqrt{2\overline{X}_t(1-\overline{X}_t)}\,\mathrm{d}B_t \qquad (t \geq 0),$$

where $2\overline{X}_t(1-\overline{X}_t) = \alpha(X_t)$ is of course the quadratic variation function we have just calculated. Solutions to this SDE are *Wright-Fisher diffusions* , i.e., Markov processes with continuous sample paths and generator

$$Gf(\overline{x}) = \overline{x}(1-\overline{x})\tfrac{\partial^2}{\partial \overline{x}^2} f(\overline{x}). \tag{3.21}$$

These calculations can be made rigorous using methods from the theory of convergence of Markov processes; see, e.g., the book [EK86]. See Figure 3.6 for a simulation of the process $\overline{X}$ when X has the generator in (3.20) and $N = 100$.

3.7 Exercises

Exercise 3.3 *Do a mean-field analysis of the process with generator*

$$\begin{aligned}Gf(x) = b|\Lambda|^{-2} &\sum_{ii'j} \big\{f\big(\texttt{coop}_{ii'j}x\big) - f\big(x\big)\big\} \\ &+ \sum_i \big\{f\big(\texttt{death}_i x\big) - f\big(x\big)\big\},\end{aligned}$$

where the maps $\texttt{coop}_{ii'j}$ *and* $\texttt{death}_i$ *are defined in (1.23) and (1.7), respectively. Do you observe a phase transition? Is it first- or second order? Hint: Figure 3.7.*

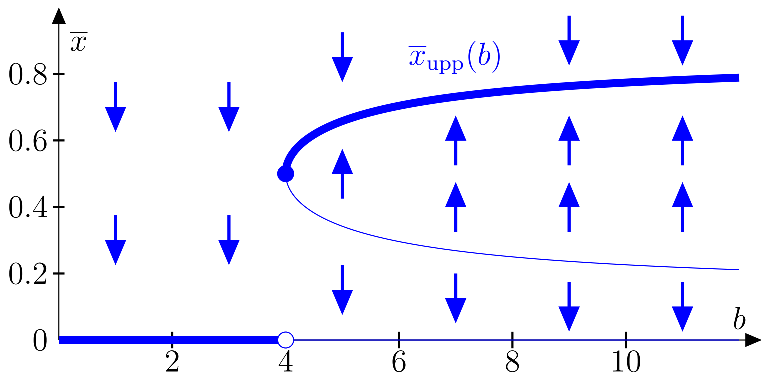


Figure 3.7: Mean-field analysis of a model with cooperative branching and deaths.

Exercise 3.4 *Same as above for the model with generator*

$$\begin{aligned} Gf(x) = b|\Lambda|^{-2} \sum_{ii'j} \big\{ f\big(\mathtt{coop}_{ii'j} x\big) - f\big(x\big) \big\} \\ + |\Lambda|^{-1} \sum_{ij} \big\{ f\big(\mathtt{rw}_{ij} x\big) - f\big(x\big) \big\}. \end{aligned}$$

Exercise 3.5 *Derive an SDE in the limit $|\Lambda| \to \infty$ for the density of the mean-field voter model with small bias and death rates, with generator*

$$\begin{aligned} Gf(x) = |\Lambda|^{-2} \sum_{ij \in \Lambda} \big\{ f\big(\mathtt{vot}_{ij} x\big) - f\big(x\big) \big\} \\ + s|\Lambda|^{-1} \sum_{ij \in \Lambda} \big\{ f\big(\mathtt{bra}_{ij} x\big) - f\big(x\big) \big\} \\ + d \sum_{i \in \Lambda} \big\{ f\big(\mathtt{death}_{i} x\big) - f\big(x\big) \big\}. \end{aligned}$$

Hint: *You should find expressions of the form*

$$\begin{aligned} \mathbb{E}^{\overline{x}}\big[(\overline{X}_t - \overline{x})\big] &= b(\overline{x}) \cdot t + O(t^2), \\ \mathbb{E}^{\overline{x}}\big[(\overline{X}_t - \overline{x})^2\big] &= a(\overline{x}) \cdot t + O(t^2), \end{aligned}$$

which leads to a limiting generator of the form

$$Gf(\overline{x}) = \tfrac{1}{2} a(\overline{x}) \tfrac{\partial^2}{\partial \overline{x}^2} f(\overline{x}) + b(\overline{x}) \tfrac{\partial}{\partial x} f(\overline{x}).$$

Exercise 3.6 *Do a mean-field analysis of the following extension of the voter model, introduced in [NP99]. In this model, the site i flips*

$$\begin{aligned} 0 \mapsto 1 \quad & \textit{with rate} \quad \big(f_0 + \alpha_{01} f_1\big) f_1, \\ 1 \mapsto 0 \quad & \textit{with rate} \quad \big(f_1 + \alpha_{10} f_0\big) f_0, \end{aligned}$$

where $\alpha_{01}, \alpha_{10} > 0$ and $f_\tau = |\mathcal{N}_i|^{-1} \sum_{j \in \mathcal{N}_i} 1_{\{x(j) = \tau\}}$ is the relative frequency of type τ in the neigborhood of i.

Find all stable and unstable fixed points of the mean-field model in the regimes: I. $\alpha_{01}, \alpha_{10} < 1$, II. $\alpha_{01} < 1 < \alpha_{10}$, III. $\alpha_{10} < 1 < \alpha_{01}$, IV. $1 < \alpha_{01}, \alpha_{10}$.

Chapter 4

Construction and ergodicity

4.1 Introduction

As explained in Chapter 1, interacting particle systems are Markov processes with a state space of the form S^Λ where S is a finite set, called the *local state space*, and Λ is a countable set, called the *lattice*. The generator of an interacting particle system can usualy be written in the form

$$Gf(x) = \sum_{m\in\mathcal{G}} r_m\big\{f\big(m(x)\big) - f\big(x\big)\big\} \qquad (x \in S^\Lambda), \tag{4.1}$$

where $\mathcal{G}$ is a set whose elements are *local maps* $m : S^\Lambda \to S^\Lambda$ and $(r_m)_{m\in\mathcal{G}}$ is a collection of nonnegative rates. If Λ is finite, then S^Λ is also a finite set and we can use Proposition 2.5 to construct a Markov process $X = (X_t)_{t\geq 0}$ with generator G in terms of a Poisson process ω.

On the other hand, if Λ is countable but infinite, then the space S^Λ is not finite, and, in fact, not even countable. Indeed, as is well-know, $\{0,1\}^{\mathbb{N}}$ has the same cardinality as the real numbers. As a result, the construction of interacting particle systems on infinite lattices is considerably more involved than in the finite case. Nevertheless, we will see that they can be constructed using more or less the same approach as in Proposition 2.5. The only complication is that the total rate of all local maps is usually infinite, so that it is no longer possible to order the elements of the Poisson set ω according to the time when they occur. However, since each map is *local*, and since in finite time intervals only finitely many local maps can influence the local state at any given site i, we will see that under certain summability assumptions, the Poisson construction still yields a well-defined process.

In practice, one usualy needs not only the Poisson construction of an interacting particle system, but also wishes to show that the process is uniquely characterized by its generator. One reason is that, as we have already seen in Section 2.4, sometimes the same process can be constructed using different Poisson constructions, and one wants to prove that these constructions are indeed equivalent.

To give a generator construction of interacting particle systems, we will apply the theory of Feller processes. We start by equipping S^Λ with the *product topology*, which says that a sequence $x_n \in S^\Lambda$ converges to a limit x if and only if

$$x_n(i) \underset{n\to\infty}{\longrightarrow} x(i) \qquad \forall i \in \Lambda.$$

Note that since S is finite, this says simply that for each $i \in \Lambda$, there is an N (which may depend on i) such that $x_n(i) = x(i)$ for all $n \geq N$. Since S is finite, it is in particular compact, so by Tychonoff's theorem, the space S^Λ is compact in the product topology. The product topology is metrizable. For example, if $(a_i)_{i\in\Lambda}$ are strictly positive constants such that $\sum_{i\in\Lambda} a_i < \infty$, then

$$d(x,y) := \sum_{i\in\Lambda} a_i 1_{\{x(i)\neq y(i)\}}$$

defines a metric that generates the product topology.

In Section 4.2, we will collect some general facts about Feller processes, which are a class of Markov processes with compact, metrizable state spaces, that are uniquely characterized by their generators. Since this is rather functional theoretic material, which is moreover well-known, we will state the main facts without proof, but give references to places where proofs can be found.

In Section 4.3, we then give the Poisson construction of interacting particle systems (including proofs). In Section 4.4, we show that our construction yields a Feller process and determine its generator.

Luckily, all this abstract theory gives us more that just the information that the systems we are interested in are well defined. In Section 4.5, we will see that as a side-result of our proofs, we can derive sufficient conditions for an interacting particle system to be ergodic, i.e., to have a unique invariant law that is the long-time limit starting from any initial state. We will apply this to derive lower bounds on the critical points of the Ising model and contact process. Applications to other interacting particle systems are directed to the exercises.

4.2 Feller processes

In Section 2.2, we gave a summary of the basic theory of continuous-time Markov processes with finite state space S. In the present section, we will see that with a bit of care, much of this theory can be generalized in a rather elegant way to Markov processes taking values in a compact metrizable state space. The basic assumption we will make is that the transition probabilities $(P_t)_{t\geq 0}$ are continuous, which means that we will be discussing *Feller processes*.

Let E be a compact metrizable space.[1] We use the notation

$$\begin{aligned}
\mathcal{B}(E) &:= \text{the Borel-}\sigma\text{-field on } E,\\
B(E) &:= \text{the space of bounded, Borel-measurable functions } f : E \to \mathbb{R},\\
\mathcal{C}(E) &:= \text{the space of continuous functions} f : E \to \mathbb{R},\\
\mathcal{M}_1(E) &:= \text{the space of probability measures } \mu \text{ on } E.
\end{aligned}$$

We equip $\mathcal{C}(E)$ with the supremumnorm

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

making $\mathcal{C}(E)$ into a Banach space. We equip $\mathcal{M}_1(E)$ with the topology of weak convergence, where by definition, μ_n converges weakly to μ, denoted $\mu_n \Rightarrow \mu$, if $\int f\,\mathrm{d}\mu_n \to \int f\,\mathrm{d}\mu$ for all $f \in \mathcal{C}(E)$. With this topology, $\mathcal{M}_1(E)$ is a compact metrizable space.

A *probability kernel* on E is a function $K : E \times \mathcal{B}(E) \to \mathbb{R}$ 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 $A \in \mathcal{B}(E)$.

This is equivalent to the statement that $x \mapsto K(x, \cdot\,)$ is a measurable map from E to $\mathcal{M}_1(E)$ (where the latter is equipped with the topology of weak convergence and the associated Borel-σ-field). By definition, a probability kernel is *continuous* if the map $x \mapsto K(x, \cdot\,)$ is continuous (with respect to the topologies which which we have equipped these spaces).

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 define the *composition* of two probability kernels K, L as

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

[1]Such spaces are always separable and complete in any metric that generates the topology; in particular, they are Polish spaces.

Then KL is again a probability kernel on E and the linear operator $(KL) : B(E) \to B(E)$ associated with this kernel is the composition of the linear operators K and L. It follows from the definition of weak convergence that a kernel K is continuous if and only if its associated linear operator maps the space $\mathcal{C}(E)$ into itself. If μ is a probability measure and K is a probability kernel, then

$$(\mu K)(A) := \int \mu(\mathrm{d}x) K(x, A) \qquad \big(A \in \mathcal{B}(E)\big)$$

defines another probability measure μK. Introducing the notation $\mu f := \int f \,\mathrm{d}\mu$, one has $(\mu K) f = \mu(K f)$ for all $f \in B(E)$.

By definition, a *continuous transition probability* on E is a collection $(P_t)_{t\geq 0}$ 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) $P_0 = 1$ and $P_s P_t = P_{s+t} \quad (s, t \geq 0)$.

In particular, (i) implies that each P_t is a continuous probability kernel, so each P_t maps the space $\mathcal{C}(E)$ into itself. One has

$$\begin{array}{rl}
\text{(i)} & \lim_{t\to 0} P_t f = P_0 f = f \qquad (f \in \mathcal{C}(E)),\\
\text{(ii)} & P_s P_t f = P_{s+t} \qquad (s, t \geq 0),\\
\text{(iii)} & f \geq 0 \text{ implies } P_t f \geq 0,\\
\text{(iv)} & P_t 1 = 1,
\end{array}$$

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*. We note that in (i), the limit is (of course) with respect to the topology we have chosen on $\mathcal{C}(E)$, i.e., with respect to the supremumnorm.

By definition, a function $w : [0, \infty) \to E$ is *cadlag* if it is right-continuous with left limits,[2] i.e.,

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

Let $(P_t)_{t\geq 0}$ be a Feller semigroup. By definition a *Feller process* with semi-

[2]The word cadlag is an abbreviation of the French **c**ontinue **à d**roit, **l**imite **à g**auche.

group $(P_t)_{t\geq 0}$ is a stochastic process $X = (X_t)_{t\geq 0}$ with cadlag sample paths[3] such that

$$\mathbb{P}\big[X_u \in \cdot\,\big|\,(X_s)_{0\leq s\leq t}\big] = P_{u-t}(X_t,\,\cdot\,) \quad \text{a.s.} \qquad (0 \leq t \leq u). \tag{4.2}$$

Here we condition on the σ-field generated by the random variables $(X_s)_{0\leq s\leq t}$. Formula (4.2) is equivalent to the statement that the finite dimensional distributions of X are given by

$$\begin{aligned}&\mathbb{P}\big[X_0 \in \mathrm{d}x_0, \ldots, X_{t_n} \in \mathrm{d}x_n\big]\\ &\quad = \mathbb{P}[X_0 \in \mathrm{d}x_0] P_{t_1 - t_0}(x_0, \mathrm{d}x_1) \cdots P_{t_n - t_{n-1}}(x_{n-1}, \mathrm{d}x_n)\end{aligned} \tag{4.3}$$

$(0 < t_1 < \cdots < t_n)$. Formula (4.3) is symbolic notation, which means that

$$\begin{aligned}&\mathbb{E}\big[f(X_0, \ldots, X_{t_n})\big]\\ &= \int \mathbb{P}[X_0 \in \mathrm{d}x_0] \int P_{t_1-t_0}(x_0, \mathrm{d}x_1) \cdots \int P_{t_n-t_{n-1}}(x_{n-1}, \mathrm{d}x_n) f(x_0, \ldots, x_n)\end{aligned}$$

for all $f \in B(E^{n+1})$. By (4.3), the law of a Feller process X is uniquely determined by its initial law $\mathbb{P}[X_0 \in \,\cdot\,]$ and its transition probabilities $(P_t)_{t\geq 0}$. Existence is less obvious than uniqueness, but the next theorem says that this holds in full generality.

Theorem 4.1 (Construction of Feller processes) *Let E be a compact metrizable space, let μ be a probability measure on E, and let $(P_t)_{t\geq 0}$ be a Feller semigroup. Then there exists a Feller process $X = (X_t)_{t\geq 0}$ with initial law $\mathbb{P}[X_0 \in \,\cdot\,] = \mu$, and such a process is unique in distribution.*

Just as in the case for finite state space, we would like to characterize a Feller semigroup by its generator. This is somewhat more complicated than in the finite setting since in general, it is not possible to make sense of the formula $P_t = e^{tG} := \sum_{n=0}^{\infty} \frac{1}{n!}(tG)^n$. This is related to the fact that if G is the generator of a Feller semigroup, then in general it is not possible to define Gf for all $f \in \mathcal{C}(E)$, as we now explain.

Let $\mathcal{V}$ be a Banach space. (In our case, the only Banach spaces that we will need is are spaces of the form $\mathcal{C}(E)$, equipped with the supremunorm.)

[3]It is possible to equip the space $\mathcal{D}_E[0,\infty)$ of cadlag functions $w : [0,\infty) \to E$ 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 $w \mapsto w_t$ $(t \geq 0)$. As a result, we can view a stochastic process $X = (X_t)_{t\geq 0}$ with cadlag sample paths as a single random variable X taking values in the space $\mathcal{D}_E[0,\infty)$. The law of such a random variable is then uniquely determined by the finite dimensional distributions of $(X_t)_{t\geq 0}$.

By definition, a *linear operator* on $\mathcal{V}$ is a pair $(A, \mathcal{D}(A))$ where $\mathcal{D}(A)$ is a linear subspace of $\mathcal{V}$, called the *domain* and A is a linear map $A : \mathcal{D}(A) \to \mathcal{V}$. Even though a linear operator is really a pair $(A, \mathcal{D}(A))$, one often writes sentences such as "let A be a linear operator" without explicitly mentioning the domain. This is similar to phrases like: let $\mathcal{V}$ be a Banach space" (without mentioning the norm) or "let M be a measurable space" (without mentioning the σ-field).

We say that a linear operator A (with domain $\mathcal{D}(A)$) on a Banach space $\mathcal{V}$ is *closed* if and only if its graph $\{(f, Af) : f \in \mathcal{D}(A)\}$ is a closed subset of $\mathcal{V} \times \mathcal{V}$. 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 $\mathcal{D}(\overline{A})$. This operator is then called the *closure* of A. We mention the following theorem.

Theorem 4.2 (Closed graph theorem) *Let $\mathcal{V}$ be a Banach space and let A be a linear operator that is everywhere defined, i.e., $\mathcal{D}(A) = \mathcal{V}$. Then the following statements are equivalent.*

(i) *A is continuous as a map from $\mathcal{V}$ into itself.*

(ii) *A is* bounded, *i.e., there exists a constant $C < \infty$ such that $\|Af\| \leq C\|f\|$ $(f \in \mathcal{V})$.*

(iii) *A is closed.*

Theorem 4.2 shows in particular that if A is an unbounded operator (i.e., there exists $0 \neq f_n \in \mathcal{D}(A)$ such that $\|Af_n\|/\|f_n\| \to \infty$) and A is closable, then its closure $\overline{A}$ will not be everywhere defined. Closed (but possibly unbounded) linear operators are in a sense "the next good thing" after bounded operators.

As before, let E be a compact metrizable space and let $(P_t)_{t\geq 0}$ be a continuous transition probability (or equivalently Feller semigroup) on E. By definition, the *generator* of $(P_t)_{t\geq 0}$ is the linear operator

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

with domain

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

Here, when we say that the limit exists, we mean (of course) with respect to the topology on $\mathcal{C}(E)$, i.e., w.r.t. the supremumnorm.

Since we cannot use the exponential formula $P_t = e^{tG}$, we need another way to characterize $(P_t)_{t\geq 0}$ in terms of G. Let A be a linear operator on $\mathcal{C}(E)$. By definition, we say that a function $[0,\infty) \ni t \mapsto u_t \in \mathcal{C}(E)$ solves the *Cauchy equation*

$$\tfrac{\partial}{\partial t} u_t = A u_t \qquad (t \geq 0) \tag{4.4}$$

if $u_t \in \mathcal{D}(A)$ for all $t \geq 0$, the maps $t \mapsto u_t$ and $t \mapsto G u_t$ are continuous (w.r.t. the topology on $\mathcal{C}(E)$), the limit $\frac{\partial}{\partial t} u_t := \lim_{s\to 0} s^{-1}(u_{t+s} - u_s)$ exists (w.r.t. the topology on $\mathcal{C}(E)$) for all $t \geq 0$, and (4.4) holds. The following proposition shows that a Feller semigroup is uniquely characterized by its generator.

Proposition 4.3 (Cauchy problem) *Let G be the generator of a Feller semigroup $(P_t)_{t\geq 0}$. Then, for each $f \in \mathcal{D}(G)$, the Cauchy equation $\frac{\partial}{\partial t} u_t = G u_t$ $(t \geq 0)$ has a unique solution $(u_t)_{t\geq 0}$ with initial state $u_0 = f$. Denoting this solution by $U_t f := u_t$ defines for each $t \geq 0$ a linear operator U_t with domain $\mathcal{D}(G)$, of which $P_t = \overline{U}_t$ is the closure.*

We need a way to check that (the closure of) a given operator is the generator of a Feller semigroup. For a given linear operator A, constant $\lambda > 0$, and $f \in \mathcal{C}(E)$, we say that a function $p \in \mathcal{C}(E)$ solves the *Laplace equation*

$$(\lambda - A)p = f \tag{4.5}$$

if $p \in \mathcal{D}(A)$ and (4.5) holds. The following lemma shows how solutions to Laplace equations typically arise.

Lemma 4.4 (Laplace equation) *Let G be the generator of a Feller semigroup $(P_t)_{t\geq 0}$ on $\mathcal{C}(E)$, let $\lambda > 0$ and $f \in \mathcal{C}(E)$. Then the Laplace equation $(\lambda - G)p = f$ has a unique solution, that is given by*

$$p = \int_0^\infty P_t f\, e^{-\lambda t} \mathrm{d}t.$$

We say that an operator A on $\mathcal{C}(E)$ with domain $\mathcal{D}(A)$ satisfies the *positive maximum principle* if, whenever a function $f \in \mathcal{D}(A)$ assumes its maximum over E in a point $x \in E$ and $f(x) \geq 0$, we have $Af(x) \leq 0$. The following proposition gives necessary and sufficient conditions for a linear operator G to be the generator of a Feller semigroup.

Theorem 4.5 (Generators of 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 positive maximum principle.*

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

(iv) *For every $f \in \mathcal{C}(E)$ and $\lambda > 0$, the Laplace equation $(\lambda - G)p = f$ has a solution.*

In practice, it is rarely possible to give an explicit description of the (full) domain of a Feller generator. Rather, one often starts with an operator that is defined on a smaller domain of "nice" functions and then takes its closure. Here the following theorem is very useful.

Theorem 4.6 (Hille-Yosida) *A linear operator A on $\mathcal{C}(E)$ with domain $\mathcal{D}(A)$ is closable and its closure $G := \overline{A}$ is the generator of a Feller semigroup if and only if*

(i) *There exist $f_n \in \mathcal{D}(A)$ such that $f_n \to 1$ and $Af_n \to 0$.*

(ii) *A satisfies the positive maximum principle.*

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

(iv) *For some (and hence for all) $\lambda \in (0, \infty)$, there exists a dense subspace $\mathcal{R} \subset \mathcal{C}(E)$ such that for every $f \in \mathcal{R}$, the Laplace equation $(\lambda - G)p = f$ has a solution p.*

Conditions (i)–(iii) are usually easy to verify for a given operator A, but condition (iv) is the "hard" condition since this means that one has to prove existence of solutions to the Laplace equation $(\lambda - G)p = f$ for a dense set of functions f.

If K is a probability kernel on E and $r > 0$, then

$$Gf := r(Kf - f) \qquad \big(f \in \mathcal{C}(E)\big) \tag{4.6}$$

defines a Feller generator that is everywhere defined (i.e., $\mathcal{D}(G) = \mathcal{C}(E)$) and hence, in view of Theorem 4.2, a bounded operator. For generators of this simple form, one can construct the corresponding semigroup by the exponential formula

$$P_t f = e^{tG} f := \sum_{n=0}^{\infty} \frac{1}{n!} (tG)^n f,$$

where the infinite sum converges in $\mathcal{C}(E)$. The corresponding Markov process has a simple description: with rate r, the process jumps from its current position x to a new position chosen according to the probability law $K(x\, \cdot)$.

As soon as Feller processes get more complicated in the sense that "the total rate of all things that can happen" is infinite (as will be the case for interacting particle systems), one needs the more complicated Hille-Yosida theory. To demonstrate the strength of Theorem 4.6, consider $E := [0,1]$ and the linear operator A defined by $\mathcal{D}(A) := \mathcal{C}^2[0,1]$ (the space of twice continuously differentiable functions on $[0,1]$) and

$$Af(x) := x(1-x)\tfrac{\partial^2}{\partial x^2}f(x) \qquad \big(x \in [0,1]\big). \tag{4.7}$$

Then one can show that A satisfies the conditions of Theorem 4.6 and hence $\overline{A}$ generates a Feller semigroup. The corresponding Markov process turns out to have continuous sample paths and is indeed the *Wright-Fisher diffusion* that we met before in formula (3.21).

Some notes on the proofs

In the remainder of this section, we indicate where proofs of the stated theorems can be found. Readers who are more interested in interacting particle systems than in functional analysis may skip from here to the next section.

The fact that there is a one-to-one correspondence between continuous transition probabilities and collections $(P_t)_{t\geq 0}$ of linear operators satisfying the assumptions (i)–(iv) of a Feller semigroup follows from [Kal97, Prop. 17.14].

Theorem 4.1 (including a proof) can be found in [Kal97, Thm 17.15] and [EK86, Thm 4.2.7]. Theorem 4.2 (the closed graph theorem and characterization of continuous linear maps) can be found on many places (including Wikipedia).

Proposition 4.3 summarizes a number of well-known facts. The fact that $u_t := P_t f$ solves the Cauchy equation if $f \in \mathcal{D}(G)$ is proved in [EK86, Prop 1.1.5 (b)], [Kal97, Thm 17.6], and [Lig10, Thm 3.16 (b)]. To see that solutions to the Cauchy equation are unique, we use the following fact.

Lemma 4.7 (Positive maximum principle) *Let A be a linear operator on $\mathcal{C}(E)$ and let $u = (u_t)_{t\geq 0}$ be a solution to the Cauchy equation $\frac{\partial}{\partial t}u_t = Au_t$ $(t \geq 0)$. Assume that A satisfies the positive maximum principle and $u_0 \geq 0$. Then $u_t \geq 0$ for all $t \geq 0$.*

Proof By linearity, we may equivalently show that $u_0 \leq 0$ implies $u_t \leq 0$. Assume that $u_t(x) > 0$ for some $x \in E$. By the compactness of E, the function $(x,t) \mapsto e^{-t}u_t(x)$ must assume its maximum over $E \times [0,t]$ in some point (y,s). Our assumptions imply that $e^{-s}u_s(y) > 0$ and hence $s > 0$. But now, since A satisfies the positive maximum principle,

$$0 \leq \tfrac{\partial}{\partial s}\big(e^{-s}u_s(y)\big) = -e^{-s}u_s(y) + e^{-s}Au_s(y) \leq -e^{-s}u_s(y) < 0,$$

so we arrive at a contradiction. ∎

By linearity, Lemma 4.7 implies that if u, v are two solutions to the same Cauchy equation and $u_0 \leq v_0$, then $u_t \leq v_t$ for all $t \geq 0$. In particular, since by Theorem 4.5, Feller generators satisfy the positive maximum principle, this implies uniqueness of solutions of the Cauchy equation in Proposition 4.3. Again by Theorem 4.5, the domain of a Feller semigroup is a dense subspace of of $\mathcal{C}(E)$, so the final statement of Proposition 4.3 follows from the following simple lemma and the fact that $\|P_t f\|_\infty \leq \|f\|_\infty$.

Lemma 4.8 (Closure of bounded operators) *Let $(\mathcal{V}, \|\cdot)$ be a Banach space and let A be a linear operator on $\mathcal{V}$ such that $\mathcal{D}(A)$ is dense and $\|Af\| \leq C\|f\|$ $(f \in \mathcal{D}(A))$ for some $C < \infty$. Then A is closable, $\mathcal{D}(\overline{A}) = \mathcal{V}$, and $\|\overline{A}f\| \leq C\|f\|$ $(f \in \mathcal{V})$.*

Proof (sketch) Since $\mathcal{D}(A)$ is dense, for each $f \in \mathcal{V}$ we can choose $\mathcal{D}(A) \ni f_n \to f$. Using the fact that A is bounded, it is easy to check that if $(f_n)_{n\geq 0}$ is a Cauchy sequence and $f_n \in \mathcal{D}(A)$ for all n, then $(Af_n)_{n\geq 0}$ is also a Cauchy sequence. By the completeness of $\mathcal{V}$, it follows that the limit $\overline{A}f := \lim_{n\to\infty} Af_n$ exists for all $f \in \mathcal{V}$. To see that this defines $\overline{A}$ unambiguously, assume that $f_n \to f$ and $g_n \to f$ and observe that $\|Af_n - Ag_n\| \leq C\|f_n - g_n\| \to 0$. The fact that $\|\overline{A}f\| \leq C\|f\|$ $(f \in \mathcal{V})$ follows from the continuity of the norm. ∎

Lemma 4.4 follows from [EK86, Prop 1.2.1]. Theorems 4.5 and 4.6 both go under the name of the Hille-Yosida theorem. Often, they are stated in a more general form without condition (i). In this generality, the operator G generates a semigroup of *subprobability kernels* $(P_t)_{t\geq 0}$, i.e., $P_t(x, \cdot\,)$ is a measure with total mass $P_t(x, E) \leq 1$. In this context, a Feller semigroup with $P_t(x, E) = 1$ for all t, x is called *conservative*. It is clear from Proposition 4.3 that condition (i) in Theorems 4.5 and 4.6 is necessary and sufficient for the Feller group to be conservative.

The versions of the Hille-Yosida theorem stated in [EK86, Kal97] are more general than Theorems 4.5 and 4.6 since they allow for the case that E is not compact but only locally compact. This is not really more general, however, since what these books basically do if E is not compact is the following. First, they construct the one-point compactification $\overline{E} = E \cup \{\infty\}$ of E. Next, they extend the transition probabilities to $\overline{E}$ by putting $P_t(\infty, \cdot\,) := \delta_\infty$ for all $t \geq 0$. Having proved that they generate a conservative Feller semigroup on $\overline{E}$ of this form, they then still need to prove that the associated Markov process does not explode in the sense that $\mathbb{P}^x[X_t \in E\ \forall t \geq 0] = 1$. In practical situations (such as when constructing Markov processes with state space $\mathbb{R}^d$) it is usually better to explicitly work with the one-point compactification of

$\mathbb{R}^d$ instead of trying to formulate theorems for locally compact spaces that try hide this compactification in the background.

Theorems 4.5 and 4.6 are special cases of more general theorems (also called Hille-Yosida theorem) for strongly continuous contraction semigroups taking values in a general Banach space. In this context, the positive maximum principle is replaced by the assumption that the operator under consideration is *dissipative*. In this more general setting, Theorems 4.5 and 4.6 correspond to [EK86, Thms 1.2.6 and 1.2.12]. In the more specific set-up of Feller semigroups, versions of Theorem 4.6 can be found in [EK86, Thm 4.2.2] and [Kal97, Thm 17.11]. There is also an account of Hille-Yosida theory for Feller semigroups in [Lig10, Chap 3], but this reference does not mention the positive maximum principle (using a dissipativity assumption instead).

Feller semigroups with bounded generators such as in (4.6) are treated in [EK86, Sect 4.2] and [Kal97, Prop 17.2]. The fact that the operator A in (4.7) satisfies the assumptions of Theorem 4.6 is proved in [EK86, Thm 8.2.8].

4.3 Poisson construction

We briefly recall the set-up introduced in Section 4.1. S is a finite set, called the *local state space*, and Λ is a countable set, called the *lattice*. We equip the product space S^Λ with the product topology, making it into a compact metrizable space. Elements of S^Λ are denoted $x = (x(i))_{i\in\Lambda}$. Given a set $\mathcal{G}$ whose elements are maps $m : S^\Lambda \to S^\Lambda$ and a collection of nonnegative rates $(r_m)_{m\in\mathcal{G}}$, we wish to give sufficient conditions so that there exists a Feller process with state space S^Λ and generator

$$Gf(x) = \sum_{m\in\mathcal{G}} r_m\big\{f\big(m(x)\big) - f\big(x\big)\big\} \qquad (x \in S^\Lambda). \tag{4.8}$$

As explained in the previous section, we cannot expect Gf to be defined for all $f \in \mathcal{C}(S^\Lambda)$, but instead define Gf first for a class of "nice" functions and then find the full generator by taking the closure.

At present, we will not follow this generator approach but instead give a Poisson construction of the processes we are interested in, in the spirit of Proposition 2.5. In the next section, it will then be shown that the process constructed in this way indeed has a generator of the form (4.8).

We will only consider processes whose generator can be represented in terms of *local* maps, i.e., maps that change the local state of finitely many sites only, using also only information about finitely many sites. For any map $m : S^\Lambda \to S^\Lambda$, 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. For $i \in \Lambda$, we write

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

We say that a map $m : S^\Lambda \to S^\Lambda$ is *local* if it satisfies the following three conditions.

(i) $\mathcal{D}(m)$ is finite.

(ii) $\mathcal{R}_i(m)$ is finite for all $i \in \Lambda$.

(iii) For each $i \in \Lambda$, if $x(j) = y(j)$ for all $j \in \mathcal{R}_i(m)$, then $m(x)(i) = m(y)(i)$.

Note that it is possible that $\mathcal{D}(m)$ is nonempty but $\mathcal{R}_i(m) = \emptyset$ for all $i \in \mathcal{D}(m)$. The following exercise shows that condition (iii) is not automatically satisfied. In fact, one can show that a map $m : S^\Lambda \to S^\Lambda$ is continuous w.r.t. the product topology if and only if it satisfies conditions (ii) and (iii) of the definition of a local map; see [SS15b, Lemma 24].

Exercise 4.9 (A discontinuous map) *Define $m : \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$ by*

$$m(x)(0) := \begin{cases} 1 & \text{if } x(i) = 1 \text{ for finitely many } i \in \mathbb{N}, \\ 0 & \text{otherwise,} \end{cases} \tag{4.9}$$

and $m(x)(k) := x(k)$ for $k > 0$. Show that m satisfies conditions (i) and (ii) of the definition of a local map, but not condition (iii).

The following exercise describes yet another way to look at local maps.

Exercise 4.10 *Show that a map $m : S^\Lambda \to S^\Lambda$ is local if and only if there exists a finite $\Delta \subset \Lambda$ and a map $m' : S^\Delta \to S^\Delta$ such that*

$$m(x)(k) = \begin{cases} m'\big((x(i))_{i\in\Delta}\big)(k) & \text{if } k \in \Delta, \\ x(k) & \text{otherwise.} \end{cases}$$

Before we continue, it is good to see a number of examples.

- The voter map $\mathtt{vot}_{ij}$ defined in (1.4) satisfies

$$\mathcal{D}(\mathtt{vot}_{ij}) = \{j\} \quad \text{and} \quad \mathcal{R}_j(\mathtt{vot}_{ij}) = \{i\},$$

  since only the type at j changes, and it suffices to know the type at i to predict the new type of j.

- The branching map $\mathtt{bra}_{ij}$ defined in (1.6) satisfies

$$\mathcal{D}(\mathtt{bra}_{ij}) = \{j\} \quad \text{and} \quad \mathcal{R}_j(\mathtt{bra}_{ij}) = \{i, j\},$$

  since only the type at j changes, but we need to know both the type at i and j to predict the new type of j since $\mathtt{bra}_{ij}(x)(j) = x(i) \vee x(j)$.

- The death map $\mathtt{death}_i$ defined in (1.7) satisfies

$$\mathcal{D}(\mathtt{death}_i) = \{i\} \quad \text{and} \quad \mathcal{R}_i(\mathtt{death}_i) = \emptyset$$

  since only the type at i changes, and the new type at i is 0 regardless of x.

- The coalescing random walk map $\mathtt{rw}_{ij}$ defined in (1.18) satisfies

$$\mathcal{D}(\mathtt{rw}_{ij}) = \{i, j\}, \quad \mathcal{R}_i(\mathtt{rw}_{ij}) = \emptyset, \quad \text{and} \quad \mathcal{R}_j(\mathtt{rw}_{ij}) = \{i, j\},$$

  since the types at both i and j can change, the new type at i is 0 regardless of the previous state, but to calculate $\mathtt{rw}_{ij}(x)(j)$ we need to know both $x(i)$ and $x(j)$.

Exercise 4.11 *Recall the exclusion map* $\mathtt{excl}_{ij}$ *defined in (1.21) and the cooperative branching map* $\mathtt{coop}_{ij}$ *defined in (1.23). For* $m = \mathtt{excl}_{ij}$ *or* $m = \mathtt{coop}_{ij}$*, determine* $\mathcal{D}(m)$*, and determine* $\mathcal{R}_i(m)$ *for all* $i \in \mathcal{D}(m)$*.*

Let $\mathcal{G}$ be a countable set whose elements are local maps $m : S^\Lambda \to S^\Lambda$, let $(r_m)_{m\in\mathcal{G}}$ be nonnegative constants, and (as in Proposition 2.5) let ω be a Poisson point set on $\mathcal{G} \times \mathbb{R}$ with intensity $r_m \mathrm{d}t$. The difficulty is that we will typically have that $\sum_{m\in\mathcal{G}} r_m = \infty$. As a result, $\{t : (m, t) \in \omega\}$ will be a dense subset of $\mathbb{R}$, so it will no longer possible to order the elements of $\omega_{s,t}$ according to their times as we did in (2.7). Nevertheless, since our maps m are local, we can hope that under suitable assumptions on the rates, only finitely many points of $\omega_{0,t}$ are needed to determine the local state $X_t(i)$ of our process at a given lattice point $i \in \Lambda$ and time $t \geq 0$.

For example, for a contact process with a generator as in (1.8), elements of ω are points of the form $(\mathtt{bra}_{ij}, t)$ or $(\mathtt{death}_i, t)$, which indicate that the

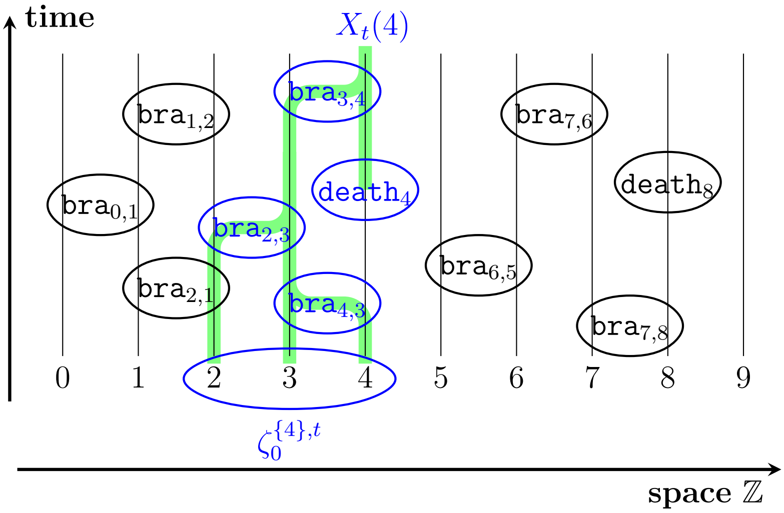


Figure 4.1: Graphical representation of a one-dimensional contact process, with paths of influence and the set $\zeta_0^{\{4\},t}$ of lattice points at time zero whose value is relevant for $X_t(4)$.

corresponding local map should be applied at time t. In Figure 4.1, we have drawn space horizontally and time vertically and vizualized one random realization of ω in such a way that for each element (m,t) of ω we draw a symbol representing the map m at the time t and at the sites that are involved in the map. Such a picture is called a *graphical representation* for an interacting particle system. In practice, various symbols (such as arrows, squares, stars etc.) are used to indicate different maps. Our aim is to find sufficient conditions under which such a graphical representation almost surely yields a well-defined process.

As a first step, we observe that for each $i \in \Lambda$, the set

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

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

$$K_0 := \sup_i \sum_{\substack{m\in\mathcal{G}\\ \mathcal{D}(m)\ni i}} r_m < \infty, \tag{4.10}$$

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.

For any $i,j\in\Lambda$ and $s<u$, by definition, a *path of influence* from (i,s) to (j,u) is a cadlag function $\gamma:(s,u]\to\Lambda$ such that $\gamma_{s-}=i$, $\gamma_u=j$, and

$$\begin{array}{rl}\text{(i)} & \text{if } \gamma_{t-}\neq\gamma_t \text{ for some } t\in(s,u]\text{, then there exists some } m\in\mathcal{G}\\ & \text{such that } (m,t)\in\omega,\ \gamma_t\in\mathcal{D}(m) \text{ and } \gamma_{t-}\in\mathcal{R}_{\gamma_t}(m),\\ \text{(ii)} & \text{for each } (m,t)\in\omega \text{ with } t\in(s,u] \text{ and } \gamma_t\in\mathcal{D}(m),\\ & \text{one has } \gamma_{t-}\in\mathcal{R}_{\gamma_t}(m).\end{array}\tag{4.11}$$

We write $(i,s)\rightsquigarrow(j,u)$ if there is a path of influence from (i,s) to (j,u). Similarly, for any $A\subset\Lambda$, we write $(i,s)\rightsquigarrow A\times\{u\}$ if there is a path of influence from (i,s) to some point (j,u) with $j\in A$. For any finite set $A\subset\Lambda$ and $s<u$, we set

$$\zeta^{A,u}_s:=\big\{i\in\Lambda:(i,s)\rightsquigarrow A\times\{u\}\big\},\tag{4.12}$$

and we let $\zeta^{A,u}_u:=A$. If we start the process at time zero, then $\zeta^{A,t}_0$ will be the set of lattice points whose values at time zero are relevant for the local state of the process in A at time t. See Figure 4.1 for a picture of $\zeta^{A,u}_u$ and the collection of all paths of influence that end in $A\times\{t\}$. The following lemma will be the cornerstone of our Poisson construction of interacting particle systems.

Lemma 4.12 (Exponential bound) *Assume that the rates $(r_m)_{m\in\mathcal{G}}$ satisfy (4.10) and that*

$$K:=\sup_{i\in\Lambda}\sum_{\substack{m\in\mathcal{G}\\ \mathcal{D}(m)\ni i}} r_m\big(|\mathcal{R}_i(m)|-1\big)<\infty.\tag{4.13}$$

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

$$\mathbb{E}\big[|\zeta^{A,u}_s|\big]\leq|A|e^{K(u-s)}\qquad(s\leq u).\tag{4.14}$$

Proof To simplify notation, we fix A and u and write

$$\xi_t:=\zeta^{A,u}_{u-t}\qquad(t\geq 0).$$

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

$$\xi^n_t:=\big\{i\in\Lambda:(i,t)\rightsquigarrow_n A\times\{u\}\big\},$$

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

$$\xi^n_t \uparrow \xi_t \qquad (t\geq 0).$$

Let $\mathcal{G}_n := \{m\in\mathcal{G} : \mathcal{D}(m)\cap\Lambda_n\neq\emptyset\}$. It follows from (4.10) that the total rate $\sum_{m\in\mathcal{G}_n} r_m$ at which maps from $\mathcal{G}_m$ are applied is finite, and only these maps have the potential to change the state of ξ^n_t. The process $(\xi^n_t)_{t\geq 0}$ is left-continuous; let $\xi^n_{t+} := \lim_{s\downarrow t}\xi^n_t$ denote its right-continuous modification. We claim that $(\xi^n_{t+})_{t\geq 0}$ is a Markov process. Indeed, one can have $\xi^n_{t+}\neq\xi^n_t$ only when $(m,u-t)\in\omega$ for some $m\in\mathcal{G}_n$ and at such an instant, if the previous state is $\xi^n_t = A$, then the new state is $\xi^n_{t+} = A^m$, where

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

Proposition 2.5 now implies that the process $(\xi^n_{t+})_{t\geq 0}$ is a Markov process taking values in the (finite) space of all subsets of Λ_n, with generator

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

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

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

It follows that

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

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

$$\mathbb{E}\big[|\xi^n_t|\big] \leq |A|e^{Kt} \qquad (t\geq 0). \tag{4.15}$$

Letting $n \uparrow \infty$ we arrive at (4.14). ∎

Recall that $\omega_{s,t} := \omega \cap (\mathcal{G} \times (s,t])$. 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 4.13 (Finitely many relevant events) *Assume that the rates $(r_m)_{m\in\mathcal{G}}$ satisfy (4.10) and that*

$$K_1 := \sup_{i\in\Lambda} \sum_{\substack{m\in\mathcal{G} \\ \mathcal{D}(m)\ni i}} r_m |\mathcal{R}_i(m)| < \infty. \tag{4.16}$$

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

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

is finite.

Proof Set

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

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 (4.11). Lemma 4.12 does not automatically imply that $|\xi^{A,u}_s| < \infty$ for all $s \leq u$. However, applying the same method of proof to the Markov process $(\xi^{A,u}_{u-t})_{t\geq 0}$, replacing (4.13) by the slightly stronger condition (4.16), we can derive an exponential bound for $\mathbb{E}[|\xi^{A,u}_s|]$, proving that $\xi^{A,u}_s$ is a.s. finite for each finite A and $s \leq u$. Since by (4.10), there are only finitely many events $(m,t) \in \omega_{s,u}$ such that $\mathcal{D}(m) \cap \xi^{A,u}_s \neq \emptyset$, our claim follows. ∎

Remark Conditions (4.10) and (4.16) can be combined in the condition

$$\sup_{i\in\Lambda} \sum_{\substack{m\in\mathcal{G} \\ \mathcal{D}(m)\ni i}} r_m\big(|\mathcal{R}_i(m)| + 1\big) < \infty. \tag{4.17}$$

We are now ready to define a stochastic flow, similar to what we did for finite state spaces in Proposition 2.5, that can be used to construct the interacting particle systems we are interested in. Fix $s \leq u$ and $i \in \Lambda$. By Lemma 4.13, the set

$$\omega^{(i,u)}_s := \big\{(m,t) \in \omega_{s,u} : \mathcal{D}(m) \times \{t\} \rightsquigarrow (i,u)\big\}$$

is a.s. finite. For any finite set $\tilde{\omega} \supset \omega_s^{(i,u)}$, we can order the elements according to the time when they occur:

$$\tilde{\omega} = \{(m_1, t_1), \ldots, (m_n, t_n)\} \quad \text{with} \quad t_1 < \cdots < t_n.$$

Then, setting

$$\mathbf{X}_{s,u}(x)(i) := m_n \circ \cdots \circ m_1(x)(i) \qquad (i \in \Lambda,\ s \leq u)$$

unambiguously defines random maps $\mathbf{X}_{s,u} : S^\Lambda \to S^\Lambda$, where the definition does not depend on the choice of the finite set $\tilde{\omega} \supset \omega_s^{(i,u)}$; indeed, since $\omega_s^{(i,u)}$ contains all Poisson events that are relevant for $\mathbf{X}_{s,u}(x)(i)$, adding more events will not change the outcome.

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

$$P_t(x,\,\cdot\,) := \mathbb{P}\big[\mathbf{X}_{0,t}(x) \in \cdot\,\big] \qquad (x \in S^\Lambda,\ t \geq 0). \tag{4.18}$$

Below is the main result of this chapter.

Theorem 4.14 (Poisson construction of particle systems) *Let $\mathcal{G}$ be a countable set whose elements are local maps $m : S^\Lambda \to S^\Lambda$, let $(r_m)_{m\in\mathcal{G}}$ be nonnegative constants satisfying (4.17), and let ω be a Poisson point set on $\mathcal{G}\times[0,\infty)$ with intensity $r_m\mathrm{d}t$. Then (4.18) 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 ω, then*

$$X_t := \mathbf{X}_{0,t}(X_0) \qquad (t \geq 0) \tag{4.19}$$

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 (4.19) 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 $\mathbf{X}_{0,t}$ 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 $[0, \infty)$.

It is straightforward to check that $(\mathbf{X}_{s,t})_{s\leq t}$ is a stochastic flow with independent increments. The proof that $(X_s)_{0\leq s\leq t}$ is a Markov process with semigroup $(P_t)_{t\geq 0}$ now follows in exactly the same way as in the proof of Proposition 2.5, with (4.3) taking the place of (2.5).

The fact that $P_sP_t = P_{s+t}$ follows from the fact that $(\mathbf{X}_{s,t})_{s\leq t}$ is a stochastic flow. 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 use negative times. (Note that we have defined ω to be a Poisson point process on $\mathcal{G} \times \mathbb{R}$, even though for (4.19) we only need points $(m, t) \in \omega$ with $t > 0$.) Since the law of ω is invariant under translations in the time direction, we have (compare (4.18))

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

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

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

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

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

for each $i \in \Lambda$. By Lemma 4.13, there exists some $\varepsilon > 0$ such that there are no events in $\omega_{-t-\varepsilon,-t+\varepsilon}$ that are relevant for $(i, 0)$, while by Lemma 4.12, $\zeta^{\{i\},0}_{-t}$ 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 $\mathbf{X}_{-t_n,0}(x_n)(i) = \mathbf{X}_{-t,0}(x)(i)$, proving the desired a.s. convergence. ∎

4.4 Generator construction

Although Theorem 4.14 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 spaces of functions by

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

We say that functions in $\mathcal{C}_{\rm sum}$ are of 'summable variation'. The next exercise shows that functions in $\mathcal{C}_{\rm fin}$ depend on finitely many coordinates only.

Exercise 4.15 *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 4.16 (Domain of pregenerator) *Assume that the rates $(r_m)_{m\in\mathcal{G}}$ satisfy (4.10). Then, for each $f \in \mathcal{C}_{\rm sum}(S^\Lambda)$,*

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

where K_0 is the constant from (4.10). In particular, for each $f \in \mathcal{C}_{\rm sum}(S^\Lambda)$, the right-hand side of (4.8) is absolutely summable and Gf is well-defined.

Proof This follows by writing

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

∎

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

Theorem 4.17 (Generator construction of particle systems) *Assume that the rates $(r_m)_{m\in\mathcal{G}}$ satisfy (4.17), let $(P_t)_{t\geq 0}$ be the Feller semigroup defined in (4.18) and let G be the linear operator with domain $\mathcal{D}(G) := \mathcal{C}_{\rm sum}$ defined by (4.8). Then G is 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 4.17 says that $\mathcal{C}_{\rm fin}$ is a core for $\overline{G}$.

To prepare for the proof of Theorem 4.17 we need a few lemmas.

Lemma 4.18 (Generator on local functions) *Under the asumptions of Theorem 4.17, 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 $\omega_{0,t}(A) := \{(m,s) \in \omega_{0,t} : \mathcal{D}(m) \cap A \neq \emptyset\}$. Then $|\omega_{0,t}(A)|$ is Poisson distributed with mean $t\sum_{m\in\mathcal{G},\ \mathcal{D}(m)\cap A\neq\emptyset} r_m$, which is finite by (4.10). Write

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

Since

$$\big|\mathbb{E}[f\big(\mathbf{X}_{0,t}(x)\big)\,\big|\, |\omega_{0,t}(A)| \geq 2\big]\big| \leq \|f\|,$$

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

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

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

Lemma 4.19 (Approximation by local functions) *Assume that the rates* $(r_m)_{m\in\mathcal{G}}$ *satisfy (4.10). 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

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

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

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

Moreover, we observe that

$$\begin{aligned}
&|Gf_n(x) - Gf(x)| \\
&= \Big|\sum_{m\in\mathcal{G}} r_m\big(f_n(m(x)) - f_n(x)\big) - \sum_{m\in\mathcal{G}} r_m\big(f(m(x)) - f(x)\big)\Big| \\
&\le \sum_{m\in\mathcal{G}} r_m\big|f(m(x)_n) - f(x_n) - f(m(x)) + f(x)\big|.
\end{aligned} \tag{4.20}$$

On the one hand, we have

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

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

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

Let $\omega \subset \Lambda$ be finite. Inserting either of our two estimates into (4.20), depending on whether $\mathcal{D}(m) \cap \omega \neq \emptyset$ or not, we find that

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

It follows that

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

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

Lemma 4.20 (Functions of summable variation) *Under the asumptions of Theorem 4.17, one has*

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

where K is the constant from (4.13). In particular, for each $t \ge 0$, 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

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

By Lemma 4.12, it follows that

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

∎

Proof of Theorem 4.17 Let H be the full generator of $(P_t)_{t\ge 0}$ and let $\mathcal{D}(H)$ denote it domain. Then Lemma 4.18 shows that $\mathcal{C}_{\rm fin} \subset \mathcal{D}(H)$ and $Gf = Hf$ for all $f \in \mathcal{C}_{\rm fin}$. By Lemma 4.19, 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 the Hille-Yosida Theorem 4.6. 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 (4.8) is nonpositive, hence $Gf(x) \le 0$. The fact that $\mathcal{C}_{\rm sum}(S^\Lambda)$ is dense follows from Exercise 4.15 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 (4.13), and for each $f \in \mathcal{C}_{\rm fin}(S^\Lambda)$, there exists a $p_r \in \mathcal{C}_{\rm sum}(S^\Lambda)$ that solves the Laplace equation $(r - G)p_r = f$. In the light of Lemma 4.4 a natural candidate for such a function is

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

and we will show that this p_r indeed satisfies $p_r \in \mathcal{C}_{\rm sum}(S^\Lambda)$ and $(r-G)p_r = f$. It follows from Theorem 4.6 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 4.20,

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

which proves that $p_r \in \mathcal{C}_{\rm sum}$. This completes the proof that $\overline{G} = H$. By Lemma 4.19, 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, that is sometimes useful.

Lemma 4.21 (Differentiation of semigroup) *Assume that the rates $(r_m)_{m\in\mathcal{G}}$ satisfy (4.17), let $(P_t)_{t\geq 0}$ be the Feller semigroup defined in (4.18) and let G be the linear operator with domain $\mathcal{D}(G) := \mathcal{C}_{\rm sum}(S^\Lambda)$ defined by (4.8). 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 $\frac{\partial}{\partial t}P_t f = GP_t f$ for each $t \geq 0$.*

Proof This is a direct consequence of Proposition 4.3, Lemma 4.20, and Theorem 4.17. 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. ∎

Some bibliographical remarks

Theorem 4.17 is similar to Liggett's [Lig85, Theorem I.3.9], but there are also some differences. Liggett does not write his generators in terms of local maps, but in terms of local transition kernels, that, using information about the total configuration $(x(i))_{i\in\Lambda}$ of the system, change the local configuration $(x(i))_{i\in\Delta}$, with Δ a finite subset of Λ, in a random way. This way of writing the generator is more general and sometimes (for example, for stochastic Ising models) more natural than our approach using local maps. It is worth noting that Liggett's construction, like ours, depends on a clever way of writing the generator that is in general not unique.

Unlike our Theorem 4.14, Liggett does not give an explicit construction of his interacting particle systems using Poisson point sets, but instead gives a direct proof that the closure of G generates a Feller semigroup $(P_t)_{t\geq 0}$, and then invokes the abstract result Theorem 4.1 about Feller processes to prove the existence of a corresponding Markov process with cadlag sample paths. Later in his book, he does use explicit Poisson constructions for some systems, such as the contact process. He does not actually prove that these Poisson constructions yield the same process as the generator construction, but apparently finds this self-evident. (Equivalence of the two constructions follows from our Theorem 4.17 but alternatively can also be proved by approximation with finite systems, using approximation results such as [Lig85, Cor. I.3.14].)

Liggett's [Lig85, Theorem I.3.9] allows for the case that the local state space S is a (not necessarily finite) compact metrizable space. This is occasionally convenient. For example, this allows one to construct voter models with infinitely many types, where at time zero, the types $(X_0(i))_{i\in\Lambda}$ are i.i.d. and uniformly distributed on $S = [0, 1]$. For simplicity, we have restricted ourselves to finite local state spaces.

4.5 Ergodicity

Luckily, our efforts in the previous chapter are not wasted on knowing only that the systems we are interested in exist, but actually allow us to prove something interesting about these systems as well.

If X is a Markov process with state space E and transition probabilities $(P_t)_{t\geq 0}$, then by definition, an *invariant law* of X is a probability measure ν on E such that

$$\nu P_t = \nu \qquad (t \geq 0).$$

This says that if we start the process in the initial law $\mathbb{P}[X_0 \in \cdot\,] = \nu$, then $\mathbb{P}[X_t \in \cdot\,] = \nu$ for all $t \geq 0$. As a consequence, one can construct a *stationary*[4] process $(X_t)_{t\in\mathbb{R}}$ such that (compare (4.2))

$$\mathbb{P}\big[X_u \in \cdot\,\big|\,(X_s)_{-\infty<s\leq t}\big] = P_{u-t}(X_t,\,\cdot\,) \quad \text{a.s.} \qquad (t \leq u), \tag{4.21}$$

and $\mathbb{P}[X_t \in \cdot\,] = \nu$ for all $t \in \mathbb{R}$. Conversely, the existence of such a stationary Markov process implies that the law at any time $\nu := \mathbb{P}[X_t \in \cdot\,]$ must be an invariant law. For this reason, invariant laws are sometimes also called *stationary laws.*

[4] Recall that a process $(X_t)_{t\in\mathbb{R}}$ is stationary if for each $s \in \mathbb{R}$, it is equal in distribution to $(X'_t)_{t\in\mathbb{R}}$ defined as $X'_t := X_{t-s}$ $(t \in \mathbb{R})$.

Theorem 4.22 (Ergodicity) *Let X be an interacting particle system with state space of the form S^Λ and generator G of the form (4.8), and assume that the rates $(r_m)_{m\in\mathcal{G}}$ satisfy (4.17).*

(a) *Assume that the constant K from (4.13) satisfies $K < 0$. Then the process ζ defined in (4.12) satisfies*

$$\lim_{s\to-\infty} \zeta_s^{\{i\},u} = \emptyset \quad \text{a.s.} \quad (i \in \Lambda,\ u \in \mathbb{R}). \tag{4.22}$$

(b) *Assume that the process ζ defined in (4.12) satisfies (4.22). Then the interacting particle system X has a unique invariant law ν, and*

$$\mathbb{P}^x\big[X_t \in \cdot\,\big] \underset{t\to\infty}{\Longrightarrow} \nu \qquad (x \in S^\Lambda). \tag{4.23}$$

Proof Part (a) is immediate from Lemma 4.12. If (4.22) holds, then the a.s. limit

$$X_t(i) := \lim_{s\to-\infty} \mathbf{X}_{s,t}(x)(i) \qquad (i \in \Lambda,\ t \in \mathbb{R}) \tag{4.24}$$

does not depend on the choice of a point $x \in S^\Lambda$, since the set $\zeta_s^{\{i\},t}$ of lattice points whose value at time s is relevant for $\mathbf{X}_{s,t}(x)(i)$ is empty for s suffiently small. As a result, (4.24) unambiguously defines a stationary process that is also Markov with respect to the transition probabilities $(P_t)_{t\geq 0}$ in the sense of (4.21). It follows that

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

does not depend on $t \in \mathbb{R}$ and defines an invariant law ν. Since

$$\mathbb{P}^x\big[X_t \in \cdot\,\big] = \mathbb{P}\big[\mathbf{X}_{-t,0}(x)(i) \in \cdot\,\big]$$

and since by (4.24), we have

$$\mathbf{X}_{-t,0}(x) \underset{t\to\infty}{\longrightarrow} X_0 \quad \text{a.s.}$$

with respect to the topology of pointwise convergence, we conclude that (4.23) holds. ∎

We note that (4.23) says that if we start the process in an arbitrary initial state x, then the law at time t converges weakly[5] as $t \to \infty$ to the invariant law ν. This property is often described by saying that the interacting particle system is *ergodic*. Indeed, this implies that the corresponding stationary

[5]Here weak convergence is of course w.r.t. our topology on S^Λ, i.e., w.r.t. the product topology.

process $(X_t)_{t\in\mathbb{R}}$ is ergodic in the usual sense of that word, i.e., has a trivial tail-σ-field. The converse conclusion cannot be drawn, however, so the traditional way of describing (4.23) as "ergodicity" is a bit of a bad habit.

We have split Theorem 4.22 into a part (a) and (b) since the condition (4.22) is sometimes satisfied even when the constant K from (4.13) is positive. Indeed, we will later see that for the contact process, the condition (4.22) is sharp but the condition $K < 0$ is not.

Theorem 4.22 is similar, but not identical to [Lig85, Thm I.4.1]. For Theorem 4.22 (a) and (b) to be applicable, one needs to be able to express the generator in terms of local maps such that the constant K from (4.13) is negative. For [Lig85, Thm I.4.1], one needs to express the generator in a convenient way in terms of local transition kernels. For certain problems, the latter approach is more natural and [Lig85, Thm I.4.1] yields sharper estimates for the regime where ergodicity holds.

4.6 Application to the Ising model

The Ising model with Glauber dynamics has been introduced in Section 1.4. So far, we have not show how to represent the generator of this interacting particle system in terms of local maps. In the present section, we will fill this gap. As an application of the theory developed so far, we will then show that the Ising model with Glauber dynamics is well-defined for all values of its parameter, and ergodic for β sufficiently small. Our construction will also prepare for the next chapter, where we discuss monotone interacting particle systems, by showing that the Ising model with Glauber dynamics can be represented in monotone maps.

We recall from Section 1.4 that the Ising model with Glauber dynamics on a graph (Λ, E) is the interacting particle system with state space $\{-1,+1\}^\Lambda$ and dynamics such that

$$\text{site } i \text{ flips to the value } \sigma \text{ with rate} \quad \frac{e^{\beta N_{x,i}(\sigma)}}{e^{\beta N_{x,i}(+1)} + e^{\beta N_{x,i}(-1)}},$$

where

$$N_{x,i}(\sigma) := \sum_{j\in\mathcal{N}_i} 1_{\{x(j) = \sigma\}} \qquad \big(\sigma \in \{-1,+1\}\big)$$

denotes the number of neighbors of i that have the spin value σ. Let

$$M_{x,i} := N_{x,i}(+) - N_{x,i}(-) = \sum_{j\in\mathcal{N}_i} x(j)$$

denote the *local magnetization* in the neighborhood $\mathcal{N}_i$ of i. Since $N_{x,i}(+) - N_{x,i}(-) = |\mathcal{N}_i|$, we can rewrite the rate of flipping to the spin value $+1$ as

$$\begin{aligned}\frac{e^{\beta N_{x,i}(+1)}}{e^{\beta N_{x,i}(+1)} + e^{\beta N_{x,i}(-1)}} &= \frac{e^{\beta(|\mathcal{N}_i|+M_{x,i})/2}}{e^{\beta(|\mathcal{N}_i|+M_{x,i})/2} + e^{\beta(|\mathcal{N}_i|-M_{x,i})/2}}\\ &= \frac{e^{\frac{1}{2}\beta M_{x,i}}}{e^{\frac{1}{2}\beta M_{x,i}} + e^{-\frac{1}{2}\beta M_{x,i}}} = 1 + \frac{e^{\frac{1}{2}\beta M_{x,i}} - e^{-\frac{1}{2}\beta M_{x,i}}}{e^{\frac{1}{2}\beta M_{x,i}} + e^{-\frac{1}{2}\beta M_{x,i}}}\\ &= 1 + \tanh(\tfrac{1}{2}\beta M_{x,i}).\end{aligned}$$

Similarly, the rate of flipping to -1 is $1 - \tanh(\frac{1}{2}\beta M_{x,i})$. Define maps $m_i^\pm : \{-1,+1\}^\Lambda \to \{-1,+1\}^\Lambda$ by

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

Then we can write the generator of the Ising model in the form

$$\begin{aligned}G_{\text{Ising}} f(x) := &\big(1 + \tanh(\tfrac{1}{2}\beta M_{x,i})\big)\big\{f\big(m_i^+(x)\big) - f\big((x)\big\}\\ &+\big(1 - \tanh(\tfrac{1}{2}\beta M_{x,i})\big)\big\{f\big(m_i^-(x)\big) - f\big((x)\big\}.\end{aligned} \tag{4.25}$$

This is not yet a representation in terms of local maps of the form (4.8), since the rates

$$r_i^\pm(x) := 1 \pm \tanh(\tfrac{1}{2}\beta M_{x,i}) \tag{4.26}$$

at which the local maps $m_i^\pm$ are applied depend on the actual state x of the model.[6]

To fix this, we proceed as follows. For (mainly notational) simplicity, let us assume that the size of the neighborhood

$$N := |\mathcal{N}_i| \qquad (i \in \Lambda)$$

does not depend on $i \in \Lambda$. For $L = -N, -N+2, \ldots, N$, we define local maps $m_{i,L}^\pm$ by

$$\begin{aligned} m_{i,L}^+(x)(j) &:= \begin{cases} +1 & \text{if } j = i \text{ and } M_{x,i} \geq L,\\ x(j) & \text{otherwise.}\end{cases}\\ m_{i,L}^-(x)(j) &:= \begin{cases} -1 & \text{if } j = i \text{ and } M_{x,i} \leq L,\\ x(j) & \text{otherwise.}\end{cases}\end{aligned} \tag{4.27}$$

[6]On the other hand, (4.25) is an expression for G_{Ising} of the form considered in Liggett's [Lig85, Theorem I.3.9].

We observe that the hyperbolic tangent is an increasing function. As a result, the rates $r^\pm_{i,L}$ defined by

$$r^+_{i,L} := \begin{cases} 1 + \tanh(-\frac{1}{2}\beta N) & \text{if } L = -N, \\ \tanh(\frac{1}{2}\beta L) - \tanh(\frac{1}{2}\beta(L-2)) & \text{otherwise.} \end{cases}$$

$$r^-_{i,L} := \begin{cases} 1 - \tanh(\frac{1}{2}\beta N) & \text{if } L = N, \\ \tanh(\frac{1}{2}\beta L) - \tanh(\frac{1}{2}\beta(L+2)) & \text{otherwise} \end{cases}$$

are all positive. We claim that we can rewrite the Ising generator (4.25) in the form

$$G_{\rm Ising} f(x) = \sum_{i\in\Lambda} \sum_{\sigma\in\{-,+\}} \sum_{L\in\{-N,-N+2,\ldots,N\}} r^\sigma_{i,L} \big\{ f\big(m^\sigma_{i,L}(x)\big) - f\big((x\big)\big\}. \quad (4.28)$$

To see this, note that if the present state is x, then the total rate at which the site i flips to the spin value $+1$ is

$$\sum_{L:\, L\leq M_{x,i}} r^+_{i,L} = 1 + \tanh(\tfrac{1}{2}\beta M_{x,i}) = r^+_i(x),$$

where $r^+_i(x)$ is defined in (4.26). The argument for flips to -1 is the same.

Theorem 4.23 (Existence and ergodicity of the Ising model) *Consider an Ising model with Glauber dynamics on a countable graph Λ in which each lattice point i has exactly $|\mathcal{N}_i| = N$ neighbors, i.e., the Markov process X with state space $\{-1,+1\}^\Lambda$ and generator $G_{\rm Ising}$ given by (4.25) or equivalently (4.28). Then, for each $\beta \geq 0$, the closure of $G_{\rm Ising}$ generates a Feller semigroup. Moreover, for each β such that*

$$e^{\beta N} < \frac{N}{N-1}, \quad (4.29)$$

the Markov process with generator $\overline{G}_{\rm Ising}$ has a unique invariant law ν, and the process started in an arbitrary initial state x satisfies

$$\mathbb{P}^x\big[X_t \in \cdot\,\big] \underset{t\to\infty}{\Longrightarrow} \nu \qquad \big(x \in \{-1,+1\}^\Lambda\big).$$

Proof We use the representation (4.28). We observe that

$$\mathcal{D}(m^\pm_{i,L}) = \{i\}$$

is the set of lattice points whose spin value can be changed by the map $m^{\pm}_{i,L}$. The set of lattice points that are $m^{\pm}_{i,L}$-relevant for i is given by

$$\mathcal{R}_i(m^{\sigma}_{i,L}) = \begin{cases} \emptyset & \text{if } \sigma = +,\ L = -N \quad \text{or} \quad \sigma = -,\ L = N, \\ \mathcal{N}_i & \text{otherwise.} \end{cases}$$

Here we have used that $-N \leq M_{x,i} \leq N$ holds always, so $m^{+}_{-N}(x)(i) = +1$ and $m^{-}_{N}(x)(i) = -1$ regardless of what x is. On the other hand, in all other cases, the value of each lattice point $j \in \mathcal{N}_i$ can potentially make a difference for the outcome $m^{\pm}_{i,L}(x)(i)$.

By Theorem 4.17, to conclude that the closure of G_{Ising} generates a Feller semigroup, it suffices to check (4.17), which in our case says that

$$\sup_{i\in\Lambda} \sum_{\sigma\in\{-,+\}} \sum_{L\in\{-N,-N+2,\ldots,N\}} r^{\sigma}_{i,L}\big(|\mathcal{R}_i(m^{\sigma}_{i,L})| + 1\big)$$

should be finite. Since $\sum_L r^{\sigma}_{i,L} \leq 1 + \tanh(\frac{1}{2}\beta N) \leq 2$ and $|\mathcal{R}_i(m^{\sigma}_{i,L})| \leq |\mathcal{N}_i| = N$, this expression is $\leq 4(N+1) < \infty$ regardless of the value of β.

To prove ergodicity for β small enough, we apply Theorem 4.22. We calculate the constant K from (4.13). By the symmetry between minus and plus spins,

$$\begin{aligned} K &= 2 \sum_{L\in\{-N,-N+2,\ldots,N\}} r^{+}_{i,L}\big(|\mathcal{R}_i(m^{\sigma}_{i,L})| - 1\big) \\ &= -2r^{+}_{i,-N} + 2 \sum_{L\in\{-N+2,\ldots,N\}} r^{+}_{i,L}\big(N-1\big) \\ &= -2\big(1 + \tanh(-\tfrac{1}{2}\beta N)\big) + 2\big(\tanh(\tfrac{1}{2}\beta N) - \tanh(-\tfrac{1}{2}\beta N)\big)\big(N-1\big), \end{aligned}$$

which is negative if and only if

$$\begin{aligned} & 1 + \tanh(-\tfrac{1}{2}\beta N) > \big(\tanh(\tfrac{1}{2}\beta N) - \tanh(-\tfrac{1}{2}\beta N)\big)\big(N-1\big) \\ \Leftrightarrow\quad & 1 + \frac{e^{-\frac{1}{2}\beta N} - e^{\frac{1}{2}\beta N}}{e^{\frac{1}{2}\beta N} + e^{-\frac{1}{2}\beta N}} > \Big(\frac{e^{\frac{1}{2}\beta N} - e^{-\frac{1}{2}\beta N}}{e^{\frac{1}{2}\beta N} + e^{-\frac{1}{2}\beta N}} - \frac{e^{-\frac{1}{2}\beta N} - e^{\frac{1}{2}\beta N}}{e^{\frac{1}{2}\beta N} + e^{-\frac{1}{2}\beta N}}\Big)\big(N-1\big) \\ \Leftrightarrow\quad & 2e^{-\frac{1}{2}\beta N} > \big(e^{\frac{1}{2}\beta N} - e^{-\frac{1}{2}\beta N}\big)\big(N-1\big) \\ \Leftrightarrow\quad & \frac{e^{-\frac{1}{2}\beta N}}{e^{\frac{1}{2}\beta N} - e^{-\frac{1}{2}\beta N}} > N-1 \quad \Leftrightarrow \quad \frac{1}{e^{\beta N} - 1} > N-1 \\ \Leftrightarrow\quad & e^{\beta N} - 1 < \frac{1}{N-1} \quad \Leftrightarrow \quad e^{\beta N} < \frac{N}{N-1}, \end{aligned}$$

which is condition (4.29). ∎

4.7 Further results

In the present section we collect a number of technical results of a general nature that will be needed in later chapters. On a first reading, readers are adviced to skip the present section and refer back to specific results when the need arises. The only result of the present section that is perhaps of some intrinsic value is Theorem 4.28 which together with Corollary 4.29 below implies that the transition probabilities of interacting particle systems on infinite lattices can be approximated by those on finite lattices, something that we have been using implicitly when doing simulations.

Let E be a compact metrizable space. By definition, a collection of functions $\mathcal{H} \subset \mathcal{C}(E)$ is *distribution determining* if

$$\mu f = \nu f \ \forall f \in \mathcal{H} \quad \text{implies} \quad \mu = \nu.$$

We say that $\mathcal{H}$ *separates points* if for all $x, y \in E$ such that $x \neq y$, there exists an $f \in \mathcal{H}$ such that $f(x) \neq f(y)$. We say that $\mathcal{H}$ is *closed under products* if $f, g \in \mathcal{H}$ implies $fg \in \mathcal{H}$.

Lemma 4.24 (Application of Stone-Weierstrass) *Let E be a compact metrizable space. Assume that $\mathcal{H} \subset \mathcal{C}(E)$ separates points and is closed under products. Then $\mathcal{H}$ is distribution determining.*

Proof If $\mu f = \nu f$ for all $f \in \mathcal{H}$, then we can add the constant function 1 to $\mathcal{H}$ and retain this property. In a next step, we can add all linear combinations of functions in $\mathcal{H}$ to the set $\mathcal{H}$; by the linearity of the integral, it will then still be true that $\mu f = \nu f$ for all $f \in \mathcal{H}$. But now $\mathcal{H}$ is an algebra that separates points and vanishes nowhere, so by the Stone-Weierstrass theorem, $\mathcal{H}$ is dense in $\mathcal{C}(E)$. If $f_n \in \mathcal{H}$, $f \in \mathcal{C}(E)$, and $\|f_n - f\|_\infty \to 0$, then $\mu f_n \to \mu f$ and likewise for ν, so we conclude that $\mu f = \nu f$ for all $f \in \mathcal{C}(E)$. If $A \subset E$ is a closed set, then the function $f(x) := d(x, A)$ is continuous, where d is a metric generating the topology on E and $d(x, A) := \inf_{y \in A} d(x, y)$ denotes the distance of x to A. Now the functions $f_n := 1 \wedge nf$ are also continuous and $f_n \uparrow 1_{A^c}$, so by the continuity of the integral with respect to increasing sequences we see that $\mu(O) = \nu(O)$ for every open set $O \subset E$. Since the open sets are closed under intersections, it follows that $\mu(A) = \nu(A)$ for every element A of the σ-algebra generated by the open sets, i.e., the Borel-σ-field $\mathcal{B}(E)$. ∎

Lemma 4.25 (Weak convergence) *Let E be a compact metrizable space. Assume that $\mu_n \in \mathcal{M}_1(E)$ have the property that $\lim_{n\to\infty} \mu_n f$ exists for all $f \in \mathcal{H}$, where $\mathcal{H} \subset \mathcal{C}(E)$ is distribution determining. Then there exists a $\mu \in \mathcal{M}_1(E)$ such that $\mu_n \Rightarrow \mu$.*

Proof By Prohorov's theorem, the space $\mathcal{M}_1(E)$, equipped with the topology of weak convergence, is compact. Therefore, to prove the statement, it suffices to show that the sequence μ_n has not more than one cluster point, i.e., it suffices to show that if μ, μ' are subsequential limits, then $\mu' = \mu$. Clearly, μ, μ' must satisfy $\mu' f = \mu f$ for all $f \in \mathcal{H}$, so the claim follows from the assumption that $\mathcal{H}$ is distribution determining. ∎

Lemma 4.26 (Continuous probability kernels) *Let E be a compact metrizable space and let K be a continuous probability kernel on E. Then, for any $\mu_n, \mu \in \mathcal{M}_1(E)$ and $f_n, f \in \mathcal{C}(E)$,*

$$\begin{aligned} &\mu_n \underset{n\to\infty}{\Longrightarrow} \mu \quad \textit{implies} \quad \mu_n K \underset{n\to\infty}{\Longrightarrow} \mu K \\ \textit{and} \quad &\|f_n - f\|_\infty \underset{n\to\infty}{\longrightarrow} 0 \quad \textit{implies} \quad \|K f_n - K f\|_\infty \underset{n\to\infty}{\longrightarrow} 0. \end{aligned}$$

Proof Since K is a continuous probability kernel, its associated operator maps the space $\mathcal{C}(E)$ into itself, so $\mu_n \Rightarrow \mu$ implies that $\mu_n(Kf) \Rightarrow \mu(Kf)$ for all $f \in \mathcal{C}(E)$, or equivalently $(\mu_n K)f \Rightarrow (\mu K)f$ for all $f \in \mathcal{C}(E)$, i.e., the measures $\mu_n K$ converge weakly to μ.

The second statement follows from the linearity and monotonicity of K and the fact that $K1 = 1$, which together imply that $\|K f_n - K f\|_\infty \leq \|f_n - f\|_\infty$. ∎

Lemma 4.27 (Long-time limits) *Let E be a compact metrizable space and let $(P_t)_{t\geq 0}$ be the transition probabilities of a Feller process in E. Let $\mu \in \mathcal{M}_1(E)$ and assume that*

$$\mu P_t \underset{t\to\infty}{\Longrightarrow} \nu$$

for some $\nu \in \mathcal{M}_1(E)$. Then ν is an invariant law of the Feller process with transition probabilities $(P_t)_{t\geq 0}$.

Proof Using Lemma 4.26, this follows by writing

$$\nu P_t = (\lim_{s\to\infty} \mu P_s) P_t = \lim_{s\to\infty} \mu P_s P_t = \lim_{s\to\infty} \mu P_{s+t} = \nu.$$

∎

The following theorem follows from [Kal97, Thm 17.25], where it is moreover shown that the condition (4.30) implies convergence in distribution of the associated Feller processes, viewed as random variables taking values in the space $\mathcal{D}_E[0,\infty)$ of cadlag paths with values in E. Note that in (4.30) below, $\to$ (of course) means convergence in the topology we have defined on $\mathcal{C}(E)$, i.e., convergence w.r.t. the supremumnorm.

Theorem 4.28 (Limits of semigroups) *Let E be a compact metrizable space and let G_n, G be generators of Feller processes in E. Assume that there exists a linear operator on $\mathcal{C}(E)$ such that $\overline{A} = G$ and*

$$\forall f \in \mathcal{D}(A)\ \exists f_n \in \mathcal{D}(G_n) \text{ such that } f_n \to f \quad \text{and} \quad G_n f_n \to Af. \tag{4.30}$$

Then the Feller semigroups $(P_t^n)_{t\geq 0}$ and $(P_t)_{t\geq 0}$ with generators G_n and G, respectively, satisfy

$$\sup_{t\in[0,T]} \|P_t^n f - P_t f\|_\infty \underset{n\to\infty}{\longrightarrow} 0 \quad \big(f \in \mathcal{C}(E),\ T < \infty\big).$$

Moreover, if $\mu_n, \mu \in \mathcal{M}_1(E)$, then

$$\mu_n \underset{n\to\infty}{\Longrightarrow} \mu \quad \text{implies} \quad \mu_n P_t^n \underset{n\to\infty}{\Longrightarrow} \mu P_t \qquad (t \geq 0).$$

We note that in the case of interacting particle systems, Theorem 4.17 implies the following.

Corollary 4.29 (Convergence of particle systems) *Let S be a finite set and let Λ be countable. Let G_n, G be generators of interacting particle systems in S^Λ and assume that G_n, G can be written in the form (4.8) with rates satisfying (4.17). Assume moreover that*

$$\|G_n f - Gf\|_\infty \underset{n\to\infty}{\longrightarrow} 0 \qquad \big(f \in \mathcal{C}_{\rm fin}(S^\Lambda)\big).$$

Then the generators G_n, G satisfy (4.30).

Theorem 4.28 has the following useful consequence.

Proposition 4.30 (Limits of invariant laws) *Let E be a compact metrizable space and let G_n, G be generators of Feller processes in E satisfying (4.30). Let $\nu_n, \nu \in \mathcal{M}_1(E)$ and assume that for each n, the measure ν_n is an invariant law of the Feller process with generator G_n. Then $\nu_n \Rightarrow \nu$ implies that ν is an invariant law of the Feller process with generator G.*

Proof Using Theorem 4.28, this follows simply by observing that

$$\nu P_t = \lim_{n\to\infty} \nu_n P_t^n = \lim_{n\to\infty} \nu_n = \nu$$

for each $t \geq 0$. ∎

Chapter 5

Monotonicity

5.1 The stochastic order

We recall that if S and T are partially ordered sets, then a function $f : S \to T$ is called *monotone* iff $x \leq y$ implies $f(x) \leq f(y)$. In particular, this definition also applies to real-valued functions (where we equip $\mathbb{R}$ with the well-known order). If the local state space S of an interacting particle system is partially ordered, then we equip the product space with the *product order*

$$x \leq y \quad \text{iff} \quad x(i) \leq y(i)\ \forall i \in \Lambda.$$

Many well-known interacting particle systems use the local state space $S = \{0,1\}$, which is of course equipped with a natural order $0 \leq 1$. Often, it is useful to prove comparison results, that say that two interacting particle systems X and Y can be coupled in such a way that $X_t \leq Y_t$ for all $t \geq 0$. Here X and Y may be different systems, started in the same initial state, or also two copies of the same interacting particle system, started in initial states such that $X_0 \leq Y_0$.

The following theorem gives necessary and sufficient conditions for it to be possible to couple two random variables X and Y such that $X \leq Y$. A *coupling* of two random variables X and Y, in the most general sense of the word, is *a way to construct X and Y together on one underlying probability space* $(\Omega, \mathcal{F}, \mathbb{P})$. More precisely, if X and Y are random variables defined on different underlying probability spaces, then a coupling of X and Y is a pair of random variables (X', Y') defined on one underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$, such that X' is equally distributed with X and Y' is equally distributed with Y. Equivalently, since the laws of X and Y are all we really care about, we may say that a *coupling* of two probability laws μ, ν defined on measurable spaces $(E, \mathcal{E})$ and $(F, \mathcal{F})$, respectively, is a probability measure

ρ on the product space $(E \times F, \mathcal{E} \otimes \mathcal{F})$ such that the first marginal of ρ is μ and its second marginal is ν.

Theorem 5.1 (Stochastic order) *Let S be a finite partially ordered set, let Λ be a countable set, and let μ, ν be probability laws on S^Λ. Then the following statements are equivalent:*

(i) $\int \mu(\mathrm{d}x) f(x) \leq \int \nu(\mathrm{d}x) f(x)$ $\forall$ *monotone* $f \in \mathcal{C}(S^\Lambda)$,

(ii) $\int \mu(\mathrm{d}x) f(x) \leq \int \nu(\mathrm{d}x) f(x)$ $\forall$ *monotone* $f \in B(S^\Lambda)$,

(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 implication (iii)$\Rightarrow$(ii) is easy: if X and Y are coupled such that $X \leq Y$ and f is monotone, then

$$\mathbb{E}\big[f(Y)\big] - \mathbb{E}\big[f(X)\big] = \mathbb{E}\big[f(Y) - f(X)\big] \geq 0,$$

since $f(Y) - f(X) \geq 0$ a.s. The implication (ii)$\Rightarrow$(i) is trivial.

For the nontrivial implication (i)$\Rightarrow$(iii) we refer to [Lig85, Theorem II.2.4]. For finite spaces, a nice intuitive proof based on the max flow min cut theorem can be found in [Pre74]. ∎

If two probability laws μ, ν satisfy the equivalent conditions of Theorem 5.1, then we say that μ and ν are *stochastically ordered* and we write[1] $\mu \leq \nu$. Clearly $\mu \leq \nu \leq \rho$ implies $\mu \leq \rho$. In light of this, the next lemma shows that the stochastic order is a bona fide partial order on $\mathcal{M}_1(S^\Lambda)$.

Lemma 5.2 (Monotone functions are distribution determining) *Let S be a finite partially ordered set and let Λ be countable. Then the set $\{f \in \mathcal{C}(S^\Lambda) : f$ is monotone $\}$ is distribution determining. In particular, $\mu \leq \nu$ and $\mu \geq \nu$ imply $\mu = \nu$.*

Proof Since the finite-dimensional distributions uniquely determine a probability measure on S^Λ, it suffices to prove the statement for finite Λ. In view of this, it suffices to show that if S is a finite partially ordered set, then the space of all monotone functions $f : S \to \mathbb{R}$ is distribution determining.

By definition, an *increasing subset* of S is a set $A \subset S$ such that $A \ni x \leq y$ implies $y \in A$. If A is increasing, then its indicator function 1_A is monotone,

[1] 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 $\mu = \nu$.

so it suffices to show that $\{1_A : A \text{ increasing}\}$ is distribution determining. By Lemma 4.24, it suffices to show that this class separates points and is closed under products.

If $x \neq y$, then either $x \not\in \{z : z \geq y\}$ or $y \not\in \{z : z \geq x\}$, so $\{1_A : A \text{ increasing}\}$ separates points. If A, B are increasing, then so is $A \cap B$, so by the fact that $1_A 1_B = 1_{A\cap B}$ we see that $\{1_A : A \text{ increasing}\}$ is closed under products. ∎

We continue to consider spaces of the form S^Λ where S is a finite partially ordered set and Λ is countable. In particular, since Λ be be a set with only one element, this includes arbitrary finite partially ordered sets. By definition, a probability kernel K on S^Λ is *monotone* if it satisfies the following equivalent conditions. Note that in (i), $\leq$ denotes the stochastic order. The equivalence of (i) and (ii) is a trivial consequence of Theorem 5.1.

(i) $K(x, \cdot\,) \leq K(y, \cdot\,)$ for all $x \leq y$.

(ii) Kf is monotone whenever $f \in \mathcal{C}(S^\Lambda)$ is monotone.

We note that if K is monotone, then

$$\mu \leq \nu \quad \text{implies} \quad \mu K \leq \nu K. \tag{5.1}$$

Indeed, for each montone $f \in B(S^\Lambda)$, the function Kf is also monotone and hence $\mu \leq \nu$ implies that $\mu K f \leq \nu K f$.

By definition, a *random mapping representation* of a probability kernel K is a random map M such that

$$K(x, \cdot\,) = \mathbb{P}[M(x) \in \cdot\,] \quad \forall x. \tag{5.2}$$

We say that K can be represented in the class of monotone maps, or that K is *monotonically representable*, if there exists a random monotone map M such that (5.2) holds. We recall from Section 2.4 that when a Markov generator G is written in the form

$$Gf(x) = \sum_{m\in\mathcal{G}} r_m \big\{ f\big(m(x)\big) - f\big(x\big) \big\}, \tag{5.3}$$

then we call (5.3) a *random mapping representation* of G. If the set $\mathcal{G}$ can be chosen such that all maps $m \in \mathcal{G}$ are monotone, then we say that G is *monotonically representable.*

Lemma 5.3 (Monotone representability) *Each monotonically representable probability kernel is monotone. If the generator of an interacting particle system is monotonically representable, then, for each $t \geq 0$, the transition probability P_t is a monotonically representable probability kernel .*

Proof If a probability kernel K can be written in the form (5.2) with M a random monotone map, then for each $x \leq y$, the random variables $M(x)$ and $M(y)$ are coupled such that $M(x) \leq M(y)$ a.s., so their laws are stochastically ordered as $K(x, \cdot\,) \leq K(y, \cdot\,)$. Since this holds for all $x \leq y$, the kernel K is monotone.

Given a random mapping representation of the form (5.3) of the generator G of an interacting particle system, we can construct a stochastic flow $(\mathbf{X}_{s,t})_{s \leq t}$ as in Theorem 4.14 based on a Poisson set $\omega \subset \mathcal{G} \times \mathbb{R}$. If all maps $m \in \mathcal{G}$ are monotone, then the random maps $\mathbf{X}_{s,t} : S^\Lambda \to S^\Lambda$ are also monotone, since they are pointwise defined as the concatenation of finitely many maps from $\mathcal{G}$. It follows that

$$P_t(x, \cdot\,) = \mathbb{P}\big[\mathbf{X}_{0,t}(x) \in \cdot\,\big]$$

is a representation of P_t in terms of the random monotone map $\mathbf{X}_{0,t}$, so P_t is monotonically representable. ∎

We say that an interacting particle system is *monotone* if its transition kernels are monotone probability kernels, and we say that it is *monotonically representable* if its generator is monotonically representable. Somewhat surprisingly, it turns out that for probability kernels, "monotonically representable" is a strictly stronger concept than being "monotone". See [FM01] for an example of a probability kernel on $\{0,1\}^2$ that is monotone but not monotonically representable. Nevertheless, it turns out that (almost) all monotone interacting particle systems that one encounters in practice are also monotonically representable.

The following maps are examples of monotone maps:

- The voter map $\texttt{vot}_{ij}$ defined in (1.4).
- The branching map $\texttt{bra}_{ij}$ defined in (1.6).
- The death map $\texttt{death}_i$ defined in (1.7).
- The coalescing random walk map $\texttt{rw}_{ij}$ defined in (1.18).
- The exclusion map $\texttt{excl}_{ij}$ defined in (1.21).
- The cooperative branching map $\texttt{coop}_{ij}$ defined in (1.23).
- The maps $m^{\pm}_{i,L}$ defined in (4.27).

As a result, the following interacting particle systems are monotonically representable (and hence, in particular, monotone):

- The voter model with generator as in (1.5).
- The contact process with generator as in (1.8).
- The ferromagnetic Ising model with Glauber dynamics, since its generator can be written as in (4.28).
- The biased voter model with generator as in (1.15).
- Systems of coalescing random walks with generator as in (1.19).
- The exclusion process with generator as in (1.22).
- Systems with cooperative branching and coalescence as in Figure 1.11.

On the other hand, the following maps are *not* monotone:

- The annihilating random walk map $\texttt{ann}_{ij}$ defined in (1.20).
- The killing map $\texttt{kill}_{ij}$ defined in (1.24).

Examples of interacting particle systems that are not monotone[2] are:

- The antiferromagnetic Ising model with Glauber dynamics.
- "Rebellious" voter models as in (1.16).
- Systems of annihilating random walks.
- The biased annihilating branching process of [Sud97, Sud99].

5.2 The upper and lower invariant laws

In the present section, we assume that the local state space is $S = \{0, 1\}$, which covers all examples of monotone interacting particle systems mentioned in the previous section. We also use the symbols 0 and 1 to denote the states in S^Λ that are identically 0 or 1, respectively. Below, δ_0 denotes the delta measure at the configuration that is identically 0, so $\delta_0 P_t$ denotes the law at time t of the process started in $X_0(i) = 0$ a.s. $(i \in \Lambda)$.

[2]Note that the fact that a given interacting particle system is represented in maps that are not monotone does not prove that the system is not monotone. Indeed, it is conceivable that the same system can also be monotonely represented.

Theorem 5.4 (Upper and lower invariant laws) *Let X be an interacting particle system with state space of the form $\{0,1\}^\Lambda$ and transition probabilities $(P_t)_{t\geq 0}$. Assume that X is monotone. Then there exist invariant laws $\underline{\nu}$ and $\overline{\nu}$ such that*

$$\delta_0 P_t \underset{t\to\infty}{\Longrightarrow} \underline{\nu} \quad \text{and} \quad \delta_1 P_t \underset{t\to\infty}{\Longrightarrow} \overline{\nu}.$$

If ν is any other invariant law, then $\underline{\nu} \leq \nu \leq \overline{\nu}$.

The invariant laws $\underline{\nu}$ and $\overline{\nu}$ from Theorem 5.4 are called *lower* and *upper invariant law*, respectively. Before we give the proof of Theorem 5.4, we start with two preparatory lemmas.

Lemma 5.5 (Equal mean) *Let μ,ν be probability laws on $\{0,1\}^S$ such that $\mu \leq \nu$ and*

$$\int \mu(\mathrm{d}x)\, x(i) \geq \int \nu(\mathrm{d}x)\, x(i) \qquad (i \in \Lambda).$$

Then $\mu = \nu$.

Proof By Theorem 5.1, we can couple random variables with laws $\mathbb{P}[X \in \cdot\,] = \mu$ and $\mathbb{P}[Y \in \cdot\,] = \nu$ in such a way that $X \leq Y$. Now $\mathbb{E}[X(i)] \geq \mathbb{E}[Y(i)]$ implies $\mathbb{E}[Y(i) - X(i)] \leq 0$. Since $Y(i) - X(i) \geq 0$ a.s., it follows that $X(i) = Y(i)$. In particular, if this holds for all $i \in \Lambda$, then $\mu = \nu$. ∎

Lemma 5.6 (Monotone convergence of probability laws) *Let $(\nu_n)_{n\geq 0}$ be a sequence of probability laws on $\{0,1\}^\Lambda$ that are stochastically ordered as $\nu_k \leq \nu_{k+1}$ $(k \geq 0)$. Then there exists a probability law ν on $\{0,1\}^\Lambda$ such that $\nu_n \Rightarrow \nu$, i.e., the ν_n's converge weakly to ν.*

Proof Since $\nu_n f$ increases to a finite limit for each monotone $f \in \mathcal{C}(\{0,1\}^\Lambda)$, this is an immediate consequece of Lemmas 5.2 and 4.25. ∎

Proof of Theorem 5.4 By symmetry, it suffices to prove the statement for $\underline{\nu}$. Since 0 is the lowest possible state, for each $t \geq 0$, we trivially have

$$\delta_0 \leq \delta_0 P_t$$

By (5.1), this implies that

$$\delta_0 P_s \leq \delta_0 P_t P_s = \delta_0 P_{t+s} \qquad (s,t \geq 0),$$

which shows that $t \mapsto \delta_0 P_t$ is nondecreasing with respect to the stochastic order. By Lemma 5.6, each monotone sequence of probability laws has a weak limit, so there exists a probability law $\underline{\nu}$ on $\{0,1\}^\Lambda$ such that

$$\delta_0 P_t \underset{t\to\infty}{\Longrightarrow} \underline{\nu}.$$

It follows from Lemma 4.27 that $\underline{\nu}$ is an invariant law.

To complete the proof of the theorem, we observe that if ν is any other invariant law, then, by (5.1), for any monotone $f \in \mathcal{C}(\{0,1\}^\Lambda)$,

$$\delta_0 \leq \nu \quad \Rightarrow \quad \delta_0 P_t \leq \nu P_t = \nu \quad (t \geq 0).$$

Letting $t \to \infty$, if follows that $\underline{\nu} f \leq \nu f$ for all monotone $f \in \mathcal{C}(\{0,1\}^\Lambda)$, which by Theorem 5.1 implies that $\underline{\nu} \leq \nu$. ∎

Theorem 5.7 (Ergodicity of monotone systems) *Let X be a monotone interacting particle system with state space $\{0,1\}^\Lambda$ and upper and lower invariant laws $\underline{\nu}$ and $\overline{\nu}$. If*

$$\int \underline{\nu}(\mathrm{d}x) x(i) = \int \overline{\nu}(\mathrm{d}x) x(i) \qquad \forall i \in \Lambda, \tag{5.4}$$

then X has a unique invariant law $\nu := \underline{\nu} = \overline{\nu}$ and is ergodic in the sense that

$$\mathbb{P}^x\big[X_t \in \cdot\,\big] \underset{t\to\infty}{\Longrightarrow} \nu \qquad (x \in \{0,1\}^\Lambda).$$

On the other hand, if (5.4) does not hold, then X has at least two invariant laws.

Proof By Lemma 5.5, (5.4) is equivalent to the condition that $\underline{\nu} = \overline{\nu}$. It is clear that if $\underline{\nu} \neq \overline{\nu}$, then X has at least two invariant laws and ergodicity cannot hold. On the other hand, by Theorem 5.4, any invariant law ν satisfies $\underline{\nu} \leq \nu \leq \overline{\nu}$, so if $\underline{\nu} = \overline{\nu}$, then $\nu = \underline{\nu} = \overline{\nu}$.

To complete the proof, we must show that $\underline{\nu} = \overline{\nu} =: \nu$ implies $\delta_x P_t \Rightarrow \nu$ as $t \to \infty$ for all $x \in \{0,1\}^\Lambda$. Since

$$\delta_0 P_t f \leq \delta_x P_t f \leq \delta_1 P_t f$$

for all monotone $f \in \mathcal{C}(\{0,1\}^\Lambda)$, we see that

$$\underline{\nu} f \leq \liminf_{t\to\infty} P_t f \leq \limsup_{t\to\infty} P_t f \leq \overline{\nu} f$$

The claim now follows from Lemmas 4.25 and 5.2. ∎

To state the final result of this section, we need a bit of theory. We observe that for any interacting particle system, the set $\mathcal{I}$ of all invariant laws is a compact, convex subset of $\mathcal{M}_1(S^\Lambda)$. Indeed, if μ and ν are invariant laws and $p \in [0,1]$, then clearly

$$P_t(p\mu + (1-p)\nu) = pP_t\mu + (1-p)P_t\nu = p\mu + (1-p)\nu \quad (t \geq 0),$$

proving that $p\mu + (1-p)\nu$ is an invariant law. The fact that $\mathcal{I}$ is closed follows from Proposition 4.30. Since $\mathcal{M}_1(S^\Lambda)$ is compact, $\mathcal{I}$ is also compact.

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.,

$$\nu = p\nu_1 + (1-p)\nu_2 \qquad (0 < p < 1,\ \nu_1, \nu_2 \in \mathcal{I}) \quad \text{implies} \quad \nu_1 = \nu_2 = \nu.$$

We let

$$\mathcal{I}_\mathrm{e} := \{\nu \in \mathcal{I} : \nu \text{ is an extremal element of } \mathcal{I}\}.$$

Since $\mathcal{I}$ is compact and convex, Choquet's theorem implies that each invariant law ν can be written as

$$\nu = \int \rho_\nu(\mathrm{d}\mu)\mu,$$

where ρ_ν is a probability measure on $\mathcal{I}_\mathrm{e}$. In practice, it happens quite often[3] that $\mathcal{I}_\mathrm{e}$ is a finite set.[4] In this case, Choquet's theorem simply says that each invariant law is a convex combination of the extremal invariant laws, i.e., each invariant law is of the form

$$\nu = \sum_{\mu \in \mathcal{I}_\mathrm{e}} p(\mu)\mu,$$

where $(p(\mu))_{\mu \in \mathcal{I}_\mathrm{e}}$ are nonnegative constants, summing up to one. In view of this, we are naturally interested in finding all extremal invariant laws of a given interacting particle system.

Lemma 5.8 (The lower and upper invariant law are extremal) *Let X be a monotone interacting particle system with state space $\{0,1\}^\Lambda$ and upper and lower invariant laws $\underline{\nu}$ and $\overline{\nu}$. Then $\underline{\nu}$ and $\overline{\nu}$ are extremal invariant laws of X.*

Proof By symmetry, it suffices to prove the statement for $\overline{\nu}$. Imagine that

$$\overline{\nu} = p\nu_1 + (1-p)\nu_2 \quad \text{for some} \quad 0 < p < 1,\ \nu_1, \nu_2 \in \mathcal{I}.$$

By Theorem 5.4, for each monotone $f \in B(\{0,1\}^\Lambda)$, one has $\nu_1 f \leq \overline{\nu} f$ and $\nu_2 f \leq \overline{\nu} f$. Since

$$p(\overline{\nu} f - \nu_1 f) + (1-p)(\overline{\nu} f - \nu_2 f) = 0,$$

it follows that $\overline{\nu} f = \overline{\nu}_1 f = \overline{\nu}_2 f$. Since this holds for each monotone f, we conclude (by Lemma 5.2) that $\overline{\nu} = \nu_1 = \nu_2$. ∎

[3]Though the voter model in dimensions $d \geq 3$ is a countexample.

[4]This may, however, be quite difficult to prove!

Exercise 5.9 *Let X be an interacting particle system with state space $\{0,1\}^\Lambda$ and generator G. Assume that G has a random mapping representation in terms of monotone maps and let $(\mathbf{X}_{s,t})_{s\leq t}$ be the corresponding stochastic flow as in Theorem 4.14. Show that the a.s. limits*

$$\left.\begin{array}{l}\underline{X}_t := \lim_{s\to-\infty} \mathbf{X}_{s,t}(0),\\ \overline{X}_t := \lim_{s\to-\infty} \mathbf{X}_{s,t}(1)\end{array}\right\} \quad (t\in\mathbb{R})$$

define stationary Markov processes $(\underline{X}_t)_{t\in\mathbb{R}}$ and $(\overline{X}_t)_{t\in\mathbb{R}}$ whose invariant laws

$$\underline{\nu} = \mathbb{P}[\underline{X}_t \in \cdot\,] \quad \textit{and} \quad \overline{\nu} = \mathbb{P}[\overline{X}_t \in \cdot\,] \qquad (t\in\mathbb{R})$$

are the lower and upper invariant law of X, respectively. Show that (5.4) implies that

$$\lim_{s\to-\infty} \mathbf{X}_{s,t}(x) = \underline{X}_t = \overline{X}_t \quad \text{a.s.} \quad (x\in\{0,1\}^\Lambda,\ t\in\mathbb{R}).$$

5.3 The contact process

We recall the definition of the contact process from (1.8). Since both the branching and death map are monotone, this is a monotonically representable interacting particle system, so by Theorem 5.4, it has a lower and upper invariant law $\underline{\nu}$ and $\overline{\nu}$. Since $\mathtt{bra}_{ij}(0) = 0$ and $\mathtt{death}_i(0) = 0$ for each $i, j \in \Lambda$, the all-zero configuration 0 is a trap for the contact process, so $\delta_0 P_t = \delta_0$ for all $t \geq 0$ and hence

$$\underline{\nu} = \delta_0.$$

Therefore, by Theorem 5.7, the contact process is ergodic if and only if the function

$$\theta(\lambda) := \int \overline{\nu}(\mathrm{d}x)\, x(i) \qquad (i\in\mathbb{Z}^d) \tag{5.5}$$

satisfies $\theta(\lambda) > 0$. Here λ denotes the infection rate and we stick to the convention to take the recovery rate δ (1.8) equal to 1. We note that by translation invariance, for the model on $\mathbb{Z}^d$ (either nearest-neighbor or range R), the density $\int \overline{\nu}(\mathrm{d}x)\, x(i)$ of the upper invariant law does not depend on $i \in \mathbb{Z}^d$. For reasons that will become clear in the next chapter, $\theta(\lambda)$ is actually the same as the survival probability started from a single occupied site, i.e., this is the function in Figure 1.4.

By definition, we say that a probability law μ on $\{0,1\}^\Lambda$ is *nontrivial* if

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

i.e., if μ gives zero probability to the all-zero configuration.

Lemma 5.10 (Nontriviality of the upper invariant law) *For the contact process, if $\overline{\nu} \neq \delta_0$, then $\overline{\nu}$ is nontrivial.*

Proof We can always write $\overline{\nu} = (1-p)\delta_0 + p\mu$ where $p \in [0,1]$ and μ is a nontrivial law. By assumption, $\overline{\nu} \neq \delta_0$, so $p > 0$. Since $\overline{\nu}$ and δ_0 are invariant laws, μ must be an invariant law too. By Lemma 5.8, $\overline{\nu}$ cannot be written as a nontrivial convex combination of other invariant laws, so we conclude that $p = 1$. ∎

Proposition 5.11 (Monotonicity in the infection rate) *Let $\overline{\nu}_\lambda$ denote the upper invariant law of the contact process with infection rate λ. Then $\lambda \leq \lambda'$ implies $\overline{\nu}_\lambda \leq \overline{\nu}_{\lambda'}$. In particular, the function $\lambda \mapsto \theta(\lambda)$ is nondecreasing.*

Proof Let X and X' be contact processes started in the initial state $X_0 = 1 = X'_0$ and with infection rates λ and λ'. It suffices to prove that X and X' can be coupled such that $X_t \leq X'_t$ for all $t \geq 0$.

We use a Poisson construction, based on the random mapping representation (1.8). We write $\mathcal{G} = \mathcal{G}_{\rm bra} \cup \mathcal{G}_{\rm death}$ where

$$\mathcal{G}_{\rm bra} := \{\texttt{bra}_{ij} : (i,j) \in \mathcal{E}^d\} \quad \text{and} \quad \mathcal{G}_{\rm death} := \{\texttt{death}_i : i \in \mathbb{Z}^d\}.$$

Then X can be constructed as in Theorem 4.14 from a Poisson point set ω on

$$\mathcal{G} \times \mathbb{R} = (\mathcal{G}_{\rm bra} \cup \mathcal{G}_{\rm death}) \times \mathbb{R},$$

with intensity measure ρ_λ given by

$$\rho_\lambda(\{m\} \times A) := \begin{cases} \lambda\ell(A) & \text{if } m \in \mathcal{G}_{\rm bra}, \\ \ell(A) & \text{if } m \in \mathcal{G}_{\rm death}, \end{cases} \qquad \big(A \in \mathcal{B}(\mathbb{R})\big),$$

where ℓ denotes the Lebesgue measure. Likewise, X' can be constructed from a Poisson point set ω' with intensity $\rho_{\lambda'}$. We claim that we can couple ω and ω' in such a way that the latter has more branching events, and the same death events as ω. This can be done as follows. Let ω'' be a Poisson point set on $\mathcal{G} \times \mathbb{R}$, independent of ω, with intensity measure $\rho'' := \rho_{\lambda'} - \rho_\lambda$, i.e.,

$$\rho''(\{m\} \times A) := \begin{cases} (\lambda' - \lambda)\ell(A) & \text{if } m \in \mathcal{G}_{\rm bra}, \\ 0 & \text{if } m \in \mathcal{G}_{\rm death}, \end{cases} \qquad \big(A \in \mathcal{B}(\mathbb{R})\big).$$

Since the sum of two independent Poisson sets yields another Poisson set, setting

$$\omega' := \omega + \omega''$$

defines a Poisson point set with intensity $\rho_{\lambda'}$. We observe that

$$\begin{aligned} x \leq x' &\quad \text{implies} \quad \mathtt{bra}_{ij}(x) \leq \mathtt{bra}_{ij}(x'), \\ x \leq x' &\quad \text{implies} \quad \mathtt{death}_i(x) \leq \mathtt{death}_i(x'), \\ x \leq x' &\quad \text{implies} \quad x \leq \mathtt{bra}_{ij}(x'). \end{aligned}$$

The first two statements just say that the maps $\mathtt{bra}_{ij}$ and $\mathtt{death}_i$ are monotone. The third statement says that if we apply a branching map only to the larger configuration x', then the order between x and x' is preserved.

Since ω' has the same branching and death events as ω, plus some extra branching events, we conclude that the stochastic flows $(\mathbf{X}_{s,t})_{s\leq t}$ and $(\mathbf{X}'_{s,t})_{s\leq t}$ constructed from ω and ω' satisfy

$$x \leq x' \quad \text{implies} \quad \mathbf{X}_{s,t}(x) \leq \mathbf{X}'_{s,t}(x') \qquad (s \leq t).$$

In particular, setting $X_t := \mathbf{X}_{0,t}(1)$ and $X'_t := \mathbf{X}'_{0,t}(1)$ yields the desired coupling between X and X'. ∎

Exercise 5.12 *Let X be a contact process on a graph Λ where each site i has exactly $|\mathcal{N}_i| = N$ neighbors. Calculate the constant K from (4.13) and apply Theorem 4.22 to conclude that*

$$\lambda N < 1 \quad \textit{implies} \quad \overline{\nu} = \delta_0.$$

In Chapter 7, we will prove that $\theta(\lambda) > 0$ for λ sufficienty large.

5.4 Other examples

The Ising model with Glauber dynamics

We have seen in (4.28) that the generator of the Ising model with Glauber dynamics is monotonicaly representable, so by Theorem 5.4, [5] it has a lower and upper invariant law $\underline{\nu}$ and $\overline{\nu}$. We let

$$m_*(\beta) := \int \overline{\nu}(\mathrm{d}x)\, x(i),$$

which is independent of i if the processes has some translation invariant structure (like the nearest neighbor or range R processes on $\mathbb{Z}^d$). For reasons that cannot be explained here, this function is actually the same as the

[5] The difference between the local state space $\{-1, 1\}$ of the ising model and $\{0, 1\}$ of Theorem 5.4 is of course entirely notational.

one defined in (1.13), i.e., this is the *spontaneous magnetization* of the Ising model, see Figure 1.6. By the symmetry between +1 and −1 spins, we clearly have

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

By Theorem 4.23, we have

$$e^{\beta N} < \frac{N}{N-1} \quad \text{implies} \quad \underline{\nu} = \overline{\nu},$$

from which we conclude that $m_*(\beta) = 0$ for β sufficiently small,

The function $\beta \mapsto m_*(\beta)$ is nondecreasing, but this cannot be used with the sort of techniques used in Proposition 5.11. The lower and upper invariant laws of the Ising model with Glauber dynamics are infinite volume Gibbs measures, and much of the analysis of the Ising model is based on this fact. In fact, the Ising model with Glauber dynamics is just one example of an interacting particle system that has these Gibbs measures as its invariant laws. In general, interacting particle systems with this property are called stochastic Ising models, and the Gibbs measures themselves are simply called the Ising model. We refer to [Lig85, Chapter IV] for an exposition of this material. In particular, in [Lig85, Thm IV.3.14], it is shown that for the nearest-neighbor model on $\mathbb{Z}^2$, one has $m_*(\beta) > 0$ for β sufficiently large.

The voter model

Consider a voter model with local state space $S = \{0, 1\}$. Since the voter maps $\mathtt{vot}_{ij}$ from (1.4) are monotone, the voter model is monotonically representable. Since both the constant configurations 0 and 1 are traps, so

$$\underline{\nu} = \delta_0 \quad \text{and} \quad \overline{\nu} = \delta_1,$$

so we conclude (recall Theorem 5.7) that the voter model is never ergodic. For the model on $\mathbb{Z}^d$, it is proved in [Lig85, Thm V.1.8] that if $d = 1, 2$, then δ_0 and δ_1 are the only extremal invariant laws. On the other hand, in dimensions $d \geq 3$, the set $\mathcal{I}_\mathrm{e}$ of extremal invariant laws is of the form $\{\nu_p : p \in [0, 1]\}$ where the invariant measure ν_n has intensity $\int \nu_p(\mathrm{d}x)\, x(i) = p$. We will give a partial proof of these statements in Chapter 6.

5.5 Exercises

Exercise 5.13 *Give an example of two probability measures μ, ν on a set of the form $\{0,1\}^\Lambda$ that satisfy*

$$\int \mu(\mathrm{d}x)x(i) \leq \int \nu(\mathrm{d}x)x(i) \qquad (i \in \Lambda),$$

but that are not *stochastically ordered as $\mu \leq \nu$.*

Exercise 5.14 *Let $(X^\lambda_t)_{t\geq 0}$ denote the contact process with infaction rate λ (and death rate one), started in $X^\lambda_0 = 1$. Apply Corollary 4.29 to prove that for each fixed $t \geq 0$, the function*

$$\theta_t(\lambda) := \mathbb{P}[\mathbf{X}^\lambda_{0,t}(1)(i) = 1] \tag{5.6}$$

depends continuously on λ. Use this to conclude that the function $\theta(\lambda)$ from (5.5) is right-continuous. Hint: *Use that the decreasing limit of continuous functions is upper semi-continuous.*

For the next exercise, let us define a *double death* map

$$\texttt{death}_{ij}x(k) := \begin{cases} 0 & \text{if } k \in \{ij\}, \\ x(k) & \text{otherwise.} \end{cases} \tag{5.7}$$

Recall the branching map $\texttt{bra}_{ij}$ defined in (1.6), the death map $\texttt{death}_i$ defined in (1.7), and the cooperative branching map $\texttt{coop}_{ij}$ defined in (1.23). Consider the cooperative branching process X with values in $\{0,1\}^\mathbb{Z}$ with generator

$$\begin{aligned} G_X f(x) = \lambda \sum_{i\in\mathbb{Z}} \sum_{\sigma\in\{-1,+1\}} &\big\{f\big(\texttt{coop}_{i,i+\sigma,i+2\sigma}x\big) - f\big(x\big)\big\} \\ + \sum_{i\in\mathbb{Z}} &\big\{f\big(\texttt{death}_i x\big) - f\big(x\big)\big\}, \end{aligned}$$

and the contact process with double deaths Y with generator

$$\begin{aligned} G_Y f(y) = \lambda \sum_{i\in\mathbb{Z}} \sum_{\sigma\in\{-1,+1\}} &\big\{f\big(\texttt{bra}_{i,i+\sigma}y\big) - f\big(y\big)\big\} \\ + \sum_{i\in\mathbb{Z}} &\big\{f\big(\texttt{death}_{i,i+1} y\big) - f\big(y\big)\big\}, \end{aligned}$$

Exercise 5.15 *Let X be the process with cooperative branching defined above and set*

$$X^{(2)}_t(i) := 1_{\{X_t(i)=1=X_t(i+1)\}} \qquad (i \in \mathbb{Z},\ t \geq 0).$$

Show that X can be coupled to a contact process with double deaths Y (with the same parameter λ) in such a way that

$$Y_0 \leq X_0^{(2)} \quad \textit{implies} \quad Y_t \leq X_t^{(2)} \quad (t \geq 0).$$

Exercise 5.16 *Show that a system $(X_t)_{t\geq 0}$ of annihilating random walks can be coupled to a system $(Y_t)_{t\geq 0}$ of coalescing random walks such that*

$$X_0 \leq Y_0 \quad \textit{implies} \quad X_t \leq Y_t \quad (t \geq 0).$$

Note that the annihilating random walks are not a monotone particle system.

Exercise 5.17 *Let X be a system of branching and coalescing random walks with generator*

$$\begin{aligned} G_X f(x) = &\tfrac{1}{2} b \sum_{i\in\mathbb{Z}} \sum_{\sigma\in\{-1,+1\}} \big\{ f\big(\mathtt{bra}_{i,i+\sigma} x\big) - f\big(x\big)\big\} \\ &+ \tfrac{1}{2} \sum_{i\in\mathbb{Z}} \big\{ f\big(\mathtt{rw}_{i,i+\sigma} x\big) - f\big(x\big)\big\}, \end{aligned}$$

and let Y be a system of coalescing random walks with positive drift, with generator

$$\begin{aligned} G_Y f(y) = &\tfrac{1}{2}(1+b) \sum_{i\in\mathbb{Z}} \big\{ f\big(\mathtt{rw}_{i,i+1} y\big) - f\big(y\big)\big\} \\ &+ \tfrac{1}{2} \sum_{i\in\mathbb{Z}} \big\{ f\big(\mathtt{rw}_{i,i-1} y\big) - f\big(y\big)\big\}. \end{aligned}$$

Show that X and Y can be coupled such that

$$Y_0 \leq X_0 \quad \textit{implies} \quad Y_t \leq X_t \quad (t \geq 0).$$

Exercise 5.18 *Let $d < d'$ and identify $\mathbb{Z}^d$ with the subset of $\mathbb{Z}^{d'}$ consisting of all $(i_1, \ldots, i_{d'})$ with $(i_{d+1}, \ldots, i_{d'}) = (0, \ldots, 0)$. Let X and X' denote the nearest-neighbor contact processes on $\mathbb{Z}^d$ and $\mathbb{Z}^{d'}$, respectively, with the same infection rate λ. Show that X and X' can be coupled such that*

$$X_0(i) \leq X'_0(i) \qquad (i \in \mathbb{Z}^d)$$

implies

$$X_t(i) \leq X'_t(i) \qquad (t \geq 0,\ i \in \mathbb{Z}^d).$$

Prove the same when X is the nearest-neighbor process and X' is the range R process (both on $\mathbb{Z}^d$).

Chapter 6

Duality

6.1 Introduction

In Figure 4.1, we have already seen an example of a graphical representation of a contact process, together with an example of the set $\zeta_s^{\{k\},u}$ of sites whose value at time s is relevant for the value of k at time u. In Lemma 4.12, we have already seen that for quite general interacting particle systems, under suitable summability conditions on the rates, the "backwards in time" process

$$(\zeta_{u-t}^{\{k\},u})_{t\geq 0}$$

is a Markov process with values in the set of finite subsets of the lattice Λ, and that the expected size of $\zeta_{u-t}^{\{k\},u}$ grows at most exponentially in t.

In the particular case of the contact process, by looking at Figure 4.1 and remembering how the maps $\mathtt{bra}_{ij}$ and $\mathtt{death}_i$ are defined, we can make some interesting observations:

(i) The set-valued process $(\zeta_{u-t}^{\{k\},u})_{t\geq 0}$, or rather the process of the corresponding indicator functions, is itself a contact process.

(ii) The site k is infected at time u if and only if at least one site in $\zeta_s^{\{k\},u}$ is infected at time s.

Observation (ii) means that we can construct X_t only by knowing the initial state X_0 and knowing the sets $\zeta_0^{\{k\},t}$ for each $k \in \Lambda$. This idea of "looking back in time" leads to the very useful concept of *duality*.

To demonstrate the usefulness of this idea, in Section 6.5, we will use "looking back in time" considerations to show that the voter model clusters in dimensions $d = 1, 2$, but not in dimensions $d \geq 3$. In Section 6.6, we use the self-duality of the contact process to prove that for processes with some sort of translation invariant structure, the upper invariant law is the limit law started from any nontrivial translation invariant initial law, and we will show that this in turn implies that the function $\theta(\lambda)$ from (5.5) is continuous everywhere, except possibly at the critical point. Finally, in Section 6.7, we use duality to show that for a model with a mixture of voter model and contact process duality, the critical points associated with survival and ontriviality of the upper invariant law coincide.

Before we come to these applications, we first develop the observations (i) and (ii) into a more general idea, which will first lead to the concept of additive systems duality, and then Markov process duality more generally.

6.2 Additive systems duality

By definition, a map $m : \{0,1\}^\Lambda \to \{0,1\}^\Lambda$ is *additive* iff

$$\begin{aligned} &\text{(i) } m(0) = 0, \\ &\text{(ii) } m(x \vee y) = m(x) \vee m(y) \qquad \big(x, y \in \{0,1\}^\Lambda\big). \end{aligned}$$

Since we will only be interested in local maps, in view of Exercise 4.10, we can assume without loss of generality that Λ is finite. For $i \in \Lambda$, let $1_{\{i\}}$ denote the indicator function of i, i.e., the element of $\{0,1\}^\Lambda$ such that $1_{\{i\}}(i) = 1$ and $1_{\{i\}}(j) = 0$ for all $i \neq j$. Since

$$m(x) = \bigvee_{i:\, x(i)=1} m(1_{\{i\}}),$$

an additive map is uniquely characterized by its action on configurations of the form $1_{\{i\}}$. It is easy to see that additive maps are monotone. Examples of additive maps are:

- The voter map $\mathtt{vot}_{ij}$ defined in (1.4).
- The branching map $\mathtt{bra}_{ij}$ defined in (1.6).
- The death map $\mathtt{death}_i$ defined in (1.7).
- The coalescing random walk map $\mathtt{rw}_{ij}$ defined in (1.18).
- The exclusion map $\mathtt{excl}_{ij}$ defined in (1.21).

On the other hand, the following maps are monotone, but not additive:

- The cooperative branching map $\mathtt{coop}_{ij}$ defined in (1.23).

- The maps $m^{\pm}_{i,L}$ used to construct the Ising model with Glauber dynamics in (4.28).

An interacting particle system is called *additive* if its generator can be represented in additive local maps. Examples of additive particle systems are

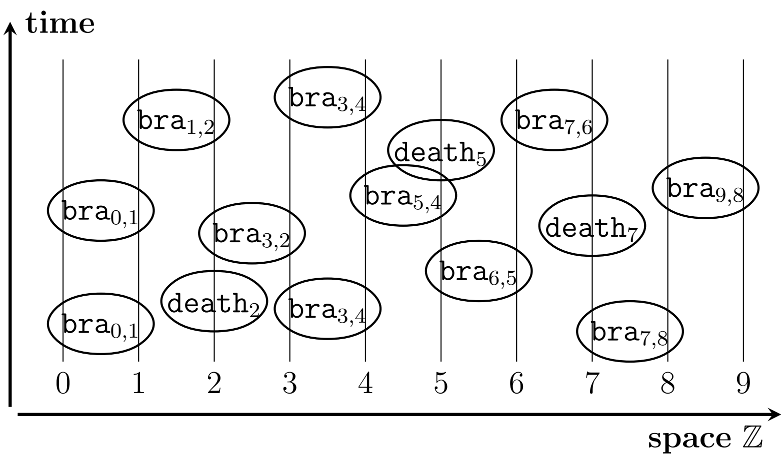
- The voter model with generator as in (1.5).
- The contact process with generator as in (1.8).
- The biased voter model with generator as in (1.15).
- Systems of coalescing random walks with generator as in (1.19).
- The exclusion process with generator as in (1.22).

In the graphical representation of an additive particle system, we visualize an event $(m,t) \in \omega$ where m is an additive local map in the following way:

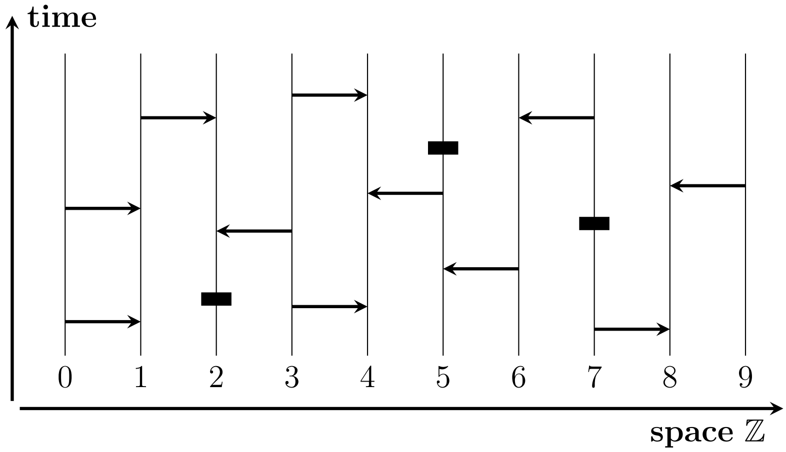
(i) For each $i \neq j$ such that $m(1_{\{i\}})(j) = 1$, we draw an arrow from (i,t) to (j,t)

(ii) For each i such that $m(1_{\{i\}})(i) = 0$, we draw a blocking symbol ▬ at (i,t).

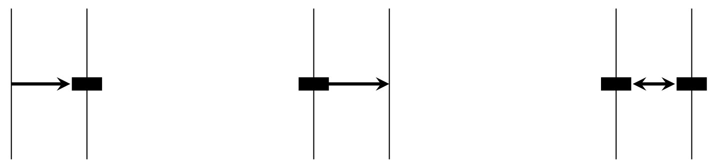
In Figure 4.1, we drew the graphical representation of a contact process in the following fashion:



With our new conventions, the same graphical representation looks as follows:



The voter model map $\mathtt{vot}_{ij}$, coalescing random walk map $\mathtt{rw}_{ij}$, and exclusion map $\mathtt{excl}_{ij}$ look in the same convention as follows:



$\mathtt{vot}_{1,2}$ $\quad$ $\mathtt{rw}_{1,2}$ $\quad$ $\mathtt{excl}_{1,2}$

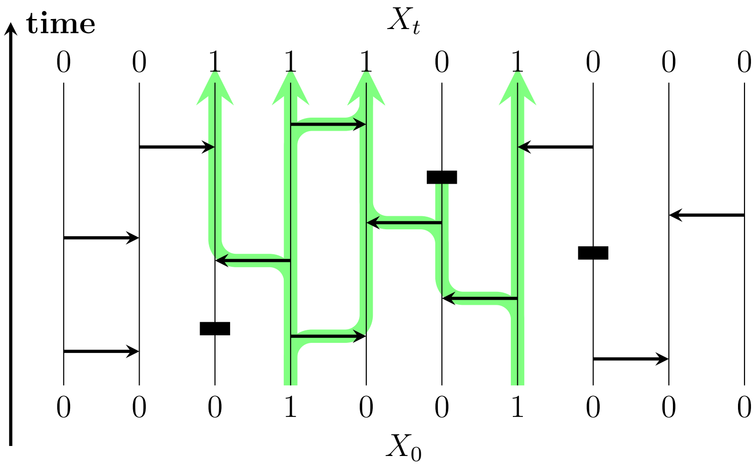
For any $i, j \in \Lambda$ and $s < u$, by definition, an *open path* from (i, s) to (j, u) is a cadlag function $\gamma : (s, u] \to \Lambda$ such that $\gamma_{s-} = i$, $\gamma_u = j$, and

$$\begin{array}{rl}\text{(i)} & \text{if } \gamma_{t-} \neq \gamma_t \text{ for some } t \in (s,u], \text{ then there is an} \\ & \text{arrow from } (\gamma_{t-}, t) \text{ to } (\gamma_t, t), \\ \text{(ii)} & \text{there exist no } t \in (s,u] \text{ such that } \gamma_{t-} = \gamma_t \\ & \text{while there is a blocking symbol at } (\gamma_t, t).\end{array} \tag{6.1}$$

In the context of additive systems, one can check that these open paths are exactly the paths of influence defined in (4.11). Moreover, the stochastic flow $(\mathbf{X}_{s,t})_{s \leq t}$ associated with the graphical representation of an additive particle system has the following simple description:

$$\mathbf{X}_{s,t}(x)(j) = 1 \quad \text{iff} \quad \begin{array}{l}\text{there exists an } i \in \Lambda \text{ such that } x(i) = 1 \\ \text{and an open path from } (i,s) \text{ to } (j,t).\end{array} \tag{6.2}$$

For example, for the graphical representation of the contact process that we earlier used as an example, the time evolution of the process $X_t := \mathbf{X}_{0,t}(X_0)$ $(t \geq 0)$ might look as follows:



Thanks to formula (6.2), there is a simple way to find out if at a given time t, the site j is in the state $X_t(j) = 1$: we simply follow all open paths ending at (j,t) backward in time, and check if at time zero one of these paths arrives at a site i with $X_0(i) = 1$. We observe that open paths backwards in time are in fact the open paths forward in time of a different graphical representation, that is obtained by turning the original graphical representation upside down and reversing the direction of all arrows. In other words, setting

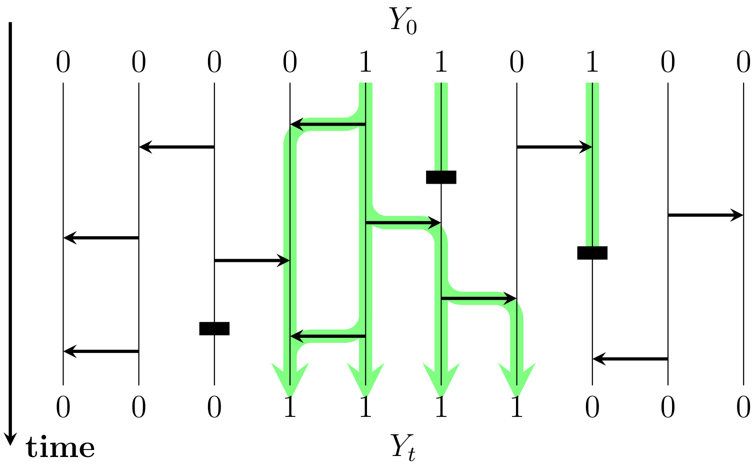
$$\mathbf{Y}_{t,s}(y)(i) = 1 \quad \text{iff} \quad \begin{array}{l} \text{there exists a } j \in \Lambda \text{ such that } x(j) = 1 \\ \text{and an open path from } (i,s) \text{ to } (j,t). \end{array} \tag{6.3}$$

defines a collection of random maps $(\mathbf{Y}_{t,s})_{t \geq s}$ that is almost a stochastic flow, except that time runs backwards; more precisely, setting

$$\hat{\mathbf{Y}}_{s,t} := \mathbf{Y}_{-s,-t} \qquad (s \leq t)$$

defines exactly[1] a stochastic flow that belongs to an (a priori) different additive particle system. For example, for the graphical representation of our contact process, reversing the direction of all arrows and letting time run downwards, the picture is as follows:

[1] Actually, this is still not completely correct, since $\hat{\mathbf{Y}}_{s,t}(y)$, for fixed y and s, is *left* continuous with *right* limits as a function of t, instead of cadlag. But since we will mostly be interested in deterministic times s,t, we can ignore this small technical complication for the moment.



We fix $t > 0$ and deterministic $X_0, Y_0 \in \{0,1\}^\Lambda$, and using the stochastic flows $(\mathbf{X}_{s,t})_{s\leq t}$ and $(\mathbf{Y}_{s,t})_{s\leq t}$ from (6.2) and (6.3), we define additive particle systems[2] X and Y by

$$\left.\begin{array}{r} X_s := \mathbf{X}_{0,s}(X_0) \\ Y_s := \mathbf{Y}_{t,t-s}(Y_0) \end{array}\right\} \quad (s \geq 0).$$

Then

$$\begin{array}{ll} X_t \wedge Y_0 = 0 & \\ \Leftrightarrow & \text{there is no open path from a point } (i,0) \text{ to a point } (j,t) \\ & \text{such that } X_0(i) = 1 \text{ and } Y_0(i) = 1 \\ \Leftrightarrow & X_0 \wedge Y_t = 0. \end{array}$$

In other words, we have coupled the processes X and Y in such a way that

$$1_{\{X_t \wedge Y_0 = 0\}} = 1_{\{X_0 \wedge Y_t = 0\}} \quad \text{a.s.}$$

In particular, taking expectations, this shows that

$$\mathbb{P}[X_t \wedge Y_0 = 0] = \mathbb{P}[X_0 \wedge Y_t = 0] \qquad (t \geq 0).$$

We note that these relations are also true for processes with random initial states X_0 and Y_0, as long as we take X_0 and Y_0 independent of each other and of the graphical representation ω. In this case X_t is independent of Y_0 and Y_t is independent of X_0.

The following proposition summarizes what we have discovered so far.

[2]The paths of Y, defined in this way, will be left continuous with right limits, but as before we ignore this small complication for the moment.

Proposition 6.1 (Additive systems duality) *For each additive local map* $m : \{0,1\}^\Lambda \to \{0,1\}^\Lambda$ *there exists a unique dual map* $m' : \{0,1\}^\Lambda \to \{0,1\}^\Lambda$ *such that*

$$1_{\{m(x) \wedge y = 0\}} = 1_{\{x \wedge m'(y) = 0\}} \qquad \big(x, y \in \{0,1\}^\Lambda\big), \tag{6.4}$$

and this dual map m' *is also an additive local map. Let* $\mathcal{G}$ *be a collection of additive local maps, let* $(r_m)_{m\in\mathcal{G}}$ *be nonnegative constants, and assume that the generators*

$$\begin{aligned} Gf(x) &:= \sum_{m\in\mathcal{G}} r_m \big\{ f\big(m(x)\big) - f\big(x\big) \big\}, \\ G'f(y) &:= \sum_{m\in\mathcal{G}} r_m \big\{ f\big(m'(y)\big) - f\big(y\big) \big\} \end{aligned} \tag{6.5}$$

both satisfy the summability condition (4.17). Let $X = (X_t)_{t\geq 0}$ *and* $Y = (Y_t)_{t\geq 0}$ *be interacting particle systems with generators* G *and* G'*, respectively. Then, for each* $t > 0$*, it is possible to couple* X *and* Y *in such a way that for each* $s \in [0,t]$*, the processes* X_s *and* Y_{t-s} *are independent, and*

$$1_{\{X_t \wedge Y_0 = 0\}} = 1_{\{X_{t-s} \wedge Y_s = 0\}} = 1_{\{X_0 \wedge Y_t = 0\}} \quad \text{a.s.} \tag{6.6}$$

In particular, if X_t *is independent of* Y_0 *and* Y_t *is independent of* X_0*, then*

$$\mathbb{P}[X_t \wedge Y_0 = 0] = \mathbb{P}[X_0 \wedge Y_t = 0] \qquad (t \geq 0). \tag{6.7}$$

Proof We have already seen that each additive local map m has a dual which can graphically be represented by reversing the arrows of m and keeping the blocking symbols in place. Knowing $1_{\{x \wedge m'(y)=0\}}$ for all $x \in \{0,1\}^\Lambda$ clearly determines $m'(y)$ uniquely, since $1_{\{i\}} \wedge m'(y) = 0$ if and only if $m'(y)(i) = 0$.

If G and G' both[3] satisfy the summability condition (4.17), then by Theorem 4.14, their graphical representations can be used to construct well-defined processes X and Y. We have alread seen how using these graphical representations, for fixed $t > 0$, it is possible to couple the processes X and Y such that

$$1_{\{X_t \wedge Y_0 = 0\}} = 1_{\{X_0 \wedge Y_t = 0\}} \quad \text{a.s.}$$

In fact, for the same coupling, if we fix any $s \in [0,t]$, then X_s is a function only of X_0 and Poisson events with times in $(0,s]$, while Y_{t-s} is a function only of Y_0 and Poisson events with times in $(s,t]$, so by the fact that the

[3]It is easy to find examples where G satisfies the summability condition (4.17) while G' does not, so in general, one has to check this condition for both G and G'.

restrictions of Poisson point sets to disjoint parts of space are independent, we see that X_s is independent of Y_{t-s}. Moreover, we observe that

$$\begin{aligned} & X_s \wedge Y_{t-s} = 0 \\ & \Leftrightarrow \quad \text{there is no open path from a point } (i,0) \text{ to a point } (j,t), \\ & \qquad \text{such that } X_0(i) = 1 \text{ and } Y_0(i) = 1 \end{aligned} \tag{6.8}$$

so (6.6) follows from our earlier arguments.[4] Taking expectations, this implies in particular (6.7). ∎

If two additive local maps m and m' are related as in (6.4), then we say that they are *dual* to each other. Using the recipe: "reverse the arrows and keep the blocking symbols in place", it is easy to find the duals of the additive local maps we have already seen. Indeed:

$$\begin{aligned} \mathtt{vot}'_{ij} &= \mathtt{rw}_{ji}, \\ \mathtt{bra}'_{ij} &= \mathtt{bra}_{ji}, \\ \mathtt{death}'_{i} &= \mathtt{death}_{i}, \\ \mathtt{excl}'_{ij} &= \mathtt{excl}_{ij}. \end{aligned} \tag{6.9}$$

We say that two additive interacting particle systems X and Y are *dual* if their generators G and G' satisfy (6.5). In particular, we see that the voter model is dual to a system of coalescing random walks, while the contact and exclusion processes are *self-dual*, i.e., they are their own duals.[5]

We note that if we known the expression in (6.7) for all finite initial states $Y_o = y$, then this determines the law of X_t uniquely. Indeed:

Lemma 6.2 (Distribution determining functions) *The functions* $\{f_y : y \in \{0,1\}^\Lambda,\ |y| < \infty\}$ *with* $f_y(x) := 1_{\{x \wedge y = 0\}}$ *are distribution determining.*

Proof Since $x \wedge 1_{\{i\}}(i) = x(i)$, the class $\{f_y : |y| < \infty\}$ separates points, and since $f_y f_{y'} = f_{y \vee y'}$, this class is closed under products. The claim now follows from Lemma 4.24. ∎

---

[4] Formula (6.6) holds a.s. for each fixed, deterministic $s \in [0,t]$. If we want (6.6) to hold a.s. for all $s \in [0,t]$ simultaneously, then we need to construct the processes X and Y in such a way that one has right-continuous sample paths while the other has left-continuous sample paths.

[5] For contact processes, this is only true provided that the process is symmetric in the sense that for each i, j, the map $\mathtt{bra}_{ij}$ is applied with the same rate as $\mathtt{bra}_{ji}$.

6.3 Cancellative systems duality

If we define a duality map $\psi : \{0,1\}^\Lambda \times \{0,1\}^\Lambda \to \{0,1\}$ by

$$\psi(x,y) := 1_{\{x \wedge y = 0\}} \qquad \big(x,y \in \{0,1\}^\Lambda\big), \tag{6.10}$$

then the additive systems duality (6.7) takes the form

$$\mathbb{E}\big[\psi(X_t, Y_0)\big] = \mathbb{E}\big[\psi(X_0, Y_t)\big] \qquad (t \geq 0), \tag{6.11}$$

where it is understood that X_t is independent of Y_0 and Y_t is independent of X_0, if the initial states are random. More generally, if (6.11) holds for a given bounded measurable function ψ and given Markov processes X and Y (for all initial states), then we say that the processes X and Y are *dual* with respect to the *duality function* ψ.

Also, if two maps m, m' satisfy

$$\psi\big(m(x), y\big) = \psi\big(x, m'(y)\big) \qquad \forall x, y, \tag{6.12}$$

then we say that m and m' are *dual* with respect to the duality function ψ. If two Markov generators G and G' are related as in (6.5), where m' denotes the dual of m with respect to some given duality function ψ, then exactly the same arguments as those leading up to (6.6) show that for each $t > 0$, it is possible to couple X and Y in such a way that for each $s \in [0,t]$, the processes X_s and Y_{t-s} are independent, and

$$\psi(X_t, Y_0) = \psi(X_s, Y_{t-s}) = \psi(X_0, Y_t) \quad \text{a.s.} \tag{6.13}$$

In this case, we say that the processes X and Y are *pathwise dual* to each other w.r.t. ψ.

To show that there are nontrivial examples of such sort of dualities, apart from additive systems duality, we start by considering *cancellative systems duality*. Let $\oplus$ denote addition modulo two, i.e.,

$$0 \oplus 0 := 0, \quad 0 \oplus 1 := 1, \quad 1 \oplus 0 := 1, \quad \text{and} \quad 1 \oplus 1 := 0.$$

By definition, a map $m : \{0,1\}^\Lambda \to \{0,1\}^\Lambda$ is *cancellative* if

$$m(0) = 0 \quad \text{and} \quad m(x \oplus y) = m(x) \oplus m(y) \qquad \big(x, y \in \{0,1\}^\Lambda\big).$$

Since

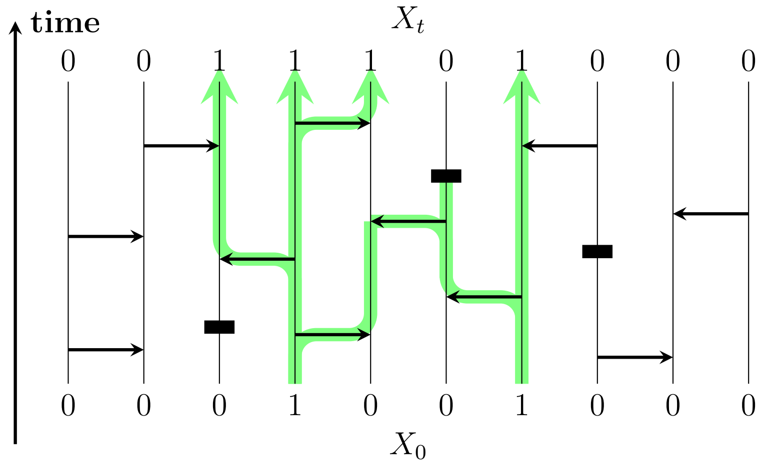
$$m(x) = \bigoplus_{i:\, x(i)=1} m(1_{\{i\}}),$$

a cancellative map is uniquely characterized by its action on configurations of the form $1_{\{i\}}$. In graphical representations of cancellative particle systems, we use the same conventions as for additive systems, i.e., we visualize an event $(m,t) \in \omega$ where m is a cancellative map as follows:

(i) For each $i \neq j$ such that $m(1_{\{i\}})(j) = 1$, we draw an arrow from (i,t) to (j,t)

(ii) For each i such that $m(1_{\{i\}})(i) = 0$, we draw a blocking symbol ▬ at (i,t).

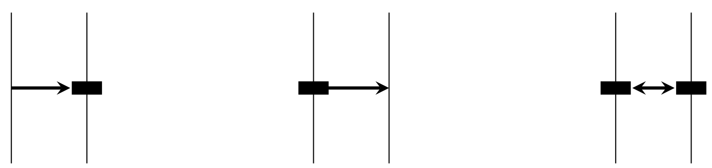
With these conventions, each graphical representation for an additive particle system can also be used to construct a cancellative system. For example, reusing the graphical representation of the contact process in this way, we obtain something that looks like this:



In this example, arrows represent the *annihilating branching map* $\texttt{bran}_{ij} : \{0,1\}^\Lambda \to \{0,1\}^\Lambda$ defined as

$$\texttt{bran}_{ij}(x)(k) := \begin{cases} x(i) \oplus x(j) & \text{if } k = j, \\ x(k) & \text{otherwise,} \end{cases} \tag{6.14}$$

and blocking symbols ▬ still correspond to the death map $\texttt{death}_i$ as before. Other cancellative maps that we have already seen are represented as follows:



$\texttt{vot}_{1,2}$ $\texttt{ann}_{1,2}$ $\texttt{excl}_{1,2}$

Here $\texttt{ann}_{ij}$ is the annihilating random walk map. The maps $\texttt{vot}_{ij}$ and $\texttt{excl}_{ij}$ are both additive and cancellative, and represented in the same way as additive and cancellative maps. The following proposition is very similar to Proposition 6.1. Below, for any configuration $x \in \{0,1\}^\Lambda$, we let

$$|x| := \sum_{i\in\Lambda} x(i)$$

denote the number of ones. For any $x, y \in \{0,1\}^\Lambda$ such that either $|x| < \infty$ or $|y| < \infty$, we define

$$\langle\!\langle x, y\rangle\!\rangle := \bigoplus_{i\in\Lambda} x(i)y(i).$$

Proposition 6.3 (Cancellative systems duality) *For each cancellative local map* $m : \{0,1\}^\Lambda \to \{0,1\}^\Lambda$ *there exists a unique dual map* $m' : \{0,1\}^\Lambda \to \{0,1\}^\Lambda$ *such that*

$$\langle\!\langle m(x), y\rangle\!\rangle = \langle\!\langle x, m'(y)\rangle\!\rangle \tag{6.15}$$

for all $x, y \in \{0,1\}^\Lambda$ *such that* $|x| \wedge |y| < \infty$, *and this dual map* m' *is also a cancellative local map. Let* $\mathcal{G}$ *be a collection of additive local maps, let* $(r_m)_{m\in\mathcal{G}}$ *be nonnegative constants, and assume that of the generators*

$$\begin{aligned} Gf(x) &:= \sum_{m\in\mathcal{G}} r_m\big\{f\big(m(x)\big) - f\big(x\big)\big\},\\ G'f(y) &:= \sum_{m\in\mathcal{G}} r_m\big\{f\big(m'(y)\big) - f\big(y\big)\big\}, \end{aligned} \tag{6.16}$$

G *satisfies the summability condition (4.17). Let* $X = (X_t)_{t\geq 0}$ *and* $Y = (Y_t)_{t\geq 0}$ *be interacting particle systems with generators* G *and* G', *respectively, and assume that* $|Y_0| < \infty$ *is a.s. Then,* $|Y_t| < \infty$ *for all* $t \geq 0$ *a.s. and for each* $t > 0$, *it is possible to couple* X *and* Y *in such a way that for each* $s \in [0,t]$, *the processes* X_s *and* Y_{t-s} *are independent, and*

$$\langle\!\langle X_t, Y_0\rangle\!\rangle = \langle\!\langle X_s, Y_{t-s}\rangle\!\rangle = \langle\!\langle X_0, Y_t\rangle\!\rangle \quad \text{a.s.} \tag{6.17}$$

In particular, if X_t *is independent of* Y_0 *and* Y_t *is independent of* X_0, *then*

$$\mathbb{P}\big[\sum_{i\in\Lambda} X_t(i)Y_0(i) \text{ is odd}\big] = \mathbb{P}\big[\sum_{i\in\Lambda} X_0(i)Y_1(i) \text{ is odd}\big] \qquad (t \geq 0). \tag{6.18}$$

Proof The proof is almost identical to the proof of Proposition 6.1, where instead of (6.8) we now have that

$$\begin{aligned} &\langle\!\langle X_s, Y_{t-s}\rangle\!\rangle = 1\\ \Leftrightarrow\quad &\text{the number of open path between points } (i,0) \text{ and } (j,t),\\ &\text{such that } X_0(i) = 1 \text{ and } Y_0(i) = 1 \text{ is odd.} \end{aligned}$$

There is also a small technical complication: the cancellative duality function

$$\psi(x,y) := \langle\!\langle x,y\rangle\!\rangle = \bigoplus_{i\in\Lambda} x(i)y(i) \tag{6.19}$$

is not well-defined for all $x,y \in \{0,1\}^\Lambda$. To overcome this, we assume that $|Y_0| < \infty$ is a.s. Since open paths and paths of relevance are the same, Lemma 4.12 tells us that if the generator G of the forward process X satisfies the summability condition (4.17), then the dual process Y satisfies $|Y_t| < \infty$ a.s. for all $t \geq 0$. In fact, by Lemma 4.13, almost surely $|Y_t| < \infty$ for all $t \geq 0$ simultaneously. ∎

We note that if we known the expression in (6.18) for all finite initial states $Y_0 = y$, then this determines the law of X_t uniquely. Indeed:

Lemma 6.4 (Distribution determining functions) *The functions $\{f_y : y \in \{0,1\}^\Lambda,\ |y| < \infty\}$ with $f_y(x) := \langle\!\langle x,y\rangle\!\rangle$ are distribution determining.*

Proof We may equivalently show that the functions

$$g_y(x) := 1 - 2f_y(x) = (-1)^{\langle\!\langle x,y\rangle\!\rangle}$$

are distribution determining. Since $\langle\!\langle x,1_{\{i\}}\rangle\!\rangle = x(i)$, the class $\{g_y : |y| < \infty\}$ separates points, and since

$$g_y g_{y'} = g_{y\oplus y'},$$

this class is closed under products. The claim now follows from Lemma 4.24. ∎

Some models that a priori do not look like cancellative systems turn out to be representable in cancellative maps. An example is the Neuhauser-Pacala model, defined by its transition rates in (1.16). We define a *rebellious map* by

$$\texttt{rebel}_{ijk}(x)(l) := \begin{cases} x(i)\oplus x(j)\oplus x(k) & \text{if } l = k,\\ x(l) & \text{otherwise.}\end{cases} \tag{6.20}$$

In words, this says that $x(k)$ changes its state if $x(i) \neq x(j)$.

Exercise 6.5 *Show that the map $\texttt{rebel}_{ijk}$ is cancellative. Show that the generator of the Neuhauser-Pacala model defined in (1.16) can be represented as*

$$\begin{aligned} G_{\rm NP} f(x) =& \frac{\alpha}{|\mathcal{N}_i|}\sum_i\sum_{j\in\mathcal{N}_i}\big\{f\big(\texttt{vot}_{ji}(x)\big) - f\big(x\big)\big\}\\ &= \frac{1-\alpha}{|\mathcal{N}_i|^2}\sum_i\sum_{\substack{j,k\in\mathcal{N}_i\\ j\neq k}}\big\{f\big(\texttt{rebel}_{kji}(x)\big) - f\big(x\big)\big\}. \end{aligned}$$

Exercise 6.6 *In the* threshold voter model*, the site i changes its type $x(i)$ from 0 to 1 with rate one as long as at least one site in its neighborhood $\mathcal{N}_i$ has type 1, and likewise, i flips from 1 to 0 with rate one as long as at least one site in $\mathcal{N}_i$ has type 0. Show that the generator of the threshold voter model can be written as*

$$G_{\rm thresh} f(x) = 2^{-|\mathcal{N}_i|+1} \sum_i \sum_{\substack{\Delta\subset\mathcal{N}_i\cup\{i\} \\ |\Delta| \text{ is even}}} \big\{ f\big(m_{\Delta,i}(x)\big) - f\big(x\big)\big\},$$

where $m_{\Delta,i}$ is the cancellative map defined by

$$m_{\Delta,i}(x)(k) := \begin{cases} x(i) \oplus \bigoplus_{j\in\Delta} x(j) & \text{if } k = i, \\ x(k) & \text{otherwise.} \end{cases}$$

Exercise 6.7 *Show that the threshold voter model is monotone.*

6.4 Other dualities

The additive systems duality function (6.10) and cancellative systems duality function (6.19) are not the only choices of ψ that lead to useful dualities. For $q \in [-1, 1)$, consider the function

$$\psi_q(x, y) := \prod_{i\in\Lambda} q^{x(i)y(i)} = q^{\langle x,\, y\rangle} \qquad \big(x, y \in \{0, 1\}^\Lambda\big), \tag{6.21}$$

where we use the conventions

$$0^0 := 1 \quad \text{and} \quad \langle x, y\rangle := \sum_{i\in\Lambda} x(i)y(i).$$

and in (6.21), if $q = -1$, we assume in addition that $|x| < \infty$ or $|y| < \infty$ to ensure that the infinite product is well-defined. The usefulness of this duality function has been discovered by Lloyd and Sudbury [SL95, SL97, Sud00]. In particular,

$$\begin{aligned} \psi_0(x, y) &= 1_{\{x \wedge y = 0\}}, \\ \psi_{-1}(x, y) &= (-1)^{\langle\langle x,\, y\rangle\rangle}, \end{aligned}$$

so ψ_0 is the additive systems duality function (6.10) and ψ_{-1} is simple reparametrization of the cancellative systems duality function from (6.19).

It seems that for $q \neq 0, -1$, particle systems are never[6] *pathwise duals* in the sense of (6.13) with respect to ψ_q, but nevertheless there are many nontrivial examples of particle systems that are (plain) *dual* with respect to ψ_q in the sense of (6.11). If two particle systems are dual w.r.t. ψ_q, then we will say that they are *q-dual*. Although a lot of the known duals of particle systems are q-duals, occasionally different duality functions are used. Examples can be found in [SL95, SL97, Sud00, Swa13a].

To give an example of q-duality with $q \neq 0, -1$, consider an interacting partice system whose dynamics are a mixture of contact process and voter model dynamics, with generator of the form:

$$
\begin{aligned}
G_{\rm covo} f(x) := \lambda \sum_{(i,j)\in\mathcal{E}^d} & \big\{ f\big(\big(\mathtt{bra}_{ij}(x)\big) - f\big(x\big)\big\} \\
+ \sum_{i\in\mathbb{Z}^d} & \big\{ f\big(\big(\mathtt{death}_{i}(x)\big) - f\big(x\big)\big\} \\
+ \gamma \sum_{(i,j)\in\mathcal{E}^d} & \big\{ f\big(\big(\mathtt{vot}_{ij}(x)\big) - f\big(x\big)\big\} \qquad (x \in \{0,1\}^{\mathbb{Z}^d}).
\end{aligned}
\tag{6.22}
$$

Such systems are studied in [DLZ14], who are especially interested in the fast-voting limit $\gamma \to \infty$. The contact-voter model is additive (but not cancellative, because the branching map is not), and by results from Section 6.2 0-dual to a system with branching, death, and coalescing random walk dynamics. Perhaps surprisingly, it is also self-dual.

Proposition 6.8 (Self-duality of the contact-voter model) *The contact-voter model with generator as in (6.22) is q-dual to itself, with*

$$
q := \frac{\gamma}{\gamma + \lambda}.
$$

Proof We first show how the result follows directly from a general theorem of [Sud00], and then sketch the steps one would have to take to prove the result oneself.

The paper [Sud00] considers interacting particle systems on graphs where the configuration along each edge makes the following transitions with the

[6] Except some very trivial and pathological cases.

following rates:[7]

$$\begin{array}{rcccl}
\text{“annihilation”} & & 11 \mapsto 00 & & \text{at rate } a, \\
\text{“branching”} & 01 \mapsto 11 & \text{and} & 01 \mapsto 11 & \text{each at rate } b, \\
\text{“coalescence”} & 11 \mapsto 01 & \text{and} & 11 \mapsto 10 & \text{each at rate } c, \\
\text{“death”} & 01 \mapsto 00 & \text{and} & 10 \mapsto 00 & \text{each at rate } d, \\
\text{“exclusion”} & 01 \mapsto 10 & \text{and} & 10 \mapsto 01 & \text{each at rate } e.
\end{array}$$

In this notation, the model in (6.22) corresponds to

$$a = 0, \quad b = \lambda + \gamma, \quad c = 1, \quad d = 1 + \gamma, \quad e = 0.$$

Now [Sud00, Thm 1] says that provided $b \neq 0$, such a model is always self-dual, with parameter

$$q = \frac{d - a - c}{b}$$

Filling in the values of a, b, c, d, e yields $q = \gamma/(\gamma + \lambda)$.

If one wants to prove such a result oneself, then as a first step one needs to use Theorem 4.28 and Corollary 4.29 to reduce the problem to finite lattices Λ. Having reduced the problem to finite spaces, one wishes to show that

$$\sum_{x'} P_t(x, x')\psi_q(x', y) = \sum_{y'} \psi_q(x, y')P'_t(y, y'),$$

where $(P_t)_{t\geq 0}$ and $(P'_t)_{t\geq 0}$ denote the transition probabilities of the process and its dual (in this case, $P_t = P'_t$ since we are looking for a self-duality). Differentiating, this is equivalent to

$$\sum_{x'} G(x, x')\psi_q(x', y) = \sum_{y'} \psi_q(x, y')G'(y, y'),$$

which can also be written as

$$G\psi(\,\cdot\,, y)(x) = G'\psi(x, \,\cdot\,)(y), \tag{6.23}$$

i.e., letting the generator G of the original process act on the first variable of $\psi(x, y)$ yields the same as letting the generator G' of the dual process act on the second variable of $\psi(x, y)$. This part of the argument is quite general and can be used to prove dualities for all kind of Markov proceses and duality functions. To actually do the calculations when $G = G' = G_{\text{covo}}$ and $\psi = \psi_q$ is somewhat cumbersome, but straightforward. These calculations can be found in [Sud00] and also in [Swa13b]. ∎

[7] The meaning of the words “annihilation”, “branching”,... here is a bit different from the way we have used these words so far. In particular, the “death” rate d refers only to “deaths while the neighboring site is empty”, while “deaths while the neighboring site is occupied” are called “coalescence”.

6.5 Invariant laws of the voter model

By (6.9) and Proposition 6.1, the voter model X is dual, in the sense of additive systems duality, to a collection Y of coalescing random walks. Mainly since $|Y_t|$ is a nonincreasing function of t (i.e., the number of walkers can only decrease), it is much easier to work with this dual system than with the voter model itself, so duality is really the key to understanding the voter model.

Proposition 6.9 (Clustering in low dimensions) *Let X be a nearest-neighbor or range R voter model on $\mathbb{Z}^d$. Assume that $d = 1, 2$. Then, regardless of the initial law,*

$$\mathbb{P}[X_t(i) = X_t(j)] \underset{t\to\infty}{\longrightarrow} 1 \qquad \forall i, j \in \mathbb{Z}^d.$$

Moreover, the delta measures δ_0 and δ_1 on the constant configurations are the only extremal invariant laws.

Proof In the graphical representation of the voter model, for each $(i,t) \in \mathbb{Z}^d \times \mathbb{R}$ and $s \geq 0$, there is a unique site

$$j =: \xi_s^{(i,t)} \in \mathbb{Z}^d \text{ such that } (j, t-s) \rightsquigarrow (i,t).$$

Here $(\xi_s^{(i,t)})_{s\geq 0}$ is the path of a random walk starting at $\xi_0^{(i,t)} = i$ and "running downwards in the graphical representation". Two such random walks started from different space-time points (i,t) and (i',t') are independent up to the first time they meet, and coalesce as soon as they meet. Moreover, if $X_t = \mathbf{X}_{0,t}(X_0)$, then

$$X_t(i) = X_{t-s}(\xi_s^{(i,t)}) \qquad (0 \leq s \leq t),$$

i.e., $\xi_s^{(i,t)}$ traces back where the site i at time t got its type from.[8]

Since the difference $\xi_s^{(i,t)} - \xi_s^{(j,t)}$ of two such random walks is a random walk with absoption in the origin, and since random walk on $\mathbb{Z}^d$ in dimensions $d = 1, 2$ is recurrent, we observe that

$$\mathbb{P}[X_t(i) = X_t(j)] \geq \mathbb{P}[\xi_t^{(i,t)} = \xi_t^{(j,t)}] = \mathbb{P}[\xi_t^{(i,0)} = \xi_t^{(j,0)}] \underset{t\to\infty}{\longrightarrow} 1 \qquad \forall i, j \in \mathbb{Z}^d.$$

This clearly implies that all invariant laws must be concentrated on constant figurations, i.e., a general invariant law is of the form $p\delta_0 + (1-p)\delta_1$ with $p \in [0, 1]$. ∎

For product initial laws we can be more precise. Although we state the following theorem for two-type processes only, it is clear from the proof that the statement generalizes basically unchanged to multitype voter models.

[8] This construction works in fact generally for multitype voter models, where the local state space S can be any finite set, and which are in general of course not additive systems. For simplicity, we will focus on the two-type voter model here.

Theorem 6.10 (Process started in product law) *Let X be a nearest neighbor or range R voter model on $\mathbb{Z}^d$. Assume that the $(X_0(i))_{i\in\mathbb{Z}^d}$ are i.i.d. with intensity $\mathbb{P}[X_0(i)=1]=p\in[0,1]$. Then*

$$\mathbb{P}[X_t\in\cdot\,]\underset{t\to\infty}{\Longrightarrow}\nu_p, \tag{6.24}$$

where ν_p is an invariant law of the process. If $d=1,2$, then

$$\nu_p=(1-p)\delta_0+p\delta_1. \tag{6.25}$$

On the other hand, if $d\geq 3$ and $0<p<1$, then the measures ν_p are concentrated on configurations that are not constant.

Proof As in the proof of Proposition 6.9, let $(\xi^{(i,t)}_s)_{s\geq 0}$ be the backward random walk in the graphical representation starting at (i,t). Define a random equivalence relation $\sim$ on $\mathbb{Z}^d$ by

$$i\sim j \quad\text{iff}\quad \xi^{(i,0)}_s=\xi^{(j,0)}_s \text{ for some } s\geq 0.$$

We claim that if we color the equivalence classes of $\sim$ in an i.i.d. fashion such that each class gets the color 1 with probability p and the color 0 with probability $1-p$, then this defines an invariant law ν_p such that (6.24) holds. Since random walk in dimensions $d=1,2$ is recurrent, there is a.s. only one equivalence class, and $\nu_p=(1-p)\delta_0+p\delta_1$. On the other hand, since random walk in dimensions $d\geq 3$ is transient, there are a.s. infinitely many[9] equivalence classes and hence for $p\neq 0,1$ the measure ν_p is concentrated on configurations that are not constant.

To prove (6.24), we use coupling. Let $(\chi(i))_{i\in\mathbb{Z}^d}$ be i.i.d. $\{0,1\}$-valued with $\mathbb{P}[\chi(i)=1]=p$. For each $t\geq 0$, w define a random equivalence relation $\sim_t$ on $\mathbb{Z}^d$ by

$$i\sim_t j \quad\text{iff}\quad \xi^{(i,0)}_s=\xi^{(j,0)}_s \text{ for some } 0\leq s\leq t.$$

We enumerate the elements of $\mathbb{Z}^d$ in some arbitrary way and define

$$\tilde X_t(i):=\chi(j)\quad\text{where j is the smallest element of } \{k\in\mathbb{Z}^d: i\sim_t k\}. \tag{6.26}$$

Then $\tilde X_t$ is equally distributed with X_t and converges a.s. as $t\to\infty$ to a random variable with law ν_p. ∎

[9] Although this is intuitively plausible, it requires a bit of work to prove this. A quick proof, that however requires a bit of ergodic theory, is as follows: since Poisson point processes are spatially ergodic, and the number N of equivalence classes is a translation-invariant random variable, this random number N must in fact be a.s. constant. Since the probability that two paths coalesce tends to zero as the distance between their starting points tends to infinity, for each finite n we can find n starting points sufficientlty far from each other so that with positive probability, none of the paths started at these points coalesce. This implies that $\mathbb{P}[N\geq n]>0$ for each finite n and hence by the fact that N is a.s. constant $\mathbb{P}[N=\infty]=1$.

6.6 Homogeneous invariant laws

In the present section, we show how the self-duality of the contact process can be used to prove that for contact processes with some sort of translation invariant structure, the upper invariant law is the limit law started from any nontrivial translation invariant initial law, and we will show that this in turn implies that the function $\theta(\lambda)$ from (5.5) is continuous everywhere, except possibly at the critical point. The methods of the present section are not restricted to additive particle systems. Applications of the technique to cancellative systems can be found in [SS08, CP14]. Applications to systems whose duals are systems of interacting diffusion processes can be found in [AS04, AS09, AS12].

We start with a simpler observation, that has been anticipated before, and which says that the functions $\theta(\lambda)$ from (1.9) and (5.5) are the same.

Lemma 6.11 (The function theta) *Let X denote the contact process with infection rate λ on a graph Λ and let $\overline{\nu}$ denote its upper invariant law. Then*

$$\int \overline{\nu}(\mathrm{d}x)\, x(i) = \mathbb{P}^{1_{\{i\}}}[X_t \neq 0 \ \forall t \geq 0] \qquad (i \in \Lambda).$$

More generally, for any $y \in \{0,1\}^\Lambda$ such that $|y| < \infty$,

$$\int \overline{\nu}(\mathrm{d}x)\, 1_{\{x \wedge y \neq 0\}} = \mathbb{P}^y[X_t \neq 0 \ \forall t \geq 0].$$

Proof By (6.9) and Proposition 6.1, the contact process X is self-dual with respect to the additive systems duality function, i.e.,

$$\mathbb{P}^x[X_t \wedge y = 0] = \mathbb{P}^y[x \wedge X_t = 0] \qquad (t \geq 0).$$

In particular, setting $x = 1$, we see that

$$\begin{aligned}\int \overline{\nu}(\mathrm{d}x)\, 1_{\{x \wedge y \neq 0\}} &= \lim_{t\to\infty} \mathbb{P}^1[X_t \wedge y \neq 0] \\ &= \lim_{t\to\infty} \mathbb{P}^y[1 \wedge X_t \neq 0] = \mathbb{P}^y[X_t \neq 0 \ \forall t \geq 0].\end{aligned}$$

∎

In what follows, we will be interested in contact processes that have some sort of translation invariant structure. For simplicity, we will concentrate on processes on $\mathbb{Z}^d$ with a nearest-neighbor or range R graph structure, even though the arguments can be generalized to other graphs such as infinite regular trees.

We define translation operators $T_i : \{0,1\}^{\mathbb{Z}^d} \to \{0,1\}^{\mathbb{Z}^d}$ by

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

We say that a probability law μ on $\{0,1\}^{\mathbb{Z}^d}$ is *homogeneous* or *translation invariant* if $\mu \circ T_i^{-1} = \mu$ for all $i \in \mathbb{Z}^d$.

The main aim of the present section is to prove the following result, which is originally due to Harris [Har76]. We can think of this result as a sort of spatial analogue of the observation in Section 3.5 that for the mean-field contact process, solutions of the differential equation (3.18) started in any nonzero initial state converge to the upper fixed point.

Theorem 6.12 (Convergence to upper invariant law) *Let $(X_t)_{t\geq 0}$ be a contact process started in a homogeneus nontrivial initial law $\mathbb{P}[X_0 \in \cdot\,]$. Then*

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

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

We start with two preparatory lemmas. We will use the graphical representation of the contact process as an additive particle system (see Section 6.2) and use the shorthand

$$X_t^x := \mathbf{X}_{0,t}(x) \qquad \big(t \geq 0,\ x \in \{0,1\}^{\mathbb{Z}^d}\big),$$

where $(\mathbf{X}_{s,t})_{s\leq t}$ is the stochastic flow constructed from the graphical representation as in (6.2). We continue to use the notation $|x| := \sum_i x(i)$. We say that x is *finite* if $|x| < \infty$.

Lemma 6.13 (Extinction versus unbounded growth) *For each finite $x \in \{0,1\}^{\mathbb{Z}^d}$, one has*

$$X_t^x = 0 \textit{ for some } t \geq 0 \quad \textit{or} \quad |X_t^x| \underset{t\to\infty}{\longrightarrow} \infty \quad \text{a.s.} \tag{6.27}$$

Proof Define

$$\rho(x) := \mathbb{P}\big[X_t^x \neq 0\ \forall t \geq 0\big] \qquad (x \in \{0,1\}^{\mathbb{Z}^d},\ |x| < \infty).$$

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

$$|x| \leq N \quad \text{implies} \quad \rho(x) \leq 1-\varepsilon. \tag{6.28}$$

We first argue why it is plausible that this implies (6.27) and then give a rigorous proof. Imagine that $|X_t^x| \not\to \infty$. Then, in view of (6.28), 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

$$\mathcal{A}_x := \{X^x_t \neq 0 \ \forall t \geq 0\} \qquad (x \in \{0,1\}^{\mathbb{Z}^d},\ |x| < \infty).$$

denote the event that the process $(X^x_t)_{t\geq 0}$ 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

$$\rho(X^x_t) = \mathbb{P}\big[\mathcal{A}_x \,\big|\, \mathcal{F}_t\big] \underset{t\to\infty}{\longrightarrow} 1_{\mathcal{A}_x} \quad \text{a.s.}, \tag{6.29}$$

where we have used an elementary result from probability theory that says that if $\mathcal{F}_n$ is an increasing sequence of σ-fields and $\mathcal{F}_\infty = \sigma(\bigcup_n \mathcal{F}_n)$, then $\lim_n \mathbb{P}[x|\mathcal{F}_n] = \mathbb{P}[x|\mathcal{F}_\infty]$ a.s. for each measurable event x. (See [Loe63, § 29, Complement 10 (b)].) In view of (6.28), formula (6.29) implies (6.27). ∎

Lemma 6.14 (Nonzero intersection) *Let $(X_t)_{t\geq 0}$ be a contact process with a homogeneus nontrivial initial law $\mathbb{P}[X_0 \in \cdot\,]$. Then for each $s,\varepsilon > 0$ there exists an $N \geq 1$ such that for any $x \in \{0,1\}^{\mathbb{Z}^d}$*

$$|x| \geq N \quad \textit{implies} \quad \mathbb{P}\big[x \wedge X_s = 0\big] \leq \varepsilon.$$

Proof By duality,

$$\mathbb{P}\big[x \wedge X_s = 0\big] = \mathbb{P}\big[X^x_s \wedge X_0 = 0\big]$$

where X_0 is independent of the graphical representation used to define X^x_s. Set $\Lambda_M := \{-M,\ldots,M\}^d$. It is not hard to see that for each $x \in \{0,1\}^{\mathbb{Z}^d}$ with $|x| \geq N$ we can find an $x' \leq x$ with $|x'| \geq N/|\Lambda_M|$ such that the sets

$$\big\{i + \Lambda_M : x'(i) = 1\big\}$$

are disjoint, where 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

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

Then, using Hölder's inequality[10] in the inequality marked with an exclama-

[10] 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 $\|\prod_{i=1}^n f_i\|_1 \leq \prod_{i=1}^n \|f_i\|_n$.

tion mark, we have

$$\begin{aligned}
\mathbb{P}\big[X^x_s \wedge X_0 = 0\big] &= \int \mathbb{P}[X_0 \in \mathrm{d}y]\,\mathbb{P}\big[X^x_s \wedge y = 0\big] \\
&\leq \int \mathbb{P}[X_0 \in \mathrm{d}y]\,\mathbb{P}\big[\bigvee_{i:\,x'(i)=1} X^{\{i\}\,(M)}_s \wedge y = 0\big] \\
&= \int \mathbb{P}[X_0 \in \mathrm{d}y] \prod_{i:\,x'(i)=1} \mathbb{P}\big[X^{\{i\}\,(M)}_s \wedge y = 0\big] \\
&\stackrel{!}{\leq} \prod_{i:\,x'(i)=1} \Big(\int \mathbb{P}[X_0 \in \mathrm{d}y]\,\mathbb{P}\big[X^{\{i\}\,(M)}_s \wedge y = 0\big]^{|x'|}\Big)^{1/|x'|} \\
&= \prod_{i:\,x'(i)=1} \Big(\int \mathbb{P}[X_0 \in \mathrm{d}y]\,\mathbb{P}\big[X^{\{0\}\,(M)}_s \wedge y = 0\big]^{|x'|}\Big)^{1/|x'|} \\
&= \int \mathbb{P}[X_0 \in \mathrm{d}y]\,\mathbb{P}\big[X^{\{0\}\,(M)}_s \wedge y = 0\big]^{|x'|},
\end{aligned}$$

where we have used the homogeneity of $\mathbb{P}[X_0 \in \cdot\,]$ in the last but one equality. Our arguments so far show that $|x| \geq N$ implies that

$$\mathbb{P}\big[x \wedge X_s = 0\big] \leq \int \mathbb{P}[X_0 \in \mathrm{d}y]\mathbb{P}\big[X^{\{0\}\,(M)}_s \wedge y = 0\big]^{N/|\Lambda_M|} =: f(N, M).$$

Here, using the fact that

$$\mathbb{P}\big[X^{\{0\}\,(M)}_s \wedge y = 0\big] < 1 \quad \text{if } y(i) = 1 \text{ for some } i \in \Lambda_M,$$

we see that

$$\lim_{N\uparrow\infty} f(N, M) = \int \mathbb{P}[X_0 \in \mathrm{d}y] 1_{\{y(i)=0\ \ \forall i \in \Lambda_M\}} = \mathbb{P}[X_0(i) = 0\ \forall i \in \Lambda_M].$$

Since $\mathbb{P}[X_0 \in \cdot\,]$ is nontrivial, we have that

$$\lim_{M\uparrow\infty} \mathbb{P}[X_0(i) = 0\ \forall i \in \Lambda_M] = \mathbb{P}[X_0 = 0] = 0.$$

Together with our previous equation, this shows 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 6.15 *Show by counterexample that the statement of Lemma 6.14 is false for $s = 0$.*

Proof of Theorem 6.12 As in the proof of Lemma 6.13, we set

$$\rho(x) := \mathbb{P}\big[X^x_t \neq 0\ \forall t \geq 0\big] \qquad (x \in \{0,1\}^{\mathbb{Z}^d},\ |x| < \infty).$$

By Lemmas 4.25, 6.2, and 6.11, it suffices to show that

$$\lim_{t\to\infty} \mathbb{P}\big[x \wedge X_t \neq 0\big] = \rho(x)$$

for all finite $x \in \{0,1\}^{\mathbb{Z}^d}$. By duality, this is equivalent to showing that

$$\lim_{t\to\infty} \mathbb{P}\big[X^x_{t-s} \wedge X_s \neq 0\big] = \rho(x) \qquad \big(x \in \{0,1\}^{\mathbb{Z}^d},\ |x| < \infty\big),$$

where $(X^x_t)_{t\geq 0}$ and $(X_t)_{t\geq 0}$ are independent and $s > 0$ is some fixed constant. For each $\varepsilon > 0$, we can choose N as in Lemma 6.14, and write

$$\begin{aligned}\mathbb{P}\big[X^x_t \wedge X_s \neq 0\big] = &\,\mathbb{P}\big[X^x_t \wedge X_s \neq 0 \,\big|\, |X^x_t| = 0\big]\, \mathbb{P}\big[|X^x_t| = 0\big] \\ &+ \mathbb{P}\big[X^x_t \wedge X_s \neq 0 \,\big|\, 0 < |X^x_t| < N\big]\, \mathbb{P}\big[0 < |X^x_t| < N\big] \\ &+ \mathbb{P}\big[X^x_t \wedge X_s \neq 0 \,\big|\, |X^x_t| \geq N\big]\, \mathbb{P}\big[|X^x_t| \geq N\big].\end{aligned}$$

Here, by Lemma 6.13 and our choice of N,

$$\begin{array}{rl}\text{(i)} & \mathbb{P}\big[X^x_t \wedge X_s \neq 0 \,\big|\, |X^x_t| = 0\big] = 0, \\ \text{(ii)} & \displaystyle\lim_{t\to\infty} \mathbb{P}\big[0 < |X^x_t| < N\big] = 0, \\ \text{(iii)} & \displaystyle\liminf_{t\to\infty} \mathbb{P}\big[X^x_t \wedge X_s \neq 0 \,\big|\, |X^x_t| \geq N\big] \geq 1 - \varepsilon, \\ \text{(iv)} & \displaystyle\lim_{t\to\infty} \mathbb{P}\big[|X^x_t| \geq N\big] = \rho(x),\end{array}$$

from which we conclude that

$$(1-\varepsilon)\rho(x) \leq \liminf_{t\to\infty} \mathbb{P}\big[X^x_t \wedge X_s \neq 0\big] \leq \limsup_{t\to\infty} \mathbb{P}\big[X^x_t \wedge X_s \neq 0\big] \leq \rho(x).$$

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

Theorem 6.12 has a simple corollary.

Corollary 6.16 (Homogeneous invariant laws) *All homogeneous invariant laws of a contact process are convex combinations of $\delta_\emptyset$ and $\overline{\nu}$.*

Proof Let ν be any homogeneous invariant law. We will show that ν is a convex combination. of $\delta_\emptyset$ and $\overline{\nu}$. If $\nu = \delta_0$ we are done. Otherwise, as in the proof of Lemma 5.10, we can write $\nu = (1-p)\delta_0 + p\mu$ where $p \in (0,1]$ and μ is a nontrivial homogeneous invariant law. But now Theorem 6.12 implies that

$$\mu = \mu P_t \underset{t\to\infty}{\Longrightarrow} \overline{\nu},$$

so we conclude that $\mu = \overline{\nu}$. ∎

Recall from Exercise 5.14 that the function $\lambda \mapsto \theta(\lambda)$ from (5.5) is right-continuous everywhere. We let

$$\lambda_{\rm c} := \inf\{\lambda \in \mathbb{R} : \theta(\lambda) > 0\} \tag{6.30}$$

denote the *critical point* of the contact process. As an application of Theorem 6.12, we prove the following result.

Proposition 6.17 (Continuity above the critical point) *The function* $\lambda \mapsto \theta(\lambda)$ *is left-continuous on* $(\lambda_{\rm c}, \infty)$.

Proof Let $\overline{\nu}_\lambda$ denote the upper invariant law of the contact process with infection rate λ. Fix $\lambda > \lambda_{\rm c}$ and choose $\lambda_n \uparrow \lambda$. Since the space $\mathcal{M}_1(\{0,1\}^{\mathbb{Z}^d})$ of probability measures on $\{0,1\}^{\mathbb{Z}^d}$, equipped with the topology of weak convergence, is compact, it suffices to show that each subsequential limit ν_* of the measures $\overline{\nu}_{\lambda_n}$ equals $\overline{\nu}_\lambda$. By Proposition 4.30, each such subsequential ν_* limit is an invariant law. It clearly is also homogeneous. Since $\lambda > \lambda_{\rm c}$, by Lemma 5.10, the measures $\overline{\nu}_{\lambda_n}$ are nontrivial for n large enough, and hence, using also Proposition 5.11, the same is true for ν_*. By Corollary 6.16, we conclude that $\nu_* = \overline{\nu}$. This argument shows that the map

$$(\lambda_{\rm c}, \infty) \ni \lambda \mapsto \overline{\nu}_\lambda$$

is left-continuous w.r.t. the topology of weak convergence. Since $x \mapsto x(i)$ is a continuous function and $\theta(\lambda)$ is its expectation under $\overline{\nu}_\lambda$, the claim follows. ∎

6.7 Equality of critical points

The contact voter model X, that has a mixture of contact process and voter model dynamics, has been introduced in (6.22). It has two parameters: the infection rate λ and the voter rate γ. We say that X *survives* if

$$\mathbb{P}^{1_{\{0\}}}[X_t \neq 0 \ \forall t \geq 0] > 0.$$

For each $\gamma \geq 0$, we define critical infection rates $\lambda_{\rm c}(\gamma)$ and $\lambda'_{\rm c}(\gamma)$ by

$$\begin{aligned}
\lambda_{\rm c}(\gamma) &:= \inf\big\{\lambda \in \mathbb{R} : \text{the upper invariant law is nontrivial}\big\},\\
\lambda'_{\rm c}(\gamma) &:= \inf\big\{\lambda \in \mathbb{R} : \text{the process survives}\big\}.
\end{aligned}$$

The paper [DLZ14] studies the asymptotics of $\lambda_{\rm c}(\gamma)$ as $\gamma \to \infty$. Here, we will use duality to prove a more simple statement, namely, that $\lambda_{\rm c}(\gamma) = \lambda'_{\rm c}(\gamma)$ for all $\gamma \geq 0$.

For $\gamma = 0$ (i.e., the pure contact process), we already know this, as it is a direct consequence of Lemma 6.11, which follows from self-duality. We will use a similar argument here using Proposition 6.8, which says that the contact voter model is q-dual to itself, with $q = \gamma/(\gamma+\lambda)$. Note that if $\gamma = 0$ (the pure contact process), then $q = 0$ which corresponds to additive systems duality.

Proposition 6.18 (Characterization of the upper invariant law) *Let $q := \gamma/(\gamma+\lambda)$. The upper invariant law $\overline{\nu}$ of the contact voter model satisfies*

$$\int \overline{\nu}(\mathrm{d}x)\, q^{\langle x,\, y\rangle} = \mathbb{P}^y\big[X_t = 0 \textit{ for some } t \geq 0\big] \tag{6.31}$$

for all finite $y \in \{0,1\}^{\mathbb{Z}^d}$. In particular, $\lambda_{\rm c}(\gamma) = \lambda'_{\rm c}(\gamma)$ for all $\gamma \geq 0$.

Proof Letting X^1 and X^y denote the processes started in $X^1_0 = 1$ and $X^y_0 = y$, we observe that by Proposition 6.8,

$$\int \overline{\nu}(\mathrm{d}x)\, q^{\langle x,\, y\rangle} = \lim_{t\to\infty} \mathbb{E}\big[q^{\langle X^1_t,\, y\rangle}\big] = \lim_{t\to\infty} \mathbb{E}\big[q^{\langle 1,\, X^y_t\rangle}\big] = \lim_{t\to\infty} \mathbb{E}\big[q^{|X^y_t|}\big].$$

The proof of Lemma 6.13 carries over without a change to the contact voter model, so

$$X^y_t = 0 \text{ for some } t \geq 0 \quad \text{or} \quad |X^y_t| \underset{t\to\infty}{\longrightarrow} \infty \quad \text{a.s.}$$

Using this, we see that

$$\lim_{t\to\infty} \mathbb{E}\big[q^{|X^y_t|}\big] = \mathbb{P}^y\big[X_t = 0 \text{ for some } t \geq 0\big],$$

completing the proof of (6.31).

Inserting $y = 1_{\{0\}}$ into (6.31), we see that

$$\int \overline{\nu}(\mathrm{d}x)\, \big(1 - (1-q)x(i)\big) = \mathbb{P}^{1_{\{0\}}}\big[X_t = 0 \text{ for some } t \geq 0\big],$$

or equivalently, using the fact that $1 - q = \lambda/(\gamma+\lambda)$,

$$\frac{\lambda}{\gamma+\lambda} \int \overline{\nu}(\mathrm{d}x)\, x(i) = \mathbb{P}^{1_{\{0\}}}\big[X_t \neq 0\ \forall t \geq 0\big].$$

This shows that $\overline{\nu} = \delta_0$ if and only if the process survives. ∎

Chapter 7

Oriented percolation

7.1 Introduction

Although we have seen phase transitions in our simulations of interacting particle systems in Chapter 1, and we have seen how phase transitions are defined and can be calculated in the mean-field limit in Chapter 3, we have not yet proved the existence of a phase transition for any of the spatial models that we have seen so far.

In the present chapter, we fill this gap by proving that the contact process on $\mathbb{Z}^d$ undergoes a phase transition by showing that the critical point λ_c defined in (6.30) is nontrivial in the sense that $0 < \lambda_c < \infty$. Note that by Lemma 6.11,

$$\begin{aligned}\lambda_c &= \inf\{\lambda \in \mathbb{R} : \text{ the contact process survives}\}\\ &= \inf\{\lambda \in \mathbb{R} : \text{ the upper invariant law is nontrivial}\}.\end{aligned}$$

In Exercise 5.12, which is based on Theorem 4.22, we have already proved for the process that

$$\frac{1}{|\mathcal{N}_0|} \leq \lambda_c,$$

where $|\mathcal{N}_0| = 2d$ or $= (2R+1)^d - 1$ is the size of the neighborhood of the origin for the nearest-neighbor process and for the range R process, respectively. In view of this, it suffices to prove that $\lambda_c < \infty$. A simple comparison argument (Exercise 5.18) shows that if the nearest-neighbor one-dimensional contact process survives for some value of λ, then the same is true for the nearest-neighbor and range R processes in dimensions $d \geq 2$. Thus, it suffices to show that $\lambda_c < \infty$ for the nearest-neighbor process in dimension one.

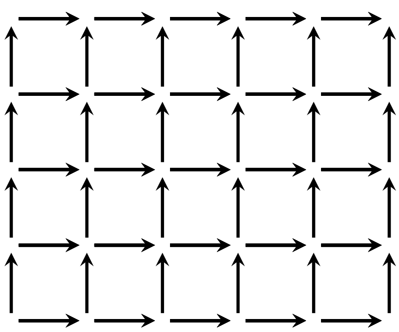
The method we will use is comparison with oriented percolation. This neither leads to a particularly short proof nor does it yield a very good upper bound on λ_c, but it has the advantage that it is a very robust method that can be applied to many other interacting particle systems. For example, in [SS08] and [SS15a], the method is applied to rebellious voter models and systems with cooperative branching and coalescencing random walk dynamics, respectively. An important paper for propagating the technique was [Dur91], where this was for the first time applied to non-monotone systems and it was shown that "basically, all one needs" to prove survival is that a particle system spreads into empty areas at a positive speed.

7.2 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

$$\mathcal{A} := \big\{(i,j) : i, j \in \mathbb{Z}^d,\ i \leq j,\ |i - j| = 1\big\}. \tag{7.1}$$

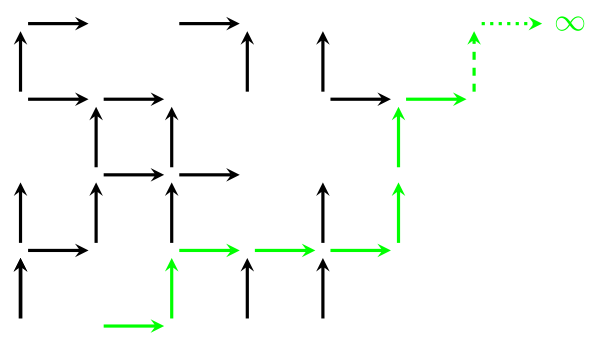
We view $\mathbb{Z}^d$ as an infinite directed graph, where elements $(i,j) \in \mathcal{A}$ represent arrows (or directed *bonds*) between neighbouring sites. Note that all arrows point 'upwards' in the sense of the natural order on $\mathbb{Z}^d$.



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 $\gamma(0) = i$, $\gamma(n) = j$, and

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

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 $0 \in \mathbb{Z}^d$.



Exercise 7.1 *Show that the number of vertices that can be reached by an open path from the origin is infinite if and only if there starts an infinite open path in the origin.*

Theorem 7.2 (Critical percolation parameter) *For oriented percolation in dimensions $d \geq 2$, there exists a critical parameter $p_c = p_c(d)$ such that $\mathbb{P}[0 \rightsquigarrow \infty] = 0$ for $p < p_c$ and $\mathbb{P}[0 \rightsquigarrow \infty] > 0$ for $p > p_c$. One has*

$$\frac{1}{d} \leq p_c(d) \leq \frac{8}{9}.$$

Proof Set

$$p_c := \inf\big\{p \in [0,1] : \mathbb{P}[0 \rightsquigarrow \infty] > 0\big\}.$$

A simple monotone coupling argument shows that $\mathbb{P}[0 \rightsquigarrow \infty] = 0$ for $p < p_c$ and $\mathbb{P}[0 \rightsquigarrow \infty] > 0$ for $p > p_c$.

To prove that $0 < p_c$, 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 $1/d \leq p_c(d)$.

To prove that $p_c(d) \leq 8/9$ 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 $p_c(1) = 1$.)

We will use a Peierls argument, named after R. Peierls who used a similar argument in 1936 for the Ising model [Pei36]. In Figure 7.1, we have drawn a piece of $\mathbb{Z}^2$ with a random collection of open arrows. Sites $i \in \mathbb{Z}^2$ such that $0 \rightsquigarrow i$ are drawn green. These sites are called *wet*. Consider the *dual lattice*

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

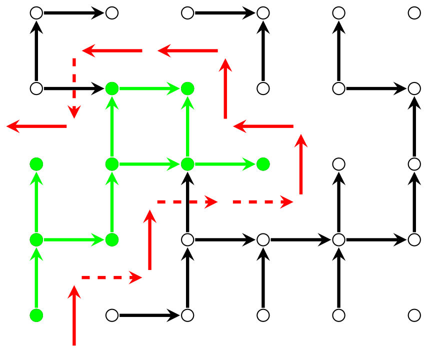


Figure 7.1: Peierls argument for oriented percolation. The green cluster of points reachable from the origin is surrounded by a red contour. The *north* and *west* steps of this contour cannot cross open arrows.

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 $(k+\frac{1}{2},-\frac{1}{2})$ and ending in some point $(-\frac{1}{2},m+\frac{1}{2})$ with $k,m\geq 0$, such that all sites immediately to the left of γ are wet, and no open arrows starting at these sites cross γ. In Figure 7.1, we have indicated γ with red 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 $(k+\frac{1}{2},-\frac{1}{2})$ $(k\geq 0)$, 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 (solid red arrows in the figure) are not crossed by an open arrow. Let M_n be the number of paths of length n with these properties. Since there are $n-1$ 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 less than $n3^n$ paths of length n with these properties. Since each path must make at least half of its steps in the north and west 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, we use a trick. (This part of the argument comes from [Dur88].) Set $D_m := \{0, \ldots, m\}^2$. Then, by the same arguments as before

$$\mathbb{P}[D_m \not\rightsquigarrow \infty] \leq \mathbb{P}\big[\sum_{n=2m}^{\infty} M_n \geq 1\big] \leq \mathbb{E}\big[\sum_{n=2m}^{\infty} M_n\big] \leq \sum_{n=2m}^{\infty} n 3^n (1-p)^{n/2},$$

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 $\mathbb{P}[0 \rightsquigarrow \infty] > 0$. ∎

7.3 Survival

The main result of the present chapter is the following theorem, which rigorously establishes the existence of a phase transition for the contact process on $\mathbb{Z}^d$.

Theorem 7.3 (Nontrivial critical point) *For the nearest-neighbor or range R contact process on $\mathbb{Z}^d$ $(d \geq 1)$, the critical infection rate satisfies $0 < \lambda_{\rm c} < \infty$.*

Proof As already mentioned in Section 7.1, the fact that $0 < \lambda_{\rm c}$ has already been proved in Exercise 5.12. By Exercise 5.18, to prove that $\lambda_{\rm c} < \infty$, it suffices to consider the one-dimensional nearest-neighbor case.

We will set up a comparison between the graphical representation of the one-dimensional nearest-neighbor contact process and oriented bond percolation on $\mathbb{Z}^2$; see Figure 7.2.

We fix $T > 0$ and define a map $\psi : \mathbb{Z}^2 \to \mathbb{Z} \times \mathbb{R}$ by

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

The points (κ_i, σ_i) with $i \in \mathbb{N}^2$ are indicated by open circles in Figure 7.2. As before, we make $\mathbb{Z}^2$ into an oriented graph by defining a collection of arrows $\mathcal{A}$ as in (7.1). 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

$$\tau_i^{\pm} := \inf\{t \geq \sigma_i : \text{at time } t \text{ there is an infection arrow from } \kappa_i \text{ to } \kappa_i \pm 1\}$$

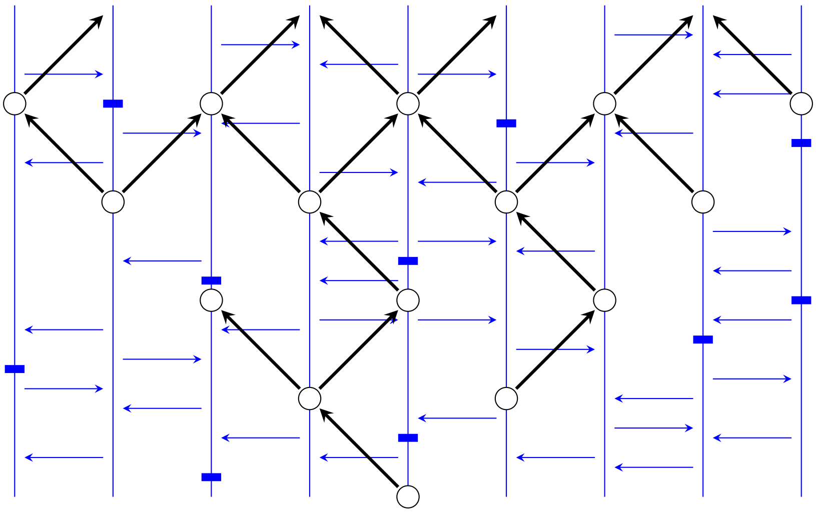


Figure 7.2: Comparison with oriented percolation. Good events in the graphical representation of the contact process (blue) correspond to open percolation arrows (black). An infinite open path along percolation arrows implies an infinite open paths in the graphical representation of the contact process.

denote the first time after σ_i that an arrow points out of κ_i to the left or right, respectively, and we define "good events"

$$\begin{aligned}\mathcal{G}_i^{\pm} := \big\{\, \tau_i^{\pm} < \sigma_i + T \text{ and there are no blocking symbols on}\\ \{\kappa_i\} \times (\sigma_i, \tau_i^{\pm}] \text{ and } \{\kappa_i \pm 1\} \times (\tau_i^{\pm}, \sigma_i + T]\big\}.\end{aligned}$$

Clearly,

$$\begin{aligned}&\mathcal{G}_i^{-} \quad \text{implies} \quad \psi(i_1, i_2) \rightsquigarrow \psi(i_1, i_2 + 1),\\ \text{and} \quad &\mathcal{G}_i^{-} \quad \text{implies} \quad \psi(i_1, i_2) \rightsquigarrow \psi(i_1 + 1, i_2).\end{aligned}$$

In view of this, we set

$$\omega_{((i_1,i_2),(i_1,i_2+1))} := 1_{\mathcal{G}_i^{-}} \quad \text{and} \quad \omega_{((i_1,i_2),(i_1+1,i_2))} := 1_{\mathcal{G}_i^{+}}.$$

Then the existence of an infinite open path in the oriented percolation model defined by the $(\omega_{(i,j)})_{(i,j)\in\mathcal{A}}$ implies the existence of an infinite open path in the graphical representation of the contact process, and hence survival of the latter.

We observe that

$$p := \mathbb{P}[\omega_{(i,j)} = 1] = \mathbb{P}(\mathcal{G}_i^{\pm}) = (1 - e^{-\lambda T})e^{-T} \qquad \big((i,j) \in \mathcal{A}\big), \tag{7.2}$$

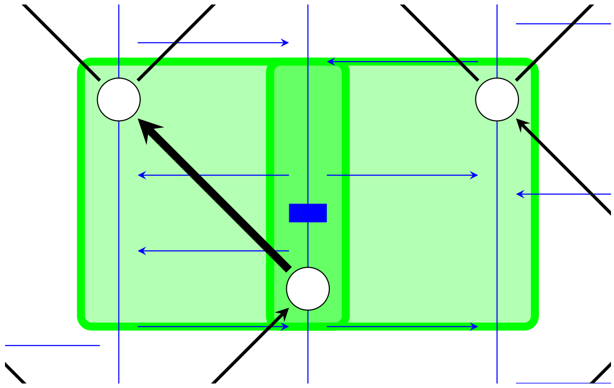


Figure 7.3: Good events use information from partially overlapping regions of space-time.

which tends to one as $\lambda \to \infty$ while $T \to 0$ in such a way that $\lambda T \to \infty$. It follows that 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, life is not quite so simple, since as shown in Figure 7.3, the good events $\mathcal{G}_i^{\pm}$ have been defined using information from partially overlapping space-time regions of the graphical representation of the contact process, and in view of this are not independent. They are, however, 3-dependent in the sense of Theorem 7.4 below, so by applying that result we can estimate the Bernoulli random variables $(\omega_{(i,j)})_{(i,j)\in\mathcal{A}}$ from below by i.i.d. 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. ∎

7.4 K-dependence

To finish the proof of Theorem 7.3 we need to provide the proof of Theorem 7.4 below, which states that k-dependent random variables with succes probability p can be estimated from below by i.i.d. random variables with a succes probability $\tilde{p}$ that tends to one as $p \to 1$.

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.

It is a bit unfortunate that the traditional definition of k-dependence is strictly tied to the integer lattice $\mathbb{Z}^d$, while the structure of $\mathbb{Z}^d$ has little to do with the essential idea. Therefore, in these lecture notes, we will deviate from tradition and replace(!) the definition above by the following definition.

Let Λ be countable and let $(X_i)_{i\in\Lambda}$ be a countable collection of randm variables. Then we will say that the $(X_i)_{i\in\Lambda}$ are K-*dependent* if for each $i \in \Lambda$ there exists a $\Delta_i \subset \Lambda$ with $i \in \Delta_i$ and $|\Delta_i| \leq K$, such that

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

Note that according to our new definition, 1-dependence means independence. The next theorem is taken from [Lig99, Thm B26], who in turn cites [LSS97].

Theorem 7.4 (K-dependence) *Let Λ be a countable set and let $p \in (0,1)$, $K < \infty$. Assume that $(\chi_i)_{i\in\Lambda}$ are K-dependent Bernoulli random variables with $P[\chi_i = 1] \geq p$ $(i \in \Lambda)$. 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*

$$P[\tilde{\chi}_i = 1] = \tilde{p} := \big(1 - (1-p)^{1/K}\big)^2 \qquad (i \in \Lambda), \tag{7.3}$$

in such a way that $\tilde{\chi}_i \leq \chi_i$ for all $i \in \Lambda$.

Proof 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

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

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,

$$P[\chi'_n = 1 \,|\, \chi'_0, \ldots, \chi'_{n-1}] \geq \tilde{p} \tag{7.4}$$

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 (7.3) 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 (7.4) holds. Set $p'_0 := P[\chi'_0 = 1]$ and

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

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

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

and define inductively

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

Then

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

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 (7.3) in such a way that $\tilde{\chi}_i \leq \chi'_i$.

What makes life complicated is that (7.4) 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

$$\begin{aligned} &P[\chi'_n = 1 \,|\, \chi'_0 = \varepsilon_0, \ldots, \chi'_{n-1} = \varepsilon_{n-1}] \\ &\quad = rP[\chi_n = 1 \,|\, \chi'_0 = \varepsilon_0, \ldots, \chi'_{n-1} = \varepsilon_{n-1}]. \end{aligned} \tag{7.5}$$

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

$$P[\chi_n = 0 \,|\, \chi'_0 = \varepsilon_0, \ldots, \chi'_{n-1} = \varepsilon_{n-1}] \leq 1 - r. \tag{7.6}$$

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

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

[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 $P[\chi_n = 1 | \chi_{n-1} = 0, \chi_{n-2} = 1] = 0$.

Then

$$\begin{aligned}
&P[\chi_n = 0 \,|\, \chi'_0 = \varepsilon_0, \ldots, \chi'_{n-1} = \varepsilon_{n-1}] \\
&\quad = P\big[\chi_n = 0 \,\big|\, \chi'_i = 0\ \forall i \in E_0,\ \chi_i = 1 = \psi_i\ \forall i \in E_1,\ \chi'_i = \varepsilon_i\ \forall i \in F\big] \\
&\quad = P\big[\chi_n = 0 \,\big|\, \chi'_i = 0\ \forall i \in E_0,\ \chi_i = 1\ \forall i \in E_1,\ \chi'_i = \varepsilon_i\ \forall i \in F\big] \\
&\quad = \frac{P\big[\chi_n = 0,\ \chi'_i = 0\ \forall i \in E_0,\ \chi_i = 1\ \forall i \in E_1,\ \chi'_i = \varepsilon_i\ \forall i \in F\big]}{P\big[\chi'_i = 0\ \forall i \in E_0,\ \chi_i = 1\ \forall i \in E_1,\ \chi'_i = \varepsilon_i\ \forall i \in F\big]} \\
&\quad \leq \frac{P\big[\chi_n = 0,\ \chi'_i = \varepsilon_i\ \forall i \in F\big]}{P\big[\psi_i = 0\ \forall i \in E_0,\ \chi_i = 1\ \forall i \in E_1,\ \chi'_i = \varepsilon_i\ \forall i \in F\big]} \\
&\quad = \frac{P\big[\chi_n = 0 \,\big|\, \chi'_i = \varepsilon_i\ \forall i \in F\big]}{P\big[\psi_i = 0\ \forall i \in E_0,\ \chi_i = 1\ \forall i \in E_1 \,\big|\, \chi'_i = \varepsilon_i\ \forall i \in F\big]} \\
&\quad \leq \frac{1-p}{(1-r)^{|E_0|} P\big[\chi_i = 1\ \forall i \in E_1 \,\big|\, \chi'_i = \varepsilon_i\ \forall i \in F\big]} \leq \frac{1-p}{(1-r)^{|E_0|}\, r^{|E_1|}},
\end{aligned} \tag{7.7}$$

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

$$P\big[\chi_i = 1\ \forall i \in E_1 \,\big|\, \chi'_i = \varepsilon_i\ \forall i \in F\big] \geq r^{|E_1|}. \tag{7.8}$$

We claim that (7.8) is a consequence of the induction hypothesis (7.6). Indeed, we may assume that the induction hypothesis (7.6) 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 (7.8) may be written as

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

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 (7.7) can be further estimated as

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

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

$$\frac{1-p}{(1-r)^{K-1}} \leq 1-r. \tag{7.9}$$

In particular, choosing $r = 1 - (1-p)^{1/K}$ yields equality in (7.9). Having proved (7.6), we see by (7.5) that (7.4) holds provided that we put $\tilde{p} := r^2$. ∎

Exercise 7.5 *Combine Theorem 7.2 and formulas (7.2) and (7.3) to derive an explicit upper bound on the critical infection rate $\lambda_{\rm c}$ of the one-dimensional contact process.*

Exercise 7.6 *The one-dimensional contact process with double deaths has been introduced just before Exercise 5.15. Use comparison with oriented percolation to prove that the one-dimensional contact process with double deaths survives with positive probability if its branching rate λ is large enough. When you apply Theorem 7.4, what value of K do you (at least) need to use?*

Exercise 7.7 *Use the previous exercise and Exercise 5.15 to conclude that for the cooperative branching process considered there, if λ is large enough, then:* 1° *If the process is started with at least two particles, then there is a positive probability that the number of particles will always be at least two.* 2° *The upper invariant law is nontrivial.*

Exercise 7.8 *Assume that there exists some $t > 0$ such that the contact process satisfies*

$$r := \mathbb{E}^{1_{\{0\}}}\big[|X_t|\big] < 1.$$

Show that this then implies that

$$\mathbb{E}^{1_{\{0\}}}\big[|X_{nt}|\big] \leq r^n \qquad (n \geq 0)$$

and the process started in any finite initial state dies out a.s. Can you use this to improve the lower bound $1/|\mathcal{N}_i| \leq \lambda_{\rm c}$ from Excercise 5.12, e.g., for the one-dimensional nearest-neighbor process?

Bibliography

[ADS15] M. Aizenman, H. Duminil-Copin, and V. Sidoravicius. Random currents and continuity of Ising model's spontaneous magnetization. *Commun. Math. Phys.* 334 (2015), 719–742.

[AS04] S.R. Athreya and J.M. Swart. Branching-coalescing particle systems. *Prob. Theory Relat. Fields* 131(3) (2005), 376–414.

[AS09] S.R. Athreya and J.M. Swart. Branching-coalescing particle systems. Erratum. *Prob. Theory Relat. Fields.* 145(3-4) (2009), 639–640.

[AS12] S.R. Athreya and J.M. Swart. Systems of branching, annihilating, and coalescing particles. *Electron. J. Probab.* 17 (2012), paper No. 80, 1–32.

[BD01] W.A. Brock and S.N. Durlauf. Discrete Choice with Social Interactions. *The Review of Economic Studies* 68(2) (2001), 235–260.

[BG90] C. Bezuidenhout and G. Grimmett. The critical contact process dies out. *Ann. Probab.* 18(4) (1990), 1462–1482.

[CP14] J.T. Cox and E.A. Perkins. A complete convergence theorem for voter model perturbations. *Ann. Appl. Probab.* 24(1) (2014), 150–197.

[CS73] P. Clifford and A. Sudbury. A model for spatial conflict. *Biometrika* 60(3) (1973), 581–588.

[DLZ14] R. Durrett, T. Liggett, and Y. Zhang. The contact process with fast voting. *Electron. J. Probab.* 19 (2014), Paper No. 28, 1–19.

[DN08] R.W.R. Darling and J.R. Norris Differential equation approximations for Markov chains *Probab. Surveys* 5 (2008), 37–79.

[Dur88] R. Durrett. *Lecture notes on particle systems and percolation.* Wadsworth & Brooks/Cole, Pacific Grove, 1988.

[Dur91] R. Durrett. A new method for proving the existence of phase transitions. *Spatial Stochastic Processes, Festschr. in Honor of Ted Harris 70th Birthday* (K.S. Alexander and J.C. Watkins eds.) 141–169. Prog. Probab. 19, Birkhäuser, Boston, 1991.

[Dur92] R. Durrett. Stochastic growth models: Bounds on critical values. *J. Appl. Probab.* 29(1) (1992), 11–20.

[EK86] S.N. Ethier and T.G. Kurtz. *Markov Processes: Characterization and Convergence.* John Wiley & Sons, New York, 1986.

[FM01] J.A. Fill and M. Machida. Stochastic monotonicity and realizable monotonicity. *Ann. Probab.* 29(2), (2001), 938–978.

[Har74] T.E. Harris. Contact interactions on a lattice. *Ann. Probab.* 2 (1974), 969–988.

[Har76] T.E. Harris. On a class of set-valued Markov processes. *Ann. Probab.* 4 (1976), 175–194.

[Isi25] E. Ising. Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik* 31 (1925), 253–258.

[Kal97] O. Kallenberg. *Foundations of modern probability.* Springer, New York, 1997.

[Kro99] S. Krone. The two-stage contact process. *Ann. Appl. Probab.* 9(2) (1999), 331–351.

[Lig85] T.M. Liggett. *Interacting Particle Systems.* Springer-Verlag, New York, 1985.

[Lig95] T.M. Liggett. Improved upper bounds for the contact process critical value. *Ann. Probab.* 23 (1995), 697–723.

[Lig99] T.M. Liggett. *Stochastic interacting systems: contact, voter and exclusion processes.* Springer-Verlag, Berlin, 1999.

[Lig10] T.M. Liggett. *Continuous time Markov processes. An introduction.* AMS, Providence, 2010.

[LSS97] T.M. Liggett, R.H. Schonmann and A.M. Stacey. Domination by product measures. *Ann. Probab.* 25(1) (1997), 71–95.

[Loe63] M. Loève. *Probability Theory, 3rd edition.* Van Nostrand, Princeton, 1963.

[Neu94] C. Neuhauser. A long range sexual reproduction process. *Stochastic Processes Appl.* 53(2) (1994), 193–220.

[Nob92] C. Noble. Equilibrium behavior of the sexual reproduction process with rapid diffusion. *Ann. Probab.* 20(2) (1992), 724–745.

[NP99] C. Neuhauser and S.W. Pacala. An explicitly spatial version of the Lotka-Volterra model with interspecific competition. *Ann. Appl. Probab.* 9(4) (1999), 1226–1259.

[Ons44] L. Onsager. Crystal statistics, I. A two-dimensional model with an order-disorder transition. *Phys. Rev.* 65 (1944), 117–149.

[Pei36] R. Peierls. On Ising's model of ferromagnetism. *Proc. Camebridge Philos. Soc.* 36 (1936), 477-481.

[Pre74] C.J. Preston A generalization of the FKG inequalities. *Comm. Math. Phys.* 36(3) (1974), 233-241.

[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) (1995), 1816–1830.

[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) (1997), 96–114.

[SS08] A. Sturm and J.M. Swart. Voter models with heterozygosity selection. *Ann. Appl. Probab.* 18(1) (2008), 59–99.

[SS15a] A. Sturm and J.M. Swart. A particle system with cooperative branching and coalescence. *Ann. Appl. Probab.* 25(3) (2015), 1616–1649.

[SS15b] A. Sturm and J.M. Swart. Monotone and additive Markov process duality. Preprint, 2015, ArXiv:1510.06284.

[Sud97] A. Sudbury. The convergence of the biased annihilating branching process and the double-flipping process in $\mathbb{Z}^d$. *Stoch. Proc. Appl.* 68 (1997), 25–264.

[Sud99] A. Sudbury. Hunting submartingales in the jumping voter model and the biased annihilating branching process. *Adv. Appl. Probab.* 31(3) (1999), 839–854.

[Sud00] A. Sudbury. Dual families of interacting particle systems on graphs. *J. Theor. Probab.* 13(3) (2000), 695–716.

[Swa13a] J.M. Swart. Noninvadability implies noncoexistence for a class of cancellative systems. *Electron. Commun. Probab.* 18 (2013), paper No. 38, 1–12.

[Swa13b] J.M. Swart. Duality and Intertwining of Markov Chains. Lecture notes for a course at the ALEA in Europe School, October 21–25 (2013) at CIRM, Luminy (Marseille). Available at `http://staff.utia.cas.cz/swart/posters_index.html`.

Index