A Course in Interacting Particle Systems

J.M. Swart

November 13, 2024

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 $\mathbb{Z}^d$. Situated on each point in this lattice, there is 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.

An attractive property of interacting particle systems is that they are easy to simulate on a computer.[1] 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 insufficiently 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

---

[1]To get started doing this yourself, look at my simulation library that is available from `http://staff.utia.cas.cz/swart/simulate.html`.

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. Lanchier's [Lan24] book makes an impressive attempt to give an overview of the modern literature on interacting particle systems in the life and social siences only, at the cost of not being able to give detailed proofs of all the known facts. Indeed, such is the sophistication of modern methods that it would be possible to write a book on each of the four classic models mentioned above alone.

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. I would like to thank Aernout van Enter for a number of corrections and suggestions that helped me improve the text from 2016. Sam Olesker-Taylor pointed out

some typoes in the version of 2020. Finally, Jan Niklas Latz helped correct many mistakes in the versions of 2021 and 2022.

These lecture notes have first been posted on the arXiv in 2017 and have been updated in 2020 and 2022. The update in 2020 corrected the proof of Lemma 4.25, which was wrong in the original version, and made some other minor corrections. In the update of 2022 several additions have been made. The main existence and uniqueness result in Chapter 4 has been formulated as a pathwise uniqueness result for solutions of the evolution equation (4.10). Chapter 6 has also been radically rewritten and in Chapter 1 a section has been added about periodic behavior. In the other chapters several smaller points have been clarified and some pictures have been added.

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. 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 can usually be written in the form

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

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

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

Then $(\mathbb{Z}^d, E^d)$ is the integer lattice equipped with the *nearest neighbor* graph structure and $(\mathbb{Z}^d, E_R^d)$ is the graph obtained by connecting all edges within $\|\cdot\|_\infty$-distance R with an edge. We let $\mathcal{E}^d$ and $\mathcal{E}_R^d$ 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 $(\mathbb{Z}^d, E_R^d)$ 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* $\mathtt{vot}_{ij} : S^\Lambda \to S^\Lambda$ is defined as

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

Applying $\mathtt{vot}_{ij}$ 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* on $\mathbb{Z}^d$ is the interacting particle system with generator

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

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 $\mathcal{E}_R^d$ 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 $\mathtt{vot}_{ij}$ 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. In this case, we see that even after a long time, there are still many different types near the origin.[2]

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 $S = \{0, 1\}$.

---

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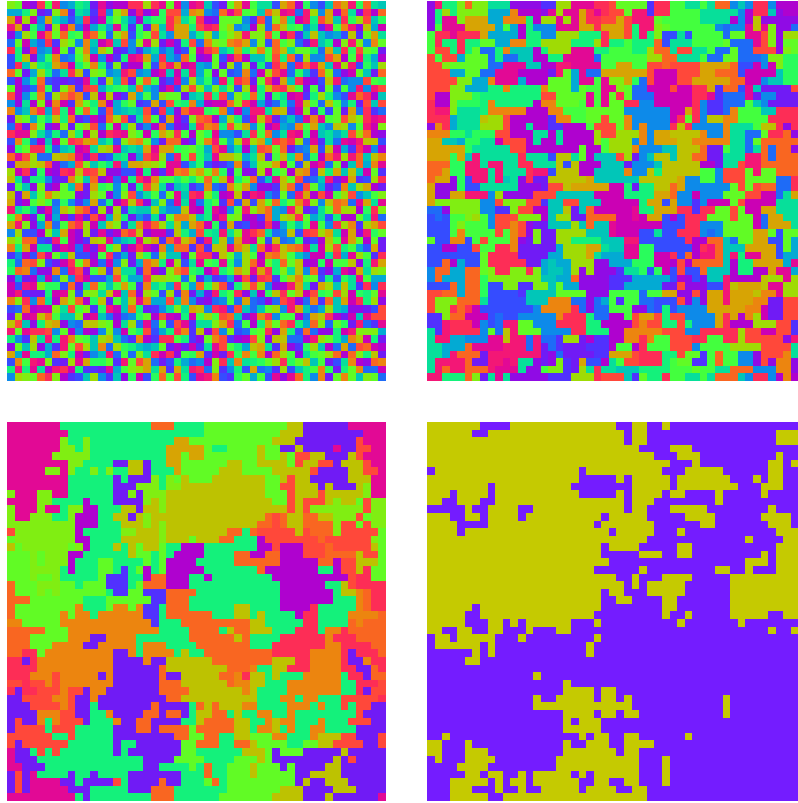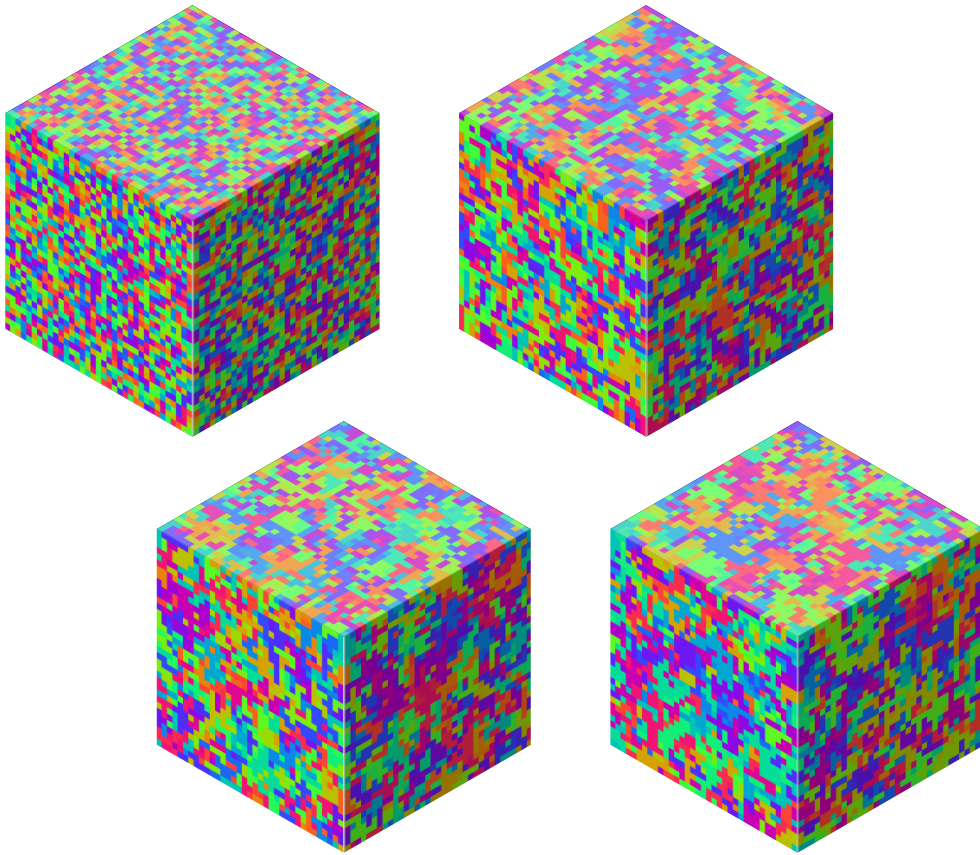
Figure 1.2: Four snapshots of a three-dimensional voter model with periodic boundary conditions. Initially, the types of sites are i.i.d. Time evolved in these pictures is 0, 4, 32, and 250.

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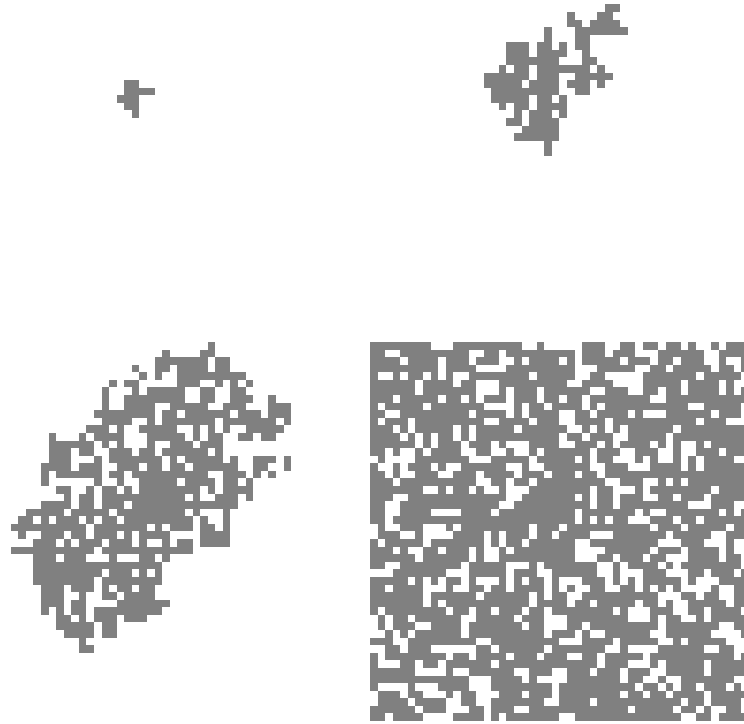
$\lambda \geq 0$ and *death rate* $\delta \geq 0$ is the interacting particle system with generator

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

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 epidemic.

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* $\lambda_c = \lambda_c(d)$ with $0 < \lambda_c < \infty$ such that $\theta(\lambda) = 0$ for $\lambda \leq \lambda_c$ and $\theta(\lambda) > 0$ for $\lambda > \lambda_c$. The function θ is continuous, strictly increasing and concave on $[\lambda_c, \infty)$ and satisfies $\lim_{\lambda\to\infty} \theta(\lambda) = 1$. One has

$$
\lambda_c(1) = 1.6489 \pm 0.0002. \tag{1.10}
$$

Proving these statements is not easy, however. For example, continuity of the function θ in the point λ_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 numerically observable.
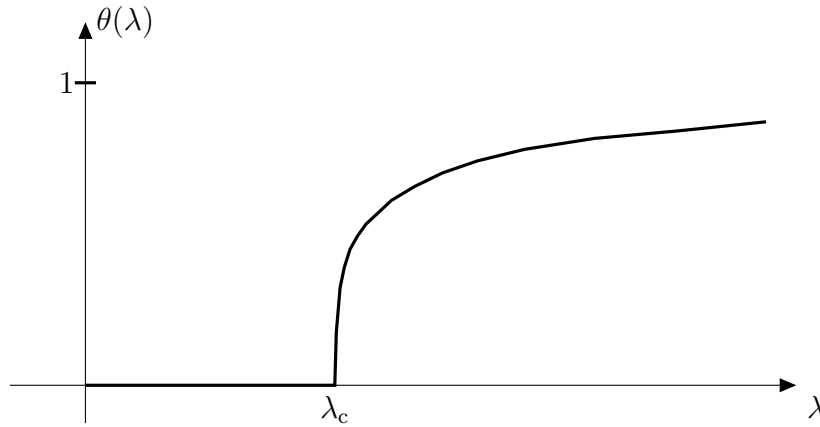
Figure 1.4: Survival probability of the one-dimensional contact process.

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].

Krone [Kro99] introduced a *two-stage contact process*. In this model, the local state space is $\{0, 1, 2\}$ where 0 represents an empty site, 1 a young organism, and 2 an adult organism. In a branching event, an adult organism produces a young organism on an empty neighboring site. In addition, young organisms can grow up. Both young and adults can die, the young possibly at a higher rate than the adults. The behavior of this model is similar to that of the contact process.

1.4   Ising and Potts models

In a *stochastic 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 *stochastic Potts models* where each "spin" can have $q \geq 2$ possible values. In this case, the local state space is traditionally denoted as $S = \{1, \ldots, q\}$,

---

[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].

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

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

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 value as their neighbors, depending on a parameter $\beta \in \mathbb{R}$. Adding a so-called *Glauber dynamics* to the model,[5] 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, the *stochastic Potts model* with *Glauber dynamics* is the interacting particle system that evolves in such a way that

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

More formally, we can write the generator as

$$G_{\text{Potts}} f(x) := \sum_{i \in \mathbb{Z}^d} \sum_{\sigma \in S} r_i^\sigma(x) \{ f(m_i^\sigma(x)) - f(x) \}, \qquad (1.13)$$

where $m_i^\sigma : S^\Lambda \to S^\Lambda$ are maps defined by

$$m_i^\sigma(x)(j) := \begin{cases} \sigma & \text{if } j = i, \\ x(j) & \text{otherwise.} \end{cases} \qquad (1.14)$$

The attentive reader may notice that the way we have written the generator in (1.13) is different from the way we have written our generators so far, since unlike the rates r_m in (1.1), the rates $r_i^\sigma(x)$ depend on the state x. This will be explained in more detail in Chapter 4. In particular, in Section 4.6, we will see that it is possible to rewrite the generator in (1.13) in a way that fits the general form (1.1) (with rates that do not depend on the state x) but for the Potts model, unlike the models we have seen so far, this way of writing the generator is less natural and more complicated.

---

[5]The terms *Ising model* and *Potts model* refer only to certain Gibbs measures. A *stochastic* Ising model or Potts model is any interacting particle system that has these Gibbs measures as its invariant laws (usually reversible). There exist several different ways to invent a dynamics with this property. This will be explained in a bit more detail in Section 5.4. In this section, we stick to Glauber dynamics.

Returning to our informal description in (1.12), we notice that for $\beta > 0$, 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].
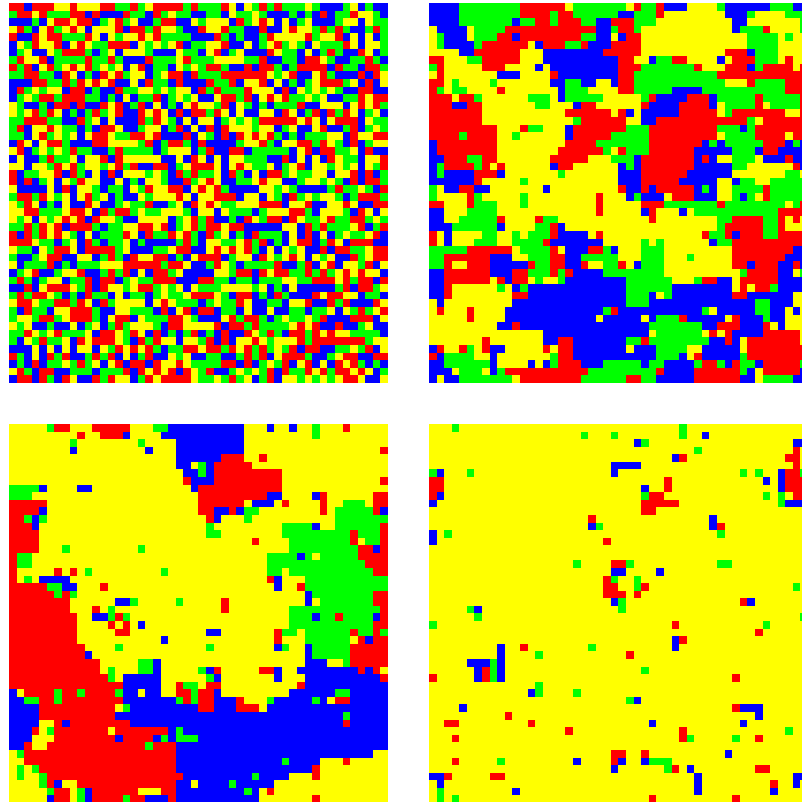
Figure 1.5: Four snapshots of a $q = 4$, $\beta = 1.2$ Potts model with Glauber dynamics and periodic boundary conditions. Initially, the types of sites are i.i.d. Time evolved in these pictures is 0, 4, 32, 500.

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 a 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 that clustering 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 $0 < \beta_c < \infty$ 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

$$
m_*(\beta) := \quad \text{the expectation of } x(0) \text{ with } +1 \text{ boundary} \\
\text{conditions, in the limit of large system size.} \tag{1.15}
$$

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 by Onsager [Ons44]. The formula is

$$
m_*(\beta) = \begin{cases} \left(1 - \sinh(\beta)^{-4}\right)^{1/8} & \text{for } \beta \geq \beta_c := \log(1 + \sqrt{2}), \\ 0 & \text{for } \beta \leq \beta_c. \end{cases} \tag{1.16}
$$

This function is plotted in Figure 1.6. In this case, the critical point β_c is known explicitly.

For Ising models in dimensions $d \geq 3$, the graph of $m_*(\beta)$ looks roughly similar to Figure 1.6, with $\beta_c \approx 0.442$ in dimension 3 [GPA01], 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 large.[6] 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

---

[6]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.
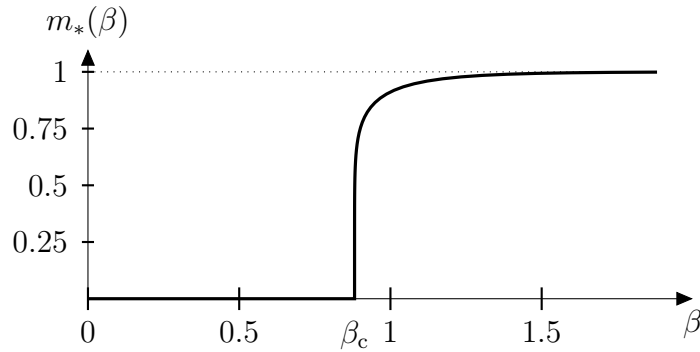
Figure 1.6: The spontaneous magnetization of the two-dimensional Ising model.

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.

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 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 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 law also appears. Potts models have q invariant laws (one corresponding to each color)
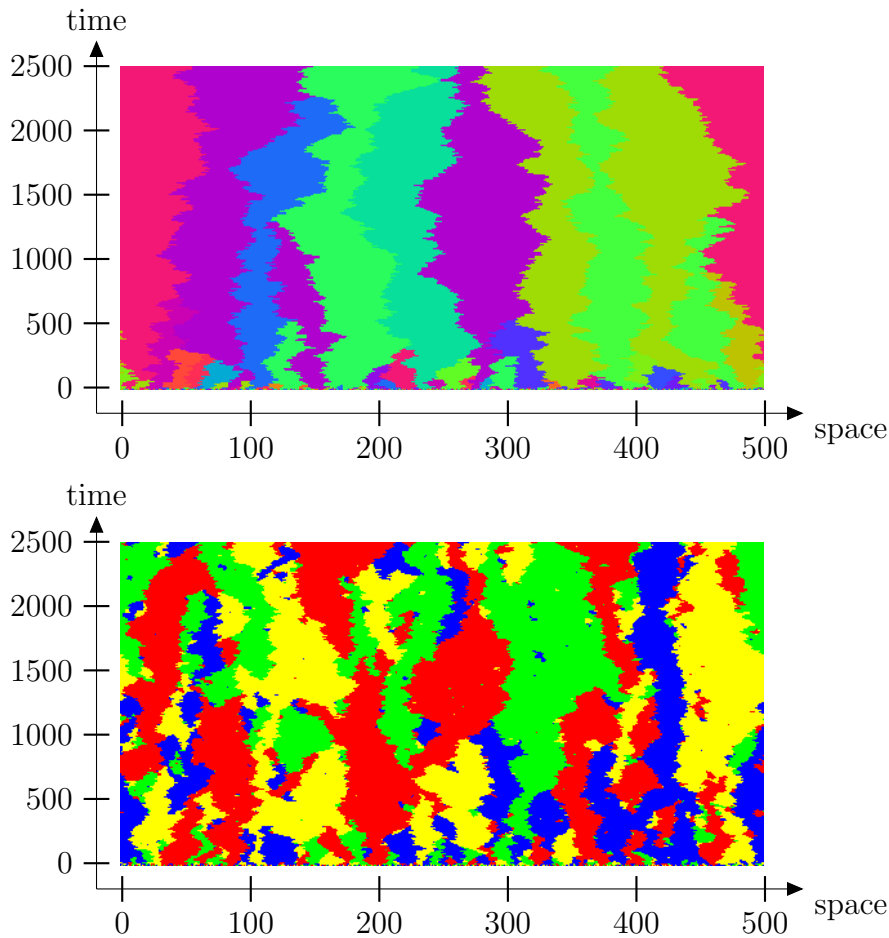
Figure 1.7: Time evolution of a one-dimensional voter model (above) and a one-dimensional Potts model (below) with a high value of β.

above the critical point.[7] Multiple invariant laws are a general phenomenon associated with phase transitions.

Phase transitions are classified into *first order* and *second order* phase transitions.[8] Second order phase transitions are also called *continuous* phase

_____

[7]More precisely, they have q invariant laws that have the additional property that they are also translation invariant in space. Depending on the dimension, there may exist additional invariant laws that are not translation invariant.

[8]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.

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 λ_c and β_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 β_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 analogous 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

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

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, it has numerically been observed that

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

with a critical exponent

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

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 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.[9] 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

_____

[9]Universality in the range R does not always hold. It has been proved that the $q = 3$ ferromagnetic Potts model in dimension two has a first order phase transition for large R [GB07], while the model with $R = 1$ is known to have a second order phase transition [DST17].

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 (that have a conformally invariant scaling limit that can be described using the Schramm-Loewner equation).

1.6 Variations on the voter model

Apart from the models discussed so far, lots of other interacting particle systems have been introduced and 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 sections, 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))

$$
\begin{aligned}
G_{\text{bias}} f(x) := & \frac{1}{2d} \sum_{(i,j) \in \mathcal{E}^d} \big\{ f\big(\texttt{vot}_{ij}(x)\big) - f\big(x\big) \big\} \\
& + \frac{s}{2d} \sum_{(i,j) \in \mathcal{E}^d} \big\{ f\big(\texttt{bra}_{ij}(x)\big) - f\big(x\big) \big\},
\end{aligned}
\tag{1.17}
$$

where $\texttt{vot}_{ij}$ and $\texttt{bra}_{ij}$ 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 or 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

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

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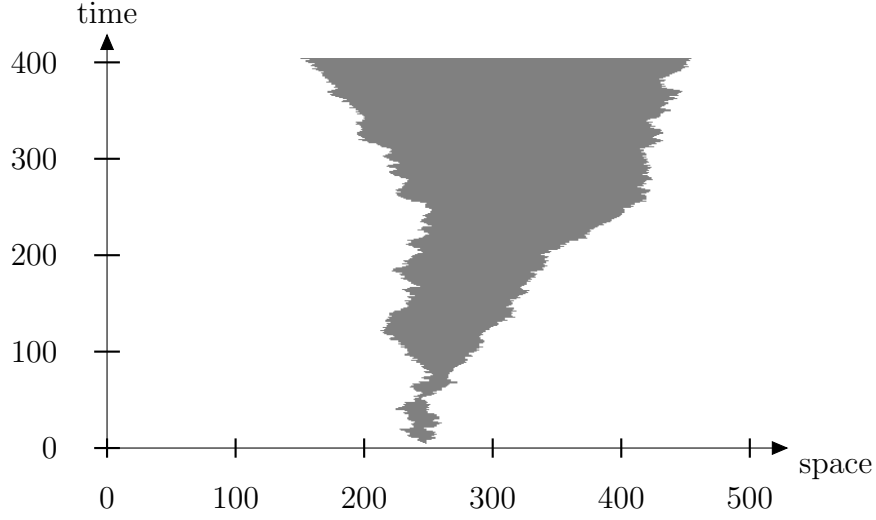
Figure 1.8: Time evolution of a one-dimensional biased voter model with bias $s = 0.2$.

following rates:

$$0 \mapsto 1 \qquad \text{with rate } f_1(x),$$
$$1 \mapsto 0 \qquad \text{with rate } f_0(x).$$

In the biased voter model, this is changed to

$$0 \mapsto 1 \qquad \text{with rate } (1+s)f_1(x),$$
$$1 \mapsto 0 \qquad \text{with rate } f_0(x).$$

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

$$0 \mapsto 1 \qquad \text{with rate } f_1(x)\big(f_0(x) + \alpha f_1(x)\big),$$
$$1 \mapsto 0 \qquad \text{with rate } f_0(x)\big(f_1(x) + \alpha f_0(x)\big), \qquad (1.18)$$

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

$$f_\tau(x) + \alpha f_{1-\tau}(x), \qquad (1.19)$$

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 (1.19) 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 (1.19) as saying that persons change their mind *less* often if they disagree with a lot of neighbors, i.e., the model in (1.18) has "rebellious" behavior.

Numerical simulations, shown in Figure 1.9, suggest that in one dimension and for ranges $R \geq 2$, the model in (1.18) 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   Branching and coalescing particles

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

$$\mathtt{rw}_{ij}(x)(k) := \begin{cases} 0 & \text{if } k = i, \\ x(i) \vee x(j) & \text{if } k = j, \\ x(k) & \text{otherwise.} \end{cases} \qquad (1.20)$$

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

$$G_{\mathrm{rw}} f(x) = \frac{1}{|\mathcal{N}_0|} \sum_{(i,j) \in \mathcal{E}^d} \left\{ f\big(\mathtt{rw}_{ij}(x)\big) - f(x) \right\} \qquad (x \in \{0,1\}^{\mathbb{Z}^d}) \qquad (1.21)$$

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

$$\mathtt{arw}_{ij}(x)(k) := \begin{cases} 0 & \text{if } k = i, \\ x(i) + x(j) \mod(2) & \text{if } k = j, \\ x(k) & \text{otherwise,} \end{cases} \qquad (1.22)$$

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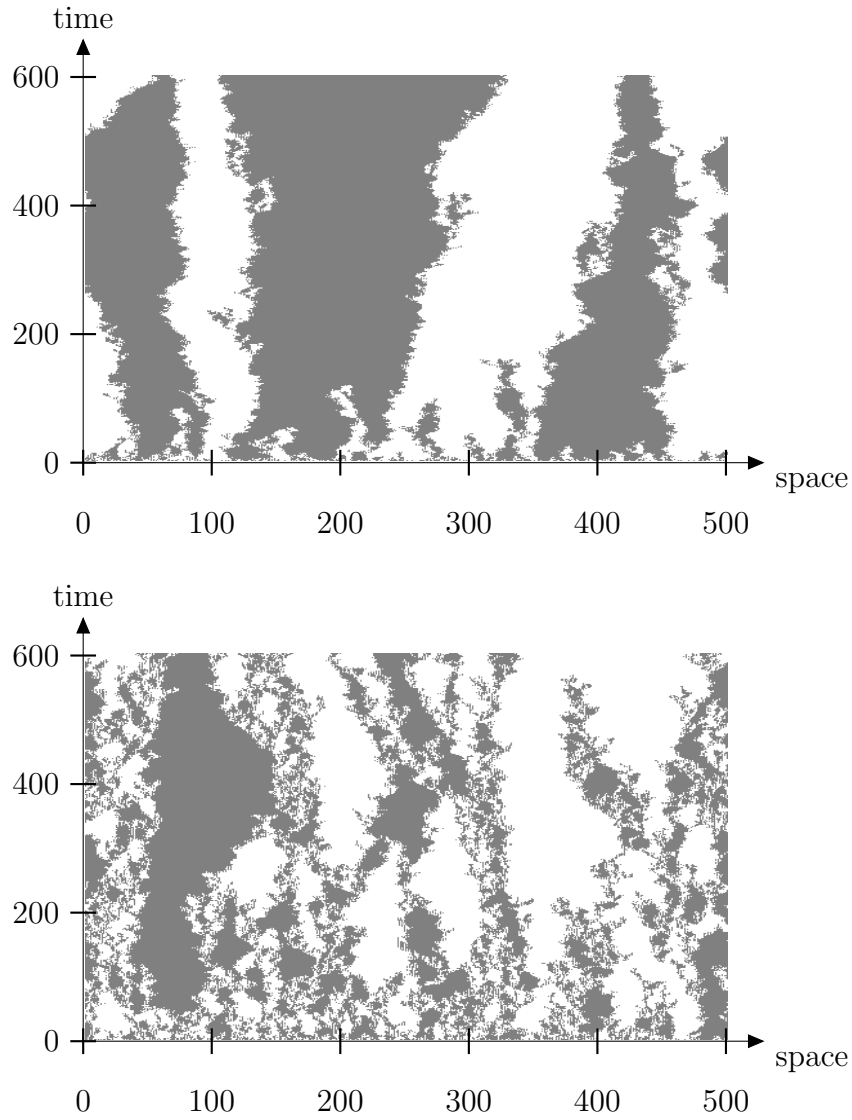
Figure 1.9: Evolution of "rebellious" voter models with $\alpha = 0.8$ and $\alpha = 0.3$, respectively.

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

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

Applying $\texttt{excl}_{ij}$ to a configuration x has the effect of interchanging the types
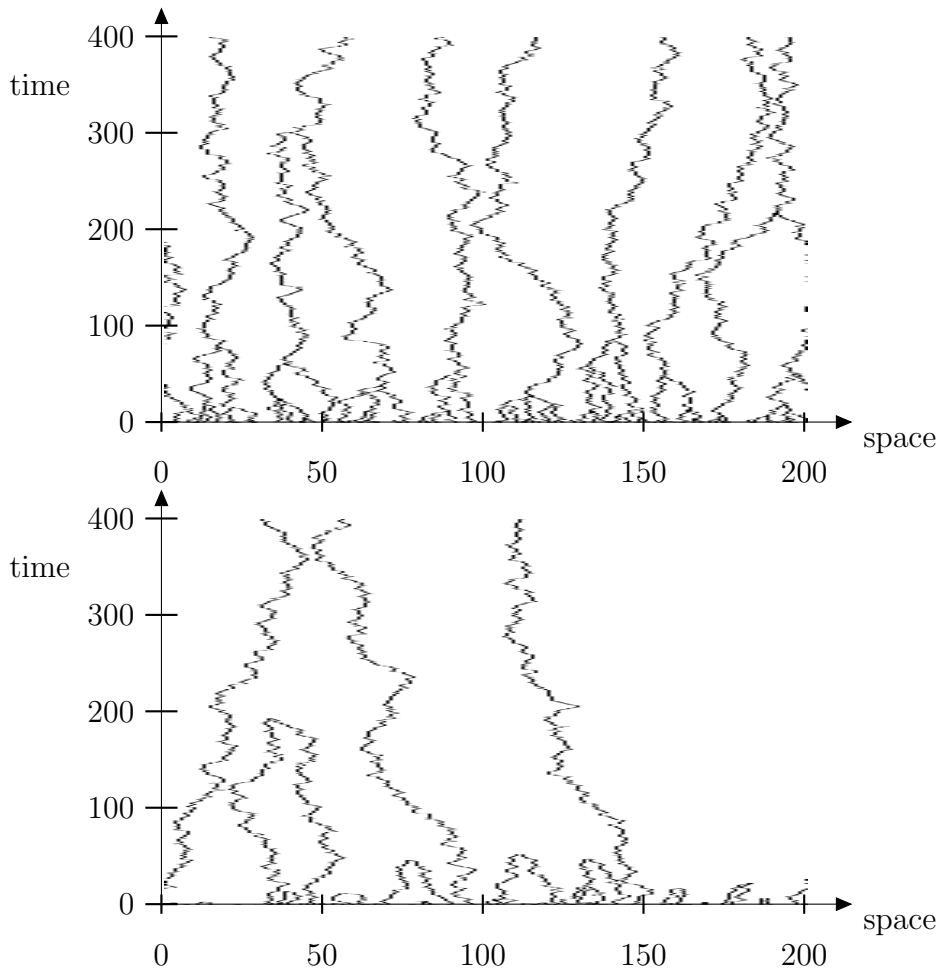
Figure 1.10: Systems of coalescing random walks (above) and annihilating random walks (below).

of j and j. The interacting particle system with state space $\{0,1\}^{\mathbb{Z}^d}$ and generator

$$G_{\mathrm{excl}}f(x) = \frac{1}{|\mathcal{N}_0|} \sum_{(i,j)\in\mathcal{E}^d} \{f(\mathtt{excl}_{ij}(x)) - f(x)\} \qquad (x \in \{0,1\}^{\mathbb{Z}^d}) \quad (1.24)$$

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 ran-

dom 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*

$$\texttt{coop}_{ijk}(x)(l) := \begin{cases} 1 & \text{if } l = k, \ x(i) = 1, \ x(j) = 1, \\ x(l) & \text{otherwise.} \end{cases} \tag{1.25}$$

See Figure 1.11 for a one-dimensional interacting particle system involving cooperative branching and coalescing random walks.
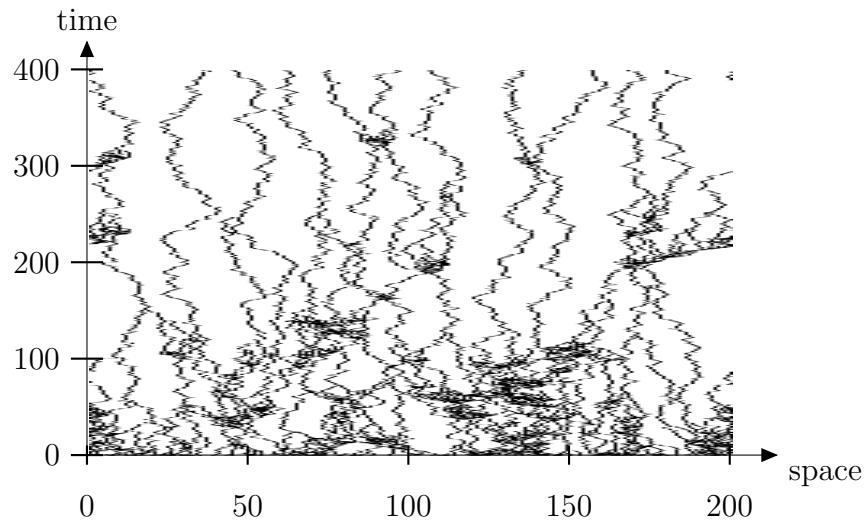
Figure 1.11: A one-dimensional interacting particle system with cooperative branching and coalescing random walk dynamics.

We define a *killing map* by

$$\texttt{kill}_{ij}(x)(k) := \begin{cases} 0 & \text{if } k = j, \ x(i) = 1, \ x(j) = 1, \\ x(k) & \text{otherwise.} \end{cases} \tag{1.26}$$

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

$$
\begin{aligned}
G_{\mathrm{babp}} f(x) := \lambda \sum_{(i,j)\in\mathcal{E}^1} & \left\{ f\big(\mathtt{bra}_{ij}(x)\big) - f\big(x\big) \right\} \\
+ \sum_{(i,j)\in\mathcal{E}^1} & \left\{ f\big(\mathtt{kill}_{ij}(x)\big) - f\big(x\big) \right\} \qquad (x \in \{0,1\}^{\mathbb{Z}}).
\end{aligned}
\tag{1.27}
$$

In the physics literature, this model is known as the Fredrickson-Andersen one spin facilitated model, see formula (26) in [RS03] (with $f = 1$). It is part of the class of *kinetically constrained models*. In the mathematical literature on this subject, a slight variant of the model has been studied [BDT19]. Figure 1.12 shows a simulation of such a system when $\lambda = 0.2$. When λ is small, in the simulations, the process seems to behave similar to systems of branching and coalescing random walks.
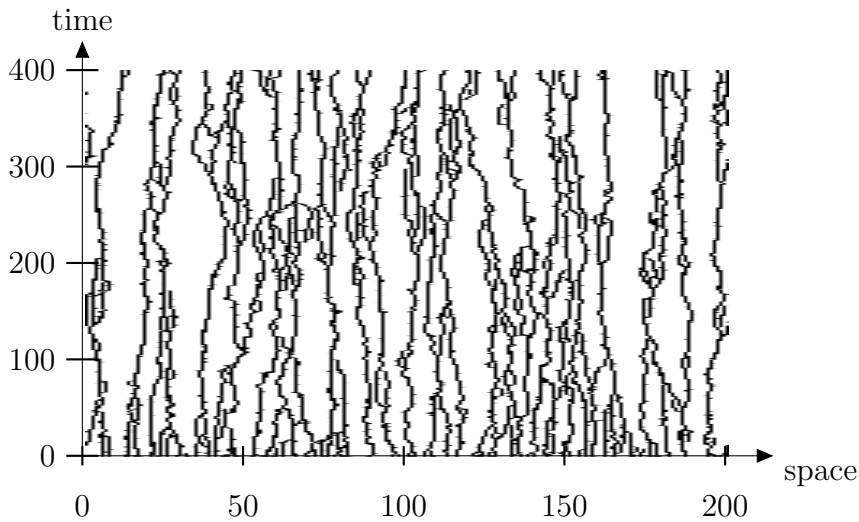
Figure 1.12: A system with branching and killing.

1.8 Periodic behavior

An invariant law of an interacting particle system is a probability distribution on the space S^{Λ} of all possible configurations with the property that if the system at time zero is distributed according to this law, then at all later times it is also distributed according to this law. Invariant laws need not be unique. For example, Potts models above the critical point have q different

invariant laws, that are characterized by the color that occupies the majority of the sites.

For all the interacting particle systems and initial states we have considered so far, the system has the property that as time tends to infinity, the distribution of the system converges to an invariant law. This need not always be the case. Perhaps the simplest way in which this can fail is if the system has a *periodic law*, i.e., a law that has the property that if the system at time zero is distributed according to this law, then it returns to this law after a finite time $T > 0$ (the *period*), but the system has a different distribution at all intermediate times $0 < t < T$.

Very little is known rigorously about interacting particle systems with periodic laws. Jahnel and Külske [JK14a] have constructed a three dimensional interacting particle system that has a periodic law. A general result due to Mountford [Mou95] implies that one dimensional systems with finite range interactings cannot have periodic laws. Beyond this, very little is known for spatial models, although there are some studies of periodic behavior in the mean-field limit (see Chapter 3). In particular, it is not known whether periodic laws are possible in two dimensions. The construction in [JK14a] is rather abstract since they do not write down the dynamics of their system explictly but only prove that such a system exists. Their system also does not have finite range interactings, although the strength of the interaction decays exponentially in the distance, which is almost as good.

Numerical simulations suggest that periodic behavior is not a rare phenomenon. Several interacting particle systems with explicit dynamics are known to exhibit periodic behavior in simulations. All known examples seem to work only in dimensions three and higher, however, which suggests that, perhaps, periodic behavior is not possible in two dimensions.

In the *mutually metastable Ising model*, based on the work of Collet, Formentin and Tovazzi [CFT16], the local state space is $S = \{-1, +1\}^2$ and the lattice is $\Lambda = \mathbb{Z}^3$. We denote an element of S^Λ as $x = (x(i))_{i \in \Lambda}$ where $x(i) = (x^1(i), x^2(i))$ with $x^1(i), x^2(i) \in \{-1, +1\}$ $(i \in \mathbb{Z}^3)$. We let

$$M_i^1(x) := \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} x^1(j) \quad \text{and} \quad M_i^2(x) := \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} x^2(j)$$

denote the average values of $x^1(j)$ and $x^2(j)$ among the neighbors of a site i. The *mutually metastable Ising model* with parameters $\alpha, \beta \geq 0$ is the interacting particle sysytem with state space S^Λ that evolves in such a way that

$$x^1(i) \text{ flips its the value with rate} \quad e^{-x^1(i)[\beta M_i^1(x) + \alpha x^2(i)]},$$

$$x^2(i) \text{ flips its the value with rate} \quad e^{-x^2(i)[\beta M_i^2(x) - \alpha x^1(i)]}.$$
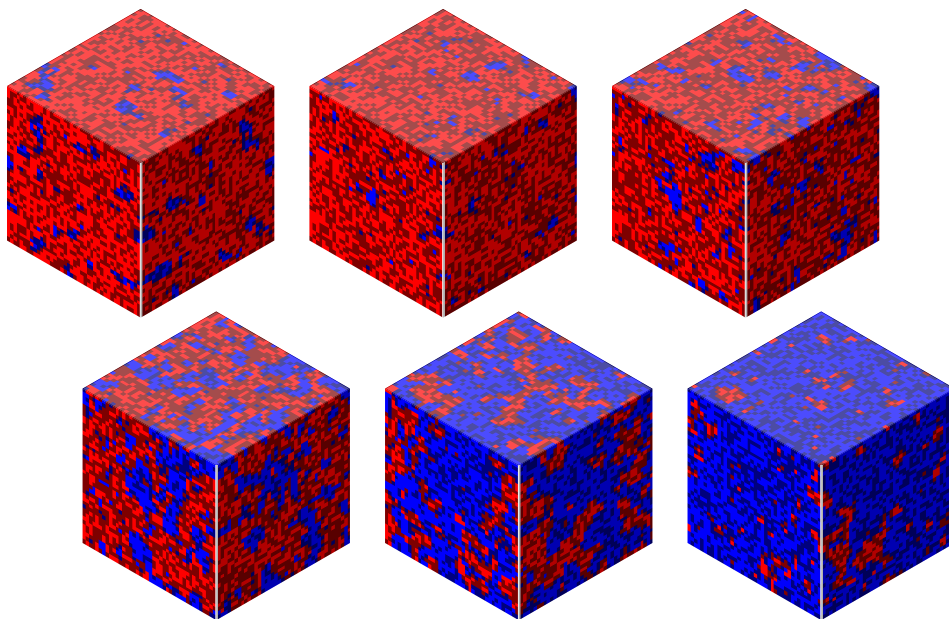
Figure 1.13: Periodic behavior of the dissipative Ising model with $\alpha = 0.2$ and $\beta = 8$ on a cube with side length 40 and periodic boundary conditions. Colors indicate the state of $x^1(i)$. Good sites have bright colors and corrupted sites have dark colors. Shown is the transition from a predominantly minus state to a predominantly plus state.

When $\alpha = 0$, there is no interaction between the two spins at a given site. In this case $X_t^1 := (X_t^1(i))_{i \in \Lambda}$ and $X_t^2 := (X_t^1(i))_{i \in \Lambda}$ evolve as independent stochastic Ising models.[10] In particular, the effect of the term $-\beta x^1(i) M_i^1(x)$ in the exponent is that the spin $x^1(i)$ does not change its value often if it agrees with most of its neighbors (so that $x^1(i) M_i^1(x) > 0$), and on the other hand changes its value with a high rate if it disagrees with most of its neighbors (so that $x^1(i) M_i^1(x) < 0$). Similarly, the spin $x^2(i)$ has a tendency to align with the majority of its neighbors.

When $\alpha > 0$, then in addition, the spin $x^1(i)$ tries to align with $x^2(i)$, but on the other hand, the spin $x^2(i)$ prefers to have a different value from $x^1(i)$. When most of the spins $x^1(i)$ have the value $+1$ and most of the spins $x^2(i)$ also have the value $+1$, then the spins $x^1(i)$ are perfectly satisfied, because they agree with most of their neighbors and also with the other spins $x^2(i)$.

---

[10]The parameter β here is defined differently from the parameter β in Section 1.4. Denoting the latter by β', one has $\frac{1}{2}\beta' = \frac{1}{|\mathcal{N}_i|}\beta$. Also, the dynamics here are different from the Glauber dynamics of Section 1.4.
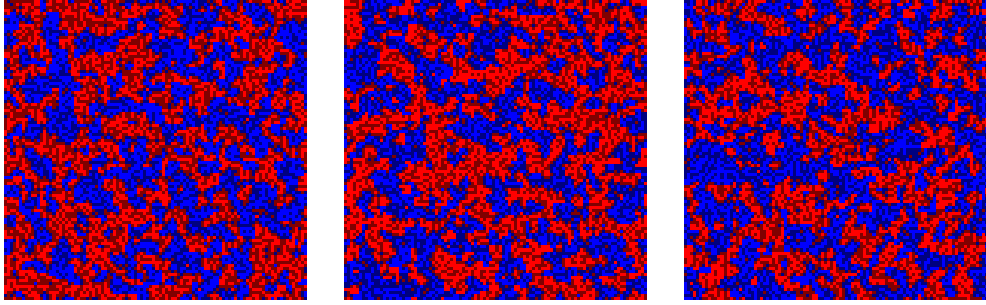
Figure 1.14: Three snapshots of a stationary dissipative Ising model with $\alpha = 0.2$ and $\beta = 8$ on a square with side length 100 and periodic boundary conditions. Although the system locally behaves in a periodic way sites do not coordinate their action over longer distances.

But the spins $x^2(i)$ are not completely satisfied. From their point of view, it would be ideal if most of the spins $x^1(i)$ have the value $+1$ but most of the spins $x^2(i)$ have the value -1. Numerical simulations suggest that the result of this is, for suitable choices of the parameters, that the systems starts to cycle, in a coordinated way, through the states

$$(+1, +1) \mapsto (+1, -1) \mapsto (-1, -1) \mapsto (-1, +1) \mapsto (+1, +1) \mapsto \cdots .$$

However, this sort of coordinated behavior seems to happen only in dimensions three and more. In two dimensions, even though individual sites still tend to pass through the states $(+1, +1) \mapsto (+1, -1) \mapsto (-1, -1) \mapsto (-1, +1) \mapsto$ in this order, the sites at larger distances from each other no longer coordinate their behavior.

In the *dissipative Ising model*, based on a model described in [DFR13, Tov19, CD+20], the local state space is again $S = \{-1, +1\}^2$. There are again two model parameters $\alpha, \beta \geq 0$ but the dynamics are different. Letting $M_i(x) := (M_i^1(x) + M_i^2(x))/2$ denote the average value of $x^1(j)$ and $x^2(j)$ for all sites j neighboring the site i, the system evolves in such a way that

$$x^1(i) \text{ flips its the value with rate } \quad e^{-\beta x^1(i) M_i(x)},$$
$$x^2(i) \text{ flips its the value with rate } \quad \alpha 1_{\{x^1(i) = x^2(i)\}}.$$

We call sites where $x^1(i) = x^2(i)$ *good* sites with spin $x^1(i)$ and sites where $x^1(i) \neq x^2(i)$ *corrupted* sites with spin $x^1(i)$. Note that corrupted sites have no influence on $M_i(x)$. Therefore, we can describe the dynamics informally as follows: sites try to align with their neighbors, but they only see the good

neighbors. After flipping their spin, sites are initially good, but they become corrupt with rate α and stay so until the next time when they flip their spin.

In simulations of the three dimensional model, for a suitable choice of the parameters, we see periodic behavior, where the system cycles through the states

$$(+1, +1) \mapsto (+1, -1) \mapsto (-1, -1) \mapsto (-1, +1) \mapsto (+1, +1) \mapsto \cdots$$

in a coordinated way. In two dimensions, we still see this behavior locally, but sites at larger distances from each other no longer coordinate their behavior.

Chapter 2

Continuous-time Markov chains

2.1 Finite state space

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

$$Af(x) := \sum_{y \in S} A(x,y)f(y) \quad \text{and} \quad fA(y) := \sum_{x \in S} f(x)A(x,y). \qquad (2.1)$$

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 probability kernel is *deterministic* if it is of the form

$$K_m(x,y) := \begin{cases} 1 & \text{if } x = m(x), \\ 0 & \text{otherwise,} \end{cases}$$

for some function $m : S \to S$. It is easy to see that the space of all probability kernels on a finite set S is convex, and the deterministic probability kernels are exactly the extremal elements of this set. It follows that each probability kernel can be written as a convex combination of deterministic probability

kernels. Another way to say this is that for each probability kernel K on S, it is possible to find a random map $M : S \to S$ such that[1]

$$K(x,y) = \mathbb{P}\big[M(x) = y\big] \qquad (x,y \in S). \tag{2.2}$$

A formula of this form is called a *random mapping representation* of the probability kernel K.

A *Markov semigroup* is a collection of probability kernels $(P_t)_{t \geq 0}$ such that

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

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, we say that a function f that is defined on an interval $I \subset \mathbb{R}$ is *piecewise constant* if each compact subinterval of I can be divided into finitely many subintervals, so that f is constant on each subinterval. 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 $(X_s)_{0 \leq s \leq t}$. One can prove that formula (2.4) is equivalent to the statement that

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

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. We say that $(P_t)_{t \geq 0}$ are the *transition kernels* of the Markov process. It is custom

---

[1]Indeed, this formula says nothing else than $K = \sum_m \mathbb{P}[M = m]K_m$, where the sum runs over all maps $m : S \to S$, and K_m is the deterministic kernel defined by the map m.

to let $\mathbb{P}^\mu$ denote the law of the Markov process with initial law μ, and to let $\mathbb{P}^x$ denote the law of the Markov process with deterministic initial state $X_0 = x$ a.s. We let $\mathbb{E}^\mu$ and $\mathbb{E}^x$ denote expectation with respect to $\mathbb{P}^\mu$ and $\mathbb{P}^x$, respectively. Recalling our notation (2.1), we see that for any probability law μ on S and function $f : S \to \mathbb{R}$,

$$\begin{aligned} \mu P_t(x) &= \mathbb{P}^\mu[X_t = x], \\ P_t f(x) &= \mathbb{E}^x[f(X_t)] \end{aligned} \qquad (t \geq 0, \ x \in S).$$

In particular, μP_t is the law of the process at time t. 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 S be a finite set. Let $\mathcal{K}$ be a finite set whose elements are probability kernels on S and let $(r_K)_{K \in \mathcal{K}}$ be nonnegative constants. Then it is straightforward to check that setting

$$Gf := \sum_{K \in \mathcal{K}} r_K \left\{ Kf - f \right\} \tag{2.6}$$

defines a Markov generator. The following exercise says that conversely, each Markov generator can be written in this form, where we can even choose the set $\mathcal{K}$ so that it has only one element.

Exercise 2.1 *Let S be a finite set. Show that each Markov generator G on S can be written in the form $Gf = r\{Kf - f\}$, where $r \geq 0$ is a constant and K is a probability kernel on S. Hint: first add a multiple of the identity matrix to G to make all diagonal entries nonnegative and then normalize.*

If all kernels in the set $\mathcal{K}$ are deterministic, then our expression for G takes the form

$$Gf(x) = \sum_{m \in \mathcal{G}} r_m \left\{ f\big(m(x)\big) - f(x) \right\}, \tag{2.7}$$

where $\mathcal{G}$ is a finite set whose elements are maps $m : S \to S$ and $(r_m)_{m \in \mathcal{G}}$ are nonnegative constants. We call (2.7) a *random mapping representation* of the generator G. Random mapping representations of generators will become important in Section 2.6 below.

Exercise 2.2 *Let S be a finite set. Show that each Markov generator G on S has a random mapping representation.*

2.2   The embedded Markov chain

Continuous-time Markov processes with countable state space (also known as continuous-time Markov chains) can in many ways be treated in the same way as those with a finite state space, but there are some complications, which mainly stem from the fact that such processes may *explode*. A good general reference for this material is [Lig10, Chapter 2].

We calculate with matrices indexed by a countably infinite set S in the same way as for finite S, provided the infinite sums are well-defined (i.e., not of the form $\infty - \infty$). Generalizing our earlier definition, we say that K is a *subprobability kernel* if $\sum_y K(x, y) \leq 1$ for all $x \in S$. Also when S is infinite, we define generators as in (2.3). Note that

$$\sum_{y \in S} G(x, y) = -G(x, x) + \sum_{y:\, y \neq x} G(x, y).$$

Since $G(x, y) \geq 0$ for $x \neq y$, the sum on the right-hand side is always well-defined, though a priori it may be infinite. The condition $\sum_y G(x, y)$ implies that it must be finite, however. In the special context of continuous-time Markov chains, a generator is traditionally called a *Q-matrix* (and denoted as Q) but we will stick to the term generator.

It is well-known [Lig10, Section 2.5.2] that one can construct a continuous-time Markov chain with generator G from its associated embedded discrete Markov chain and independent, exponentially distributed holding times. We now recall this construction. Let G be a generator, let $c(x) := -G(x, x)$ $(x \in S)$, and let K be the probability kernel on S defined by

$$K(x, y) := \begin{cases} c(x)^{-1} G(x, y) & \text{if } c(x) > 0, \ x \neq y, \\ 1 & \text{if } c(x) = 0, \ x = y, \\ 0 & \text{otherwise.} \end{cases}$$

For each $x \in S$, let $(Y_k^x)_{k \geq 0}$ be the discrete-time Markov chain with initial state $Y_0^x = x$ and transition kernel K. Set $N := \inf\{n \geq 0 : c(Y_n^x) = 0\}$, which may be infinite, and let $(\sigma_k)_{k \geq 0}$ be i.i.d. exponentially distributed random variables with mean one, independent of $(Y_k^x)_{k \geq 0}$. We define $(\tau_k)_{0 \leq k \leq N+1}$ by

$$\tau_0 := 0, \quad \tau_n := \sum_{k=0}^{n-1} \sigma_k / c(Y_k^x) \quad (1 \leq n \leq N+1),$$

where we use the conventions that $\sigma_k / c(Y_k^x) := \infty$ if $c(Y_k^x) = 0$ and $N + 1 := \infty$ if $N = \infty$. We set $\tau := \tau_{N+1}$. Note that $\tau = \infty$ on the event that $N < \infty$,

but τ may be finite on the event that $N = \infty$. We define a stochastic process $(X_t^x)_{t \geq 0}$ with values in $S_\infty := S \cup \{\infty\}$ by

$$X_t^x := \begin{cases} Y_k^x & \text{if } t \in [\tau_k, \tau_{k+1}), \ 0 \leq k < N+1, \\ \infty & \text{if } t \geq \tau. \end{cases}$$

We call τ the *explosion time*. We set $X_t^\infty := \infty$ $(t \geq 0)$ and define probability kernels $(\overline{P}_t)_{t \geq 0}$ on S_∞ by

$$\overline{P}_t(x, y) := \mathbb{P}\big[X_t^x = y\big] \qquad (t \geq 0, \ x, y \in S_\infty).$$

We call $(X_t^x)_{t \geq 0}$ the *continuous-time Markov chain* with generator G and we call $(Y_k^x)_{k \geq 0}$ its associated *embedded discrete-time Markov chain*. It is well-known that $(X_t^x)_{t \geq 0}$ is a Markov process (in the sense of (2.4) and (2.5)) with state space S_∞ and transition kernels $(\overline{P}_t)_{t \geq 0}$. In Exercise 2.3 below, you will be be asked to prove this. The random times

$$\eta_k := \sigma_k / c(Y_k^x) \quad (0 \leq k < N)$$

are called the *holding times*. Note that conditional on the embedded chain $(Y_k^x)_{k \geq 0}$, the holding times $(\eta_k)_{0 \leq k < N}$ are independent exponentially distributed such that η_k has mean $1/c(Y_k^x)$. If $\tau = \infty$ a.s. for each initial state $x \in S$, then we say that the continuous-time Markov chain with generator G is *nonexplosive*. In the opposite case, it is *explosive*. We let

$$P_t(x, y) := \overline{P}_t(x, y) \qquad (t \geq 0, \ x, y \in S) \tag{2.8}$$

denote the restrictions of the transition kernels $(\overline{P}_t)_{t \geq 0}$ to S. If G is explosive, then these are only subprobability kernels.

Exercise 2.3 *Show that $(X_t^x)_{t \geq 0}$ is a Markov process in the sense of (2.4) with state space S_∞ and transition kernels $(\overline{P}_t)_{t \geq 0}$. Hint: let M be the number of jumps of the process $(X_s^x)_{0 \leq s \leq t}$, which may be infinite. Then after conditioning on $(X_s^x)_{0 \leq s \leq t}$, you know M as well as $(Y_k^x)_{0 \leq k < M+1}$ and $(\sigma_k)_{0 \leq k < M}$, plus in the case that $M < \infty$ you have the information that $\sum_{k=0}^{M} \sigma_k / c(Y_k^x) > t$. Given all this information, what do you know about the process $(X_u^x)_{u \geq t}$?*

2.3 Generator construction

Let G be the generator of a continuous-time Markov chain with countable state space S. Generalising our earlier definition to countable state spaces,

we call the collection of subprobability kernels $(P_t)_{t \geq 0}$ defined in (2.8) the *Markov semigroup* with generator G. In this section we make a more direct link between $(P_t)_{t \geq 0}$ and G.

Let $u : S \times [0, \infty) \to [0, \infty)$ be a function. We say that u solves the *Kolmogorov backward equation*

$$\tfrac{\partial}{\partial t} u_t(x) = \sum_y G(x, y) u_t(y) \qquad (t \geq 0, \ x \in S), \qquad (2.9)$$

if the function $t \mapsto u_t(x)$ is continuously differentiable for each $x \in S$ and (2.9) holds. Note that the sum on the right-hand side of (2.9) is always well-defined since $u \geq 0$ and all terms except one are nonnegative. A priori, the sum might be ∞, but the equality in (2.9) implies that it must be finite. We say that u is a *minimal* solution to (2.9) if any other solution u' with the same initial condition $u_0' = u_0$ satisfies $u_t(x) \leq u_t'(x)$ for all $t \geq 0$ and $x \in S$. Note that for a given initial condition, there can be at most one minimal solution to (2.9). The following theorem is the main result of this section.

Theorem 2.4 (Generator construction) *Let G be the generator of a continuous-time Markov chain with countable state space S and let $(P_t)_{t \geq 0}$ be the Markov semigroup with generator G. Then for each bounded function $f : S \to [0, \infty)$, the function*

$$u_t(x) := P_t f(x) \qquad (t \geq 0, \ x \in S)$$

is the minimal solution to the Kolmogorov backward equation (2.9) with initial condition f.

To prepare for the proof of Theorem 2.4, as a first step, we set $c(x) := -G(x, x) \ (x \in S)$, and we consider the equation

$$u_t(x) = u_0(x) e^{-c(x)t} + \int_0^t \mathrm{d}s \, e^{-c(x)s} \sum_{y : \, y \neq x} G(x, y) u_{t-s}(y). \qquad (2.10)$$

By definition, a solution to (2.10) is a function $u : S \times [0, \infty) \to [0, \infty)$ such that $t \mapsto u_t(x)$ is measurable for all $x \in S$ and (2.10) holds for all $t \geq 0$ and $x \in S$. Note that since $G(x, y) \geq 0$ for $x \neq y$, the sum over y and consequently also the integral over s are well-defined, even though a priori the outcome may be ∞ (a posteriori, of course, (2.10) implies that the outcome must be finite).

Lemma 2.5 (First jump decomposition) *Under the assumptions of Theorem 2.4, the function u solves (2.10) with $u_0 = f$.*

Proof Since f is bounded $u_t(x) := P_t f(x) < \infty$ for all $t \geq 0$ and $x \in S$. (This is the only place in the proof where we use the boundedness of f.) Let $(X_t^x)_{t \geq 0}$ be the continuous-time Markov chain with generator G and initial state $x \in S$, constructed from the embedded discrete-time Markov chain $(Y_k^x)_{k \geq 0}$ and i.i.d. standard exponential random variables $(\sigma_k)_{k \geq 0}$ as in the previous section. We extend f to S_∞ by setting $f(\infty) := 0$ so that $P_t f(x) = \mathbb{E}[f(X_t^x)]$. Let τ_k denote the time when $(X_t^x)_{t \geq 0}$ makes its k-th jump. If $c(x) = 0$, then $\tau_1 = \infty$ a.s. and $X_t = x$ a.s. for all $t \geq 0$, which implies $u_t(x) = P_t f(x) = \mathbb{E}[f(X_t)] = f(x)$ $(t \geq 0)$, so (2.10) is trivially satisfied with $u_0 = f$. We assume, therefore, from now on that $c(x) > 0$. In this case, conditional on $\tau_1 = s$ and $X_{\tau_1}^x = y$, the process $(X_{\tau_1+t}^x)_{t \geq 0}$ is equally distributed with $(X_t^y)_{t \geq 0}$, which allows us to write

$$\mathbb{E}[f(X_t^x)] = f(x)\mathbb{P}[\tau_1 > t] + \int_0^t \mathbb{P}[\tau_1 \in \mathrm{d}t] \sum_{y:\, y \neq x} \mathbb{P}[X_{\tau_1}^x = y]\mathbb{P}\big[f(X_{t-s}^y)\big]$$

$$= f(x)e^{-c(x)t} + \int_0^t c(x)e^{-c(x)s}\mathrm{d}s\, c(x)^{-1} \sum_{y:\, y \neq x} G(x,y)P_{t-s}f(y),$$

which shows that $u_t := P_t f$ solves (2.10) with $u_0 = f$. ∎

Lemma 2.6 (The backward equation) *If a function $u : S \times [0, \infty) \to [0, \infty)$ solves (2.9), then it solves (2.10). Conversely, each bounded solution to (2.10) also solves (2.9).*

Proof If u solves (2.9), then

$$\tfrac{\partial}{\partial t}u_t(x) + c(x)u_t(x) = \sum_{y:\, y \neq x} G(x,y)u_t(y),$$

which implies

$$\tfrac{\partial}{\partial t}\Big(e^{c(x)t}u_t(x)\Big) = e^{c(x)t} \sum_{y:\, y \neq x} G(x,y)u_t(y) \qquad (t \geq 0,\ x \in S).$$

Integrating and then multiplying both sides of the equation by $e^{-c(x)t}$, we obtain

$$u_t(x) = u_0(x)e^{-c(x)t} + \int_0^t \mathrm{d}s\, e^{-c(x)(t-s)} \sum_{y:\, y \neq x} G(x,y)u_s(y),$$

which after the substitution $s \mapsto t - s$ yields (2.10).

Conversely, if (2.10) holds and u is bounded, then the right-hand side is continuous in t for each x and hence so is the left-had side. But then te right-hand side must actually be continuously differentiable as a function of t and the same must be true for the left-hand side. We can then reverse the argument above (differentiating instead of integrating) to obtain (2.9). ∎

Lemma 2.7 (Comparison principle) *Let G be the generator of a continuous-time Markov chain with countable state space S and let $(P_t)_{t\geq 0}$ be the Markov semigroup with generator G. Assume that $u : S \times [0,\infty) \to [0,\infty)$ satisfies*

$$\tfrac{\partial}{\partial t} u_t(x) \geq \sum_y G(x,y) u_t(y) \qquad (t \geq 0, \ x \in S),$$

where $t \mapsto u_t(x)$ is continuously differentiable for each $x \in S$. Then

$$P_t u_0(x) \leq u_t(x) \qquad (t \geq 0, \ x \in S).$$

Proof Let $(X_t^x)_{t\geq 0}$ be the continuous-time Markov chain with generator G and initial state x and as in Section 2.2, let τ_k denote the time of its k-th jump, for $1 \leq k \leq N$, where N denotes the total number of jumps, which may be finite or infinite, and let τ denote the explosion time. Then

$$P_t u_0(x) = \mathbb{E}\big[u_0(X_t^x) 1_{\{t < \tau\}}\big] \qquad (t \geq 0).$$

On the event that $N < \infty$ we set $\tau_k := \infty$ for $k > N$ and we define

$$u_0^{(n)}(x) = \mathbb{E}\big[u_0(X_t^x) 1_{\{t < \tau_n\}}\big] \qquad (t \geq 0).$$

Then

$$u_t^{(n)}(x) \underset{n\to\infty}{\longrightarrow} P_t u_0(x) \qquad (t \geq 0).$$

Using the same argument as in the proof of Lemma 2.5, we see that

$$u_t^{(n+1)}(x) = u_0(x) e^{-c(x)t} + \int_0^t \mathrm{d}s \, e^{-c(x)s} \sum_{y:\, y\neq x} G(x,y) u_{t-s}^{(n)}(y).$$

By the same argument as in the proof of Lemma 2.6, with all equalities replaced by inequalities,

$$u_t(x) \geq u_0(x) e^{-c(x)t} + \int_0^t \mathrm{d}s \, e^{-c(x)s} \sum_{y:\, y\neq x} G(x,y) u_{t-s}(y).$$

We claim that $u_t^{(n)}(x) \leq u_t(x)$ $(n \geq 0,\ t \geq 0)$. The proof is by induction. Clearly $u_t^{(0)}(x) = 0 \leq u_t(x)$ $(t \geq 0)$. Assuming that the statement holds for n, we have

$$u_t(x) \geq u_0(x)e^{-c(x)t} + \int_0^t \mathrm{d}s\, e^{-c(x)s} \sum_{y:\, y \neq x} G(x,y)u_{t-s}(y)$$

$$\geq u_0(x)e^{-c(x)t} + \int_0^t \mathrm{d}s\, e^{-c(x)s} \sum_{y:\, y \neq x} G(x,y)u_{t-s}^{(n)}(y) = u_t^{(n+1)}(x).$$

Letting $n \to \infty$ we obtain $u_t(x) \geq P_t(x)$. ∎

Proof of Theorem 2.4 By Lemmas 2.5 and 2.6, u solves the Kolmogorov backward equation (2.9) with initial condition f. If u' is another solution, then Lemma 2.7 implies that $u \leq u'$, showing that u is minimal. ∎

We conclude this section with the following observation.

Proposition 2.8 (Uniqueness of solutions) *The Kolmogorov backward equation (2.9) has a unique bounded solution u with initial condition $u_0 = f$ for each bounded function $f : S \to [0, \infty)$ if and only if the continuous-time Markov chain with generator G is nonexplosive.*

Proof For each $r \in \mathbb{R}$, let $\underline{r} : S \to \mathbb{R}$ denote the function that is constantly equal to r. If G is explosive, then $u_t := P_t\underline{1}$ and $u_t' := \underline{1}$ $(t \geq 0)$ are two different bounded solutions of the Kolmogorov backward equation (2.9) with initial condition $u_0 = u_0' = \underline{1}$, proving that solutions are not unique.

On the other hand, assume that G is nonexplosive and that u is a bounded solution with initial condition $u_0 = f$. Since u is bounded, there exist an $r \geq 0$ such that $u_t \leq \underline{r}$ $(t \geq 0)$. Lemma 2.7 tells us that $P_tf \leq u_t$ $(t \geq 0)$. Also, since $\underline{r} - u$ solves (2.9) with initial condition $\underline{r} - f$, Lemma 2.7 tells us that $\underline{r} - u_t \geq P_t(\underline{r} - f) = \underline{r} - P_tf$ $(t \geq 0)$ where in the last step we have used that G is nonexplosive. Combining these inequalities, we see that $u_t = P_tf$ $(t \geq 0)$. ∎

2.4 Lyapunov functions

It is tempting to think of explosive continuous-time Markov chains as pathological, but there exist very natural chains that are explosive. In fact, each transient chain can with a suitable random time transform be transformed into an explosive chain, so from this point of view the distinction between transient and recurrent chains would appear to be more fundamental than

the distinction between explosive and nonexplosive chains. Nevertheless, it is useful to have at our disposal a technique for proving that a given chain is nonexplosive. In the present section, we will show how nonexplosiveness can be proved with the help of Lyapunov functions. Below is the main result of this section. The term "Lyapunov function" originates in the stability theory of ordinary differential equations but is sometimes also used for certain functions occurring in Foster's theorem, that gives necessary and sufficient conditions for positive recurrence of a Markov chain. The role of the function L in the following theorem is similar, so using the term in a general sense, we may call it a Lyapunov function too.

Theorem 2.9 (Sufficient conditions for nonexplosiveness) *Let G be the generator of a continuous-time Markov chain with countable state space S and let $c(x) := -G(x,x)$ $(x \in S)$. Assume that there exists a function $L : S \to [0, \infty)$ and constant $\lambda \in \mathbb{R}$ such that:*

(i) $\sup\{c(x) : x \in S, \ L(x) < C\}$ *is finite for all $C < \infty$,*

(ii) $GL \leq \lambda L$.

Then the continuous-time Markov chain $(X_t)_{t \geq 0}$ with generator G is nonexplosive and

$$\mathbb{E}^x\big[L(X_t)\big] \leq e^{\lambda t} L(x) \qquad (t \geq 0, \ x \in S).$$

The proof of Theorem 2.9 depends on two lemmas.

Lemma 2.10 (Exponential bound) *Let $(X_t)_{t \geq 0}$ be a continuous-time Markov chain with generator G, started in $X_0 = x$. Assume that $L : S \to [0, \infty)$ satisfies $GL \leq \lambda L$ for some $\lambda \in \mathbb{R}$. Then*

$$\mathbb{E}^x\big[L(X_t)1_{\{t<\tau\}}\big] \leq e^{\lambda t} L(x) \qquad (t \geq 0), \qquad (2.11)$$

where τ denotes the explosion time of $(X_t)_{t \geq 0}$.

Proof The function $u_t(x) := L(x)e^{\lambda t}$ satisfies $\frac{\partial}{\partial t} u_t \geq G u_t$ $(t \geq 0)$, so Lemma 2.7 tells us that $P_t u_0 \leq u_t$ $(t \geq 0)$, which is the same as (2.11). ∎

Lemma 2.11 (Bounded jump rates) *Let G be the generator of a continuous-time Markov chain with countable state space S and let $c(x) := -G(x,x)$ $(x \in S)$. Assume that $\sup_{x \in S} c(x) < \infty$. Then G is nonexplosive.*

Proof Let $(X_t^x)_{t \geq 0}$ be the continuous-time Markov chain with generator G and initial state x, constructed from its embedded discrete-time Markov chain $(Y_k^x)_{k \geq 0}$ and i.i.d. standard exponential random variables $(\sigma_k)_{k \geq 0}$ as in Section 2.2. Let N be the total number of jumps, which may be finite or infinite, and let τ denote the explosion time. On the event that $N < \infty$ we have $\tau = \infty$ while on the event that $N = \infty$ we have

$$\tau = \sum_{k=0}^{\infty} \sigma_k / c(Y_k^x).$$

By our assumption that $C := \sup_{x \in S} c(x) < \infty$ we can estimate this from below by

$$\tau \geq C^{-1} \sum_{k=1}^{\infty} \sigma_k$$

which is ∞ a.s. by the strong law of large numbers. ∎

Proof of Theorem 2.9 It suffices to prove that G is nonexplosive, since the statement about the expectation of $L(X_t)$ then follows from Lemma 2.10. We set

$$S_C := \big\{ x \in S : L(x) < C \big\}$$

and define a generator G_C by

$$G_C(x, y) := \begin{cases} G(x, y) & \text{if } x \in S_C, \\ 0 & \text{if } x \notin S_C \end{cases}$$

We let $(X_t^x)_{t \geq 0}$ and $(X_t^{x,C})_{t \geq 0}$ denote the continuous-time Markov chains with generators G and G_C respectively. It follows from the construction of these processes in terms of their embedded Markov chains that we can naturally couple these processes such that

$$X_t^x = X_t^{x,C} \qquad \forall t \leq \tau_C := \inf \big\{ t \geq 0 : X_t^x \notin S_C \big\}.$$

In fact, we then have

$$X_t^{x,C} = X_{t \wedge \tau_C}^x \qquad (t \geq 0),$$

i.e., $(X_t^{x,C})_{t \geq 0}$ corresponds to the process $(X_t^x)_{t \geq 0}$ stopped as soon as it leaves S_C.

Let τ denote the explosion time of $(X_t^x)_{t \geq 0}$. Lemma 2.11 and assumption (i) of the theorem imply that $(X_t^{x,C})_{t \geq 0}$ is nonexplosive. Since the processes $(X_t^x)_{t \geq 0}$ and $(X_t^{x,C})_{t \geq 0}$ are equal up to time τ_C it follows that $\tau_C \leq \tau$.

Making λ larger if necessary, we can without loss of generality assume that $\lambda \geq 0$. Assumption (ii) of the theorem then implies that also $G_C L \leq \lambda L$. Indeed, $G_C L(x) = GL(x)$ if $L(x) < C$ and $G_C L(x) = 0$ otherwise. We can therefore use Lemma 2.10 and the fact that G_C is nonexplosive to conclude that

$$C\mathbb{P}^x[\tau_C \leq t] \leq \mathbb{E}\big[L(X_t^{x,C})\big] \leq e^{\lambda t} L(x) \qquad (t \geq 0).$$

Since $\tau_C \leq \tau$, it follows that

$$\mathbb{P}^x[\tau \leq t] \leq e^{\lambda t} L(x)/C \qquad (t \geq 0),$$

so letting $C \to \infty$ we see that $\tau = \infty$ a.s. ∎

2.5   Poisson point sets

The construction of a continuous-time Markov chain from its embedded discrete-time Markov chain is useful for theoretical purposes, but for the purpose of studying interacting particle systems a different construction will turn out to be much more useful, that is based on Poisson point sets. To prepare for this, in the present section, we recall the definition of Poisson point sets and some of their basic properties.

Let S be a σ-compact[2] metrizable space. We will mainly be interested in the case that $S = \mathcal{G} \times \mathbb{R}$ where $\mathcal{G}$ 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.

---

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

The following result follows from [Kal97, Lemma 10.1 and Prop. 10.4].[3] Below, $\mathcal{S}_{\mathrm{loc}} := \{A \in \mathcal{S} : \overline{A} \text{ is compact}\}$ denotes the set of measurable subsets of S whose closure is compact.

Proposition 2.12 (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 $A_1, \ldots, A_n \in \mathcal{S}_{\mathrm{loc}}$, 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 (2.12) 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 $A \in \mathcal{S}_{\mathrm{loc}}$. 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 $\mathrm{supp}(\nu)$, 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 if $\mu(\{x\}) > 0$. 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 2.13 (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

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

---

[3]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.

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 2.14 (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*

$$\mathbb{E}\big[(1 - f)^\xi\big] = e^{-\int f \mathrm{d}\mu} \tag{2.12}$$

for each measurable $f : S \to [0, 1]$. Conversely, if ξ is a random counting measure and (2.12) 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 (2.12) is proved in [Kal97, Lemma 10.2], which is written in terms of $-\log f$, rather than f. The fact that (2.12) determines the law of ξ uniquely follows from [Kal97, Lemma 10.1]. ∎

Formula (2.12) 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 (2.12) 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 2.14, 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 2.15 (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}.*

2.6 Poisson construction of Markov processes

In the present section we will show how a continuous-time Markov chain with countable state space S can be constructed by applying certain maps $m : S \to S$ at the times of a Poisson point process. We start with the following observation.

Lemma 2.16 (Random mapping representation) *Let S be a countable set, let $\mathcal{G}$ be a countable collection of maps $m : S \to S$, and let $(r_m)_{m\in\mathcal{G}}$ be nonnegative real numbers such that*

$$c(x) := \sum_{m:\,m(x)\neq x} r_m < \infty \quad \text{for all } x \in S. \tag{2.13}$$

Then there exists a generator G of a continuous-time Markov chain with state space S such that

$$Gf(x) = \sum_{m\in\mathcal{G}} r_m\big\{f\big(m(x)\big) - f\big(x\big)\big\} \tag{2.14}$$

for all bounded $f : S \to [0, \infty)$. Conversely, each generator G of a continuous-time Markov chain with state space S can be written in the form (2.14) for a suitable choice of a collection $\mathcal{G}$ of maps $m : S \to S$ and nonnegative real numbers $(r_m)_{m\in\mathcal{G}}$ satisfying (2.13).

Proof Let $\mathcal{G}$ be a countable collection of maps $m : S \to S$, let $(r_m)_{m\in\mathcal{G}}$ be nonnegative real numbers satisfying (2.13), and let $c(x)$ be as defined in (2.13). Then it is straightforward to check that

$$G(x, y) := \begin{cases} \displaystyle\sum_{\substack{m\in\mathcal{G} \\ m(x)=y}} r_m & \text{if } x \neq y, \\ -c(x) & \text{if } x = y \end{cases} \qquad (x, y \in S)$$

defines a generator such that (2.14) holds. To see that each generator G can be written in this form, we define for each $x, y \in S$ with $x \neq y$ a map

$m_{x,y} : S \to S$ by

$$
m_{x,y}(z) := \begin{cases} y & \text{if } z = x, \\ z & \text{otherwise,} \end{cases}
$$

we set $\mathcal{G} := \{m_{x,y} : x, y \in S, \ x \neq y\}$ and $r_{m_{x,y}} := G(x,y)$. It is then straightforward to check that the rates $(r_m)_{m \in \mathcal{G}}$ satisfy (2.13) and (2.14) holds. ∎

We call the way of writing G as in (2.14) a *random mapping representation* of G. Recall that random mapping representations of probability kernels have been defined in Section 2.1. There is a close connection between the two. Indeed, we will see that if $(P_t)_{t \geq 0}$ is the Markov semigroup with generator G, then using a random mapping representation of G we can for each $t \geq 0$ construct a random mapping representation of P_t. We will do this by giving a Poisson construction of the continuous-time Markov chain with generator G. In the context of interacting particle systems, such Poisson construction are known as *graphical representations* and they have many applications.

From now on, we assume that G is the generator of a continuous-time Markov chain with countable state space S and we fix a random mapping representation of G of the form (2.14) in terms of rates $(r_m)_{m \in \mathcal{G}}$ satisfying (2.13). We equip the space $\mathcal{G} \times \mathbb{R}$ with the measure

$$
\rho(\{m\} \times A) := r_m \, \ell(A) \qquad (m \in \mathcal{G}, \ A \in \mathcal{B}(\mathbb{R})), \tag{2.15}
$$

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. We call ω the *graphical representation* associated with the random mapping representation (2.14). We claim that for each $t \in \mathbb{R}$, there exists at most one $m \in \mathcal{G}$ such that $(m, t) \in \omega$. To see this, we note that for each $m \in \mathcal{G}$, the set

$$
\xi_m := \sum_{t : (m,t) \in \omega} \delta_t
$$

is a Poisson point measure on $\mathbb{R}$ with intensity $r_m \ell$. Since the sets $\mathbb{R} \times \{m\}$ $(m \in \mathcal{G})$ are disjoint, the random measures ξ_m $(m \in \mathcal{G})$ are independent, and hence for each $m \neq m'$, the measure $\xi_m + \xi_{m'}$ is a Poisson point measure on $\mathbb{R}$ with intensity $(r_m + r_{m'})\ell$. Since the Lebesgue measure is atomless, by Lemma 2.13, this Poisson point measure is simple, so there are no times $t \in \mathbb{R}$ for which both $(m, t) \in \omega$ and $(m', t) \in \omega$. In view of this, we can unambiguously define a random function $\mathbb{R} \ni t \mapsto \mathfrak{m}_t^\omega$ by setting

$$
\mathfrak{m}_t^\omega := \begin{cases} m & \text{if } (m, t) \in \omega, \\ 1 & \text{otherwise,} \end{cases}
$$

where we write 1 to denote the identity map.

As before, we write $S_\infty := S \cup \{\infty\}$. We equip S_∞ with a topology such that a set $A \subset S_\infty$ is closed if and only if A is either finite or A is infinite and $\infty \in A$. One can check that S_∞ is compact in this topology and a sequence $x_n \in S$ converges to ∞ if and only if it leaves every finite subset of S, i.e., for each finite $S' \subset S$, there exists an m such that $x_n \notin S'$ for all $n \geq m$. The topological space S_∞ is known as the *one-point compactification* of S.

Fix $s \in \mathbb{R}$. By definition, we say that a random function $X : [s, \infty) \to S_\infty$ solves the evolution equation

$$X_t = \mathfrak{m}_t^\omega(X_{t-}) \qquad (t > s) \tag{2.16}$$

if

(i) $\lim_{r \downarrow t} X_r = X_t$ $(t \geq s)$ and $\lim_{r \uparrow t} X_r =: X_{t-}$ exists $(t > s)$,

(ii) $X_t = \infty$ for all $t \geq \tau := \inf\{t \geq s : X_t = \infty\}$,

(iii) if $\tau < \infty$, then $\lim_{t \uparrow \tau} X_t = \infty$,

and (2.16) holds. Below is the main result of this section.

Theorem 2.17 (Poisson construction) *Let G be the generator of a continuous-time Markov chain with countable state space S and let ω be the graphical representation associated with a random mapping representation of G. Then almost surely, for each $s \in \mathbb{R}$ and $x \in S_\infty$, there exists a unique solution $(X_t^{s,x})_{t \geq s}$ to the evolution equation (2.16) with initial state $X_s^{s,x} = x$. Setting*

$$\mathbf{X}_{s,t}(x) := X_t^{s,x} \qquad (s \leq t, \ x \in S_\infty) \tag{2.17}$$

defines a collection of maps $(\mathbf{X}_{s,t})_{s \leq t}$ from S_∞ into itself such that

$$\mathbf{X}_{s,s} = 1 \quad and \quad \mathbf{X}_{t,u} \circ \mathbf{X}_{s,t} = \mathbf{X}_{s,u} \qquad (s \leq t \leq u). \tag{2.18}$$

If G is nonexplosive, then almost surely, $\mathbf{X}_{s,t}$ maps S into itself for all $s \leq t$. If $s \in \mathbb{R}$ and X_0 is an S-valued random variable with law μ, independent of ω, then the process $(X_t)_{t \geq 0}$ defined as

$$X_t := \mathbf{X}_{s,s+t}(X_0) \qquad (t \geq 0) \tag{2.19}$$

is distributed as the continuous-time Markov chain with generator G and initial law μ.

Formula (2.18) says that the random maps $(\mathbf{X}_{s,t})_{s \leq t}$ form a *stochastic flow*. This stochastic flow is *stationary* in the sense that

$$X_{0,t} \text{ is equally distributed with } X_{s,s+t} \quad (s \in \mathbb{R}).$$

We note that since $\mathbf{X}_{s,t}$ is constructed using only Poisson points of the form (m, r) with $s < r \leq t$, and restrictions of a Poisson point set to disjoint parts of the space are independent, it follows that the stochastic flow $(\mathbf{X}_{s,t})_{s \leq t}$ has *independent increments* in the sense that

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

We now that (2.19) implies that

$$P_t(x, y) = \mathbb{P}\big[\mathbf{X}_{s,s+t}(x) = y\big] \qquad (s, t \in \mathbb{R}, \ x, y \in S),$$

so indeed, as announced, we have found a random mapping representation of the subprobability kernels $(P_t)_{t \geq 0}$.

Proof of Theorem 2.17 We start by proving that for each $s \in \mathbb{R}$ and $x \in S_\infty$, there exists a unique solution $(X_t^{s,x})_{t \geq s}$ to the evolution equation (2.16) with initial state $X_s^{s,x} = x$. If $x = \infty$, then clearly $X_t^\infty = \infty$ $(t \geq s)$ is the unique solution of (2.16) so without loss of generality we assume from now on that $x \in S$. For each $x \in S$, the set

$$\big\{ t \in \mathbb{R} : \mathfrak{m}_t^\omega(x) \neq x \big\}$$

is a Poisson point set on $\mathbb{R}$ with intensity $c(x)$ defined in (2.13), which is finite by assumption. This allows us to inductively define times $(\tau_n)_{n \geq 0}$ and a discrete chain $(Y_n^x)_{n \geq 0}$ by setting $\tau_0 := 0$, $Y_0^x := x$,

$$\tau_{n+1} := \begin{cases} \inf \big\{ t > \tau_n : \mathfrak{m}_t^\omega(x_n) \neq x_n \big\} & \text{if } \tau_n < \infty, \\ \infty & \text{if } \tau_n = \infty \end{cases}$$

and

$$Y_{n+1}^x := \begin{cases} \mathfrak{m}_{\tau_{n+1}}^\omega(Y_n^x) & \text{if } \tau_n < \infty, \\ Y_n^x & \text{if } \tau_n = \infty. \end{cases}$$

We set $\tau := \lim_{n \to \infty} \tau_n$ and

$$N := \inf \big\{ n \geq 0 : \tau_{n+1} = \infty \big\} = \inf \big\{ n \geq 0 : c(Y_n^x) = 0 \big\}.$$

We claim that $(X_t^{s,x})_{t \geq s}$ defined as

$$X_t^x := \begin{cases} Y_k^x & \text{if } t \in [\tau_k, \tau_{k+1}), \ 0 \leq k < N + 1, \\ \infty & \text{if } t \geq \tau \end{cases}$$

solves the evolution equation (2.16). It is easy to see that it satisfies properties (i) and (ii) of a solution and that (2.16) holds. To see that it also satisfies (iii), imagine that on the contrary that $\tau < \infty$ while X_t does not converge to ∞ as $t \to \tau$. By the definition of the one-point compactification, this implies that there exist a finite set $S' \subset S$ and times $s_n \to \tau$ such that $X_{s_n} \in S'$. This, in turn, implies that during the finite time interval $[0, \tau)$, the function X_t makes infinitely many jumps that start at some point in S' and end in some other point in S. But this is impossible, since by (2.13),

$$\bigcup_{y \in S'} \left\{ t \in \mathbb{R} : \mathfrak{m}_t^\omega(y) \neq y \right\}$$

is a locally finite subset of $\mathbb{R}$. This completes the proof that $(X_t^{s,x})_{t \geq s}$ solves (2.16). If $(X_t')_{t \geq s}$ is another solution, then we see by induction that $X_t' = X_t^{s,x}$ for all $0 \leq t \leq \tau_n$ and for all $n \geq 0$. By property (iii) this implies that $X_\tau' = \infty$ and hence $X_t' = \infty$ for all $t \geq \tau$ by property (ii), which shows that solutions to (2.16) are unique.

Let $(\mathbf{X}_{s,t})_{s \leq t}$ be defined in (2.17). Then it is straightforward to check that $(\mathbf{X}_{s,t})_{s \leq t}$ is a stochastic flow in the sense of (2.18). By the argument we have already given, $(\mathbf{X}_{s,t})_{s \leq t}$ has independent increments. It is also clearly *stationary* in the sense that

$$\mathbf{X}_{0,t} \text{ is equally distributed with } \mathbf{X}_{s,s+t} \qquad (s, t \in \mathbb{R}).$$

Let $(\overline{P}_t)_{t \geq 0}$ be the probability kernels on S_∞ defined as

$$\overline{P}_t(x, y) := \mathbb{P}\big[\mathbf{X}_{0,t}(x) = y\big] \qquad (t \geq 0, \ x, y \in S_\infty).$$

If $(X_t)_{t \geq 0}$ is defined as in (2.19), then using the fact that $(\mathbf{X}_{s,t})_{s \leq t}$ is a stationary stochastic flow with independent increments, we see that

$$\mathbb{P}\big[X_0 = x_0, \dots, X_{t_n} = x_n\big] \quad = \mathbb{P}[X_0 = x_0] P_{t_1 - t_0}(x_0, x_1) \cdots P_{t_n - t_{n-1}}(x_{n-1}, x_n)$$

for all $0 < t_1 < \cdots < t_n$ and $x_0, \dots, x_n \in S_\infty$, which shows that $(X_t)_{t \geq 0}$ is a Markov process with transition kernels $(\overline{P}_t)_{t \geq 0}$.

Let $(P_t)_{t \geq 0}$ denote the restrictions of the probability kernels $(\overline{P}_t)_{t \geq 0}$ to S. We claim that $(P_t)_{t \geq 0}$ is the Markov semigroup with generator G. Using the properties of Poisson point processes, it is easy to see that

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

where 1 denotes the identity matrix, so $1(x, z) = 1$ if $x = z$ and $= 0$ otherwise. One can prove that this property, together with the Markov property of the

process $(X_t)_{t\geq 0}$, suffices to uniquely identify $(X_t)_{t\geq 0}$ defined in (2.19) as the continuous-time Markov chain with generator G and initial law μ. We refer the interested reader to [Lig10, Section 2.3] for the arguments, which are elegant but a bit lengthy.

An alternative way to prove that $(X_t)_{t\geq 0}$ defined in (2.19) is the continuous-time Markov chain with generator G and initial law μ is as follows. First, by conditioning on X_0 (which is independent of everything else) we can reduce the problem to the case that $X_0 = x$ is deterministic. The case $x = \infty$ is trivial so without loss of generality we may assume that $x \in S$. It then suffices to show that $(Y_n^x)_{n\geq 0}$ defined above is the correct embedded discrete-time Markov chain and that

$$\sigma_n := c(Y_n^x)(\tau_{n+1} - \tau_n) \qquad (0 \leq n < N)$$

are i.i.d. standard exponentially distributed random variables. More precisely, we want to show that for each $x = x_0, \ldots, x_n$ with $c(x_k) > 0$ for all $0 \leq k < n$ and for all $s_1, \ldots, s_n \geq 0$, one has

$$\mathbb{P}\big[\tau_1 - \tau_0 > s_1, \ Y_1^x = x_1, \ldots, \tau_n - \tau_{n-1} > s_n, \ Y_n^x = x_n\big]$$
$$= e^{-c(x_0)s_1}c(x_0)^{-1}G(x_0, x_1)\cdots e^{-c(x_{n-1})s_n}c(x_{n-1})^{-1}G(x_{n-1}, x_n).$$

The easiest way to prove this is via discrete approximation. For $r_1, \ldots, r_n \in \varepsilon\mathbb{N} := \{\varepsilon n : n \geq 0\}$, setting $t_k := \sum_{n=1}^{k}(r_k + \varepsilon)$ $(0 \leq k \leq n)$, we have

$$\mathbb{P}\big[\tau_1 \in (t_1, t_1 + \varepsilon], \ Y_1^x = x_1, \ldots, \tau_n \in (t_n, t_n + \varepsilon], \ Y_n^x = x_n\big]$$
$$= e^{-c(x_0)r_1}\varepsilon G(x_0, x_1)\cdots e^{-c(x_{n-1})r_n}\varepsilon G(x_{n-1}, x_n) + O(\varepsilon^2),$$

where $e^{-c(x_{k-1})r_k}$ is the probability that $m_t^\omega(x_{k-1}) = x_{k-1}$ for all $t \in (t_{k-1} + \varepsilon, t_k]$ and $\varepsilon G(x_{k-1}, x_k) + O(\varepsilon^2)$ is the probability that there exists a unique $t \in (t_k, t_k + \varepsilon]$ such that $m_t^\omega(x_{k-1}) = x_k$, while $m_{t'}^\omega(x_{k-1}) = x_{k-1}$ for all $t' \in (t_k, t]$ and $m_{t'}^\omega(x_k) = x_k$ for all $t' \in (t, t_k + \varepsilon]$. Since

$$\sum_{\substack{r\in\varepsilon\mathbb{N} \\ r\geq s_k}} \varepsilon e^{-c(x_{k-1})r} \approx \int_{s_k}^{\infty}\mathrm{d}r\, e^{-c(x_{k-1})r} = c(x_{k-1})^{-1}e^{-c(x_{k-1})s_k},$$

by letting $\varepsilon \to 0$ one can check that the joint law of the embedded Markov chain and the holding times are precisely what they should be. For brevity, we leave the details to the reader.

To complete the proof, we must show that if G is nonexplosive, then almost surely, $\mathbf{X}_{s,t}$ maps S into itself for all $s \leq t$. In other words, we must show that

$$\mathbf{X}_{s,t}(x) \in S \quad \forall s, t \in \mathbb{R} \text{ with } s \leq t \text{ and } x \in S \quad \text{a.s.}$$

If s is deterministic, then by what we have proved $X_t := \mathbf{X}_{s,s+t}(x)$ $(t \geq 0)$ is the continuous-time Markov chain with generator G and initial state x, so if this process is nonexplosive, then

$$\mathbf{X}_{s,t}(x) \in S \quad \forall t \in [s,\infty) \quad \text{a.s.} \quad (s \in \mathbb{R}, \ x \in S).$$

Since $s \mapsto \mathbf{X}_{s,t}(x)$ is constant between the times of the Poisson point process

$$\left\{ s \in \mathbb{R} : \mathfrak{m}_s^\omega(x) \neq x \right\}$$

which has finite intensity by (2.13), we can improve our previous statement to

$$\mathbf{X}_{s,t}(x) \in S \quad \forall s,t \in \mathbb{R} \text{ with } s \leq t \quad \text{a.s.} \quad (x \in S),$$

and since S is countable, we see that the statement holds for all $s,t \in \mathbb{R}$ with $s \leq t$ and $x \in S$ simultaneously. ∎

2.7 Local maps

The state space of the interacting particle systems we are interested is 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*. Here S^Λ denotes the cartesian product of Λ copies of S, i.e., this is the space of all functions $x : \Lambda \to S$. Typically, we can write the generator G of an interacting particle system in the form (2.14), i.e., there is a natural random mapping representation for G. Sometimes, there are two or more random mapping representations for the same generator that each have their own advantages and disadvantages. Sometimes, like in the case of stochastic Ising models, it is not immediately obvious how to find a random mapping representation of the generator, but in Section 4.6 we will show that one can found for this process as well.

If the lattice Λ is finite, then so is S^Λ and we can immediately apply Theorem 2.17 to construct our interacting particle system from a graphical representation ω. If Λ is infinite, then S^Λ is uncountable (as long as S has at least two elements), so Theorem 2.17 is not applicable. In Chapter 4 we will develop the methods needed to construct interacting particle systems on infinite lattices from a graphical representation. There is one important special case where Theorem 2.17 is applicable, however, even if the lattice is finite. Many interacting particle systems with a local state space of the form $S = \{0, 1\}$ have the property that if we start the system in an initial configuration with finitely many ones, then the system stays in such states

for all times. Examples are the voter model, the contact process, and various systems of branching and coalescing particles. A counterexample are stochastic Ising models.

In Section 2.8 we will apply Theorem 2.17 to such interacting particle systems on infinite lattices, but started in an initial configuration with finitely many ones. To prepare for this, in the present section, we take a closer look at the sort of maps that are typically used to construct an interacting particle system, such as the voter map in (1.4) and the branching and death maps in (1.6) and (1.7).

We will always equip the state space S^Λ of an interacting particle system 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 simply says 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.

Let S and T be finite sets, let Λ be a countable set, and let $f : S^\Lambda \to T$ be a function. Then we say that a point $j \in \Lambda$ is *f-relevant* if

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

i.e., changing the value of x in j may change the value of $f(x)$. We write

$$\mathcal{R}(f) := \big\{ j \in \Lambda : j \text{ is } f\text{-relevant} \big\}.$$

The following lemma (which we have taken from [SS15b, Lemma 24]) says that a function $f : S^\Lambda \to T$ is continuous with respect to the product topology if and only if it depends on finitely many coordinates.

Lemma 2.18 (Continuous maps) *Let S and T be finite sets and let Λ be a countable set. Then a function $f : S^\Lambda \to T$ is continuous with respect to the product topology if and only if the following two conditions are satisfied:*

(i) *$\mathcal{R}(f)$ is finite,*

(ii) *If $x, y \in S^\Lambda$ satisfy $x(j) = y(j)$ for all $j \in \mathcal{R}(f)$, then $f(x) = f(y)$.*

Before we give the proof of Lemma 2.18, we first make some observations. The following exercise shows how continuity can fail if condition (i) of Lemma 2.18 does not hold.

Exercise 2.19 (A discontinuous map) *Let* $2\mathbb{N} := \{2n : n \in \mathbb{N}\}$ *and* $2\mathbb{N} + 1 := \{2n + 1 : n \in \mathbb{N}\}$. *Define* $f : \{0,1\}^{\mathbb{N}} \to \{0,1\}$ *by*

$$f(x) := \begin{cases} 1 & \text{if } \inf\{i \in \mathbb{N} : x(i) = 1\} \in 2\mathbb{N} \cup \{\infty\}, \\ 0 & \text{if } \inf\{i \in \mathbb{N} : x(i) = 1\} \in 2\mathbb{N} + 1. \end{cases} \tag{2.20}$$

Show that f satisfies condition (ii) of Lemma 2.18 but not condition (i). Show that f is not continuous.

The following exercise shows that contrary to what one might initially have guessed, condition (ii) of Lemma 2.18 is not automatically satisfied, even when condition (i) holds.

Exercise 2.20 (Another discontinuous map) *Define $f : \{0,1\}^{\mathbb{N}} \to \{0,1\}$ by*

$$f(x) := \begin{cases} 1 & \text{if } \{i \in \mathbb{N} : x(i) = 1\} \text{ is finite,} \\ 0 & \text{if } \{i \in \mathbb{N} : x(i) = 1\} \text{ is infinite.} \end{cases} \tag{2.21}$$

Show that f satisfies condition (i) of Lemma 2.18 but not condition (ii). Show that f is not continuous.

Proof of Lemma 2.18 Let $(\alpha_j)_{j \in \Lambda}$ be strictly positive constants such that $\sum_{j \in \Lambda} \alpha_j < \infty$. Then the metric

$$d(x,y) := \sum_{j \in \Lambda} \alpha_j 1_{\{x(j) \neq y(j)\}} \qquad (x, y \in S^{\Lambda}) \tag{2.22}$$

generates the product topology on S^{Λ}. By Tychonoff's theorem, S^{Λ} is compact, so the function f is uniformly continuous. Since the target space T is finite, this means that there exists an $\varepsilon > 0$ such that $d(x,y) < \varepsilon$ implies $f(x) = f(y)$. Since $\sum_{j \in \Lambda} \alpha_j < \infty$, there exists some finite $\Lambda' \subset \Lambda$ such that $\sum_{j \in \Lambda \setminus \Lambda'} \alpha_j < \varepsilon$. It follows that

(ii)' If $x, y \in S^{\Lambda}$ satisfy $x(j) = y(j)$ for all $j \in \Lambda'$, then $f(x) = f(y)$.

We conclude from this that $\mathcal{R}(f) \subset \Lambda'$, proving (i). If this is a strict inclusion, then we can inductively remove those points from Λ' that are not elements of $\mathcal{R}(f)$ while preserving the property (ii)', until in a finite number of steps we see that (ii) holds.

Conversely, if (i) and (ii) hold and $x_k \to x$ pointwise, then by (i) there exists some n such that $x_k(j) = x(j)$ for all $j \in \mathcal{R}(f)$ and hence by (ii) $f(x_k) = f(x)$ for all $k \geq n$, proving that f is continuous. ∎

For any map $m : S^{\Lambda} \to S^{\Lambda}$ and $i \in \Lambda$, we define $m[i] : S^{\Lambda} \to S$ by $m[i](x) := m(x)(i)$ $(x \in S^{\Lambda}, \ i \in \Lambda)$. It follows immediately from the definition of the product topology that m is continuous if and only if $m[i]$ is

continuous for all $i \in \Lambda$. We 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 i for which $m[i]$ is not the identity map.

By definition, a *local map* is a function $m : S^{\Lambda} \to S^{\Lambda}$ such that:

(i) m is continuous,

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

In view of Lemma 2.18, this says that m is local if m changes the values of at most finitely many lattice points using information from finitely many lattice points only. The following exercise describes yet another way to look at local maps.

Exercise 2.21 (Local maps) *Show that a map $m : S^{\Lambda} \to S^{\Lambda}$ is local if and only if there exists a finite set $\Delta \subset \Lambda$ 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

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

  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

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

  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

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

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

- For each $i \in \Lambda$, we can similarly define a *birth map* $\texttt{birth}_i : \{0,1\}^\Lambda \to \{0,1\}^\Lambda$ as

$$\texttt{birth}_i(x)(k) := \begin{cases} 1 & \text{if } k = i, \\ x(k) & \text{otherwise.} \end{cases} \tag{2.23}$$

  Then

$$\mathcal{D}(\texttt{birth}_i) = \{i\} \quad \text{and} \quad \mathcal{R}(\texttt{birth}_i[i]) = \emptyset.$$

- The coalescing random walk map $\texttt{rw}_{ij}$ defined in (1.20) satisfies

$$\mathcal{D}(\texttt{rw}_{ij}) = \{i,j\}, \quad \mathcal{R}(\texttt{rw}_{ij}[i]) = \emptyset, \quad \text{and} \quad \mathcal{R}(\texttt{rw}_{ij}[j]) = \{i,j\},$$

  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 $\texttt{rw}_{ij}(x)(j)$ we need to know both $x(i)$ and $x(j)$.

Exercise 2.22 (Exclusion and cooperative branching maps) *Recall the exclusion map $\texttt{excl}_{ij}$ defined in (1.23) and the cooperative branching map $\texttt{coop}_{ijk}$ defined in (1.25). For $m = \texttt{excl}_{ij}$ or $m = \texttt{coop}_{ijk}$, determine $\mathcal{D}(m)$, and determine $\mathcal{R}(m[i])$ for all $i \in \mathcal{D}(m)$.*

2.8 Systems of finitely many particles

Throughout this section we assume that S is a finite set containing a special element that we denote by 0. For $x \in S^\Lambda$ we introduce the notation

$$|x| := \big|\{i \in \Lambda : x(i) \neq 0\}\big| \qquad (x \in S^\Lambda)$$

and we write

$$\mathcal{S}(\Lambda) := S^\Lambda \quad \text{and} \quad \mathcal{S}_{\text{fin}}(\Lambda) := \big\{x \in \mathcal{S}(\Lambda) : |x| < \infty\big\}.$$

It is easy to see that $\mathcal{S}_{\text{fin}}(\Lambda)$ is countable. We let $\underline{0} \in \mathcal{S}(\Lambda)$ denote the configuration that is identically zero, i.e., this is the constant function defined as $\underline{0}(i) := 0$ $(i \in \Lambda)$. We note that if a local map $m : \mathcal{S}(\Lambda) \to \mathcal{S}(\Lambda)$ satisfies

$$m(\underline{0}) = \underline{0}, \tag{2.24}$$

then by the fact that it is a local map

$$x \in \mathcal{S}_{\text{fin}}(\Lambda) \quad \text{implies} \quad m(x) \in \mathcal{S}_{\text{fin}}(\Lambda).$$

Almost all the examples of local maps listed in the previous section satisfy (2.24). Indeed, this holds for the local maps

$$\mathtt{vot}_{ij}, \quad \mathtt{bra}_{ij}, \quad \mathtt{death}_i, \quad \mathtt{rw}_{ij}, \quad \mathtt{excl}_{ij}, \quad \text{and} \quad \mathtt{coop}_{ijk},$$

while $\mathtt{birth}_i$ is the only local map listed in the previous section that does not map $\underline{0}$ into itself.

Let $\mathcal{G}$ is a countable collection of local maps $m : \mathcal{S}(\Lambda) \to \mathcal{S}(\Lambda)$ that all satisfy (2.24) and let $(r_m)_{m \in \mathcal{G}}$ be rates. Then under suitable assumptions on the rates, we may expect that

$$Gf(x) := \sum_{m \in \mathcal{G}} r_m \big\{ f\big(m(x)\big) - f(x) \big\} \tag{2.25}$$

is the generator of a nonexplosive continuous-time Markov chain with countable state space $\mathcal{S}_{\mathrm{fin}}(\Lambda)$. To formulate sufficient conditions for this to be true, for any local map $m : S^\Lambda \to S^\Lambda$ we introduce the following notation:

$$\mathcal{R}(m) := \big\{ (i, j) \in \Lambda^2 : j \in \mathcal{D}(m), \ i \in \mathcal{R}(m[j]) \big\},$$

$$\mathcal{R}_i^\uparrow(m) := \big\{ j \in \Lambda : (i, j) \in \mathcal{R}(m) \big\}, \quad \mathcal{R}_j^\downarrow(m) := \big\{ i \in \Lambda : (i, j) \in \mathcal{R}(m) \big\}.$$

Here is the main result of this section. The form of condition (2.26) is originally due to [Lat24].

Theorem 2.23 (Finite particle configurations) *Let S and Λ be a finite and countable set, respectively, and assume that S contains a special element denoted as 0. Let $\mathcal{G}$ be a countable collection of local maps $m : S^\Lambda \to S^\Lambda$ such that $m(\underline{0}) = \underline{0}$ for all $m \in \mathcal{G}$ and let $(r_m)_{m \in \mathcal{G}}$ be nonnegative rates. Assume that*

$$\sup_{i \in \Lambda} \sum_{m \in \mathcal{G}} r_m \big(1_{\mathcal{D}(m)}(i) + |\mathcal{R}_i^\uparrow(m)| \big) < \infty. \tag{2.26}$$

Then G defined in (2.25) is the generator of a nonexplosive continuous-time Markov chain with state space $\mathcal{S}_{\mathrm{fin}}(\Lambda)$.

Proof We start by checking condition (2.13) which is necessary and sufficient for (2.25) to define the generator of a (possibly explosive) continuous-time Markov chain. In our present setting, (2.13) reads

$$\sum_{m : m(x) \neq x} r_m < \infty \quad \text{for all } x \in \mathcal{S}_{\mathrm{fin}}(\Lambda).$$

Let $x \in \mathcal{S}_{\mathrm{fin}}(\Lambda)$ and let $A := \{ i \in \Lambda : x(i) \neq 0 \}$ which is finite by the definition of $\mathcal{S}_{\mathrm{fin}}(\Lambda)$. If $m(x) \neq x$, then $m(x)(j) \neq x(j)$ for some some

$j \in \mathcal{D}(m)$. If $j \notin A$, then by the fact that $m(\underline{0}) = \underline{0}$, there must exist an $i \in A$ such that $i \in \mathcal{R}(m[j])$. This allows us to estimate

$$\sum_{m:\, m(x) \neq x} r_m \leq \sum_{j \in A} \sum_{m \in \mathcal{G}} 1_{\mathcal{D}(m)}(j) r_m + \sum_{j \in \Lambda \setminus A} \sum_{i \in A} \sum_{m \in \mathcal{G}} 1_{\mathcal{R}_i^{\uparrow}(m)}(j) r_m$$
$$= \sum_{i \in A} \Big(\sum_{m \in \mathcal{G}} 1_{\mathcal{D}(m)}(i) r_m + \sum_{m \in \mathcal{G}} |\mathcal{R}_i^{\uparrow}(m) \setminus A| r_m \Big),$$

which is finite by (2.26).

It remains to prove that G is nonexplosive. We apply Theorem 2.9 to the Lyapunov function

$$L(x) := |x| \quad \big(x \in \mathcal{S}_{\mathrm{fin}}(\Lambda) \big). \tag{2.27}$$

Note that $c(x) := -G(x, x) = \sum_{m:\, m(x) \neq x} r_m$, so our previous calculation shows that the function L satisfies condition (i) of Theorem 2.9. It therefore remains to check condition (ii) of Theorem 2.9. We estimate

$$GL(x) = \sum_{m \in \mathcal{G}} r_m \big\{ L\big(m(x) \big) - L\big(x \big) \big\}$$
$$\leq \sum_{m \in \mathcal{G}} r_m \big| \big\{ j \in \Lambda : x(j) = 0, \ m(x)(j) \neq 0 \big\} \big|$$
$$\leq \sum_{m \in \mathcal{G}} r_m \big| \big\{ (i, j) \in \mathcal{R}(m) : x(i) \neq 0, \ x(j) = 0 \big\} \big|$$
$$\leq L(x) \sup_{i \in \Lambda} \sum_{m \in \mathcal{G}} r_m |\mathcal{R}_i^{\uparrow}(m)| \qquad (x \in \mathcal{S}_{\mathrm{fin}}(\Lambda)),$$

where in the second inequality we have used that $m(\underline{0}) = \underline{0}$ for all $m \in \mathcal{G}$. Our calculation shows that condition (ii) of Theorem 2.9 is satisfied with

$$\lambda := \sup_{i \in \Lambda} \sum_{m \in \mathcal{G}} r_m |\mathcal{R}_i^{\uparrow}(m)|,$$

which is finite by (2.26). ∎

It is instructive to see some concrete examples of interacting particle systems to which Theorem 2.23 is applicable. Generalising (1.5), if $\lambda : \Lambda^2 \to [0, \infty)$ is a function, then we can define a *voter model* generator by

$$G_{\mathrm{vot}} f(x) := \sum_{i,j \in \Lambda^2} \lambda(i, j) \big\{ f\big(\mathtt{vot}_{ij}(x) \big) - f\big(x \big) \big\} \qquad (x \in S^{\Lambda}), \tag{2.28}$$

where $\lambda(i, j) \geq 0$ is the Poisson rate at which site j adopts the type of site i.

Exercise 2.24   *Show that the generator G_{vot} satisfies the assumptions of Theorem 2.23 if*

$$\sup_{i \in \Lambda} \Big[\sum_{j \in \Lambda} \lambda(j, i) + \sum_{j \in \Lambda} \lambda(i, j) \Big] < \infty.$$

Similarly, generalising (1.8), we can define a *contact process* generator by

$$\begin{aligned} G_{\mathrm{cont}} f(x) := &\sum_{i,j \in \Lambda} \lambda(i, j) \big\{ f\big(\mathtt{bra}_{ij}(x)\big) - f\big(x\big) \big\} \\ &+ \delta \sum_{i \in \Lambda} \big\{ f\big(\mathtt{death}_i(x)\big) - f\big(x\big) \big\}, \end{aligned} \tag{2.29}$$

where $\lambda(i, j) \geq 0$ is the *infection rate* from i to j and $\delta \geq 0$ is the *death rate*.

Exercise 2.25   *Show that the generator G_{cont} satisfies the assumptions of Theorem 2.23 if*

$$\sup_{i \in \Lambda} \Big[\sum_{j \in \Lambda} \lambda(j, i) + \sum_{j \in \Lambda} \lambda(i, j) \Big] < \infty.$$

For processes with some sort of translation invariant structure the expression $\sum_{j \in \Lambda} \lambda(j, i) + \sum_{j \in \Lambda} \lambda(i, j)$ does not depend on $i \in \Lambda$. Using this, one can check that for translation invariant voter models, the condition in Exercise 2.24 is optimal. Indeed, if we start the process with a single one at i, then $\sum_{j \in \Lambda} \lambda(j, i)$ is the rate at which this one becomes a zero while $\sum_{j \in \Lambda} \lambda(i, j)$ is the rate at which this one produces another one somewhere. In case of the contact process, we can actually do a bit better than Exercise 2.25.

Proposition 2.26 (Finite contact processes)   *Assume that*

$$r := \sup_{i \in \Lambda} \sum_{j \in \Lambda} \lambda(i, j) < \infty. \tag{2.30}$$

Then G_{cont}, defined in (2.29), is the generator of a nonexplosive continuous-time Markov chain with state space $\mathcal{S}_{\mathrm{fin}}(\Lambda)$. Moreover,

$$\mathbb{E}^x\big[|X_t|\big] \leq e^{(r - \delta)t}|x| \qquad \big(t \geq 0, \ x \in \mathcal{S}_{\mathrm{fin}}(\Lambda)\big). \tag{2.31}$$

Proof   For any $x \in \mathcal{S}_{\mathrm{fin}}(\Lambda)$, we can estimate the quantity $c(x)$ from (2.13) by

$$\sum_{m: \, m(x) \neq x} r_m = \delta|x| + \sum_{i: \, x(i)=1} \sum_{j: \, x(j)=0} \lambda(i, j) \leq (\delta + r)|x|,$$

where r is the quantity in (2.30). Since this is finite for each $x \in \mathcal{S}_{\mathrm{fin}}(\Lambda)$, G_{cont} is the generator of a (possibly explosive) continuous-time Markov chain with state space $\mathcal{S}_{\mathrm{fin}}(\Lambda)$.

To see that G_{cont} is nonexplosive we apply Theorem 2.9 to the Lyapunov function in (2.27). Our previous calculation shows that L satisfies condition (i) of Theorem 2.9 so it remains to check condition (ii). We estimate

$$
\begin{aligned}
GL(x) &= \sum_{m \in \mathcal{G}} r_m \big\{ L\big(m(x)\big) - L\big(x\big) \big\} \\
&\leq \sum_{i,j \in \Lambda} \lambda(i,j) 1_{\{x(i)=1,\ x(j)=0\}} - \delta \sum_{i \in \Lambda} 1_{\{x(i)=1\}} \\
&\leq (r - \delta)|x|,
\end{aligned}
$$

from which we see that condition (ii) of Theorem 2.9 is satisfied with $\lambda = r - \delta$. Theorem 2.9 now tells us that G_{cont} is nonexplosive and (2.31) holds. ∎

In particular, Proposition 2.26 tells us that if $r < \delta$, then the contact process *dies out* in the sense that

$$
\mathbb{P}^x\big[X_t = \underline{0}\big] \underset{t \to \infty}{\longrightarrow} 1 \qquad \forall x \in \mathcal{S}_{\text{fin}}(\Lambda).
$$

This is quite natural since r, defined in (2.30), is the maximal reproduction rate of an individual (assuming all other sites are vacant). If this is less than the death rate, then each individual produces on average less than one offspring before it dies, leading to an exponential decay of the population size.

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 *Curie-Weiss model*, 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

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

and jumps

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

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

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

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

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

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. (This argument will be made rigorous in Section 3.4 below.) Cancelling a common factor $e^{\beta/2}$ in the nominator and denominator of the rates, we can simplify our formulas a bit to

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

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

$$(3.2)$$

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

$$\mathbb{E}^{\overline{x}}[\overline{X}_t] = \overline{x} + t g_\beta(\overline{x}) + O(t^2) \quad \text{where} \quad g_\beta(\overline{x}) := \overline{G}_{N,\beta} f(\overline{x}) \quad \text{with} \quad f(\overline{x}) := \overline{x}.$$
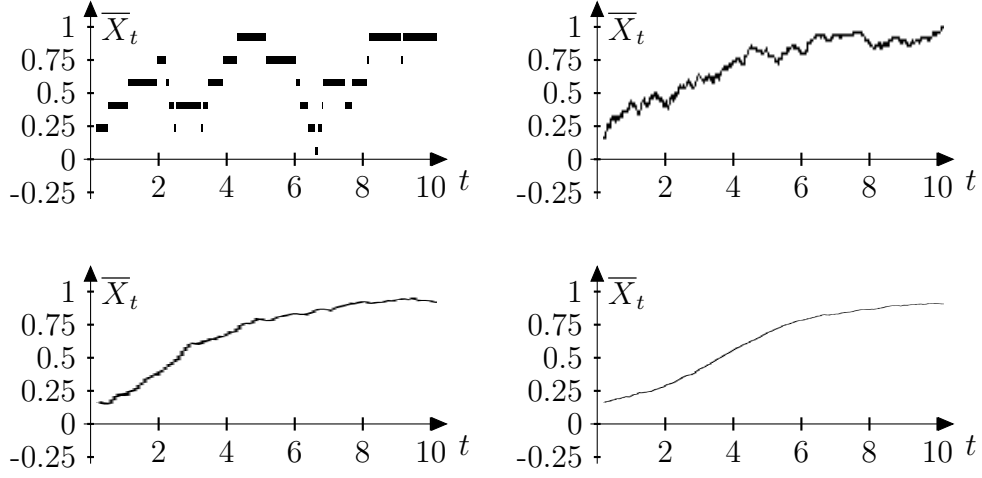
Figure 3.1: The mean-field Ising model on lattice with $N = 10$, 100, 1000, and $10,000$ sites, respectively. In these simulations, the parameter is $\beta = 3$, and the initial state is $\overline{X}_0 = 0.1$, except in the first picture, where $\overline{X}_0 = 0.2$.

We calculate

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

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 $\beta > 2$. Here, with a *fixed point* of the differential equation,

we mean a point z such that $\overline{x}_0 = z$ implies $\overline{x}_t = z$ for all $t \geq 0$, i.e., this is a point such that $g_\beta(z) = 0$.
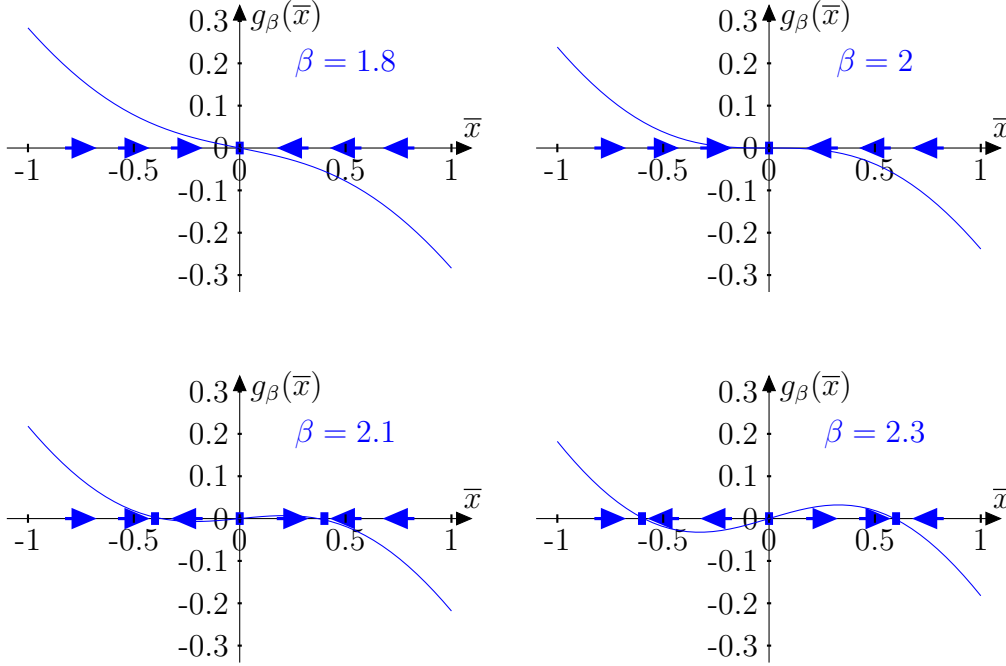
Figure 3.2: The drift function g_β for $\beta = 1.8$, 2, 2.1, and 2.3, respectively. For $\beta > 2$, the fixed point $\overline{x} = 0$ becomes unstable and two new fixed points appear.

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

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

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.

plays a similar role as the spontaneous magnetization $m_*(\beta)$ for the Ising model on $\mathbb{Z}^d$ (see formula (1.15)). More precisely, for mean-field processes started in initial states $\overline{X}_0 > 0$, the quantity x_{upp} describes the double limit
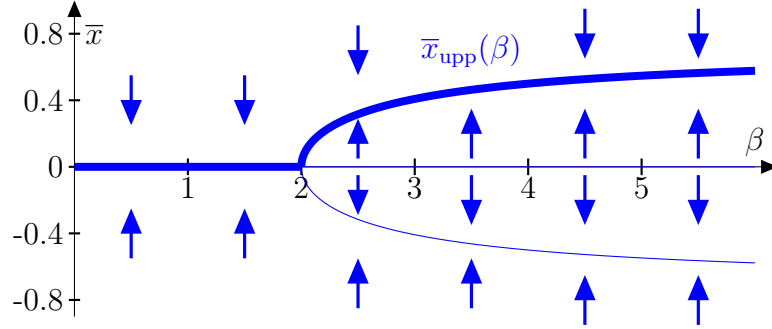
$$\lim_{t \to \infty} \lim_{N \to \infty} \overline{X}_t = x_{\mathrm{upp}}. \tag{3.6}$$

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 $\beta_{\mathrm{c}} = 2$. Since

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

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,

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

---

[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.

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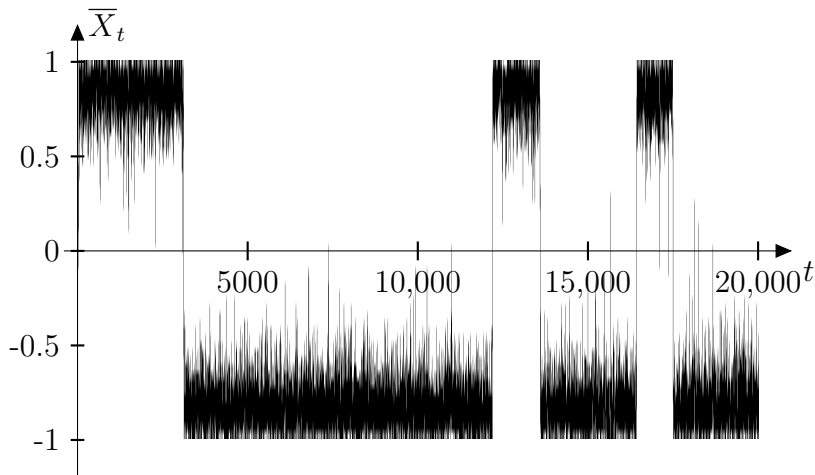
Figure 3.4: Metastable behavior of a mean-field Ising model with $N = 50$ and $\beta = 3$. Note the different time scale compared to Figure 3.1.

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_{upp} and $-x_{\mathrm{upp}}$, 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_{upp} and $-x_{\mathrm{upp}}$ 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

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

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

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

is itself a Markov process.

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

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

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. Conversely, if for each initial law of the process X, it is true that Y is a Markov process with generator H, then (3.9) must hold.

Proof of Proposition 3.1 Let us define $\mathcal{H}(x, y')$ as in (3.7) also when $f(x) = y'$. We start by noting that if (3.9) holds for all $x \in S$ and $y' \in T$ such that $f(x) \neq y'$, then it also holds when $f(x) = y'$. To see this, we write

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

$$= - \sum_{x':\, f(x') \neq f(x)} G(x, x') = \sum_{x':\, f(x')=f(x)} G(x, x'),$$

where we have used that since H and G are Markov generators, one has $\sum_{y' \in T} H(f(x), y') = 0$ and $\sum_{x' \in S} G(x, x') = 0$. We have thus shown that (3.9) is equivalent to

$$H\big(f(x), y'\big) = \sum_{x':\, f(x')=y'} G(x, x') \qquad \big(x \in S,\ y' \in T\big). \qquad (3.10)$$

We claim that this is equivalent to

$$Q_t\big(f(x), y'\big) = \sum_{x': f(x')=y'} P_t(x, x') \qquad \big(t \geq 0, \ x \in S, \ y' \in T\big), \qquad (3.11)$$

where $(Q_t)_{t\geq 0}$ is the semigroup generated by H. To prove this, we start by observing that for any function $g : T \to \mathbb{R}$,

$$G(g \circ f)(x) = \sum_{x'} G(x, x')g(f(x')) = \sum_{y'} \sum_{x': f(x')=y'} G(x, x')g(y'),$$

$$(Hg) \circ f(x) = \sum_{y'} H(f(x), y')g(y').$$

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

$$G(g \circ f) = (Hg) \circ f \qquad (g : T \to \mathbb{R}). \qquad (3.12)$$

By exactly the same argument with G replaced by P_t and H replaced by Q_t, we see that (3.11) is equivalent to

$$P_t(g \circ f) = (Q_t g) \circ f \qquad (t \geq 0, \ g : T \to \mathbb{R}). \qquad (3.13)$$

To see that (3.12) and (3.13) are equivalent, we write

$$P_t = e^{Gt} = \sum_{n=0}^{\infty} \frac{1}{n!} t^n G^n \quad \text{and} \quad Q_t = e^{Ht} = \sum_{n=0}^{\infty} \frac{1}{n!} t^n H^n. \qquad (3.14)$$

We observe that (3.12) implies

$$G^2(g \circ f) = G\big((Hg) \circ f\big) = (H^2 g) \circ f,$$

and similarly, by induction, $G^n(g \circ f) = (H^n g) \circ f$ for all $n \geq 0$, which by (3.14) implies (3.13). Conversely, if (3.13) holds for all $t \geq 0$, then it must hold up to first order in t as $t \downarrow 0$, which implies (3.12). This completes the proof that (3.9) is equivalent to (3.11).

If (3.11) holds, then by (2.5), the finite dimensional distributions of Y are given by

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

$$(3.15)$$

$(0 = t_0 < \cdots < t_n)$. Again by (2.5), this implies that Y is a Markov process with generator H.

Conversely, if Y is a Markov process with generator H for each initial state of X, then for each $x_0 \in S$, (3.15) must hold when $X_0 = x_0$ a.s. and for $n = 1$, from which we see that (3.11) and hence (3.9) hold.   ∎

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).

Remark For the final statement of the proposition, it is essential that Y is a Markov process for *each* initial law X. There exist interesting examples of functions of Markov processes that are not autonomous Markov processes, but nonetheless are Markov processes for some *special* initial laws of the original Markov process. This is closely related to the concept of intertwining of Markov processes that will briefly be mentioned in Section 6.5 below.

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 (3.19) below. An alternative, more probabilistic approach to mean-field equations is described in [MSS20].

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

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

The idea is that if α tends to zero and β approximates a nice, Lipschitz continuous 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

$$\sup_{x \in S_N} \big| \beta_N(x) - b\big(f_N(x)\big) \big| \xrightarrow[N \to \infty]{} 0, \tag{3.16}$$

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

$$b(I_-) \geq 0 \quad \text{if } I_- > -\infty \quad \text{and} \quad b(I_+) \leq 0 \quad \text{if } I_+ < -\infty, \tag{3.17}$$

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 as well as (3.16), one moreover has*

$$\sup_{x \in S_N} \alpha_N(x) \xrightarrow[N \to \infty]{} 0. \tag{3.18}$$

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

$$\mathbb{P}\big[|f_N(X_t^N) - y_t| \leq \varepsilon \ \forall t \in [0, T] \big] \xrightarrow[N \to \infty]{} 1. \tag{3.19}$$

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 $f_N(x)$ is not further than ε from a possible value the solution of the differential equation. In our case, these more general assumptions are automatically satisfied. Set $\delta := \tfrac{1}{3} \varepsilon e^{-LT}$. 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

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

Then [DN08, Thm 4.1] tells us that

$$\mathbb{P}\Big[\sup_{t \in [0,T]} |f(X_t) - y_t| > \varepsilon \Big] \le 4KT\delta^{-2} + \mathbb{P}\big(\Omega_0^{\mathrm{c}} \cup \Omega_1^{\mathrm{c}} \cup \Omega_{K,2}^{\mathrm{c}} \big). \qquad (3.20)$$

Our assumption that $f_N(X_0^N) \to y_0$ in probability implies that $\mathbb{P}(\Omega_0^{\mathrm{c}}) \to 0$ 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 (3.18) and $B_N \to 0$ by (3.16). Since

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

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

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

we see that $\mathbb{P}(\Omega_{A_N,2}^{\mathrm{c}}) = 0$ for all N. Inserting $K = A_N$ in (3.20), 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 $X^N = (X_t^N)_{t \ge 0}$ 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 (3.16) is already verified in (3.3). To check also (3.18), 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

$$
\begin{aligned}
G_{\mathrm{cont}} f(x) := & \lambda \sum_{i \in \mathbb{Z}^d} \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} \big\{ f\big((\mathtt{bra}_{ij}(x))\big) - f(x) \big\} \\
& + \sum_{i \in \mathbb{Z}^d} \big\{ f\big((\mathtt{death}_i(x))\big) - f(x) \big\} \qquad (x \in \{0,1\}^\Lambda),
\end{aligned}
\tag{3.21}
$$

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 $E = E_N := \{\{i,j\} : i,j \in \Lambda_N\}$ and corresponding set of oriented edges $\mathcal{E} = \mathcal{E}_N$. We will be interested in the fraction of infected sites

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

which jumps with the following rates

$$
\begin{aligned}
\overline{x} \mapsto \overline{x} + \tfrac{1}{N} \qquad & \text{with rate} \qquad r_+(\overline{x}) := \lambda N \overline{x}(1 - \overline{x}), \\
\overline{x} \mapsto \overline{x} - \tfrac{1}{N} \qquad & \text{with rate} \qquad r_-(\overline{x}) := N \overline{x}.
\end{aligned}
\tag{3.22}
$$

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 are 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.

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

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

$$
\tfrac{\partial}{\partial t} \overline{X}_t = b_\lambda(\overline{X}_t) \quad (t \geq 0), \quad \text{where} \quad b_\lambda(\overline{x}) := \lambda \overline{x}(1 - \overline{x}) - \overline{x}.
\tag{3.23}
$$

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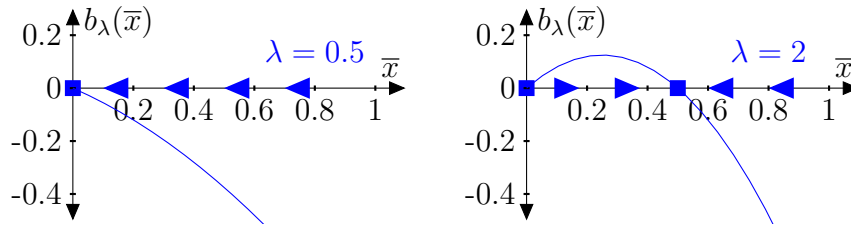
We see from this that the fixed point $\overline{x} = 0$ is stable for $\lambda \leq 1$ but becomes unstable for $\lambda > 1$, in which case $\overline{x} = 1 - \lambda^{-1}$ is the only stable fixed point that attracts all solutions started in a nonzero initial state. The situation is summarized in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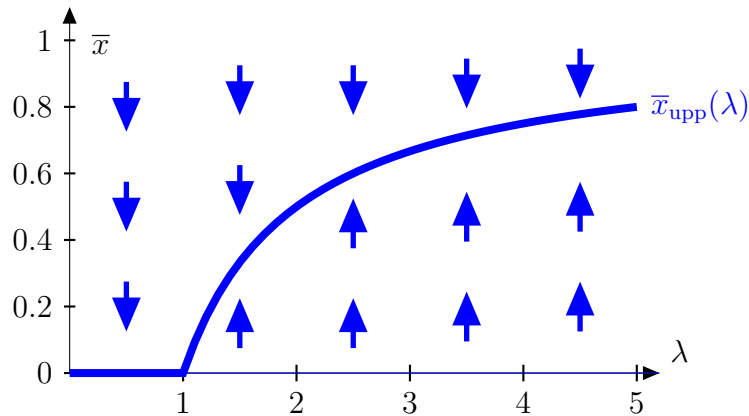
Figure 3.5: Mean-field analysis of the contact process.

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

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

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

$$G_{\text{vot}}f(x) = \frac{1}{|\Lambda|} \sum_{(i,j) \in \mathcal{E}} \left\{ f((\text{vot}_{ij}(x)) - f(x) \right\} \qquad (x \in \{0, 1\}^\Lambda). \qquad (3.24)$$

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 (3.22))

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

$$\overline{x} \mapsto \overline{x} - \tfrac{1}{N} \qquad \text{with rate} \qquad r_-(\overline{x}) := N\overline{x}(1 - \overline{x}).$$

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.

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

$$\beta_N(x) = r_+(\overline{x})\tfrac{1}{N} - r_-(\overline{x})\tfrac{1}{N} = 0.$$

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.

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

$$G_{\text{vot}}f(x) = \sum_{(i,j) \in \mathcal{E}} \left\{ f((\text{vot}_{ij}(x)) - f(x) \right\} \qquad (x \in \{0, 1\}^\Lambda), \qquad (3.25)$$
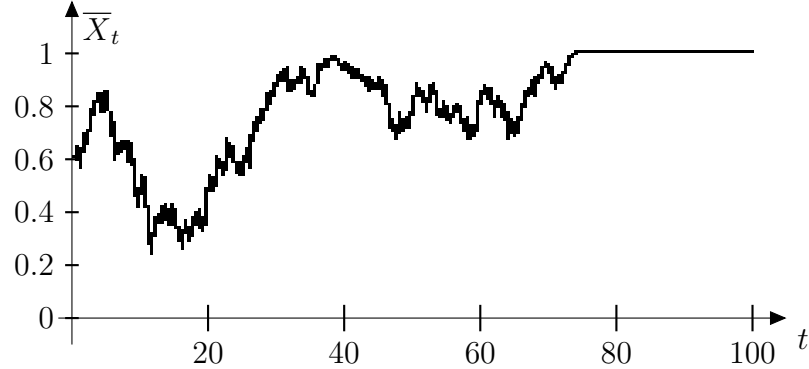
Figure 3.6: The fraction of type 1 individuals in the mean-field voter model from (3.25) on $N = 100$ sites. This process approximates the Wright-Fisher diffusion.

then the drift and quadractic variation are given by

$$\alpha_N(x) = 2\overline{x}(1 - \overline{x}),$$

$$\beta_N(x) = 0.$$

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

$$Gf(\overline{x}) = \overline{x}(1 - \overline{x})\tfrac{\partial^2}{\partial \overline{x}^2}f(\overline{x}). \tag{3.26}$$

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 (3.25) and $N = 100$.

3.7 Exercises

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

$$Gf(x) = b|\Lambda|^{-2} \sum_{ii'j} \left\{ f\left(\texttt{coop}_{ii'j}x\right) - f(x) \right\}$$
$$+ \sum_{i} \left\{ f\left(\texttt{death}_{i}x\right) - f(x) \right\},$$

where the maps $\texttt{coop}_{ii'j}$ *and* $\texttt{death}_{i}$ *are defined in (1.25) and (1.7), respectively. Do you observe a phase transition? Is it first- or second order? Hint: Figure 3.7.*
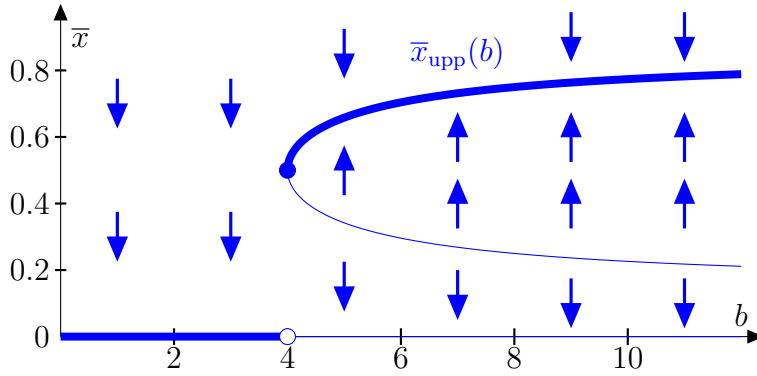
Figure 3.7: Mean-field analysis of a model with cooperative branching and deaths.

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

$$Gf(x) = b|\Lambda|^{-2} \sum_{ii'j} \left\{ f\left(\texttt{coop}_{ii'j}x\right) - f(x) \right\}$$
$$+ |\Lambda|^{-1} \sum_{ij} \left\{ f\left(\texttt{rw}_{ij}x\right) - f(x) \right\}.$$

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*

$$Gf(x) = \sum_{ij \in \Lambda} \left\{ f\left(\texttt{vot}_{ij}x\right) - f(x) \right\}$$
$$+ s|\Lambda|^{-1} \sum_{ij \in \Lambda} \left\{ f\left(\texttt{bra}_{ij}x\right) - f(x) \right\}$$
$$+ d \sum_{i \in \Lambda} \left\{ f\left(\texttt{death}_{i}x\right) - f(x) \right\}.$$

Hint: *You should find expressions of the form*

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

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*

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

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

Throughout this chapter, S is a finite set called the *local state space*, Λ is a countable set called the *lattice*, $\mathcal{G}$ is a countable collection of local maps $m : S^\Lambda \to S^\Lambda$, and $(r_m)_{m \in \mathcal{G}}$ are nonnegative rates. Our aim is to construct a Markov process with state space S^Λ and formal generator of the form

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

Recall that $\mathcal{D}(m)$, defined in Section 2.7, is the set of lattice points whose values can be changed by m. Under the assumption

$$\sum_{m : \, \mathcal{D}(m) \ni i} r_m < \infty \quad (i \in \Lambda)$$

one can show that Gf is well-defined for all functions $f : S^\Lambda \to \mathbb{R}$ that depend on finitely many coordinates (see Lemma 4.23 below). In general, we will need stronger conditions on the rates $(r_m)_{m \in \mathcal{G}}$ to ensure that G generates the Markov process. Most of the interacting particle systems introduced in Chapter 1 have a generator that can naturally be written in the form (4.1). The only processes for which this is not so obvious are the stochastic Ising and Potts models. Nevertheless, we will show in Section 4.6 below that also the generator of the stochastic Ising model with Glauber dynamics can be written in the form (4.1), and we will use this random mapping representation of G to prove ergodicity for small values of β.

    The space S^Λ is uncountable except in the trivial case that S has cardinality one. This means that we cannot use the theory of continuous-time Markov chains. Instead, we will rely on the theory of Feller processes. 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 analytic 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 Poisson construction of interacting particle systems (including proofs) that is similar to the Poisson construction of continuous-time Markov chains. To some degree, this provides a probabilistic alternative to the functional analytic approach via Feller processes. To get the full picture, however, one needs both approaches, so in Section 4.4, we show that our Poisson construction yields a Feller process and determine its generator.

Luckily, all this abstract theory gives us more than 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. In Section 4.6 we apply this to derive lower bounds on the critical points of the Ising model.

4.2   Feller processes

In Section 2.1, we showed how the semigroup of a Markov process on a finite state space can be characterised in terms of its generator and in Section 2.3 we generalized this to countable state spaces. In the present section, we will treat a class of Markov processes with compact metrizable state spaces. 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*. We will later apply the theory of Feller processes to state spaces of the form S^Λ equipped with the product topology, which are compact by Tychonoff's theorem. It is easy to see that the product topology on S^Λ 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.

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

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

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,[2] 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 E 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 with which we have equipped these spaces). A probability kernel is *deterministic* if it is of the form $K(x, \cdot) = \delta_{m(x)}$ for some measurable map $m : E \to E$, where $\delta_{m(x)}$ denotes the delta-measure at $m(x)$. It is easy to see that a deterministic kernel is continuous if and only if m is a continuous map. A *random mapping representation* of a probability kernel K is a random measurable map[3] $M : E \to E$ such that $K(x, \cdot) = \mathbb{P}[M(x) \in \cdot]$ $(x \in E)$.[4]

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

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

---

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

[2]More precisely, the topology of weak convergence is the unique *metrizable* topology with this property. Since in metrizable spaces, convergent subsequences uniquely characterize the topology, such a definition is unambiguous.

[3]More formally, this means that $M : \Omega \times E \to E$ is measurable with respect to the product-σ-field $\mathcal{F} \otimes \mathcal{B}(E)$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is the underlying probability space.

[4]For infinite spaces, it is not so clear if every probability kernel has a random mapping representation. One could also ask if every continuous probability kernel has a representation in terms of continuous maps. Although these questions are interesting, we will neglect them here.

defines a linear operator $K : B(E) \to B(E)$. We define the *composition* of two probability kernels K, L as

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

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 $(\mu K)f = \mu(Kf)$ 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 $P_s P_t = P_{s+t}$ $(s, t \geq 0)$.

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

(i) $\lim_{t \to 0} P_t f = P_0 f = f$ $(f \in \mathcal{C}(E))$,

(ii) $P_s P_t f = P_{s+t} f$ $(s, t \geq 0)$,

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

(iv) $P_t 1 = 1$,

and conversely, each collection of linear operators $P_t : \mathcal{C}(E) \to \mathcal{C}(E)$ with these properties corresponds to a unique continuous transition probability on E. Such a collection of linear operators $P_t : \mathcal{C}(E) \to \mathcal{C}(E)$ is called a *Feller semigroup*. 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,[5] i.e.,

(i) $\lim_{t \downarrow s} w_t = w_s$ $(s \geq 0)$,

(ii) $\lim_{t \uparrow s} w_t =: w_{s-}$ exists $(s > 0)$.

---

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

Let $(P_t)_{t \geq 0}$ be a Feller semigroup. By definition a *Feller process* with semigroup $(P_t)_{t \geq 0}$ is a stochastic process $X = (X_t)_{t \geq 0}$ with cadlag sample paths[6] 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

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

$(0 < t_1 < \cdots < t_n)$. 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 exponential 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.

---

[6]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 $(X_t)_{t \geq 0}$.

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

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

with domain

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

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. The following lemma says that generators are closed, densely defined operators.

Lemma 4.3 (Elementary properties of generators) *Let G be the generator of a Feller semigroup $(P_t)_{t \geq 0}$. Then G is closed and $\mathcal{D}(G)$ is a dense subspace of $\mathcal{C}(E)$.*

Since G is closed, Theorem 4.2 tells us that G is everywhere defined (i.e., $\mathcal{D}(G) = \mathcal{C}(G)$) if and only if G is bounded. For bounded generators, it is not hard to show that the exponential formula $e^{tG} := \sum_{n=0}^{\infty} \frac{1}{n!}(tG)^n$ converges in the norm on $\mathcal{C}(G)$ and that the Feller semigroup with generator G is given by $P_t = e^{tG}$. On the other hand, if G is unbounded, then it is in general not possible to make sense of the exponential formula.[7] In the context of interacting particle systems, it is not hard to show that a generator of the form (4.1) is bounded if $\sum_{m \in \mathcal{G}} r_m < \infty$. For the particle systems we will be interested in, this sum will usually be infinite and the generator will be unbounded.

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. Similar to what we did in Section 2.3, we will use the backward equation instead. 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*

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

if $u_t \in \mathcal{D}(A)$ for all $t \geq 0$, the maps $t \mapsto u_t$ and $t \mapsto A u_t$ 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 4.4 (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*

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

---

[7] In order for $\sum_{n=0}^{\infty} \frac{1}{n!} t^n G^n f$ to make sense, we need that $G^n f$ is well-defined for all $n \geq 0$. For $n = 1$ this already requires that $f \in \mathcal{D}(G)$ but for higher n we need even more since it is in general not true that G maps $\mathcal{D}(G)$ into itself. Thus, it is not even clear for which class of functions we can make sense of each term in the expansion seperately.

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

Lemma 4.5 (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 4.6 (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. In general, if G is a closed linear operator and $\mathcal{D}' \subset \mathcal{D}(G)$, then we let $G|_{\mathcal{D}'}$ denote the restriction of G to $\mathcal{D}'$, i.e., $G|_{\mathcal{D}'}$ is the linear operator with domain $\mathcal{D}(G|_{\mathcal{D}'}) := \mathcal{D}'$ defined as $G|_{\mathcal{D}'} f := Gf$ for all $f \in \mathcal{D}'$. We say that $\mathcal{D}'$ is a *core* for G if $\overline{G|_{\mathcal{D}'}} = G$.

Lemma 4.7 (Core of a generator) *Let G be the generator of a Feller semigroup and let $\mathcal{D}' \subset \mathcal{D}(G)$. Assume that $\mathcal{D}'$ is dense in $\mathcal{C}(E)$. Then the following conditions are equivalent:*

 (i) $\mathcal{D}'$ *is a core for* G,

 (ii) *the set $\{(\lambda - A)p : p \in \mathcal{D}'\}$ is dense in $\mathcal{C}(E)$ for some $\lambda > 0$,*

 (iii) *the set $\{(\lambda - A)p : p \in \mathcal{D}'\}$ is dense in $\mathcal{C}(E)$ for all $\lambda > 0$.*

Note that by condition (ii) of Lemma 4.7, to check that a dense set $\mathcal{D}' \subset \mathcal{C}(E)$ is a core for G, it suffices to show that for some $\lambda > 0$, there exists a dense subspace $\mathcal{R} \subset \mathcal{C}(E)$ such that for every $f \in \mathcal{R}$, the Laplace equation $(\lambda - A)p = f$ has a solution $p \in \mathcal{D}'$. Using Lemma 4.7, one can prove the following alternative version of the Hille-Yosida theorem.

Theorem 4.8 (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 $(\lambda - A)p = f$ 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 4.8, 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

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

One can show that A satisfies the conditions of Theorem 4.8 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 (3.26).

Exercise 4.9 (Brownian motion) *Let $(P_t)_{t \geq 0}$ denote the transition kernels of Brownian motion on $\mathbb{R}^d$. Let $E := \mathbb{R}^d \cup \{\infty\}$ denote the one-point compactification of $\mathbb{R}^d$ and extend P_t $(t \geq 0)$ to probability kernels on E by setting $P_t(\infty, \cdot) := \delta_\infty$. Show that $(P_t)_{t \geq 0}$ is a Feller semigroup.*

In Chapter 2, we viewed (possibly explosive) continuous-time Markov chains with a countable state space S as Markov processes on the extended state space S_∞, where S_∞ is the one-point compactification of S. It is natural to ask if they are in fact Feller processes on S_∞. The answer is, in general, negative. The reason is that the extended transition kernels $(\overline{P}_t)_{t \geq 0}$ on S_∞ may fail to be continuous at ∞, i.e., $\overline{P}_t(x_n, \cdot)$ may fail to converge to $\overline{P}_t(\infty, \cdot)$ if $x_n \to \infty$. In many cases where this problem occurs, it can be solved by choosing another compactification of S (i.e., by adding more points at infinity). Whether this can be done in general I don't know.

Exercise 4.10 (Wright-Fisher diffusion) *Show that the operator A defined in (4.7) satisfies the conditions of Theorem 4.8. Hint: show that if f is a polynomial of order n, then so is Af. Use this to show that the Cauchy equation $\frac{\partial}{\partial t}u_t = Au_t$ has a solution for each initial state $u_0 = f$ that is a polynomial. Then show that $p := \int_0^\infty u_t\, e^{-\lambda t}\, \mathrm{d}t$ solves the Laplace equation $(\lambda - A)p = f$.*

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 in many places (including Wikipedia).

Lemma 4.3 follows from [EK86, Corollary I.1.6]. The statements of this lemma can also easily be derived from the Hille-Yosida theorem (see below). Proposition 4.4 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 4.11 (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 $\frac{\partial}{\partial t} u_t = A u_t$ ($t \geq 0$). 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,

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

so we arrive at a contradiction. ∎

By linearity, Lemma 4.11 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 4.6, Feller generators satisfy the positive maximum principle, this implies uniqueness of solutions of the Cauchy equation in Proposition 4.4. Again by Theorem 4.6, the domain of a Feller semigroup is a dense subspace of of $\mathcal{C}(E)$, so the final statement of Proposition 4.4 follows from the following simple lemma and the fact that $\|P_t f\|_\infty \leq \|f\|_\infty$.

Lemma 4.12 (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 $\overline{A}f :=$

$\lim_{n \to \infty} A f_n$ 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 $\|A f_n - A g_n\| \leq C \|f_n - g_n\| \to 0$. The fact that $\|\overline{A} f\| \leq C \|f\|$ ($f \in \mathcal{V}$) follows from the continuity of the norm. ∎

Lemma 4.5 follows from [EK86, Prop 1.2.1]. Theorems 4.6 and 4.8 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 4.4 that condition (i) in Theorems 4.6 and 4.8 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 4.6 and 4.8 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 4.6 and 4.8 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 4.6 and 4.8 correspond to [EK86, Thms 1.2.6 and 1.2.12]. Lemma 4.7 follows from [EK86, Lemma 1.2.11 and Prop 1.3.1]. In the more specific set-up of Feller semigroups, versions of Theorem 4.8 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 4.8 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 $x = (x(i))_{i \in \Lambda}$. We fix a countable set $\mathcal{G}$ whose elements are local maps $m : S^\Lambda \to S^\Lambda$ as well as nonnegative rates $(r_m)_{m \in \mathcal{G}}$. Our aim is to construct a Markov process with formal generator of the form (4.1). Our approach is very similar to the Poisson construction of continuous-time Markov chains described in Section 2.6. We equip the space $\mathcal{G} \times \mathbb{R}$ with the measure

$$\rho\big(\{m\} \times A\big) := r_m \, \ell(A) \qquad \big(m \in \mathcal{G}, \ A \in \mathcal{B}(\mathbb{R})\big), \tag{4.8}$$

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. We call ω the *graphical representation* associated with the random mapping representation (2.14). Since $\mathcal{G}$ is countable, by the argument used in Section 2.6, the time coordinates of points $(m, t) \in \omega$ are all different. Therefore, as we did in the case of continuous-time Markov chains, we can unambiguously define a random function $\mathbb{R} \ni t \mapsto \mathfrak{m}_t^\omega$ by setting

$$\mathfrak{m}_t^\omega := \begin{cases} m & \text{if } (m, t) \in \omega, \\ 1 & \text{otherwise,} \end{cases} \tag{4.9}$$

where we write 1 to denote the identity map. By definition, we say that a random function $X : [s, \infty) \to S^\Lambda$ solves the evolution equation

$$X_t = \mathfrak{m}_t^\omega(X_{t-}) \qquad (t > s), \tag{4.10}$$

if $[s, \infty) \ni t \mapsto X_t \in S^\Lambda$ is cadlag and (4.10) holds. We recall that for any local map m and site $i \in \Lambda$, the set $\mathcal{R}_i^\downarrow(m)$ has been defined in Section 2.8. Here is the main result of the present section.

Theorem 4.13 (Poisson construction) *Assume that the rates* $(r_m)_{m \in \mathcal{G}}$ *satisfy*

$$\sup_{i \in \Lambda} \sum_{m \in \mathcal{G}} r_m \big(1_{\mathcal{D}(m)}(i) + |\mathcal{R}_i^\downarrow(m)|\big) < \infty. \tag{4.11}$$

Then almost surely, for each $s \in \mathbb{R}$ *and* $x \in S^\Lambda$*, there exists a unique solution* $(X_t^{s,x})_{t \geq s}$ *to the evolution equation (4.10) with initial state* $X_s^{s,x} = x$*. Setting*

$$\mathbf{X}_{s,t}(x) := X_t^{s,x} \qquad (s \leq t, \ x \in S^\Lambda) \tag{4.12}$$

defines a collection of continuous maps $(\mathbf{X}_{s,t})_{s \leq t}$ *from* S^Λ *into itself such that*

$$\mathbf{X}_{s,s} = 1 \quad and \quad \mathbf{X}_{t,u} \circ \mathbf{X}_{s,t} = \mathbf{X}_{s,u} \qquad (s \leq t \leq u). \qquad (4.13)$$

Setting

$$P_t(x, \cdot) := \mathbb{P}\big[\mathbf{X}_{0,t}(x) \in \cdot\big] \qquad (t \geq 0, \ x \in S^\Lambda) \qquad (4.14)$$

defines the semigroup of a Feller process with state space S^Λ. *If* $s \in \mathbb{R}$ *and* X_0 *is an* S^Λ-*valued random variable with law* μ, *independent of* ω, *then the process* $(X_t)_{t \geq 0}$ *defined as*

$$X_t := \mathbf{X}_{s,s+t}(X_0) \qquad (t \geq 0) \qquad (4.15)$$

is distributed as the Feller process with semigroup $(P_t)_{t \geq 0}$ *and initial law* μ.

The proof of Theorem 4.13 will take up the rest of this section. At first, it may be surprising that solutions of the evolution equation (4.10) with a given initial state are unique. After all, if we replace the compact set S^Λ by $[0, 1]$, then there are many cadlag functions $X : [s, \infty) \to [0, 1]$ with a given initial state that make no jumps at all. The following exercise shows that at least in the case when $\omega = \emptyset$, the equation (4.10) has a unique solution.

Exercise 4.14 (Total disconnectedness) *A topological space* E *is* totally disconnected *if for each* $x_1, x_2 \in E$ *with* $x_1 \neq x_2$, *there exist open sets* $O_1 \ni x_1$ *and* $O_2 \ni x_2$ *such that* $O_1 \cap O_2 = \emptyset$ *and* $O_1 \cup O_2 = E$. *Prove that* S^Λ *is totally disconnected. Prove that if* E *is a totally disconnected space, then each continuous function* $f : [0, \infty) \to E$ *is constant.*

The difficulty with proving existence and uniqueness of solutions of (4.10) 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 ω according to their times, and solutions $(X_t)_{t \geq s}$ to (4.10) will not be piecewise constant. However, most of the jumps of $(X_t)_{t \geq s}$ will involve sites that are far away, and $t \mapsto X_t(i)$ will still be piecewise constant for each fixed $i \in \Lambda$.

The trick for proving uniqueness of solutions of (4.10) is to look backwards in time. One may notice the similarity between condition (4.11) and the condition (2.26) from Chapter 2, the only difference being that $\mathcal{R}_i^\uparrow(m)$ is replaced by $\mathcal{R}_i^\downarrow(m)$, which looks "downwards" in time. (Here we use our usual convention of plotting time upwards in pictures of graphical representations so that downwards means back in time.) We recall from Lemma 2.18 that if T is a finite set and $f : S^\Lambda \to T$ is a continuous function, then f depends on finitely many coordinates. A consequence of this is that the space $\mathcal{C}(S^\Lambda, T)$

of continous functions $\phi : S^\Lambda \to T$ is countable. It turns out that if we fix $\phi \in \mathcal{C}(S^\Lambda, T)$ and a time $u \in \mathbb{R}$, then the stochastic flow $(\mathbf{X}_{s,t})_{s \leq t}$ that we are about to construct has the property that the process

$$\Phi_t := \phi \circ \mathbf{X}_{u-t,u} \qquad (t \geq 0)$$

is a nonexplosive continuous-time Markov chain with countable state space $\mathcal{C}(S^\Lambda, T)$. We call $(\Phi_t)_{t \geq 0}$ the *backward in time process.* The first step towards proving Theorem 4.13 is showing that this continuous-time Markov chain is nonexplosive.

Proposition 4.15 (Backward in time process) *Assume (4.11). Let T be a finite set. Then there exists a generator H of a nonexplosive continuous-time Markov chain with state space $\mathcal{C}(S^\Lambda, T)$ such that*

$$Hf(\phi) = \sum_{m \in \mathcal{G}} r_m \{ f(\phi \circ m) - f(\phi) \} \qquad (4.16)$$

for all bounded $f : \mathcal{C}(S^\Lambda, T) \to [0, \infty)$. Almost surely, for each $u \in \mathbb{R}$ and $\phi \in \mathcal{C}(S^\Lambda, T)$, there exists a unique cadlag function $\Phi^{u,\phi} : (-\infty, u] \to \mathcal{C}(S^\Lambda, T)$ such that $\Phi_u^{u,\phi} = \phi$ and

$$\Phi_{t-}^{u,\phi} = \begin{cases} \Phi_t^{u,\phi} \circ m & \text{if } (m, t) \in \omega, \\ \Phi_t^{u,\phi} & \text{otherwise} \end{cases} \qquad (t \leq u). \qquad (4.17)$$

One has

$$\mathbb{E}\big[|\mathcal{R}(\Phi_t^{u,\phi})| \big] \leq |\mathcal{R}(\phi)| e^{-K(u-t)} \qquad \big(t \leq u, \ \phi \in \mathcal{C}(S^\Lambda, T) \big), \qquad (4.18)$$

where

$$K := \sup_{i \in \Lambda} \sum_{m \in \mathcal{G}} r_m \big(|\mathcal{R}(m[i])| - 1 \big) < \infty. \qquad (4.19)$$

Proof The general idea of the proof is illustrated in Figure 4.1. The function $\Phi_t^{u,\phi}$, that will in (4.30) below be shown to be equal to $\phi \circ \mathbf{X}_{t,u}$, evolves backwards in time in such a way that for each point $(m, t) \in \omega$ of the graphical representation, at time t, the function $\Phi_t^{u,\phi}$ is replaced by $\Phi_t^{u,\phi} \circ m$. We will show that the size of the set of relevant points $\mathcal{R}(\Phi_t^{u,\phi})$ grows at most exponentially in time, so that $\Phi_t^{u,\phi}$ corresponds to a nonexplosive continuous-time Markov chain.

To see that H is the generator of a (possibly explosive) continuous-time Markov chain with state space $\mathcal{C}(S^\Lambda, T)$, we must check condition (2.13), which in the present context says that

$$c(\phi) := \sum_{m : \phi \circ m \neq \phi} r_m < \infty \quad \text{for all } \phi \in \mathcal{C}(S^\Lambda, T).$$
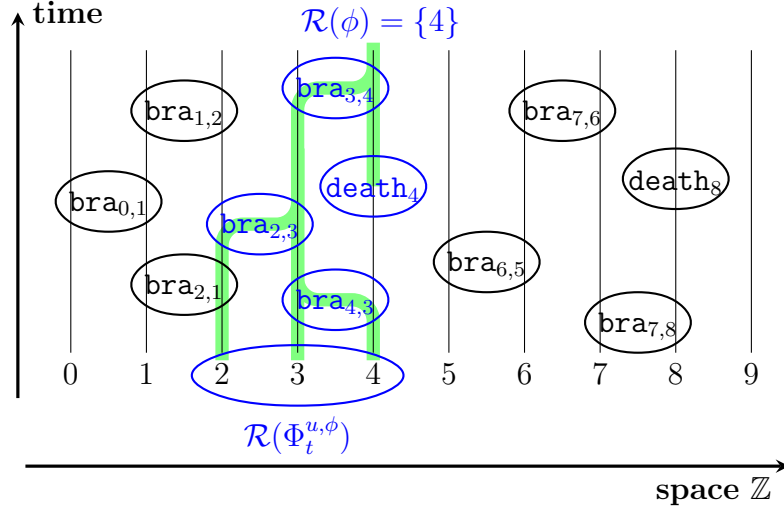
Figure 4.1: Graphical representation of a one-dimensional contact process, with the backwards in time process $(\Phi_t^{u,\phi})_{t\in(-\infty,u]}$ started in the initial condition $\phi(x) := x(4)$. The green lines show the evolution backwards in time of the set $\mathcal{R}(\Phi_t^{u,\phi})$ of relevant lattice points for the map $\Phi_t^{u,\phi} = \phi \circ \mathbf{X}_{t,u}$.

We observe that $\phi(m(x)) \neq \phi(x)$ for some $x \in S^\Lambda$ implies that $\mathcal{D}(m) \cap \mathcal{R}(\phi) \neq \emptyset$, so we can estimate

$$\sum_{m:\,\phi\circ m\neq\phi} r_m \leq \sum_{m:\,\mathcal{D}(m)\cap\mathcal{R}(\phi)\neq\emptyset} r_m \leq \sum_{i\in\mathcal{R}(\phi)}\sum_m 1_{\mathcal{D}(m)}(i)r_m,$$

which is finite by (4.11) and the finiteness of $\mathcal{R}(\phi)$. To prove that H is nonexplosive we apply Theorem 2.9 to the Lyapunov function

$$L(\phi) := |\mathcal{R}(\phi)| \qquad \big(\phi \in \mathcal{C}(S^\Lambda, T)\big).$$

Our previous calculation shows that L satisfies condition (i) of Theorem 2.9, so it remains to check condition (ii). We have

$$HL(\phi) = \sum_{m\in\mathcal{G}} r_m\big\{L(\phi\circ m) - L(\phi)\big\} = \sum_{m\in\mathcal{G}} r_m\big\{|\mathcal{R}(\phi\circ m)| - |\mathcal{R}(\phi)|\big\}.$$

Since

$$\mathcal{R}(\phi\circ m) \subset \bigcup_{j\in\mathcal{R}(\phi)} \mathcal{R}(m[j]) = \mathcal{R}(\phi)\backslash\mathcal{D}(m) \cup \bigcup_{j\in\mathcal{D}(m)\cap\mathcal{R}(\phi)} \mathcal{R}_j^\downarrow(m),$$

we can estimate

$$|\mathcal{R}(\phi \circ m)| - |\mathcal{R}(\phi)| \leq \sum_{j \in \mathcal{D}(m) \cap \mathcal{R}(\phi)} \big(|\mathcal{R}_j^\downarrow(m)| - 1\big).$$

This gives

$$HL(\phi) \leq \sum_{m \in \mathcal{G}} r_m \sum_{j \in \mathcal{D}(m) \cap \mathcal{R}(\phi)} \big(|\mathcal{R}_j^\downarrow(m)| - 1\big) = \sum_{j \in \mathcal{R}(\phi)} \sum_{\substack{m \in \mathcal{G} \\ \mathcal{D}(m) \ni j}} r_m\big(|\mathcal{R}_j^\downarrow(m)| - 1\big)$$

$$\leq |\mathcal{R}(\phi)| \sup_{j \in \Lambda} \sum_{\substack{m \in \mathcal{G} \\ \mathcal{D}(m) \ni j}} r_m\big(|\mathcal{R}_j^\downarrow(m)| - 1\big) = K|\mathcal{R}(\phi)|,$$

where we have used that the terms with $i \notin \mathcal{D}(m)$ in (4.19) are zero since in this case $\mathcal{R}(m[i]) = \{i\}$. It follows from (4.11) that

$$K = \sup_{i \in \Lambda} \sum_{\substack{m \in \mathcal{G} \\ \mathcal{D}(m) \ni i}} r_m\big(|\mathcal{R}_j^\downarrow(m)| - 1\big) \leq \sup_{i \in \Lambda} \sum_{m \in \mathcal{G}} r_m|\mathcal{R}_j^\downarrow(m)|,$$

so condition (ii) of Theorem 2.9 is satisfied with $\lambda = K$. It follows that H is nonexplosive and the continuous-time Markov chain $(\Phi_t)_{t \geq 0}$ with generator H satisfies

$$\mathbb{E}^\phi\big[|\mathcal{R}(\Phi_t)|\big] \leq |\mathcal{R}(\phi)|e^{-Kt} \qquad \big(t \geq 0, \ \phi \in \mathcal{C}(S^\Lambda, T)\big). \tag{4.20}$$

By Theorem 2.17, we can construct $(\Phi_t)_{t \geq 0}$ from a graphical representation, by applying the maps $\phi \mapsto \phi \circ m$ at the times of a Poisson point process. In fact, we can use the Poisson point process ω that we already have, provided we "read it off downwards in time" and instead of applying the map $x \mapsto m(x)$ at time t for each $(m, t) \in \omega$, we apply the map $\phi \mapsto \phi \circ m$. Thus, Theorem 2.17 implies that for each $u \in \mathbb{R}$ and $\phi \in \mathcal{C}(S^\Lambda, T)$, there exists a unique cadlag function $(\Phi_t)_{t \geq 0}$ such that $\Phi_0 = \phi$ and

$$\Phi_t = \begin{cases} \Phi_{t-} \circ m & \text{if } (m, u - t) \in \omega, \\ \Phi_{t-} & \text{otherwise} \end{cases} \qquad (t > 0).$$

Moreover, Theorem 2.17 tells us that $(\Phi_t)_{t \geq 0}$ is distributed as the continuous-time Markov chain with initial state ϕ and generator H. We would now like to say that the process $(\Phi_t^{u,\phi})_{t \in (-\infty, u]}$ from the proposition is given by

$$\Phi_t^{u,\phi} := \Phi_{u-t} \qquad \big(t \in (-\infty, u]\big) \quad (?), \tag{4.21}$$

but this is not quite right yet, since the process $(\Phi_t)_{t\geq 0}$ is right-continuous with left limits and as a result, because we are reversing time, the process defined in (4.21) is left-continuous with right limits. In view of this, we have to be a bit more careful. To get a right-continuous process, we have to define $(\Phi_t^{u,\phi})_{t\in(-\infty,u]}$ as the right-continuous modification of $(\Phi_{u-t})_{t\in(-\infty,u]}$. Moreover, to satisfy the evolution equation (4.17) also at $t = u$, in the special case that ω contains a point of the form (m, u), we have to start the process $(\Phi_t)_{t\geq 0}$ in $\Phi_0 = \phi \circ m$ instead of $\Phi_0 = \phi$. Apart from these little complications, the idea works, however, so Theorem 2.17 implies existence and uniqueness of the function $(\Phi_t^{u,\phi})_{t\in(-\infty,u]}$. Since the process a.s. does not jump at deterministic times, (4.20) remains true if we replace $(\Phi_t)_{t\geq 0}$ by its left-continuous modification. It also does not matter if we modify its initial state on an event of probability zero, so (4.20) implies (4.18).   ∎

Proposition 4.15 is the cornerstone of the proof of Theorem 4.13. We can now start to reap its benefits. For Markov processes that have a Poisson construction in terms of a set of maps $\mathcal{G}$ that is finite, we can simply construct the process by ordering the elements (m, t) of the Poisson set ω according to their times and then apply them in the right order. In our present setting, the set of times t for which ω contains an element of the form (m, t) is typically dense, so we can no longer simply concatenate the maps associated with elements of ω in the order described by their times. Since ω is countable, however, it is natural to approximate ω with finite sets $\omega' \subset \omega$ and hope that the construction converges to a limit as $\omega' \uparrow \omega$. We will prove that under the assumptions of Proposition 4.15, this idea indeed works.

If ω is a finite subset of $\mathcal{G} \times \mathbb{R}$ such that there exist no two elements $(m, t), (m', t') \in \omega$ with $m \neq m'$ and $t = t'$, then we define a maps $\mathbf{X}_{s,u}^{\omega} : \mathcal{S}^{\Lambda} \to S^{\Lambda}$ with $s \leq u$ by

$$\mathbf{X}_{s,u}^{\omega} := m_n \circ \cdots \circ m_1$$
$$\text{where} \quad \{(m, t) \in \omega : t \in (s, u]\} = \{(m_1, t_1), \ldots, (m_n, t_n)\} \qquad (4.22)$$
$$\text{with} \quad t_1 < \cdots < t_n.$$

In words, this says that $\mathbf{X}_{s,u}^{\omega}$ is the concatenation of the maps m for which $(m, t) \in \omega$ with $t \in (s, u]$, ordered by their times. We recall from Section 2.7 that for any $m : S^{\Lambda} \to S^{\Lambda}$ and $i \in \Lambda$, the map $m[i] : S^{\Lambda} \to S$ is defined as $m[i](x) := m(x)(i)$.

Proposition 4.16 (Finitely many relevant maps) *Assume (4.11). Then almost surely, for each $s \leq u$ and $i \in \Lambda$, there exists a finite set $\omega' \subset \omega$ such that $\mathbf{X}_{s',u'}^{\omega'}[i] = \mathbf{X}_{s',u'}^{\omega''}[i]$ for each finite $\omega'' \supset \omega'$ and for all $s \leq s' \leq u' \leq u$.*

Proof It suffices to prove the statement for $s, t \in \mathbb{Q}$ only. Since $\mathbb{Q}$ and Λ are countable, it therefore suffices to prove that the statement holds almost surely for deterministic $s \le u$ and $i \in \Lambda$. We fix $s \le u$ and $i \in \Lambda$. We define $\phi \in \mathcal{C}(S^\Lambda, S)$ by $\phi(x) := x(i)$ $(x \in S^\Lambda)$ and let $(\Phi^{u,\phi})_{t \in (-\infty, u]}$ be the backwards in time process of Proposition 4.15. Since this backwards in time process is a nonexplosive continuous-time Markov chain with state space $\mathcal{C}(S^\Lambda, S)$, the set $\{t \in (s, u] : \Phi^{u,\phi}_{t-} \ne \Phi^{u,\phi}_t\}$ is finite. We can order its elements as $t_1 < \cdots < t_n$ for some $n \ge 0$. Then there exist unique $m_1, \ldots, m_n \in \mathcal{G}$ such that

$$\omega' := \big\{(t_1, m_1), \ldots, (t_n, m_n)\big\} \subset \omega. \tag{4.23}$$

Clearly

$$\mathbf{X}^{\omega'}_{t,u}[i] = \Phi^{u,\phi}_t \qquad \big(t \in [s, u]\big). \tag{4.24}$$

We claim that the same is true with ω' replaced by any finite $\omega'' \subset \omega$ such that $\omega' \subset \omega''$. Indeed, it follows from (4.17) that

$$\Phi^{u,\phi}_{t-} = \Phi^{u,\phi}_t \circ m \quad \forall (m, t) \in \omega \backslash \omega' \text{ with } t \in (s, u],$$

so adding points in $\omega \backslash \omega'$ to ω' one at a time has no effect on $(\Phi^{u,\phi})_{t \in [s,u]}$ or on the maps $\mathbf{X}^{\omega'}_{t,u}[i]$ with $t \in [s, u]$.

Our arguments so far show that there exists a finite set $\omega' \subset \omega$ such that $\mathbf{X}^{\omega'}_{s',u'}[i] = \mathbf{X}^{\omega''}_{s',u'}[i]$ for each finite $\omega'' \supset \omega'$ and for all $s \le s' \le u' = u$. To generalise this to $u' \le u$, for each $u' \in (s, u]$, let $\omega'(u')$ be defined as in (4.23) but with $(\Phi^{u,\phi})_{t \in (-\infty, u]}$ replaced by $(\Phi^{u',\phi})_{t \in (-\infty, u']}$. To complete the proof, it suffices to show that

$$\bigcup_{u' \in (s,u]} \omega'(u') \tag{4.25}$$

is finite. It follows from (4.11) that the set

$$\xi := \big\{t \in (s, u] : \exists (m, t) \in \omega \text{ s.t. } m \in \mathcal{D}(m[i])\big\}$$

is a.s. finite. If $s < u' < u'' \le u$ are such that $\xi \cap (u', u''] =$, then $(\Phi^{u',\phi})_{t \in (-\infty, u']}$ is the restriction of $(\Phi^{u'',\phi})_{t \in (-\infty, u'']}$ to $(-\infty, u']$ and as a result $\omega'(u') = \omega'(u'')$. This means that the function $u' \mapsto \omega'(u')$ is piecewise constant. By what we have already proved, $\omega'(u')$ is a.s. finite for each fixed $u' \in (s, u]$, so we conclude that the union in (4.25) is a.s. finite. ∎

For each finite $\omega' \subset \omega$, the maps $(\mathbf{X}^{\omega'}_{s,u})_{s \le u}$ form a flow in the sense of (4.13). Proposition 4.16 says that these flows converge to a limit as $\omega' \uparrow \omega$. To formalize this, let Ω denote the set of all finite subsets of ω and for each $\omega' \in \Omega$, let $\mathbf{X}^{\omega'}$ denote the function from the set $\mathcal{X} := \{(s, u) : s, u \in \mathbb{R}, \ s \le u\} \times S^\Lambda$ into S^Λ given by $(s, u, x) \mapsto \mathbf{X}^{\omega'}_{s,u}(x)$. The collection $(\mathbf{X}^{\omega'})_{\omega' \in \Omega}$ is a

net in the topological sense of the word.[8] Proposition 4.16 says that this net almost surely has a limit $\mathbf{X}$ in a locally uniform sense. Indeed, if d is any metric generating the topology on S^Λ, then:

> Almost surely, for every compact $C \subset \mathcal{X}$ and $\varepsilon > 0$, there exists an $\omega' \in \Omega$ such that
> $$\sup_{(s,u,x) \in C} d\big(\mathbf{X}_{s,u}^{\omega''}(x), \mathbf{X}_{s,u}(x)\big) \leq \varepsilon \quad \forall \omega' \subset \omega'' \in \Omega.$$

We will denote this somewhat less formally as

$$\lim_{\omega' \uparrow \omega} (\mathbf{X}_{s,u}^{\omega'})_{s \leq u} = (\mathbf{X}_{s,u})_{s \leq u} \quad \text{locally uniformly a.s.} \tag{4.26}$$

The important thing to remember is that for any finite $\Delta \subset \Lambda$ and $S \leq U$, there exists a finite $\omega' \subset \omega$ such that

$$\mathbf{X}_{s,u}^{\omega''}(x)(i) = \mathbf{X}_{s,u}(x)(i) \quad \text{for all } \omega'' \supset \omega', \ S \leq s \leq u \leq U, \ i \in \Delta.$$

We now start to prove the various statements of Theorem 4.13.

Lemma 4.17 (The forward evolution) *Assume (4.11) and let* $(\mathbf{X}_{s,u})_{s \leq u}$ *be defined in (4.26). Then almost surely, for each* $s \in \mathbb{R}$ *and* $x \in S^\Lambda$, *the function* $(X_t^{s,x})_{t \geq s}$ *defined as*

$$X_t^{s,x} = \mathbf{X}_{s,t}(x) \qquad (t \geq s) \tag{4.27}$$

solves the evolution equation (4.10) with initial state $X_s^{s,x} = x$.

Proof Fix $s \in \mathbb{R}$ and $x \in S^\Lambda$ and choose finite $\omega_n \uparrow \omega$. For each n, there exists a unique piecewise constant, right-continuous function $X^n : [s, \infty) \to S^\Lambda$ such that

$$X_s^n = x \quad \text{and} \quad X_t^n = \mathfrak{m}_t^{\omega_n}(X_{t-}^n) \qquad (t > s),$$

and this function is given by

$$X_t^n = \mathbf{X}_{s,t}^{\omega_n}(x) \qquad (t \geq s).$$

---

[8]Readers unfamiliar with this term should look it up in any textbook on topology or on Wikipedia. Nets generalize sequences. Some statements for metrizable spaces, such as the statement that a space is compact if and only if each sequence has a convergent subsequence, generalize to arbitrary topological spaces only if sequences are replaced by the more general nets.

By (4.26) the function X^n converges locally uniformly to the function $X :$ $[s, \infty) \to S^\Lambda$ given by

$$X_t = \mathbf{X}_{s,t}(x) \qquad (t \geq s).$$

In particular, this implies that $X_t^n \to X_t$ and $X_{t-}^n \to X_{t-}$ for all $t > s$. If $(m, t) \in \omega$, then $\mathfrak{m}_t^{\omega_n} = m$ for all n large enough and hence $X_t^n = m(X_{t-}^n)$ for all n large enough. Taking the limit $n \to \infty$, using the continuity of m, it follows that $X_t = m(X_{t-})$. On the other hand, if there exists no $m \in \omega$ such that $(m, t) \in \omega$, then $\mathfrak{m}_t^{\omega_n} = 1$ for all n and hence $X_t^n = X_{t-}^n$ for all n, so letting $n \to \infty$ we see that $X_t = X_{t-}$. This shows that X solves (4.10) with the initial state x. ∎

Lemma 4.18 (Forward and backward equations) *Assume (4.11) and let T be a finite set. Almost surely, if a function $X : [s, \infty) \to S^\Lambda$ solves (4.10) for some $s \in \mathbb{R}$, then*

$$\phi(X_u) = \Phi_s^{u,\phi}(X_s) \qquad \big(u \geq s, \ \phi \in \mathcal{C}(S^\Lambda, T)\big), \tag{4.28}$$

where $(\Phi_t^{u,\phi})_{t \in (-\infty, u]}$ is the backward in time process defined in (4.17).

Proof Assume that $X : [s, \infty) \to S^\Lambda$ solves (4.10) and that $u \geq s$ and $f \in \mathcal{C}(S^\Lambda, T)$. By the definition of a solution to (4.10), the function $[s, u] \ni t \mapsto X_t$ is cadlag. The function $[s, u] \ni t \mapsto \Phi_s^{u,\phi}$ is cadlag with values in the countable set $\mathcal{C}(S^\Lambda, T)$. It follows[9] that the function $[s, u] \ni t \mapsto \Phi_t^{u,\phi}(X_t) \in T$ is cadlag, and that

$$\lim_{t' \uparrow t} \Phi_{t'}^{u,\phi}(X_{t'}) = \Phi_{t-}^{u,\phi}(X_{t-}) \quad \big(t \in (s, u]\big).$$

We claim that

$$\Phi_{t-}^{u,\phi}(X_{t-}) = \Phi_t^{u,\phi}(X_t) \quad \big(t \in (s, u]\big). \tag{4.29}$$

Indeed, if $(m, t) \in \omega$ for some $m \in \mathcal{G}$, (4.29) follows from the fact that

$$\Phi_{t-}^{u,\phi}(X_{t-}) = \Phi_t^{u,\phi} \circ m(X_{t-}) = \Phi_t^{u,\phi}(X_t),$$

while at all other times it follows fron the fact that $\Phi_{t-}^{u,\phi} = \Phi_t^{u,\phi}$ and $X_{t-} = X_t$. Since $[s, u] \ni t \mapsto \Phi_s^{u,\phi}$ takes values in the finite set T, we conclude from (4.29)

---

[9]A very formal way to see this is to observe that the function $\mathcal{C}(S^\Lambda, T) \times S^\Lambda \ni (\phi, x) \mapsto \phi(x) \in T$ is continuous. Since $\mathcal{C}(S^\Lambda, T)$ is countable one can also use that $[s, u] \ni t \mapsto \Phi_s^{u,\phi}$ is piecewise constant for a more elementary argument.

that this function is constant, so in particular its values in s and u are equal, which gives us (4.28). ∎

We note that Lemmas 4.17 and 4.18 combine to give us

$$\Phi_s^{t,\phi}(x) = \phi(X_t^{s,x}) = \phi\big(\mathbf{X}_{s,t}(x)\big),$$

proving that

$$\Phi_s^{t,\phi} = \phi \circ \mathbf{X}_{s,t} \qquad \big(s \leq t, \ \phi \in \mathcal{C}(S^\Lambda, T)\big). \tag{4.30}$$

Before we can prove Theorem 4.13 we need one more lemma.

Lemma 4.19 (Almost sure continuity) *Assume (4.11). Let $x_n, x \in S^\Lambda$ and $t_n, t \in [0, \infty)$ satisfy $x_n \to x$ (pointwise) and $t_n \to t$. Then*

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

Proof Since we equip S^Λ with the product topology, we need to show that

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

for each $i \in \Lambda$. Let $I := \{s \in \mathbb{R} : \exists\, (m, s) \in \omega \text{ s.t. } i \in \mathcal{D}(m)\}$ and let $t_- := \sup\{s : s \in I, \ s \leq t\}$ and $t_+ := \inf\{s : s \in I, \ s \geq t\}$. Since t is a deterministic time, $t_- < t < t_+$ a.s. Then $\mathbf{X}_{0,t_n}(x_n)(i) = \mathbf{X}_{0,t}(x)(i)$ for all n large enough such that $t_- < t_n < t_+$ and $x_n = x$ on the finite set $\mathcal{R}(\mathbf{X}_{0,t}[i])$, proving the desired a.s. convergence. ∎

Proof of Theorem 4.13 We claim that almost surely, for each $s \in \mathbb{R}$ and $x \in S^\Lambda$, there exists a unique solution $(X_t^{s,x})_{t \geq s}$ to the evolution equation (4.10) with initial state $X_s^{s,x} = x$, which is given by

$$X_t^{s,x} = \mathbf{X}_{s,t}(x) \qquad (t \geq s), \tag{4.31}$$

where $(\mathbf{X}_{s,u})_{s \leq u}$ is defined in (4.26). Indeed, Lemma 4.17 says that $(X_t^{s,x})_{t \geq s}$ defined in (4.31) solves (4.10), and Lemma 4.18 tells us that any solution $(X_t)_{t \geq s}$ to (4.10) with initial state $X_s = x$ satisfies

$$\phi(X_t) = \Phi_s^{t,\phi}(x) \qquad \big(t \geq s, \ \phi \in \mathcal{C}(S^\Lambda, T)\big),$$

which shows that solutions to (4.10) with a given initial state are unique.

By (4.31), the random maps $(\mathbf{X}_{s,t})_{s \leq t}$ defined in (4.12) coincide with those defined in (4.26). It follows from (4.30) that if T is any finite set, then

$$\phi \circ \mathbf{X}_{s,t} \in \mathcal{C}(S^\Lambda, T) \quad \forall \phi \in \mathcal{C}(S^\Lambda, T).$$

Using this for $\phi(x) := x(i)$ shows that $x \mapsto \mathbf{X}_{s,t}(x)(i)$ is continuous for each i and hence $x \mapsto \mathbf{X}_{s,t}(x)$ is continuous with respect to the product topology. The fact that $(\mathbf{X}_{s,t})_{s\leq t}$ form a stochastic flow in the sense of (4.13) follows from (4.26) and the analogue property of $(\mathbf{X}_{s,t}^{\omega'})_{s\leq t}$. This stochastic flow is clearly stationary. Using the fact that restrictions of a Poisson point set to disjoint parts of the space are independent, we also see that $(\mathbf{X}_{s,t})_{s\leq t}$ has independent increments.

We claim that the probability kernels $(P_t)_{t\geq 0}$ defined in (4.14) form a semigroup. Indeed, for each bounded measurable $f : S^\Lambda \to \mathbb{R}$,

$$P_0 f(x) = \mathbb{E}\big[f(\mathbf{X}_{0,0}(x))\big] = f(x)$$

and, for every measurable $A \subset S^\Lambda$,

$$
\begin{aligned}
P_{s+t}(x, A) &= \mathbb{P}\big[\mathbf{X}_{s,s+t} \circ \mathbf{X}_{0,s}(x) \in A\big] \\
&= \int \mathbb{P}\big[\mathbf{X}_{0,s}(x) \in \mathrm{d}y\big]\, \mathbb{P}\big[\mathbf{X}_{s,s+t}(y) \in A \,|\, \mathbf{X}_{0,s}(x) = y\big] \\
&= \int \mathbb{P}\big[\mathbf{X}_{0,s}(x) \in \mathrm{d}y\big]\, \mathbb{P}\big[\mathbf{X}_{s,s+t}(y) \in A\big] \\
&= \int P_s(x, \mathrm{d}y) P_t(y, A) = P_s P_t(x, A),
\end{aligned}
$$

where we have used the stationarity and independent increments of the stochastic flow. To prove that $(P_t)_{t\geq 0}$ is a Feller semigroup, we need to show that

$$P_{t_n}(x_n, \,\cdot\,) \underset{n\to\infty}{\Longrightarrow} P_t(x, \,\cdot\,) \quad \forall (x_n, t_n) \to (x, t),$$

where $\Rightarrow$ denotes weak convergence of probability measures. By the definition of weak convergence, this means that we need to show that

$$P_{t_n} f(x_n) \underset{n\to\infty}{\longrightarrow} P_t f(x) \quad \text{as } (x_n, t_n) \to (x, t)$$

for each continuous $f : S^\Lambda \to \mathbb{R}$. This follows from Lemma 4.19, which implies that

$$P_{t_n} f(x_n) = \mathbb{E}\big[f\big(\mathbf{X}_{0,t_n}(x_n)\big)\big] \underset{n\to\infty}{\longrightarrow} \mathbb{E}\big[f\big(\mathbf{X}_{0,t}(x)\big)\big] = P_t f(x).$$

Let $(X_t)_{t\geq 0}$ be defined as in (4.15), let $0 = t_0 \leq \cdots \leq t_n$, and let $f : (S^\Lambda)^n \to \mathbb{R}$ be bounded and measurable. Then using the stationarity and independent increments of the stochastic flow, we see that

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

which proves that $(X_t)_{t\geq 0}$ is the Feller process with semigroup $(P_t)_{t\geq 0}$ and inital law μ. ∎

Exercise 4.20 *The definitions of the generators G_{vot} and G_{cont} of the voter model and the contact process in (2.28) and (2.29) are of the general form (4.1), where in the first case $\mathcal{G} = \{\mathtt{vot}_{ij} : i, j \in \Lambda\}$ and $r_{\mathtt{vot}_{ij}} = \lambda(i, j)$, and in the second case $\mathcal{G} = \{\mathtt{cont}_{ij} : i, j \in \Lambda\}$ and $r_{\mathtt{cont}_{ij}} = \lambda(i, j)$. Give necessary and sufficient conditions on the constants $\lambda(i, j)$ for these rates $(r_m)_{m\in\mathcal{G}}$ to satisfy (4.11).*

4.4   Generator construction

Although Theorem 4.13 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.[10]

Lemma 4.21 (Variation of a function) *Let $f \in \mathcal{C}(S^\Lambda)$. Then*

$$\big|f(x) - f(y)\big| \leq \sum_{i:\,x(i)\neq y(i)} \delta f(i) \qquad \big(f \in \mathcal{C}(S^\Lambda), \ x, y \in S^\Lambda\big). \qquad (4.32)$$

Proof Let n be the number of sites i where x and y differ. Enumerate these sites as $\{i : x(i) \neq y(i)\} = \{i_1, \ldots, i_n\}$ or $= \{i_1, i_2, \ldots\}$ depending on whether n is finite or not. For $0 \leq k < n + 1$, set

$$z_k(i) := \begin{cases} y(i) & \text{if } i \in \{i_1, \ldots, i_k\}, \\ x(i) & \text{otherwise.} \end{cases}$$

If n is finite, then

$$\big|f(x) - f(y)\big| \leq \sum_{k=1}^{n} \big|f(z_k) - f(z_{k-1})\big| \leq \sum_{k=1}^{n} \delta f(i_k)$$

---

[10]This definition is similar to, but different from the more usual definition of the (total) variation of a function of one real variable. With functions of one real variable, the total variation is the maximal sum of all changes in the value of the function as one gradually increases the real variable. For functions on S^Λ, the idea is similar but instead of increasing a real variable we will gradually change a configuration x by modifying its coordinates one by one.

and we are done. If n is infinite, then the same argument gives

$$\big|f(x) - f(z_m)\big| \leq \sum_{k=1}^{m} \delta f(i_k) \qquad (m \geq 1).$$

Since $z_m \to y$ pointwise and f is continuous, (4.32) now follows by letting $m \to \infty$. ∎

We define spaces of functions by

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

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

Exercise 4.22 *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*

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

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

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

and that $\mathcal{C}_{\mathrm{fin}}(S^\Lambda)$ 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.

In what follows, we assume that $\mathcal{G}$ is a countable collection of local maps $m : S^\Lambda \to S^\Lambda$ and that $(r_m)_{m\in\mathcal{G}}$ are nonnegative rates. We define constants K_0 and K by

$$\begin{aligned}
K_0 &:= \sup_{i\in\Lambda} \sum_{m\in\mathcal{G}} r_m \mathbb{1}_{\mathcal{D}(m)}(i), \\
K &:= \sup_{i\in\Lambda} \sum_{m\in\mathcal{G}} r_m \big(|\mathcal{R}(m[i])| - 1\big).
\end{aligned} \tag{4.33}$$

Note that K is the same as the constant defined in formula (4.19) of Proposition 4.15, while K_0 is an upper bound on the maximal rate at which a site can change its type. Under the main assumption (4.11) of Theorem 4.13, both K_0 and K are finite.

Lemma 4.23 (Domain of pregenerator) *Assume that the constant K_0 defined in (4.33) is finite. Then, for each $f \in \mathcal{C}_{\mathrm{sum}}(S^\Lambda)$,*

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

In particular, for each $f \in \mathcal{C}_{\mathrm{sum}}(S^\Lambda)$ and $x \in S^\Lambda$, the right-hand side of (4.1) is absolutely summable and Gf is well-defined.

Proof This follows by writing

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

∎

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

Theorem 4.24 (Generator construction of particle systems) *Assume that the rates $(r_m)_{m \in \mathcal{G}}$ satisfy (4.11), let $(P_t)_{t \geq 0}$ be the Feller semigroup defined in (4.14) and let G be the linear operator with domain $\mathcal{D}(G) := \mathcal{C}_{\mathrm{sum}}$ defined by (4.1). Then G is closable and its closure $\overline{G}$ is the generator of $(P_t)_{t \geq 0}$. Moreover, $\mathcal{C}_{\mathrm{fin}}$ is a core for G.*

To prepare for the proof of Theorem 4.24 we need a few lemmas.

Lemma 4.25 (Generator on local functions) *Under the asumptions of Theorem 4.24, one has $\lim_{t \downarrow 0} t^{-1}(P_t f - f) = Gf$ for all $f \in \mathcal{C}_{\mathrm{fin}}$, where the limit exists in the topology on $\mathcal{C}(S^\Lambda)$.*

Proof Since $f \in \mathcal{C}_{\mathrm{fin}}$, by Exercise 4.22, there exists some finite $A \subset \Lambda$ such that f depends only on the coordinates in A. Let $\mathcal{G}^A := \{m \in \mathcal{G} : \mathcal{D}(m) \cap A \neq \emptyset\}$ denote the set of maps $m \in \mathcal{G}$ that can potentially change the state in A and for $B \supset A$, let $\mathcal{G}^{A,B} := \{m \in \mathcal{G}^A : \mathcal{R}(m[i]) \subset B \; \forall i \in A\}$ denote those maps who only need information from B to update the state in A. We introduce the notation

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

For $s \leq t$, we write $\omega_{s,t}^A$ and $\omega_{s,t}^{A,B}$ to denote the sets of Poisson points $(m,r) \in \omega_{s,t}$ with $m \in \mathcal{G}^A$ resp. $m \in \mathcal{G}^{A,B}$. If $\omega_{0,t}^A = \emptyset$, then $f(\mathbf{X}_{0,t}(x)) = f(x)$. Also, if

$\omega_{0,t}^{A,B}$ contains a single element (m, s) while $\omega_{0,t}^{B}$ contains no further elements, then $f(\mathbf{X}_{0,t}(x)) = f(m(x))$. Therefore

$$P_t f(x) = \mathbb{E}\big[f\big(\mathbf{X}_{0,t}(x)\big)\big] = f(x)\mathbb{P}[\omega_{0,t}^{A} = \emptyset]$$
$$+ \sum_{m \in \mathcal{G}^{A,B}} f\big(m(x)\big)\mathbb{P}\big[\omega_{0,t}^{A,B} = \omega_{0,t}^{B} = \{(m,s)\} \text{ for some } 0 < s \leq t\big]$$
$$+ \mathbb{E}\big[f\big(\mathbf{X}_{0,t}(x)\big)1_{\{\omega_{0,t}^{A}\backslash\omega_{0,t}^{A,B} \neq \emptyset \text{ or } |\omega_{0,t}^{B}| \geq 2\}}\big].$$

Here $\mathbb{P}[\omega_{0,t}^{A} = \emptyset] = e^{-Rt}$ with $R := \sum_{m \in \mathcal{G}^{A}} r_m$, so

$$f(x)\mathbb{P}[\omega_{0,t}^{A} = \emptyset] = f(x) - t\sum_{m \in \mathcal{G}^{A}} r_m f(x) + O(t^2)f(x),$$

where $O(t^2)$ is a function such that $\limsup_{t\to\infty} t^{-2}|O(t^2)| < \infty$. Similarly,

$$\sum_{m \in \mathcal{G}^{A,B}} f\big(m(x)\big)\mathbb{P}\big[\omega_{0,t}^{A,B} = \omega_{0,t}^{B} = \{(m,s)\} \text{ for some } 0 < s \leq t\big]$$
$$= t\sum_{m \in \mathcal{G}^{A,B}} r_m f\big(m(x)\big) + O_B(x, t^2) = t\sum_{m \in \mathcal{G}^{A}} r_m f\big(m(x)\big) + t\varepsilon_B(x) + O_B(x, t^2),$$

where the error terms satisfy $\limsup_{t\to\infty} t^{-2}\sup_{x \in S}|O_B(x, t^2)| < \infty$ for each fixed $B \supset A$, and $\lim_{B\uparrow\Lambda}\sup_{x \in S}|\varepsilon_B(x)| = 0$. Similarly

$$\mathbb{P}\big[\omega_{0,t}^{A}\backslash\omega_{0,t}^{A,B} \neq \emptyset \text{ or } |\omega_{0,t}^{B}| \geq 2\big] = t\varepsilon_B + O_B(t^2),$$

where the term of order t comes from the event that $\omega_{0,t}^{A}$ contains exactly one element, which is not in $\omega_{0,t}^{A,B}$. Combining our last three formulas, we obtain

$$P_t f(x) - f(x) = t\sum_{m \in \mathcal{G}^{A}} r_m\big\{f\big(m(x)\big) - f\big(x\big)\big\} + t\varepsilon_B(x) + O_B(x, t^2).$$

It follows that for each fixed $B \supset A$,

$$\limsup_{t\to\infty} \big\|t^{-1}\big(P_t f(x) - f(x)\big) - Gf(x)\big\| \leq \|\varepsilon_B\|.$$

Since $\|\varepsilon_B\| \to 0$ as $B \uparrow \Lambda$, the claim of the lemma follows. ∎

Lemma 4.26 (Approximation by local functions) *Assume that the constant K_0 defined in (4.33) is finite. Then for all $f \in \mathcal{C}_{\text{sum}}$ there exist $f_n \in \mathcal{C}_{\text{fin}}$ 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

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

Fix $f \in \mathcal{C}_{\text{sum}}$ and define $f_n(x) := f(x_n)$ $(x \in S^\Lambda)$. Then f_n depends only on the coordinates in Λ_n, hence $f_n \in \mathcal{C}_{\text{fin}}$. Formula (4.32) tells us that for any $x \in S^\Lambda$,

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

Since $f \in \mathcal{C}_{\text{sum}}$, it follows that

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

Moreover, we observe that

$$\begin{aligned}
&|Gf_n(x) - Gf(x)| \\
&= \Big| \sum_{m \in \mathcal{G}} r_m \big(f_n(m(x)) - f_n(x) \big) - \sum_{m \in \mathcal{G}} r_m \big(f(m(x)) - f(x) \big) \Big| \\
&\leq \sum_{m \in \mathcal{G}} r_m \big| f(m(x)_n) - f(x_n) - f(m(x)) + f(x) \big|.
\end{aligned} \qquad (4.34)$$

On the one hand, we have

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

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

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

Let $A \subset \Lambda$ be finite. Inserting either of our two estimates into (4.34), depending on whether $\mathcal{D}(m) \cap A \neq \emptyset$ or not, we find that

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

It follows that

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

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

Lemma 4.27 (Functions of summable variation) *Under the asumptions of Theorem 4.24, one has*

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

where K is the constant from (4.33). In particular, for each $t \geq 0$, P_t maps $\mathcal{C}_{\mathrm{sum}}(S^\Lambda)$ into itself.

Proof Fix $i \in \Lambda$. For each $j \in \Lambda$, define $\phi_j \in \mathcal{C}(S^\Lambda, S)$ by $\phi_j(x) := x(j)$ $(x \in S^\Lambda)$, and let $(\Phi_s^{t,\phi_j})_{s \in (-\infty,t]}$ denote the backward in time process started at time $t \geq 0$ in the function ϕ_j. Then for each $x, y \in S^\Lambda$ such that $x(k) = y(k)$ for all $k \neq i$, we can estimate using (4.32) and (4.30)

$$\begin{aligned}
|P_t f(x) - P_t f(y)| &= \left|\mathbb{E}[f(\mathbf{X}_{0,t}(x))] - \mathbb{E}[f(\mathbf{X}_{0,t}(y))]\right| \\
&\leq \mathbb{E}\left[|f(\mathbf{X}_{0,t}(x)) - f(\mathbf{X}_{0,t}(y))|\right] \\
&\leq \mathbb{E}\left[\sum_{j: \mathbf{X}_{0,t}(x)(j) \neq \mathbf{X}_{0,t}(y)(j)} \delta f(j)\right] \\
&= \sum_j \mathbb{P}\left[\mathbf{X}_{0,t}(x)(j) \neq \mathbf{X}_{0,t}(y)(j)\right] \delta f(j) \\
&= \sum_j \mathbb{P}\left[\Phi_0^{t,\phi_j}(x) \neq \Phi_0^{t,\phi_j}(y)\right] \delta f(j) \\
&\leq \sum_j \mathbb{P}\left[i \in \mathcal{R}(\Phi_0^{t,\phi_j})\right] \delta f(j).
\end{aligned}$$

By formula (4.18) of Proposition 4.15, it follows that

$$\begin{aligned}
\sum_i \delta P_t f(i) &\leq \sum_{ij} \mathbb{P}\left[i \in \mathcal{R}(\Phi_0^{t,\phi_j})\right] \delta f(j) \\
&= \sum_j \mathbb{E}\left[|\mathcal{R}(\Phi_0^{t,\phi_j})|\right] \delta f(j) \leq e^{Kt} \sum_j \delta f(j).
\end{aligned}$$

∎

Proof of Theorem 4.24 Let H be the full generator of $(P_t)_{t \geq 0}$ and let $\mathcal{D}(H)$ denote it domain. Then Lemma 4.25 shows that $\mathcal{C}_{\mathrm{fin}} \subset \mathcal{D}(H)$ and $Gf = Hf$

for all $f \in \mathcal{C}_{\text{fin}}$. By Lemma 4.26, it follows that $\mathcal{C}_{\text{sum}} \subset \mathcal{D}(H)$ and $Gf = Hf$ for all $f \in \mathcal{C}_{\text{sum}}$. To complete the proof, it suffices to show that $\mathcal{C}_{\text{fin}}$, and hence also the larger $\mathcal{C}_{\text{sum}}$, is a core for H.

We first prove that $\mathcal{C}_{\text{sum}}$ is a core for H. We will apply Lemma 4.7. We will show that for each $r > K$, where K is the constant from (4.33), and for each $f \in \mathcal{C}_{\text{sum}}(S^\Lambda)$, there exists a $p_r \in \mathcal{C}_{\text{sum}}(S^\Lambda)$ that solves the Laplace equation $(r - G)p_r = f$. Since $\mathcal{C}_{\text{sum}}(S^\Lambda)$ is dense in $\mathcal{C}(S^\Lambda)$ by Exercise 4.22, the claim then follows from Lemma 4.7 (ii).

Fix $r > K$ and $f \in \mathcal{C}_{\text{sum}}(S^\Lambda)$. We need to find a $p_r \in \mathcal{C}_{\text{sum}}(S^\Lambda)$ that solves the Laplace equation $(r - G)p_r = f$. In the light of Lemma 4.5 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 $p_r \in \mathcal{C}_{\text{sum}}(S^\Lambda)$ and $(r - G)p_r = f$. It follows from Theorem 4.8 that $p_r \in \mathcal{D}(H)$ and $(r - H)p_r = f$. Thus, it suffices to show that $p_r \in \mathcal{C}_{\text{sum}}$. To see this, note that if $x(j) = y(j)$ for all $j \neq i$, then

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

and therefore, by Lemma 4.27,

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

which proves that $p_r \in \mathcal{C}_{\text{sum}}$. This completes the proof that $\mathcal{C}_{\text{sum}}$ is a core for H, i.e., the closure of $G|_{\mathcal{C}_{\text{sum}}}$ is H. By Lemma 4.26, the closure of $G|_{\mathcal{C}_{\text{fin}}}$ contains $G|_{\mathcal{C}_{\text{sum}}}$, so we see that $\mathcal{C}_{\text{fin}}$ is also a core for H. ∎

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

Lemma 4.28 (Differentiation of semigroup)   *Assume that the rates $(r_m)_{m \in \mathcal{G}}$ satisfy (4.11), let $(P_t)_{t \geq 0}$ be the Feller semigroup defined in (4.14) and let G be the linear operator with domain $\mathcal{D}(G) := \mathcal{C}_{\text{sum}}(S^\Lambda)$ defined by (4.1). Then, for each $f \in \mathcal{C}_{\text{sum}}(S^\Lambda)$, $t \mapsto P_t f$ is a continuously differentiable function from $[0, \infty)$ to $\mathcal{C}(S^\Lambda)$ satisfying $P_0 f = f$, $P_t f \in \mathcal{C}_{\text{sum}}(S^\Lambda)$, and $\frac{\partial}{\partial t} P_t f = G P_t f$ for each $t \geq 0$.*

Proof This is a direct consequence of Proposition 4.4, Lemma 4.27, and Theorem 4.24. 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

In 1972, several authors published results of various degree of generality showing that interacting particle systems on infinite lattices are well-defined. Harris [Har72] used the Poisson approach. His result applies only to nearest-neighbor interactions on $\mathbb{Z}^d$. Instead of using the backward in time process he argued forwards in time, using percolation theory to show that if t is small enough, then the lattice can randomly be divided into finite pieces that mutually do not interact with each other during the time interval $(0, t]$.

Liggett [Lig72], on the other hand, gave a direct proof that the closure of G generates a Feller semigroup $(P_t)_{t \geq 0}$, and then invoked the abstract result Theorem 4.1 about Feller processes to prove the existence of a corresponding Markov process with cadlag sample paths. This result is more widely applicable than Harris' result and made it to Liggett's famous book [Lig85, Theorem I.3.9]. This is similar to our Theorem 4.24 but there are also some differences. Liggett does not write his generators in terms of local maps, but in terms of local probability kernels. 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.

Liggett's book [Lig85] does not treat graphical representations in the generality of our Theorem 4.13 but he does use explicit Poisson constructions for some specific 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 4.24 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.

The backward in time process of Proposition 4.15 will come back in Chapter 6 when we discuss duality of interacting particle systems. It is also interesting to look at the mean-field limit of this process. One can show that in the mean-field limit, the process

$$\left(\mathcal{R}(\Phi^{u,\phi}_{u-t}) \right)_{t \geq 0}$$

behaves as a branching process. As a result, solutions to the mean-field equation can be represented in terms of a stochastic process on the genealogical tree of a branching process. This is explained in [MSS20].

4.5   Ergodicity

The proof of Theorems 4.13 and 4.24 were quite long. Luckily, they yield more information than just the fact that the interacting particle systems we are interested in are well-defined. The basic phenomenon that motivates the study of interacting particle systems is collective behavior. The general picture is that for weak strengths of the interaction, different parts of space behave essentially independently, but for sufficiently strong interaction it may happen that all sites start to coordinate their behavior, giving rise to multiple invariant laws or even more exotic phenomena such as periodic behavior.

As a result of the methods of the previous sections, we will be able to prove results that confirm the "easy" part of this picture, namely the absence of collective behavior for weak strengths of the interaction.

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*[11] process $(X_t)_{t \in \mathbb{R}}$ such that (compare (4.2))

$$\mathbb{P}\big[X_u \in \,\cdot\, \big| (X_s)_{-\infty < s \leq t}\big] = P_{u-t}(X_t, \,\cdot\,) \quad \text{a.s.} \qquad (t \leq u), \tag{4.35}$$

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*.

Theorem 4.29 (Ergodicity) *Let X be an interacting particle system with state space of the form S^Λ and generator G of the form (4.1), and assume that the rates $(r_m)_{m \in \mathcal{G}}$ satisfy (4.11). Let T be a finite set with at least two elements.*

(a) *Assume that the constant K from (4.33) satisfies $K < 0$. Then the backward in time process satisfies*

$$\lim_{t \to -\infty} |\mathcal{R}(\Phi_t^{u,\phi})| = 0 \quad \text{a.s.} \quad \big(u \in \mathbb{R}, \ \phi \in \mathcal{C}(S^\Lambda, T)\big). \tag{4.36}$$

---

[11]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 $X_t' := X_{t-s}$ $(t \in \mathbb{R})$.

(b) *Assume that the backward in time process satisfies (4.36). Then the interacting particle system X has a unique invariant law ν, and*

$$\mathbb{P}^x\big[X_t \in \cdot\big] \underset{t\to\infty}{\Longrightarrow} \nu \qquad (x \in S^\Lambda). \tag{4.37}$$

Moreover, there exists a unique cadlag process $(X_t)_{t\in\mathbb{R}}$ such that

$$X_t = \mathfrak{m}_t^\omega(X_{t-}) \qquad (t \in \mathbb{R}), \tag{4.38}$$

and $(X_t)_{t\in\mathbb{R}}$ is distributed as the stationary Markov process corresponding to the invariant law ν.

Proof Part (a) is immediate from formula (4.18) of Proposition 4.15. Let Ψ_i denote the set of functions $\phi : S^\Lambda \to T$ that depend only on $x(i)$. Note that this set is finite. By (4.30) and the assumption that T has at least two elements

$$\mathcal{R}\big(\mathbf{X}_{s,t}[i]\big) \subset \bigcup_{\phi\in\Psi_i} \mathcal{R}\big(\Phi_s^{t,\phi}\big) \quad (s \le t, \; i \in \Lambda),$$

so (4.36) implies that

$$\lim_{s\to-\infty} \big|\mathcal{R}(\mathbf{X}_{s,t}[i])\big| = 0 \quad \text{a.s.} \quad \big(t \in \mathbb{R}, \; i \in \Lambda\big). \tag{4.39}$$

It follows from the definition of $(\mathbf{X}_{s,t})_{s\le t}$ in (4.12) that the function $t \mapsto \mathbf{X}_{s,t}[i]$ jumps only at times for which there exists a $(m,t) \in \omega$ such that $i \in \mathcal{D}(m)$. Sinc this set is locally finite by (4.11), we can replace the order of the "almost sure" and "for all t" statements, i.e., (4.39) holds almost surely for all t simultaneously. Formula (4.39) says that for small enough s, the function $\mathbf{X}_{s,t}[i]$ is constant, which implies that for each $z \in S^\Lambda$ the a.s. limit

$$X_t(i) := \lim_{s\to-\infty} \mathbf{X}_{s,t}(z)(i) \qquad (i \in \Lambda, \; t \in \mathbb{R}) \tag{4.40}$$

exists and does not depend on the choice of the configuration $z \in S^\Lambda$. Using the continuity of $\mathbf{X}_{s,u}$ (which is proved in Theorem 4.13) and the flow property, we see that a.s.

$$\mathbf{X}_{t,u}(X_t) = \lim_{s\to-\infty} \mathbf{X}_{t,u} \circ \mathbf{X}_{s,t}(z) = X_u \qquad (t \le u),$$

which implies that $(X_t)_{t\in\mathbb{R}}$ solves (4.38). If $(X'_t)_{t\in\mathbb{R}}$ is another solution, then for all s small enough so that $\mathcal{R}(\mathbf{X}_{s,t}[i]) = \emptyset$,

$$X'_t(i) = \mathbf{X}_{s,t}(X'_s)(i) = \mathbf{X}_{s,t}(z)(i) = X_t(i),$$

which shows that solutions to (4.38) are unique.

We claim that $X = (X_t)_{t \in \mathbb{R}}$ is Markov with respect to the transition probabilities $(P_t)_{t \geq 0}$ in the sense of (4.35). Indeed, for almost every trajectory $(x_s)_{-\infty < s \leq t}$ with respect to the law of $(X_s)_{-\infty < s \leq t}$, we have

$$
\begin{aligned}
\mathbb{P}\big[X_u &\in \cdot \,\big|\, (X_s)_{-\infty < s \leq t} = (x_s)_{-\infty < s \leq t}\big] \\
&= \mathbb{P}\big[\lim_{s \to -\infty} \mathbf{X}_{t,u} \circ \mathbf{X}_{s,t}(z) \in \cdot \,\big|\, (X_s)_{-\infty < s \leq t} = (x_s)_{-\infty < s \leq t}\big] \\
&\overset{1}{=} \mathbb{P}\big[\mathbf{X}_{t,u}(X_t) \in \cdot \,\big|\, (X_s)_{-\infty < s \leq t} = (x_s)_{-\infty < s \leq t}\big] \\
&= \mathbb{P}\big[\mathbf{X}_{t,u}(x_t) \in \cdot \,\big|\, (X_s)_{-\infty < s \leq t} = (x_s)_{-\infty < s \leq t}\big] \\
&\overset{2}{=} \mathbb{P}\big[\mathbf{X}_{t,u}(x_t) \in \cdot \big] = P_{u-t}(x_t, \cdot),
\end{aligned}
$$

where in step 1 we have used the continuity of the map $\mathbf{X}_{t,u}$ and in step 2 we have used that the random variables $\mathbf{X}_{t,u}$ and $(X_s)_{-\infty < s \leq t}$ are independent, since they are functions of the restriction of the Poisson set ω to the disjoint sets $\mathcal{G} \times (t, u]$ and $\mathcal{G} \times (-\infty, t]$, respectively. By stationarity,

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

does not depend on $t \in \mathbb{R}$, and since X is Markov this defines an invariant law ν. Since

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

and since by (4.40), we have

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

with respect to the topology of pointwise convergence, we conclude that (4.37) holds. ∎

We note that (4.37) says that if we start the process in an arbitrary initial state x, then the law at time t converges weakly[12] 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 process $(X_t)_{t \in \mathbb{R}}$ is ergodic in the usual sense of that word, i.e., the σ-field of events that are invariant under translations in time is trivial. The converse conclusion cannot be drawn, however, so the traditional way of describing (4.37) as "ergodicity" is a bit of a bad habit.

We have split Theorem 4.29 into a part (a) and (b) since the condition (4.36) is sometimes satisfied even when the constant K from (4.33) is positive. Indeed, we will later see that for the contact process, the condition (4.36)

---

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

is sharp but the condition $K < 0$ is not. In Exercise 5.12 below, we will calculate the constant K for the contact process and deduce that this process is ergodic for small values of the infection rate.

Theorem 4.29 is similar, but not identical to [Lig85, Thm I.4.1]. For Theorem 4.29 (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 (4.33) 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 shown 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

site i flips to the value σ with rate $\quad r_i^\sigma(x) := \dfrac{e^{\beta N_{x,i}(\sigma)}}{e^{\beta N_{x,i}(+1)} + e^{\beta N_{x,i}(-1)}},$

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. For each $i \in \Lambda$, let K_i^β denote the local probability kernel on $\{-1, +1\}^\Lambda$ defined as

$$K_i^\beta(x, y) := \begin{cases} r_i^\sigma(x) & \text{if } y = m_i^\sigma(x) \quad \big(\sigma \in \{-1, +1\}\big), \\ 0 & \text{otherwise,} \end{cases}$$

where $m_i^\sigma(x)$ is defined in (1.14). Then the generator (1.13) of the Ising model takes the form

$$G_{\text{Ising}}f = \sum_{i \in \Lambda} \big\{K_i^\beta f - f\big\}, \tag{4.41}$$

which is an expression of the form (2.6) but not a random mapping representation of the form (4.1). To find a random mapping representation for G_{Ising} in terms of local maps as in (4.1), it suffices to find a random mapping representation for the kernels K_i^β. This needs some preparations. 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

$$r_i^+(x) = \frac{e^{\beta N_{x,i}(+1)}}{e^{\beta N_{x,i}(+1)} + e^{\beta N_{x,i}(-1)}} = \frac{e^{\beta(|\mathcal{N}_i| + M_{x,i})/2}}{e^{\beta(|\mathcal{N}_i| + M_{x,i})/2} + e^{\beta(|\mathcal{N}_i| - M_{x,i})/2}}$$

$$= \frac{e^{\frac{1}{2}\beta M_{x,i}}}{e^{\frac{1}{2}\beta M_{x,i}} + e^{-\frac{1}{2}\beta M_{x,i}}} = \frac{1}{2}\left(1 + \frac{e^{\frac{1}{2}\beta M_{x,i}} - e^{-\frac{1}{2}\beta M_{x,i}}}{e^{\frac{1}{2}\beta M_{x,i}} + e^{-\frac{1}{2}\beta M_{x,i}}}\right)$$

$$= \frac{1}{2}\left(1 + \tanh(\tfrac{1}{2}\beta M_{x,i})\right).$$

Similarly, the rate of flipping to -1 is $r_i^-(x) = \frac{1}{2}(1 - \tanh(\frac{1}{2}\beta M_{x,i}))$.

For (mainly notational) simplicity, let us assume that each site i has the same number of neighbors in the graph (Λ, E), so that the size of the neighborhood

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

does not depend on $i \in \Lambda$. Then $M_{x,i}$ takes values in $\{-N, -N+2, \ldots, N\}$. We observe that the function $z \mapsto \frac{1}{2}(1 + \tanh(\frac{1}{2}\beta z))$ is increasing (see Figure 4.2). Inspired by this, for $L = -N, -N+2, \ldots, N$, we define local maps $m_{i,L}^\pm$ by

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

$$m_{i,L}^-(x)(j) := \begin{cases} -1 & \text{if } j = i \text{ and } M_{x,i} \leq L, \\ x(j) & \text{otherwise,} \end{cases} \qquad (4.42)$$

and we try a generator of the form

$$G_{\mathrm{Ising}}f(x) = \sum_{i \in \Lambda} \sum_{\sigma \in \{-,+\}} \sum_{L \in \{-N, -N+2, \ldots, N\}} r_{i,L}^\sigma \{f(m_{i,L}^\sigma(x)) - f(x)\}, \quad (4.43)$$

where $r_{i,L}^\sigma \geq 0$ are probabilities that need to be chosen in such a way that

$$K_i^\beta(x,y) = \sum_{\sigma \in \{-,+\}} \sum_{L \in \{-N, -N+2, \ldots, N\}} r_{i,L}^\sigma 1\{m_{i,L}^\sigma(x) = y\} \qquad (4.44)$$

is a random mapping representation of the kernel K_i^β. Let us fix x^-, x^+ such that $x^-(i) = -1$, $x^+(i) = +1$, and $x^-(j) = x^+(j)$ for all $j \neq i$. Setting $x = x^-$ and $y = x^+$ in (4.44) yields the equation

$$r_i^+(x) = \sum_{L=-N}^{M_{x,i}} r_{i,L}^+$$

Similarly, setting $x = x^+$ and $y = x^-$ in (4.44) yields the equation

$$r_i^-(x) = \sum_{L=M_{x,i}}^{N} r_{i,L}^-.$$

From this, we see that setting

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

(4.45)

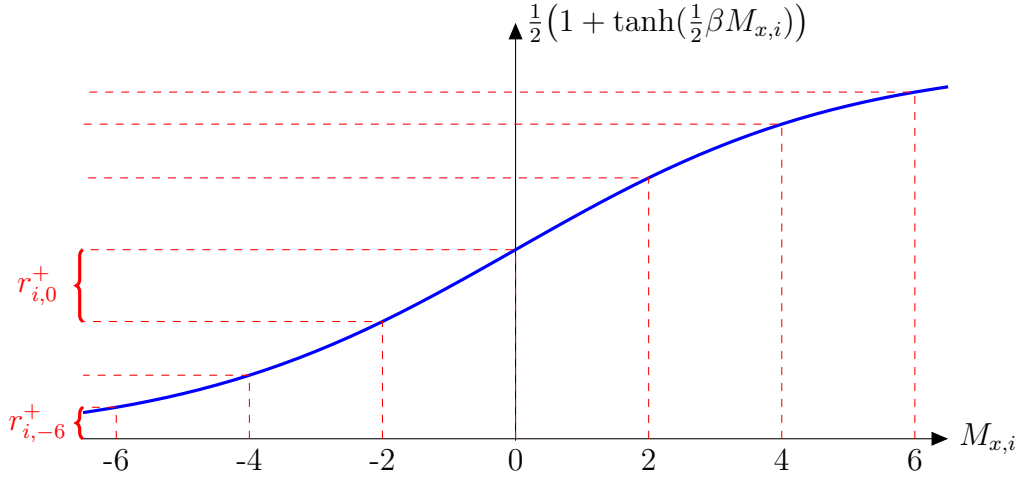has the effect that the generator in (4.43) equals the one in (4.41).

Figure 4.2: Definition of the rates $r_{i,L}^+$ from (4.45). In this example $N = 6$ and $\beta = 0.4$.

Theorem 4.30 (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_{Ising} given by (4.41). Then, for each $\beta \geq 0$, the closure of G_{Ising} generates a Feller semigroup. Moreover, for each β such that*

$$e^{\beta N} < \frac{N}{N-1}, \tag{4.46}$$

the Markov process with generator $\overline{G}_{\mathrm{Ising}}$ 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 (4.43). We observe that

$$\mathcal{D}(m_{i,L}^\pm) = \{i\}$$

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

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

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 $m_{i,L}^\pm(x)(i)$.

By Theorem 4.24, to conclude that the closure of G_{Ising} generates a Feller semigroup, it suffices to check (4.11), which in our case says that

$$\sup_{i\in\Lambda} \sum_{\sigma\in\{-,+\}} \sum_{L\in\{-N,-N+2,\ldots,N\}} r_{i,L}^\sigma\big(1 + |\mathcal{R}(m_{i,L}^\sigma[i])|\big)$$

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

To prove ergodicity for β small enough, we apply Theorem 4.29. We calculate the constant K from (4.33). By the symmetry between minus and plus spins,

$$\begin{aligned}
K &= 2 \sum_{L\in\{-N,-N+2,\ldots,N\}} r_{i,L}^+\big(|\mathcal{R}(m_{i,L}^+[i])| - 1\big) \\
&= -2r_{i,-N}^+ + 2 \sum_{L\in\{-N+2,\ldots,N\}} r_{i,L}^+\big(N-1\big) \\
&= -\big(1 + \tanh(-\tfrac{1}{2}\beta N)\big) + \big(\tanh(\tfrac{1}{2}\beta N) - \tanh(-\tfrac{1}{2}\beta N)\big)\big(N-1\big),
\end{aligned}$$

which is negative if and only if

$$1 + \tanh(-\tfrac{1}{2}\beta N) > \big(\tanh(\tfrac{1}{2}\beta N) - \tanh(-\tfrac{1}{2}\beta N)\big)\big(N-1\big)$$

$$\Leftrightarrow \quad 1 + \frac{e^{-\frac{1}{2}\beta N} - e^{\frac{1}{2}\beta N}}{e^{\frac{1}{2}\beta N} + e^{-\frac{1}{2}\beta N}} > \Big(\frac{e^{\frac{1}{2}\beta N} - e^{-\frac{1}{2}\beta N}}{e^{\frac{1}{2}\beta N} + e^{-\frac{1}{2}\beta N}} - \frac{e^{-\frac{1}{2}\beta N} - e^{\frac{1}{2}\beta N}}{e^{\frac{1}{2}\beta N} + e^{-\frac{1}{2}\beta N}}\Big)\big(N-1\big)$$

$$\Leftrightarrow \quad 2e^{-\frac{1}{2}\beta N} > 2\big(e^{\frac{1}{2}\beta N} - e^{-\frac{1}{2}\beta N}\big)\big(N-1\big)$$

$$\Leftrightarrow \quad \frac{e^{-\frac{1}{2}\beta N}}{e^{\frac{1}{2}\beta N} - e^{-\frac{1}{2}\beta N}} > N-1 \quad \Leftrightarrow \quad \frac{1}{e^{\beta N} - 1} > N-1$$

$$\Leftrightarrow \quad e^{\beta N} - 1 < \frac{1}{N-1} \quad \Leftrightarrow \quad e^{\beta N} < \frac{N}{N-1},$$

which is condition (4.46). ∎

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 4.35 which together with Corollary 4.36 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 4.31 (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 4.32 (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 4.33 (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 $f_n, f \in \mathcal{C}(E)$,*

$$\mu_n \underset{n \to \infty}{\Longrightarrow} \mu \quad \text{implies} \quad \mu_n K \underset{n \to \infty}{\Longrightarrow} \mu K$$
$$\text{and} \quad \|f_n - f\|_\infty \underset{n \to \infty}{\longrightarrow} 0 \quad \text{implies} \quad \|Kf_n - Kf\|_\infty \underset{n \to \infty}{\longrightarrow} 0.$$

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 μK.

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

Lemma 4.34 (Long-time limits) *Let E be a compact metrizable space and let $(P_t)_{t\geq0}$ 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 $(P_t)_{t\geq0}$.

Proof Using Lemma 4.33, 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 (4.47) 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 (4.47) 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 4.35 (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 A on $\mathcal{C}(E)$ such that $\overline{A} = G$ and*

$$\forall f \in \mathcal{D}(A) \ \exists f_n \in \mathcal{D}(G_n) \text{ such that } f_n \to f \quad and \quad G_n f_n \to Af. \quad (4.47)$$

Then the Feller semigroups $(P_t^n)_{t\geq0}$ and $(P_t)_{t\geq0}$ 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

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

We note that in the case of interacting particle systems, Theorem 4.24 implies the following.

Corollary 4.36 (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 (4.1) with rates satisfying (4.11). Assume moreover that*

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

Then the generators G_n, G satisfy (4.47).

Theorem 4.35 has the following useful consequence.

Proposition 4.37 (Limits of invariant laws) *Let E be a compact metrizable space and let G_n, G be generators of Feller processes in E satisfying (4.47). 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 4.35, 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]. The theorem holds for more general spaces than spaces of the form S^Λ. For example, it holds also for $\mathbb{R}^n$; see [KKO77].   ∎

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.

---

[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 $\mu = \nu$.

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, so it suffices to show that $\{1_A : A \text{ increasing}\}$ is distribution determining. By Lemma 4.31, it suffices to show that this class separates points and is closed under products.

If $x \neq y$, then either $x \notin \{z : z \geq y\}$ or $y \notin \{z : z \geq x\}$, 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 can 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) below, $\leq$ 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 $\mu Kf \leq \nu Kf$.

Recall from (2.2) that 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. In Chapter 4 we based our construction of an interacting particle system on a random mapping representation of its generator G in terms of local maps, of the form

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

where the rates satisfy (4.11). If there exists such a random mapping representation for which all local maps $m \in \mathcal{G}$ are monotone, then we say that G is *monotonically representable*.

Lemma 5.3 (Monotone representability) *Each monotonically represent-*
able 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 gener-
ator G of an interacting particle system, we can construct a stochastic flow
$(\mathbf{X}_{s,t})_{s \leq t}$ as in Theorem 4.13 based on a graphical representation ω. If all
maps $m \in \mathcal{G}$ are monotone, then for each finite $\omega' \subset \omega$, the maps $(\mathbf{X}_{s,t}^{\omega'})_{s \leq t}$ de-
fined in (4.22) are also monotone, since they are the concatenation of finitely
many maps from $\mathcal{G}$. By (4.26), this implies that the maps $(\mathbf{X}_{s,t})_{s \leq t}$ are also
monotone. 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 *monotoni-*
cally representable if its generator is monotonically representable. Somewhat
surprisingly, it turns out that for probability kernels, "monotonically repre-
sentable" 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 $\mathtt{vot}_{ij}$ 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 (1.20).

- The exclusion map $\mathtt{excl}_{ij}$ defined in (1.23).

- The cooperative branching map $\mathtt{coop}_{ij}$ defined in (1.25).

- The maps $m_{i,L}^{\pm}$ defined in (4.42).

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 (4.43).

- The biased voter model with generator as in (1.17).

- Systems of coalescing random walks with generator as in (1.21).

- The exclusion process with generator as in (1.24).

- 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 $\mathtt{ann}_{ij}$ defined in (1.22).

- The killing map $\mathtt{kill}_{ij}$ defined in (1.26).

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

- The antiferromagnetic Ising model with Glauber dynamics.

- "Rebellious" voter models as in (1.18).

- Systems of annihilating random walks.

- The biased annihilating branching process of [Sud97, Sud99].

---

[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 monotonically represented.

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 use the symbols $\underline{0}$ and $\underline{1}$ to denote the states in S^Λ that are identically 0 or 1, respectively. Below, $\delta_{\underline{0}}$ denotes the delta measure at the configuration that is identically 0, so $\delta_{\underline{0}} P_t$ denotes the law at time t of the process started in $X_0(i) = 0$ a.s. $(i \in \Lambda)$.

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*

$$\delta_{\underline{0}} P_t \underset{t \to \infty}{\Longrightarrow} \underline{\nu} \quad and \quad \delta_{\underline{1}} P_t \underset{t \to \infty}{\Longrightarrow} \overline{\nu}.$$

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 4.32. ∎

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

$$\delta_{\underline{0}} \leq \delta_{\underline{0}} P_t$$

By (5.1), this implies that

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

which shows that $t \mapsto \delta_{\underline{0}}P_t$ 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

$$\delta_{\underline{0}}P_t \underset{t\to\infty}{\Longrightarrow} \underline{\nu}.$$

It follows from Lemma 4.34 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 $f \in \mathcal{C}(\{0,1\}^\Lambda)$,

$$\delta_{\underline{0}} \leq \nu \quad \Rightarrow \quad \delta_{\underline{0}}P_t \leq \nu P_t = \nu \quad (t \geq 0).$$

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

$$\delta_{\underline{0}}P_t f \leq \delta_x P_t f \leq \delta_{\underline{1}}P_t f$$

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 4.32 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

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

proving that $p\mu + (1 - p)\nu$ is an invariant law. The fact that $\mathcal{I}$ is closed follows from Proposition 4.37. 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.*

---

[3]The the voter model in dimensions $d \geq 3$ is a counterexample. The Ising model in dimensions $d \geq 3$ is also a counterexample, although for the Ising model, it is still true that $\underline{\nu}$ and $\overline{\nu}$ are the only extremal invariant measures that are moreover translation invariant.

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

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

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 4.13. Show that the a.s. limits*

$$\left.\begin{aligned}
\underline{X}_t &:= \lim_{s\to-\infty} \mathbf{X}_{s,t}(\underline{0}), \\
\overline{X}_t &:= \lim_{s\to-\infty} \mathbf{X}_{s,t}(\underline{1})
\end{aligned}\right\} \quad (t \in \mathbb{R})$$

define stationary Markov processes $(\underline{X}_t)_{t\in\mathbb{R}}$ and $(\overline{X}_t)_{t\in\mathbb{R}}$ whose invariant laws

$$\underline{\nu} = \mathbb{P}[\underline{X}_t \in \cdot] \quad \text{and} \quad \overline{\nu} = \mathbb{P}[\overline{X}_t \in \cdot] \qquad (t \in \mathbb{R})$$

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 $\mathtt{bra}_{ij}(\underline{0}) = \underline{0}$ and $\mathtt{death}_i(\underline{0}) = \underline{0}$ for each $i,j \in \Lambda$, the all-zero configuration $\underline{0}$ is a trap for the contact process, so $\delta_{\underline{0}} P_t = \delta_{\underline{0}}$ for all $t \geq 0$ and hence

$$\underline{\nu} = \delta_{\underline{0}}.$$

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

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

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 $\overline{\nu} \neq \delta_{\underline{0}}$, then $\overline{\nu}$ is nontrivial.*

Proof We can always write $\overline{\nu} = (1-p)\delta_{\underline{0}} + p\mu$ where $p \in [0,1]$ and μ is a nontrivial law. By assumption, $\overline{\nu} \neq \delta_{\underline{0}}$, so $p > 0$. Since $\overline{\nu}$ and $\delta_{\underline{0}}$ 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 $X_0 = \underline{1} = X'_0$ 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 $\mathcal{G} = \mathcal{G}_{\mathrm{bra}} \cup \mathcal{G}_{\mathrm{death}}$ where

$$\mathcal{G}_{\mathrm{bra}} := \{\mathtt{bra}_{ij} : (i,j) \in \mathcal{E}^d\} \quad \text{and} \quad \mathcal{G}_{\mathrm{death}} := \{\mathtt{death}_i : i \in \mathbb{Z}^d\}.$$

Then X can be constructed as in Theorem 4.13 from a Poisson point set ω on

$$\mathcal{G} \times \mathbb{R} = (\mathcal{G}_{\mathrm{bra}} \cup \mathcal{G}_{\mathrm{death}}) \times \mathbb{R},$$

with intensity measure ρ_λ given by

$$\rho_\lambda(\{m\} \times A) := \begin{cases} \lambda \ell(A) & \text{if } m \in \mathcal{G}_{\mathrm{bra}}, \\ \ell(A) & \text{if } m \in \mathcal{G}_{\mathrm{death}}, \end{cases} \qquad \big(A \in \mathcal{B}(\mathbb{R})\big),$$

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 incidents, and the same death incidents 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.,

$$\rho''(\{m\} \times A) := \begin{cases} (\lambda' - \lambda)\ell(A) & \text{if } m \in \mathcal{G}_{\text{bra}}, \\ 0 & \text{if } m \in \mathcal{G}_{\text{death}}, \end{cases} \qquad (A \in \mathcal{B}(\mathbb{R})).$$

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

$$\begin{aligned} x \le x' \quad &\text{implies} \quad \mathtt{bra}_{ij}(x) \le \mathtt{bra}_{ij}(x'), \\ x \le x' \quad &\text{implies} \quad \mathtt{death}_i(x) \le \mathtt{death}_i(x'), \\ x \le x' \quad &\text{implies} \quad x \le \mathtt{bra}_{ij}(x'). \end{aligned}$$

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 incidents as ω, plus some extra branching incidents, using (4.26) we conclude that the stochastic flows $(\mathbf{X}_{s,t})_{s \le t}$ and $(\mathbf{X}'_{s,t})_{s \le t}$ constructed from ω and ω' satisfy

$$x \le x' \quad \text{implies} \quad \mathbf{X}_{s,t}(x) \le \mathbf{X}'_{s,t}(x') \qquad (s \le t).$$

In particular, setting $X_t := \mathbf{X}_{0,t}(\underline{1})$ and $X'_t := \mathbf{X}'_{0,t}(\underline{1})$ 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 (4.33) and apply Theorem 4.29 to conclude that*

$$\lambda N < 1 \quad \text{implies} \quad \overline{\nu} = \delta_{\underline{0}}.$$

In Chapter 7, we will prove that $\theta(\lambda) > 0$ for λ sufficiently large.

5.4   Other examples

The Ising model with Glauber dynamics

We have seen in (4.43) 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 one defined in (1.15), 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 4.30, 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 proved 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 rep-

---

[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.

resentable. Since both the constant configurations $\underline{0}$ and $\underline{1}$ are traps,

$$\underline{\nu} = \delta_{\underline{0}} \quad \text{and} \quad \overline{\nu} = \delta_{\underline{1}},$$

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 $\delta_{\underline{0}}$ and $\delta_{\underline{1}}$ are the only extremal invariant laws. On the other hand, in dimensions $d \geq 3$, the set $\mathcal{I}_e$ 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 $(X_t^\lambda)_{t \geq 0}$ denote the contact process with infection rate λ (and death rate one), started in $X_0^\lambda = 1$. Apply Corollary 4.36 to prove that for each fixed $t \geq 0$, the function*

$$\theta_t(\lambda) := \mathbb{P}[\mathbf{X}_{0,t}^\lambda(1)(i) = 1] \tag{5.6}$$

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

$$\mathtt{death}_{ij}x(k) := \begin{cases} 0 & \text{if } k \in \{i, j\}, \\ x(k) & \text{otherwise.} \end{cases} \tag{5.7}$$

Recall the branching map $\mathtt{bra}_{ij}$ defined in (1.6), the death map $\mathtt{death}_i$ defined in (1.7), and the cooperative branching map $\mathtt{coop}_{ij}$ defined in (1.25). Consider the cooperative branching process X with values in $\{0, 1\}^\mathbb{Z}$ with generator

$$G_X f(x) = \lambda \sum_{i \in \mathbb{Z}} \sum_{\sigma \in \{-1, +1\}} \left\{ f\left(\mathtt{coop}_{i, i+\sigma, i+2\sigma}x\right) - f(x) \right\}$$
$$+ \sum_{i \in \mathbb{Z}} \left\{ f\left(\mathtt{death}_i x\right) - f(x) \right\},$$

and the contact process with double deaths Y with generator

$$
\begin{aligned}
G_Y f(y) = \lambda \sum_{i \in \mathbb{Z}} \sum_{\sigma \in \{-1,+1\}} & \left\{ f\left(\mathtt{bra}_{i,i+\sigma} y\right) - f(y) \right\} \\
+ \sum_{i \in \mathbb{Z}} & \left\{ f\left(\mathtt{death}_{i,i+1} y\right) - f(y) \right\},
\end{aligned}
$$

Exercise 5.15   *Let X be the process with cooperative branching defined above and set*

$$
X_t^{(2)}(i) := 1_{\{X_t(i) = 1 = X_t(i+1)\}} \qquad (i \in \mathbb{Z}, \ t \geq 0).
$$

Show that X can be coupled to a contact process with double deaths Y (with the same parameter λ) in such a way that

$$
Y_0 \leq X_0^{(2)} \quad \text{implies} \quad Y_t \leq X_t^{(2)} \quad (t \geq 0).
$$

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*

$$
X_0 \leq Y_0 \quad \text{implies} \quad X_t \leq Y_t \quad (t \geq 0).
$$

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*

$$
\begin{aligned}
G_X f(x) = \tfrac{1}{2} b \sum_{i \in \mathbb{Z}} \sum_{\sigma \in \{-1,+1\}} & \left\{ f\left(\mathtt{bra}_{i,i+\sigma} x\right) - f(x) \right\} \\
+ \tfrac{1}{2} \sum_{i \in \mathbb{Z}} \sum_{\sigma \in \{-1,+1\}} & \left\{ f\left(\mathtt{rw}_{i,i+\sigma} x\right) - f(x) \right\},
\end{aligned}
$$

and let Y be a system of coalescing random walks with positive drift, with generator

$$
\begin{aligned}
G_Y f(y) = \tfrac{1}{2}(1+b) \sum_{i \in \mathbb{Z}} & \left\{ f\left(\mathtt{rw}_{i,i+1} y\right) - f(y) \right\} \\
+ \tfrac{1}{2} \sum_{i \in \mathbb{Z}} & \left\{ f\left(\mathtt{rw}_{i,i-1} y\right) - f(y) \right\}.
\end{aligned}
$$

Show that X and Y can be coupled such that

$$
Y_0 \leq X_0 \quad \text{implies} \quad Y_t \leq X_t \quad (t \geq 0).
$$

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 Basic definitions

Let S be a finite set, let Λ be countable, let $\mathcal{G}$ be a collection of local maps $m : S^\Lambda \to S^\Lambda$, and let $(r_m)_{m \in \mathcal{G}}$ be nonnegative rates satusfying (4.11). Then Theorem 4.13 tells us how the interacting particle system with generator G defined in (4.1) can be constructed from a graphical representation ω, which is a Poisson point set on $\mathcal{G} \times \mathbb{R}$ with intensity as in (4.8). More precisely, in (4.12) or alternatively (and equivalently) in (4.26), we have seen how in terms of ω it possible to define a stationary stochastic flow $(\mathbf{X}_{s,t})_{s \leq t}$ with independent increments, so that if $s \in \mathbb{R}$ and X_0 is an S^Λ-valued random variable with law μ independent of ω, then setting

$$X_t := \mathbf{X}_{s,s+t}(X_0) \qquad (t \geq 0) \tag{6.1}$$

defines a Feller process $(X_t)_{t \geq 0}$ whose generator is (the closure of) G defined in (4.1). We call this the interacting particle system with generator G.

Key to the proof of Theorem 4.13 was the backward in time process. Let T be a finite set and let $\mathcal{C}(S^\Lambda, T)$ be the space of continuous maps $\phi : S^\Lambda \to T$. We have seen in Lemma 2.18 that elements of $\mathcal{C}(S^\Lambda, T)$ are functions that depend on finitely many coordinates. As a result, $\mathcal{C}(S^\Lambda, T)$ is countable. By Proposition 4.15 and formula (4.30), for each $u \in \mathbb{R}$ and $\phi \in \mathcal{C}(S^\Lambda, T)$, setting

$$\Phi_t := \phi \circ \mathbf{X}_{u-t,u} \qquad (t \geq 0) \tag{6.2}$$

defines a nonexplosive continuous-time Markov chain $(\Phi_t)_{t \geq 0}$ with state space $\mathcal{C}(S^\Lambda, T)$ and generator as in (4.16). We call this the *backward in time process*. Somewhat unusually, the process $(\Phi_t)_{t \geq 0}$ has left-continuous sample paths. This is a result of the fact that in (4.22), we concatenate all maps

m for which $(m, t) \in \omega'$ with $t \in (s, u]$. We could alternatively have con-
catenated the maps with $t \in [s, u)$. This would make the process in (6.2)
right-continuous, but the forward process in (6.1) left-continuous. We will
stick to the definitions we have and just accept the fact that $(\Phi_t)_{t \geq 0}$ is left-
continuous.

Clearly, a lot can be learned about the interacting particle system $(X_t)_{t \geq 0}$
from a good understanding of the backward in time process $(\Phi_t)_{t \geq 0}$. It follows
from (6.1) and (6.2) that

$$\phi(X_{u-s}) = \phi \circ \mathbf{X}_{s,u}(X_0) = \Phi_{u-s}(X_0) \qquad (s \leq u),$$

so Φ_{u-s} tells us how a local function ϕ of the process at time $u - s$ depends
on the initial state X_0. A lot of the most studied interacting particle systems
have the property that the backward in time process is particularly simple,
which allows for a good understanding of these interacting particle systems.
This idea can be formalized via the concept of *duality*, which we now explain.

Recall the definition of a stochastic flow in (2.18). We can similarly define
a *backward stochastic flow* on a set R to be a collection $(\mathbf{Y}_{t,s})_{t \geq s}$ of random
maps[1] $\mathbf{Y}_{t,s} : R \to R$ such that

$$\mathbf{Y}_{s,s} = 1 \quad \text{and} \quad \mathbf{Y}_{t,s} \circ \mathbf{Y}_{u,t} = \mathbf{Y}_{s,u} \qquad (u \geq t \geq s).$$

Stationarity and independent increments are defined for backward stochas-
tic flows just as in the forward case. If $(\mathbf{Y}_{t,s})_{t \geq s}$ is a stationary backward
stochastic flow with independent increments, $u \in \mathbb{R}$, and Y_0 is an R-valued
random variable, independent of $(\mathbf{Y}_{t,s})_{t \geq s}$, then setting

$$Y_t := \mathbf{Y}_{u,u-t}(Y_0) \qquad (t \geq 0) \tag{6.3}$$

defines a Markov process $(Y_t)_{t \geq 0}$ in R with initial state Y_0.

Let S, R, and T be sets and let $\psi : S \times R \to T$ be a function. Then we
say that two maps $m : S \to S$ and $\hat{m} : R \to R$ are *dual* to each other with
respect to the *duality function* ψ if

$$\psi\big(m(x), y\big) = \psi\big(x, \hat{m}(y)\big) \qquad (x \in S, \ y \in R).$$

If $(\mathbf{X}_{s,t})_{s \leq t}$ is a stochastic flow on S and $(\mathbf{Y}_{t,s})_{t \geq s}$ is a backward stochastic
flow on R, then we say that $(\mathbf{X}_{s,t})_{s \leq t}$ and $(\mathbf{Y}_{t,s})_{t \geq s}$ are *dual* to each other
with respect to the duality function ψ if

$$\psi\big(\mathbf{X}_{s,t}(x), y\big) = \psi\big(x, \mathbf{Y}_{t,s}(y)\big) \qquad (s \leq t, \ x \in S, \ y \in R). \tag{6.4}$$

---

[1]We are deliberately a bit vague here. If R is a finite set, then our definitions are
complete. More generally, R could be any measurable space in which case one would want
$\mathbf{Y}_{s,t}$ to be jointly measurable both as a function of R and of the underlying probability
space. Usually, one also imposes some conditions on the way $\mathbf{Y}_{t,s}$ depends on s and t,
such as right-continuity.

Fix $s < u$ and let $(X_t)_{t \geq 0}$ be defined in terms of the stochastic flow $(\mathbf{X}_{s,t})_{s \leq t}$ as in (6.1) and let $(Y_t)_{t \geq 0}$ be defined in terms of the backward stochastic flow $(\mathbf{Y}_{t,s})_{t \geq s}$ as in (6.3). Then we claim that

$$\text{the function } [s, u] \ni t \mapsto \psi(X_{t-s}, Y_{u-t}) \text{ is constant.} \tag{6.5}$$

Indeed, the duality of $(\mathbf{X}_{s,t})_{s \leq t}$ and $(\mathbf{Y}_{t,s})_{t \geq s}$ implies

$$\psi(X_{t-s}, Y_{u-t}) = \psi\big(\mathbf{X}_{s,t}(X_0), \mathbf{Y}_{u,t}(Y_0)\big)$$
$$= \psi\big(\mathbf{X}_{t,u} \circ \mathbf{X}_{s,t}(X_0), Y_0\big) = \psi\big(\mathbf{X}_{s,u}(X_0), Y_0\big),$$

which clearly does not depend on t. A relation of the form (6.5) is called a *pathwise duality*.[2] In particular, setting $t = s, u$ in (6.5) we see that

$$\psi(X_u, Y_0) = \psi(X_0, Y_u).$$

In the special case that ψ takes values in a linear space such as $\mathbb{R}$ or $\mathbb{C}$, we can take expectations and conclude that

$$\mathbb{E}\big[\psi(X_u, Y_0)\big] = \mathbb{E}\big[\psi(X_0, Y_u)\big] \qquad (u \geq 0), \tag{6.6}$$

whenever X_u is independent of Y_0 and X_0 is independent of Y_u. A relation of the form (6.6) is called a *duality* between the Markov processes $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$.

We claim that any interacting particle system of the type described at the beginning of this section has a pathwise dual, which is the backward in time process. To see this, we define a backward stochastic flow $(\mathbb{F}_{t,s})_{t \geq s}$ on $\mathcal{C}(S^\Lambda, T)$ by

$$\mathbb{F}_{t,s}(\phi) := \phi \circ \mathbf{X}_{s,t} \qquad \big(s \leq t, \ \phi \in \mathcal{C}(S^\Lambda, T)\big), \tag{6.7}$$

and we define a duality function $\psi : S^\Lambda \times \mathcal{C}(S^\Lambda, T) \to T$ by

$$\psi(x, \phi) := \phi(x) \qquad \big(x \in S^\Lambda, \ \phi \in \mathcal{C}(S^\Lambda, T)\big). \tag{6.8}$$

Then it is straightforward to check that the stochastic flow $(\mathbf{X}_{s,t})_{s \leq t}$ and the backward stochastic flow $(\mathbb{F}_{t,s})_{t \geq s}$ are dual with respect to the duality function ψ, and hence the interacting particle system $(X_t)_{t \geq 0}$ and the backward in time process $(\Phi_t)_{t \geq 0}$ are pathwise dual.

Although formally, we have now found a pathwise dual for each interacting particle system, this dual is of little use in practice since the backward in time process is in general very complicated and the space $\mathcal{C}(S^\Lambda, T)$ is very large. As we will see in the coming sections, however, it sometimes happens that $\mathcal{C}(S^\Lambda, T)$ contains subspaces of "nice" functions that are mapped into themselves under the backward stochastic flow $(\mathbb{F}_{t,s})_{t \geq s}$, and this leads to more useful pathwise dualities.

---

[2]This terminology was first introduced in [JK14b].

6.2 Additive systems

Recall that if E is any topological space, then a function $f : E \to \mathbb{R}$ is called *lower semi-continuous* if one (and hence both) of the following equivalent conditions are satisfied:

  (i) $\liminf_{n\to\infty} f(x_n) \geq f(x)$ whenever $x_n \to x$,

  (ii) the level set $\{x \in E : f(x) \leq a\}$ is closed for each $a \in \mathbb{R}$.

Using (ii), it is easy to see that if a sequence f_n of lower semi-continuous functions has an increasing limit $f_n \uparrow f$, then f is lower semi-continuous. Let Λ and Δ be countable sets. We equip $\{0,1\}^\Lambda$ and $\{0,1\}^\Delta$ with the product topology and we let $\mathcal{C} = \mathcal{C}(\{0,1\}^\Lambda, \{0,1\}^\Delta)$ and $\mathcal{L} = \mathcal{L}(\{0,1\}^\Lambda, \{0,1\}^\Delta)$ denote the spaces of functions $m : \{0,1\}^\Lambda \to \{0,1\}^\Delta$ that are continuous and lower semi-continuous, respectively. We set

$$\mathcal{C}_{\mathrm{mon}} := \big\{m \in \mathcal{C} : m \text{ is monotone}\big\},$$
$$\mathcal{L}_{\mathrm{mon}} := \big\{m \in \mathcal{L} : m \text{ is monotone}\big\}.$$

We observe that

$$m \in \mathcal{L}_{\mathrm{mon}} \text{ and } x_n \uparrow x \quad \text{imply} \quad m(x_n) \xrightarrow[n\to\infty]{} m(x). \qquad (6.9)$$

Indeed, $\limsup_{n\to\infty} m(x_n) \leq m(x)$ by monotonicity while the reverse inequality follows from lower semi-continuity. By definition, a map $m : \{0,1\}^\Lambda \to \{0,1\}^\Delta$ is *additive* iff

  (i) $m(\underline{0}) = \underline{0}$,

  (ii) $m(x \vee y) = m(x) \vee m(y) \qquad \big(x, y \in \mathcal{S}(\Lambda)\big)$.

Each additive map is monotone, as follows by observing that $y \geq x$ implies that $m(y) = m(x \vee y) = m(x) \vee m(y) \geq m(x)$. We set

$$\mathcal{C}_{\mathrm{add}} := \big\{m \in \mathcal{C} : m \text{ is additive}\big\} \quad \text{and} \quad \mathcal{L}_{\mathrm{add}} := \big\{m \in \mathcal{L} : m \text{ is additive}\big\}.$$

We observe that

$$m \in \mathcal{L}_{\mathrm{add}} \quad \text{implies} \quad m\Big(\bigvee_{k=1}^{\infty} x_k\Big) = \bigvee_{k=1}^{\infty} m(x_k), \qquad (6.10)$$

which follows from (6.9) by writing

$$m\Big(\bigvee_{k=1}^{\infty} x_k\Big) = m\Big(\lim_{n\to\infty} \bigvee_{k=1}^{n} x_k\Big) = \lim_{n\to\infty} m\Big(\bigvee_{k=1}^{n} x_k\Big)$$
$$= \lim_{n\to\infty} \bigvee_{k=1}^{n} m(x_k) = \bigvee_{k=1}^{\infty} m(x_k).$$

A lot of local maps we have already seen are additive. Examples 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 (1.20).

- The exclusion map $\mathtt{excl}_{ij}$ defined in (1.23).

On the other hand, the following local maps are monotone, but not additive:

- The cooperative branching map $\mathtt{coop}_{ijk}$ defined in (1.25).

- The maps $m^{\pm}_{i,L}$ used to construct the Ising model with Glauber dynamics in (4.43).

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 (1.17).

- Systems of coalescing random walks with generator as in (1.21).

- The exclusion process with generator as in (1.24).

Define $e_i \in \mathcal{S}(\Lambda)$ by $e_i(j) := 1$ if $i = j$ and $:= 0$ otherwise. By (6.10), each $m \in \mathcal{L}_{\mathrm{add}}(\{0,1\}^\Lambda, \{0,1\}^\Delta)$ satisfies

$$m(x) = \bigvee_{i:x(i)=1} m(e_i) = \bigvee_{i\in\Lambda} x(i)m(e_i) \qquad \left(x \in \{0,1\}^\Lambda\right), \tag{6.11}$$

and hence is uniquely characterized by its action on the "basis vectors" e_i.

Lemma 6.1 (Matrix formulation) *Let Λ and Δ be countable sets and let $M : \Lambda \times \Delta \to \{0,1\}$ be a function. Then there exists a unique $m \in \mathcal{L}_{\mathrm{add}}(\{0,1\}^\Lambda, \{0,1\}^\Delta)$ such that $m(e_i)(j) = M(i,j)$ ($i \in \Lambda$, $j \in \Delta$), which is given by*

$$m(x)(j) := \bigvee_{i\in\Lambda} x(i)M(i,j) \qquad \left(j \in \Delta, \ x \in \{0,1\}^\Lambda\right). \tag{6.12}$$

This map m is continuous if and only if $\sum_{i\in\Lambda} M(i,j) < \infty$ for all $j \in \Delta$.

Proof Given M, we define m by (6.12). It is easy to see that m is additive and satisfies $m(e_i)(j) = M(i, j)$ ($i \in \Lambda$, $j \in \Delta$). We need to prove that m is lower semi-continuous, and continuous if and only if $\sum_{i \in \Lambda} M(i, j) < \infty$ for all $j \in \Delta$. The latter claim follows from Lemma 2.18 and the observation that $m(x)(j)$ depends on finitely many coordinates if and only if $\sum_{i \in \Lambda} M(i, j) < \infty$. To see that m defined in (6.12) is in general lower semi-continuous, let Λ_n be finite sets such that $\Lambda_n \uparrow \Lambda$, define $M_n(i, j) := M(i, j)$ if $i \in \Lambda_n$ and $:= 0$ otherwise, and define m_n as in (6.12) with M replaced by M_n. Then the functions m_n are continuous and increase to m, which implies that m is lower semi-continuous. By (6.11), the function m in (6.12) is the unique function in $\mathcal{L}_{\mathrm{add}}(\{0,1\}^\Lambda, \{0,1\}^\Delta)$ such that $m(e_i)(j) = M(i, j)$ ($i \in \Lambda$, $j \in \Delta$), completing the proof. ∎

We now specialize to the case $\Lambda = \Delta$. There is a useful graphical way to describe a lower semi-continuous additive map $m : \{0,1\}^\Lambda \to \{0,1\}^\Lambda$, that works as follows:

- For each $i, j \in \Lambda$ with $i \neq j$ such that $m(e_i)(j) = 1$, we draw an arrow from i to j.

- For each $i \in \Lambda$ such that $m(e_i)(i) = 0$, we draw a blocking symbol ▬ at i.

The following lemma says that additive local maps are fully described by their arrows and blocking symbols.

Lemma 6.2 (Graphical description) *If $m \in \mathcal{L}_{\mathrm{add}}(\{0,1\}^\Lambda, \{0,1\}^\Lambda)$ and $x \in \{0,1\}^\Lambda$, then $m(x)(j) = 1$ if and only if at least one of the following conditions is satisfied:*

(i) *there is an $i \in \Lambda$ such that $x(i) = 1$ and there is an arrow from i to j,*

(ii) *$x(j) = 1$ and there is no blocking symbol at j.*

A map $m \in \mathcal{L}_{\mathrm{add}}(\{0,1\}^\Lambda, \{0,1\}^\Lambda)$ is local if and only if its description involves only finitely many arrows and blocking symbols.

Proof The desciption of m in points (i) and (ii) is just a reformulation of (6.11). A map m is local if and only it is continous and $\mathcal{D}(m)$ is finite. By Lemma 6.1 m is continuous if and only if the number of incoming arrows at each $i \in \Lambda$ is finite, while $\mathcal{D}(m)$ is the set of all $i \in \Lambda$ that are the endpoint of an arrow or where there is a blocking symbol. ∎

In terms of arrows and blocking symbols, the branching map $\mathtt{bra}_{ij}$, the death map $\mathtt{death}_i$, the voter map $\mathtt{vot}_{ij}$, the coalescing random walk map $\mathtt{rw}_{ij}$, and the exclusion map $\mathtt{excl}_{ij}$ look as follows:

$$\texttt{bra}_{1,2} \qquad \texttt{death}_1$$

$$\texttt{vot}_{1,2} \qquad \texttt{rw}_{1,2} \qquad \texttt{excl}_{1,2}$$
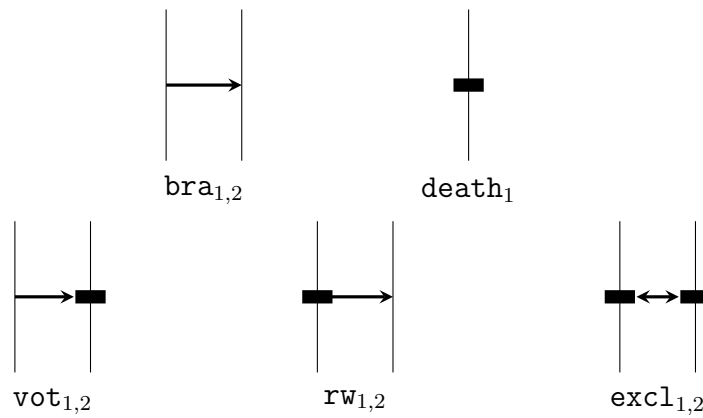
We use our conventions of representing additive maps in terms of arrows and blocking symbols to depict the graphical representation of an additive interacting particle system in a more suggestive way. In Figure 4.1, we drew the graphical representation of a contact process in the following fashion:
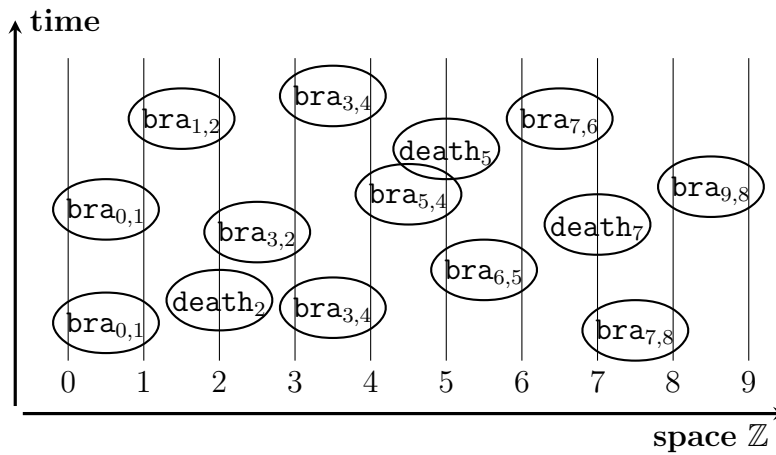
With our new conventions, the same graphical representation looks as follows:
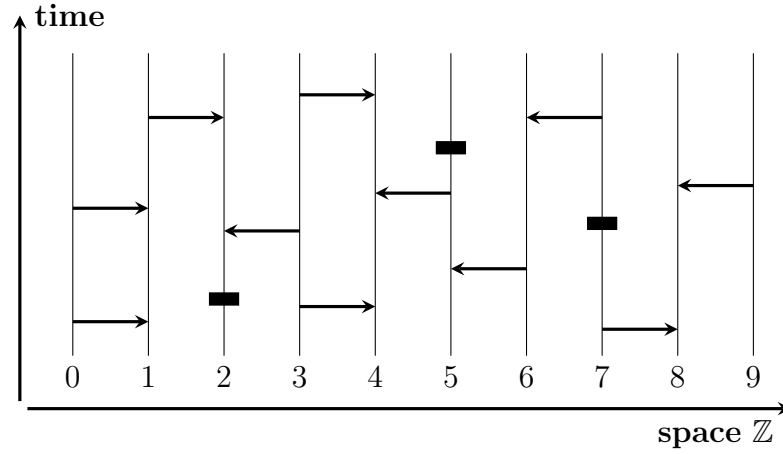
It is easy to see that the concatenation of two additive maps is again additive. As a result, using (4.26), we see that if $(\mathbf{X}_{s,t})_{s \leq t}$ is the stochastic flow associated with the graphical representation of an additive particle system, then the functions $\mathbf{X}_{s,t} : \{0,1\}^{\Lambda} \to \{0,1\}^{\Lambda}$ are additive maps. By Theorem 4.13 they are also continuous. We claim that $\mathbf{X}_{s,t}(x)(j) = 1$ if and only if there is an $i \in \Lambda$ with $x(i) = 1$ and it is possible to walk through the graphical representation from the space-time point (i, s) to the space time point (j, t) along a path that may use arrows, but must avoid the blocking symbols. We now make this claim more precise.

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

(i) if $\gamma_{t-} \neq \gamma_t$ for some $t \in (s, u]$, then there is an arrow from (γ_{t-}, t) to (γ_t, t),

(ii) there exist no $t \in (s, u]$ such that $\gamma_{t-} = \gamma_t$ while there is a blocking symbol at (γ_t, t).

$$(6.13)$$

We write $(i, s) \rightsquigarrow (j, u)$ if there exists an open path from (i, s) to (j, u). With these definitions, we can make our earlier claim precise. We claim that:

$$\mathbf{X}_{s,t}(x)(j) = 1 \quad \text{iff} \quad \exists i \in \Lambda \text{ s.t. } x(i) = 1 \text{ and } (i, s) \rightsquigarrow (j, t). \qquad (6.14)$$

To prove (6.14), it suffices to observe that if we define $X_t(j) := 1$ iff the condition on the right-hand side of (6.14) is satisfied, then the function $(X_t)_{t \geq s}$ solves the evolution equation (4.10). 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:
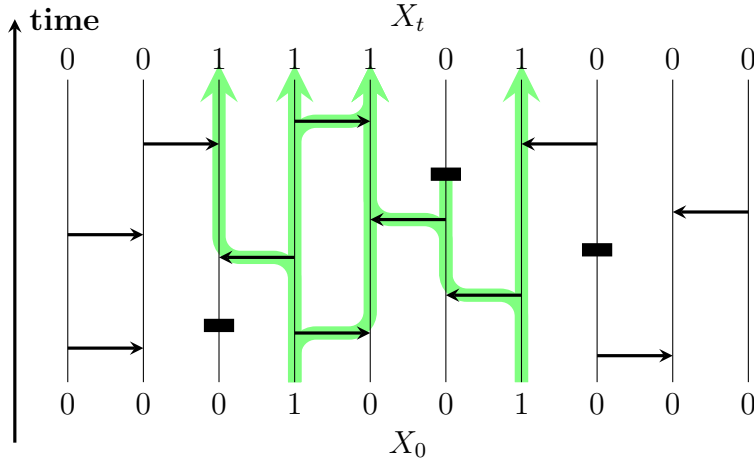
6.3 Additive duality

There exists a useful duality theory for additive systems. We first formulate the main results and then set out to prove them. Throughout this section, we set

$$\mathcal{S}(\Lambda) := \{0,1\}^{\Lambda} \quad \text{and} \quad \mathcal{S}_{\text{fin}}(\Lambda) := \big\{ x \in \mathcal{S}(\Lambda) : |x| < \infty \big\},$$

where in line with notation introduced in Section 2.8, we set

$$|x| := \sum_{i \in \Lambda} x(i) \qquad \big(x \in \mathcal{S}(\Lambda) \big).$$

We define a function $\psi_{\text{add}} : \mathcal{S}(\Lambda) \times \mathcal{S}(\Lambda) \to \{0,1\}$ by

$$\psi_{\text{add}}(x,y) := 1_{\{x \wedge y \neq \underline{0}\}} \qquad \big(x, y \in \mathcal{S}(\Lambda) \big), \tag{6.15}$$

where $x \wedge y$ denotes the pointwise minimum of x and y and $\underline{0}$ denotes the all zero configuration. With this notation, the main results of this section are as follows.

Lemma 6.3 (Dual maps) *Each lower semi-continuous additive map* $m : \mathcal{S}(\Lambda) \to \mathcal{S}(\Lambda)$ *has a unique dual map* $\hat{m}$ *with respect to the duality function* ψ_{add}. *The dual map* $\hat{m} : \mathcal{S}(\Lambda) \to \mathcal{S}(\Lambda)$ *is also a lower semi-continuous additive map and given by*

$$\hat{m}(e_j)(i) = m(e_i)(j) \qquad (i, j \in \Lambda). \tag{6.16}$$

If m *is local, then so is* $\hat{m}$.

Theorem 6.4 (Additive duality) *Let G be the generator of an interacting particle system $(X_t)_{t\geq 0}$ with state space $\mathcal{S}(\Lambda)$. Assume that G has a random mapping representation of the form (4.1) such that all local maps $m \in \mathcal{G}$ are additive and the rates $(r_m)_{m\in\mathcal{G}}$ satisfy (4.11). Set $\hat{\mathcal{G}} := \{\hat{m} : m \in \mathcal{G}\}$ and $r_{\hat{m}} := r_m$ $(m \in \mathcal{G})$. Then*

$$\hat{G} := \sum_{\hat{m}\in\hat{\mathcal{G}}} r_{\hat{m}}\{f(\hat{m}(y)) - f(y)\} \qquad (y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)) \qquad (6.17)$$

is the generator of a nonexplosive continuous-time Markov chain $(Y_t)_{t\geq 0}$ with state space $\mathcal{S}_{\mathrm{fin}}(\Lambda)$. The Markov processes $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ are dual with respect to the duality function ψ_{add}, i.e.,

$$\mathbb{E}[\psi_{\mathrm{add}}(X_t, Y_0)] = \mathbb{E}[\psi_{\mathrm{add}}(X_0, Y_t)] \qquad (t \geq 0) \qquad (6.18)$$

whenever X_t is independent of Y_0 and X_0 is independent of Y_t. If the dual rates $(r_{\hat{m}})_{\hat{m}\in\hat{\mathcal{G}}}$ also satisfy (4.11) so that $\hat{G}$ defined as in (6.17) for all $y \in \mathcal{S}(\Lambda)$ generates an interacting particle system, then (6.18) remains true if $(Y_t)_{t\geq 0}$ is this interacting particle system started in an arbitrary initial state in $\mathcal{S}(\Lambda)$.

We observe that in terms of our graphical way of depicting additive maps, formula (6.16) has the following interpretation:

$$\begin{array}{l} \hat{m} \text{ is obtained from } m \text{ by keeping the blocking symbols} \\ \text{and reversing the direction of all arrows.} \end{array} \qquad (6.19)$$

In applications of Theorem 6.4 and in particular of the duality relation (6.18), if is important to know that knowing $\mathbb{E}[\psi_{\mathrm{add}}(X_t, y)]$ for all $y \in \mathcal{S}(\Lambda)$ determines the law of X_t uniquely. This is guaranteed by the following lemma, which shows that in fact it suffices to know $\mathbb{E}[\psi_{\mathrm{add}}(X_t, y)]$ for all $y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)$.

Lemma 6.5 (Distribution determining functions) *The functions $\{f_y : y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)\}$ with $f_y(x) := 1_{\{x\wedge y\neq \underline{0}\}}$ are distribution determining on $\mathcal{S}(\Lambda)$.*

Proof We may equivalently prove that the functions $g_y(x) := 1 - f_y(x) = 1_{\{x\wedge y=\underline{0}\}}$ are distribution determining. Since $x \wedge e_i(i) = x(i)$, the class $\{g_y : y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)\}$ separates points, and since $g_y g_{y'} = g_{y\vee y'}$, this class is closed under products. The claim now follows from Lemma 4.31. ∎

We now set out to prove Lemma 6.3 and Theorem 6.4. We start with the lemma.

Proof of Lemma 6.3 It follows from Lemma 6.1 that there exists a unique $\hat{m} \in \mathcal{L}_{\mathrm{add}}(\{0,1\}^\Lambda, \{0,1\}^\Lambda)$ such that (6.16) holds. Now, using (6.11) we see that for each $x, y \in \mathcal{S}(\Lambda)$

$$\psi_{\mathrm{add}}\big(m(x), y\big) = 1_{\{m(x) \wedge y \neq \underline{0}\}} = \bigvee_{j \in \Lambda} m(x)(j)y(j)$$
$$= \bigvee_{i,j \in \Lambda} x(i)\, m(e_i)(j)\, y(j) = \bigvee_{i,j \in \Lambda} x(i)\, \hat{m}(e_j)(i)\, y(j)$$
$$= \bigvee_{i \in \Lambda} x(i)\hat{m}(y)(i) = 1_{\{x \wedge \hat{m}(y) \neq \underline{0}\}} = \psi_{\mathrm{add}}\big(x, \hat{m}(y)\big),$$

which proves that m and $\hat{m}$ are dual with respect to the duality function ψ_{add}. The latter implies in particular that

$$\hat{m}(y)(i) = \psi_{\mathrm{add}}\big(e_i, \hat{m}(y)\big) = \psi_{\mathrm{add}}\big(m(e_i), y\big) \quad (i \in \Lambda),$$

which shows that $\hat{m}$ is the unique map from $\mathcal{S}(\Lambda)$ into itself with this property. (In particular, each such map must be lower semi-continuous and additive.) It is clear from (6.19) that if m is local, then so is $\hat{m}$. ∎

Proof of Theorem 6.4 Let $(\mathbb{F}_{t,s})_{t \geq s}$ be defined as

$$\mathbb{F}_{t,s}(\phi) := \phi \circ \mathbf{X}_{s,t} \qquad \big(s \leq t, \ \phi \in \mathcal{C}(\{0,1\}^\Lambda, \{0,1\})\big),$$

i.e., this is the backward stochastic flow defined in (6.7) (with $S = T = \{0,1\}$). We have already seen that $\mathbf{X}_{s,t}$ is a continuous additive map for each $s \leq t$. Therefore, since the concatenation of two continuous additive maps is again continuous and additive, we see that

$$\mathbb{F}_{t,s}(\phi) \in \mathcal{C}_{\mathrm{add}}\big(\{0,1\}^\Lambda, \{0,1\}\big) \qquad \big(s \leq t, \ \phi \in \mathcal{C}_{\mathrm{add}}(\{0,1\}^\Lambda, \{0,1\})\big). \quad (6.20)$$

By Proposition 4.15 and formula (4.30), it follows that for each $u \in \mathbb{R}$ and $\phi \in \mathcal{C}_{\mathrm{add}}(\{0,1\}^\Lambda, \{0,1\})$, setting

$$\Phi_t := \mathbb{F}_{u,u-t}(\phi) \qquad (t \geq 0) \tag{6.21}$$

defines a nonexplosive continuous-time Markov chain $(\Phi_t)_{t \geq 0}$ with left-continuous sample paths and state space $\mathcal{C}_{\mathrm{add}}(\{0,1\}^\Lambda, \{0,1\})$.

It follows from (6.11) that each $\phi \in \mathcal{L}_{\mathrm{add}}(\{0,1\}^\Lambda, \{0,1\})$ can be written in the form

$$\phi(x) = \psi_{\mathrm{add}}(x, y) = \bigvee_{i \in \Lambda} x(i)y(i) \qquad \big(x \in \mathcal{S}(\Lambda)\big), \tag{6.22}$$

where $y \in \mathcal{S}(\Lambda)$ is defined as $y(i) := \phi(e_i)$ $(i \in \Lambda)$. Conversely, by Lemma 6.1, each $y \in \mathcal{S}(\Lambda)$ defines via (6.22) a function $\phi \in \mathcal{L}_{\mathrm{add}}(\{0,1\}^\Lambda, \{0,1\})$, and this function is continuous if and only if $y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)$. This means that

$$\mathcal{S}_{\mathrm{fin}}(\Lambda) \ni y \mapsto \psi_{\mathrm{add}}(\,\cdot\,, y) \in \mathcal{C}_{\mathrm{add}}(\{0,1\}^\Lambda, \{0,1\}) \qquad (6.23)$$

is a bijection. We can use this bijection to identify the continuous-time Markov chain $(\Phi_t)_{t\geq 0}$ with state space $\mathcal{C}_{\mathrm{add}}(\{0,1\}^\Lambda, \{0,1\})$ from (6.21) with a continuous-time Markov chain $(Y_t)_{t\geq 0}$ with state space $\mathcal{S}_{\mathrm{fin}}(\Lambda)$. We will show that this is precisely the process with generator $\hat{G}$ as in (6.17).

Formally, we proceed as follows. Combining (6.20) with the fact that the map in (6.23) is a bijection, we see that we can define a backward stochastic flow $(\mathbf{Y}_{t,s})_{t\geq s}$ on $\mathcal{S}_{\mathrm{fin}}(\Lambda)$ by

$$\mathbb{F}_{t,s}(\psi_{\mathrm{add}}(\,\cdot\,, y)) =: \psi_{\mathrm{add}}(\,\cdot\,, \mathbf{Y}_{t,s}(y)) \qquad (t \geq s, \ y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)). \qquad (6.24)$$

We claim that for each $u \in \mathbb{R}$ and $y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)$, setting

$$Y_t := \mathbf{Y}_{u,u-t}(y) \qquad (t \geq 0)$$

defines a nonexplosive continuous-time Markov chain $(Y_t)_{t\geq 0}$ with left-continuous sample paths and state space $\mathcal{S}_{\mathrm{fin}}(\Lambda)$. Indeed, if we identify $\mathcal{S}_{\mathrm{fin}}(\Lambda)$ with $\mathcal{C}_{\mathrm{add}}(\{0,1\}^\Lambda, \{0,1\})$ via the bijection in (6.23), then $(\mathbf{Y}_{t,s})_{t\geq s}$ corresponds to $(\mathbb{F}_{t,s})_{t\geq s}$ and the process $(Y_t)_{t\geq 0}$ corresponds to $(\Phi_t)_{t\geq 0}$. Recall from Proposition 4.15 that the process $(\Phi_t)_{t\geq 0}$ jumps from ϕ to $\phi \circ m$ with rate r_m, for each $m \in \mathcal{G}$. In view of the bijection in (6.23), this means that the process $(Y_t)_{t\geq 0}$ jumps from y with rate r_m to $m'(y)$ defined as

$$\psi_{\mathrm{add}}(\,\cdot\,, y) \circ m =: \psi_{\mathrm{add}}(\,\cdot\,, m'(y)) \qquad (y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)).$$

This says that

$$\psi_{\mathrm{add}}(x, m'(y)) = \psi_{\mathrm{add}}(\,\cdot\,, y) \circ m(x) = \psi_{\mathrm{add}}(m(x), y)$$

$(x \in \mathcal{S}(\Lambda), \ y \in \mathcal{S}_{\mathrm{fin}}(\Lambda))$. Knowing $\psi_{\mathrm{add}}(x, m'(y))$ for all x clearly uniquely determines $m'(y)$, so we see that $m' = \hat{m}$, the dual map from Lemma 6.3. This proves that $(Y_t)_{t\geq 0}$ is the continuous-time Markov chain with state space $\mathcal{S}_{\mathrm{fin}}(\Lambda)$ and generator $\hat{G}$ from (6.17). In particular, the latter is nonexplosive since $(\Phi_t)_{t\geq 0}$ is nonexplosive by Proposition 4.15.

We claim that the stochastic flow $(\mathbf{X}_{s,t})_{s\leq t}$ and the backward stochastic flow $(\mathbf{Y}_{t,s})_{t\geq s}$ are dual with respect to ψ_{add}, i.e.,

$$\psi_{\mathrm{add}}(\mathbf{X}_{s,t}(x), y) = \psi_{\mathrm{add}}(x, \mathbf{Y}_{t,s}(y)) \qquad (s \leq t, \ x \in \mathcal{S}(\Lambda), \ y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)). \qquad (6.25)$$

Indeed, this follows from (6.7) by writing

$$\psi_{\mathrm{add}}\big(x, \mathbf{Y}_{t,s}(y)\big) = \psi_{\mathrm{add}}\big(\,\cdot\,, \mathbf{Y}_{t,s}(y)\big)(x) = \mathbb{F}_{t,s}\big(\psi_{\mathrm{add}}(\,\cdot\,,y)\big)(x)$$
$$= \psi_{\mathrm{add}}(\,\cdot\,,y) \circ \mathbf{X}_{s,t}(x) = \psi_{\mathrm{add}}\big(\mathbf{X}_{s,t}(x), y\big).$$

By the general arguments given in Section 6.1, the duality of the flows $(\mathbf{X}_{s,t})_{s \le t}$ and $(\mathbf{Y}_{t,s})_{t \ge s}$ implies that $(X_t)_{t \ge 0}$ is pathwise dual to $(Y_t)_{t \ge 0}$ and (6.18) holds.

We note that for each $u \in \mathbb{R}$ and $y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)$, if we set

$$Y_t^{u,y} := \mathbf{Y}_{u,t} \qquad \big(t \in (-\infty, u],\ y \in \mathcal{S}(\Lambda)\big),$$

then by formula (4.17) of Proposition 4.15, translated via the bijection in (6.23) into our present setting, the function $(Y_t^{u,y})_{t \in (-\infty,u]}$ is the unique solution with initial state $Y_u^{u,y} = y$ to the evolution equation

$$Y_{t-}^{u,y} = \hat{\mathfrak{m}}_t^\omega(Y_t^{u,y}) \qquad (t \le u), \tag{6.26}$$

where $\hat{\mathfrak{m}}_t^\omega$ is the dual of the map $\mathfrak{m}_t^\omega$ defined in (4.9) (with $\hat{1} = 1$). If the dual rates $(r_{\hat{m}})_{\hat{m} \in \hat{\mathcal{G}}}$ satisfy condition (4.11) of Theorem 4.13, then[3] the evolution equation (6.26) has a unique solution for each $y \in \mathcal{S}(\Lambda)$, and setting

$$\mathbf{Y}_{u,t} := Y_t^{u,y} \qquad \big(u \ge t,\ y \in \mathcal{S}(\Lambda)\big)$$

defines a backward stochastic flow $(\mathbf{Y}_{u,t})_{u \ge t}$ on $\mathcal{S}(\Lambda)$ that extends the backward stochastic flow on $\mathcal{S}_{\mathrm{fin}}(\Lambda)$ that we had so far. We claim that (6.25) now holds for all $x, y \in \mathcal{S}(\Lambda)$. To see this, we choose $y_n \in \mathcal{S}_{\mathrm{fin}}(\Lambda)$ such that $y_n \uparrow y$. By the continuity and monotonicity of $\mathbf{Y}_{t,s}$ this implies that $\mathbf{Y}_{t,s}(y_n) \uparrow \mathbf{Y}_{t,s}(y)$ which in turn implies that $\psi_{\mathrm{add}}\big(x, \mathbf{Y}_{t,s}(y_n)\big) \uparrow \psi_{\mathrm{add}}\big(x, \mathbf{Y}_{t,s}(y)\big)$. On the other side of the equation, $\psi_{\mathrm{add}}\big(\mathbf{X}_{s,t}(x), y_n\big) \uparrow \psi_{\mathrm{add}}\big(\mathbf{X}_{s,t}(x), y\big)$, proving the claim. It follows that (6.18) holds for general $\mathcal{S}(\Lambda)$ valued initial states Y_0, completing the proof of Theorem 6.4. ∎

Remark The key to the proof of Theorem 6.4 was formula (6.20), which says that the backward in time process from Propostion 4.15 leaves the space $\mathcal{C}_{\mathrm{add}}\big(\{0,1\}^\Lambda, \{0,1\}\big)$ of additive functions invariant. Via the bijection in (6.23), the dual process $(Y_t)_{t \ge 0}$ with generator $\hat{G}$ simply corresponds to the backward in time process restricted to the space of additive functions. We will see later that for cancellative duality, a very similar picture holds, except

---

[3]We apply Theorem 4.13 backwards in time here and use a "left-continuous" version of Theorem 4.13. This "left-continuous" version of Theorem 4.13 can easily be derived from Theorem 4.13 in the same way as in the proof of Proposition 4.15 when we used Theorem 2.17 to show that (4.17) has a unique solution.

that here the space of cancellative functions is an invariant subspace for the backward in time process. Generalizing the results of the present section, one can show that each interacting particle system that can be represented in monotone local maps has a dual that arises from the fact that in this case, the space of monotone functions is an invariant subspace for the backward in time process. This idea is worked out in [LS23b]. The earlierst formulation of this monotone duality is due to Gray [Gra86].

The proof of Theorem 6.4 yields a useful corollary.

Corollary 6.6 (Pathwise additive duality) *Under the assumptions of Theorem 6.4, for each $u \in \mathbb{R}$ and $y \in \mathcal{S}_{\mathrm{fin}}$, there exists a unique solution $(Y_t^{u,y})_{t \in (-\infty, u]}$ with initial state $Y_u^{u,y} = y$ to the evolution equation*

$$Y_{t-}^{u,y} = \hat{\mathfrak{m}}_t^{\omega}(Y_t^{u,y}) \qquad (t \le u).$$

Setting

$$\mathbf{Y}_{u,t} := Y_t^{u,y} \qquad \big(u \ge t, \ y \in \mathcal{S}(\Lambda)\big)$$

defines a backward stochastic flow $(\mathbf{Y}_{u,t})_{u \ge t}$ on $\mathcal{S}_{\mathrm{fin}}(\Lambda)$ that is dual to $(\mathbf{X}_{s,t})_{s \le t}$ with respect to ψ_{add}, i.e.,

$$\psi_{\mathrm{add}}\big(\mathbf{X}_{s,t}(x), y\big) = \psi_{\mathrm{add}}\big(x, \mathbf{Y}_{t,s}(y)\big) \qquad \big(s \le t, \ x \in \mathcal{S}(\Lambda), \ y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)\big).$$

If $u \in \mathbb{R}$ and Y_0 is an $\mathcal{S}_{\mathrm{fin}}(\Lambda)$-valued random variable, independent of ω, then setting

$$Y_t := \mathbf{Y}_{u,u-t}(Y_0) \qquad (t \ge 0)$$

defines a continuous-time Markov chain $(Y_t)_{t \ge 0}$ with left-continuous sample paths, generator $\hat{G}$ as in (6.17), and state space $\mathcal{S}_{\mathrm{fin}}(\Lambda)$. If the dual rates $(r_{\hat{m}})_{\hat{m} \in \hat{\mathcal{G}}}$ satisfy (4.11), then all previous statements remain true if $\mathcal{S}_{\mathrm{fin}}(\Lambda)$ is replaced by $\mathcal{S}(\Lambda)$ and $(Y_t)_{t \ge 0}$ is the interacting particle system with generator $\hat{G}$.

The additive duality of Theorem 6.4 and its pathwise formulation in Corollary 6.6 have a simple graphical interpretation. We recall from (6.19) that we can obtain the dual of an additive local map by reversing the direction of all arrows and keeping all blocking symbols. In Section 6.2, we constructed a contact process $(X_t)_{t \ge 0}$ from its graphical representation in terms of open paths, as follows:
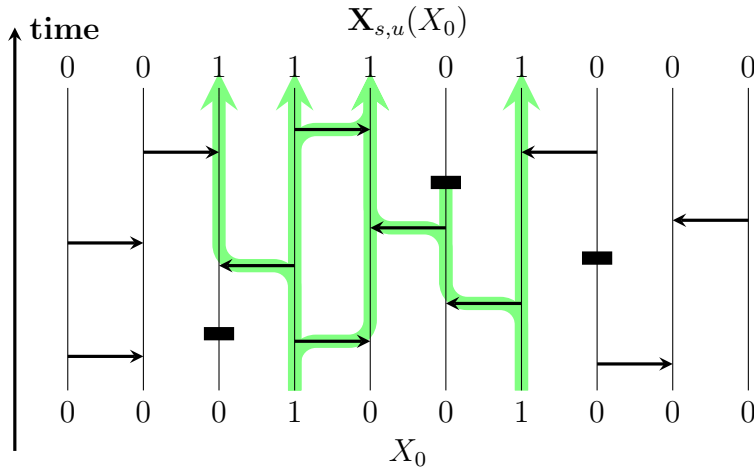
Using the recipe "reverse the arrows, keep the blocking symbols" we can construct the dual process $(Y_t)_{t\geq 0}$ as follows.
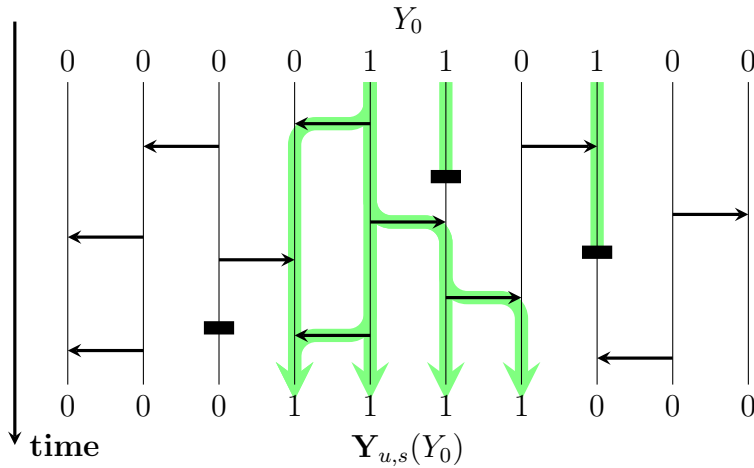
The duality relation between the forward and backward stochastic flows has a simple interpretation in terms of open paths. Indeed

$$\psi_{\mathrm{add}}\big(\mathbf{X}_{s,u}(x), y\big) = 1$$
$$\Leftrightarrow \exists i, j \in \Lambda \text{ s.t. } x(i) = 1, \ y(j) = 1, \ (i,s) \rightsquigarrow (j,u)$$
$$\Leftrightarrow \psi_{\mathrm{add}}\big(x, \mathbf{Y}_{u,s}(y)\big) = 1.$$

In our previous example of the contact process, the dual process is also a contact process, but in general, the dual process can have a different dynamics from the forward in time process. Applying the principle "reverse the arrows and keep the blocking symbols" to the voter model map $\mathtt{vot}_{i,j}$, we obtain the coalescing random walk map $\mathtt{rw}_{j,i}$, and vice versa:

$$\texttt{vot}_{1,2} \qquad\qquad \texttt{rw}_{2,1}$$

We see from this that the additive dual of the voter model is a system of coalescing random walks.

Exercise 6.7 *Give an example of an additive particle system for which the forward generator G satisfies the summability condition (4.11) but the dual generator $\hat{G}$ does not satisfy (4.11). Hint: consider a contact process and its dual contact process on a binary tree where for the forward process all infections point away from the root and for the dual process all infections point in the direction of the root.*

6.4   Cancellative systems

Let $\oplus$ denote addition modulo two, i.e.,

$$0 \oplus 0 := 0, \quad 0 \oplus 1 := 1, \quad 1 \oplus 0 := 1, \quad \text{and} \quad 1 \oplus 1 := 0.$$

Let Λ and Δ be countable sets. For $x, y \in \mathcal{S}(\Lambda)$, we define $(x \oplus y)(i) := x(i) \oplus y(i)$ $(i \in \Lambda)$ in a pointwise way. By definition, a map $m : \mathcal{S}(\Lambda) \to \mathcal{S}(\Delta)$ is *cancellative* iff

(i) $m(\underline{0}) = \underline{0}$,

(ii) $m(x \oplus y) = m(x) \oplus m(y) \qquad \big(x, y \in \mathcal{S}(\Lambda)\big)$.

We will later also sometimes need maps $m : \mathcal{S}_{\mathrm{fin}}(\Lambda) \to \mathcal{S}(\Delta)$ that satisfy (i) and (ii) but cannot reasonably be extended to $\mathcal{S}(\Lambda)$. We call these maps cancellative too. An interacting particle system is called *cancellative* if its generator can be represented in cancellative local maps. Examples of cancellative maps are:

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

- The annihilating branching map $\texttt{bran}_{ij}$ defined in (6.27) below.

- The death map $\texttt{death}_i$ defined in (1.7).

- The annihilating random walk map $\texttt{arw}_{ij}$ defined in (1.22).

- The exclusion map $\mathtt{excl}_{ij}$ defined in (1.23).

Here, we define an *annihilating branching map* $\mathtt{bran}_{ij} : \{0,1\}^\Lambda \to \{0,1\}^\Lambda$ as

$$\mathtt{bran}_{ij}(x)(k) := \begin{cases} x(i) \oplus x(j) & \text{if } k = j, \\ x(k) & \text{otherwise.} \end{cases} \tag{6.27}$$

We let $\mathcal{C} = \mathcal{C}(\{0,1\}^\Lambda, \{0,1\}^\Delta)$ as before denote the space of continuous functions $m : \{0,1\}^\Lambda \to \{0,1\}^\Delta$, and se set

$$\mathcal{C}_{\mathrm{canc}} = \mathcal{C}_{\mathrm{canc}}\big(\{0,1\}^\Lambda, \{0,1\}^\Delta\big) := \big\{m \in \mathcal{C} : m \text{ is cancellative}\big\}.$$

The following lemma is similar to Lemma 6.1.

Lemma 6.8 (Matrix formulation) *Let Λ and Δ be countable sets and let $M : \Lambda \times \Delta \to \{0,1\}$ be a function such that $\sum_{i \in \Lambda} M(i,j) < \infty$ for all $j \in \Delta$. Then there exists a unique $m \in \mathcal{C}_{\mathrm{canc}}(\{0,1\}^\Lambda, \{0,1\}^\Delta)$ such that $m(e_i)(j) = M(i,j)$ $(i \in \Lambda, \ j \in \Delta)$, which is given by*

$$m(x)(j) := \bigoplus_{i \in \Lambda} x(i) M(i,j) \qquad \big(j \in \Delta, \ x \in \{0,1\}^\Lambda\big). \tag{6.28}$$

Proof Given M, we define m by (6.28). It is easy to see that m is cancellative and satisfies $m(e_i)(j) = M(i,j)$ $(i \in \Lambda, \ j \in \Delta)$. The condition $\sum_{i \in \Lambda} M(i,j) < \infty$ guarantees that $m(x)(j)$ depends on finitely many coordinates, so m is continuous by Lemma 2.18. To prove uniqueness, assume that $m' \in \mathcal{C}_{\mathrm{add}}(\{0,1\}^\Lambda, \{0,1\}^\Delta)$ also satisfies $m'(e_i)(j) = M(i,j)$ $(i \in \Lambda, \ j \in \Delta)$. Then

$$m'(x)(j) = \bigoplus_{i:\, x(i)=1} m(e_i)(j) = \bigoplus_{i \in \Lambda} x(i) M(i,j) \qquad \big(j \in \Delta, \ x \in \mathcal{S}_{\mathrm{fin}}(\Lambda)\big).$$

Since $\mathcal{S}_{\mathrm{fin}}(\Lambda)$ is dense in $\mathcal{S}(\Lambda)$, by the continuity of m and m', it follows that $m' = m$. ∎

We now specialize to the case $\Lambda = \Delta$. We represent continuous cancellative maps by arrows and blocking symbols exactly in the same way as we did for additive maps:

- For each $i, j \in \Lambda$ with $i \neq j$ such that $m(e_i)(j) = 1$, we draw an arrow from i to j.

- For each $i \in \Lambda$ such that $m(e_i)(i) = 0$, we draw a blocking symbol ■ at i.

The following lemma, which is similar to Lemma 6.2, shows that cancellative local maps are fully described by their arrows and blocking symbols.

Lemma 6.9 (Cancellative local maps) *Let m be a continuous cancellative map, let $j \in \Lambda$, and let $J \subset \Lambda$ be defined as:*

$$J := \big\{i \in \Lambda : i \neq j \text{ and there is an arrow from } i \text{ to } j$$
$$\text{or } i = j \text{ and there is no blocking symbol at } j\big\}.$$

Then for each $x \in \mathcal{S}(\Lambda)$, one has $m(x)(j) = 1$ if and only if $|\{i \in J : x(i) = 1\}|$ is odd. Moreover, a cancellative map $m : \{0,1\}^{\Lambda} \to \{0,1\}^{\Lambda}$ is local if and only if its description involves only finitely many arrows and blocking symbols.

Proof The proof is very similar to the the proof of Lemma 6.2. In particular, the definition of a continuous cancellative map in terms of arrows and blocking symbols is just a reformulation of (6.28). ∎

Every graphical representation involving arrows and blocking symbols that can be used to define an additive particle system can also be used to define a cancellative particle system. The cancellative maps mentioned above have the following representations in terms of arrows and blocking symbols:

$$\texttt{bran}_{1,2} \qquad\qquad \texttt{death}_1$$

$$\texttt{vot}_{1,2} \qquad\qquad \texttt{arw}_{1,2} \qquad\qquad \texttt{excl}_{1,2}$$
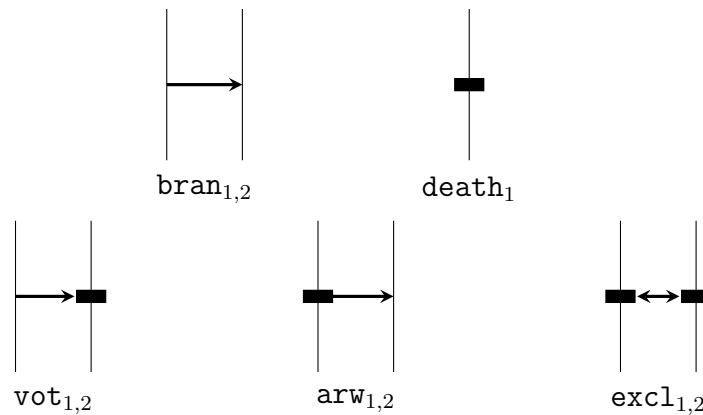
If we interpret the graphical representation of a contact process in a cancellative way, then it becomes a graphical representation for an interacting particle system involving the annihilating branching map $\texttt{bran}_{ij}$ and the death map $\texttt{death}_i$, that has been studied in [BDD91].
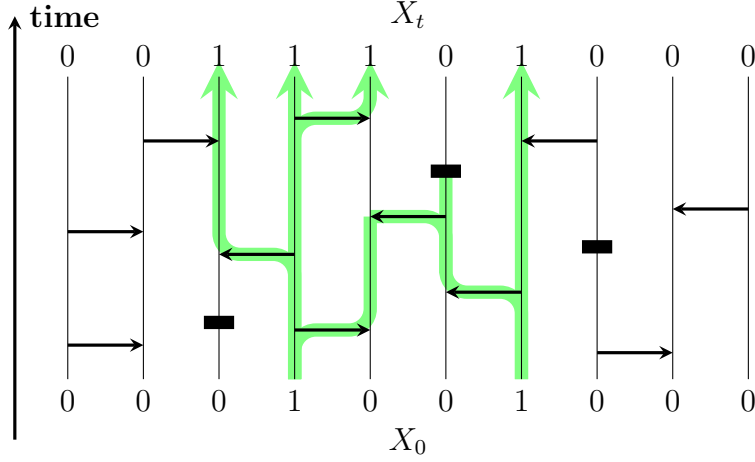
We define a *cancellative duality function* $\psi_{\mathrm{canc}} : \mathcal{S}(\Lambda) \times \mathcal{S}_{\mathrm{fin}}(\Lambda) \to \{0,1\}$ by

$$\psi_{\mathrm{canc}}(x,y) := \bigoplus_{i \in \Lambda} x(i)y(i) \qquad \big(x \in \mathcal{S}(\Lambda), \ y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)\big). \qquad (6.29)$$

Note that since $y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)$, all but finitely many of the summands are zero, so the infinite sum modulo two is well-defined. Unlike in the additive case, there is no way to make sense of $\psi_{\mathrm{canc}}(x,y)$ for general $x,y \in \mathcal{S}(\Lambda)$.[4] The following lemma is similar to Lemma 6.3 but there are some complications since infinite sums modulo two are in general undefined.

Lemma 6.10 (Dual cancellative maps) *For each continuous cancellative map $m : \mathcal{S}(\Lambda) \to \mathcal{S}(\Lambda)$, there exists a unique dual map $\hat{m} : \mathcal{S}_{\mathrm{fin}}(\Lambda) \to \mathcal{S}_{\mathrm{fin}}(\Lambda)$ such that*

$$\psi_{\mathrm{canc}}\big(m(x),y\big) = \psi_{\mathrm{canc}}\big(x,\hat{m}(y)\big) \qquad \big(x \in \mathcal{S}(\Lambda), \ y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)\big). \qquad (6.30)$$

The dual map $\hat{m}$ is cancellative and can be extended to a continuous cancellative map $\hat{m} : \mathcal{S}(\Lambda) \to \mathcal{S}(\Lambda)$ if and only if the matrix M of m satisfies $\sum_j M(i,j) < \infty$ for all $i \in \Lambda$. If m is local then so is $\hat{m}$.

Proof Let $\mathcal{M}$ be the space of $\{0,1\}$-valued matrices $(M(i,j))_{i,j \in \Lambda}$ indexed by Λ and let

$$\mathcal{M}_1 := \big\{ M \in \mathcal{M} : \sum_{i \in \Lambda} M(i,j) < \infty \ \forall j \in \Lambda \big\},$$

$$\mathcal{M}_2 := \big\{ M \in \mathcal{M} : \sum_{j \in \Lambda} M(i,j) < \infty \ \forall i \in \Lambda \big\}.$$

---

[4] For interacting particle systems on $\Lambda = \mathbb{Z}$, it is sometimes useful to consider the case that $\sup\{i : x(i) = 1\} < \infty$ and $\inf\{i : y(i) = 1\} > -\infty$. Clearly, $\psi_{\mathrm{canc}}(x,y)$ is well-defined for such x,y, even though both may be infinite.

Lemma 6.8 says that there is a one-to-one correspondence between $\mathcal{M}_1$ and the space $\mathcal{C}_{\mathrm{canc}}(\mathcal{S}(\Lambda), \mathcal{S}(\Lambda))$ of continuous cancellative maps $m : \mathcal{S}(\Lambda) \to \mathcal{S}(\Lambda)$. We claim that similarly, there is a one-to-one correspondence between $\mathcal{M}_2$ and the space $\mathcal{F}_{\mathrm{canc}}(\mathcal{S}_{\mathrm{fin}}(\Lambda), \mathcal{S}_{\mathrm{fin}}(\Lambda))$ of all cancellative maps $m : \mathcal{S}_{\mathrm{fin}}(\Lambda) \to \mathcal{S}_{\mathrm{fin}}(\Lambda)$. Indeed, this follows by writing

$$m(x)(j) = m\Big(\bigoplus_{i:\, x(i)=1} e_i \Big) = \bigoplus_{i \in \Lambda} x(i) m(e_i)(j) =: \bigoplus_{i \in \Lambda} x(i) M(i,j),$$

and observing that $m(x) \in \mathcal{S}_{\mathrm{fin}}(\Lambda)$ for all $x \in \mathcal{S}_{\mathrm{fin}}(\Lambda)$ if and only if $M \in \mathcal{M}_2$.

Inserting $x = e_i$ and $y = e_j$ in (6.30) yields the condition

$$m(e_i)(j) = \hat{m}(e_j)(i) \qquad (i, j \in \Lambda), \tag{6.31}$$

which suggests that we define

$$\hat{m}(y) := \sum_{j \in \Lambda} y(j) \hat{M}(j,i) \quad \text{with} \quad \hat{M}(j,i) := M(i,j) \quad (i, j \in \Lambda).$$

Since $M \in \mathcal{M}_1$ we have $\hat{M} \in \mathcal{M}_2$ so by our earlier remarks, this defines a cancellative map $m : \mathcal{S}_{\mathrm{fin}}(\Lambda) \to \mathcal{S}_{\mathrm{fin}}(\Lambda)$. This map can be extended to a continuous cancellative map $\hat{m} : \mathcal{S}(\Lambda) \to \mathcal{S}(\Lambda)$ if and only if $M \in \mathcal{M}_2$.

Since

$$\psi_{\mathrm{canc}}\big(m(x), y\big) = \sum_{i,j} x(i) M(i,j) y(j) = \psi_{\mathrm{canc}}\big(x, \hat{m}(y)\big)$$

for all $x \in \mathcal{S}(\Lambda)$ and $y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)$ we see that $\hat{m}$ satisfies (6.30). Formula (6.30) implies that

$$\hat{m}(y)(i) = \psi_{\mathrm{canc}}\big(e_i, \hat{m}(y)\big) = \psi_{\mathrm{canc}}\big(m(e_i), y\big) \qquad \big(i \in \Lambda,\ y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)\big),$$

so $\hat{m}$ is the only map from $\mathcal{S}_{\mathrm{fin}}(\Lambda)$ into itself that satisfies (6.30). By (6.31), the graphical representation of $\hat{m}$ can be constructed from the graphical representation of m by reversing all arrows and keeping the blocking symbols. It follows that if m is local then so is $\hat{m}$. ∎

The duality theory for cancellative systems is very similar to the additive case. We note that the duality of stochastic flows (6.35) below implies that The Markov processes $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ are dual with respect to the duality function ψ_{canc}, i.e.,

$$\mathbb{E}\big[\psi_{\mathrm{canc}}(X_t, Y_0)\big] = \mathbb{E}\big[\psi_{\mathrm{canc}}(X_0, Y_t)\big] \qquad (t \geq 0) \tag{6.32}$$

whenever X_t is independent of Y_0 and X_0 is independent of Y_t.

Theorem 6.11 (Cancellative duality) *Let G be the generator of an interacting particle system $(X_t)_{t \geq 0}$ with state space $\mathcal{S}(\Lambda)$. Assume that G has a random mapping representation of the form (4.1) such that all local maps $m \in \mathcal{G}$ are cancellative and the rates $(r_m)_{m \in \mathcal{G}}$ satisfy (4.11). Set $\hat{\mathcal{G}} := \{\hat{m} : m \in \mathcal{G}\}$ and $r_{\hat{m}} := r_m$ $(m \in \mathcal{G})$, where $\hat{m}$ denotes the cancellatove dual of m. Then*

$$\hat{G} := \sum_{\hat{m} \in \hat{\mathcal{G}}} r_{\hat{m}} \{ f(\hat{m}(y)) - f(y) \} \qquad (y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)) \qquad (6.33)$$

is the generator of a nonexplosive continuous-time Markov chain $(Y_t)_{t \geq 0}$ with state space $\mathcal{S}_{\mathrm{fin}}(\Lambda)$. Almost surely, for each $u \in \mathbb{R}$ and $y \in \mathcal{S}_{\mathrm{fin}}$, there exists a unique solution $(Y_t^{u,y})_{t \in (-\infty,u]}$ with initial state $Y_u^{u,y} = y$ to the evolution equation

$$Y_{t-}^{u,y} = \hat{\mathbf{m}}_t^\omega(Y_t^{u,y}) \qquad (t \leq u).$$

Setting

$$\mathbf{Y}_{u,t} := Y_t^{u,y} \qquad (u \geq t, \ y \in \mathcal{S}(\Lambda)) \qquad (6.34)$$

defines a backward stochastic flow $(\mathbf{Y}_{u,t})_{u \geq t}$ on $\mathcal{S}_{\mathrm{fin}}(\Lambda)$ that is dual to $(\mathbf{X}_{s,t})_{s \leq t}$ with respect to ψ_{canc}, i.e.,

$$\psi_{\mathrm{canc}}(\mathbf{X}_{s,t}(x), y) = \psi_{\mathrm{canc}}(x, \mathbf{Y}_{t,s}(y)) \qquad (s \leq t, \ x \in \mathcal{S}(\Lambda), \ y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)).$$
$$(6.35)$$

If $u \in \mathbb{R}$ and Y_0 is an $\mathcal{S}_{\mathrm{fin}}(\Lambda)$-valued random variable, independent of ω, then setting

$$Y_t := \mathbf{Y}_{u,u-t}(Y_0) \qquad (t \geq 0)$$

defines a continuous-time Markov chain $(Y_t)_{t \geq 0}$ with left-continuous sample paths, generator $\hat{G}$ as in (6.33), and state space $\mathcal{S}_{\mathrm{fin}}(\Lambda)$.

Proof The proof is basically identical to the proof of Theorem 6.4 and Corollary 6.6. The basic observation is that the backward stochastic flow defined in (6.7) maps the space $\mathcal{C}_{\mathrm{canc}}(\mathcal{S}(\Lambda), \{0,1\})$ into itself. By Lemma 6.8,

$$\mathcal{S}_{\mathrm{fin}}(\Lambda) \ni y \mapsto \psi_{\mathrm{canc}}(\,\cdot\,, y) \in \mathcal{C}_{\mathrm{canc}}(\mathcal{S}(\Lambda), \{0,1\})$$

is a bijection. Using this bijection, we can identify the backward in time process $(\Phi_t)_{t \geq 0}$ restricted to $\mathcal{C}_{\mathrm{canc}}(\mathcal{S}(\Lambda), \{0,1\})$ with the continuous-time Markov chain $(Y_t)_{t \geq 0}$ with generator $\hat{G}$ as in (6.33). Since everything is completely analogous to the proof of Theorem 6.4, we skip the details. ∎

The following lemma is similar to Lemma 6.5.

Lemma 6.12 (Distribution determining functions) *The functions* $\{f_y :$
$y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)\}$ *with* $f_y(x) := \bigoplus_{i \in \Lambda} x(i)y(i)$ *are distribution determining on*
$\mathcal{S}(\Lambda)$.

Proof We may equivalently prove that the functions $g_y(x) := 1 - 2f_y =$
$(-1)^{f_y(x)} = (-1)^{\sum_i x(i)y(i)}$ are distribution determining. Since $g_y g_{y'} = g_{y \oplus y'}$,
the class $\{g_y : y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)\}$ is closed under products and since $g_{e_i}(x) =$
$(-1)^{x(i)}$ this class separates points. The claim now follows from Lemma 4.31.
∎

Formulas (6.31) and (6.34) say that as in the additive case, the graphical
representation of the dual process can be obtained from the graphical repre-
sentation of the forward process by reversing all arrows, keeping the blocking
symbols, and then reading the graphical representation backwards in time
(downwards in our pictures). From this, we see that

$$\widehat{\mathrm{bran}}_{ij} = \mathrm{bran}_{ji}, \quad \widehat{\mathrm{death}}_i = \mathrm{death}_i, \quad \widehat{\mathrm{vot}}_{ij} = \mathrm{arw}_{ji}, \quad \widehat{\mathrm{excl}}_{ij} = \mathrm{excl}_{ij}.$$

In terms of the graphical representation, the duality between the forward and
backward stochastic flows has a nice interpretation in terms of open paths.
Indeed

$$\psi_{\mathrm{add}}\big(\mathbf{X}_{s,u}(x), y\big) = 1$$
$$\Leftrightarrow \text{ the number of open path } \gamma : [s,t] \to \Lambda$$
$$\text{s.t. } X_0(\gamma_s) = 1 \text{ and } Y_0(\gamma_u) = 1 \text{ is odd}$$
$$\Leftrightarrow \psi_{\mathrm{add}}\big(x, \mathbf{Y}_{u,s}(y)\big) = 1.$$

The similarity between additive and cancellative duality is so large that one
wonders if they can be treated in a unified way. This is indeed the case. For
local state spaces with three or more elements, an obvious thing one can do
is to replace the addition modulo two from cancellative systems by addition
modulo three or more. There are also less obvious other possibilities. The
paper [LS21] explores dualities where $(\{0, 1\}, \vee)$ or $(\{0, 1\}, \oplus)$ are replaced
by commutative monoids or semirings.

Exercise 6.13 *Show that the map* rebel_{ijk} *is cancellative. Show that the*
generator of the Neuhauser-Pacala model defined in (1.18) can be represented
as

$$G_{\mathrm{NP}}f(x) = \frac{\alpha}{|\mathcal{N}_i|} \sum_i \sum_{j \in \mathcal{N}_i} \big\{ f\big(\mathrm{vot}_{ji}(x)\big) - f\big(x\big) \big\}$$
$$= \frac{1-\alpha}{|\mathcal{N}_i|^2} \sum_i \sum_{\substack{j,k \in \mathcal{N}_i \\ j \neq k}} \big\{ f\big(\mathrm{rebel}_{kji}(x)\big) - f\big(x\big) \big\}.$$

Exercise 6.14 *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*

$$G_{\mathrm{thresh}} f(x) = 2^{-|\mathcal{N}_i|+1} \sum_i \sum_{\substack{\Delta \subset \mathcal{N}_i \cup \{i\} \\ |\Delta| \ is \ even}} \big\{ f\big(m_{\Delta,i}(x)\big) - f(x) \big\},$$

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 6.15 *Show that the threshold voter model is monotone.*

6.5   Other dualities

The additive systems duality function (6.15) and cancellative systems duality function (6.29) are not the only choices of ψ that lead to useful dualities. There are two approaches to finding useful duality functions: the pathwise approach, that aims to find dualities between stochastic flows in the sense of (6.4), and the algebraic approach, that only aims to prove distributional relations of the form (6.6). There has been a lot of recent work using the algebraic approach, starting with the paper [GK+09], linking dualities to representations of Lie algebras. For an overview of this work, we refer to [GF24]. We will below present some older results, based on the algebraic approach, due to Lloyd and Sudbury [SL95, SL97, Sud99].

The pathwise approach always depends on finding a clever stochastic flow and then finding a suitable space of functions on S^Λ that is mapped into itself by the stochastic flow of the backward in time process. As we have seen, for additive and cancellative systems, the spaces $\mathcal{C}_{\mathrm{add}}(\{0,1\}^\Lambda, \{0,1\})$ and $\mathcal{C}_{\mathrm{canc}}(\{0,1\}^\Lambda, \{0,1\})$ are invariant, and this naturally leads to additive and cancellative duality.

To explain a bit about the algebraic approach, which only aims to prove relations of the form (6.6) without proving a duality of stochastic flows, for technical simplicity, for the remainder of this section we will restrict ourselves to finite state spaces. In general, when trying to prove a duality for interacting particle systems on infinite lattices, it is often a good idea to first prove the result on finite lattices and then extend it to infinite lattices

using approximation results such as Theorem 4.35 and Corollary 4.36. We will demonstrate this method in Section 6.6 below.

In the following lemma, $\mathbb{E}^x$ (resp. E^y) denotes expectation with respect to the law of the process X started in $X_0 = x$ (resp. Y started in $Y_0 = y$).
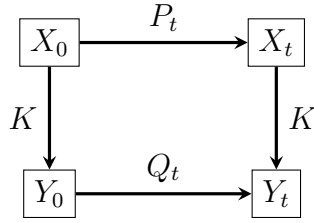
Lemma 6.16 (Duality of finite Markov processes) *Let $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ be Markov processes with finite state spaces S and R, generators G and H, and Markov semigroups $(P_t)_{t\geq 0}$ and $(Q_t)_{t\geq 0}$. Then one has*

$$\mathbb{E}^x\big[\psi(X_t, y)\big] = \mathbb{E}^y\big[\psi(x, Y_t)\big] \qquad (x \in S,\ y \in R,\ t \geq 0) \tag{6.36}$$

if and only if

$$G\psi(\,\cdot\,, y)(x) = H\psi(x, \,\cdot\,)(y). \tag{6.37}$$

Proof The duality relation (6.36) says that

$$\sum_{x' \in S} P_t(x, x')\psi(x', y) = \sum_{y' \in R} \psi(x, y')Q_t(y, y') \qquad (x \in S,\ y \in R,\ t \geq 0),$$

which can in matrix form be written as

$$P_t\psi = \psi Q_t^\dagger \qquad (t \geq 0), \tag{6.38}$$

where $Q_t^\dagger(y', y) := Q_t(y, y')$ denotes the transpose of Q_t. Differentiating with respect to t and setting $t = 0$, it follows that

$$G\psi = \psi H^\dagger$$

which is just a more formal way of writing (6.37). Conversely, if (6.37) holds, then $G^2\psi = G\psi H^\dagger = \psi(H^\dagger)^2$ and by induction $G^n\psi = \psi(H^\dagger)^n$ for all $n \geq 0$. Using the fact that

$$P_t = \sum_{n=0}^{\infty} \frac{1}{n!}t^n G^n \quad \text{and} \quad Q_t = \sum_{n=0}^{\infty} \frac{1}{n!}t^n H^n,$$

it follows that $P_t\psi = \psi Q_t^\dagger$ $(t \geq 0)$ and hence (6.36) holds. $\blacksquare$

Let $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ be Markov processes with finite state spaces S and R, generators G and H, and Markov semigroups $(P_t)_{t\geq 0}$ and $(Q_t)_{t\geq 0}$. Let K be a probability kernel from S to R. A relation of the form (compare (6.38))

$$P_t K = K Q_t \qquad (t \geq 0) \tag{6.39}$$

is called an *intertwining* of Markov processes. Note that (6.39) says that the following two procedures are equivalent for each S-valued random variable X_0:

- Evolve the state X_0 for time t under the evolution of the Markov process $(X_t)_{t \geq 0}$, then map the outcome X_t into a random variable Y_t using the kernel K.

- Map X_0 into a random variable Y_0 using the kernel K, then evolve Y_0 for time t under the evolution of the Markov process $(Y_t)_{t \geq 0}$.

We can summarize the situation in the following commutative diagram:

$$
\begin{array}{ccc}
\boxed{X_0} & \xrightarrow{\;\;P_t\;\;} & \boxed{X_t} \\
{\scriptstyle K}\Big\downarrow & & \Big\downarrow{\scriptstyle K} \\
\boxed{Y_0} & \xrightarrow{\;\;Q_t\;\;} & \boxed{Y_t}
\end{array}
$$

Lemma 6.17 (Intertwining of finite Markov processes) *The inter-twining relation (6.39) is equivalent to*

$$GK = KH. \tag{6.40}$$

Proof Analogue to the proof of Lemma 6.16. ∎

As one might guess, there is a close relationship between duality and intertwining. If $(P_t)_{t \geq 0}$, $(Q_t)_{t \geq 0}$, and $(R_t)_{t \geq 0}$ are Markov semigroups, K is a probability kernel, and ψ a duality function such that

$$P_t K = K Q_t \quad \text{and} \quad Q_t \psi = \psi R_t^\dagger \qquad (t \geq 0),$$

then trivially

$$P_t(K\psi) = KQ_t\psi = (K\psi)R_t^\dagger \qquad (t \geq 0), \tag{6.41}$$

which says that the Markov processes with semigroups $(P_t)_{t \geq 0}$ and $(R_t)_{t \geq 0}$ are dual with duality function $K\psi$.

To see these general principles at work, let us look at interacting particle systems with state space of the form $\{0,1\}^\Lambda$ where Λ is finite. For each $r > 0$, we let ψ_r denote the duality function

$$\psi_r(x, y) := \prod_{i \in \Lambda} (1 - r)^{x(i)y(i)} \qquad \left(x, y \in \{0,1\}^\Lambda \right). \tag{6.42}$$

Using the fact that $0^n = 1_{\{n=0\}}$, we observe that

$$
\left.
\begin{aligned}
\psi_1(x, y) &= 1 - \psi_{\mathrm{add}}(x, y) \\
\psi_2(x, y) &= (-1)^{\psi_{\mathrm{canc}}(x, y)}
\end{aligned}
\right\}
\qquad \left(x, y \in \{0,1\}^\Lambda \right).
$$

Therefore, duality with respect to duality functions of the form (6.42) includes additive and cancellative duality as special cases. Duality functions of the form (6.42) arose from the work of Lloyd and Sudbury [SL95, SL97]. Accordingly, we will call ψ_r the *Lloyd-Sudbury duality function.*

There is a close connection between duality functions of the form (6.42) and *thinning*, as we now explain. Thinning has already been introduced in Section 2.5 but for convenience we repeat the definition here. Let $(\chi_p(i))_{i \in \Lambda}$ be i.i.d. with $\mathbb{P}[\chi_p(i) = 1] = p$ and $\mathbb{P}[\chi_p(i) = 0] = 1 - p$. Then

$$K_p(x, y) := \mathbb{P}\big[y(i) = \chi_p(i)x(i) \ \forall i \in \Lambda\big] \qquad \big(x, y \in \{0, 1\}^\Lambda\big) \qquad (6.43)$$

defines a *thinning kernel.* Note that if we interpret sites i with $x(i) = 1$ as being occupied by a particle, then the effect of K_p is to independently throw away some of these particles, where each particle has a probability p to remain. We claim that

$$K_p K_q = K_{pq} \quad \text{and} \quad K_p \psi_r = \psi_{pr} \qquad (0 \le p, q \le 1, \ r > 0). \qquad (6.44)$$

The first relation is clear from the interpretation in terms of thinning, while the second relation follows by writing

$$K_p \psi_r(x, z) = \sum_y K_p(x, y) \prod_{i \in \Lambda} (1 - r)^{y(i)z(i)} = \mathbb{E}\big[\prod_{i \in \Lambda} (1 - r)^{\chi_p(i)x(i)z(i)}\big]$$
$$= \prod_{i \in \Lambda} \mathbb{E}\big[(1 - r)^{\chi_p(i)x(i)z(i)}\big] = \prod_{i \in \Lambda} (1 - pr)^{x(i)z(i)} = \psi_{pr}(x, z).$$

The following lemma says that if a particle system has two duals, one with respect to the duality function ψ_{r_1} and the other with respect to the duality function ψ_{r_2}, then one of these duals is a thinning of the other.

Lemma 6.18 (Lloyd-Sudbury duals and thinning) *Let G, H_1 and H_2 be generators of Markov processes with state space $\{0, 1\}^\Lambda$ where Λ is finite. Let $0 < r_1 \le r_2$ and set $p := r_1/r_2$. Then of the relations*

(i) $H_1 \psi_{r_1} = \psi_{r_1} G^\dagger$,   (ii) $H_2 \psi_{r_2} = \psi_{r_2} G^\dagger$,   *and*   (iii) $H_1 K_p = K_p H_2$

any two imply the third.

Proof Using all tree relations (i)–(iii) as well as (6.44), we have the "circular" sequence of equalities:

$$H_1 \psi_{r_1} \overset{\text{(i)}}{=} \psi_{r_1} G^\dagger = K_p \psi_{r_2} G^\dagger \overset{\text{(ii)}}{=} K_p H_2 \psi_{r_2} \overset{\text{(iii)}}{=} H_1 K_p \psi_{r_2} = H_1 \psi_{r_1}.$$

From this, we immediately see that of the relations

(i) $H_1\psi_{r_1} = \psi_{r_1}G^\dagger$,   (ii)' $K_pH_2\psi_{r_2} = K_p\psi_{r_2}G^\dagger$,   (iii)' $H_1K_p\psi_{r_2} = K_pH_2\psi_{r_2}$

any two imply the third. To complete the proof, it suffices to show that K_p and ψ_r are invertible as matrices for all $p \in (0,1]$ and $r > 0$, since we can then multiply (ii)' from the left with K_p^{-1} and (iii)' from the right with $\psi_{r_2}^{-1}$ to obtain (ii) and (iii).

We can view the linear space of all functions $f : \{0,1\}^\Lambda \to \mathbb{R}$ as the tensor product $\bigotimes_{i\in\Lambda} \mathbb{R}^{\{0,1\}}$. In this picture, the matrices K_p and ψ_r are the tensor product over Λ of single-site matrices of the form

$$\begin{pmatrix} 1 & 0 \\ 1-p & p \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 1 \\ 1 & 1-r \end{pmatrix},$$

respectively. These single-site matrices are invertible for all $p \in (0,1]$ and $r \in (0,\infty)$ and hence the same is true for their tensor products K_p and ψ_r. ∎

It is useful to look at a concrete example. Let (Λ, E) be a finite graph, as in (1.2) let $\mathcal{N}_i := \{j \in \Lambda : \{i,j\} \in E\}$ denote the neighborhood of a vertex $i \in \Lambda$, let and assume that $N := |\mathcal{N}_i|$ does not depend on $i \in \Lambda$. In line with notation introduced in Section 1.1, we let $\mathcal{E} := \{(i,j) \in \Lambda^2 : \{i,j\} \in E\}$ denote the set of directed edges associated with E. Let $G_{\text{vot}}, G_{\text{rw}}\ G_{\text{arw}}$ be the Markov generators defined by

$$G_{\text{vot}}f(x) := N^{-1} \sum_{(i,j)\in\mathcal{E}} \{f(\texttt{vot}_{ij}(x)) - f(x)\},$$

$$G_{\text{rw}}f(x) := N^{-1} \sum_{(i,j)\in\mathcal{E}} \{f(\texttt{rw}_{ij}(x)) - f(x)\},$$

$$G_{\text{arw}}f(x) := N^{-1} \sum_{(i,j)\in\mathcal{E}} \{f(\texttt{arw}_{ij}(x)) - f(x)\},$$

where the voter model map $\texttt{vot}_{ij}$, the coalescing random walk map $\texttt{rw}_{ij}$, and the annihilating random walk map $\texttt{arw}_{ij}$ are defined in (1.4), (1.20), and (1.22), respectively. In words, G_{vot} is the generator of a voter model in which each site $i \in \Lambda$ adopts with rate one the type of a randomly chosen neighbor. The processes with generators G_{rw} and G_{arw} consist of coalescing and annihilating particles that jump to a randomly chosen neighboring site with rate one. We claim that

(i) $G_{\text{rw}}\psi_1 = \psi_1 G_{\text{vot}}^\dagger$,   (ii) $G_{\text{arw}}\psi_2 = \psi_2 G_{\text{vot}}^\dagger$,   (iii) $G_{\text{rw}}K_{1/2} = K_{1/2}G_{\text{arw}}$.

Indeed, the voter model map is both additive and cancellative, so (i) follows from Theorem 6.4 since the coalescing random walk map is the additive dual

of the voter model map and likewise (ii) follows from Theorem 6.11 since the annihilating random walk map is the cancellative dual of the voter model map. By Lemma 6.18, (i) and (ii) imply (iii), which says that annihilating random walks are a 1/2-thinning of coalescing random walks. In other words, for each $t \geq 0$, the following two procedures are equivalent:

- Run coalescing random walk dynamics for time t and then thin the resulting configuration with 1/2.

- Thin the initial configuration with 1/2 and then run coalescing random walk dynamics for time t.

One can also verify this directly and use this to deduce (ii) from (i) or vice versa. In Proposition 6.21 below we will see a less trivial example of a thinning relation between two interacting particle systems.

Surprisingly, there exist many duality relations between interacting particle systems with respect to the Lloyd-Sudbury duality function ψ_r for other values of r than $r = 1, 2$. These dualities can usually not be obtained as pathwise dualities.

Let (Λ, E) be a finite graph. The paper [Sud00] considers interacting particle systems on graphs where the configuration along each edge makes the following transitions with the following rates:[5]

$$
\begin{array}{lccl}
\text{``annihilation''} & \multicolumn{2}{c}{11 \mapsto 00} & \text{at rate } a, \\
\text{``branching''} & 01 \mapsto 11 & \text{and} \quad 10 \mapsto 11 & \text{each at rate } b, \\
\text{``coalescence''} & 11 \mapsto 01 & \text{and} \quad 11 \mapsto 10 & \text{each at rate } c, \\
\text{``death''} & 01 \mapsto 00 & \text{and} \quad 10 \mapsto 00 & \text{each at rate } d, \\
\text{``exclusion''} & 01 \mapsto 10 & \text{and} \quad 10 \mapsto 01 & \text{each at rate } e.
\end{array}
$$

More formally, for each $i, j \in \Lambda$, we can define a map $m_{ij}^{01 \mapsto 11}$ on $\{0,1\}^{\Lambda}$ as follows:

$$
m_{ij}^{01 \mapsto 11}(x)(k) = \begin{cases} 1 & \text{if } k = i \text{ and } \big(x(i), x(j)\big) = (0,1), \\ 1 & \text{if } k = j \text{ and } \big(x(i), x(j)\big) = (0,1), \\ x(k) & \text{in all other cases.} \end{cases}
$$

---

[5]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".

Defining $m_{ij}^{11\mapsto 00}$ etc. in a similar way, the generator of the process we are interested in can be written as

$$
\begin{aligned}
Gf(x) = \sum_{\{i,j\}} a\big\{ f\big(m_{ij}^{11\mapsto 00}(x)\big) - f(x)\big\} \\
+ \sum_{(i,j)} \Big[b\big\{ f\big(m_{ij}^{01\mapsto 11}(x)\big) - f(x)\big\} + c\big\{ f\big(m_{ij}^{11\mapsto 01}(x)\big) - f(x)\big\} \\
+ d\big\{ f\big(m_{ij}^{01\mapsto 00}(x)\big) - f(x)\big\} + e\big\{ f\big(m_{ij}^{01\mapsto 10}(x)\big) - f(x)\big\} \Big],
\end{aligned}
$$

(6.45)

where the first sum runs over all (unordered) edges $\{i,j\} \in E$ and the second sum runs over all ordered pairs (i,j) such that $\{i,j\} \in E$.

Theorem 6.19 (Lloyd-Sudbury duality) *Let G and G' be defined as in (6.45) in terms of rates a,b,c,d,e and a',b',c',d',e', respectively, and let $r > 0$. Then one has*

$$G\psi_r = \psi_r G'^{\dagger} \qquad (6.46)$$

if and only if $a' = a + 2(1-r)\gamma$, $b' = b + \gamma$, $c' = c - (2-r)\gamma$, $d' = d + \gamma$, and $e' = e - \gamma$, where $\gamma := (a + c - d + (1-r)b)/r$.

Proof This follows from Lemma 6.16 by checking (6.37). The calculations are a bit tedious, so we omit them here. They can be found in [Sud00, formula (9)], which is a simplification of [SL95, formula (21)]. ∎

6.6 The contact-voter model

As we have already seen, $\psi_1 = 1 - \psi_{\mathrm{add}}$ and $\psi_2 = (-1)^{\psi_{\mathrm{canc}}}$ correspond to additive and cancellative duality. It seems that for $r \neq 1, 2$, dualities of the form (6.46) are almost never[6] pathwise dualities. To give an example with $r \neq 1, 2$, consider an interacting particle system on a (possibly infinite) graph (Λ, E) whose dynamics are a mixture of contact process and voter model dynamics, with generator of the form:

$$
\begin{aligned}
G_{\mathrm{covo}} f(x) := \lambda \sum_{(i,j)\in\mathcal{E}} \big\{ f\big(\mathtt{bra}_{ij}(x)\big) - f(x)\big\} \\
+ \sum_{i\in\Lambda} \big\{ f\big(\mathtt{death}_i(x)\big) - f(x)\big\} \\
+ \alpha \sum_{(i,j)\in\mathcal{E}} \big\{ f\big(\mathtt{vot}_{ij}(x)\big) - f(x)\big\} \qquad (x \in \{0,1\}^{\Lambda}),
\end{aligned}
$$

(6.47)

---

[6]Except some very trivial and pathological cases.

where $\mathcal{E}$ denotes the set of oriented edges associated with E. Letting $\mathcal{N}_i := \{j \in \Lambda : \{i,j\} \in E\}$ denote the set of neighbors of i, we assume that Λ is countable and

$$\sup_{i \in \Lambda} |\mathcal{N}_i| < \infty,$$

which implies that the generator in (6.47) satisfies the summability condition (4.11) of Theorem 4.13 and hence corresponds to a well-defined interacting particle system. Such systems are studied in [DLZ14], who are especially interested in the fast-voting limit $\alpha \to \infty$. The contact-voter model is additive (but not cancellative, because the branching map is not), and by Theorem 6.4 dual with respect to the duality function $\psi_1 = 1 - \psi_{\mathrm{add}}$ to the interacting particle system with generator

$$\begin{aligned} G_{\mathrm{corw}} f(y) := &\lambda \sum_{(i,j) \in \mathcal{E}} \big\{ f\big(\mathtt{bra}_{ij}(y)\big) - f\big(y\big) \big\} \\ &+ \sum_{i \in \Lambda} \big\{ f\big(\mathtt{death}_i(y)\big) - f\big(y\big) \big\} \\ &+ \alpha \sum_{(i,j) \in \mathcal{E}} \big\{ f\big(\mathtt{rw}_{ij}(y)\big) - f\big(y\big) \big\} \qquad (y \in \{0,1\}^{\Lambda}), \end{aligned} \tag{6.48}$$

which corresponds to a system of branching and coalescing random walks. Perhaps surprisingly, the contact-voter model is also self-dual.

Proposition 6.20 (Self-duality of the contact-voter model) *Assume that $\lambda > 0$. Then the contact-voter model with generator as in (6.47) is self-dual with respect to the duality function ψ_r with $r := \lambda/(\alpha + \lambda)$.*

Proof We first consider the case that the graph (Λ, E) is finite. The generator G_{covo} is a special case of the generators considered in Theorem 6.19 and corresponds to the choice of parameters

$$a = 0, \quad b = \lambda + \alpha, \quad c = 1, \quad d = 1 + \alpha, \quad e = 0.$$

We observe that setting $r := \lambda/(\alpha + \lambda)$ makes the parameter γ from Theorem 6.19 zero, which has the effect that $a' = a$, $b' = b$, $c' = c$, $d' = d$, and $e' = e$, i.e., we have found a self-duality.

To extend the result to infinite graphs, we use an approximation argument. We need to show that

$$\mathbb{E}\big[\psi_r\big(\mathbf{X}_{0,t}(x), x'\big)\big] = \mathbb{E}\big[\psi_r\big(x, \mathbf{X}_{0,t}(x')\big)\big] \qquad \big(t \geq 0, \ x, x' \in \{0,1\}^{\Lambda}\big), \tag{6.49}$$

where $(\mathbf{X}_{s,u})_{s \leq u} = (\mathbf{X}^+_{s,u})_{s \leq u}$ denotes the stochastic flow defined by the graphical representation of the contact-voter model, and

$$\psi_r(x,y) := \prod_{i \in \Lambda} (1-r)^{x(i)y(i)} \qquad \big(x, y \in \{0,1\}^{\Lambda}\big). \tag{6.50}$$

Let (Λ_n, E_n) be finite subgraphs of (Λ, E) that increase to the whole graph. For each n, let $x_n(i) := x(i)$ if $i \in \Lambda_n$ and $:= 0$ otherwise, and define x'_n similarly in terms of x'. Note that if x and y are zero outside Λ_n, then in (6.50) it does not matter if we take the product over Λ or Λ_n. If we restrict the sums in the definition of the generator (6.47) to vertices and edges in (Λ_n, E_n), then for the restricted process $(X_t^n)_{t \geq 0}$, the coordinates $X_t^n(i)$ with $i \in \Lambda \backslash \Lambda_n$ do not evolve (keep their initial value) while the coordinates in Λ_n form a contact-voter model on a finite graph for which Theorem 6.19 is applicable. Thus, letting $(\mathbf{X}_{s,u}^n)_{s \leq u}$ denote the stochastic flow of the restricted process, we see that for each n, we have

$$\mathbb{E}\big[\psi_r\big(\mathbf{X}_{0,t}^n(x_n), x'_n\big)\big] = \mathbb{E}\big[\psi_r\big(x_n, \mathbf{X}_{0,t}^n(x'_n)\big)\big] \qquad \big(t \geq 0, \; x, x' \in \{0,1\}^\Lambda\big). \tag{6.51}$$

Since the contact-voter model is additive, we can express its stochastic flow in terms of open paths as in (6.14). There is a natural coupling of the graphical representations of the finite processes and the infinite process. In this coupling, $\mathbf{X}_{s,t}^n(x_n)(j) = 1$ if and only if there exists an $i \in \Lambda_n$ such that $x(i) = 1$ and an open path from (i, s) to (j, t) that never leaves the finite set Λ_n. Using this, we see that in this coupling, $\mathbf{X}_{s,t}^n(x_n)$ a.s. increases to $\mathbf{X}_{s,t}(x)$, i.e.,

$$\mathbf{X}_{s,t}^n(x_n) \leq \mathbf{X}_{s,t}^{n+1}(x_{n+1}) \; \forall n \quad \text{and} \quad \mathbf{X}_{s,t}^n(x_n) \underset{n \to \infty}{\longrightarrow} \mathbf{X}_{s,t}(x) \quad \text{pointwise.} \tag{6.52}$$

Since $r \leq 1$, it is easy to check that if x_n increases to x and y_n increases to y, then $\psi_r(x_n, y_n)$ decreases to $\psi_r(x, y)$. Using this, taking the limit in (6.51), we arrive at (6.49).

Note that although the duality function ψ_r is continuous with respect to increasing sequences in the way we have just described, it is in general not true that $x_n \to x$ and $y_n \to y$ pointwise imply that $\psi_r(x_n, y_n) \to \psi_r(x, y)$. This is why we based our approximation argument on a clever monotone coupling and did not use the more general Corollary 4.36 to approximate infinite systems with finite systems. ∎

We have already seen in Lemma 6.18 that there is a close connection between the Llyod-Sudbury duality functions ψ_r and thinning. The following proposition demonstrates this on our example of the contact-voter model.

Proposition 6.21 (Thinning of the contact-voter model) *Let $(P_t)_{t \geq 0}$ and $(Q_t)_{t \geq 0}$ denote the semigroups of the contact-voter model with generator as in (6.47) and the system of branching and coalescing random walks with generator as in (6.48), respectively. Let K_r denote the thinning kernel defined in (6.43) with $p := \lambda/(\alpha + \lambda)$. Then*

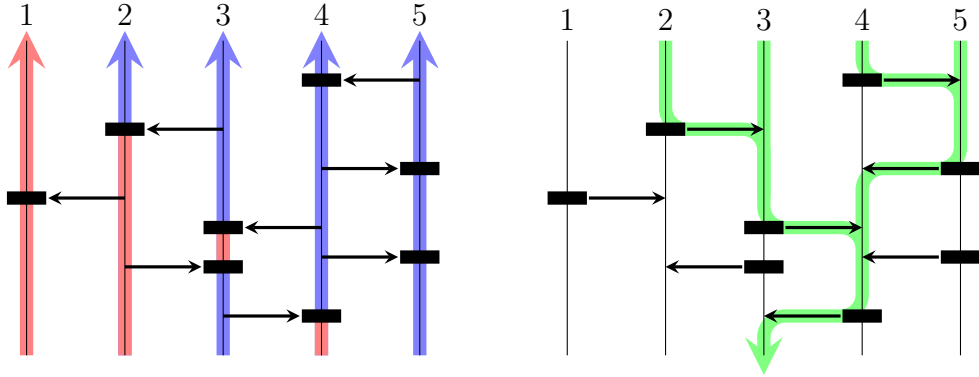$$P_t K_p = K_p Q_t \qquad (t \geq 0). \tag{6.53}$$

Figure 6.1: Graphical representation of a one-dimensional voter model and its dual system of coalescing random walks. At the final time, the points 2,3,4, and 5 have the same type, because they descend from the same ancestor.

Proof We first prove the statement for finite graphs. Additive duality tells us that (i) $G_{\mathrm{covo}}\psi_1 = \psi_1 G_{\mathrm{corw}}^\dagger$, and Proposition 6.20 tells us that (ii) $G_{\mathrm{covo}}\psi_p = \psi_p G_{\mathrm{covo}}^\dagger$. By Lemma 6.18, this implies (iii) $G_{\mathrm{covo}}K_p = K_p G_{\mathrm{corw}}$, which implies (6.53).

To also get the result for infinite graphs, we use approximation with finite graphs. In this case, the argument is simpler than in the proof of Proposition 6.20 since thinning is a continuous operation in the sense that if X^n are random variables with values in $\{0,1\}^\Lambda$ that converge weakly in law to X, and Y^n and Y are obtained from X^n and X by thinning with the kernel K_r, then the Y_n converge weakly in law to Y. As a result, we can use Corollary 4.36 to approximate infinite systems with finite systems and take the limit to get the result for infinite systems. ∎

We will continue our study of the contact-voter model in Section 6.9.

6.7   Invariant laws of the voter model

By Theorem 6.4 and Corollary 6.6,, the voter model X is pathwise dual, with respect to the additive duality function ψ_{add} from (6.15), to a collection Y of coalescing random walks. Due to the fact that $|Y_t|$ 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 6.22 (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 $\delta_{\underline{0}}$ and $\delta_{\underline{1}}$ 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, as demonstrated in Figure 6.1,

$$X_t(i) = X_{t-s}(\xi_s^{(i,t)}) \qquad (0 \leq s \leq t),$$

i.e., $\xi_s^{(i,t)}$ traces back where the site i at time t got its type from.[7]

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 $p\delta_{\underline{0}} + (1 - p)\delta_{\underline{1}}$ 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.

Theorem 6.23 (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*

$$\mathbb{P}[X_t \in \cdot] \underset{t \to \infty}{\Longrightarrow} \nu_p, \tag{6.54}$$

---

[7]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.

where ν_p is an invariant law of the process. If $d = 1, 2$, then

$$\nu_p = (1 - p)\delta_{\underline{0}} + p\delta_{\underline{1}}. \tag{6.55}$$

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 6.22, let $(\xi_s^{(i,t)})_{s \geq 0}$ be the backward random walk in the graphical representation starting at (i, t). Define a random equivalence relation $\sim$ on $\mathbb{Z}^d$ by

$$i \sim j \quad \text{iff} \quad \xi_s^{(i,0)} = \xi_s^{(j,0)} \text{ for some } s \geq 0.$$

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 (6.54) holds. Since random walk in dimensions $d = 1, 2$ is recurrent, there is a.s. only one equivalence class, and $\nu_p = (1 - p)\delta_{\underline{0}} + p\delta_{\underline{1}}$. On the other hand, since random walk in dimensions $d \geq 3$ is transient, there are a.s. infinitely many[8] equivalence classes and hence for $p \neq 0, 1$ the measure ν_p is concentrated on configurations that are not constant.

To prove (6.54), 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 $t \geq 0$, we define a random equivalence relation $\sim_t$ on $\mathbb{Z}^d$ by

$$i \sim_t j \quad \text{iff} \quad \xi_s^{(i,0)} = \xi_s^{(j,0)} \text{ for some } 0 \leq s \leq t.$$

We enumerate the elements of $\mathbb{Z}^d$ in some arbitrary way and define

$$\tilde{X}_t(i) := \chi(j) \quad \text{where } j \text{ is the smallest element of } \{k \in \mathbb{Z}^d : i \sim_t k\}. \tag{6.56}$$

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. ∎

Remark In dimensions $d \geq 3$, it is in fact known that the measures ν_p are extremal, and each extremal invariant law of the voter model is of this form. See [Lig85, Thm V.1.8].

---

[8]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 sufficiently 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 $\mathbb{P}[N = \infty] = 1$.

6.8 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 [BDD91, SS08, CP14] while [LS23a] treats a coupling of an additive and a cancellative system.

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 6.24 (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*

$$\int \overline{\nu}(\mathrm{d}x)\, x(i) = \mathbb{P}^{e_i}[X_t \neq \underline{0}\ \forall t \geq 0] \qquad (i \in \Lambda).$$

More generally, for any $y \in \{0,1\}^{\Lambda}$ such that $|y| < \infty$,

$$\int \overline{\nu}(\mathrm{d}x)\, 1_{\{x \wedge y \neq \underline{0}\}} = \mathbb{P}^{y}[X_t \neq \underline{0}\ \forall t \geq 0].$$

Proof By Theorem 6.4, the contact process X is self-dual with respect to the additive systems duality function, i.e.,

$$\mathbb{P}^{x}[X_t \wedge y = \underline{0}] = \mathbb{P}^{y}[x \wedge X_t = \underline{0}] \qquad (t \geq 0).$$

In particular, setting $x = \underline{1}$, we see that

$$\int \overline{\nu}(\mathrm{d}x)\, 1_{\{x \wedge y \neq \underline{0}\}} = \lim_{t \to \infty} \mathbb{P}^{\underline{1}}[X_t \wedge y \neq \underline{0}]$$
$$= \lim_{t \to \infty} \mathbb{P}^{y}[\underline{1} \wedge X_t \neq \underline{0}] = \mathbb{P}^{y}[X_t \neq \underline{0}\ \forall t \geq 0],$$

where in the first step we have used that $|y| < \infty$ and the measures $\mathbb{P}^{\underline{1}}[X_t \in \cdot]$ converge weakly to $\overline{\nu}$ as $t \to \infty$. The condition $|y| < \infty$ can be removed by doing this step more carefully, using monotone convergence instead of weak convergence, but since contact processes started in infinite initial states a.s. do not die out in finite time, this case is less interesting. ∎

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 (3.23) started in any nonzero initial state converge to the upper fixed point. Recall from Chapter 5 that a probability law μ on $\{0,1\}^{\mathbb{Z}^d}$ is *nontrivial* if $\mu(\{\underline{0}\}) = 0$, i.e., if μ gives zero probability to the all-zero configuration.

Theorem 6.25 (Convergence to upper invariant law) *Let $(X_t)_{t \geq 0}$ be a contact process started in a homogeneous 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 (6.14). We continue to use the notation $|x| := \sum_i x(i)$. We say that x is *finite* if $|x| < \infty$.

Lemma 6.26 (Extinction versus unbounded growth) *For each finite $x \in \{0,1\}^{\mathbb{Z}^d}$, one has*

$$X_t^x = \underline{0} \text{ for some } t \geq 0 \quad or \quad |X_t^x| \underset{t \to \infty}{\longrightarrow} \infty \quad \text{a.s.} \tag{6.57}$$

Proof Define

$$\rho(x) := \mathbb{P}\big[X_t^x \neq \underline{0} \ \forall t \geq 0\big] \qquad (x \in \{0,1\}^{\mathbb{Z}^d},\ |x| < \infty).$$

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

$$|x| \leq N \quad \text{implies} \quad \rho(x) \leq 1 - \varepsilon. \tag{6.58}$$

We first argue why it is plausible that this implies (6.57) and then give a rigorous proof. Imagine that $|X_t^x| \not\to \infty$. Then, in view of (6.58), 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

$$\mathcal{A}_x := \{X_t^x \neq \underline{0} \; \forall t \geq 0\} \qquad (x \in \{0,1\}^{\mathbb{Z}^d}, \; |x| < \infty).$$

denote the event that the process $(X_t^x)_{t \geq 0}$ 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

$$\rho(X_t^x) = \mathbb{P}\big[\mathcal{A}_x \,\big|\, \mathcal{F}_t\big] \underset{t \to \infty}{\longrightarrow} 1_{\mathcal{A}_x} \quad \text{a.s.,} \tag{6.59}$$

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 $\lim_n \mathbb{P}[\mathcal{A}|\mathcal{F}_n] = \mathbb{P}[\mathcal{A}|\mathcal{F}_\infty]$ a.s. for each measurable event $\mathcal{A}$. (See [Loe63, § 29, Complement 10 (b)].) In view of (6.58), formula (6.59) implies (6.57). ∎

Lemma 6.27 (Nonzero intersection) *Let $(X_t)_{t \geq 0}$ be a contact process with a homogeneous 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 $x \in \{0,1\}^{\mathbb{Z}^d}$*

$$|x| \geq N \quad \text{implies} \quad \mathbb{P}\big[x \wedge X_s = \underline{0}\big] \leq \varepsilon.$$

Proof By duality,

$$\mathbb{P}\big[x \wedge X_s = \underline{0}\big] = \mathbb{P}\big[X_s^x \wedge X_0 = \underline{0}\big]$$

where X_0 is independent of the graphical representation used to define X_s^x. 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

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

Then, using Hölder's inequality[9] in the inequality marked with an exclamation mark, we have

$$
\begin{aligned}
\mathbb{P}\big[X_s^x \wedge X_0 = \underline{0}\big] &= \int \mathbb{P}[X_0 \in \mathrm{d}y]\,\mathbb{P}\big[X_s^x \wedge y = \underline{0}\big] \\
&\leq \int \mathbb{P}[X_0 \in \mathrm{d}y]\,\mathbb{P}\big[\bigvee_{i:\,x'(i)=1} X_s^{\{i\}\,(M)} \wedge y = \underline{0}\big] \\
&= \int \mathbb{P}[X_0 \in \mathrm{d}y] \prod_{i:\,x'(i)=1} \mathbb{P}\big[X_s^{\{i\}\,(M)} \wedge y = \underline{0}\big] \\
&\overset{!}{\leq} \prod_{i:\,x'(i)=1} \Big(\int \mathbb{P}[X_0 \in \mathrm{d}y]\,\mathbb{P}\big[X_s^{\{i\}\,(M)} \wedge y = \underline{0}\big]^{|x'|}\Big)^{1/|x'|} \\
&= \prod_{i:\,x'(i)=1} \Big(\int \mathbb{P}[X_0 \in \mathrm{d}y]\,\mathbb{P}\big[X_s^{\{0\}\,(M)} \wedge y = \underline{0}\big]^{|x'|}\Big)^{1/|x'|} \\
&= \int \mathbb{P}[X_0 \in \mathrm{d}y]\,\mathbb{P}\big[X_s^{\{0\}\,(M)} \wedge y = \underline{0}\big]^{|x'|},
\end{aligned}
$$

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

$$
\mathbb{P}\big[x \wedge X_s = \underline{0}\big] \leq \int \mathbb{P}[X_0 \in \mathrm{d}y]\mathbb{P}\big[X_s^{\{0\}\,(M)} \wedge y = \underline{0}\big]^{N/|\Lambda_M|} =: f(N, M).
$$

Here, using the fact that

$$
\mathbb{P}\big[X_s^{\{0\}\,(M)} \wedge y = \underline{0}\big] < 1 \quad \text{if } y(i) = 1 \text{ for some } i \in \Lambda_M,
$$

we see that

$$
\lim_{N\uparrow\infty} f(N, M) = \int \mathbb{P}[X_0 \in \mathrm{d}y]\mathbb{1}_{\{y(i)=0\ \forall i \in \Lambda_M\}} = \mathbb{P}[X_0(i) = 0\ \forall i \in \Lambda_M].
$$

Since $\mathbb{P}[X_0 \in \cdot\,]$ is nontrivial, we have that

$$
\lim_{M\uparrow\infty} \mathbb{P}[X_0(i) = 0\ \forall i \in \Lambda_M] = \mathbb{P}[X_0 = \underline{0}] = 0.
$$

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. ∎

---

[9]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 $\|\prod_{i=1}^n f_i\|_1 \leq \prod_{i=1}^n \|f_i\|_n$.

Exercise 6.28 *Show by counterexample that the statement of Lemma 6.27 is false for $s = 0$.*

Proof of Theorem 6.25 As in the proof of Lemma 6.26, we set

$$\rho(x) := \mathbb{P}\big[X_t^x \neq \underline{0} \ \forall t \geq 0\big] \qquad (x \in \{0,1\}^{\mathbb{Z}^d}, \ |x| < \infty).$$

By Lemmas 4.32, 6.5, and 6.24, it suffices to show that

$$\lim_{t \to \infty} \mathbb{P}\big[x \wedge X_t \neq \underline{0}\big] = \rho(x)$$

for all finite $x \in \{0,1\}^{\mathbb{Z}^d}$. By duality, this is equivalent to showing that

$$\lim_{t \to \infty} \mathbb{P}\big[X_{t-s}^x \wedge X_s \neq \underline{0}\big] = \rho(x) \qquad \big(x \in \{0,1\}^{\mathbb{Z}^d}, \ |x| < \infty\big),$$

where $(X_t^x)_{t \geq 0}$ 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 6.27, and write

$$\begin{aligned}
\mathbb{P}\big[X_t^x \wedge X_s \neq \underline{0}\big] &= \mathbb{P}\big[X_t^x \wedge X_s \neq \underline{0} \,\big|\, |X_t^x| = 0\big]\, \mathbb{P}\big[|X_t^x| = 0\big] \\
&\quad + \mathbb{P}\big[X_t^x \wedge X_s \neq \underline{0} \,\big|\, 0 < |X_t^x| < N\big]\, \mathbb{P}\big[0 < |X_t^x| < N\big] \\
&\quad + \mathbb{P}\big[X_t^x \wedge X_s \neq \underline{0} \,\big|\, |X_t^x| \geq N\big]\, \mathbb{P}\big[|X_t^x| \geq N\big].
\end{aligned}$$

Here, by Lemma 6.26 and our choice of N,

- (i) $\mathbb{P}\big[X_t^x \wedge X_s \neq \underline{0} \,\big|\, |X_t^x| = 0\big] = 0,$
- (ii) $\lim_{t \to \infty} \mathbb{P}\big[0 < |X_t^x| < N\big] = 0,$
- (iii) $\liminf_{t \to \infty} \mathbb{P}\big[X_t^x \wedge X_s \neq \underline{0} \,\big|\, |X_t^x| \geq N\big] \geq 1 - \varepsilon,$
- (iv) $\lim_{t \to \infty} \mathbb{P}\big[|X_t^x| \geq N\big] = \rho(x),$

from which we conclude that

$$(1 - \varepsilon)\rho(x) \leq \liminf_{t \to \infty} \mathbb{P}\big[X_t^x \wedge X_s \neq \underline{0}\big] \leq \limsup_{t \to \infty} \mathbb{P}\big[X_t^x \wedge X_s \neq \underline{0}\big] \leq \rho(x).$$

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

Theorem 6.25 has a simple corollary.

Corollary 6.29 (Homogeneous invariant laws) *All homogeneous invariant laws of a contact process are convex combinations of $\delta_{\underline{0}}$ and $\overline{\nu}$.*

Proof Let ν be any homogeneous invariant law. We will show that ν is a convex combination of $\delta_{\underline{0}}$ and $\overline{\nu}$. If $\nu = \delta_{\underline{0}}$ we are done. Otherwise, as in the proof of Lemma 5.10, we can write $\nu = (1 - p)\delta_{\underline{0}} + p\mu$ where $p \in (0, 1]$ and μ is a nontrivial homogeneous invariant law. But now Theorem 6.25 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

$$\lambda_{\mathrm{c}} := \inf\{\lambda \in \mathbb{R} : \theta(\lambda) > 0\} \qquad (6.60)$$

denote the *critical point* of the contact process. As an application of Theorem 6.25, we prove the following result.

Proposition 6.30 (Continuity above the critical point) *The function* $\lambda \mapsto \theta(\lambda)$ *is left-continuous on* $(\lambda_{\mathrm{c}}, \infty)$.

Proof Let $\overline{\nu}_\lambda$ denote the upper invariant law of the contact process with infection rate λ. Fix $\lambda > \lambda_{\mathrm{c}}$ 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 4.37, each such subsequential ν_* limit is an invariant law. It clearly is also homogeneous. Since $\lambda > \lambda_{\mathrm{c}}$, 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 6.29, we conclude that $\nu_* = \overline{\nu}$. This argument shows that the map

$$(\lambda_{\mathrm{c}}, \infty) \ni \lambda \mapsto \overline{\nu}_\lambda$$

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.   ∎

Exercise 6.31 *Let* $(X_t)_{t \geq 0}$ *be a additive interacting particle system and let* $(Y_t)_{t \geq 0}$ *be it additive dual. Show that the upper invariant law* $\overline{\nu}$ *of* $(X_t)_{t \geq 0}$ *is uniquely characterised by*

$$\int \overline{\nu}(\mathrm{d}x)\, 1_{\{x \wedge y \neq \underline{0}\}} = \mathbb{P}^y[Y_t \neq \underline{0}\ \forall t \geq 0] \qquad \big(y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)\big). \qquad (6.61)$$

Exercise 6.32 *Let $(X_t)_{t\geq 0}$ be a cancellative interacting particle system and let $(Y_t)_{t\geq 0}$ be it cancellative dual. Let $\pi_{1/2}$ denote product measure with intensity $1/2$. Show that*

$$\mathbb{P}^{\pi_{1/2}}\big[X_t \in \,\cdot\,\big] \underset{t\to\infty}{\Longrightarrow} \nu_{1/2},$$

where $\nu_{1/2}$ is an invariant law that is uniquely characterised by the relation

$$\int \nu_{1/2}(\mathrm{d}x)1_{\{|x \wedge y| \ is \ odd\}} = \tfrac{1}{2}\mathbb{P}^y\big[Y_t \neq \underline{0} \ \forall t \geq 0\big] \qquad \big(y \in \mathcal{S}_{\mathrm{fin}}(\Lambda)\big).$$

Because of the simililarity of this formula to the characterisation of the upper invariant law of an additive interacting particle system in (6.61), the measure $\nu_{1/2}$ is sometimes called the odd upper invariant law.

6.9 Equality of critical points

The contact voter model X on $\mathbb{Z}^d$, that has a mixture of contact process and voter model dynamics, has been introduced in Section 6.6. It has two parameters: the infection rate λ and the voter rate α. We say that X *survives* if

$$\mathbb{P}^{1_{\{0\}}}[X_t \neq \underline{0} \ \forall t \geq 0] > 0.$$

For each $\alpha \geq 0$, we define critical infection rates $\lambda_{\mathrm{c}}(\alpha)$ and $\lambda'_{\mathrm{c}}(\alpha)$ by

$$\lambda_{\mathrm{c}}(\alpha) := \inf\big\{\lambda \in \mathbb{R} : \text{the upper invariant law is nontrivial}\big\},$$

$$\lambda'_{\mathrm{c}}(\alpha) := \inf\big\{\lambda \in \mathbb{R} : \text{the process survives}\big\}.$$

The paper [DLZ14] studies the asymptotics of $\lambda_{\mathrm{c}}(\alpha)$ as $\alpha \to \infty$. Here, we will use duality to prove a more simple statement, namely, that $\lambda_{\mathrm{c}}(\alpha) = \lambda'_{\mathrm{c}}(\alpha)$ for all $\alpha \geq 0$.

For $\alpha = 0$ (i.e., the pure contact process), we already know this, as it is a direct consequence of Lemma 6.24, which follows from the self-duality of the contact process. We will use a similar argument here using Proposition 6.20, which says that the contact voter model is self-dual with respect to the duality function ψ_r from (6.42) with $r := \lambda/(\alpha + \lambda)$. Note that if $\alpha = 0$ (the pure contact process), then $r = 1$ which corresponds to additive systems duality.

Proposition 6.33 (Characterization of the upper invariant law) *Let $r := \lambda/(\alpha+\lambda)$. The upper invariant law $\overline{\nu}$ of the contact voter model satisfies*

$$\int \overline{\nu}(\mathrm{d}x)\psi_r(x, y) = \mathbb{P}^y\big[X_t = \underline{0} \ for \ some \ t \geq 0\big] \tag{6.62}$$

for all finite $y \in \{0, 1\}^{\mathbb{Z}^d}$. In particular, $\lambda_{\mathrm{c}}(\alpha) = \lambda'_{\mathrm{c}}(\alpha)$ for all $\alpha \geq 0$.

Proof Letting X^1 and X^y denote the processes started in $X_0^1 = 1$ and $X_0^y = y$, we observe that by Proposition 6.20,

$$\int \overline{\nu}(\mathrm{d}x)\, \psi_r(x, y) = \lim_{t\to\infty} \mathbb{E}\big[\psi_r(X_t^1, y)\big]$$
$$= \lim_{t\to\infty} \mathbb{E}\big[\psi_r(1, X_t^y)\big] = \lim_{t\to\infty} \mathbb{E}\big[(1-r)^{|X_t^y|}\big].$$

The proof of Lemma 6.26 carries over without a change to the contact voter model, so

$$X_t^y = \underline{0} \text{ for some } t \geq 0 \quad \text{or} \quad |X_t^y| \xrightarrow[t\to\infty]{} \infty \quad \text{a.s.}$$

Using this, we see that

$$\lim_{t\to\infty} \mathbb{E}\big[(1-r)^{|X_t^y|}\big] = \mathbb{P}^y\big[X_t = \underline{0} \text{ for some } t \geq 0\big],$$

completing the proof of (6.62).

Inserting $y = 1_{\{0\}}$ into (6.62), we see that

$$\int \overline{\nu}(\mathrm{d}x)\,(1-r)^{x(0)} = \mathbb{P}^{1_{\{0\}}}\big[X_t = \underline{0} \text{ for some } t \geq 0\big],$$

or equivalently, using the fact that $1 - (1-r)^{x(0)} = r x(0)$ with $r = \lambda/(\alpha+\lambda)$,

$$\frac{\lambda}{\alpha + \lambda} \int \overline{\nu}(\mathrm{d}x)\, x(0) = \mathbb{P}^{1_{\{0\}}}\big[X_t \neq \underline{0} \; \forall t \geq 0\big].$$

This shows that $\overline{\nu} \neq \delta_{\underline{0}}$ 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 (6.60) is nontrivial in the sense that $0 < \lambda_c < \infty$. Note that by Lemma 6.24,

$$\lambda_c = \inf\{\lambda \in \mathbb{R} : \text{ the contact process survives}\}$$
$$= \inf\{\lambda \in \mathbb{R} : \text{ the upper invariant law is nontrivial}\}.$$

In Exercise 5.12, which is based on Theorem 4.29, 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 up-

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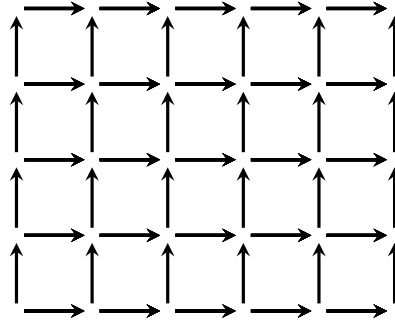
Now fix some *percolation parameter* $p \in [0, 1]$ and let $(\omega_{(i,j)})_{(i,j)\in\mathcal{A}}$ be a collection of i.i.d. Bernoulli random variables with $\mathbb{P}[\omega_{(i,j)} = 1] = p$. We say that there is an *open path* from a site $i \in \mathbb{Z}^d$ to $j \in \mathbb{Z}^d$ if there exist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

$$p_c := \inf \left\{ p \in [0,1] : \mathbb{P}[0 \rightsquigarrow \infty] > 0 \right\}.$$

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

$$\mathbb{P}[N_n \neq 0] \leq \mathbb{E}[N_n] = d^n p^n.$$

Since the events $\{N_n \neq 0\}$ decrease as $n \to \infty$ to the event $\{0 \rightsquigarrow \infty\}$, taking the limit, we see that $\mathbb{P}[0 \rightsquigarrow \infty] = 0$ for all $p < 1/d$, 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 $0 \not\rightarrow \infty$ 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

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

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 $D_m := \{0, \dots, m\}^2$ and let us say that a set i is "wet" if $j \leadsto i$ for some $j \in D_m$. If $D_m \not\leadsto \infty$, then the set of wet sites must be finite, and, just as before, there must be a dual path surrounding this set of wet sites. Then, by the same arguments as before

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

where now the sum starts at $2m$ since the dual path must surround D_m and hence have length $2m$ at least. If $p > \frac{8}{9}$, then the sum is finite so it can be made arbitrarily small by choosing m suffiently large. It follows that $\mathbb{P}[D_m \leadsto \infty] > 0$ for some m, hence $\mathbb{P}[i \leadsto \infty] > 0$ for at least one $i \in D_m$, and therefore, by translation invariance, also $\mathbb{P}[0 \leadsto \infty] > 0$. ∎

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 $0 < \lambda_c < \infty$.*

Proof As already mentioned in Section 7.1, the fact that $0 < \lambda_c$ has already been proved in Exercise 5.12. By Exercise 5.18, to prove that $\lambda_c < \infty$, 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

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

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

$$\omega_{(i,j)} = 1 \quad \text{implies} \quad (\kappa_i, \sigma_i) \leadsto (\kappa_j, \sigma_j) \qquad ((i,j) \in \mathcal{A}).$$
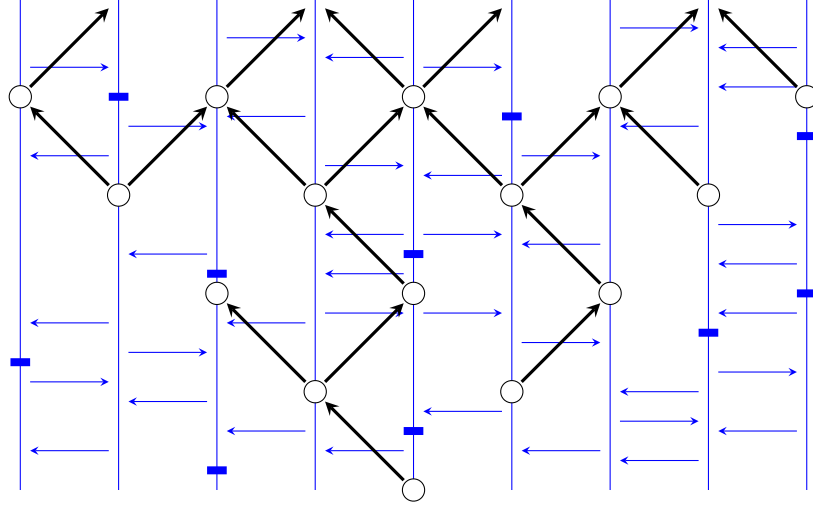
Figure 7.2: Comparison with oriented percolation. Good events in the graphical representation of the contact process (blue) correspond to open percolation arrows (black). An infinite open path along percolation arrows implies an infinite open path in the graphical representation of the contact process.

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

$$\tau_i^{\pm} := \inf\{t \geq \sigma_i : \text{at time } t \text{ there is an infection arrow from } \kappa_i \text{ to } \kappa_i \pm 1\}$$

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"

$$\mathcal{G}_i^{\pm} := \big\{ \tau_i^{\pm} < \sigma_i + T \text{ and there are no blocking symbols on}$$
$$\{\kappa_i\} \times (\sigma_i, \tau_i^{\pm}] \text{ and } \{\kappa_i \pm 1\} \times (\tau_i^{\pm}, \sigma_i + T]\big\}.$$

Clearly,
$$\mathcal{G}_i^{-} \quad \text{implies} \quad \psi(i_1, i_2) \rightsquigarrow \psi(i_1, i_2 + 1),$$
$$\text{and} \quad \mathcal{G}_i^{+} \quad \text{implies} \quad \psi(i_1, i_2) \rightsquigarrow \psi(i_1 + 1, i_2).$$

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.
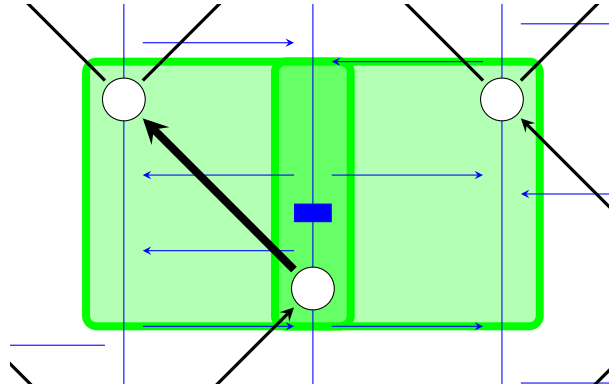
Figure 7.3: Good events use information from partially overlapping regions of space-time.

We observe that

$$p := \mathbb{P}[\omega_{(i,j)} = 1] = \mathbb{P}(\mathcal{G}_i^{\pm}) = (1 - e^{-\lambda T})e^{-T} \qquad ((i,j) \in \mathcal{A}), \qquad (7.2)$$

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 success 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

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

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 $P[\chi_i = 1] \geq p$ $(i \in \Lambda)$, and that*

$$\tilde{p} := \left(1 - (1-p)^{1/K}\right)^2 \geq \tfrac{1}{4}.$$

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

$$P[\tilde{\chi}_i = 1] = \tilde{p} \qquad (i \in \Lambda), \tag{7.3}$$

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,

$$P[\chi'_n = 1 \mid \chi'_0, \ldots, \chi'_{n-1}] \geq \tilde{p} \tag{7.4}$$

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

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

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
$$P[\chi'_n = \varepsilon_n, \ldots, \chi'_0 = \varepsilon_0] = p'_n(\varepsilon_0, \ldots, \varepsilon_{n-1}) \cdots p'_1(\varepsilon_0) \cdot p'_0.$$

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

$$\begin{aligned} P[\chi'_n = 1 \mid \chi'_0 &= \varepsilon_0, \ldots, \chi'_{n-1} = \varepsilon_{n-1}] \\ &= r P[\chi_n = 1 \mid \chi'_0 = \varepsilon_0, \ldots, \chi'_{n-1} = \varepsilon_{n-1}]. \end{aligned} \qquad (7.5)$$

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

$$P[\chi_n = 0 \mid \chi'_0 = \varepsilon_0, \ldots, \chi'_{n-1} = \varepsilon_{n-1}] \leq 1 - r. \qquad (7.6)$$

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 3-dependent with $P[\chi_n = 1] = p$, but $P[\chi_n = 1 \mid \chi_{n-1} = 0, \chi_{n-2} = 1] = 0$.

Then

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

$$(7.7)$$

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

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

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 λ_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

$$\mathbb{E}^{1_{\{0\}}}\big[|X_{nt}|\big] \le r^n \qquad (n \ge 0)$$

and the process started in any finite initial state dies out a.s. Can you use this to improve the lower bound $1/|\mathcal{N}_i| \le \lambda_c$ 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.

[AU29] P. Alexandroff and P. Urysohn. Mémoire sur les espaces topologiques compacts, dédié à Monsieur D. Egoroff. *Verhandelingen Amsterdam* 14(1) (1929) 1–93.

[BD01] W.A. Brock and S.N. Durlauf. Discrete choice with social interactions. *The Review of Economic Studies* 68(2) (2001), 235–260.

[BDD91] M. Bramson, W. Ding, and R. Durrett. Annihilating branching processes. *Stoch. Process. Appl.* 37 (1991), 1–17,

[BDT19] O. Blondel, A. Deshayes, and C. Toninelli. Front evolution of the Fredrickson-Andersen one spin facilitated model. *Electron. J. Probab.* 24 (2019), paper No. 1, 1–32.

[BG90] C. Bezuidenhout and G. Grimmett. The critical contact process dies out. *Ann. Probab.* 18(4) (1990), 1462–1482.

[CD+20] R. Cerf, P. Dai Pra, M. Formentin, and D. Tovazzi. Rhythmic behavior of an Ising Model with dissipation at low temperature. Preprint (2020), arXiv:2002.08244.

[CFT16] F. Collet, M. Formentin, and D. Tovazzi. Rhythmic behavior in a two-population mean-field Ising model. *Phys. Rev. E* 94 (2016), 042139.

[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.

[DFR13] P. Dai Pra, M. Fischer, and D. Regoli. A Curie-Weiss model with dissipation. *J. Stat. Phys.* 152(1) (2013), 37–53.

[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.

[DST17] H. Duminil-Copin, V. Sidoravicius, and V. Tassion. Continuity of the phase transition for planar random-cluster and Potts models with $1 \leq q \leq 4$. *Commun. Math. Phys.* 349(1) (2017), 47–107.

[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.

[GB07]   T. Gobron and I. Merola. First-order phase transition in Potts models with finite-range interactions. *J. Stat. Phys.* 126(3) (2007), 507–583.

[GF24]   C. Giardinà and F. Redig *Duality for Markov processes: a Lie algebraic approach.* In preparation, 2024.

[GK+09] C. Giardinà, J. Kurchan, F. Redig, and K. Vafayi. Duality and hidden symmetries in interacting particle systems. *J. Stat. Phys.* 135(1) (2009), 25–55.

[GPA01] M. Ghaemi, G.A. Parsafar, and M. Ashrafizaadeh. Calculation of the critical temperature for 2- and 3-dimensional Ising models and for 2-dimensional Potts models using the transfer matrix method. *J. Phys. Chem. B* 105(42) (2001), 10355–10359.

[Gra86]   L. Gray Duality for general attractive spin systems with applications in one dimension. *Ann. Probab.* 14(2) (1986), 371–396.

[Har72]   T.E. Harris. Nearest-neighbor Markov interaction processes on multidimensional lattices. *Adv. Math.* 9 (1972), 66–89.

[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.

[JK14a]   B. Jahnel and C. Külske. A class of nonergodic interacting particle systems with unique invariant measure. *Ann. Appl. Probab.* 24(6) (2014), 2595–2643.

[JK14b]   S. Jansen and N. Kurt. On the notion(s) of duality for Markov processes. *Prob. Surveys* 11 (2014), 59–120.

[Kal97]   O. Kallenberg. *Foundations of modern probability.* Springer, New York, 1997.

[Kol56]   A.N. Kolmogorov. On Skorohod convergence. *Theor. Probability Appl.* 1 (1956), 213–222.

[KKO77]   T. Kamae, U. Krengel, and G.L. O'Brien. Stochastic inequalities on partially ordered spaces. *Ann. Probab.* 5 (1977), 899–912.

[Kro99]   S. Krone. The two-stage contact process. *Ann. Appl. Probab.* 9(2) (1999), 331–351.

[Lan24]   N. Lanchier. *Stochastic Interacting Systems in Life and Social Sciences.* Walter de Gruyter, Berlin, 2024.

[Lat24]   J.N. Latz. *Pathwise Duality of Interacting Particle Systems.* PhD thesis, Faculty of Mathematics and Physics, Charles University, Prague, 2024. `https://library.utia.cas.cz/separaty/2024/SI/latz-0598658.pdf`

[Law05]   G.F. Lawler. *Conformally Invariant Processes in the Plane.* Mathematical Surveys and Monographs 114. American Mathematical Society, Providence, 2005.

[Lig72]   T.M. Liggett. Existence theorems for infinite particle systems. *Trans. Am. Math. Soc.* 165 (1972), 471–481.

[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.

[LS21]   J.N. Latz and J.M. Swart. Commutative monoid duality. *J. Theor. Probab.* 36 (2023), 1088–1115.

[LS23a]   J.N. Latz and J.M. Swart. Applying monoid duality to a double contact process. *Electron. J. Probab.* 28 (2023), paper no. 70, 1–26.

[LS23b]   J.N. Latz and J.M. Swart. Monotone duality of interacting particle systems. Preprint (2023), arXiv:2312.00595.

[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.

[Mou95]   T.S. Mountford. A coupling of infinite particle systems. *J. Math. Kyoto Univ.* 35(1) (1995), 43–52.

[MSS20]   T. Mach, A. Sturm, and J.M. Swart. Recursive tree processes and the mean-field limit of stochastic flows. *Electron. J. Probab.* 25 (2020) paper No. 61, 1–63.

[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.

[RS03]   F. Ritort and P. Sollich. Glassy dynamics of kinetically constrained models. *Adv. Phys.* 52(4) (2003), 219–342.

[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. *J. Theor. Probab.* 31(2) (2018), 932–983.

[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.

[Tov19]   D. Tovazzi. *Self-sustained periodic behaviors in interacting systems: macroscopic limits and fluctuations.* PhD Thesis University of Padua, 2019.

Index