Interacting particles and stochastic PDEs

by Nicolas Perkowski

Introduction

Physical systems are typically described by partial differential equations (PDEs). For example, the flow of heat through a medium is governed by the heat equation, and the dynamics of viscous fluids are described by the Navier-Stokes equation. These equations are typically derived with physical arguments that operate on a macroscopic (observable) scale and do not involve the behavior of single molecules. On the other hand, if we would zoom into the physical systems, we would see only single molecules, moving seemingly chaotically. The question we try to approach in this course is whether we can derive the equations governing the observable behavior of certain physical systems by considering simple random dynamics for the single molecules and then zooming out to the observable scale.

The first question that then arises is how we should model the dynamics of the single molecules. We will do this using *interacting particle systems*. Interacting particle systems are certain Markov processes that model the dynamic behavior of indistinguishable members of a population (these could be molecules, individuals in an animal population, agents in a social network, ...).

Example 1. (Exclusion process)

On the lattice $\mathbb{Z}^d$ we introduce particles, which perform independent random walks. However, there is at most one particle per site allowed, so if another particle tries to jump to an already occupied site, then the jump is suppressed. We could imagine that the particles are molecules in a gas, and two molecules cannot occupy the same space.

Example 2. (Contact process)

Consider a graph $G=(V,E)$. The links between the vertices are contacts between individuals in a population, and we try to track the spread of a disease through the population. At every vertex, one of two states can occur: 1 (meaning the individual is sick) or 0 (meaning the individual is healthy). We then include two mechanisms for changing the state: every sick individual that is in contact with a healthy individual may infect the healthy one with a certain probability, and every sick individual can cure itself with a certain probability.

Example 3. (Voter model)

Again we consider a graph $G=(V, E)$, where the links between the vertices are contacts between individuals in a population. This time we try to track the behavior of potential voters, and in particular the formation of opinion clusters. Assume that there are two possible votes, 0 and 1, and at every time every individual chooses one of these two options. However, the individuals influence each other and each individual convinces each of its neighbors with a certain probability to convert to its own opinion, while it itself is also convinced with a certain probability to switch to the opinion of the neighbor.

Example 4. (Zero range process)

This time the particles evolve on the graph $\mathbb{Z}^d$, and again each particle is performing an independent random walk. However, now there is no exclusion rule and instead the particles attract each other: If a particle sits in a site with k other particles, then it will leave that site with a probability that decreases with k. It is not clear to me whether this is a realistic model from a physical viewpoint, but it allows for nice mathematics which is why we will study it.

Ultimately, we are interested in the limit of infinitely many particles (if the particles are molecules in a gas, then there would be of the order 10^{23} of them, so infinitely many seems like a pretty good approximation). We will also try to understand the error that we make if we apply the infinitely many particle approximation to a system with finitely many particles, and see that the difference can be described by a *stochastic* partial differential equation.

But of course, the first problem we face is to write down mathematical models for the heuristic dynamics described above. And due to the interaction, this is a bit subtle. Consider for example the exclusion process on $\mathbb{Z}$. If the particle at site $n-1$ wants to move to n and there is already a particle at n, then the jump should be suppressed. But what if the particle at n wants to move at $n+1$, too? Then the behavior of the particle at $n-1$ matters on which one of these two moves is carried out first, which was not part of our description of the dynamics. Similar problems occur in the other examples. Therefore, it is much more convenient to make the system evolve in continuous time, where such an interaction of two simultaneous jumps can never occur. So before we look into interacting particle systems, we first need to understand some basics of continuous time Markov processes.

Notation and conventions

Positive and *negative* are used in the European sense, i.e. meaning $\geqslant 0$ and $\leqslant 0$. If we want to stress that a quantity is >0 we call it *strictly positive*. $\mathbb{N}=\{1,2,\ldots\}$ and $\mathbb{N}_0=\{0,1,\ldots\}$ and $\mathbb{R}_+=[0,\infty)$. $a\lesssim b$ and $b\gtrsim a$ mean that $a\leqslant Cb$ for a constant $C>0$ that does not depend on the variables and parameters under consideration. $a\simeq b$ means $a\lesssim b$ and $a\gtrsim b$. For inner products on $\mathbb{R}^d$ we write

$$\langle x, y\rangle := x\cdot y := \sum_{j=1}^{d} x_j y_j.$$

1 Continuous time Markov processes on Polish state spaces

The first two chapters follow [Eberle, *Markov Processes*, Universität Bonn lecture notes, 2017], and [Liggett, *Continuous Time Markov Processes*, AMS, 2010].

Throughout this section we fix a probability space $(\Omega,\mathcal{F},\mathbb{P})$ and S will denote a Polish state space (i.e. S is a separable metric space), equipped with its Borel sigma algebra $\mathcal{B}$. A stochastic process $(X_t)_{t\geqslant 0}$ with values in S is a family of random variables $X_t\colon\Omega\to S$. We will typically write

$$\mathcal{F}_t^X=\sigma(X_s\colon s\in[0,t]).$$

Definition 1.1. *Let $(\mathcal{F}_t)_{t\geqslant 0}$ be a filtration and $X=(X_t)_{t\geqslant 0}$ be an adapted stochastic process. Then X is called a Markov process with respect to $(\mathcal{F}_t)_{t\geqslant 0}$ if for all $0\leqslant s\leqslant t$ and all $B\in\mathcal{B}$ we have almost surely (a.s.)*

$$\mathbb{P}(X_t\in B|\mathcal{F}_s)=\mathbb{P}(X_t\in B|X_s).$$

This is equivalent to having for all bounded and measurable functions f

$$\mathbb{E}[f(X_t)|\mathcal{F}_s]=\mathbb{E}[f(X_t)|X_s].$$

If X is a Markov process with respect to $(\mathcal{F}_t)_{t\geqslant 0}$, then by the tower property X is also a Markov process with respect to $(\mathcal{F}_t^X)_{t\geqslant 0}$. So we will simply call Markov process with respect to $(\mathcal{F}_t^X)_{t\geqslant 0}$ a Markov process.

Remark 1.2. Since S is a polish space, for any pair of random variables X,Y with values in S the regular conditional probability distribution of Y given X exists. This is a map $\mu\colon S\times\mathcal{B}\to[0,1]$, such that

i. for each $B\in\mathcal{B}$ the map $x\mapsto\mu(x,B)$ is measurable and satisfies for almost all $\omega\in\Omega$:

$$\mu(X(\omega),B)=\mathbb{E}[Y\in B|X](\omega);$$

ii. for all $x\in S$ the map $B\mapsto\mu(x,B)$ defines a probability measure.

See [Durrett, *Probability. Theory and Examples*, CUP, 2010]. If μ satisfies the properties above, then we call it a probability kernel (on $S\times\mathcal{B}$).

Let now $X = (X_t)_{t \geqslant 0}$ be a Markov process with values in S. Then there exists a family of probability kernels $(p_{s,t})_{0 \leqslant s \leqslant t}$ such that for all $0 \leqslant s \leqslant t$ and all $B \in \mathcal{B}$ we have almost surely

$$p_{s,t}(X_s, B) = \mathbb{P}(X_t \in B | X_s).$$

Let us write

$$\mu f(x) := \int_S f(y) \mu(x, \mathrm{d}y)$$

whenever μ is a probability kernel on $S \times \mathcal{B}$ and $f: S \to \mathbb{R}$ is a bounded measurable function. Note that then μf is again a bounded measurable function.

Lemma 1.3. *Let $(p_{s,t})_{0 \leqslant s \leqslant t}$ be the probability kernels associated to a Markov process $X = (X_t)_{t \geqslant 0}$ with values in S. Then the following consisteny condition holds: for any $0 \leqslant s \leqslant u \leqslant t$ and any bounded and measurable $f: S \to \mathbb{R}$ we have*

$$p_{s,t} f = p_{s,u} p_{u,t} f, \qquad \mathbb{P} \circ X_s^{-1} \text{ a.s..}$$

Moreover, of course we have $p_{t,t} f(x) = f(x)$ for $\mathbb{P} \circ X_s^{-1}$-almost all x.

Proof. Using the Markov property and the tower property, we get

$$\begin{aligned} p_{s,t} f(X_s) &= \mathbb{E}[f(X_t) | X_s] = \mathbb{E}[\mathbb{E}[f(X_t) | \mathcal{F}_u] | \mathcal{F}_s] = \mathbb{E}[\mathbb{E}[f(X_t) | X_u] | \mathcal{F}_s] \\ &= \mathbb{E}[p_{u,t} f(X_u) | X_s] = p_{s,u} p_{u,t} f(X_s), \end{aligned}$$

which proves the first claim. The second claim is easy. □

Definition 1.4. *A family of probability kernels $(p_{s,t})_{0 \leqslant s \leqslant t}$ is called a transition function if*

i. $p_{t,t}(x, \cdot) = \delta_x$ for all $t \geqslant 0$ and $x \in S$, where $\delta_x(B) = 1$ if $x \in B$ and $\delta_x(B) = 0$ otherwise;

ii. the Chapman-Kolmogorov equation holds: $p_{s,t} = p_{s,u} p_{u,t}$ for all $0 \leqslant s \leqslant u \leqslant t$, where the composition of the two probability kernels $p_{s,u}$ and $p_{u,t}$ is defined as

$$p_{s,u} p_{u,t}(x, B) = \int_S p_{s,u}(x, \mathrm{d}y) p_{u,t}(y, B), \qquad x \in S, B \in \mathcal{B}.$$

The transition function is called time-homogeneous if $p_{s,t} = p_{s+r,t+r}$ for all $r \geqslant 0$, and in that case we also write p_{t-s} instead of $p_{s,t}$.

Note that for a time-homogeneous transition function the Chapman-Kolmogorov equation becomes simpler: $p_{s+t} = p_s p_t$ for all $s, t \geqslant 0$.

We saw above that any Markov process morally gives rise to a transition function. However, there is a subtle difference between the conditions derived in Lemma 1.3 and the conditions required in Definition 1.4: Now we request that i. and ii. hold for all $x \in S$ and not just $\mathbb{P} \circ X_s^{-1}$ almost surely. This will allow us to link a transition function to a Markov process with arbitrary starting point.

Example 1.5. Assume that S is countable and let $(P(x, y))_{x,y \in S}$ be the transition matrix of a time-homogeneous discrete time Markov chain $(X_n)_{n \in \mathbb{N}_0}$. Then for $m, n \in \mathbb{N}_0$ with $m \leqslant n$

$$q_{m,n}(x, B) = \sum_{y \in B} P^{n-m}(x, y)$$

defines a "discrete time transition function". Indeed, $q_{m,m}(x, B) = \sum_{y \in B} P^0(x, y) = \delta_x(B)$ and

$$\begin{aligned} q_{m,k} q_{k,n}(x, B) &= \sum_{y \in S} P^{k-m}(x, y) \sum_{z \in B} P^{n-k}(y, z) = \sum_{z \in B} (P^{k-m} P^{n-k})(x, z) \\ &= P^{n-m}(x, B) = q_{m,n}(x, B). \end{aligned}$$

Setting $p_{s,t} = q_{\lfloor s \rfloor, \lfloor t \rfloor}$, we get a (continuous time) transition function. Note however that this transition function is in general not time-homogeneous, despite the fact that q is time-homogeneous on $\mathbb{N}_0$. Indeed, we have $p_{0,1/2} = q_{0,0} = \mathrm{id}$ but $p_{1/2,1} = q_{0,1}$.

Definition 1.6. *Let $(p_{s,t})_{0\leqslant s\leqslant t}$ be a transition function and let $(\mathcal{F}_t)_{t\geqslant 0}$ be a filtration. An adapted process $X=(X_t)_{t\geqslant 0}$ with values in S is called a Markov process with respect to $(\mathcal{F}_t)$ with transition function $(p_{s,t})$ if for all $0\leqslant s\leqslant t$ and all $B\in\mathcal{B}$ we have a.s.*

$$\mathbb{P}(X_t\in B|\mathcal{F}_s)=p_{s,t}(X_s,B).$$

If the transition function is time-homogeneous, then X is called time-homogeneous.

Remark 1.7. Let $(X_t)_{t\geqslant 0}$ be a Markov process with transition function $(p_{s,t})$. Then

$$\big(\hat{X}_t=(t,X_t)\big)_{t\geqslant 0}$$

is also a Markov process, with time-homogeneous transition function

$$\hat{p}_t=((s,x),\cdot)=\delta_{s+t}\otimes p_{s,s+t}(x,\cdot).$$

Therefore, we will from now on only focus on time-homogeneous Markov processes, and often we will omit the term "time-homogeneous" in the description.

Example 1.8. Let $W=(W_t)_{t\geqslant 0}$ be a Brownian motion. Then W is a Markov process with transition function

$$p_t f(x)=\int_{\mathbb{R}} f(y)\frac{1}{\sqrt{2\pi t}}\exp\bigg(-\frac{(x-y)^2}{2t}\bigg)\mathrm{d}y.$$

Indeed, since W has independent increments we have for any bounded and measurable function f

$$\mathbb{E}[f(W_t)|\mathcal{F}_s^X]=\mathbb{E}[f(W_t-W_s+W_s)|\mathcal{F}_s^X]=\Phi(W_s)$$

with

$$\begin{aligned}\Phi(x)=\mathbb{E}[f(W_t-W_s+x)]&=\int_{\mathbb{R}} f(y+x)\frac{1}{\sqrt{2\pi(t-s)}}\exp\bigg(-\frac{y^2}{2(t-s)}\bigg)\mathrm{d}y\\&=\int_{\mathbb{R}} f(y)\frac{1}{\sqrt{2\pi(t-s)}}\exp\bigg(-\frac{(x-y)^2}{2(t-s)}\bigg)\mathrm{d}y.\end{aligned}$$

Example 1.9. Assume that S is countable, let $Y=(Y_n)_{n\in\mathbb{N}_0}$ be a time-homogeneous discrete time Markov chain with transition matrix $(P(x,y))_{x,y\in S}$, and let $(N_t)_{t\geqslant 0}$ be an independent Poisson process with intensity $\lambda>0$. Then $(X_t=Y_{N_t})_{t\geqslant 0}$ is a Markov process with transition function

$$p_t f(x)=\sum_{k=0}^{\infty} e^{-\lambda t}\frac{(\lambda t)^k}{k!}P^k f(x).$$

Indeed, since the increment N_t-N_s is independent of $\mathcal{F}_s^X$, we have

$$\begin{aligned}\mathbb{E}[f(X_t)|\mathcal{F}_s^X]&=\sum_{k=0}^{\infty}\mathbb{E}[f(Y_{N_s+k})\mathbb{I}_{N_t-N_s=k}|\mathcal{F}_s^X]=\sum_{k=0}^{\infty}\mathbb{E}[f(Y_{N_s+k})|\mathcal{F}_s^X]\mathbb{P}(N_t-N_s=k)\\&=\sum_{k=0}^{\infty}\left(\sum_{x\in S}P^k(Y_{N_s},x)f(x)\right)e^{-\lambda(t-s)}\frac{(\lambda(t-s))^k}{k!}=\sum_{k=0}^{\infty}e^{-\lambda(t-s)}\frac{(\lambda(t-s))^k}{k!}P^k f(X_s).\end{aligned}$$

Note that here it was very important that (N_t) has independent increments, and the argument does not work if we replace the Poisson process by another Markov process with values in $\mathbb{N}_0$.

Our first aim is to construct Markov processes from transition functions. For that purpose we introduce the canonical space of trajectories $S^{\mathbb{R}_+}$, equipped with the product sigma algebra $\mathcal{B}^{\otimes\mathbb{R}_+}$, i.e. the smallest sigma algebra on $S^{\mathbb{R}_+}$ such that all the projections

$$X_t\colon S^{\mathbb{R}_+}\to S,\qquad X_t(\omega):=\omega(t),$$

are $\mathcal{B}^{\otimes\mathbb{R}_+}$-measurable.

Theorem 1.10. *(Construction of Markov processes)*

Let S be a Polish space, let $(p_t)_{t\geqslant 0}$ be a time-homogeneuous transition function, and let μ be a probability distribution on S. Then there exists a unique probability measure $\mathbb{P}_\mu$ on $(S^{\mathbb{R}_+}, \mathcal{B}^{\otimes \mathbb{R}_+})$, such that X is a Markov process with transition function (p_t) and initial distribution $\mathbb{P}_\mu \circ X_0^{-1} = \mu$.

Proof. This follows from Kolmogorov's extension theorem: If $\mathbb{P}_\mu$ was already given, then we would get for $0 \leqslant t_1 \leqslant \ldots \leqslant t_n$ and $B_1, \ldots, B_n \in \mathcal{B}$ from the Markov property:

$$\begin{aligned}
&\mathbb{P}_\mu(X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n) = \mathbb{E}_\mu[\mathbb{I}_{X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n}] \\
&= \mathbb{E}_\mu[\mathbb{E}_\mu[\mathbb{I}_{X_{t_n} \in B_n} | X_{t_1}, \ldots, X_{t_{n-1}}] \mathbb{I}_{X_{t_1} \in B_1, \ldots, X_{t_{n-1}} \in B_{n-1}}] \\
&= \mathbb{E}_\mu\left[\int_{B_n} p_{t_n - t_{n-1}}(X_{t_{n-1}}, \mathrm{d}x_n) \mathbb{I}_{X_{t_1} \in B_1, \ldots, X_{t_{n-1}} \in B_{n-1}}\right] \\
&= \ldots = \int_{S \times B_1 \times \ldots \times B_n} \mu(\mathrm{d}x_0) p_{t_1}(x_0, \mathrm{d}x_1) p_{t_2 - t_1}(x_1, \mathrm{d}x_2) \ldots p_{t_n - t_{n-1}}(x_{n-1}, \mathrm{d}x_n).
\end{aligned}$$

So we can define the finite-dimensional distributions

$$\Pi_{t_1, \ldots, t_n}(B_1, \ldots, B_n) := \int_{S \times B_1 \times \ldots \times B_n} \mu(\mathrm{d}x_0) p_{t_1}(x_0, \mathrm{d}x_1) p_{t_2 - t_1}(x_1, \mathrm{d}x_2) \ldots p_{t_n - t_{n-1}}(x_{n-1}, \mathrm{d}x_n),$$

and to apply Kolmogorov's extension theorem we need to check e.g. that

$$\Pi_{t_1, t_2, \ldots, t_{n+1}}(S, B_2, \ldots, B_{n+1}) = \Pi_{t_2, \ldots, t_{n+1}}(B_2, \ldots, B_{n+1})$$

(we would have to replace every B_k by S, but the argument is the same only the notation gets more complicated). But

$$\begin{aligned}
&\Pi_{t_1, t_2, \ldots, t_{n+1}}(S, B_2, \ldots, B_{n+1}) \\
&= \int_{S \times S \times B_2 \ldots \times B_{n+1}} \mu(\mathrm{d}x_0) p_{t_1}(x_0, \mathrm{d}x_1) p_{t_2 - t_1}(x_1, \mathrm{d}x_2) \ldots p_{t_{n+1} - t_n}(x_n, \mathrm{d}x_{n+1}) \\
&= \int_{S \times B_2 \ldots \times B_{n+1}} \mu(\mathrm{d}x_0) p_{t_2}(x_1, \mathrm{d}x_2) \ldots p_{t_{n+1} - t_n}(x_n, \mathrm{d}x_{n+1}) \\
&= \Pi_{t_2, \ldots, t_{n+1}}(B_2, \ldots, B_{n+1}),
\end{aligned}$$

where in the second step we applied the Chapman-Kolmogorov equation. Therefore, Kolmogorov's theorem applies and our claim follows. □

Corollary 1.11. *It follows from the proof that for all $0 \leqslant t_1 \leqslant \ldots \leqslant t_n$ and all bounded and measurable functions $f: S^n \to \mathbb{R}$ we have*

$$\begin{aligned}
&\mathbb{E}_\mu[f(X_{t_1}, \ldots, X_{t_n})] \\
&= \int_{S^{n+1}} f(x_1, \ldots, x_n) \mu(\mathrm{d}x_0) p_{t_1}(x_0, \mathrm{d}x_1) p_{t_2 - t_1}(x_1, \mathrm{d}x_2) \ldots p_{t_n - t_{n-1}}(x_{n-1}, \mathrm{d}x_n).
\end{aligned} \tag{1.1}$$

In general we might want to realize a Markov process on another probability space than the canonical space. For that purpose the following result is useful:

Proposition 1.12. *Let $(p_t)_{t\geqslant 0}$ be a transition function and let $(X_t)_{t\geqslant 0}$ be a stochastic process with values in S. Then the following statements are equivalent:*

i. *(X_t) is a Markov process with transition function (p_t) and initial distribution $\mathbb{P} \circ X_0^{-1} = \mu$.*

ii. *For all $0 \leqslant t_1 \leqslant \ldots \leqslant t_n$ and all bounded and measurable functions $f_1, \ldots, f_n$ equality (1.11) holds.*

iii. $\mathbb{P} \circ X^{-1} = \mathbb{P}_\mu$.

iv. $\mathbb{P} \circ X_0^{-1} = \mu$ *and for all bounded and measurable $F: S^{\mathbb{R}_+} \to \mathbb{R}$ and all $0 \leqslant s \leqslant t$ we have a.s.*

$$\mathbb{E}[F((X_t)_{t\geqslant s}) | \mathcal{F}_s^X] = \mathbb{E}_{\delta_{X_s}}[F],$$

where $\mathbb{E}_{\delta_{X_s}}$ denotes the expectation with respect to the measure $\mathbb{P}_{\delta_{X_s}}$ on $(S^{\mathbb{R}_+}, \mathcal{B}^{\otimes \mathbb{R}_+})$ which is constructed in Theorem 1.10.

Proof. i.⇒ii. and ii.⇒iii. we already showed in the proof of Theorem 1.10.

iii.⇒iv.: Assume first that $F((x(t))_{t\geqslant 0}) = f_1(x(t_1)) \times \ldots \times f_n(x(t_n))$. Then the Markov property gives

$$\begin{aligned}\mathbb{E}[F((X_t)_{t\geqslant s})|\mathcal{F}_s^X] &= \mathbb{E}[f_1(X_{s+t_1}) \times \ldots \times f_n(X_{s+t_n})|\mathcal{F}_s^X]\\ &= \int_{S^n} f_1(x_1) \times \ldots \times f_n(x_n) p_{t_1}(X_s, \mathrm{d}x_1) p_{t_2-t_1}(x_1, \mathrm{d}x_2) \ldots p_{t_n-t_{n-1}}(x_{n-1}, \mathrm{d}x_n)\\ &= \int_{S^{n+1}} f_1(x_1) \times \ldots \times f_n(x_n) \delta_{X_s}(\mathrm{d}x_0) p_{t_1}(x_0, \mathrm{d}x_1) p_{t_2-t_1}(x_1, \mathrm{d}x_2) \ldots p_{t_n-t_{n-1}}(x_{n-1}, \mathrm{d}x_n)\\ &= \mathbb{E}_{\delta_{X_s}}[F].\end{aligned}$$

By a continuity argument (e.g. monotone class theorem) this extends to general F.

iv.⇒i.: Let $0 \leqslant s \leqslant t$ and $B \in \mathcal{B}$ and define $F((x(r))_{r\geqslant 0}) = \mathbb{I}_{x(t-s)\in B}$. Then

$$\mathbb{P}(X_t \in B|\mathcal{F}_s^X) = \mathbb{E}[F((X_r)_{r\geqslant s})|\mathcal{F}_s^X] = \mathbb{E}_{\delta_{X_s}}[F],$$

and since the right hand side is measurable with respect to X_s (Exercise!) and $\sigma(X_s) \subset \mathcal{F}_s^X$, the right hand side must be the conditional expectation given X_s. □

2 Continuous time Markov chains on countable state spaces

2.1 The Q-matrix

From now on we assume that S is a countable state space, and we try to get a more intuitive understanding of Markov processes with values in S. Note that for countable S, every transition function can be interpreted as a matrix by setting

$$p_t(x, y) := p_t(x, \{y\}).$$

Since in general it is tedious or even impossible to write down explicit transition functions, we are looking for a description of the infinitesimal rate of change of a Markov process respectively its transition function.

Definition 2.1. *If S is a countable state space, then we make the convention that every transition function on S must satisfy in addition to the properties in Definition 1.4 also*

$$\lim_{t\to 0} p_t(x, x) = 1$$

for all $x \in S$.

Example 2.2. Let $p \in [0, 1]$ and let $(Y_n)_{n\in\mathbb{N}_0}$ be a random walk with

$$P(x, x+1) = \mathbb{P}(Y_{n+1} = x+1|Y_n = x) = p = 1 - \mathbb{P}(Y_{n+1} = x-1|Y_n = x) = P(x, x-1)$$

and $P(x, y) = 0$ otherwise. Let $(N_t)_{t\geqslant 0}$ be a Poisson process with intensity $\lambda > 0$, and let (X_t) be the continuous time Markov process of Example 1.9. Then

$$p_h(x, y) = \sum_{k=0}^{\infty} e^{-\lambda h} \frac{(\lambda h)^k}{k!} P^k(x, y) = e^{-\lambda h} \mathbb{I}_{y=x} + e^{-\lambda h} \lambda h P(x, y) + O(h^2).$$

So if we subtract $p_0(x, y) = \mathbb{I}_{y=x}$, divide by h and send $h \to 0$, we get

$$\begin{aligned}\partial_t p_t(x, y)|_{t=0} &= -\lambda \mathbb{I}_{y=x} + \lambda P(x, y) = \lambda(P - \mathrm{id})(x, y)\\ &= \lambda(p\mathbb{I}_{y=x+1} + (1-p)\mathbb{I}_{y=x-1} - \mathbb{I}_{y=x}) =: q(x, y).\end{aligned}$$

Moreover, by the Chapman-Kolmogorov equation we also get

$$\partial_t p_t|_{t=t_0} = \lim_{h\to 0} \frac{p_{t_0+h} - p_{t_0}}{h} = \lim_{h\to 0} \frac{p_{t_0}(p_h - \mathrm{id})}{h} = p_{t_0} q,$$

as well as

$$\partial_t p_t|_{t=t_0} = \lim_{h\to 0} \frac{p_{t_0+h}(x,y) - p_{t_0}(x,y)}{h} = \lim_{h\to 0} \frac{(p_h - \mathrm{id})p_{t_0}}{h} = q p_{t_0}.$$

The matrix q satisfies:

i. $q(x,y) \geqslant 0$ for all $y \neq x$ and $q(x,x) \leqslant 0$;

ii. $\sum_{y\in S} q(x,y) = 0$, i.e. $q(x,x) = -\sum_{y\neq x} q(x,y)$.

Note that q encodes how the state of the process is changing within an infinitesimal time interval: for $x \neq y$ we have

$$\mathbb{P}(X_{t+h} = y | X_t = x) = h\, q(x,y) + o(h),$$

and therefore $q(x,y)$ is the rate with which X transititions from x to y. Moreover,

$$\mathbb{P}(X_{t+h} \neq x | X_t = x) = \sum_{y\neq x} h\, q(x,y) + o(h) = -h\, q(x,x) + o(h),$$

so that $c(x) := -q(x,x)$ measures the rate with which X is leaving the state x.

In the previous example the matrix q encodes all the information about the process X (with which rate is it leaving a site, once it leaves, where does it go), so it seems to be an important object. This motivates the following definition:

Definition 2.3. *Let S be a countable state space. A Q-matrix on S is a matrix $(q(x,\, y))_{x,y\in S}$ such that*

i. $q(x,y) \geqslant 0$ for all $y \neq x$ and $q(x,x) \leqslant 0$ for all x;

ii. $\sum_{y\in S} q(x,y) = 0$, i.e. $q(x,x) = -\sum_{y\neq x} q(x,y)$.

In that case we often write $c(x) := -q(x,x)$.

We would like to associate a transition function to a given Q-matrix. In the example above we have $\partial_t p_t = q p_t$ with $p_0 = \mathrm{id}$, and therefore we would formally expect that $p_t = e^{tq}$ is a transition function. This works if the state space is finite, but in the countable case we have to be more careful because the matrix exponential e^{tq} might not converge.

Lemma 2.4. *(Exercise) Let S be a finite state space and let q be a Q-matrix on S. Then the matrix exponentials $p_t = e^{tq}$ define a transition function with*

$$\partial_t p_t|_{t=t_0} = p_{t_0} q = q p_{t_0}, \qquad p_0 = \mathrm{id}.$$

Moreover, (p_t) is the only solution to this equation.

Conversely, if $(p_t)_{t\geqslant 0}$ is a transition function, then $q = \partial_t p_t|_{t=0}$ is a Q-matrix and $\partial_t p_t|_{t=t_0} = p_{t_0} q = q p_{t_0}$ for all t_0.

Proof. The uniqueness is clear and also that p solves the equation, because we are considering a linear ordinary differential equation (ODE) on a finite-dimensional space. Let us show that p is a transition function. Since S is countable, $x \mapsto p_t(x,B)$ is of course measurable, and since we can write

$$e^{tq} = e^{t(A-\lambda \mathrm{id})} = e^{tA} e^{-t\lambda \mathrm{id}} = e^{tA} e^{-t\lambda}$$

for some matrix A with positive entries and $\lambda > 0$, we see that all entries of e^{tq} are positive. Furthermore,

$$\partial_t p_t(x,S) = \sum_{y\in S} p_t q(x,y) = \sum_{z\in S} p_t(x,z) \sum_{y\in S} q(z,y) = 0$$

by property ii. in the definition of a Q-matrix, and therefore $p_t(x,S) = p_0(x,S) = \mathrm{id}(x,S) = 1$ for all $t \geqslant 0$. Finally, the Chapman-Kolmogorov equation trivially holds because $e^{(t+s)q} = e^{tq} e^{sq}$.

Conversely, if (p_t) is a transition function, then

$$\lim_{h\to 0}\frac{p_{t_0+h}-p_{t_0}}{h}=\lim_{h\to 0}\frac{p_{t_0}(p_h-\mathrm{id})}{h}=p_{t_0}q$$

and analogously $\partial_t p_t|_{t=t_0}=qp_{t_0}$. Note that for $t_0>0$ also

$$\lim_{h\to 0}\frac{p_{t_0}-p_{t_0-h}}{h}=\lim_{h\to 0}\frac{p_{t_0-h}(p_h-\mathrm{id})}{h}=p_{t_0}q,$$

so we are not restricted to taking the right derivative. Let us show that q satisfies the properties of a Q-matrix:

$$0=\partial_t\sum_y p_t(x,y)|_{t=0}=\sum_y q(x,y)$$

and

$$q(x,y)=\lim_{h\to 0}\frac{p_h(x,y)-\delta_{x,y}}{h}\begin{cases}\geqslant 0, & x\neq y,\\ \leqslant 0, & x=y.\end{cases}$$

□

In the countable case we can use the same approach, but the arguments become significantly more involved. First, let us go from transition function to Q-matrix:

2.2 From transition function to infinitesimal description

Assume now that we are given a transition function (p_t) on the countable state space S. Our aim is to find a Q-matrix q such that $\partial_t p_t|_{t=0}=q$. However, this needs additional assumptions.

Lemma 2.5. *Let (p_t) be a transition function on the countable state space S.*

i. Then $p_t(x,x)>0$ for all $x\in S$ and $t\geqslant 0$.

ii. For all $x,y\in S$ the map $t\mapsto p_t(x,y)$ is continuous.

Before we get to the proof, note that by the Chapman-Kolmogorov equation we have for $s,t\geqslant 0$

$$p_{s+t}(x,y)\geqslant p_s(x,x)p_t(x,y). \tag{2.1}$$

Proof.

i. For $t>0$ sufficiently small the inequality holds by definition, and using (2.1) this extends to arbitrary $t>0$.

ii. By the Chapman-Kolmogorov equation we have

$$\begin{aligned}|p_{t+s}(x,y)-p_t(x,y)| &= \left|\sum_{z\in S} p_s(x,z)p_t(z,y)-p_t(x,y)\right|\\ &\leqslant p_t(x,y)|p_s(x,x)-1|+\sum_{z\neq x} p_s(x,z)p_t(z,y)\\ &= p_t(x,y)(1-p_s(x,x))+\sum_{z\in S} p_s(x,z)-p_s(x,x)\\ &= p_t(x,y)(1-p_s(x,x))+1-p_s(x,x),\end{aligned}$$

and by definition the right hand side converges to 0 for $s\to 0$. □

Theorem 2.6. *Let (p_t) be a transition function on the countable state space S.*

i. For all $x\in S$ the derivative

$$c(x):=-\partial_t p_t(x,x)|_{t=0}\in[0,\infty]$$

exists and is such that $p_t(x,x)\geqslant e^{-tc(x)}$ for all $t\geqslant 0$.

ii. If $c(x) < \infty$, then for this x and all $y \neq x$ also

$$\partial_t p_t(x, y)|_{t=0} =: q(x, y) \in [0, \infty)$$

exists and we have $\sum_y q(x, y) \leqslant 0$.

iii. If $c(x) < \infty$ and $\sum_y q(x, y) = 0$, then $t \mapsto p_t(x, y)$ is continuously differentiable for all $y \in S$, and the Kolmogorov backward equation holds:

$$\partial_t p_t(x, y) = \sum_{z \in S} q(x, z) p_t(z, y).$$

Proof.

i. Set $f(t) = -\log(p_t(x, x))$, which is well defined and continuous by Lemma 2.5. Moreover, by (2.1) we get $f(s+t) \leqslant f(s) + f(t)$. By the following exercise,

$$c(x) = \lim_{t \to 0} \frac{f(t)}{t} \in [0, \infty]$$

exists and $f(t) \leqslant c(x)t$ for all $t \geqslant 0$. If $c(x) < \infty$ we can combine this with the fact that $-\log(1 - h) = h + O(h^2)$ as $h \to 0$ to get i.. If $c(x) = \infty$, then for every $C > 0$ we have $p_t(x, x) \leqslant e^{-Ct}$ for all small $t > 0$. Therefore, $\limsup_{h \to 0} \frac{1 - p_h(x, x)}{h} \geqslant C$, and since $C > 0$ was arbitrary the claim follows.

ii. If $c(x) < \infty$ we have

$$\sum_{y \neq x} p_t(x, y) = 1 - p_t(x, x) \leqslant 1 - e^{-t c(x)} \leqslant t c(x),$$

and therefore

$$\limsup_{t \to 0} \frac{p_t(x, y)}{t} \leqslant c(x) < \infty$$

for all $y \neq x$. We define $q(x, y)$ as this limsup. It remains to show $\liminf_{t \to 0} \frac{p_t(x, y)}{t} \geqslant q(x, y)$. For that purpose not that for all $\delta > 0$ and $n \in \mathbb{N}$ we have similarly to (2.1)

$$\begin{aligned} p_{n\delta}(x, y) &= \sum_{z_1, \ldots, z_{n-1} \in S} p_\delta(x, z_1) p_\delta(z_1, z_2) \ldots p_\delta(z_{n-1}, y) \\ &\geqslant \sum_{z_1, \ldots, z_{n-1} \in S} \sum_{k=0}^{n-1} \mathbb{I}_{z_1 = \ldots = z_k = x} \mathbb{I}_{z_{k+1} = y} p_\delta(x, z_1) p_\delta(z_1, z_2) \ldots p_\delta(z_{n-1}, y) \\ &= \sum_{k=0}^{n-1} p_\delta^k(x, x) p_\delta(x, y) p_{(n-k-1)\delta}(y, y) \geqslant n e^{-n\delta c(x)} p_\delta(x, y) \times \inf_{s \in [0, n\delta]} p_s(y, y), \end{aligned}$$

and now we let $\delta_m \to 0$ be such that $\frac{p_{\delta_m}(x, y)}{\delta_m} \to q(x, y)$ and we let $n_m \to \infty$ be such that $n_m \delta_m \to t$ to obtain

$$\frac{p_t(x, y)}{t} = \lim_{m \to \infty} \frac{p_{n_m \delta_m}(x, y)}{n_m \delta_m} \geqslant e^{-t c(x)} q(x, y) \times \inf_{s \in [0, t]} p_s(y, y),$$

and now it suffices to send $t \to 0$ to obtain the claimed convergence. The inequality $\sum_{y \neq x} q(x, y) \leqslant 0$ follows from Fatou's lemma:

$$\sum_{y \neq x} q(x, y) = \sum_{y \neq x} \liminf_{t \to 0} \frac{p_t(x, y)}{t} \leqslant \liminf_{t \to 0} \sum_{y \neq x} \frac{p_t(x, y)}{t} \leqslant c(x) = -q(x, x).$$

iii. Note that

$$\frac{1}{h}(p_{t+h}(x, y) - p_t(x, y)) = \sum_{z \in S} \frac{1}{h}(p_h(x, z) - \delta_{x,z}) p_t(z, y).$$

Let now $T \subset S$ be a finite subset. Then

$$\lim_{h \to 0} \sum_{z \in T} \frac{1}{h}(p_h(x, z) - \delta_{x,z}) p_t(z, y) = \sum_{z \in T} q(x, z) p_t(z, y).$$

If we can show that $\limsup_{h\to 0}\left|\sum_{z\in T^c}\frac{1}{h}(p_h(x,z)-\delta_{x,z})p_t(z,y)\right|$ vanishes as $T^c\downarrow\emptyset$, then the claimed convergence will follow by sending $T\uparrow S$. But as soon as T^c does not contain x, we have

$$\left|\sum_{z\in T^c}\frac{1}{h}(p_h(x,z)-\delta_{x,z})p_t(z,y)\right|\leqslant\sum_{z\in T^c}\frac{1}{h}p_h(x,z)=\frac{1}{h}\left(1-\sum_{z\in T}p_h(x,z)\right)\to-\sum_{z\in T}q(x,z),$$

and since $\sum_{z\in S}q(x,z)=0$, the right hand side can be made arbitrarily small by making T large. This proves the existence and the claimed formula for the right derivative, the left differentiability follows by the same arguments. □

Exercise 2.1. Let $f\colon\mathbb{R}_+\to\mathbb{R}$ be right-continuous with $f(0)=0$ and assume that $f(s+t)\leqslant f(s)+f(t)$ for all $s,t\geqslant 0$. Then

$$\lim_{t\to 0}\frac{f(t)}{t}=\sup_{t>0}\frac{f(t)}{t}\in(-\infty,\infty]$$

exists.

Solution. *Let $t>0$. It suffices to show that*

$$\frac{f(t)}{t}\leqslant\liminf_{s\to 0}\frac{f(s)}{s}=c.$$

Let $(s_n)_{n\in\mathbb{N}}$ be such that $s_n\to 0$ and $f(s_n)/s_n\to c$. We can decompose

$$t=k_n s_n+\varepsilon_n$$

with $k_n\in\mathbb{N}_0$ and $\varepsilon_n\in[0,s_n)$. Then the subadditivity of f gives

$$\frac{f(t)}{t}=\frac{f(k_n s_n+\varepsilon)}{t}\leqslant\frac{k_n f(s_n)}{k_n s_n+\varepsilon_n}+\frac{f(\varepsilon_n)}{t}=\frac{k_n s_n}{k_n s_n+\varepsilon_n}\frac{f(s_n)}{s_n}+\frac{f(\varepsilon_n)}{t}=\left(1-\frac{\varepsilon_n}{t}\right)\frac{f(s_n)}{s_n}+\frac{f(\varepsilon_n)}{t},$$

and since f is right-continuous with $f(0)=0$, the right hand side converges to c as $n\to\infty$. This concludes the proof.

We just showed that, modulo some technical conditions, we can associate a Q-matrix to any transition function. We also saw that the q-matrix encodes the infinitesimal rate of change of our process and is maybe easier to understand than the transition function itself. So our next aim is to construct a Markov process/transition function starting from a Q-matrix.

2.3 From infinitesimal description to Markov process

In Theorem 2.6 we saw that a transition function and its Q-matrix are linked through the Kolmogorov backward equation, $\partial_t p_t=qp_t$ with initial condition $p_0=\mathrm{id}$. So given a Q-matrix we could try to construct a transition function by solving this equation. Under appropriate conditions this is indeed possible, but it is quite technical, see Section 2.5.1 in [Ligget]. So instead we will give a nice probabilistic construction which directly associates a Markov process to a given Q-matrix, again under suitable assumptions. To guess how the probabilistic construction should look like, we first study the distribution of the first jump of X if we know that $X_0=x$.

Notation. *For $x\in S$ we will from now on mostly write $\mathbb{E}_x[F]:=\mathbb{E}_{\delta_x}[F]$. Note that for any probability distribution μ on S and any bounded and measurable $F\colon S^{\mathbb{R}_+}\to\mathbb{R}$ we have from the Markov property*

$$\mathbb{E}_\mu[F]=\mathbb{E}_\mu[\mathbb{E}_\mu[F|X_0]]=\mathbb{E}_\mu[\mathbb{E}_{X_0}[F]]=\int_S\mathbb{E}_x[F]\mu(\mathrm{d}x).$$

Exercise 2.2. Let $(X_t)_{t\geqslant 0}$ be a Markov process with values in the countable state space S and assume that X has right-continuous trajectories and that $X_0=x\in S$ a.s.. Let

$$\tau=\inf\{t\geqslant 0\colon X_t\neq X_0\}.$$

Show that τ is exponentially distributed.

(*Hint:* Show first that $\mathbb{P}(\tau > t+s) = \mathbb{P}(\tau > t)\mathbb{P}(\tau > s)$).

Solution. *Let* $\tau_r = \inf\{t \geqslant r : X_t \neq X_r\}$*, so that* $\tau = \tau_0$*. Then*

$$\mathbb{P}(\tau > t+s) = \mathbb{E}[\mathbb{I}_{\tau_t > t+s}\mathbb{I}_{\tau > t}] = \mathbb{E}[\mathbb{E}[\mathbb{I}_{\tau_t > t+s}|\mathcal{F}_t^X]\mathbb{I}_{\tau > t}]$$
$$= \mathbb{E}[\mathbb{E}_{X_t}[\mathbb{I}_{\tau_0 > s}]\mathbb{I}_{\tau > t}] = \mathbb{E}[\mathbb{E}_x[\mathbb{I}_{\tau > s}]\mathbb{I}_{\tau > t}] = \mathbb{P}_x(\tau > s)\mathbb{P}(\tau > t),$$

but of course $\mathbb{P}_x(\tau > s) = \mathbb{P}(\tau > s)$ *because* $\mathbb{P}(X_0 = x) = 1$*. From here we get for* $k, n \in \mathbb{N}$

$$\mathbb{P}\left(\tau > \frac{k}{n}\right) = \mathbb{P}\left(\tau > \frac{1}{n}\right)^k = \mathbb{P}(\tau > 1)^{\frac{k}{n}}.$$

Moreover, $\lim_{\varepsilon \to 0}\mathbb{P}(\tau > \varepsilon) = \mathbb{P}(\tau > 0) = 1$ *because* X *has right-continuous trajectories, and combining this with the multiplicativity we get that* $t \mapsto \mathbb{P}(\tau > t)$ *is continuous, which means that*

$$\mathbb{P}(\tau > t) = \mathbb{P}(\tau > 1)^t$$

for all $t \geqslant 0$*. So setting* $c = -\log(\mathbb{P}(\tau > 1))$ *we get* $\mathbb{P}(\tau > t) = e^{-\mathrm{ct}}$ *for all* $t \geqslant 0$*, and therefore* τ *is exponentially distributed with parameter* $c \in [0, \infty)$*. Note that* $\mathbb{P}(\tau > 1) > 0$ *because*

$$\mathbb{P}(\tau > 1) = \mathbb{P}\left(\tau > \frac{1}{n}\right)^n,$$

and for large n *the right hand side is strictly positive; but it could happen that* $\mathbb{P}(\tau > 1) = 1$ *and then* $c = 0$*.*

Remark 2.7. If $\mathbb{P} \circ X_0^{-1} = \mu$ (we also write $\mathrm{law}(X_0) = \mu$ or $X_0 \sim \mu$) and μ is supported in more than one state, then τ is in general not exponentially distributed: We just showed that for all $x \in S$ there exists $c(x) \in [0, \infty)$ with $\mathbb{P}_x(\tau > t) = e^{-t\,c(x)}$. But then

$$\mathbb{P}(\tau > t) = \int_S \mathbb{P}_x(\tau > t)\mu(\mathrm{d}x) = \int_S e^{-t\,c(x)}\mu(\mathrm{d}x)$$

is only a mixture of exponential variables, but in general not exponentially distributed itself.

To construct a Markov process from a Q-matrix q we start by constructing a discrete time transition matrix P from q. If $c(x) = 0$ we set $P(x,x) = 1$ and $P(x,y) = 0$ for $y \neq x$. If $c(x) > 0$ we define

$$P(x,y) = \begin{cases} \dfrac{q(x,y)}{c(x)}, & y \neq x, \\ 0, & y = x. \end{cases}$$

Note that $\sum_y P(x,y) = 1$ by definition of $c(x)$. Let μ be a probability distribution on S and let $(Y_n)_{n \in \mathbb{N}_0}$ be the discrete time Markov chain with initial distribution μ and transition matrix P. To construct a continuous time transition mechanism, recall that by Exercise 2.2 the jump times must be exponentially distributed. So consider an i.i.d. sequence of standard exponential random variables $(\sigma_n)_{n \in \mathbb{N}}$, independent of Y, and set $\tau_n := \frac{\sigma_n}{c(Y_{n-1})}$ with $\frac{\sigma_n}{0} = \infty$. Note that conditionally on Y_n the random variable τ_n is exponentially distributed with parameter $c(Y_n)$. Next, let $T_0 := 0$ and $T_n := \tau_1 + \ldots + \tau_n$ for $n \in \mathbb{N}$, and then

$$N_t := \#\{n \in \mathbb{N}_{>0} : T_n \leqslant t\} = \sum_{n=1}^{\infty} n\mathbb{I}_{[T_n, T_{n+1})}(t) + \infty \times \mathbb{I}_{\sup_n T_n \leqslant t}.$$

Finally, set $X_t := Y_{N_t}$ for $t \geqslant 0$, where we introduce a cemetery state $\Delta \notin S$ and define $Y_\infty := \Delta$. Note that X has right-continuous paths as long as there is no explosion and Δ is not reached.

Lemma 2.8. $(T_n, Y_n)_{n \in \mathbb{N}_0}$ *is a (discrete time) Markov process with transition probability*

$$\mathbb{P}(T_{n+1} \in A, Y_{n+1} = y | (T_0, Y_0), \ldots, (T_n, Y_n))$$
$$= \mathbb{I}_{c(Y_n) > 0} P(Y_n, y) \int_{T_n}^{\infty} \mathbb{I}_A(s) c(Y_n) e^{-c(Y_n)(s - T_n)} \mathrm{d}s + \mathbb{I}_{c(Y_n) = 0}(\delta_{Y_n} \otimes \delta_\infty)(\{y\} \times A).$$

Proof. Let F be a bounded measurable function on $S\times[0,\infty]$. Then we have on thet set $\{c(Y_n)>0\}$

$$\begin{aligned}\mathbb{E}[F(T_{n+1},Y_{n+1})|(T_0,Y_0),\dots,(T_n,Y_n)] &= \mathbb{E}\left[F\left(T_n+\frac{\sigma_n}{c(Y_n)},Y_{n+1}\right)\middle|(T_0,Y_0),\dots,(T_n,Y_n)\right]\\ &= \mathbb{E}\left[\mathbb{E}\left[F\left(T_n+\frac{\sigma_n}{c(Y_n)},Y_{n+1}\right)\middle|(T_0,Y_0),\dots,(T_n,Y_n),Y_{n+1}\right]\middle|(T_0,Y_0),\dots,(T_n,Y_n)\right]\\ &= \mathbb{E}\left[\int_{T_n}^{\infty}F(Y_{n+1},s)c(Y_n)e^{-c(Y_n)(s-T_n)}\mathrm{d}s\middle|(T_0,Y_0),\dots,(T_n,Y_n)\right]\\ &= \sum_{y\in S}P(Y_n,y)\int_{T_n}^{\infty}F(y,s)c(Y_n)e^{-c(Y_n)(s-T_n)}\mathrm{d}s,\end{aligned}$$

where in the last step we used the Markov property of (Y_n). On the set $\{c(Y_n)=0\}$ we trivially have $T_n\geqslant\tau_n=\infty$ and $Y_{n+1}=Y_n$. □

Lemma 2.9. *(Loss of memory property)*

Assume that $\mathbb{P}(N_s=\infty)=0$ for all $s\geqslant 0$. The conditional distribution of T_{N_s+1} given $\mathcal{F}_s^X$ is given by

$$\mathbb{P}(T_{N_s+1}>t|\mathcal{F}_s^X)=e^{-(t-s)c(X_s)},$$

i.e. for predicting the next jump after time s it does matter how long X stayed in X_s before time s.

Proof. Let $A\in\mathcal{F}_s^X$ and $n\in\mathbb{N}_0$. From the construction of X it is intuitively clear (and can be rigorously shown) that there exists $A_n\in\sigma((T_0,Y_0),\dots,(T_n,Y_n))$ such that

$$A\cap\{N_s=n\}=A_n\cap\{N_s=n\}.$$

Therefore, Lemma 2.8 gives together with $T_{N_s}\leqslant s\leqslant t$

$$\begin{aligned}\mathbb{P}(\{T_{N_s+1}>t\}\cap A\cap\{N_s=n\}) &= \mathbb{P}(\{T_{n+1}>t\}\cap A_n\cap\{N_s=n\})\\ &= \mathbb{E}\left[\left(\mathbb{I}_{c(Y_n)>0}\int_{T_n}^{\infty}\mathbb{I}_{r>t}c(Y_n)e^{-c(Y_n)(r-T_n)}\mathrm{d}r+\mathbb{I}_{c(Y_n)=0}\right)\mathbb{I}_{A_n\cap\{N_s=n\}}\right]\\ &= \mathbb{E}\left[\left(\mathbb{I}_{c(Y_n)>0}\int_{t}^{\infty}c(Y_n)e^{-c(Y_n)(r-T_n)}\mathrm{d}r+\mathbb{I}_{c(Y_n)=0}\right)\mathbb{I}_{A_n\cap\{N_s=n\}}\right]\\ &= \mathbb{E}[(\mathbb{I}_{c(Y_n)>0}e^{-c(Y_n)(t-T_n)}+\mathbb{I}_{c(Y_n)=0})\mathbb{I}_{A_n\cap\{N_s=n\}}]\\ &= \mathbb{E}[e^{-c(Y_n)(t-T_n)}\mathbb{I}_{A\cap\{N_s=n\}}].\end{aligned}$$

Now

$$e^{-c(Y_n)(t-T_n)}=e^{-c(Y_n)(t-s)}e^{-c(Y_n)(s-T_n)}=e^{-c(Y_n)(t-s)}\mathbb{P}(T_{n+1}>s|(T_0,Y_0),\dots,(T_n,Y_n)),$$

and since $T_{N_s+1}\geqslant s$ by definition we obtain

$$\begin{aligned}\mathbb{E}[e^{-c(Y_n)(t-T_n)}\mathbb{I}_{A\cap\{N_s=n\}}] &= \mathbb{E}[e^{-c(Y_n)(t-s)}\mathbb{P}(T_{n+1}>(s-T_n)|(T_0,Y_0),\dots,(T_n,Y_n))\mathbb{I}_{A\cap\{N_s=n\}}]\\ &= \mathbb{E}[e^{-c(Y_n)(t-s)}\mathbb{I}_{T_{n+1}>s}\mathbb{I}_{A\cap\{N_s=n\}}]=\mathbb{E}[e^{-c(Y_n)(t-s)}\mathbb{I}_{A\cap\{N_s=n\}}]\\ &= \mathbb{E}[e^{-c(X_s)(t-s)}\mathbb{I}_{A\cap\{N_s=n\}}].\end{aligned}$$

Now the claim follows by summing over $n\in\mathbb{N}_0$. □

Remark 2.10. The same proof shows that also

$$\mathbb{P}(T_{N_s+1}>t|\mathcal{F}_s^X,Y_{N_s+1})=e^{-(t-s)c(X_s)}.$$

Note that $(X_t)_{t\geqslant 0}$ is given as a deterministic map of the Markov chain $(Y_n,T_n)_{n\in\mathbb{N}_0}$:

$$X_t=\Phi_t(T_0,Y_0,T_1,Y_1,\dots),$$

where for $t\geqslant t_0$

$$\Phi_t(t_0,y_0,t_1,y_1,\dots):=\begin{cases}y_n, & t\in[t_n,t_{n+1}),\\ \Delta, & t\geqslant\sup_n t_n.\end{cases}$$

Theorem 2.11. *Assume that* $\mathbb{P}(N_t=\infty)=0$ *for all* $t\geqslant 0$*. Then* X *is a time-homogeneous Markov process.*

Proof. Since $t\geqslant s$ we have outside the null set $\{N_s=\infty\}$

$$X_t=\Phi_t(s,Y_{N_s},T_{N_s+1},Y_{N_s+1},T_{N_s+2},Y_{N_s+2},\ldots)$$

for the same map Φ as above. So if $f\colon S\to\mathbb{R}$ is bounded and measurable, then

$$\begin{aligned}\mathbb{E}[f(X_t)|\mathcal{F}_s^X]&=\mathbb{E}[f(\Phi_t(s,Y_{N_s},T_{N_s+1},Y_{N_s+1},T_{N_s+2},Y_{N_s+2},\ldots))|\mathcal{F}_s^X]\\&=\mathbb{E}[\mathbb{E}[f(\Phi_t(s,Y_{N_s},T_{N_s+1},Y_{N_s+1},T_{N_s+2},Y_{N_s+2},\ldots))|(T_k,Y_k)_{0\leqslant k\leqslant N_s}]|\mathcal{F}_s^X]\\&=\mathbb{E}[\Gamma_{s,t}(Y_{N_s},T_{N_s+1},Y_{N_s+1})|\mathcal{F}_s^X]\end{aligned}$$

with

$$\Gamma_{s,t}(y_0,t_1,y_1)=\mathbb{E}[f(\Phi_t(s,y_0,t_1,y_1,T_1,Y_1,\ldots))|(T_0,Y_0)=(t_1,y_1)],$$

where we applied the strong Markov property of the discrete time Markov chain $(Y_n,T_n)_{n\in\mathbb{N}_0}$ (if you do not feel comfortable with this, for example because we did not check that N_s is a stopping time, we could also multiply with the indicator function $\mathbb{I}_{\{N_s=n\}}$ and sum over n in the end). By Remark 2.10 we have

$$\begin{aligned}\mathbb{E}[\Gamma_{s,t}(Y_{N_s},T_{N_s+1},Y_{N_s+1})|\mathcal{F}_s^X]&=\int_s^\infty\mathbb{E}[\Gamma_{s,t}(Y_{N_s},r,Y_{N_s+1})|\mathcal{F}_s^X]c(X_s)e^{-(r-s)c(X_s)}\mathrm{d}r\\&=\int_s^\infty\mathbb{E}\left[\sum_{y\in S}P(Y_{N_s},y)\Gamma_{s,t}(Y_{N_s},r,y)\middle|\mathcal{F}_s^X\right]c(X_s)e^{-(r-s)c(X_s)}\mathrm{d}r\\&=\int_s^\infty\sum_{y\in S}P(X_s,y)\Gamma_{s,t}(X_s,r,y)c(X_s)e^{-(r-s)c(X_s)}\mathrm{d}r\\&=\int_0^\infty\sum_{y\in S}P(X_s,y)\Gamma_{s,t}(X_s,r+s,y)c(X_s)e^{-rc(X_s)}\mathrm{d}r,\end{aligned}$$

which proves the Markov property. Since

$$\begin{aligned}\Gamma_{s,t}(y_0,t_1+s,y_1)&=\mathbb{E}[f(\Phi_t(s,y_0,t_1+s,y_1,T_1,Y_1,\ldots))|(T_0,Y_0)=(t_1+s,y_1)]\\&=\mathbb{E}[f(\Phi_{t-s}(0,y_0,t_1,y_1,T_1,Y_1,\ldots))|(T_0,Y_0)=(t_1,y_1)]=\Gamma_{0,t-s}(y_0,t_1,y_1),\end{aligned}$$

the process is also time-homogeneous. □

To apply this theorem, we need to make sure that no explosion occurs, i.e. that $\mathbb{P}(N_t=\infty)=0$ for all $t\geqslant 0$. If this is the case, then $(X_t)_{t\geqslant 0}$ has càdlàg trajectories (right-continuous and limits from the left exist) by construction. But the following example shows that explosions might happen:

Example 2.12. Consider the Q-matrix on $S=\mathbb{N}$ with

$$q(x,x+1)=x^2,\qquad q(x,x)=-x^2$$

and $q(x,y)=0$ for $y\neq x,x+1$ and let X be the process constructed above, with initial distribution $X_0\sim\delta_1$. Then $\mathbb{P}(N_t=\infty)>0$ for all sufficiently large $t>0$. Indeed, we have $Y_n=n+1$ and therefore τ_n is exponentially distributed with parameter n^2. Thus,

$$\mathbb{E}\left[\sum_{n=1}^\infty\tau_n\right]=\sum_{n=1}^\infty\frac{1}{n^2}<\infty,$$

and therefore

$$\mathbb{P}(N_t<\infty)=\mathbb{P}\left(\sum_{n=1}^\infty\tau_n>t\right)\leqslant\frac{\sum_{n=1}^\infty\frac{1}{n^2}}{t}<1$$

for all sufficiently large $t>0$. Indeed, we even have

$$\mathbb{P}(N_t<\infty\text{ for all }t\geqslant 0)\leqslant\liminf_{m\to\infty}\mathbb{P}(N_m<\infty)=0.$$

So at some point there is an infinitely fast succession of jumps and the process diverges to $+\infty$ in finite time.

A simple way of ruling out explosions is the following:

Lemma 2.13. *Let q be a Q-matrix and let μ be an initial distribution on S. Let $(Y_n)_{n\in\mathbb{N}_0}$ and $(N_t)_{t\geqslant 0}$ be the processes constructed above and assume that a.s.*

$$\sum_{n=0}^{\infty}\frac{1}{c(Y_n)}=\infty.$$

Then $\mathbb{P}(N_t=\infty)=0$ for all $t\geqslant 0$. This holds for any initial distribution μ if we assume that for any infinite subset $T\subset S$ we have

$$\sum_{x\in T}\frac{1}{c(x)}=\infty,$$

and in particular if $\sup_{x\in S}c(x)<\infty$.

Proof. It suffices to show that

$$\mathbb{P}(N_t=\infty|Y_0,Y_1,\ldots)=\mathbb{P}\left(\sum_{n=1}^{\infty}\tau_n\leqslant t\middle|Y_0,Y_1,\ldots\right)\leqslant\mathbb{P}\left(\sum_{n=1}^{\infty}\tau_n<\infty\middle|Y_0,Y_1,\ldots\right)=0.$$

But conditionally on $(Y_0,Y_1,\ldots)$ the (τ_n) are independent exponentially distributed random variables with parameters $(c(Y_{n-1}))$. Therefore,

$$\begin{aligned}\mathbb{P}\left(\sum_{n=1}^{\infty}\tau_n<\infty\middle|Y_0,Y_1,\ldots\right)&=\lim_{\lambda\downarrow 0}\mathbb{E}(e^{-\lambda\sum_{n=1}^{\infty}\tau_n}|Y_0,Y_1,\ldots)=\lim_{\lambda\downarrow 0}\prod_{n=0}^{\infty}\frac{c(Y_n)}{\lambda+c(Y_n)}\\&=\lim_{\lambda\downarrow 0}\prod_{n=0}^{\infty}\left(1-\frac{\lambda}{\lambda+c(Y_n)}\right).\end{aligned}$$

Now the product on the right hand side is zero if and only if

$$\sum_{n=0}^{\infty}\log\left(1-\frac{\lambda}{\lambda+c(Y_n)}\right)=-\infty.$$

But $\log(1-x)\leqslant -x$, and therefore

$$\sum_{n=0}^{\infty}\log\left(1-\frac{\lambda}{\lambda+c(Y_n)}\right)\leqslant(-\lambda)\sum_{n=0}^{\infty}\frac{1}{\lambda+c(Y_n)}.$$

By assumption we have $\sum_{n=0}^{\infty}\frac{1}{c(Y_n)}=\infty$, so the proof is complete once we show that then also for all $\lambda>0$

$$\sum_{n=0}^{\infty}\frac{1}{\lambda+c(Y_n)}=\infty.$$

But either there exists $N\in\mathbb{N}$ with $c(Y_n)\geqslant 1$ for all $n\geqslant N$, in which case

$$\sum_{n=0}^{\infty}\frac{1}{\lambda+c(Y_n)}\geqslant\sum_{n=N}^{\infty}\frac{1}{\lambda+c(Y_n)}=\sum_{n=N}^{\infty}\frac{1}{c(Y_n)}\frac{c(Y_n)}{\lambda+c(Y_n)}\geqslant\sum_{n=N}^{\infty}\frac{1}{c(Y_n)}\frac{1}{\lambda+1}=\infty,$$

or $c(Y_n)\leqslant 1$ for infinitely many n, say $n\in T\subset\mathbb{N}_0$, and then

$$\sum_{n=0}^{\infty}\frac{1}{\lambda+c(Y_n)}\geqslant\sum_{n\in T}\frac{1}{\lambda+1}=\frac{|T|}{\lambda+1}=\infty.$$

□

Example 2.14. (Birth-death chain)

Consider a population where each individual dies with rate 1 and gives rise to an offspring with rate 1, independently of all the other individuals. In case there is an offspring it also evolves independently of its ancestor and all other individuals. We track the total number of particles. Then the state space is $S=\mathbb{N}_0$ and the rates of transition are

$$q(x,y)=x\delta_{y,x+1}+x\delta_{y,x-1},$$

and in particular $c(x)=2x$. We do not have $\sum_{x\in T}\frac{1}{c(x)}=\infty$ for all infinite subsets $T\subset S$ (take for example $T=\{x^2 : x\in\mathbb{N}\}$). But nonetheless we have for any initial distribution μ that $\mathbb{P}(N_t=\infty)=0$: It suffices to show this for $\mu=\delta_x$ with $x\in\mathbb{N}_0$, in which case

$$\sum_{n=0}^{\infty}\frac{1}{c(Y_n)}\geqslant\sum_{y=x}^{\infty}\frac{1}{2y}=\infty.$$

So the dynamics are well defined.

If q is a Q-matrix with $\sum_{x\in T}\frac{1}{c(x)}=\infty$ for any infinite subset $T\subset S$, then we can construct for any $x\in S$ the Markov process $(X_t)_{t\geqslant 0}$ with initial distribution $X_0=x$ a.s., and we write $\mathbb{P}_x$ for its distribution. It is then straightforward to check that

$$p_t(x,y):=\mathbb{P}_x(X_t=y)$$

defines a transition function.

Exercise 2.3. Show that (p_t) solves the Kolmogorov backward equation $\partial_t p_t=qp_t$.

Solution. *For small $h>0$ have*

$$p_h(x,y)=\mathbb{P}_x(\{X_h=y\}\cap\{N_h=0\})+\mathbb{P}_x(\{X_h=y\}\cap\{N_h=1\})+\mathbb{P}_x(\{X_h=y\}\cap\{N_h\geqslant 2\})$$

with

$$\begin{aligned}\mathbb{P}_x(\{X_h=y\}\cap\{N_h\geqslant 2\})&\leqslant\mathbb{P}_x(N_h\geqslant 2)=\mathbb{P}_x(\tau_1+\tau_2\leqslant h)\\&\leqslant\mathbb{P}_x(\{\tau_1\leqslant h\}\cap\{\tau_2\leqslant h\})=\mathbb{E}_x[(1-e^{-c(x)h})(1-e^{-c(Y_1)h})]\\&=(1-e^{-c(x)h})\sum_{y\neq x}\frac{q(x,y)}{c(x)}(1-e^{-c(y)h})\\&\leqslant c(x)h\sum_{y\neq x}\frac{q(x,y)}{c(x)}(1-e^{-c(y)h})=h\sum_{y\neq x}q(x,y)(1-e^{-c(y)h}),\end{aligned}$$

and by the dominated convergence theorem the right hand side converges to zero after dividing it by h and sending $h\to 0$. In other words, on an interval of length h the probability of making at least two jumps is of order $o(h)$. It remains to consider the cases with 0 or 1 jump:

$$\mathbb{P}_x(\{X_h=y\}\cap\{N_h=0\})=\delta_{x,y}\mathbb{P}_x(\tau_1>h)=\delta_{x,y}e^{-c(x)h}$$

and since on the set $c(Y_0)\neq 0$ we have $Y_1\neq Y_0$, we get

$$\begin{aligned}\mathbb{P}_x(\{X_h=y\}\cap\{N_h=1\})&=\mathbb{P}_x(\{Y_1=y\}\cap\{N_h=1\})\\&=\mathbb{P}_x(\{Y_1=y\}\cap\{N_h\geqslant 1\})-\mathbb{P}(\{Y_1=y\}\cap\{N_h\geqslant 2\})\\&=(1-\delta_{x,y})\mathbb{P}_x(\{Y_1=y\}\cap\{\tau_1\leqslant h\})+o(h)\\&=(1-\delta_{x,y})\frac{q(x,y)}{c(x)}(1-e^{-c(x)h})+o(h).\end{aligned}$$

Thus,

$$\frac{p_h(x,y)-p_0(x,y)}{h}=\delta_{x,y}\frac{e^{-c(x)h}-1}{h}+(1-\delta_{x,y})\frac{q(x,y)(1-e^{-c(x)h})}{c(x)h}+\frac{o(h)}{h},$$

and therefore

$$\partial_t p_t(x,y)|_{t=0}=q(x,y).$$

Since q and (p_t) satisfy all the properties required in Theorem 2.6, iii., it follows that p solves the Kolmogorov backward equation.

We end this discussion with an auxiliary result that will be very useful later:

Lemma 2.15. *Let $(X_t)_{t\geqslant 0}$ be a Markov process with respect to $(\mathcal{F}_t)$ with values in the countable state space S and assume that X has right-continuous trajectories. Then the strong Markov property holds: If τ is a stopping time and $F: S^{\mathbb{R}_+}\to\mathbb{R}$ is a bounded and measurable function, then a.s.*

$$\mathbb{E}[F((X_{\tau+t})_{t\geqslant 0})|\mathcal{F}_\tau]\mathbb{I}_{\tau<\infty}=\mathbb{E}_{X_\tau}[F]\mathbb{I}_{\tau<\infty},$$

where $\mathbb{E}_{X_\tau}$ *denotes the expectation with respect to the measure* $\mathbb{P}_{\delta_{X_\tau}}$ *on* $(S^{\mathbb{R}_+}, \mathcal{B}^{\otimes \mathbb{R}_+})$ *which is constructed in Theorem 1.10.*

Proof. Assume first that τ only takes finitely many values $0 \leqslant t_1 \leqslant \ldots \leqslant t_n < \infty$ and ∞, and let $A \in \mathcal{F}_\tau$. Then for all $k = 1, \ldots, n$ we have $A \cap \{\tau = t_k\} = (A \cap \{\tau \leqslant t_k\}) \setminus (A \cap \{\tau \leqslant t_{k-1}\}) \in \mathcal{F}_{t_k}$, and therefore by the ("weak") Markov property

$$\mathbb{E}[F((X_{\tau+t})_{t\geqslant 0})\mathbb{I}_{A\cap\{\tau<\infty\}}] = \sum_{k=1}^n \mathbb{E}[F((X_{t_k+t})_{t\geqslant 0})\mathbb{I}_{A\cap\{\tau=t_k\}}]$$
$$= \sum_{k=1}^n \mathbb{E}[\mathbb{E}[F((X_{t_k+t})_{t\geqslant 0})|\mathcal{F}_{t_k}]\mathbb{I}_{A\cap\{\tau=t_k\}}] = \sum_{k=1}^n \mathbb{E}[\mathbb{E}_{X_{t_k}}[F]\mathbb{I}_{A\cap\{\tau=t_k\}}]$$
$$= \sum_{k=1}^n \mathbb{E}[\mathbb{E}_{X_\tau}[F]\mathbb{I}_{A\cap\{\tau=t_k\}}] = \mathbb{E}[\mathbb{E}_{X_\tau}[F]\mathbb{I}_{A\cap\{\tau<\infty\}}].$$

Since this holds for all $A \in \mathcal{F}_\tau$, the claim follows in the special case where τ only takes finitely many values. Now let τ be a general stopping time and consider a sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$, such that each τ_n only takes finitely many values and such that $\tau_n \downarrow \tau$. Then $\mathcal{F}_\tau \subset \mathcal{F}_{\tau_n}$, and therefore the considerations above yield for $A \in \mathcal{F}_\tau$

$$\mathbb{E}[F((X_{\tau_n+t})_{t\geqslant 0})\mathbb{I}_{A\cap\{\tau_n<\infty\}}] = \mathbb{E}[\mathbb{E}_{X_{\tau_n}}[F]\mathbb{I}_{A\cap\{\tau_n<\infty\}}]. \tag{2.2}$$

We would like to send $n \to \infty$ in this equality. On the right hand side this is easy, because since X is right-continuous in a discrete space we have $\lim_{n\to\infty}\mathbb{P}(X_{\tau_n} \neq X_\tau) = 0$, and thus

$$\lim_{n\to\infty} \mathbb{E}[\mathbb{E}_{X_{\tau_n}}[F]\mathbb{I}_{A\cap\{\tau_n<\infty\}}] = \lim_{n\to\infty} (\mathbb{E}[\mathbb{E}_{X_\tau}[F]\mathbb{I}_{A\cap\{\tau_n<\infty\}}] + \varepsilon_n),$$

with

$$|\varepsilon_n| = |\mathbb{E}[(\mathbb{E}_{X_{\tau_n}}[F] - \mathbb{E}_{X_\tau}[F])\mathbb{I}_{A\cap\{\tau_n<\infty\}\cap\{X_{\tau_n}\neq X_\tau\}}]| \leqslant 2 \sup_{\omega\in S^{\mathbb{R}_+}} |F(\omega)| \times \mathbb{P}(X_{\tau_n} \neq X_\tau) \to 0.$$

To obtain the convergence on the left hand side of (2.2) we need to assume that F is of a special form: If $F(\omega) = f(\omega(s_1), \ldots, \omega(s_m))$ for a bounded function $f: S^m \to \mathbb{R}$ (which is automatically continuous because S^m is equipped with the discrete topology), then the same argument as for the right hand side of (2.2) gives

$$\mathbb{E}[F((X_{\tau+t})_{t\geqslant 0})\mathbb{I}_{A\cap\{\tau<\infty\}}] = \lim_{n\to\infty} \mathbb{E}[F((X_{\tau_n+t})_{t\geqslant 0})\mathbb{I}_{A\cap\{\tau_n<\infty\}}]$$
$$= \lim_{n\to\infty} \mathbb{E}[\mathbb{E}_{X_{\tau_n}}[F]\mathbb{I}_{A\cap\{\tau_n<\infty\}}] = \mathbb{E}[\mathbb{E}_{X_\tau}[F]\mathbb{I}_{A\cap\{\tau<\infty\}}].$$

The claim for general F follows from a monotone class argument. □

2.4 The martingale problem

We just showed that for every Q-matrix we can construct a Markov process and a corresponding transition function, provided that we can rule out explosions (e.g. if $\sup_{x\in S} c(x) < \infty$). In this section we present an alternative description of $(X_t)_{t\geqslant 0}$ which will be very useful in what follows. In the next theorem we will need a simple observation:

Lemma 2.16. *If* $\|c\|_\infty := \sup_{x\in S} |c(x)| < \infty$, *then* $f \mapsto qf$ *is a bounded linear operator on*

$$\mathcal{M}(S) := \{f: S \to \mathbb{R} \,|\, \|f\|_\infty < \infty\},$$

with operator norm bounded by

$$\|q\|_{L(\mathcal{M}(S),\mathcal{M}(S))} := \sup_{\substack{f\in\mathcal{M}(S),\\ \|f\|_\infty\leqslant 1}} \|qf\|_\infty \leqslant 2\|c\|_\infty.$$

Proof. We have

$$|qf(x)| = \left|\sum_{y\in S} q(x,y)f(y)\right| \leqslant \sum_{y\in S} |q(x,y)f(y)| \leqslant \sum_{y\in S} |q(x,y)| \times \|f\|_\infty = 2c(x)\|f\|_\infty \leqslant 2\|c\|_\infty \|f\|_\infty.$$

□

Theorem 2.17. *Assume that* $\|c\|_\infty < \infty$ *and let* $f: S \to \mathbb{R}$ *be bounded. Then*

i. *Strong continuity holds: the function* $t \mapsto p_t f$ *is continuous with respect to the supremum norm.*

ii. *The transition function has the analytic representation*

$$p_t = e^{tq}, \qquad t \geqslant 0.$$

iii. *The Kolmogorov forward equation holds:*

$$\partial_t p_t f(x) = p_t q f(x) = p_t(qf)(x) = \sum_{y,z} p_t(x,y) q(y,z) f(z)$$

for all $x \in S$.

iv. *For any initial distribution the associated Markov process* $(X_t)_{t \geqslant 0}$ *solves the martingale problem:*

$$M_t^f := f(X_t) - f(X_0) - \int_0^t qf(X_s) \mathrm{d}s$$

is a martingale.

Proof.

i. By Theorem 2.6 the Kolmogorov backward equation holds, and integrating it in time we get

$$p_t(x,y) = \delta_{x,y} + \int_0^t (qp_s)(x,y) \mathrm{d}s,$$

and multiplying with f on both sides gives

$$\begin{aligned} |p_t f(x) - f(x)| &= \left| \sum_y \int_0^t (qp_s)(x,y) f(y) \mathrm{d}s \right| \leqslant \int_0^t \sum_y |(qp_s)(x,y) f(y)| \mathrm{d}s \\ &\leqslant \int_0^t \sum_{y,z} |q(x,z)| p_s(z,y) |f(y)| \mathrm{d}s \leqslant 2c(x) \|f\|_\infty. \end{aligned}$$

Together with the Chapman-Kolmogorov equation we obtain

$$\|p_{s+t} f - p_s f\|_\infty = \|(p_t - \mathrm{id}) p_s f\|_\infty \leqslant 2\|c\|_\infty \|p_s f\|_\infty \leqslant 2\|c\|_\infty \|f\|_\infty.$$

ii. From Lemma 2.16 we know that q is a bounded linear operator. Therefore, the unique continuous solution to $p_t = \mathrm{id} + \int_0^t q p_s \mathrm{d}s$ is given by the operator exponential $p_t = e^{tq}$ (to prove this use the Banach fixed point theorem to get uniqueness of solutions, then differentiate e^{tq} in $t = 0$ to see that the equation is satisfied in this point, then use $e^{(t+s)q} = e^{tq} e^{sq}$.

iii. Since $p_t = e^{tq}$, we get

$$p_t q = e^{tq} q = \left(\sum_{k=0}^\infty \frac{t^k}{k!} q^k \right) q = \sum_{k=0}^\infty \frac{t^k}{k!} q^{k+1} = q \left(\sum_{k=0}^\infty \frac{t^k}{k!} q^k \right) = q e^{tq} = q p_t.$$

Therefore, the Kolmogorov forward equation holds as a consequence of the Kolmogorov backward equation.

iv. Since f is bounded M^f is integrable. Moreover, the Kolmogorov forward equation gives

$$\begin{aligned} \mathbb{E}[f(X_t) | \mathcal{F}_s^X] &= p_{t-s} f(X_s) = f(X_s) + \int_0^{t-s} \partial_r p_r f(X_s) \mathrm{d}r \\ &= f(X_s) + \int_0^{t-s} p_r q f(X_s) \mathrm{d}r = f(X_s) + \int_0^{t-s} \mathbb{E}[qf(X_{s+r}) | \mathcal{F}_s^X] \mathrm{d}r \\ &= f(X_s) + \mathbb{E}\left[\int_0^{t-s} qf(X_{s+r}) \mathrm{d}r \middle| \mathcal{F}_s^X \right] = f(X_s) + \mathbb{E}\left[\int_s^t qf(X_r) \mathrm{d}r \middle| \mathcal{F}_s^X \right], \end{aligned}$$

and therefore

$$\mathbb{E}\left[f(X_t) - f(X_0) - \int_0^t qf(X_r)\mathrm{d}r \middle| \mathcal{F}_s^X\right] = f(X_s) - f(X_0) - \int_0^s qf(X_r)\mathrm{d}r.$$

□

Our next aim is to show that the martingale problem characterizes the law of X uniquely. This will be very useful in what follows, because it gives us a probabilistic description of X and because it allows us to describe many interesting functionals of X in terms of martingales, for which we have many nice properties.

To do so, we need an auxiliary result:

Lemma 2.18. *Let $\|c\|_\infty < \infty$. Then for all $\alpha > 0$ and all bounded $g: S \to \mathbb{R}$ the resolvent equation*

$$(\alpha - q)f = g$$

has a unique solution $f: S \to \mathbb{R}$ with $\|f\|_\infty \leqslant \alpha^{-1}\|g\|_\infty$. We write $f := R_\alpha g = (\alpha - q)^{-1}q$.

Proof. We showed in the last section that there exists a transition function $(p_t)_{t \geqslant 0}$ associated to q. Define

$$R_\alpha g(x) := \int_0^\infty e^{-\alpha t} p_t g(x)\mathrm{d}t$$

with $|R_\alpha g(x)| \leqslant \alpha^{-1}\|g\|_\infty$, so that R_α is a bounded linear operator. Moreover, as a bounded linear operator q commutes with the time integral, and therefre

$$qR_\alpha g = \int_0^\infty e^{-\alpha t} q p_t g \mathrm{d}t = \int_0^\infty e^{-\alpha t} \partial_t p_t g \mathrm{d}t = \int_0^\infty [\partial_t (e^{-\alpha t} p_t g) - (-\alpha) e^{-\alpha t} p_t g]\mathrm{d}t = -g + \alpha R_\alpha g,$$

and therefore $(\alpha - q)R_\alpha g = g$. Furthermore,

$$qR_\alpha g = \int_0^\infty e^{-\alpha t} q p_t g \mathrm{d}t = \int_0^\infty e^{-\alpha t} p_t q g \mathrm{d}t = R_\alpha q g,$$

where the commutation of p_t and q is possible because the transition function solves both the Kolmogorov forward and backward equations. Thus, we have also

$$R_\alpha(\alpha - q)g = g,$$

and therefore R_α is a bijection with inverse $(\alpha - q)$. □

Exercise 2.4. Note that the condition $\alpha > 0$ was used in essentially every step in the argument above. For $\alpha \leqslant 0$ the statement in general fails completely:

Let q be a Q-matrix on a finite state space S and let λ be an eigenvalue of q. Use Lemma 2.18 to show that $\lambda \in \mathbb{C} \setminus (0, \infty)$.

Solution. *If $\lambda \in \mathbb{C}$ is an eigenvalue with eigenvector $f \neq 0$, then $qf = \lambda f$ and therefore $(\lambda - q)f = 0$. On the other hand for $\alpha \in (0, \infty)$ we know by Lemma 2.18 that $(\alpha - q)g = 0$ only for $g = 0$, and thus $\lambda \notin (0, \infty)$.*

Theorem 2.19. *Let μ be a probability distribution on S and let q be a Q-matrix with $\|c\|_\infty < \infty$. Let $(X_t)_{t \geqslant 0}$ be a càdlàg stochastic process with values in S such that $X_0 \sim \mu$ and for all bounded $f: S \to \mathbb{R}$ the process*

$$M_t^f := f(X_t) - f(X_0) - \int_0^t qf(X_s)\mathrm{d}s, \qquad t \geqslant 0,$$

is a martingale with respect to $(\mathcal{F}_t^X)_{t \geqslant 0}$. Then X is the unique (in law) Markov process with initial distribution μ and Q-matrix q.

Proof. If X solves the martingale problem, then by Itô's formula we get for $\alpha > 0$

$$e^{-\alpha t} f(X_t) - f(X_0) = \int_0^t f(X_s)(-\alpha e^{-\alpha s})\mathrm{d}s + \int_0^t e^{-\alpha s} qf(X_s)\mathrm{d}s + \int_0^t e^{-\alpha s}\mathrm{d}M_s^f, \tag{2.3}$$

and therefore

$$M_t^{f,\alpha} := e^{-\alpha t} f(X_t) - f(X_0) + \int_0^t e^{-\alpha s}(\alpha - q) f(X_s)\mathrm{d}s = \int_0^t e^{-\alpha s}\mathrm{d}M_s^f$$

is a martingale. Indeed, we can simply define $\int_0^t e^{-\alpha s}\mathrm{d}M_s^f$ by formally applying the integration by parts rule and using that $e^{-\alpha\cdot}$ is of finite variation:

$$\int_0^t e^{-\alpha s}\mathrm{d}M_s^f := \left(\int_0^t e^{-\alpha s}\mathrm{d}s\right) M_t^f + \int_0^t M_s^f \alpha e^{-\alpha s}\mathrm{d}s, \tag{2.4}$$

so that (2.3) follows from the integration by parts rule for absolutely continuous functions. Since also $|M_t^f| \leqslant 2\|f\|_\infty + t2\|c\|_\infty\|f\|_\infty$, we can exchange the conditional expectation with the time integral in (2.4) to show that $M^{f,\alpha}$ is a martingale.

Using the martingale property of $M^{f,\alpha}$, we have

$$0 = \mathbb{E}[M_{s+t}^{f,\alpha} - M_s^{f,\alpha} | \mathcal{F}_s^X] = e^{-\alpha(s+t)}\mathbb{E}[f(X_{s+t})|\mathcal{F}_s^X] - e^{-\alpha s} f(X_s) \\ + \mathbb{E}\left[\int_s^{s+t} e^{-\alpha r}(\alpha - q) f(X_r)\mathrm{d}r \middle| \mathcal{F}_s^X\right],$$

and after setting $f = R_\alpha g$ and multiplying both sides with $e^{\alpha s}$ and sending $t \to \infty$ we get

$$R_\alpha g(X_s) = \mathbb{E}\left[\int_0^\infty e^{-\alpha r} g(X_{s+r})\mathrm{d}r \middle| \mathcal{F}_s^X\right] = \int_0^\infty e^{-\alpha r}\mathbb{E}[g(X_{s+r})|\mathcal{F}_s^X]\mathrm{d}r.$$

On the other hand, the explicit construction of R_α that we gave in Lemma 2.18 shows that

$$R_\alpha g(X_s) = \int_0^\infty e^{-\alpha r} p_r g(X_s)\mathrm{d}r.$$

Since $r \mapsto p_r g(X_s)$ and $r \mapsto \mathbb{E}[g(X_{s+r})|\mathcal{F}_s^X]$ are càdlàg (in the second case because of the dominated convergence theorem) and the identity holds for all $\alpha > 0$, the uniqueness of the Laplace transform gives

$$p_r g(X_s) = \mathbb{E}[g(X_{s+r})|\mathcal{F}_s^X]$$

for all $r \geqslant 0$, which is the Markov property and which also shows that X has the transition function (p_t). □

Exercise 2.5. Let $(X_t)_{t\geqslant 0}$ be a Markov process with Q-matrix q with $\|c\|_\infty < \infty$ and let $\lambda > 0$. Show that $\left(X_t^{(\lambda)} = X_{\lambda t}\right)_{t\geqslant 0}$ is a Markov process with Q-matrix λq.

Solution. We know that

$$f(X_{\lambda t}) - f(X_0) - \int_0^{\lambda t} q f(X_s)\mathrm{d}s = f\big(X_t^{(\lambda)}\big) - f\big(X_0^{(\lambda)}\big) - \int_0^t \lambda q f\big(X_s^{(\lambda)}\big)\mathrm{d}s$$

is a martingale in the time changed filtration $(\mathcal{F}_{\lambda t}^X)_{t\geqslant 0} = \big(\mathcal{F}_t^{X^{(\lambda)}}\big)_{t\geqslant 0}$, so the claim follows directly from Theorem 2.19. Alternatively simply differentiate $p_{t\lambda}$ in $t = 0$.

We conclude this section by giving a nice expression for qf: Since $\sum_y q(x,y) = 0$ we have

$$qf(x) = \sum_{y\in S} q(x,y) f(y) = \sum_{y\in S} q(x,y)(f(y) - f(x)).$$

Exercise 2.6. The martingale problem also makes sense for processes in uncountable state spaces:

i. Let $(X_t)_{t\geqslant 0}$ be a continuous stochastic process with values in $\mathbb{R}$ and with $X_0 = 0$. Show that X is a Brownian motion if and only if for all $f \in C^2$ with polynomial growth of f, f', f'', the process

$$f(X_t) - f(x) - \frac{1}{2}\int_0^t f''(X_s)\mathrm{d}s$$

is a martingale.

ii. (Voluntary exercise, difficult):

Let $(X_t)_{t\geqslant 0}$ be a continuous stochastic process with values in $\mathbb{R}$ and with $X_0 = x \in \mathbb{R}$. Let σ, b: $\mathbb{R} \to \mathbb{R}$ be Lipschitz continuous functions and assume also that $\inf_{x\in\mathbb{R}}\sigma(x) > 0$. Show that there exists an $(\mathcal{F}_t^X)$-Brownian motion B with

$$X_t = x + \int_0^t b(X_s)\mathrm{d}s + \int_0^t \sigma(X_s)\mathrm{d}B_s$$

if and only if for all $f \in C^2$ with polynomial growth of f, f', f'', the process

$$f(X_t) - f(x) - \int_0^t \left(b(X_s)f'(X_s) + \frac{1}{2}\sigma^2(X_s)f''(X_s)\right)\mathrm{d}s$$

is a martingale.

Solution. *Of course i. is a special case of ii. (with $b = 0$ and $\sigma = 1$), but let us still provide the arguments for i. and ii. separately:*

i. By taking $f(x) = x$ we see that X is a continuous martingale, and by taking $f(x) = x^2$ we obtain $\langle X\rangle_t = t$. By Lévy's characterization X is a Brownian motion.

If conversely X is a Brownian motion, then by Itô's formula

$$M_t^f = f(X_t) - f(x) - \frac{1}{2}\int_0^t f''(X_s)\mathrm{d}s = \int_0^t f'(X_s)\mathrm{d}X_s$$

is a local martingale for every $f \in C^2$. If f' has polynomial growth, then for some $C, p > 0$

$$\mathbb{E}\left[\int_0^t f'(X_s)^2\mathrm{d}s\right] \leqslant tC \sup_{s\in[0,t]} \mathbb{E}[(1 + |X_s|^p)] < \infty,$$

and therefore M^f is a true martingale.

ii. Let $M_t = X_t - \int_0^t b(X_s)\mathrm{d}s$ and $N_t^2 = X_t^2 - \int_0^t (2b(X_s)X_s + \sigma^2(X_s))\mathrm{d}s$ which are both martingales. Let us compute the quadratic variation of M. We have

$$\begin{aligned}
M_t^2 &= X_t^2 + \left(\int_0^t b(X_s)\mathrm{d}s\right)^2 - 2X_t\int_0^t b(X_s)\mathrm{d}s\\
&= N_t + \int_0^t (2b(X_s)X_s + \sigma^2(X_s))\mathrm{d}s + \left(\int_0^t b(X_s)\mathrm{d}s\right)^2 - 2\left[\int_0^t X_s b(X_s)\mathrm{d}s + \int_0^t\int_0^s b(X_r)\mathrm{d}r\mathrm{d}X_s\right]\\
&= N_t + \int_0^t \sigma^2(X_s)\mathrm{d}s + \left(\int_0^t b(X_s)\mathrm{d}s\right)^2 - 2\int_0^t\int_0^s b(X_r)\mathrm{d}r\, b(X_s)\mathrm{d}s + \int_0^t\int_0^s b(X_r)\mathrm{d}r\mathrm{d}M_s\\
&= N_t + \int_0^t\int_0^s b(X_r)\mathrm{d}r\mathrm{d}M_s + \int_0^t \sigma^2(X_s)\mathrm{d}s.
\end{aligned}$$

Since the first two terms on the right hand side are local martingales, we have

$$\langle M\rangle_t = \int_0^t \sigma^2(X_s)\mathrm{d}s.$$

Define now

$$B_t = \int_0^t \frac{1}{\sigma(X_s)}\mathrm{d}M_s,$$

which is a continuous local martingale (recall that $\sigma \geqslant c > 0$) with quadratic variation

$$\langle B\rangle_t = \int_0^t \frac{1}{\sigma^2(X_s)}\sigma^2(X_s)\mathrm{d}s = t,$$

and therefore B is a Brownian motion. Moreover, by definition we have

$$X_t = x + \int_0^t b(X_s)\mathrm{d}s + (M_t - x) = x + \int_0^t b(X_s)\mathrm{d}s + \int_0^t \sigma(X_s)\mathrm{d}B_s,$$

which is the claim.

Conversely, if X solves the stochastic SDE for some Brownian motion, then by Itô's formula

$$f(X_t) - f(x) - \int_0^t \left[b(X_s) f'(X_s) + \frac{1}{2}\sigma^2(X_s) f''(X_s) \right] \mathrm{d}s = \int_0^t \sigma(X_s) f'(X_s) \mathrm{d}B_s$$

is a local martingale, and we need to check that it is a real martingale. Since $|\sigma(y) f'(y)| \leqslant C(1+|y|^p)$ for some $p \in \mathbb{N}$, it suffices to show that X has locally bounded moments. But for $p \geqslant 2$ we have by the Burkholder-Davis-Gundy inequality and Jensen's inequality

$$\begin{aligned}\mathbb{E}[|X_t|^p] &\lesssim |x|^p + \mathbb{E}\left[\left| \int_0^t b(X_s) \mathrm{d}s \right|^p \right] + \mathbb{E}\left[\left| \int_0^t \sigma(X_s) \mathrm{d}B_s \right|^p \right] \\ &\lesssim |x|^p + t^{p-1} \mathbb{E}\left[\int_0^t |b(X_s)|^p \mathrm{d}s \right] + \mathbb{E}\left[\left| \int_0^t \sigma^2(X_s) \mathrm{d}s \right|^{p/2} \right] \\ &\lesssim |x|^p + t^{p-1} \int_0^t (1 + \mathbb{E}[|X_s|^p]) \mathrm{d}s + t^{p/2-1} \int_0^t \mathbb{E}[|\sigma(X_s)|^p] \mathrm{d}s \\ &\lesssim |x|^p + (t^{p-1} + t^{p/2-1}) \int_0^t (1 + \mathbb{E}[|X_s|^p]) \mathrm{d}s.\end{aligned}$$

So by Gronwall's lemma $\mathbb{E}[|X_t|^p] \leqslant C(T,p)$ for all $t \in [0,T]$, which completes the proof.

2.5 Stationary measures

Throughout this section we fix a Q-matrix q with associated transition function (p_t).

Definition 2.20. *A measure μ on S is called stationary or invariant for the transition function $(p_t)_{t \geqslant 0}$ if*

$$\mu p_t(x) = \sum_{y \in S} \mu(y) p_t(y,x) = \mu(x)$$

for all $x \in S$. We also write $\mu p_t = \mu$. An equivalent characterization is that for all $f \in \mathcal{M}(S)$

$$\mu(p_t f) = \mu f.$$

Lemma 2.21. *Let μ be a probability distribution and let $(X_t)_{t \geqslant 0}$ be the Markov process with initial distribution $X_0 \sim \mu$ and with transition function (p_t). Then μ is a stationary measure for (p_t) if and only if X is stationary, i.e. for all $s \geqslant 0$ the process*

$$Y_t = X_{s+t}, \qquad t \geqslant 0,$$

has the same law as X. Therefore, we also call μ a stationary (or invariant) measure for X.

Proof. We have for bounded and measurable $F: S^{\mathbb{R}_+} \to \mathbb{R}$

$$\mathbb{E}[F(Y)] = \mathbb{E}[\mathbb{E}[F((X_{t+s})_{t \geqslant 0}) | \mathcal{F}_s^X]] = \mathbb{E}[\mathbb{E}_{X_s}[F]] = \sum_{x_0, x_1 \in S} \mu(x_0) p_s(x_0, x_1) \mathbb{E}_{x_1}[F].$$

If μ is stationary for (p_t), then the summation over x_0 just gives $\mu(x_1)$, and therefore we get

$$\mathbb{E}[F(Y)] = \sum_{x_1 \in S} \mu(x_1) \mathbb{E}_{x_1}[F] = \mathbb{E}_\mu[F] = \mathbb{E}[F(X)].$$

If X is stationary, then we know that $\mathbb{E}[F(Y)] = \mathbb{E}[F(X)]$ and therefore

$$\sum_{x_0, x_1 \in S} \mu(x_0) p_s(x_0, x_1) \mathbb{E}_{x_1}[F] = \sum_{x_1 \in S} \mu(x_1) \mathbb{E}_{x_1}[F].$$

Now let $x \in S$ and set $F((\omega(t))_{t \geqslant 0}) = \mathbb{I}_{\omega(0) = x}$, so $\mathbb{E}_{x_1}[F] = \delta_{x, x_1}$. Then we get

$$\sum_{x_0} \mu(x_0) p_s(x_0, x) = \mu(x),$$

and therefore μ is stationary for (p_t). □

Lemma 2.22. *Assume that $\sum_x c(x)\mu(x) < \infty$. Then μ is stationary for the transition function (p_t) associated to q if and only if $\mu q = 0$, i.e. if*

$$\sum_{y\in S} \mu(y)q(y,x) = 0 \tag{2.5}$$

for all $x \in S$.

Proof. If μ is stationary for (p_t), then

$$0 = \left(\sum_{y\in S} \mu(y)p_t(y,x)\right) - \mu(x) = \sum_{y\in S} \mu(y)(p_t(y,x) - \delta_{x,y}),$$

and after dividing the right hand side by t we get the term $t^{-1}(p_t(y,x) - \delta_{x,y})$ which converges to $q(y,x)$ and is bounded by $c(y)$ whenever $\mu(y) > 0$, see (the proof of) Theorem 2.6. Therefore, the dominated convergence theorem applies and we obtain that $\sum_y \mu(y)q(y,x) = 0$.

If conversely $\sum_y \mu(y)q(y,x) = 0$, then

$$\sum_{y\in S} \mu(y)\frac{p_{t+h}(y,x) - p_t(y,x)}{h} = \sum_{y\in S} \mu(y)\frac{1}{h}\int_t^{t+h} qp_s(y,x)\mathrm{d}s,$$

and by the dominated convergence theorem and Fubini's theorem we conclude that

$$\partial_t\left(\sum_{y\in S} \mu(y)p_t(y,x)\right) = \sum_{y\in S} \mu(y)\sum_{z\in S} q(y,z)p_t(z,x) = \sum_{z\in S}\left(\sum_{y\in S} \mu(y)q(y,z)\right)p_t(z,x) = 0,$$

and therefore

$$\sum_{y\in S} \mu(y)p_t(y,x) = \sum_{y\in S} \mu(y)p_0(y,x) = \mu(x),$$

which shows that μ is stationary. □

Definition 2.23. *A measure μ on S is called reversible for the transition function $(p_t)_{t\geqslant 0}$ if the detailed balance condition*

$$\mu(x)p_t(x,y) = \mu(y)p_t(y,x)$$

holds for all $x, y \in S$. Summing both sides in y shows that every reversible measure is stationary.

As the name suggests, if we start our Markov process in the reversible probability measure μ, then the time reversed process has the same distribution:

Lemma 2.24. *Let μ be a probability distribution and let $(X_t)_{t\geqslant 0}$ be the Markov process with initial distribution $X_0 \sim \mu$ and with transition function (p_t). Then μ is a reversible measure for (p_t) if and only if X is reversible, i.e. for all $T \geqslant 0$ the process*

$$Y_t = X_{T-t}, \qquad t \in [0,T],$$

has the same law as $(X_t)_{t\in[0,T]}$. Therefore, we also call μ a reversible measure for X.

Proof. Assume first that μ is reversible for (p_t). Let $f\colon S^n \to \mathbb{R}$ be a bounded function and let $0 \leqslant t_1 \leqslant \ldots \leqslant t_n \leqslant T$ for some $T > 0$. Applying the stationarity of μ in the first step, and then repeatedly the reversibility gives

$$\begin{aligned}
&\mathbb{E}[f(X_{T-t_1}, \ldots, X_{T-t_n})] = \\
&\sum_{x_0,\ldots,x_n\in S} f(x_n,\ldots,x_1)\mu(x_0)p_{T-t_n}(x_0,x_1)p_{t_n-t_{n-1}}(x_1,x_2)\ldots p_{t_2-t_1}(x_{n-1},x_n)\\
&= \sum_{x_1,\ldots,x_n\in S} f(x_n,\ldots,x_1)\mu(x_1)p_{t_n-t_{n-1}}(x_1,x_2)\ldots p_{t_2-t_1}(x_{n-1},x_n)\\
&= \sum_{x_1,\ldots,x_n\in S} f(x_n,\ldots,x_1)p_{t_n-t_{n-1}}(x_2,x_1)\mu(x_2)p_{t_{n-1}-t_{n-2}}(x_2,x_2)\ldots p_{t_2-t_1}(x_{n-1},x_n)\\
&= \ldots = \sum_{x_1,\ldots,x_n\in S} f(x_n,\ldots,x_1)p_{t_n-t_{n-1}}(x_2,x_1)p_{t_{n-1}-t_{n-2}}(x_3,x_2)\ldots p_{t_2-t_1}(x_n,x_{n-1})\mu(x_n).
\end{aligned}$$

Now rename $y_k = x_{n-k+1}$ to obtain

$$= \sum_{y_1, \ldots, y_n \in S} f(y_1, \ldots, y_n) p_{t_n - t_{n-1}}(y_{n-1}, y_n) p_{t_{n-1} - t_{n-2}}(y_{n-2}, y_{n-1}) \ldots p_{t_2 - t_1}(y_1, y_2) \mu(y_1)$$
$$= \mathbb{E}[f(X_{t_1}, \ldots, X_{t_n})].$$

If conversely $(X_t)_{t \geqslant 0}$ is reversible, then we get by reversing time at $T = t$:

$$\mu(x) p_t(x, y) = \mathbb{P}(X_0 = x, X_t = y) = \mathbb{P}(X_t = x, X_0 = y) = \mu(y) p_t(y, x),$$

which is the detailed balance condition. □

Lemma. *(Not shown in lecture)*
If μ is a probability measure, then it is reversible if and only if

$$\mu((p_t f)\, g) = \sum_{x \in S} \mu(x) p_t f(x) g(x) = \sum_{x \in S} \mu(x) f(x) p_t g(x) = \mu(f p_t g) \tag{2.6}$$

for all $f, g \in \mathcal{M}(S)$, i.e. if p_t is a symmetric operator in $L^2(\mu)$.

Proof. If μ is a probability measure, then everything is absolutely summable and we can apply Fubini's theorem at will. If μ is also reversible, then for $f, g \in \mathcal{M}(S)$

$$\mu((p_t f)\, g) = \sum_{x \in S} \mu(x) \left(\sum_{y \in S} p_t(x, y) f(y) \right) g(x) = \sum_{x, y \in S} \mu(y) p_t(y, x) f(y) g(x)$$
$$= \sum_{x, y \in S} \mu(x) f(x) p_t(x, y) g(y) = \mu(f p_t g).$$

If μ satisfies (2.6), fix $x, y \in S$ and set $f(z) = \delta_{y,z}$ and $g(z) = \delta_{x,z}$, for which

$$\mu(x) p_t(x, y) = \sum_{z \in S} \mu(z) p_t(z, y) \delta_{x,z} = \mu((p_t f)\, g) = \mu(f p_t g) = \sum_{z \in S} \mu(z) \delta_{y,z} p_t(z, x) = \mu(z) p_t(y, x).$$

□

Lemma 2.25. *Assume that $\|c\|_\infty + \|\mu\|_\infty < \infty$. Then μ is reversible for (p_t) if and only if*

$$\mu(x) q(x, y) = \mu(y) q(y, x)$$

for all $x, y \in S$.

Proof. Differentiating the equality $\mu(x) p_t(x, y) = \mu(y) p_t(y, x)$ in $t = 0$ gives

$$\mu(x) q(x, y) = \mu(y) q(y, x).$$

Conversely, if this equality holds and $\|c\|_\infty + \|\mu\|_\infty < \infty$, then

$$\mu(x) p_t(x, y) = \mu(x) e^{tq}(x, y) = \mu(x) \sum_{k=0}^{\infty} \frac{t^k}{k!} q^k(x, y)$$
$$= \mu(x) \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{z_1, \ldots, z_{k-1}} q(x, z_1) q(z_1, z_2) \ldots q(z_{k-1}, y)$$
$$= \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{z_1, \ldots, z_{k-1}} q(z_1, x) \mu(z_1) q(z_1, z_2) \ldots q(z_{k-1}, y)$$
$$= \ldots = \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{z_1, \ldots, z_{k-1}} q(z_1, x) q(z_2, z_1) \ldots q(y, z_{k-1}) \mu(y)$$
$$= \sum_{k=0}^{\infty} \frac{t^k}{k!} q^k(y, x) \mu(y) = p_t(y, x) \mu(y).$$

□

Reversibility is stronger than stationarity. Yet it is easier to verify if a given measure satisfies the detailed balance condition than whether it is stationary. But of course a reversible measure may not exist, even when there is a stationary measure.

Exercise 2.7. Consider a random walk on $\mathbb{Z}$ which moves up with rate $p\in(0,1)$ and moves down with rate $1-p$. Show that the counting measure $\mu(x)\equiv 1$ is stationary, but not reversible unless $p=\frac{1}{2}$. Deduce that for $p\neq\frac{1}{2}$ no reversible measure exists.

Solution. *The Q-matrix is*

$$q(x,y)=\begin{cases} p, & y=x+1,\\ -1, & y=x,\\ 1-p, & y=x-1,\end{cases}$$

and therefore the counting measure $\mu(x)\equiv 1$ is formally stationary:

$$\sum_x \mu(x)q(x,y)=\sum_x q(x,y)=q(y-1,y)+q(y,y)+q(y+1,y)=p-1+(1-p)=0.$$

But of course $\sum_x \mu(x)c(x)=\infty$ and therefore we cannot apply Lemma 2.22. To rigorously show the invariance it suffices to note that since $q(x,y)=q(x-z,y-z)$ for all $x,y,z\in\mathbb{Z}$ we also have $p_t(x,y)=p_t(0,y-x)$ for all x,y,z, and therefore

$$\sum_x \mu(x)p_t(x,y)=\sum_x p_t(0,y-x)=\sum_z p_t(0,z)=1.$$

But if $p\neq\frac{1}{2}$, then μ is not reversible: $\mu(x)q(x,x+1)=p$ but $\mu(x+1)q(x+1,x)=1-p$. This is also intuitively clear, because if say $p>\frac{1}{2}$, then the random walk has a tendency to move upwards, while the time reversed process would move more downwards. As Corollary 2.27 below shows, q admits only one stationary measure up to multiplication with a constant (since $p\neq 0,1$ the random walk is clearly irreducible), and therefore no reversible measure exists.

Remark 2.26. Recall that in the construction of Markov processes from a given a Q-matrix q we we first defined the transition matrix P of a discrete time Markov process, based on which we constructed the continuous time process with the help of a random time change.

i. A measure μ with $\sum_x \mu(x)c(x)<\infty$ is invariant for (p_t) if and only if

$$\nu(x)=\mu(x)c(x)$$

is invariant for the transition matrix P. Indeed, we have $\mu q=0$ if and only if for all $y\in S$

$$\mu(x)c(x)=\sum_{y\neq x}\mu(y)q(y,x)=\sum_{y\neq x}\mu(y)c(y)P(y,x)=\sum_{y\neq x}\nu(y)P(y,x)$$

Since the left hand side is simply $\nu(x)$, this is equivalent to $\nu=\nu P$.

ii. (Not in lecture)

Assume that $\|c\|_\infty<\infty$ and $\sum_x c(x)\mu(x)<\infty$. Then μ is reversible for (p_t) if and only if ν is reversible for P. Indeed, it suffices to note that

$$\mu(x)q(x,y)=\mu(x)c(x)P(x,y)=\nu(x)P(x,y).$$

Corollary 2.27. *We call a transition function irreducible if the associated discrete time transition matrix is irreducible, i.e. if for all $x,y\in S$ there exists $n\in\mathbb{N}$ with $P^n(x,y)>0$. By the preceding remark and Stochastic Processes I, any irreducible transition function has at most one stationary measure, up to multiplication with a constant.*

Exercise 2.8. Let q be a Q-matrix. Show that q is irreducible if and only if for all $x\neq y\in S$ there exist $n\in\mathbb{N}$ and $x_0,\ldots,x_n$ with $x_0=x$ and $x_n=y$, such that $q(x_k,x_{k+1})>0$ for all $k=0,\ldots,n-1$.

Solution. *We have*

$$P^n(x,y) = \sum_{x_1,\ldots,x_{n-1}} \frac{q(x,x_1)}{c(x)} \times \frac{q(x_1,x_2)}{c(x_1)} \times \ldots \times \frac{c(x_{n-1},x)}{c(x_{n-1})},$$

and the sum is strictly positive if and only each one of its addends is strictly positive.

3 Interacting particle systems

In this section we mainly follow [Kipnis, Landim, *Scaling Limits of Interacting Particle Systems,* Springer, 1999] and [Sethuraman, *Large scale stochastic dynamics*, University of Arizona lecture notes, 2012].

3.1 First examples

Let $G=(V,E)$ be a finite graph with vertex set V and edge set E. We will write x,y,z for elements of V, and given $x,y\in V$ we write $x\sim y$ if $\{x,y\}\in E$. Let T be a countable set. Then we write

$$S := T^V := \{\eta | \eta\colon V \to T\}$$

for the configuration space. So $\eta\in S$ encodes a configuration in T for very vertex $v\in V$. We will write η or ζ for elements of S. Usually we will take

$$T=\{0,1\} \qquad \text{or} \qquad T=\mathbb{N}_0.$$

An interacting particle system is simply a Markov process that describes an evolution in the configuration space.

Example 3.1. (Random walk)

Let $p\colon \mathbb{Z}^d \to [0,1]$ be the step probability of a random walk, that is $\sum_{z\in\mathbb{Z}^d} p(z) = 1$. We can associate periodic transition probabilities on $\mathbb{Z}_N^d = (\mathbb{Z}/(N\mathbb{Z}))^d$ to p by setting

$$p^N(x,y) := \sum_{k\in\mathbb{Z}^d} p(y+kN-x).$$

Then we can construct a periodic continuous time random walk $(X_t)_{t\geqslant 0}$ with the following Q-matrix:

$$q^N(x,y) := \begin{cases} p^N(x,y), & y\neq x,\\ -\sum_{y\neq x} p^N(x,y), & y=x. \end{cases}$$

Taking $T=\{0,1\}$ and $V=\mathbb{Z}_N^d$ we associate an "interacting" particle system to $(X_t)_{t\geqslant 0}$ by setting

$$\eta_t(x) := \mathbb{I}_{X_t}(x) := \begin{cases} 1, & X_t = x,\\ 0, & X_t\neq x. \end{cases}$$

Of course, η lives only in a small subspace of $S=\{0,1\}^{\mathbb{Z}_N^d}$ where every η satisfies $\eta(x)=1$ for exactly one $x\in\mathbb{Z}_N^d$. Also, there is only one particle, and it does not "interact" with anything. The Q-matrix of the new Markov process $(\eta_t)_{t\geqslant 0}$ is more easily described through the action on a test function $f\colon S\to\mathbb{R}$, but of course we can recover the matrix from this by taking $f(\zeta)=\mathbb{I}_{\zeta=\zeta'}$:

$$qf(\eta) = \sum_{x,y\in\mathbb{Z}_N^d} \eta(x) p^N(x,y)(f(\mathbb{I}_y) - f(\eta)).$$

Convince yourself of this! Note that the sum in x collapses to one single term.

Example 3.2. (Independent random walks)

In the same setting as in Example 3.1 we now take several independent copies of the same random walk, say X^1, ..., X^m. Each X^k describes the motion of a particle, and we assume the particles to be indistinguishable. Therefore we do not care about the position of each individual particle, but only about the total number of particles in each site, and we define $\eta_t \in S := \mathbb{N}_0^{\mathbb{Z}_N^d}$,

$$\eta_t(x) = \sum_{k=1}^{m} \mathbb{I}_{X_t^k}(x) = \sum_{k=1}^{m} \eta_t^k(x),$$

as the number of particles at time t in the point $x \in \mathbb{Z}_N^d$, where η^k is defined from X^k as in Example 3.1. A priori it is not obvious that $(\eta_t)_{t \geqslant 0}$ is still a Markov process, because we cannot reconstruct $(X^1, ..., X^m)$ from η: after all we got rid of the labels $1, ..., m$. But since the particles all follow the same dynamics, independently of each other, there is no useful information in the labels and (η_t) is indeed a Markov process. A formal proof follows in the remark below.

Let us compute the Q-matrix of (η_t): Let $\eta \in S$ be a configuration with $\sum_{x \in \mathbb{Z}_N^d} \eta(x) = m$, let $f: S \to \mathbb{R}$ be bounded. We claim that

$$q f(\eta) = \sum_{x, y \in \mathbb{Z}_N^d} \eta(x) p^N(x, y)(f(\eta^{x, y}) - f(\eta)), \tag{3.1}$$

where

$$\eta^{x, y} := \eta + \mathbb{I}_y - \mathbb{I}_x$$

is the configuration with one particle moved from x to y. Note that the rate is multiplied with $\eta(x)$, so whenever the rate $\eta(x) p^N(x, y)$ is non-zero the configuration $\eta^{x, y}$ has only positive occupation variables; we can therefore simply set $f(\eta^{x, y}) = 0$ whenever $\eta^{x, y}$ has negative occupation variables, this will not affect q. To prove the claimed form of q we use induction. For $m = 1$ this is just Example 3.1, so assume (3.1) holds for m and let us try to establish it for $m + 1$. Let $(X^1, ..., X^{m+1})$ be the random walks, and define $\eta_t^1 := \sum_{k=1}^m \mathbb{I}_{X_t^k}$ and $\eta_t^2 = \mathbb{I}_{X_t^{m+1}}$. Then η^1 and η^2 are independent Markov processes with Q-matrices of the form (3.1), q^1 for η^1 and q^2 for η^2, and we have $\eta = \eta^1 + \eta^2$. Moreover, let $\eta_0 = \eta$ and note that

$$\begin{aligned} \mathbb{E}[f(\eta_t)] = \mathbb{E}[f(\eta_t^1 + \eta_t^2)] &= \mathbb{E}[f(\eta_t^1 + \eta_t^2)\mathbb{I}_{\eta_t^1 = \eta_0^1}] + \mathbb{E}[f(\eta_t^1 + \eta_t^2)\mathbb{I}_{\eta_t^2 = \eta_0^2}] \\ &- \mathbb{E}[f(\eta_t^1 + \eta_t^2)\mathbb{I}_{\eta_t^1 = \eta_0^1, \eta_t^2 = \eta_0^2}] + \mathbb{E}[f(\eta_t^1 + \eta_t^2)\mathbb{I}_{\eta_t^1 \neq \eta_t^1, \eta_t^2 \neq \eta_0^2}], \end{aligned}$$

and the last term on the right hand side is bounded by

$$\begin{aligned} &|\mathbb{E}[f(\eta_t^1 + \eta_t^2)\mathbb{I}_{\eta_t^1 \neq \eta_0^1, \eta_t^2 \neq \eta_0^2}]| \leqslant \|f\|_\infty \mathbb{P}(\eta_t^1 \neq \eta_0^1)\mathbb{P}(\eta_t^2 \neq \eta_0^2) \\ &= \|f\|_\infty \left(t \sum_{\zeta \neq \eta_0^1} q^1(\eta_0^1, \zeta) + o(t)\right)\left(t \sum_{\zeta \neq \eta_0^2} q^2(\eta_0^2, \zeta) + o(t)\right) = o(t). \end{aligned}$$

The remaining terms are

$$\begin{aligned} &\mathbb{E}[f(\eta_t^1 + \eta_t^2)\mathbb{I}_{\eta_t^1 = \eta_0^1}] + \mathbb{E}[f(\eta_t^1 + \eta_t^2)\mathbb{I}_{\eta_t^2 = \eta_0^2}] - \mathbb{E}[f(\eta_t^1 + \eta_t^2)\mathbb{I}_{\eta_t^1 = \eta_0^1, \eta_t^2 = \eta_0^2}] \\ &= \mathbb{E}[f(\eta_0^1 + \eta_t^2)]\mathbb{P}(\eta_t^1 = \eta_0^1) + \mathbb{E}[f(\eta_t^1 + \eta_0^2)\mathbb{P}(\eta_t^2 = \eta_0^2)] - f(\eta)\mathbb{P}(\eta_t^1 = \eta_0^1)\mathbb{P}(\eta_t^2 = \eta_0^2) \\ &= (f(\eta) + t q^1 f(\eta_0^1 + \eta_0^2) + o(t))(1 + t q^1(\eta_0^1, \eta_0^1) + o(t)) \\ &+ (f(\eta) + t q^2 f(\eta_0^1 + \eta_0^2) + o(t))(1 + t q^2(\eta_0^2, \eta_0^2) + o(t)) \\ &- f(\eta)(1 + t q^1(\eta_0^1, \eta_0^1) + o(t))(1 + t q^2(\eta_0^2, \eta_0^2) + o(t)) \\ &= f(\eta)(1 + t q^1(\eta_0^1, \eta_0^1)) + t q^1 f(\eta_0^1 + \eta_0^2) + o(t) \\ &+ f(\eta)(1 + t q^2(\eta_0^2, \eta_0^2)) + t q^2 f(\eta_0^1 + \eta_0^2) + o(t) \\ &- f(\eta)(1 + t q^1(\eta_0^1, \eta_0^1) + t q^2(\eta_0^2, \eta_0^2)) + o(t) \\ &= f(\eta) + t q^1 f(\eta) + t q^2 f(\eta) + o(t) \end{aligned}$$

Now $f(\eta_0^1 + \mathbb{I}_y - \mathbb{I}_x + \eta_0^2) = f(\eta_0^2 + \mathbb{I}_y - \mathbb{I}_x + \eta_0^1)$, and therefore

$$\begin{aligned} q^1 f(\eta) + q^2 f(\eta) &= \sum_{x, y \in \mathbb{Z}_N^d} [\eta_0^1(x) + \eta_0^2(x)] p^N(x, y)(f(\eta_0^1 + \mathbb{I}_y - \mathbb{I}_x + \eta_0^2) - f(\eta)) \\ &= \sum_{x, y \in \mathbb{Z}_N^d} \eta(x) p^N(x, y)(f(\eta^{x, y}) - f(\eta)), \end{aligned}$$

so that finally

$$\mathbb{E}[f(\eta_t)] = f(\eta) + t \sum_{x,y \in \mathbb{Z}_N^d} \eta(x) p^N(x,y)(f(\eta^{x,y}) - f(\eta)) + o(t),$$

which proves the claimed form (3.1) of the Q-matrix for $(\eta_t)_{t \geqslant 0}$.

Remark. The process $(\eta_t)_{t \geqslant 0}$ from the previous example is indeed a Markov process:

The Markov property of the random walks gives

$$\mathbb{E}[f(\eta_t)|\mathcal{F}_s^{\eta}] = \mathbb{E}\Big[\mathbb{E}\big[f(\eta_t)|\mathcal{F}_s^{(X^1,\ldots,X^k)}\big]\Big|\mathcal{F}_s^{\eta}\Big] = \mathbb{E}[\mathbb{E}[f(\eta_t)|(X_s^1,\ldots,X_s^k)]|\mathcal{F}_s^{\eta}],$$

so the Markov property of (η_t) follows once we show that $\mathbb{E}[f(\eta_t)|(X_s^1,\ldots,X_s^k)] = \mathbb{E}[f(\eta_t)|\eta_s]$. For that purpose let $(x_1,\ldots,x_m)$ and $(\tilde{x}_1,\ldots,\tilde{x}_m)$ be two configurations of the random walks with

$$\eta^{(x_1,\ldots,x_m)} := \sum_{k=1}^{m} \mathbb{I}_{x_k} = \sum_{k=1}^{m} \mathbb{I}_{\tilde{x}_k} = \eta^{(\tilde{x}_1,\ldots,\tilde{x}_m)}.$$

If we can show that

$$\mathbb{E}[f(\eta_t)|(X_s^1,\ldots,X_s^m) = (x_1,\ldots,x_m)] = \mathbb{E}[f(\eta_t)|(X_s^1,\ldots,X_s^m) = (\tilde{x}_1,\ldots,\tilde{x}_m)],$$

then $\mathbb{E}[f(\eta_t)|(X_s^1,\ldots,X_s^m)] = \mathbb{E}[f(\eta_t)|\eta_s]$. But since $\eta^{(x_1,\ldots,x_m)} = \eta^{(\tilde{x}_1,\ldots,\tilde{x}_m)}$, there exists a permutation σ of $\{1,\ldots,m\}$ such that $(\tilde{x}_1,\ldots,\tilde{x}_m) = (x_{\sigma(1)},\ldots,x_{\sigma(m)})$, and therefore

$$\begin{aligned}\mathbb{E}[f(\eta_t)|(X_s^1,\ldots,X_s^m) = (\tilde{x}_1,\ldots,\tilde{x}_m)] &= \mathbb{E}\big[f(\eta_t)|\big(X_s^{\sigma^{-1}(1)},\ldots,X_s^{\sigma^{-1}(m)}\big) = (x_1,\ldots,x_m)\big] \\ &= \mathbb{E}\big[f(\eta_t^{\sigma^{-1}})|\big(X_s^{\sigma^{-1}(1)},\ldots,X_s^{\sigma^{-1}(m)}\big) = (x_1,\ldots,x_m)\big],\end{aligned}$$

where we wrote $\eta_t^{\sigma^{-1}} := \sum_{k=1}^{m} \mathbb{I}_{X_t^{\sigma^{-1}(k)}}(x) = \eta_t$. Since $(X^{\sigma^{-1}(1)},\ldots,X^{\sigma^{-1}(m)})$ has the same distribution as $(X^1,\ldots,X^m)$, the right hand side is equal to $\mathbb{E}[f(\eta_t)|(X_s^1,\ldots,X_s^m) = (x_1,\ldots,x_m)]$, which proves the Markov property of $(\eta_t)_{t \geqslant 0}$.

Example 3.3. ((Simple) exclusion process)

Let us introduce some interaction among the random walks from Example 3.2: We consider again particles moving through $V = \mathbb{Z}_N^d$ that follow independent random walks, but now there is an exclusion rule that prevents each site from being occupied by more than one particle. Without interaction, the transition rate of particles from x to y was given by $\eta(x) p^N(x,y)$, but now we should set this to zero if there already is a particle in y, i.e. we change the Q-matrix to

$$\begin{aligned}qf(\eta) &= \sum_{x,y \in \mathbb{Z}_N^d} \eta(x) \mathbb{I}_{\eta(y)=0} p^N(x,y)(f(\eta^{x,y}) - f(\eta)) \\ &= \sum_{x,y \in \mathbb{Z}_N^d} \eta(x)(1 - \eta(y)) p^N(x,y)(f(\eta^{x,y}) - f(\eta)),\end{aligned}$$

where as before $\eta^{x,y} = \eta + \mathbb{I}_y - \mathbb{I}_x$ is the configuration η with one particle moved from x to y. Since now there is at most one particle per site we can actually take $T = \{0,1\}$ rather than $T = \mathbb{N}_0$, so $S = \{0,1\}^{\mathbb{Z}_N^d}$.

Example 3.4. (Zero range process)

We consider again particles in $V = \mathbb{Z}_N^d$ that follow the random walk with transition probabilities p^N as defined in Example 3.2. Now the particles interact by wanting to cluster together (or avoid each other). We encode this with the help of a function $g: \mathbb{N}_0 \to [0,\infty)$ with $g(0) = 0$ that describes the rate with which particles leave site x: each particle at site x leaves with rate $g(\eta(x)) / \eta(x)$. The name "zero range" comes from the fact that the motion of each particle only depends on the configuration at its own site (with "zero distance" from it), and not on the configuration around it. So the overall transition of particles from x to y has rate $\eta(x)(g(\eta(x)) / \eta(x)) p^N(x,y)$, and therefore the process has the Q-matrix

$$qf(\eta) = \sum_{x,y \in \mathbb{Z}_N^d} g(\eta(x)) p^N(x,y)(f(\eta^{x,y}) - f(\eta)).$$

Note that if $g(k)=k$, then the particles move independently and we recover Example 3.2. If $g(k)=\mathbb{I}_{k\geqslant 1}$, then we can imagine a queue at each site, where each particle is serviced with rate 1 and after being serviced it leaves to a random location according to the transition mechanism p^N, and newly arriving particles enter at the back of the queue. In the following discussions we will always assume that $g(k)>0$ for all $k\in\mathbb{N}\setminus\{0\}$ and that there exists $C>0$ with $|g(k+1)-g(k)|\leqslant C$ for all $k\in\mathbb{N}_0$.

Example 3.5. (Contact process)

We take again $T=\{0,1\}$, but now we consider an arbitrary finite graph (V,E). Now $\eta(x)=0$ means that the individual at site x is healthy, while $\eta(x)=1$ means the individual is sick. This time there is no motion of particles, but a disease spreads through the population, where neighbors can infect each other. An infected particle cures itself with rate 1, while a healthy particle is infected by each infected neighbor with rate $\lambda\geqslant 0$. Mathematically we encodes this as follows: We write

$$\eta^x(z):=\begin{cases} 1-\eta(x), & z=x,\\ \eta(z), & \text{else,}\end{cases}$$

for the configuration where the state in x is flipped and all other states remain the same. If $\eta(x)=1$, then η transitions to η^x with rate 1. If $\eta(x)=0$, then the transition to η^x happens with rate $\lambda\sum_{y\sim x}\eta(y)$ (i.e. λ times the number of infected neighbors). The q-matrix is

$$qf(\eta)=\sum_{x\in V}\left[\eta(x)+(1-\eta(x))\lambda\sum_{y\sim x}\eta(y)\right](f(\eta^x)-f(\eta)).$$

Example 3.6. (Voter model)

Again $T=\{0,1\}$ and (V,E) is a finite graph, and now $\eta(x)=0$ means that the individual x has opinion 0 while $\eta(x)=1$ means that individual x has opinion 1. Each particle adapts the opinions of its neighbor with rate 1, and therefore with the same η^x notation as in the previous example

$$qf(\eta)=\sum_{x\in V}\left[\sum_{y\sim x}\mathbb{I}_{\eta(x)\neq\eta(y)}\right](f(\eta^x)-f(\eta)).$$

Note that if there are k neighbors y that disagree with x, then x is changing its opinion with rate k. Even if there are $\ell\gg k$ neighbors that agree with x, their agreement does not have any effect towards x keeping its opinion.

All of the above models intuitively make sense also on infinite graphs, for example we could replace $\mathbb{Z}_N^d$ by $\mathbb{Z}^d$. The transition rates would be the same in this case, and for example we could simply sum over $x,y\in\mathbb{Z}^d$ in the q-matrix of the exclusion process,

$$\sum_{x,y\in\mathbb{Z}_N^d}\eta(x)(1-\eta(y))p^N(x,y)(f(\eta^{x,y})-f(\eta)),$$

and consider functions $f\colon\{0,1\}^{\mathbb{Z}^d}\to\mathbb{R}$. The difficulty is that then we have to worry about infinite sums, and more importantly also about uncountable state spaces, so things become much more technical. We therefore restrict our attention to interacting particle systems on finite graphs and refer to Chapter 4 of [Liggett, *Continuous Time Markov Processes*, AMS, 2010] or Chapter 1 of [Liggett, *Interacting Particle Systems*, Springer, 2004] for the construction of interacting particle systems on countable graphs.

3.2 Stationary measures

Let us derive some explicit stationary measures for the interacting particle systems we just constructed. Given a probability distribution μ on S and $f\in\mathcal{M}(S)$ we write

$$E_\mu[f]=\int_S f(\eta)\mu(\mathrm{d}\eta),$$

while the notation $\mathbb{E}_\mu$ will be reserved for the expectation with respect to the Markov process $(\eta_t)_{t\geqslant 0}$ with initial distribution μ (and the dynamics of (η_t), i.e. the Q-matrix, will be clear from the context).

Example 3.7. (Random walk)
The actual state space of the particle description of our random walk from Example 3.1 is not $\{0,1\}^{\mathbb{Z}_N^d}$ but

$$S=\left\{\eta\in\{0,1\}^{\mathbb{Z}_N^d}: \sum_x \eta(x)=1\right\}.$$

The uniform distribution μ on S is a stationary probability measure: By Lemma 2.22 it suffices to check that for all $\eta\in S$

$$E_\mu[q(\cdot,\eta)]=\sum_\zeta \mu(\zeta)q(\zeta,\eta)=\sum_y \mu(\mathbb{I}_y)q(\mathbb{I}_y,\eta)=0,$$

but for $\eta=\mathbb{I}_x$ and $\zeta=\mathbb{I}_y$ we have

$$q(\mathbb{I}_y,\mathbb{I}_x)=\begin{cases} p^N(y,x), & y\neq x,\\ -\sum_{z\neq x} p^N(x,z), & y=x,\end{cases}$$

and since $\mu(\mathbb{I}_y)=|S|^{-1}$ for every y we get

$$\sum_y \mu(\mathbb{I}_y)q(\mathbb{I}_y,\eta)=\frac{1}{|S|}\left(\sum_{y\neq x} p^N(y,x)-\sum_{z\neq x} p^N(x,z)\right).$$

But since $p^N(y,x)=p^N(x-y)$ only depends on $x-y$, we have

$$=\frac{1}{|S|}\left(\sum_{y\neq x} p^N(x-y)-\sum_{z\neq x} p^N(z-x)\right)=\frac{1}{|S|}\left(\sum_{y\neq 0} p^N(y)-\sum_{z\neq 0} p^N(z)\right)=0.$$

By Lemma 2.25 the uniform distribution is reversible if and only if $p^N(z)=p^N(-z)$ for all $z\in\mathbb{Z}_N^d$. If the random walk is irreducible, i.e. if it can reach every state in $\mathbb{Z}_N^d$ from 0 in finitely many steps, then by Corollary 2.27 there is a unique stationary probability measure.

In what follows we always assume that p is irreducible. This is for example the case if $p(z)>0$ for all $z\in\mathbb{Z}^d$ with $|z|=1$.

Example 3.8. (Zero range process)
Since the independent random walks of Example 3.2 are a special case of the zero range process of Example 3.4, we only consider the latter. Unlike the random walk in the last example, the zero range process is never irreducible because the number of particles is preserved by the dynamics. So there are infinitely many irreducibility classes, one for each given number of particles (because we assumed that the random walk itself is irreducible). It would therefore be natural to describe the unique invariant probability measure for each irreducibility class.

However, a different class of invariant measures where each measure assigns positive mass to all irreducibility classes is easier to work with. We define for $\alpha>0$ the formal power series

$$Z(\alpha)=\sum_{k=0}^\infty \frac{\alpha^k}{g(k)!},\qquad g(k)!=\prod_{j=1}^k g(j),\qquad g(0)!=1,$$

and we write $\alpha^*\geqslant 0$ for its convergence radius, which we assume to be strictly positive from now on. Note that $\alpha^*\geqslant \liminf_{k\to\infty} g(k)$ and therefore $\alpha^*>0$ whenever $\liminf_k g(k)>0$ (indeed, for $\alpha<\liminf_k g(k)$ we have $\frac{\alpha^k}{g(k)!}\leqslant C\beta^k$ for some $\beta<1$ and $C>0$ and all large k). For $\alpha<\alpha^*$ we have $Z(\alpha)<\infty$, and therefore we can define a probability measure m_α on $\mathbb{N}_0$ by

$$m_\alpha(k)=\frac{1}{Z(\alpha)}\frac{\alpha^k}{g(k)!},$$

and then $\tilde{\nu}_\alpha$ on $S = \mathbb{N}_0^{\mathbb{Z}_N^d}$ by

$$\tilde{\nu}_\alpha(\eta) = \prod_{x \in \mathbb{Z}_N^d} m_\alpha(\eta(x)) = \prod_{x \in \mathbb{Z}_N^d} \frac{1}{Z(\alpha)} \frac{\alpha^{\eta(x)}}{g(\eta(x))!}.$$

Lemma 3.9. *For $\alpha < \alpha^*$ the measure $\tilde{\nu}_\alpha$ is stationary for the zero range process.*

Proof. We want to apply Lemma 2.22. For that purpose we first need to show that $E_{\tilde{\nu}_\alpha}[c] < \infty$. Since

$$c(\eta) = \sum_{\zeta \neq \eta} q(\eta, \zeta) = \sum_{\zeta} q(\eta, \zeta)(\mathbb{I}_{\zeta \neq \eta} - \mathbb{I}_{\eta \neq \eta}) = (q\mathbb{I}_{\cdot \neq \eta})(\eta),$$

it suffices to show that $E_{\tilde{\nu}_\alpha}[|qf(\eta)|] < \infty$ for all $f \in \mathcal{M}(S)$, which is a more general statement that will be useful in what follows. By definition of $\tilde{\nu}_\alpha$ we have

$$E_{\tilde{\nu}_\alpha}[|qf(\eta)|] = \sum_{\eta \in S} \tilde{\nu}_\alpha(\eta)|qf(\eta)| \leqslant \sum_{x, y \in \mathbb{Z}_N^d} E_{\tilde{\nu}_\alpha}[g(\eta(x))] p^N(x, y) 2\|f\|_\infty$$
$$\lesssim \sum_{\eta \in S} \sum_{x \in \mathbb{Z}_N^d} \tilde{\nu}_\alpha(\eta) g(\eta(x)) = \sum_{x \in \mathbb{Z}_N^d} \sum_{\eta \in S} \tilde{\nu}_\alpha(\eta - \mathbb{I}_x) \alpha \mathbb{I}_{\eta(x) > 0} \leqslant \sum_{x \in \mathbb{Z}_N^d} \alpha < \infty.$$

This will allow us to apply Fubini's theorem in the following computations. We have

$$E_{\tilde{\nu}_\alpha}[qf] = \sum_{x, y \in \mathbb{Z}_N^d} E_{\tilde{\nu}_\alpha}[g(\eta(x))(f(\eta^{x,y}) - f(\eta))] p^N(x, y),$$

and since

$$\tilde{\nu}_\alpha(\eta^{y,x}) = \frac{g(\eta(y))}{g(\eta(x) + 1)} \tilde{\nu}_\alpha(\eta),$$

the change of variables $\eta^{x,y} \to \eta$ leads to

$$E_{\tilde{\nu}_\alpha}[g(\eta(x)) f(\eta^{x,y})] = \sum_{\eta \in S} \tilde{\nu}_\alpha(\eta) g(\eta(x)) f(\eta^{x,y}) = \sum_{\eta \in S} \mathbb{I}_{\eta^{y,x} \in S} \tilde{\nu}_\alpha(\eta^{y,x}) g(\eta^{y,x}(x)) f(\eta)$$
$$= \sum_{\eta \in S} \mathbb{I}_{\eta^{y,x} \in S} \frac{g(\eta(y))}{g(\eta(x) + 1)} \tilde{\nu}_\alpha(\eta) g(\eta(x) + 1) f(\eta) = E_{\tilde{\nu}_\alpha}[g(\eta(y) f(\eta))].$$

As in Example 3.7 we have $\sum_x p^N(x, y) = \sum_z p^N(z) = 1 = \sum_y p^N(x, y)$, and therefore

$$E_{\tilde{\nu}_\alpha}[qf] = \sum_{y \in \mathbb{Z}_N^d} E_{\tilde{\nu}_\alpha}[g(\eta(y)) f(\eta)] - \sum_{x \in \mathbb{Z}_N^d} E_{\tilde{\nu}_\alpha}[g(\eta(x)) f(\eta)] = 0.$$

Therefore, all conditions in Lemma 2.22 are satisfied and $\tilde{\nu}_\alpha$ is stationary. □

Corollary 3.10. *In the example of the independent random walks, Example 3.2, for every $\alpha > 0$ the product Poisson distribution with parameter α,*

$$\tilde{\nu}_\alpha(\eta) = \prod_{x \in \mathbb{Z}_N^d} e^{-\alpha} \frac{\alpha^{\eta(x)}}{\eta(x)!}$$

is invariant.

Exercise 3.1. Convince yourself that the condition $\tilde{\nu}_\alpha(\eta) q(\eta, \zeta) = \tilde{\nu}_\alpha(\zeta) q(\zeta, \eta)$ of Lemma 2.25 is satisfied if and only if $p^N(z) = p^N(-z)$ for all $z \in \mathbb{Z}_N^d$. This suggests that for such p^N the zero range process is reversible under ν_α (which is true), but of course the condition $\|c\|_\infty < \infty$ of Lemma 2.25 is in general not satisfied.

Solution. *For $\zeta \neq \eta$ we have $q(\eta, \zeta) = 0$ unless $\zeta = \eta^{x,y}$ for some $x, y \in \mathbb{Z}_N^d$ with $\eta(x) > 0$. In that case*

$$\tilde{\nu}_\alpha(\eta) q(\eta, \zeta) = \prod_{z \in \mathbb{Z}_N^d} \frac{1}{Z(\alpha)} \frac{\alpha^{\eta(z)}}{g(\eta(z))!} g(\eta(x)) p^N(x, y)$$
$$= \tilde{\nu}_\alpha(\eta^{x,y}) g(\eta(y) + 1) p^N(x, y) = \tilde{\nu}_\alpha(\eta^{x,y}) q(\eta^{x,y}, \eta) \frac{p^N(x, y)}{p^N(y, x)},$$

which proves the claim.

The measure m_α has the expected value

$$R(\alpha)=\sum_{k=0}^{\infty} k m_\alpha(k)=\frac{1}{Z(\alpha)}\sum_{k=0}^{\infty} k\frac{\alpha^k}{g(k)!}=\frac{\alpha}{Z(\alpha)}\sum_{k=0}^{\infty} k\frac{\alpha^{k-1}}{g(k)!}=\frac{\alpha}{Z(\alpha)}Z'(\alpha)=\alpha\partial_\alpha\log(Z(\alpha)), \tag{3.2}$$

where we used that for $\alpha<\alpha^*$ the function Z is analytic. Moreover,

$$\begin{aligned}R'(\alpha)=&-\frac{Z'(\alpha)}{Z(\alpha)^2}\sum_{k=0}^{\infty} k\frac{\alpha^k}{g(k)!}+\frac{1}{Z(\alpha)}\sum_{k=0}^{\infty} k^2\frac{\alpha^{k-1}}{g(k)!}=\frac{1}{Z(\alpha)}\sum_{k=0}^{\infty}\frac{\alpha^k}{g(k)!}\left(\frac{k^2}{\alpha}-k\frac{Z'(\alpha)}{Z(\alpha)}\right)\\ =&E_{m_\alpha}\left[\frac{K^2}{\alpha}-K\frac{E_{m_\alpha}[K]}{\alpha}\right]=\frac{1}{\alpha}E_{m_0}[(K^2-E_{m_\alpha}[K])^2]>0,\end{aligned}$$

where $E_{m_\alpha}[f(K)]:=\sum_k f(k)m_\alpha(k)$. Therefore, R is strictly increasing on $[0,\alpha^*)$.

Lemma 3.11. *If* $\lim_{\alpha\to\alpha^*}Z(\alpha)=\infty$, *then* $\lim_{\alpha\to\alpha^*}R(\alpha)=\infty$. *Moreover, in that case there exists for every* $\alpha\in[0,\alpha^*)$ *some* $\theta(\alpha)>0$ *with* $E_{m_\alpha}[e^{\theta(\alpha)K}]<\infty$, *i.e.* m_α *has some exponential moments.*

Proof. We distinguish the two cases $\alpha^*=\infty$ and $\alpha^*<\infty$.

1. Assume first that $\alpha^*=\infty$ and that $R(\alpha)\leqslant C$. By (3.2) we have for all $\alpha>1$

$$\partial_\alpha\log(Z(\alpha))=\frac{R(\alpha)}{\alpha}\geqslant\frac{C}{\alpha}\qquad\Rightarrow\qquad\log(Z(\alpha))-\log(Z(1))\leqslant\int_1^\alpha\frac{C}{\beta}\mathrm{d}\beta=C\times\log(\alpha)$$
$$\Rightarrow Z(\alpha)\leqslant Z(1)\alpha^C.$$

   But on the other hand we also know that $Z(\alpha)\geqslant\alpha^k/(g(k)!)$ for all $k\in\mathbb{N}$ and all $\alpha>0$. Taking $k>C$ and sending $\alpha\to\infty$, we obtain a contradiction.

2. Assume now that $\alpha^*<\infty$ and that $R(\alpha)\leqslant C$. Let $0<\alpha_0<\alpha^*$ and note that for $\alpha\in(\alpha_0,\alpha^*)$

$$\begin{aligned}\log(Z(\alpha))=&\log(Z(\alpha_0))+\int_{\alpha_0}^{\alpha}\partial_\beta\log(Z(\beta))\mathrm{d}\beta\leqslant\log(Z(\alpha_0))+\frac{1}{\alpha_0}\int_{\alpha_0}^{\alpha}\beta\partial_\beta\log(Z(\beta))\mathrm{d}\beta\\ =&\log(Z(\alpha_0))+\frac{1}{\alpha_0}\int_{\alpha_0}^{\alpha}R(\beta)\mathrm{d}\beta\leqslant\log(Z(\alpha_0))+\frac{C}{\alpha_0}(\alpha-\alpha_0),\end{aligned}$$

   and sending $\alpha\to\alpha^*$ we obtain a contradiction to the assumption $\lim_{\alpha\to\alpha^*}Z(\alpha)=\infty$.

3. It remains to show that with $E_{m_\alpha}[e^{\theta(\alpha)K}]<\infty$ for some $\theta(\alpha)>0$. But this is easy, because

$$E_{m_\alpha}[e^{\theta K}]=\frac{1}{Z(\alpha)}\sum_{k=0}^{\infty}e^{\theta k}\frac{\alpha^k}{g(k)!}=\frac{Z(e^\theta\alpha)}{Z(\alpha)},$$

   and for $e^\theta\alpha<\alpha^*$ the right hand side is finite. □

Under the conditions of Lemma 3.11 the function $R\colon[0,\alpha^*)\to[0,\infty)$ is bijective and therefore invertible. In the following it will be convenient to choose the measure $\tilde{\nu}_\alpha$ not by the (essentially meaningless) parameter α, but rather by the "density of particles" ρ, so we define for the zero range process

$$\nu_\rho:=\tilde{\nu}_{R^{-1}(\rho)}.$$

Since the Poisson distribution with parameter λ has expectation λ, we get $\nu_\rho=\tilde{\nu}_\rho$ for the case of independent random walks, $g(k)=k$. Note that by the law of large numbers we have

$$\lim_{N\to\infty}\nu_\rho^N\left(\left|\frac{1}{|\mathbb{Z}_N^d|}\sum_{x\in\mathbb{Z}_N^d}\eta(x)-\rho\right|>\varepsilon\right),$$

where we wrote ν_ρ^N instead of ν_ρ to stress that the measure also changes with N because the base space $\mathbb{Z}_N^d$ changes with N.

Exercise 3.2. Consider the exclusion process of Example 3.3 and show that for any $\rho \in [0,1]$ the product Bernoulli measure

$$\mu_\rho(\eta) = \rho^{\sum_x \eta(x)}(1-\rho)^{\sum_x (1-\eta(x))}$$

is invariant. Show that it is reversible if and only if $p^N(z) = p^N(-z)$ for all $z \in \mathbb{Z}_N^d$.

Solution.

1. *We first show stationarity. The argument is similar as for the zero range process, but easier. Since our state space is finite, we trivially have* $\|c\|_\infty < \infty$ *and therefore* $E_{\mu_\rho}[c] < \infty$. *Furthermore*

$$E_{\mu_\rho}[qf] = \sum_{x,y \in \mathbb{Z}_N^d} E_{\mu_\rho}[\eta(x)(1-\eta(y))p^N(x,y)(f(\eta^{x,y}) - f(\eta))]$$
$$= \sum_{x,y \in \mathbb{Z}_N^d} p^N(x,y)(E_{\mu_\rho}[\eta(x)(1-\eta(y))f(\eta^{x,y})] - E_{\mu_\rho}[\eta(x)(1-\eta(y))f(\eta)]).$$

But now $\eta^{x,y}$ *is just the configuration* η *with the occupation variables in* x *and* y *exchanged, and therefore* $\mu_\rho(\eta^{y,x}) = \mu_\rho(\eta)$, *so that renaming the summation variable* $\eta^{x,y} \to \eta$, *we get*

$$E_{\mu_\rho}[\eta(x)(1-\eta(y))f(\eta^{x,y})] = \sum_{\eta \in S} \mu_\rho(\eta)\eta(x)(1-\eta(y))f(\eta^{x,y})$$
$$= \sum_{\eta \in S} \mu_\rho(\eta^{x,y})\eta(x)(1-\eta(y))f(\eta^{x,y})$$
$$= \sum_{\eta \in S} \mu_\rho(\eta)\eta^{x,y}(x)(1-\eta^{x,y}(y))f(\eta)$$
$$= \sum_{\eta \in S} \mu_\rho(\eta)\eta(y)(1-\eta(x))f(\eta) = E_{\mu_\rho}[\eta(y)(1-\eta(x))f(\eta)],$$

and exchanging the names of the summation variables x, y, *we obtain*

$$\sum_{x,y \in \mathbb{Z}_N^d} p^N(x,y)E_{\mu_\rho}[\eta(x)(1-\eta(y))f(\eta^{x,y})] = \sum_{x,y \in \mathbb{Z}_N^d} p^N(x,y)E_{\mu_\rho}[\eta(y)(1-\eta(x))f(\eta)]$$
$$= \sum_{x,y \in \mathbb{Z}_N^d} p^N(y,x)E_{\mu_\rho}[\eta(x)(1-\eta(y))f(\eta)].$$

This leads to

$$E_{\mu_\rho}[qf] = \sum_{x,y \in \mathbb{Z}_N^d} (p^N(y,x) - p^N(x,y))E_{\mu_\rho}[\eta(x)(1-\eta(y))f(\eta)],$$

and by symmetry under changing the variable names x, y *we get*

$$\sum_{x,y \in \mathbb{Z}_N^d} (p^N(y,x) - p^N(x,y))E_{\mu_\rho}[\eta(x)\eta(y)f(\eta)] = 0.$$

Therefore,

$$E_{\mu_\rho}[qf] = \sum_{x,y \in \mathbb{Z}_N^d} (p^N(y,x) - p^N(x,y))E_{\mu_\rho}[\eta(x)f(\eta)] = 0$$

because $\sum_y (p^N(y,x) - p^N(x,y)) = 1 - 1 = 0$.

2. *Next, we need to show that* μ_ρ *is reversible if and only if* $p^N(z) = p^N(-z)$ *for all* z. *Since* $\|c\|_\infty < \infty$ *Lemma 2.25 shows that reversibility is equivalent to the detailed balance condition* $\mu_\rho(\eta)q(\eta,\zeta) = \mu_\rho(\zeta)q(\zeta,\eta)$. *But* $q(\eta,\zeta) = q(\zeta,\eta) = 0$ *unless* $\zeta = \eta^{x,y}$ *for some* $x, y \in \mathbb{Z}_N^d$, *and then* $\mu_\rho(\eta) = \mu_\rho(\eta^{x,y})$, *which means that* μ_ρ *is reversible if and only if* $q(\eta,\eta^{x,y}) = q(\eta^{x,y},\eta)$. *Now*

$$q(\eta,\eta^{x,y}) = \eta(x)(1-\eta(y))p^N(x,y)$$

and $q(\eta^{x,y},\eta)$ *is the rate with which, in the configuration* $\eta^{x,y}$, *a particle moves from* y *to* x, *i.e.*

$$q(\eta^{x,y},\eta) = \eta^{x,y}(y)(1-\eta^{x,y}(x))p^N(y,x) = \eta(x)(1-\eta(y))p^N(y,x),$$

from where the claim follows.

Since the Bernoulli measure with parameter ρ has expected value ρ, there is no need to change the parametrization as for the zero range process.

Example 3.12. For the contact process there is one trivial invariant measure, the Dirac distribution δ_0 that assigns probability 1 to the configuration $\mathbf{0}$ with $\mathbf{0}(x) = 0$ for all $x \in V$. Moreover, for all $\eta \in S$ we have $p_1(\eta, \mathbf{0}) > 0$, and since S is finite we also have

$$\min_{\eta \in S} p_1(\eta, \mathbf{0}) > 0 \qquad \Leftrightarrow \qquad p^* := \max_{\eta \in S} \left(\sum_{\eta' \neq \mathbf{0}} p_1(\eta, \eta') \right)$$

Hence, we get for any $\eta \in S$

$$\mathbb{P}_\eta(\eta_n \neq \mathbf{0}) = \sum_{\eta_1, \ldots, \eta_n \neq \mathbf{0}} p_1(\eta, \eta_1) p_1(\eta_1, \eta_2) \times \ldots \times p_1(\eta_{n-1}, \eta_n) \leqslant (p^*)^n,$$

which converges to 0 for $n \to \infty$. So η is almost surely absorbed in $\mathbf{0}$ for large times (the disease dies out), and δ_0 is the only invariant measure. If V is infinite rather than finite, than it is rather subtle whether the disease dies out, it depends on the geometry of V and also on the parameter λ. This is nicely explained in Chapter 4.4 of [Liggett, *Continuous Time Markov Processes*, AMS, 2010].

Example 3.13. For the voter model there are always at least two invariant measures, δ_0 and δ_1, where δ_0 is the Dirac distribution in the $\mathbf{0}$ configuration, and δ_1 is the Dirac distribution in the $\mathbf{1}$ configuration with $\mathbf{1}(x) = 1$ for all $x \in V$. The same argument as for the contact process shows that no matter the starting configuration, with probability 1 the process reaches unanimity at some point, i.e. all individuals adapt the same opinion. This shows that any invariant measure has to be supported only on $\mathbf{0}$ and $\mathbf{1}$, and since the convex combination of two stationary measures is stationary (convince yourself of this, it holds for general Markov processes!), all invariant measures are of the form $\lambda\delta_0 + (1 - \lambda)\delta_1$ with $\lambda \in [0, 1]$. On infinite graphs this is again much more complicated, see Chapter 4.3 of [Liggett, *Continuous Time Markov Processes*, AMS, 2010] for details.

The two examples of the contact process and the voter model show a particularly long time behavior, as we just discussed. To understand the behavior of the zero range process or the exclusion process for large times, we could develop a similar ergodic theory as for discrete time Markov chains, and prove analogous results. But we are more interested how the processes behave on large scales, which means that we should appropriately rescale space as we look at the large time limit.

4 First hydrodynamic limits

The *hydrodynamic limit* of an interacting particle system describes its behavior on large space-time scales. Here we picture the particles as molecules in a fluid, and the hydrodynamic limit describes the collective behavior on the observable scale. We will first study this problem for our toy problem of independent random walks. In this section we follow mostly [Kipnis, Landim, *Scaling limits of interacting particle systems*, Springer 1999], as well as the nice lecture notes [Sethuraman, *Large scale stochastic dynamics*, University of Arizona lecture notes, 2012].

4.1 Warm-up: one particle

The large scale behavior of an individual particle is very easy to understand. Indeed, let again p: $\mathbb{Z}^d \to [0, 1]$ be the step probability of a random walk, that is $\sum_{z \in \mathbb{Z}^d} p(z) = 1$, and now we assume that $p(z) \neq 0$ only for finitely many z, and since it sometimes simplifies the notation in what follows also that $p(0) = 0$. As before we set

$$p^N(x, y) := \sum_{k \in \mathbb{Z}^d} p(y + kN - x),$$

and we let $(X_t^N)_{t\geqslant 0}$ be the continuous time random walk on $\mathbb{Z}_N^d$ with q-matrix

$$q^N(x,y):=\begin{cases} p^N(x,y), & y\neq x,\\ -\sum_{y\neq x} p^N(x,y), & y=x,\end{cases}$$

started at 0. Recall that we can construct $X_t^N=Y_{R_t}^N$, where $(Y_n^N)_{n\in\mathbb{N}_0}$ is a discrete time Markov chain with transition function p^N, started at 0, and $(R_t)_{t\geqslant 0}$ is an independent Poisson process with intensity 1 (because $c(x)=\sum_{y\neq x}p^N(x,y)=1$ for all x).

The large scale behavior of X^N is quite simple. First we rescale the space variable and "zoom out". Since $\mathbb{Z}_N^d$ has a finite size, we cannot zoom out arbitrarily much for X^N, and in fact the scale in which we study the process should be adapted to $\mathbb{Z}_N^d$. For example, we can rescale $x\to Nx$, which means that we should take $x\in N^{-1}\mathbb{Z}_N^d$ (so that $Nx\in\mathbb{Z}_N^d$). Here we identify $N^{-1}\mathbb{Z}_N^d$ with a subset of $\mathbb{T}^d:=(\mathbb{R}/\mathbb{Z})^d$. Of course, for a finite time $t\geqslant 0$ the process X_t^N has with high probability a distance oN from its starting point, so on the spatial scale N it did not move: $N^{-1}X_t^N$ converges to 0 almost surely. To see nontrivial dynamics, we should also increase the speed of our process, for example by scaling $t\to Nt$.

Note that since we want to go to large scales, we have to send $N\to\infty$. This means that we have to change our process X^N as we change N, because we have to change the image space $\mathbb{Z}_N^d$. To clarify this, we write X^N here. Later we will often omit the superscript N.

Lemma 4.1. *For all $t\geqslant 0$ we have*

$$\lim_{N\to\infty}\frac{1}{N}X_{Nt}^N=vt,\qquad v:=\sum_{x\in\mathbb{Z}^d}xp(x),$$

where we interpret $\frac{1}{N}X_{Nt}^N$ and vt as elements of $\mathbb{T}^d$.

Remark 4.2. The distance on $\mathbb{T}^d$ is

$$d_{\mathbb{T}^d}(u,w):=\min\{|u+k-w|:k\in\mathbb{Z}^d\},$$

where $|u+k-w|$ is the Euclidean distance on $\mathbb{R}^d$. So the claim of the lemma is that for all $\varepsilon>0$ we have

$$\lim_{N\to\infty}\mathbb{P}\left(d_{\mathbb{T}^d}\left(\frac{1}{N}X_{Nt}^N,vt\right)>\varepsilon\right)=0.$$

To simplify notation we will usually write $|u-w|$ instead of $d_{\mathbb{T}^d}(u,w)$, but as long as we are in a periodic setting we always mean the periodic distance.

Proof. Let $(Y_n)_{n\in\mathbb{N}_0}$ be a random walk on $\mathbb{Z}^d$, started in 0 and with transition probability p, independent of R, and let $X_t:=Y_{R_t}$, $t\geqslant 0$. Let $\Phi_N:\mathbb{Z}^d\to\mathbb{Z}_N^d$ be the canonical map that sends $x\in\mathbb{Z}^d$ to "$x(\mathrm{mod}\,(N\mathbb{Z})^d)$" in $\mathbb{Z}_N^d$. Then $\Phi_N(X)$ has the same law as X^N, and moreover

$$d_{\mathbb{T}^d}\left(\Phi_N\left(\frac{1}{N}X_{Nt}\right),tv\right)=d_{\mathbb{T}^d}\left(\Phi_N\left(\frac{1}{N}X_{Nt}\right),\Phi_N\left(t\sum_{x\in\mathbb{Z}^d}xp(x)\right)\right)\leqslant\left|\frac{1}{N}X_{Nt}-t\bar{v}\right|,$$

where we wrote $\bar{v}:=\sum_{x\in\mathbb{Z}^d}xp(x)\in\mathbb{R}^d$. Therefore, it suffices to show that $\frac{1}{N}X_{Nt}$ converges in probability to $t\bar{v}$ as $N\to\infty$. But this is just a slightly more complicated version of the weak law of large numbers, so we use the same proof as for the weak law of large numbers: Since $|a+b|^2\leqslant 2(a^2+b^2)$ for $a,b\geqslant 0$ (which follows e.g. from the Cauchy-Schwarz inequality) we have

$$\begin{aligned}\mathbb{E}\left[\left|\frac{1}{N}X_{Nt}-\bar{v}t\right|^2\right]&\lesssim\mathbb{E}\left[\left|\frac{1}{N}X_{Nt}-\bar{v}\frac{R_{Nt}}{N}\right|^2\right]+\frac{|\bar{v}|^2}{N^2}\mathbb{E}[|R_{Nt}-Nt|^2]\\&=\frac{1}{N^2}\mathbb{E}[\mathbb{E}[|Y_{R_{Nt}}-R_{Nt}|^2|(R_t)_{t\geqslant 0}]]+\frac{|v|^2}{N^2}\mathbb{E}[|R_{Nt}-Nt|^2].\end{aligned}$$

Now recall that Y and R are independent, and that R_{Nt} is Poisson distributed with parameter Nt, so that $\mathbb{E}[R_{Nt}] = \mathrm{var}(R_{Nt}) = Nt$. Moreover, $\mathbb{E}[Y_k] = k\bar{v}$ and $\mathbb{E}[|Y_k - k\bar{v}|^2] = k\sigma^2$ for $\sigma^2 = \sum_{x\in\mathbb{Z}^d} |x-\bar{v}|^2 p(x)$ because Y is a random walk (convince yourself of this!), which leads to

$$\mathbb{E}\left[\left|\frac{1}{N}X_{Nt} - \bar{v}\,t\right|^2\right] \lesssim \frac{1}{N^2}\mathbb{E}[R_{Nt}\sigma^2] + \frac{|v|^2}{N^2}\mathrm{var}(R_{Nt}) = \frac{t}{N}(\sigma^2 + |v|^2) \longrightarrow 0,$$

$N\to\infty$. □

So on large space-time scales we see our particle move in $\mathbb{T}^d$ with constant speed $v\in\mathbb{T}^d$. It is not difficult to strengthen the previous result to obtain also uniform convergence in $t\in[0,T]$, but we will not bother to do so.

Before we get to studying independent particles, let us translate the convergence we just proved from the random walk point of view to the particle system point of view. Recall that the particle "system" associated to X^N was defined as $\eta_t^N \in \{0,1\}^{\mathbb{Z}_N^d}$,

$$\eta_t^N(x) := \mathbb{I}_{X_t^N}(x) := \begin{cases} 1, & X_t^N = x, \\ 0, & X_t^N \neq x. \end{cases}$$

But now we are interested in the rescaled process $N^{-1}X_t^N$, so we define $\tilde{\eta}_t^N \in \{0,1\}^{N^{-1}\mathbb{Z}_N^d}$ by

$$\tilde{\eta}_t^N(u) := \mathbb{I}_{N^{-1}X_t^N}(u) := \begin{cases} 1, & N^{-1}X_t^N = u, \\ 0, & N^{-1}X_t^N \neq u. \end{cases}$$

We can again identify $N^{-1}\mathbb{Z}_N^d$ with a subset of $\mathbb{T}^d$, so that we can interpret $\tilde{\eta}_t^N$ as a map from $\mathbb{T}^d$ to $\{0,1\}$, with $\tilde{\eta}_t^N(u) = 1$ for exactly one $u\in\mathbb{T}^d$. We could now pass to the limit as before, and obtain that $\tilde{\eta}_{Nt}^N$ converges in probability to the map $\zeta_t\colon \mathbb{T}^d \to \{0,1\}$ with $\zeta_t(vt) = 1$ and $\zeta_t(u) = 0$ for $u \neq vt$. But mathematically this is not a very convenient object to work with, and when we will deal with particle systems with a number of particles that grows to infinity as $N\to\infty$ it would not be clear how to make sense of ζ_t in the limit. Another way of encoding the same information is to identify $\tilde{\eta}_t^N$ with a measure on $\mathbb{T}^d$, more precisely with the Dirac delta in $N^{-1}X_t^N$, defined as

$$\int_{\mathbb{T}^d} \varphi(u)\tilde{\eta}_t^N(\mathrm{d}u) := \sum_{x\in N^{-1}\mathbb{Z}_N^d} \varphi(u)\tilde{\eta}_t^N(u) = \sum_{x\in\mathbb{Z}_N^d} \varphi\Big(\frac{x}{N}\Big)\eta_t^N(x) = \varphi(N^{-1}X_t^N),$$

for $\varphi\colon\mathbb{T}^d\to\mathbb{R}$. We can then show that $\tilde{\eta}_{Nt}^N$ converges for $N\to\infty$ in probabiltiy to the Dirac delta in vt. To state such a convergence, of course we have to work with a topology on the space of measures, and here the convergence holds in the weak sense.

4.2 Independent particles

Let now η^N be the particle system of Example 3.2, i.e. η^N corresponds to independent random walks on $\mathbb{Z}_N^d$. We will derive the hydrodynamic limit of the independent random walks in two different ways. First, let us try to guess what the limit should be: Let $\rho_0\colon\mathbb{T}^d\to[0,\infty)$ be in $C^1(\mathbb{T}^d)$, where

$$C^k(\mathbb{T}^d) := \{f\in C^k(\mathbb{R}^d),\, f \text{ is } 1-\text{periodic}\},$$

and let $\eta_0^N(x) = \rho_0\big(\frac{x}{N}\big)$, meaning that we start $\rho_0\big(\frac{x}{N}\big)$ independent walkers in x. As we have seen above, at time Nt each of these walkers will be (approximately, and at the macroscopic scale) in $\frac{x}{N} + vt$. So defining for $u\in\mathbb{T}^d$

$$\tilde{\eta}_t^N(u) = \eta_t^N([Nu]),$$

where $[Nu] = ([Nu_1],\ldots,[Nu_d])$ and $[Nu_i]$ is the Gauss bracket, we would expect that

$$\lim_{N\to\infty} \tilde{\eta}_{Nt}^N(u) = \rho_0(u - vt) =: \rho(t,u),$$

which solves the *transport equation*

$$\partial_t\rho(t,u) = -v\cdot\nabla\rho(t,u), \qquad \rho(0,u) = \rho_0(u).$$

This is a first link between particle systems and partial differential equations!

Now let us prove the claimed convergence rigorously. To simplify the arguments, we will only consider initial configurations which are "locally stationary" in the sense that they are product Poisson distributions with a space-dependent density of particles.

We first derive a more complicated version of the hydrodynamic limit theorem that introduces the important concept of *local equilibrium*. Then we will give a simpler proof of a "more global" version of the result.

Notation. *Recall that for all $\rho>0$ the product Poisson measure ν_ρ^N on $\mathbb{N}_0^{\mathbb{Z}_N^d}$ with density ρ is invariant for η^N. If $\rho\colon \mathbb{T}^d\to\mathbb{R}_+$ is a function rather than a constant, then we also write ν_ρ^N for the product measure with*

$$\nu_\rho^N(\eta)=\prod_{x\in\mathbb{Z}_N^d} e^{-\rho(\frac{x}{N})}\frac{\rho\big(\frac{x}{N}\big)^{\eta(x)}}{\eta(x)!}.$$

We show now that if the function ρ is continuous, then for large N and on the macroscopic scale close to $u\in\mathbb{T}^d$ (i.e. around $x=[Nu]$ on the microscopic scale) the measure ν_ρ^N is "close" to a Poisson distribution with parameter $\rho(u)$.

Definition 4.3. *A sequence of probability measures μ_N on $\mathbb{N}_0^{\mathbb{Z}_N^d}$ is a local equilibrium of profile ρ: $\mathbb{T}^d\to\mathbb{R}_+$ if for all continuity points of ρ and all bounded $\varphi\colon\mathbb{N}_0^{\mathbb{Z}^d}\to\mathbb{R}$ that only depend on finitely many coordinates $\eta(x)$ with $|x|\leqslant\ell$ for some $m>0$*

$$\lim_{N\to\infty}\int_{\mathbb{N}_0^{\mathbb{Z}_N^d}}\varphi(\eta(\cdot+[uN]))\mu^N(\mathrm{d}\eta)=\int_{\mathbb{N}_0^{\mathbb{Z}^d}}\varphi(\eta)\nu_{\rho(u)}(\mathrm{d}\eta)=E_{\nu_{\rho(u)}}[\varphi],$$

where $\nu_{\rho(u)}$ is the product Poisson distribution on $\mathbb{N}_0^{\mathbb{Z}^d}$ with parameter $\rho(u)$.

Lemma 4.4. *Let $\rho\colon\mathbb{T}^d\to\mathbb{R}_+$ and let ρ^N be a sequence of functions that converges uniformly to ρ. Then the sequence $(\nu_{\rho^N}^N)$ is a local equilibrium of profile ρ.*

Proof. It suffices to consider the characteristic function,

$$\varphi(\eta)=\exp\Bigg(i\sum_{|x|\leqslant\ell}\theta(x)\eta(x)\Bigg),$$

for which (if $N>\ell$)

$$\int_{\mathbb{N}_0^{\mathbb{Z}_N^d}}\exp\Bigg(i\sum_{|x|\leqslant\ell}\theta(x)\eta(x+[uN])\Bigg)\nu_{\rho^N}^N(\mathrm{d}\eta)=\prod_{|x|\leqslant\ell}\int_{\mathbb{N}_0^{\mathbb{Z}_N^d}}\exp(i\theta(x)\eta(x+[uN]))\nu_{\rho^N}^N(\mathrm{d}\eta)$$
$$=\prod_{|x|\leqslant\ell}\int_{\mathbb{N}_0}\exp(i\theta(x)k)m_{\rho^N((x+[uN])/N)}(\mathrm{d}k)=\prod_{|x|\leqslant\ell}\exp\Bigg(i\rho^N\bigg(\frac{x+[uN]}{N}\bigg)(e^{i\theta(x)}-1)\Bigg),$$

where we wrote $m_{\rho^N((x+[uN])/N)}$ for the Poisson distribution with parameter $\rho^N((x+[uN])/N)$. Of course

$$\lim_{N\to\infty}\frac{x+[uN]}{N}=u,$$

and if ρ is continuous in u we have

$$\left|\rho^N\bigg(\frac{x+[uN]}{N}\bigg)-\rho(u)\right|\leqslant\left|\rho^N\bigg(\frac{x+[uN]}{N}\bigg)-\rho\bigg(\frac{x+[uN]}{N}\bigg)\right|+\left|\rho\bigg(\frac{x+[uN]}{N}\bigg)-\rho(u)\right|,$$

which converges to zero because ρ^N converges uniformly to ρ. Thus, we get

$$\lim_{N\to\infty}\int_{\mathbb{N}_0^{\mathbb{Z}_N^d}}\varphi(\eta(\cdot+[uN]))\mu^N(\mathrm{d}\eta)=\lim_{N\to\infty}\exp\Bigg(i\sum_{|x|\leqslant\ell}\rho^N\bigg(\frac{x+[uN]}{N}\bigg)(e^{i\theta(x)}-1)\Bigg)$$
$$=\exp\Bigg(i\sum_{|x|\leqslant\ell}\rho(u)(e^{i\theta(x)}-1)\Bigg)=E_{\nu_{\rho(u)}}[\varphi],$$

which is the claim. □

Lemma 4.5. *Consider $(\eta_t^N)_{t\geqslant 0}$ with initial distribution $\nu_{\rho_0}^N$ for some $\rho_0: \mathbb{T}^d \to \mathbb{R}_+$. Then η_t^N has distribution $\nu_{\psi_t^N}^N$ for $\psi_t^N: \mathbb{T}^d \to \mathbb{R}_+$ defined by*

$$\psi_t^N(x) := \sum_{y\in\mathbb{Z}_N^d} \rho_0\Big(\frac{y}{N}\Big) p_t^N(y,x).$$

Proof. We work again with the characteristic function. Let us write $X^{y,k}$, for $y \in \mathbb{Z}_N^d$ and $k = \{1, ..., \eta_0(y)\}$, for the k-th random walk started in y. Then

$$\sum_{x\in\mathbb{Z}_N^d} \theta(x)\eta_t^N(x) = \sum_{x\in\mathbb{Z}_N^d} \theta(x) \sum_{y\in\mathbb{Z}_N^d} \sum_{k=1}^{\eta_0(y)} \mathbb{I}_{X_t^{y,k}=x} = \sum_{y\in\mathbb{Z}_N^d} \sum_{k=1}^{\eta_0(y)} \theta(X_t^{y,k}),$$

and therefore, since $\nu_{\rho_0}^N$ is a product measure,

$$\begin{aligned}\mathbb{E}_{\nu_{\rho_0}^N}\Bigg[\exp\Bigg(i\sum_{x\in\mathbb{Z}_N^d}\theta(x)\eta_t^N(x)\Bigg)\Bigg] &= \prod_{y\in\mathbb{Z}_N^d}\mathbb{E}_{\nu_{\rho_0}^N}\Bigg[\exp\Bigg(i\sum_{k=1}^{\eta_0(y)}\theta(X_t^{y,k})\Bigg)\Bigg]\\ &= \prod_{y\in\mathbb{Z}_N^d}\mathbb{E}_{\nu_{\rho_0}^N}\Bigg[\mathbb{E}_{\nu_\rho^N}\Bigg[\exp\Bigg(i\sum_{k=1}^{\eta_0(y)}\theta(X_t^{y,k})\Bigg)\Bigg|\eta_0\Bigg]\Bigg]\\ &= \prod_{y\in\mathbb{Z}_N^d}\mathbb{E}_{\nu_{\rho_0}^N}[\mathbb{E}[\exp(i\theta(X_t^{y,1}))]^{\eta_0(y)}].\end{aligned}$$

Now if K is Poisson distributed with parameter ρ, then

$$E[a^K] = \sum_{k=0}^{\infty} a^k \frac{\rho^k}{k!} e^{-\rho} = e^{a\rho-\rho} = e^{\rho(a-1)},$$

and therefore

$$\begin{aligned}\prod_{y\in\mathbb{Z}_N^d}\mathbb{E}_{\nu_{\rho_0}^N}[\mathbb{E}[\exp(i\theta(X_t^{y,1}))]^{\eta_0(y)}] &= \prod_{y\in\mathbb{Z}_N^d}\exp\Big(\rho_0\Big(\frac{y}{N}\Big)(\mathbb{E}[\exp(i\theta(X_t^{y,1}))]-1)\Big)\\ &= \prod_{y\in\mathbb{Z}_N^d}\exp\Bigg(\rho_0\Big(\frac{y}{N}\Big)\Bigg(\sum_{z\in\mathbb{Z}_N^d}p_t^N(y,z)e^{i\theta(x)}-1\Bigg)\Bigg)\\ &= \exp\Bigg(\sum_{y\in\mathbb{Z}_N^d}\rho_0\Big(\frac{y}{N}\Big)\sum_{z\in\mathbb{Z}_N^d}p_t^N(y,z)(e^{i\theta(x)}-1)\Bigg)\\ &= \exp\Bigg(\sum_{z\in\mathbb{Z}_N^d}\psi_t^N(z)(e^{i\theta(x)}-1)\Bigg).\end{aligned}$$

Since this is the characteristic function of the product Poisson distribution with varying intensity ψ_t^N, the proof is complete. □

We know now that for all N and all t the distribution of η_t^N is still product Poisson, with an explicitly given profile ψ_t^N. To understand the behavior of η_t^N on large spatial scales, we therefore only need to understand how ψ_t^N looks on large spatial scales. For $u \in \mathbb{T}^d$ and $\varepsilon > 0$ there exists $C > 0$ such that

$$\begin{aligned}|\psi_t^N([uN]) - \rho(u)| &= \Bigg|\sum_{y\in\mathbb{Z}_N^d}\rho_0\Big(\frac{y}{N}\Big)p_t^N(y,[uN]) - \rho_0(u)\Bigg| = \Bigg|\sum_{y\in\mathbb{Z}_N^d}\Big(\rho_0\Big(\frac{y}{N}\Big)-\rho_0(u)\Big)p_t^N([uN]-y)\Bigg|\\ &\leqslant \Bigg|\sum_{y:|[uN]-y|\leqslant C}\Big(\rho_0\Big(\frac{y}{N}\Big)-\rho_0(u)\Big)p_t^N([uN]-y)\Bigg| + \varepsilon,\end{aligned}$$

so if ρ_0 is continuous we obtain $\lim_{N\to\infty}\psi_t^N([uN]) = \rho_0(u)$, i.e. the density profile does not evolve in time. This is not surprising, because we already saw for the single random walk that if we observe it on large spatial scales, then we have to speed up time to see a nontrivial evolution.

Lemma 4.6. *Let $\rho_0 \in C^1(\mathbb{T}^d, \mathbb{R}_+)$. Then we have for all $t \geqslant 0$ uniformly in $u \in \mathbb{T}^d$*

$$\lim_{N \to \infty} \psi^N_{Nt}([uN]) = \rho(t,u) := \rho_0(u - tv).$$

The function ρ is the unique solution to the transport equation

$$\partial_t \rho = -v \cdot \nabla \rho, \qquad \rho(0) = \rho_0. \tag{4.1}$$

Proof. We simply plug in the definition:

$$\begin{aligned}\psi^N_{Nt}([uN]) &= \sum_{y \in \mathbb{Z}^d_N} \rho_0\Big(\frac{y}{N}\Big) p^N_{Nt}(y, [uN]) = \sum_{y \in \mathbb{Z}^d_N} \rho_0\Big(\frac{y}{N}\Big) p^N_{Nt}([uN] - y) \\ &= \sum_{z \in \mathbb{Z}^d_N} \rho_0\Big(\frac{[uN] - z}{N}\Big) p^N_{Nt}(z) = \mathbb{E}\Big[\rho_0\Big(\frac{[uN] - X^N_{Nt}}{N}\Big)\Big],\end{aligned}$$

where X^N is the random walk with transition function p^N, started at 0. Now $\mathbb{T}^d$ is compact, and therefore ρ_0 is uniformly continuous and bounded, and given $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\left|\mathbb{E}\Big[\rho_0\Big(\frac{[uN] - X^N_{Nt}}{N}\Big)\Big] - \rho_0(u - vt)\right| \leqslant \varepsilon + 2\|\rho_0\|_\infty \mathbb{P}\Big(d_{\mathbb{T}^d}\Big(\frac{[uN] - X^N_{Nt}}{N}, (u - vt)\Big) > \delta\Big),$$

and by Lemma 4.1 $N^{-1} X^N_{Nt}$ converges in probability to vt, which means that the right hand side converges to 0, uniformly in $u \in \mathbb{T}^d$.

It remains to show that ρ solves the transport equation (which immediately follows from the chain rule) and that it is the only solution to the transport equation. But if $\tilde{\rho}$ is another solution, then we get

$$\partial_t \tilde{\rho}(t, u + tv) = (\partial_t \tilde{\rho})(t, u + tv) + v \cdot (\nabla \tilde{\rho})(t, u + tv) = -v \cdot (\nabla \tilde{\rho})(t, u + tv) + v \cdot (\nabla \tilde{\rho})(t, u + tv) = 0,$$

and therefore $\tilde{\rho}(t, u + tv) = \tilde{\rho}(0, u) = \rho_0(u)$ and then $\tilde{\rho}(t, u) = \rho_0(u - tv) = \rho(t, u)$. ☐

If $v = 0$, then even on the time scale Nt the profile does not change, and we have to look at even larger time scales to see nontrivial dynamics:

Lemma 4.7. *Let $\rho_0 \in C^2(\mathbb{T}^d, \mathbb{R}_+)$ and assume that $v = \sum_{x \in \mathbb{Z}^d} x p(x) = 0$ and that $\det(C) \neq 0$ for the covariance matrix*

$$C_{ij} = \sum_{x \in \mathbb{Z}^d} x_i x_j p(x).$$

Then we have for all $t \geqslant 0$ uniformly in $u \in \mathbb{T}^d$

$$\lim_{N \to \infty} \psi^N_{N^2 t}([uN]) = \rho(t, u) := \int_{\mathbb{R}^d} \bar{\rho}_0(w) G_t(u - w) \mathrm{d}w,$$

where $\bar{\rho}_0$ is the periodic extension of ρ_0 from $\mathbb{T}^d$ to $\mathbb{R}^d$ and where G_t is the density of a centered Gaussian random variable with covariance tC. Let us write $\Delta_C := \sum_{i,j=1}^d C_{ij} \partial_{u_i u_j}$, so the function ρ is the unique solution to the heat equation

$$\partial_t \rho = \frac{1}{2} \Delta_C \rho, \qquad \rho(0) = \rho_0. \tag{4.2}$$

Proof. We have again

$$\psi^N_{N^2 t}([uN]) = \mathbb{E}\Big[\rho_0\Big(\frac{[uN] - X^N_{N^2 t}}{N}\Big)\Big] = \mathbb{E}\Big[\bar{\rho}_0\Big(\frac{[uN] - X_{N^2 t}}{N}\Big)\Big],$$

where X is a non-periodic continuous time random walk with transition rates p. By the central limit theorem of the following exercise, we therefore have

$$\lim_{N \to \infty} \mathbb{E}\Big[\bar{\rho}_0\Big(\frac{[uN] - X_{N^2 t}}{N}\Big)\Big] = \int_{\mathbb{R}^d} \bar{\rho}_0(u - w) G_t(w) \mathrm{d}w \int_{\mathbb{R}^d} \bar{\rho}_0(w) G_t(u - w) \mathrm{d}w,$$

and as in the previous lemma we can show that the convergence is uniform in $u \in \mathbb{T}^d$

We know from stochastic analysis or functional analysis that ρ solves the heat equation, or we can check it by direct computation using the explicit form of the Gaussian density. Uniqueness of solutions to the heat equation is easy to show, for example by using the Fourier transform. Later we will study more complicated equations and prove the uniqueness of their solutions, so here we omit the uniqueness proof. □

Exercise 4.1. Let $p: \mathbb{Z}^d \to \mathbb{R}_+$ be a transition probability with $v = \sum_x x p(x) = 0$ and $C_{jk} = \sum_{x \in \mathbb{Z}^d} x_j x_k p(x) < \infty$ for all j, k and let X be the continuous time random walk with transition rates p. Show that X satisfies a central limit theorem, in the sense that for all $\theta_1, \ldots, \theta_d \in \mathbb{R}$

$$\lim_{N \to \infty} \mathbb{E}\left[\exp\left(i \sum_{j=1}^d \theta_j N^{-1} X^j_{N^2 t}\right)\right] = \exp\left(-\frac{1}{2} \sum_{j,k=1}^d t\theta_j \theta_k C_{jk}\right).$$

Hint: Recall that X can be constructed as $X_t = Y_{R_t}$, where $(Y_n)_{n \in \mathbb{N}_0}$ is a discrete time random walk with transition probabilites p and where $(R_t)_{t \geqslant 0}$ is an independent Poisson process with rate 1. You may use without proof that the multivariate central limit theorem holds for Y.

Solution. *Since R and Y are independent, we have*

$$\mathbb{E}\left[\exp\left(i \sum_{j=1}^d \theta_j N^{-1} X^j_{N^2 t}\right)\right] = \mathbb{E}\left[\exp\left(i \frac{\theta}{N} \cdot Y_{R_{N^2 t}}\right)\right] = \sum_{m=0}^\infty e^{-N^2 t} \frac{(N^2 t)^m}{m!} \mathbb{E}\left[\exp\left(i \frac{\theta}{N} \cdot Y_m\right)\right].$$

Let $(\xi_\ell)_{\ell \in \mathbb{N}}$ be i.i.d. variables with distribution p with $Y_m = \sum_{\ell=1}^m \xi_\ell$, and write ϕ for the characteristic function of ξ_1. Then

$$\sum_{m=0}^\infty e^{-N^2 t} \frac{(N^2 t)^m}{m!} \mathbb{E}\left[\exp\left(i \frac{\theta}{N} \cdot Y_m\right)\right] = \sum_{k=0}^\infty e^{-N^2 t} \frac{(N^2 t)^m}{m!} \phi\left(\frac{\theta}{N}\right)^m = \exp\left(-N^2 t\left(\phi\left(\frac{\theta}{N}\right) - 1\right)\right),$$

and using the same arguments as in the characteristic-function-based proof of the one-dimensional central limit theorem we obtain

$$\lim_{N \to \infty} N^2 t\left(\phi\left(\frac{\theta}{N}\right) - 1\right) = \frac{1}{2} \sum_{j,k=1}^d C_{jk} \theta_j \theta_k,$$

which concludes the proof.

Exercise 4.2. Show that if C is an invertible covariance matrix and G_t is the density of a centered Gaussian variable with covariance tC, then G solves the heat equation

$$\partial_t G_t(u) = \frac{1}{2} \Delta_C G_t(u), \qquad t > 0.$$

Solution. *This follows from a direct (and not very inspired) computation:*

$$G_t(u) = (2\pi t)^{-d/2} \det(C)^{-1/2} \exp\left(-\frac{1}{2} u^T t^{-1} C^{-1} u\right),$$

and therefore

$$\partial_t G_t(u) = G_t(u)\left(-\frac{d}{2t} + \frac{1}{2t^2} u^T C^{-1} u\right),$$

$$\partial_{u_i} G_t(u) = G_t(u)\left(-t^{-1} \sum_{k=1}^d C^{-1}_{ik} u_k\right),$$

$$\partial_{u_i u_j} G_t(u) = G_t(u)\left(t^{-2} \sum_{k,\ell=1}^d C^{-1}_{ik} C^{-1}_{j\ell} u_k u_\ell - t^{-1} C^{-1}_{ij}\right),$$

which leads to

$$\begin{aligned}\frac{1}{2}\sum_{i,j=1}^{d} C_{ij}\partial_{u_iu_j}G_t(u) &= \frac{G_t(u)}{2}\left(t^{-2}\sum_{i,j,k,\ell=1}^{d} C_{ij}C^{-1}_{ik}C^{-1}_{j\ell}u_ku_\ell - t^{-1}\sum_{i,j=1}^{d} C_{ij}C^{-1}_{ij}\right)\\ &= \frac{G_t(u)}{2}\left(t^{-2}\sum_{i,k,\ell=1}^{d}(\mathrm{Id})_{i\ell}C^{-1}_{ik}u_ku_\ell - t^{-1}\sum_{i=1}^{d}(\mathrm{Id})_{ii}\right)\\ &= \frac{G_t(u)}{2}\left(t^{-2}\sum_{i,k=1}^{d} C^{-1}_{ik}u_ku_i - t^{-1}d\right)\\ &= \frac{G_t(u)}{2}(t^{-2}u^TC^{-1}u - t^{-1}d) = \partial_tG_t(u).\end{aligned}$$

Let us summarize our results:

Theorem 4.8. *Let $\rho_0 \in C^2(\mathbb{T}^d, \mathbb{R}_+)$ and $s(N) = N$ if $v \neq 0$ and $s(N) = N^2$ if $v = 0$ and $\det(C) \neq 0$. If we start η^N in $\nu^N_{\rho_0}$, then $(\mathrm{law}(\eta^N_{s(N)t}))$ is a local equilibrium sequence with respect to the profile $\rho(t, \cdot)$, where ρ is the unique solution to the transport equation (4.1) if $v \neq 0$ and to the heat equation (4.2) if $v = 0$ and $\det(C) \neq 0$.*

Proof. This follows by combining Lemma 4.4 with Lemma 4.6 respectively Lemma 4.7. □

Theorem 4.8 gives us a local description how $\eta^N_{s(N)t}([uN])$ behaves for a fixed macroscopic point $u \in \mathbb{T}^d$. We can also study the *empirical measure* on $\mathbb{T}^d$, defined by

$$\pi^N_t := \frac{1}{N^d}\sum_{x \in \mathbb{Z}^d_N} \delta_{x/N}\eta^N_t(x), \tag{4.3}$$

where $\delta_{x/N}$ denotes the Dirac delta in x/N. Note that $N^d = |\mathbb{Z}^d_N|$, so the normalization constant in (4.3) is just dividing by the number of points. For $\varphi: \mathbb{T}^d \to \mathbb{R}$ we have

$$\pi^N_t(\varphi) := \int_{\mathbb{T}^d}\varphi(u)\pi^N_t(\mathrm{d}u) = \frac{1}{N^d}\sum_{x \in \mathbb{Z}^d_N}\varphi\Big(\frac{x}{N}\Big)\eta^N_t(x).$$

The empirical measure π^N_t completely encodes the configuration η^N_t, but by averaging by integrating against a test function φ rather than keeping track of the behavior in one point we obtain a slightly different description in the limit:

Proposition 4.9. *In the setting of Theorem 4.8 we have for $\varphi \in C(\mathbb{T}^d)$*

$$\lim_{N\to\infty}\pi^N_{s(N)t}(\varphi) = \int_{\mathbb{T}^d}\varphi(u)\rho(t,u)\mathrm{d}u,$$

both if $v \neq 0$ and if $v = 0$ and $\det(C) \neq 0$.

Proof. Taking expectations, we have since $\eta^N_{s(N)t}(x)$ is Poisson distributed:

$$\mathbb{E}_{\nu^N_{\rho_0}}[\pi^N_{s(N)t}(\varphi)] = \frac{1}{N^d}\sum_{x \in \mathbb{Z}^d_N}\varphi\Big(\frac{x}{N}\Big)\mathbb{E}_{\nu^N_{\rho_0}}[\eta^N_{s(N)t}(x)] = \frac{1}{N^d}\sum_{x \in \mathbb{Z}^d_N}\varphi\Big(\frac{x}{N}\Big)\psi^N_{s(N)t}(x),$$

and since φ is bounded and $x = \left[\frac{x}{N}N\right]$ we have

$$\left|\frac{1}{N^d}\sum_{x \in \mathbb{Z}^d_N}\varphi\Big(\frac{x}{N}\Big)\psi^N_{s(N)t}\Big(\Big[\frac{x}{N}N\Big]\Big) - \frac{1}{N^d}\sum_{x \in \mathbb{Z}^d_N}\varphi\Big(\frac{x}{N}\Big)\rho\Big(t,\frac{x}{N}\Big)\right| \leqslant \|\varphi\|_\infty\|\psi^N_{s(N)t}([\cdot N]) - \rho(t,\cdot)\|_\infty,$$

which converges to zero. On the other hand, since ρ and φ are continuous we have

$$\lim_{N\to\infty}\frac{1}{N^d}\sum_{x \in \mathbb{Z}^d_N}\varphi\Big(\frac{x}{N}\Big)\rho\Big(t,\frac{x}{N}\Big) = \int_{\mathbb{T}^d}\varphi(u)\rho(t,u)\mathrm{d}u,$$

because the left hand side is a Riemann sum. This proves that the expectation of $\pi^N_{s(N)t}(\varphi)$ converges to the claimed limit. It remains to show that in the limit $\pi^N_{s(N)t}(\varphi)$ does not deviate its expectation. But since $\eta^N_{s(N)t}$ is product Poisson distributed we get

$$\begin{aligned}\mathrm{Var}_{\nu^N_{\rho_0}}(\pi^N_{s(N)t}(\varphi)) &= \mathrm{Var}_{\nu^N_{\rho_0}}\left(\frac{1}{N^d}\sum_{x\in\mathbb{Z}^d_N}\varphi\left(\frac{x}{N}\right)\eta^N_{s(N)t}(x)\right)\\ &= \frac{1}{N^{2d}}\sum_{x\in\mathbb{Z}^d_N}\varphi\left(\frac{x}{N}\right)^2\mathrm{Var}_{\nu^N_{\rho_0}}(\eta^N_{s(N)t}(x))\\ &= \frac{1}{N^{2d}}\sum_{x\in\mathbb{Z}^d_N}\varphi\left(\frac{x}{N}\right)^2\psi^N_{s(N)t}(x) \leqslant N^{-d}\|\varphi\|^2_\infty\|\psi^N_{s(N)t}\|_\infty \longrightarrow 0,\end{aligned}$$

proving the claim. □

We could improve this and show that the measure valued process $(\pi^N_{s(N)t})_{t\geqslant 0}$ converges in the space of càdlàg functions with values in measures on $\mathbb{T}^d$, with respect to the locally uniform topology. We will derive such results in the following chapters and in particular introduce the necessary tools for that, but for the independent random walks we leave the result as it is and rather continue with more interesting problems.

5 Hydrodynamics for symmetric exclusion processes

5.1 Heuristic derivation

When we derived the explicit distribution of the independent random walkers started in a product Poisson configuration with varying density (Lemma 4.5), we crucially used the independence of the individual walkers. For the exclusion process the distribution at later times is not so obvious, not even if the initial distribution is a product Bernoulli distribution with space-varying parameter ρ (except in the special case where ρ is constant and therefore μ_ρ is invariant). We therefore need another argument to derive the hydrodynamic limit of the simple exclusion process.

Notation. *Recall that for all $\rho\in[0,1]$ the product Bernoulli measure μ^N_ρ on $\{0,1\}^{\mathbb{Z}^d_N}$ with density ρ is invariant for η^N. If $\rho\colon\mathbb{T}^d\to[0,1]$ is a function rather than a constant, then we also write μ^N_ρ for the product measure with*

$$\mu^N_\rho(\eta) = \prod_{x\in\mathbb{Z}^d_N}\rho\left(\frac{x}{N}\right)^{\eta(x)}\left(1-\rho\left(\frac{x}{N}\right)\right)^{1-\eta(x)}.$$

For the exclusion process we will not study the local behavior, but start right away with the "more global" empirical measure, defined as for the independent walkers by

$$\pi^N_t := \frac{1}{N^d}\sum_{x\in\mathbb{Z}^d_N}\delta_{x/N}\eta^N_t(x),$$

where η^N is an exclusion process with Q-matrix

$$q^N f(\eta) = \sum_{x,y\in\mathbb{Z}^d_N}\eta(x)(1-\eta(y))p^N(x,y)(f(\eta^{x,y})-f(\eta)).$$

As before, we have for all $\varphi\colon\mathbb{T}^d\to\mathbb{R}$

$$\pi^N_t(\varphi) = \frac{1}{N^d}\sum_{x\in\mathbb{Z}^d_N}\varphi\left(\frac{x}{N}\right)\eta^N_t(x),$$

and therefore (since $\pi^N_t(\varphi)$ is a function of η^N_t)

$$\pi^N_{s(N)t}(\varphi) = \pi^N_0(\varphi) + \int_0^{s(N)t} q^N\pi^N_r(\varphi)\mathrm{d}r + M^{N,\varphi}_{s(N)t}, \tag{5.1}$$

where $s(N)$ is the speed in which we let the process evolve and where $M^{N,\varphi}_{s(N)t}$ is a martingale.

Lemma 5.1. *For the map* $\eta \mapsto (\pi^N(\varphi))(\eta) := \frac{1}{N^d}\sum_{x\in\mathbb{Z}_N^d}\varphi\left(\frac{x}{N}\right)\eta(x)$ *we have*

$$q^N\pi^N(\varphi) = \frac{1}{N^{d+1}}\sum_{x,y\in\mathbb{Z}_N^d}[\nabla^N_{x,y}\varphi]\eta(x)(1-\eta(x+y))p^N(y), \tag{5.2}$$

where

$$\nabla^N_{x,y}\varphi := N\Big[\varphi\Big(\frac{x+y}{N}\Big) - \varphi\Big(\frac{x}{N}\Big)\Big] \simeq \nabla\varphi\Big(\frac{x}{N}\Big)\cdot y. \tag{5.3}$$

If p^N *is symmetric, i.e.* $p^N(y) = p^N(-y)$ *for all* $y \in \mathbb{Z}_N^d$*, then the above simplifies to*

$$q^N\pi^N(\varphi) = \frac{1}{2N^{d+2}}\sum_{x,y\in\mathbb{Z}_N^d}[\Delta^N_{x,y}\varphi]\eta(x)p^N(y) = \frac{1}{2N^2}\sum_{y\in\mathbb{Z}_N^d}p^N(y)\pi^N(\Delta^N_{N\cdot,y}\varphi), \tag{5.4}$$

where

$$\Delta^N_{x,y}\varphi := N^2\Big[\varphi\Big(\frac{x+y}{N}\Big) + \varphi\Big(\frac{x-y}{N}\Big) - 2\varphi\Big(\frac{x}{N}\Big)\Big] \simeq y^T\mathrm{D}^2\varphi\Big(\frac{x}{N}\Big)y. \tag{5.5}$$

Proof. If $\eta(x)(1-\eta(y)) \neq 0$, then $\eta(x) = 1$ and $\eta(y) = 0$ and therefore $\eta^{x,y}(x) = \eta(x) - 1$ and $\eta^{x,y}(y) = \eta(y) + 1$, and thus

$$\begin{aligned}
q^N(\pi^N(\varphi))(\eta) &= \sum_{x,y}\eta(x)(1-\eta(y))p^N(y-x)\Bigg(\frac{1}{N^d}\sum_z\varphi\Big(\frac{z}{N}\Big)\eta^{x,y}(z) - \frac{1}{N^d}\sum_z\varphi\Big(\frac{z}{N}\Big)\eta(z)\Bigg)\\
&= \sum_{x,y}\eta(x)(1-\eta(y))p^N(y-x)\Bigg(\frac{1}{N^d}\sum_z\varphi\Big(\frac{z}{N}\Big)\eta(z) + \frac{1}{N^d}\Big(\varphi\Big(\frac{y}{N}\Big)-\varphi\Big(\frac{x}{N}\Big)\Big) - \frac{1}{N^d}\sum_z\varphi\Big(\frac{z}{N}\Big)\eta(z)\Bigg)\\
&= \frac{1}{N^d}\sum_{x,y}\eta(x)(1-\eta(y))p^N(y-x)\Big(\varphi\Big(\frac{y}{N}\Big)-\varphi\Big(\frac{x}{N}\Big)\Big)\\
&= \frac{1}{N^{d+1}}\sum_{x,y\in\mathbb{Z}_N^d}\eta(x)(1-\eta(y))p^N(y-x)[\nabla^N_{x,y-x}\varphi]\\
&= \frac{1}{N^{d+1}}\sum_{x,y\in\mathbb{Z}_N^d}[\nabla^N_{x,y}\varphi]\eta(x)(1-\eta(x+y))p^N(y).
\end{aligned}$$

If p^N is symmetric, we have furthermore

$$\begin{aligned}
\sum_{x,y\in\mathbb{Z}_N^d}[\nabla^N_{x,y}\varphi]\eta(x)\eta(x+y)p^N(y) &= \sum_{x,y\in\mathbb{Z}_N^d}[\nabla^N_{x,-y}\varphi]\eta(x)\eta(x-y)p^N(-y)\\
&\overset{z=x-y}{=\!=} \sum_{z,y\in\mathbb{Z}_N^d}[-\nabla^N_{z,y}\varphi]\eta(z+y)\eta(z)p^N(-y)\\
&\overset{p^N(-y)=p^N(y)}{=\!=\!=} \sum_{z,y\in\mathbb{Z}_N^d}[-\nabla^N_{z,y}\varphi]\eta(z+y)\eta(z)p^N(y) = 0,
\end{aligned}$$

where in the last step we used that we can rename z to x and therefore the sum is equal to minus itself. We remain with

$$\begin{aligned}
\frac{1}{N^{d+1}}\sum_{x,y\in\mathbb{Z}_N^d}[\nabla^N_{x,y}\varphi]\eta(x)p^N(y) &= \frac{1}{2N^{d+1}}\sum_{x,y\in\mathbb{Z}_N^d}\{[\nabla^N_{x,y}\varphi]\eta(x)p^N(y) + [\nabla^N_{x,-y}\varphi]\eta(x)p^N(-y)\}\\
&\overset{p^N(-y)=p^N(y)}{=\!=\!=} \frac{1}{2N^{d+2}}\sum_{x,y\in\mathbb{Z}_N^d}[\Delta^N_{x,y}\varphi]\eta(x)p^N(y),
\end{aligned}$$

which concludes the proof. □

Consequently, we obtain for <u>symmetric</u> p^N

$$\pi^N_{s(N)t}(\varphi) = \pi^N_0(\varphi) + s(N)\int_0^t\frac{1}{2N^2}\sum_{y\in\mathbb{Z}_N^d}p^N(y)\pi^N_{s(N)r}(\Delta^N_{N\cdot,y}\varphi)\mathrm{d}r + M^{N,\varphi}_{s(N)t},$$

and for $s(N)=N^2$ this simplifies to

$$\pi^N_{N^2t}(\varphi)=\pi^N_0(\varphi)+\int_0^t \frac{1}{2}\sum_{y\in\mathbb{Z}^d_N} p^N(y)\pi^N_{N^2r}(\Delta^N_{\cdot,y}\varphi)\mathrm{d}r+M^{N,\varphi}_{N^2t},$$

and replacing $\Delta^N_{x,y}\varphi$ formally by $y^T\mathrm{D}^2\varphi\big(\frac{x}{N}\big)y$ we get

$$\begin{aligned}\pi^N_{N^2t}(\varphi)&\sim\pi^N_0(\varphi)+\int_0^t \frac{1}{2}\pi^N_{N^2r}\left(\sum_{y\in\mathbb{Z}^d_N} p^N(y)y^T\mathrm{D}^2\varphi(\cdot)y\right)\mathrm{d}r+M^{N,\varphi}_{N^2t},\\ &=\pi^N_0(\varphi)+\int_0^t \frac{1}{2}\pi^N_{N^2r}\left(\sum_{i,j=1}^d\sum_{y\in\mathbb{Z}^d_N} p^N(y)y_iy_j\partial_{u_iu_j}\varphi(\cdot)\right)\mathrm{d}r+M^{N,\varphi}_{N^2t}\\ &=\pi^N_0(\varphi)+\int_0^t \frac{1}{2}\pi^N_{N^2r}(\Delta_C\varphi)\mathrm{d}r+M^{N,\varphi}_{N^2t}.\end{aligned}$$

If we could show now that the martingale $M^{N,\varphi}_{N^2t}$ converges to zero and that $(\pi^N_{N^2t})_{t\geqslant 0}$ converges to some measure valued process $(\pi_t)_{t\geqslant 0}$, then we would obtain from the formal arguments above that

$$\pi_t(\varphi)=\pi_0(\varphi)+\int_0^t\frac{1}{2}\pi_r(\Delta_C\varphi)\mathrm{d}r$$

for all $\varphi\in C^2(\mathbb{T}^d)$, which shows that π is a "weak solution" to the heat equation $\partial_t\pi=\frac{1}{2}\Delta_C\pi$, so the hydrodynamic limit of the empirical measures $(\pi^N_{N^2t})_{t\geqslant 0}$ is given by the heat equation.

To prove this, we need a tool to deal with the martingale $M^{N,\varphi}$, and we need to make the heuristic argumentation above rigorous. The next lemma allows us to control M:

Lemma 5.2. *Let $(X_t)_{t\geqslant 0}$ be a Markov process on a countable state space S, with Q-matrix q and $\|c\|_\infty<\infty$, and let $f\in\mathcal{M}(S)$. Then*

$$M^f_t=f(X_t)-f(X_0)-\int_0^t qf(X_s)\mathrm{d}s,\qquad t\geqslant 0,$$

and

$$|M^f_t|^2-\int_0^t((qf^2)(X_s)-2f(X_s)qf(X_s))\mathrm{d}s,\qquad t\geqslant 0,$$

are martingales. We also write

$$\langle M^f\rangle_t:=\int_0^t((qf^2)(X_s)-2f(X_s)qf(X_s))\mathrm{d}s.$$

Remark 5.3. In stochastic analysis we encountered the quadratic variation of a continuous local martingale N, given by $\langle N\rangle_t=\lim_{N\to\infty}\sum_{k=0}^{N-1}(N_{t(k+1)/N}-N_{tk/N})^2$, or alternatively defined as the unique increasing and adapted process for which $N^2-\langle N\rangle$ is a local martingale.

But in our setting $\langle M^f\rangle_t\neq\lim_{N\to\infty}\sum_{k=0}^{N-1}(M^f_{t(k+1)/N}-M^f_{tk/N})^2$. Indeed, since M^f is the sum of an absolutely continuous process and a process that performs finitely many jumps, we have

$$\lim_{N\to\infty}\sum_{k=0}^{N-1}(M^f_{t(k+1)/N}-M^f_{tk/N})^2=\sum_{s\in[0,t]}(\Delta M^f_s)^2\neq\langle M^f\rangle_t,$$

where $\Delta\Gamma_r=\Gamma_r-\lim_{s\uparrow r}\Gamma_s$ is the jump of the right-continuous function Γ in r (convince yourself of this!). The explanation is that there exist two types of quadratic variation for "locally square integrable" local martingales:

$$[N]_t:=\lim_{N\to\infty}\sum_{k=0}^{N-1}(N_{t(k+1)/N}-N_{tk/N})^2$$

and $\langle N\rangle$ is the unique increasing predictable process for which $[N]-\langle N\rangle$ is a local martingale, and therefore often called the *predictable quadratic variation of* N. Since for continuous N also $[N]$ is continuous and therefore predictable, we have $[N]=\langle N\rangle$, but in general this identity fails as the example here shows (see [Jacod, Shiryaev] or [Protter] for details).

Proof. To slightly simplify the notation we assume that X_0 is deterministic (otherwise we could argue conditionally on $\mathcal{F}_0$ or simply write a bit more complicated equations). Then we can set $\tilde{f} = f - f(X_0)$ and thus assume without loss of generality that $f(X_0) = 0$. We only need to show the claim about $\langle M^f \rangle$ since we know from the martingale problem that M^f is a martingale:

$$
\begin{aligned}
|M_t^f|^2 &= f^2(X_t) - 2f(X_t)\int_0^t qf(X_s)\mathrm{d}s + \left(\int_0^t qf(X_s)\mathrm{d}s\right)^2 \\
&= \int_0^t qf^2(X_s)\mathrm{d}s - 2f(X_t)\int_0^t qf(X_s)\mathrm{d}s + \left(\int_0^t qf(X_s)\mathrm{d}s\right)^2 + \left(f^2(X_t) - \int_0^t qf^2(X_s)\mathrm{d}s\right) \\
&= \int_0^t qf^2(X_s)\mathrm{d}s - 2M_t^f\int_0^t qf(X_s)\mathrm{d}s - \left(\int_0^t qf(X_s)\mathrm{d}s\right)^2 + \text{mart.},
\end{aligned}
$$

where we write "mart." for a martingale which may change from line to line (but will always stay a martingale). Now

$$
N_t^f := M_t^f\int_0^t qf(X_s)\mathrm{d}s - \int_0^t M_s^f qf(X_s)\mathrm{d}s
$$

is a martingale: This follows from the integration by parts rule for the finite variation jump process M^f and the finite variation absolutely continuous process $\int_0^t qf(X_s)\mathrm{d}s$, which allows us to write N^f as a stochastic integral with respect to a martingale; but to be self-contained we give a direct proof:

$$
\begin{aligned}
\mathbb{E}[N_t^f|\mathcal{F}_r] &= M_r^f\int_0^r qf(X_s)\mathrm{d}s + \mathbb{E}\left[M_t^f\int_r^t qf(X_s)\mathrm{d}s\middle|\mathcal{F}_r\right] \\
&\quad - \int_0^r M_s^f qf(X_s)\mathrm{d}s - \mathbb{E}\left[\int_r^t M_s^f qf(X_s)\mathrm{d}s\middle|\mathcal{F}_r\right] \\
&= N_r^f + \int_r^t \mathbb{E}[M_t^f qf(X_s)|\mathcal{F}_r]\mathrm{d}s - \int_r^t \mathbb{E}[M_s^f qf(X_s)|\mathcal{F}_r]\mathrm{d}s \\
&= N_r^f + \int_r^t \mathbb{E}[\mathbb{E}[M_t^f qf(X_s)|\mathcal{F}_s]|\mathcal{F}_r]\mathrm{d}s - \int_r^t \mathbb{E}[M_s^f qf(X_s)|\mathcal{F}_r]\mathrm{d}s = N_r^f.
\end{aligned}
$$

This leads to

$$
\begin{aligned}
|M_t^f|^2 &= \int_0^t qf^2(X_s)\mathrm{d}s - 2\int_0^t M_s^f qf(X_s)\mathrm{d}s - \left(\int_0^t qf(X_s)\mathrm{d}s\right)^2 + \text{mart.} \\
&= \int_0^t qf^2(X_s)\mathrm{d}s - 2\int_0^t f(X_s)qf(X_s)\mathrm{d}s + 2\int_0^t\int_0^s qf(x_r)\mathrm{d}r\, qf(X_s)\mathrm{d}s - \left(\int_0^t qf(X_s)\mathrm{d}s\right)^2 + \text{mart.},
\end{aligned}
$$

and integrating the term $(\int_0^t qf(X_s)\mathrm{d}s)^2$ by parts we see that it cancels with the third term on the right hand side. Thus, we finally have

$$
|M_t^f|^2 = \int_0^t [qf^2(X_s) - 2f(X_s)qf(X_s)]\mathrm{d}s + \text{mart.},
$$

which is what we claimed. □

Let us apply this lemma to bound the martingale $M^{N,\varphi}$ that appears in (5.1):

$$
\mathbb{E}[(M_t^{N,\varphi})^2] = \mathbb{E}[\langle M^{N,\varphi}\rangle_t] = \mathbb{E}\left[\int_0^t (q^N(\pi_s^N(\varphi)^2) - 2\pi_s^N(\varphi)q^N\pi_s^N(\varphi))\mathrm{d}s\right]. \tag{5.6}
$$

For this to be useful we need to compute $q^N(\pi^N(\varphi)^2) - 2\pi^N(\varphi)q^N\pi^N(\varphi)$. Since we will need this also later in the rigorous proof, let us write a lemma for this computation:

Lemma 5.4. *For the map* $\eta \mapsto (\pi^N(\varphi))(\eta) := \frac{1}{N^d}\sum_{x\in\mathbb{Z}_N^d}\varphi\left(\frac{x}{N}\right)\eta(x)$ *we have*

$$
q^N(\pi^N(\varphi)^2) - 2\pi^N(\varphi)q^N\pi^N(\varphi) = \frac{1}{N^{2d+2}}\sum_{x,y}\eta(x)(1-\eta(x+y))p^N(y)[\nabla_{x,y}^N\varphi]^2. \tag{5.7}
$$

Proof. Let us start by computing the action of q^N on $\pi^N(\varphi)^2$: If $\eta(x)(1-\eta(y))\neq 0$, then $\eta(x)=1$ and $\eta(y)=0$ and therefore $\eta^{x,y}(x)=\eta(x)-1$ and $\eta^{x,y}(y)=\eta(y)+1$, and thus

$$\begin{aligned}&\eta(x)(1-\eta(y))\left(\frac{1}{N^d}\sum_z \varphi\left(\frac{z}{N}\right)\eta^{x,y}(z)\right)^2\\ =&\eta(x)(1-\eta(y))\left(\frac{1}{N^d}\sum_z \varphi\left(\frac{z}{N}\right)\eta(z)+\frac{1}{N^d}\left(\varphi\left(\frac{y}{N}\right)-\varphi\left(\frac{x}{N}\right)\right)\right)^2\\ =&\eta(x)(1-\eta(y))\left(\frac{1}{N^d}\sum_z \varphi\left(\frac{z}{N}\right)\eta(z)\right)^2+2\left(\frac{1}{N^d}\sum_z \varphi\left(\frac{z}{N}\right)\eta(z)\right)\frac{1}{N^d}\left(\varphi\left(\frac{y}{N}\right)-\varphi\left(\frac{x}{N}\right)\right)\\ &+\left(\frac{1}{N^d}\left(\varphi\left(\frac{y}{N}\right)-\varphi\left(\frac{x}{N}\right)\right)\right)^2.\end{aligned}$$

From here we get

$$\begin{aligned}q^N(\pi^N(\varphi)^2)=&\sum_{x,y}\eta(x)(1-\eta(y))p^N(y-x)\left(2\left(\frac{1}{N^d}\sum_z \varphi\left(\frac{z}{N}\right)\eta(z)\right)\frac{1}{N^d}\left(\varphi\left(\frac{y}{N}\right)-\varphi\left(\frac{x}{N}\right)\right)\right)\\ &+\sum_{x,y}\eta(x)(1-\eta(y))p^N(y-x)\left(\frac{1}{N^d}\left(\varphi\left(\frac{y}{N}\right)-\varphi\left(\frac{x}{N}\right)\right)\right)^2\\ =&2\pi^N(\varphi)q^N\pi^N(\varphi)+\frac{1}{N^{2d+2}}\sum_{x,y}\eta(x)(1-\eta(y))p^N(y-x)[\nabla^N_{x,y-x}\varphi]^2,\end{aligned}$$

and thus our claim follows. □

From equation (5.7) together with (5.6) we get

$$\mathbb{E}[(M_t^{N,\varphi})^2]=\frac{1}{N^{2d+2}}\sum_{x,y}\int_0^t \mathbb{E}[\eta_s^N(x)(1-\eta_s^N(x+y))p^N(y)[\nabla^N_{x,y}\varphi]^2]\mathrm{d}s,$$

and since $[\nabla^N_{x,y}\varphi]^2\lesssim 1$ and also $|\eta_s^N(x)(1-\eta_s^N(x+y))|\leqslant 1$ (uniformly in ω), we get

$$\mathbb{E}[(M_t^{N,\varphi})^2]\lesssim\frac{1}{N^{2d+2}}\sum_{x,y}tp^N(y)\lesssim t\frac{1}{N^{2d+2}}\sum_{x\in\mathbb{Z}_N^d}1\lesssim tN^{-d-2}.$$

So we can control $M_t^{N,\varphi}$ for all $t\ll N^{d+2}$ and in particular for $t=N^2$. This completes our heuristic argument. Our next goal is make the formal derivation above rigorous.

Remark 5.5. We did not say anything about the asymmetric case. If $v=\sum_x xp(x)\neq 0$ (in $\mathbb{T}^d$) then we have on the slower time scale $s(N)=N$

$$\begin{aligned}\pi^N_{Nt}(\varphi)=&\pi^N_0(\varphi)+\int_0^t Nq^N\pi^N_{Nr}(\varphi)\mathrm{d}r+M^{N,\varphi}_{Nt}\\ =&\pi^N_0(\varphi)+\int_0^t\frac{1}{N^d}\sum_{x,y\in\mathbb{Z}_N^d}[\nabla^N_{x,y}\varphi]\eta^N_{Nr}(x)(1-\eta^N_{Nr}(x+y))p^N(y)\mathrm{d}r+M^{N,\varphi}_{Nt}.\end{aligned}$$

The martingale vanishes as $N\to\infty$ (the arguments for that did not use that p^N is symmetric), but the difference to the symmetric case is that we cannot write the drift as a function of the empirical measure π^N. However, arguing purely formal we get (assuming some "continuity" of η)

$$\begin{aligned}&\int_0^t\frac{1}{N^d}\sum_{x,y\in\mathbb{Z}_N^d}[\nabla^N_{x,y}\varphi]\eta^N_{Nr}(x)(1-\eta^N_{Nr}(x+y))p^N(y)\mathrm{d}r\\ \sim&\int_0^t\frac{1}{N^d}\sum_{x,y\in\mathbb{Z}_N^d}[\nabla^N_{x,y}\varphi]\eta^N_{Nr}(x)(1-\eta^N_{Nr}(x))p^N(y)\mathrm{d}r\\ \sim&\int_0^t\pi^N\left(\sum_y p^N(y)\nabla^N_{\cdot,y}\varphi\right)[\eta^N_{Nr}(1-\eta^N_{Nr})]\mathrm{d}r\\ \sim&\int_0^t\pi^N(v\cdot\nabla\varphi)[\eta^N_{Nr}(1-\eta^N_{Nr})]\mathrm{d}r,\end{aligned}$$

so that we would expect the limiting measure π to be a weak solution to the *Burgers equation*

$$\partial_t \pi = v \cdot \nabla(\pi(1-\pi)).$$

However, to proof this is much more involved than the proof in the symmetric case, and another problem is that Burgers equation has more than one weak solution, so we would still need an argument to show that $\pi^N_{N\cdot}$ converges to the "right" solution. Maybe we will get to this later, for now we focus on the symmetric case.

Strategy for the proof. Here we describe the strategy to show that (in the symmetric case) $\pi^N_{N^2\cdot}$ converges to the solution π of the heat equation $\partial_t \pi = \frac{1}{2}\Delta_C \pi$. This general strategy actually applies to a huge class of problems and is the standard approach for deriving weak limits of stochastic processes. It rests on the following steps:

1. **Show tightness**
   We will work in the space $D([0,T],\mathcal{M}_+)$, where $\mathcal{M}_+$ is the space of finite (positive) measures on $\mathbb{T}^d$, equipped with the weak topology, and $D([0,T],\mathcal{M}_+)$ is the space of càdlàg functions from $[0,T]$ to $\mathcal{M}_+$, equipped with the *Skorohod topology* (which we will define below).
2. **Show that every limit point has certain features**
   We will show that any limiting point is a weak solution of the heat equation $\partial_t \pi = \frac{1}{2}\Delta_C \pi$, started in $\pi_0 = \lim_N \pi^N_0$, and that any limiting point is absolutely continuous.
3. **Show that these features characterize a unique process**
   We will show that for every initial condition there is a unique absolutely continuous measure which solves the heat equation in the weak sense.

Combining 1.-3. the convergence follows.

5.2 Tightness in the Skorohod space

Here we present the main tools for proving tightness of measures supported on the Skorohod space of càdlàg functions with values in a Polish (i.e. complete and separable metric) space. Ultimately we are interested in the space of positive measures on $\mathbb{T}^d$, but since most results that follow hold for general Polish spaces, we state them in the general case. Also, we mostly skip the proofs and refer to the literature (an excellent reference on these questions is Chapter 3 of [Ethier, Kurtz, *Markov Processes. Characterization and Convergence*, Wiley 1986]).

Exercise 5.1. To see that we are really in this general setting, show that if K is a compact set and $(f_k)_{k\in\mathbb{N}}$ is a sequence of dense functions in $(C(K),\|\cdot\|_\infty)$, then

$$\delta(\pi,\nu) := \sum_{k=1}^{\infty} 2^{-k}(|\pi(f_k)-\nu(f_k)|\wedge 1)$$

defines a metric on $\mathcal{M}_+$, the space of finite (positive) measures on K. Next, show that a sequence $(\pi_n)\subset\mathcal{M}_+$ converges weakly to $\pi\in\mathcal{M}_+$ if and only if $\lim_{n\to\infty}\delta(\pi_n,\pi)=0$.

Solution. *It is easy to see that δ is symmetric and satisfies the triangle inequality. If $\delta(\pi,\nu)=0$, then $\pi(f_k)=\nu(f_k)$ for all k. If $f\in C(K)$ is an arbitrary continuous function and $\varepsilon>0$, then there exists k with $\|f_k-f\|_\infty<\varepsilon$, and therefore*

$$|\pi(f)-\nu(f)| \leqslant |\pi(f-f_k)| + |\pi(f_k)-\nu(f_k)| + |\nu(f_k-f)| \leqslant |\pi(1)|\varepsilon + |\nu(1)|\varepsilon,$$

so for $\varepsilon\to 0$ we see that $\pi(f)=\nu(f)$. Since this holds for all $f\in C(K)$, we must have $\pi=\nu$.

Next, we need to show that convergence with respect to δ is equivalent to convergence in the weak topology. Obviously we have $\lim_{n\to\infty}\delta(\pi_n,\pi)=0$ whenever π_n converges weakly to π, because then $|\pi_n(f_k)-\pi(f_k)|$ converges to zero for all k. Conversely, let $\lim_{n\to\infty}\delta(\pi_n,\pi)=0$ and let $f\in C(K)$. Then for all k

$$\begin{aligned}|\pi_n(f)-\pi(f)| &\leqslant |\pi_n(f-f_k)| + |\pi_n(f_k)-\pi(f_k)| + |\pi(f_k-f)|\\ &\leqslant (\pi_n(1)+\pi(1))\|f-f_k\|_\infty + 2^k\delta(\pi_n,\pi).\end{aligned}$$

Assume for now that the sequence $(\pi_n(1))_n$ *is bounded. Then*

$$\limsup_{n\to\infty} |\pi_n(f) - \pi(f)| \lesssim \limsup_{n\to\infty} [\|f - f_k\|_\infty + 2^k \delta(\pi_n, \pi)] = \|f - f_k\|_\infty$$

for all k*, and since the right hand side can be made arbitrarily small because* (f_k) *is dense, we obtain that* $\pi_n(f)$ *converges to* $\pi(f)$*, i.e.* π_n *converges weakly to* π*. It remains to show that* $(\pi_n(1))$ *is bounded. For that purpose let* $k \in \mathbb{N}$ *be such that* $\gamma = \inf_{x\in\mathbb{T}^d} f_k > 0$ *(which e.g. holds as soon as* $\|f_k - 1\|_\infty < 1/2$*). Then*

$$0 \leqslant \pi_n(1) = \frac{1}{\gamma}\pi_n(\gamma) \leqslant \frac{1}{\gamma}\pi_n(f_k),$$

and since $(\pi_n(f_k))_n$ *converges, it must be a bounded sequence so also* $(\pi_n(1))_n$ *is bounded.*

In the following we let (E, δ) be a Polish space and we consider the two spaces

$$C([0,T], E) = \{\varphi: [0,T] \to E, \varphi \text{ is continuous}\}$$

and the Skorohod space

$$D([0,T], E) = \{\varphi: [0,T] \to E, \varphi \text{ is càdlàg}\},$$

where we recall that φ is called *càdlàg*, sometimes also *RCLL (right continuous with left limits)* if for all $t > 0$ the limit $\lim_{s\uparrow t}\varphi(s)$ exists and for all $t \geqslant 0$ we have $\varphi(t) = \lim_{s\downarrow t}\varphi(s)$.

Exercise 5.2. For $\varphi \in D([0,T], E)$ and $t > 0$ we write $\varphi(t-) := \lim_{s\uparrow t}\varphi(t)$. Show that there exist at most countably many $t \in [0,T]$ with $\varphi(t-) \neq \varphi(t)$ (i.e. φ has only countably many jumps on $[0,T]$).

Solution. *For all* $\varepsilon > 0$ *the set* $A_\varepsilon = \{t \in (0,T]: \delta(\varphi(t-), \varphi(t)) \geqslant \varepsilon\}$ *is finite, because otherwise we could find a sequence of strictly increasing* $(t_k)_{k\in\mathbb{N}}$ *with* $t_k \in A_\varepsilon$ *for all* k*. Then* $t^* = \lim_{k\to\infty} t_k \in (0,T]$ *and for every* $\gamma > 0$ *there exists* $t_k, t_k' \in (t^* - \gamma, t^*)$ *such that* $\delta(\varphi(t_k'), \varphi(t_k)) \geqslant \varepsilon$*, which means that* $\lim_{s\uparrow t}\varphi(s)$ *does not exist, a contradiction to* φ *being càdlàg! Therefore,* A_ε *is finite, and the set of jumps of* φ *is* $\bigcup_{n\in\mathbb{N}} A_{1/n}$ *and thus countable.*

Since we want to consider measures on these spaces, we need to equip them with σ-algebras. These will always be Borel σ-algebras, so we need appropriate topologies. A natural choice is the uniform topology, defined from the metric

$$d_\infty(\varphi, \psi) := \sup_{t\in[0,T]} \delta(\varphi(t), \psi(t)),$$

and indeed we will always use this metric on $C([0, T], E)$, and $(C([0, T], E), d_\infty)$ is a complete separable metric space whenever E is a vector space: To show completeness first use that E is complete, so if (φ_n) is a Cauchy sequence with respect to d_∞, then for every $t \in [0,T]$ there exists $\varphi(t)$ with $\delta(\varphi_n(t), \varphi(t)) \to 0$; then show that $\varphi \in C([0, T], E)$ and that $d_\infty(\varphi_n, \varphi) \to 0$. To show separability we have to use that E is a vector space, take a set of dense points $(x_k)_{k\in\mathbb{N}} \subset E$ and consider all functions φ for which there exist $m \in \mathbb{N}$ and rational numbers $0 = t_0 < \ldots < t_m = T$ and $k(1), \ldots, k(m) \in \mathbb{N}$ such that $\varphi(t_i) = x_{k(i)}$ for $i = 1, \ldots, m$ and such that φ is piecewise linear between t_i and t_{i+1}; it is easy to show that the functions of this form are dense, and of course there are only countably many of them.

However, for $D([0,T], E)$ we have a problem: d_∞ still makes this space complete, but now it is no longer separable:

Example 5.6. Let $E = \mathbb{R}$ and define for $s \in [0,T]$ the function $\varphi_s(t) = \mathbb{I}_{[s,T]}(t) \in D([0,T], \mathbb{R})$. Then for $s \neq s'$

$$d_\infty(\varphi_s, \varphi_{s'}) = 1,$$

but $(\varphi_s)_{s\in[0,T]}$ is an uncountable family of functions. This shows that $(D([0, T], \mathbb{R}), d_\infty)$ is not separable (it is a good exercise to convince yourself of this).

But it is very important for us to work with measures on Polish spaces. The reason for this is Prohorov's criterion (which is known from Stochastics II):

Lemma 5.7. *(Prohorov's criterion)*

Let $(\mathbb{P}_i)_{i\in I}$ be a family of probability measures on (the Borel sets of) a Polish space F. Then $(\mathbb{P}_i)_{i\in I}$ is relatively compact with respect to the weak topology if and only if it is tight, i.e. for all $\varepsilon>0$ there exists a compact set $K_\varepsilon\subset F$ such that

$$\sup_{i\in I}\mathbb{P}_i(K_\varepsilon^c)\leqslant\varepsilon.$$

So to make $D([0,T],E)$ a Polish space we should find a weaker metric, under which it is easier to be dense (but under which the space still remains complete). Example 5.6 provides us with some intuition how that metric should look like: If $s_n\to s$, then the functions φ_{s_n} "look more and more like" φ_s, and ideally we would want them to converge to φ_s, because then the counterexample would break down. But φ_{s_n} is just a time changed version of φ_s, so the intuition we get from this example is to allow for time changes in an appropriate way.

Definition 5.8. *Let Λ be the set of all $\lambda\colon[0,T]\to[0,T]$ that are bijective and strictly increasing (we interpret them as time changes). For $\lambda\in\Lambda$ we define*

$$\|\lambda\|:=\sup_{s\neq t}\left|\log\frac{\lambda(t)-\lambda(s)}{|t-s|}\right|,$$

which may take the value ∞ (this is just the absolute value of the logarithm of the Lipschitz constant of λ). The Skorohod metric on $D([0,T],E)$ is defined as

$$d_S(\varphi,\psi):=\inf_{\lambda\in\Lambda}\max\{\|\lambda\|,d_\infty(\varphi,\psi\circ\lambda)\}.$$

See [Ethier-Kurtz, p.117] for a proof that d_S is indeed a metric.

Obviously we have $d_S\leqslant d_\infty$, because we can always take $\lambda(t)=t$. Note that (φ_n) converges to φ with respect to d_S if and only if there exists a sequence $\lambda_n\in\Lambda$ with $\|\lambda_n\|\to0$ and such that $d_\infty(\varphi_n,\varphi\circ\lambda_n)\to0$. Note that if $\|\lambda_n\|\to0$, then for all $0<t\leqslant T$

$$0=\lim_{n\to\infty}\left|\log\frac{\lambda_n(t)}{t}\right|,$$

and therefore $\lambda_n(t)\to t$. One can strengthen this to get uniform convergence, see again [Ethier-Kurtz, p.117].

Lemma 5.9. *If E is a Polish space, then $(D([0,T],E),d_S)$ is Polish (even if E is no vector space).*

Proof. See [Ethier-Kurtz, Theorem 3.5.6]. □

Now we can apply Prohorov's criterion to characterize relatively compact families of measures on (the Borel σ-algebra of) $D([0,T],E)$. However, to show tightness we first need to understand how the compact sets in $D([0,T],E)$ look like. For that purpose we first introduce:

Definition 5.10. *For $\varphi\colon[0,T]\to E$ we define the modulus of continuity as*

$$w(\varphi,r):=\sup_{\substack{s,t\in[0,T],\\|t-s|\leqslant r}}\delta(\varphi(t),\varphi(s)),\qquad r>0.$$

Note that (since $[0,T]$ is compact) $\varphi\in C([0,T],E)$ if and only $\lim_{r\to0}w(\varphi,r)=0$. We also set

$$w'(\varphi,r):=\inf_{\{t_i\}_{i=0,\ldots,k}}\max_{0\leqslant i\leqslant k-1}\sup_{s,t\in[t_i,t_{i+1})}\delta(\varphi(t),\varphi(s)),\qquad r>0,$$

where $\{t_i\}_{i=0,\ldots,k}$ denotes a partition $0=t_0<\ldots<t_k=T$ for $k\in\mathbb{N}_0$, such that $\min_i|t_{i+1}-t_i|>r$. One can show that $\varphi\in D([0,T],E)$ if and only if $\lim_{r\to0}w'(\varphi,r)=0$.

Exercise 5.3. Show that for all $\varphi\colon[0,T]\to E$ and $r>0$ we have $w'(\varphi,r)\leqslant w(\varphi,2r)$.

Solution. *There exists a partition of* $[0, T]$ *such that each interval has length* $\in (r, 2r)$*. This partition is admissible in the definition of* $w'(\varphi, r)$*, and*

$$w'(\varphi, r) \leqslant \max_{0 \leqslant i \leqslant k} \sup_{s,t \in [t_i, t_{i+1})} \delta(\varphi(t), \varphi(s)) \leqslant \sup_{|s,t| \leqslant 2r} \delta(\varphi(t), \varphi(s)) = w(\varphi, r).$$

Exercise 5.4. Show that $\lim_{r \to 0} w'(\mathbb{I}_{[T/2,T]}, r) = 0$ but $\lim_{r \to 0} w(\mathbb{I}_{[T/2,T]}, r) = 1$.

Solution. *Considering* $s = T/2 - r$ *and* $t = T/2$ *we easily see that* $\lim_{r \to 0} w(\mathbb{I}_{[T/2,T]}, r) = 1$*. However, taking the partition* $0 < T/2 < T$ *(which is admissible in the definition of* $w'(\mathbb{I}_{[T/2,T]}, r)$ *as soon as* $r < T/2$*, we get for all* $r < T/2$

$$w'(\mathbb{I}_{[T/2,T]}, r) \leqslant \max_{0 \leqslant i \leqslant 1} \sup_{s,t \in [iT/2, (i+1)T/2)} |\varphi(t) - \varphi(s)| = 0.$$

Now recall the Arzela-Ascoli theorem from Functional Analysis or Stochastics II

Proposition 5.11. *(Arzelà-Ascoli)*
A family of functions $(\varphi_i)_{i \in I} \subset C([0, T], E)$ *is relatively compact if and only if the following two conditions hold:*

i. For all $t \in [0, T]$ *the family* $(\varphi_i(t))_{i \in I}$ *is relatively compact;*

ii. The family is uniformly equicontinuous:

$$\lim_{r \to 0} \sup_{i \in I} w(\varphi_i, r) = 0.$$

The nice thing about the "modulus of continuity" is that it equips us with a criterion that is completely analogous to the Arzelà-Ascoli criterion:

Proposition 5.12. *("Arzelà-Ascoli criterion for càdlàg functions")*
A family of functions $(\varphi_i)_{i \in I} \subset D([0, T], E)$ *is relatively compact if and only if the following two conditions hold:*

i. For all $t \in [0, T]$ *the family* $(\varphi_i(t))_{i \in I}$ *is relatively compact;*

ii. The family is uniformly equicontinuous with respect to w'*:*

$$\lim_{r \to 0} \sup_{i \in I} w'(\varphi_i, r) = 0.$$

Proof. See [Ethier-Kurtz, Theorem 3.6.3]. ☐

Write now $(X_t)_{t \in [0,T]}$ for the coordinate projection on $D([0, T], E)$, i.e. $X_t(\omega) = \omega(t)$. The combination of the results above gives us the following criterion for relative compactness:

Corollary 5.13. *Let* $(\mathbb{P}_n)_{n \in \mathbb{N}}$ *be a family of probability measures on (the Borel sets of)* $D([0, T], E)$*. Then* $(\mathbb{P}_n)$ *is relatively compact with respect to the weak topology if and only if the following two conditions are satisfied:*

i. For all $t \in [0, T]$ *the family* $(\mathbb{P}_n \circ X_t^{-1})_{n \in \mathbb{N}}$ *is tight;*

ii. For all $\varepsilon > 0$ *we have*

$$\inf_{r > 0} \limsup_{n \to \infty} \mathbb{P}_n(w'(X, r) \geqslant \varepsilon) = 0.$$

Proof. Combining Proposition 5.12 with the Prohorov criterion 5.7 it is easy to see that the conditions are necessary. Sufficiency is shown in [Ethier-Kurtz, Corollary 3.7.4]. ☐

Remark 5.14. Since $w'(\varphi, r) \leqslant w(\varphi, 2r)$ by Exercise 5.3 we can replace condition ii. in Corollary 5.13 by the stronger condition
ii.' For all $\varepsilon > 0$ we have

$$\inf_{r > 0} \limsup_{n \to \infty} \mathbb{P}_n(w(X, r) \geqslant \varepsilon) = 0.$$

In that case any limiting point of $(\mathbb{P}_n)$ is supported on $C([0,T],E)$: Indeed, if $\mathbb{P}$ is a limit point,

$$\mathbb{P}\Big(\lim_{r\to 0} w(X,r)>0\Big)=\mathbb{P}\Bigg(\bigcup_{\varepsilon>0}\bigcap_{r>0}\{w(X,r)\geqslant\varepsilon\}\Bigg)=\sup_{\varepsilon>0}\mathbb{P}\Bigg(\bigcap_{r>0}\{w(X,r)\geqslant\varepsilon\}\Bigg)$$
$$\leqslant\sup_{\varepsilon>0}\limsup_{n\to\infty}\mathbb{P}_n\Bigg(\bigcap_{r>0}\{w(X,r)\geqslant\varepsilon\}\Bigg)=\sup_{\varepsilon>0}\limsup_{n\to\infty}\inf_{r>0}\mathbb{P}_n(w(X,r)\geqslant\varepsilon)$$
$$\leqslant\sup_{\varepsilon>0}\inf_{r>0}\limsup_{n\to\infty}\mathbb{P}_n(w(X,r)\geqslant\varepsilon)=0,$$

where we used that the set $\bigcap_{r>0}\{\varphi: w(\varphi,r)\geqslant\varepsilon\}$ of functions that have a jump of size at least ε is closed in $D([0,T],E)$ (convince yourself of this!), and we applied the Portmanteau theorem in the third step.

In our case we will have $E=\mathcal{M}_+$, the space of finite measures on $\mathbb{T}^d$, which is an infinite dimensional space and not very easy to work with. To simplify the proof of tightness, we use the following auxiliary result:

Proposition 5.15. *Let $(\mathbb{P}_n)_{n\in\mathbb{N}}$ be a family of probability measures on (the Borel sets of) $D([0,T],\mathcal{M}_+)$ and write $\pi_t(\omega):=\omega(t)$ for the coordinate projection. Let Φ be a dense family of functions in $(C(\mathbb{T}^d),\|\cdot\|_\infty)$. Then $(\mathbb{P}_n)_{n\in\mathbb{N}}$ satisfies the conditions of Remark 5.14 if and only if for all $f\in\Phi$ the family of measures $(\mathbb{P}_n\circ(\pi_t(f))_{t\geqslant 0})_{n\in\mathbb{N}}$ on (the Borel sets of) $D([0,T],\mathbb{R})$ satisfies the conditions of Remark 5.14.*

Remark 5.16. We will apply this for example with $\Phi=C^\infty(\mathbb{T}^d)$ (which can be shown to be dense by using mollifiers). The analogous result with the conditions of Corollary 5.13 is also true, but its proof is more involved because $w'(\varphi+\psi,r)\nleqslant w'(\varphi,r)+w'(\psi,r)$ in general (and we will use the inequality $w(\varphi+\psi,r)\leqslant w(\varphi,r)+w(\psi,r)$ in the following proof).

Proof. Necessity is easy because $\pi\mapsto\pi(f)$ is continuous, so let us focus on sufficiency.

Note that tightness is defined in terms of compact sets, which do not change if we take another metric that induces the topology of weak convergence. So we can take the sequence $(f_k)_{k\in\mathbb{N}}$ in the definition of the metric in Exercise 5.1 as a subset of Φ.

Since $\mathbb{T}^d$ is compact, a family of measures $(\mu_i)_{i\in I}$ is tight if and only if $\sup_{i\in I}\mu_i(1)<\infty$. Let $f\in\Phi$ with $\gamma=\inf_{x\in\mathbb{T}^d}f(x)>0$. Since $(\mathbb{P}_n\circ(\pi_t(f))_{t\geqslant 0})_{n\in\mathbb{N}}$ is tight, we know by Corollary 5.13 that for all $t\in[0,T]$ and all $\varepsilon>0$ there exists $C=C(t,\varepsilon)>0$ with

$$\sup_{n\in\mathbb{N}}\mathbb{P}_n(\pi_t(f)>C)\leqslant\varepsilon.$$

But then

$$\sup_{n\in\mathbb{N}}\mathbb{P}_n\Big(\pi_t(1)>\frac{C}{\gamma}\Big)=\sup_{n\in\mathbb{N}}\mathbb{P}_n\Big(\pi_t(\gamma)>\gamma\frac{C}{\gamma}\Big)\leqslant\sup_{n\in\mathbb{N}}\mathbb{P}_n(\pi_t(f)>C)\leqslant\varepsilon,$$

so that $\mathbb{P}_n\circ\pi_t$ is tight for all $t\in[0,T]$ and condition i. in Corollary 5.13 is satisfied for $\mathbb{P}_n$ (and not just for $(\mathbb{P}_n\circ(\pi_t(f))_{t\geqslant 0})_{n\in\mathbb{N}}$).

It remains to prove condition ii.' in Remark 5.14 for $(\mathbb{P}_n)_n$, where we recall that w is defined in terms of the metric δ that uses functions $(f_k)\subset\Phi$. Let $\varepsilon>0$ and let $k(\varepsilon)\in\mathbb{N}$ be such that

$$\sum_{k=k(\varepsilon)+1}^{\infty}2^{-k}\leqslant\varepsilon.$$

Then

$$\delta(\pi,\nu)\leqslant\sum_{k=1}^{k(\varepsilon)}2^{-k}(|\pi(f_k)-\nu(f_k)|\wedge 1)+\varepsilon,$$

and therefore

$$w(\pi,r)=\sup_{|s-t|\leqslant r}\delta(\pi_t,\pi_s)\leqslant\sup_{|s-t|\leqslant r}\Bigg(\sum_{k=1}^{k(\varepsilon)}2^{-k}(|\pi_t(f_k)-\pi_s(f_k)|\wedge 1)+\varepsilon\Bigg)$$
$$\leqslant\sum_{k=1}^{k(\varepsilon)}2^{-k}\sup_{|s-t|\leqslant r}|\pi_t(f_k)-\pi_s(f_k)|+\varepsilon=\sum_{k=1}^{k(\varepsilon)}2^{-k}w(\pi(f_k),r)+\varepsilon$$

Let now $\beta > 0$. By assumption there exists r with

$$\limsup_n \mathbb{P}_n(w(\pi(f_k), r) \geqslant \varepsilon) \leqslant 2^{-k}\beta, \qquad k = 1, \ldots, k(\varepsilon),$$

which yields

$$\begin{aligned}\limsup_n \mathbb{P}_n(w(\pi, r) \geqslant 2\varepsilon) &\leqslant \limsup_n \mathbb{P}_n\left(\sum_{k=1}^{k(\varepsilon)} 2^{-k} w(\pi(f_k), r) + \varepsilon \geqslant 2\varepsilon\right) \\ &\leqslant \limsup_n \left(\sum_{k=1}^{k(\varepsilon)} \mathbb{P}_n(w(\pi(f_k), r) \geqslant \varepsilon)\right) \leqslant \beta,\end{aligned}$$

and since $\varepsilon > 0$ and $\beta > 0$ were arbitrary, the claim follows. □

5.3 Step 1: Tightness of rescaled symmetric exclusion

We are now going to apply the tools developed in the previous subsection to complete Step 1 in the convergence proof for the symmetric simple exclusion process. To simplify notation let us write $\bar{\pi}_t^N := \pi_{N^2 t}^N$ and $\bar{M}_t^{N,\varphi} := M_{N^2 t}^{N,\varphi}$, so that $\bar{M}^{N,\varphi}$ is a martingale in the filtration $(\bar{\mathcal{F}}_t^N)_{t \geqslant 0} = \sigma(\eta_s^N : s \leqslant N^2 t)$. More precisely,

$$\bar{\pi}_t^N(\varphi) := \frac{1}{N^d} \sum_{x \in \mathbb{Z}_N^d} \varphi\left(\frac{x}{N}\right) \eta_{N^2 t}^N(x),$$

and then we have by the computations in Subsection 5.1

$$\bar{\pi}_t^N(\varphi) = \bar{\pi}_0^N(\varphi) + \int_0^t \frac{1}{2} \sum_{y \in \mathbb{Z}_N^d} p^N(y) \bar{\pi}_r^N(\Delta_{N\cdot, y}^N \varphi) \mathrm{d}r + \bar{M}_t^{N,\varphi}. \tag{5.8}$$

To prove tightness of $\bar{\pi}^N$ we apply Proposition 5.15 together with Remark 5.16, so that it suffices to show for all $\varphi \in C^\infty(\mathbb{T}^d)$ the following two steps:

i. For all $t \in [0, T]$ the family $(\bar{\pi}_t^N(\varphi))_{N \in \mathbb{N}}$ is tight in $\mathbb{R}$;

ii. For all $\varepsilon > 0$ we have

$$\inf_{r > 0} \limsup_{N \to \infty} \mathbb{P}_{\mu_\rho^N}(w(\bar{\pi}_t^N(\varphi), r) \geqslant \varepsilon) = 0.$$

The proof of i. is fairly easy: Since $\eta_{N^2 t}^N(x) \in \{0, 1\}$ for all x and t we have for all N and t

$$|\bar{\pi}_t^N(\varphi)| = \left| \frac{1}{N^d} \sum_{x \in \mathbb{Z}_N^d} \varphi\left(\frac{x}{N}\right) \eta_{N^2 t}^N(x) \right| \leqslant \frac{1}{N^d} \sum_{x \in \mathbb{Z}_N^d} \left| \varphi\left(\frac{x}{N}\right) \right| \leqslant \|\varphi\|_\infty,$$

so with probability 1 the random variable $\bar{\pi}_t^N(\varphi)$ is contained in the compact set $[-\|\varphi\|_\infty, \|\varphi\|_\infty]$.

To prove step ii. note that for two processes X and Y we have

$$\mathbb{P}(w(X + Y, r) \geqslant \varepsilon) \leqslant \mathbb{P}\left(w(X, r) \geqslant \frac{\varepsilon}{2}\right) + \mathbb{P}\left(w(Y, r) \geqslant \frac{\varepsilon}{2}\right),$$

so in our case it suffices to control the modulus of continuity of the finite variation part and the martingale part in (5.8) separately. For the finite variation part, we have

$$\begin{aligned}\left| \int_s^t \frac{1}{2} \sum_{y \in \mathbb{Z}_N^d} p^N(y) \bar{\pi}_r^N(\Delta_{N\cdot, y}^N \varphi) \mathrm{d}r \right| &\leqslant \int_s^t \frac{1}{2} \sum_{y \in \mathbb{Z}_N^d} p^N(y) |\bar{\pi}_r^N(\Delta_{N\cdot, y}^N \varphi)| \mathrm{d}r \\ &\leqslant \int_s^t \frac{1}{2} \sum_{y \in \mathbb{Z}_N^d} p^N(y) \|\Delta_{N\cdot, y}^N \varphi\|_\infty \mathrm{d}r,\end{aligned}$$

and now recall that $\Delta_{x,y}^N \varphi = N^2 \left[\varphi\left(\frac{x+y}{N}\right) + \varphi\left(\frac{x-y}{N}\right) - 2\varphi\left(\frac{x}{N}\right) \right] \leqslant \|\mathrm{D}^2 \varphi\|_\infty |y|^2$, which leads to

$$\left| \int_s^t \frac{1}{2} \sum_{y \in \mathbb{Z}_N^d} p^N(y) \bar{\pi}_r^N(\Delta_{N\cdot, y}^N \varphi) \mathrm{d}r \right| \leqslant \frac{1}{2} \sum_y p^N(y) |y|^2 |t - s| \lesssim |t - s|, \tag{5.9}$$

where in the last step we used that p has only a finite range and therefore of course a finite variance. For the martingale part we use the following general result:

Lemma 5.17. *Let $r,p>0$ and let $(X_t)_{t\in[0,T]}$ be a stochastic process with $\mathbb{E}[\sup_{r\in[s,t]}|X_r-X_s|^p]\leqslant C|t-s|$ for all $|t-s|\leqslant r$. Then*

$$\mathbb{E}[w(X,r)^p]\lesssim C.$$

Proof. We decompose $[0,T]$ into equal-sized intervals of length r, writing $t_k^r=(kr)\wedge T$ and get

$$\mathbb{E}[w(X,r)^p]=\mathbb{E}\Bigg[\sup_{s\leqslant t,|t-s|\leqslant r}|X_t-X_s|^p\Bigg]=\sum_{k=0}^{\lfloor T/r\rfloor}\mathbb{E}\Bigg[\sup_{s\leqslant t,|t-s|\leqslant r}\mathbb{I}_{s\in[t_k^r,t_{k+1}^r)}|X_t-X_s|^p\Bigg].$$

For fixed k we distinguish two cases: if $t\in[t_k^r,t_{k+1}^r]$, then

$$|X_t-X_s|^p\lesssim|X_s-X_{t_k^r}|^p+|X_t-X_{t_k^r}|^p\lesssim\sup_{s'\in[t_k^r,t_{k+1}^r]}|X_{s'}-X_{t_k^r}|^p,$$

while for $t\in(t_{k+1}^r,t_{k+2}^r]$

$$\begin{aligned}|X_t-X_s|^p&\lesssim|X_s-X_{t_k^r}|^p+|X_{t_{k+1}^r}-X_{t_k^r}|^p+|X_t-X_{t_k^r}|^p\\&\lesssim\sup_{s'\in[t_k^r,t_{k+1}^r]}|X_{s'}-X_{t_k^r}|^p+\sup_{s'\in[t_{k+1}^r,t_{k+2}^r]}|X_{s'}-X_{t_{k+1}^r}|^p.\end{aligned}$$

Therefore, we obtain overall

$$\begin{aligned}\sum_{k=0}^{\lfloor T/r\rfloor}\mathbb{E}\Bigg[\sup_{s\leqslant t,|t-s|\leqslant r}\mathbb{I}_{s\in[t_k^r,t_{k+1}^r)}|X_t-X_s|^p\Bigg]&\lesssim\sum_{k=0}^{\lfloor T/r\rfloor}\mathbb{E}\Bigg[\sup_{s'\in[t_k^r,t_{k+1}^r]}|X_{s'}-X_{t_k^r}|^p\Bigg]\\&\leqslant(\lfloor T/r\rfloor+1)Cr\lesssim C,\end{aligned}$$

which concludes the proof. □

We want to apply this result to our martingale $\bar{M}_t^{N,\varphi}$ and for $p=2$, for which we have by Doob's inequality

$$\begin{aligned}\mathbb{E}_{\mu_\rho^N}\Bigg[\sup_{r\in[s,t]}|\bar{M}_r^{N,\varphi}-\bar{M}_s^{N,\varphi}|^2\Bigg]&\lesssim\mathbb{E}_{\mu_\rho^N}[|\bar{M}_t^{N,\varphi}-\bar{M}_s^{N,\varphi}|^2]\\&\overset{\text{mart. property}}{=\!=}\mathbb{E}_{\mu_\rho^N}[|\bar{M}_t^{N,\varphi}|^2-|\bar{M}_s^{N,\varphi}|^2]=\mathbb{E}_{\mu_\rho^N}[\langle\bar{M}^{N,\varphi}\rangle_t-\langle\bar{M}^{N,\varphi}\rangle_s],\end{aligned}$$

where the last step follows from Lemma 5.2. By Lemma 5.4 we further have

$$\begin{aligned}\langle\bar{M}^{N,\varphi}\rangle_t-\langle\bar{M}^{N,\varphi}\rangle_s&=\int_s^t\frac{1}{N^{2d}}\sum_{x,y}\eta_{N^2r}^N(x)(1-\eta_{N^2r}^N(x+y))p^N(y)[\nabla_{x,y}^N\varphi]^2\mathrm{d}r\\&\leqslant\int_s^t\frac{1}{N^{2d}}\sum_{x,y}p^N(y)[\nabla_{x,y}^N\varphi]^2\mathrm{d}r\leqslant|t-s|\frac{1}{N^{2d}}\sum_{x,y}p^N(y)\|\nabla\varphi\|_\infty^2|y|^2\lesssim N^{-d}|t-s|,\end{aligned}$$

which together with Lemma 5.17 gives for all $\varepsilon>0$

$$\inf_{r>0}\limsup_{N\to\infty}\mathbb{P}_{\mu_\rho^N}[w(\bar{M}^{N,\varphi},r)\geqslant\varepsilon]\lesssim\inf_{r>0}\limsup_{N\to\infty}\frac{\mathbb{E}_{\mu_\rho^N}[w(\bar{M}^{N,\varphi},r)^2]}{\varepsilon^2}\lesssim\inf_{r>0}\limsup_{N\to\infty}\frac{N^{-d}}{\varepsilon^2}=0.$$

To summarize, we just showed:

Lemma 5.18. *The family $(\bar{\pi}^N)_{N\in\mathbb{N}}$ is tight in $D([0,T],\mathcal{M}_+)$.*

5.4 Step 2: Description of the possible limit points

By the results of the previous step we know that there exists a subsequence $(N(k))_{k\in\mathbb{N}}$ along which $(\bar{\pi}^{N(k)})_{k\in\mathbb{N}}$ converges to some limiting $\bar{\pi}$. Our aim is to identify $\bar{\pi}$ as a solution to the heat equation. For that purpose we first need an auxiliary lemma.

Lemma 5.19. *Let $N\in\mathbb{N}$ and $\varphi\in C^\infty(\mathbb{T}^d)$. Then*

$$\mathbb{E}_{\mu_\rho^N}\left[\sup_{t\in[0,T]}\left|\bar{\pi}_t^N(\varphi)-\bar{\pi}_0^N(\varphi)-\int_0^t\frac{1}{2}\bar{\pi}_r^N(\Delta_C\varphi)\mathrm{d}r\right|^2\right]\lesssim N^{-2}+N^{-d}.$$

Proof. We have

$$\begin{aligned}&\mathbb{E}_{\mu_\rho^N}\left[\sup_{t\in[0,T]}\left|\bar{\pi}_t^N(\varphi)-\bar{\pi}_0^N(\varphi)-\int_0^t\frac{1}{2}\bar{\pi}_r^N(\Delta_C\varphi)\mathrm{d}r\right|^2\right]\\ \lesssim&\mathbb{E}_{\mu_\rho^N}\left[\sup_{t\in[0,T]}\left|\int_0^t\frac{1}{2}\left(\sum_{y\in\mathbb{Z}_N^d}p^N(y)\bar{\pi}_r^N(\Delta_{N\cdot,y}^N\varphi)-\bar{\pi}_r^N(\Delta_C\varphi)\right)\mathrm{d}r\right|^2\right]+\mathbb{E}_{\mu_\rho^N}\left[\sup_{t\in[0,T]}|\bar{M}_t^{N,\varphi}|^2\right]\\ \lesssim&\mathbb{E}_{\mu_\rho^N}\left[\sup_{t\in[0,T]}\left|\int_0^t\frac{1}{2}\bar{\pi}_r^N\left(\sum_{y\in\mathbb{Z}_N^d}p^N(y)\Delta_{N\cdot,y}^N\varphi-\Delta_C\varphi\right)\mathrm{d}r\right|^2\right]+N^{-d},\end{aligned}$$

where the last step follows from Lemma 5.17 with $r=T$ (or more precisely from the computations based on the martingale quadratic variation in the previous subsection). Now for all $x\in\mathbb{T}^d$ and all N that are larger than the range of our random walk

$$\begin{aligned}&\left|\sum_{y\in\mathbb{Z}_N^d}p^N(y)\Delta_{x,y}^N\varphi-\Delta_C\varphi\left(\frac{x}{N}\right)\right|\\ =&\left|\sum_{y\in\mathbb{Z}^d}p(y)\left(N^2\left[\varphi\left(\frac{x+y}{N}\right)+\varphi\left(\frac{x-y}{N}\right)-2\varphi\left(\frac{x}{N}\right)\right]-y^T\mathrm{D}^2\varphi\left(\frac{x}{N}\right)y\right)\right|\\ \leqslant&\sum_{y\in\mathbb{Z}^d}p(y)N^{-1}\|\mathrm{D}^3\varphi\|_{L^\infty}|y|^3\lesssim N^{-1},\end{aligned}$$

and therefore, since $\eta_{N^2r}^N(x)\in\{0,1\}$ and thus $|\bar{\pi}_r^N(\psi)|\leqslant\|\psi\|_\infty$,

$$\mathbb{E}_{\mu_\rho^N}\left[\sup_{t\in[0,T]}\left|\int_0^t\frac{1}{2}\bar{\pi}_r^N\left(\sum_{y\in\mathbb{Z}_N^d}p^N(y)\Delta_{N\cdot,y}^N\varphi-\Delta_C\varphi\right)\mathrm{d}r\right|^2\right]\lesssim\frac{T^2}{N^2}.$$

□

Corollary 5.20. *Let $(N(k))_{k\in\mathbb{N}}$ be such that $(\bar{\pi}^{N(k)})_{k\in\mathbb{N}}$ converges in distribution in $D([0,T],\mathcal{M}_+)$ to a limit $\bar{\pi}\in C([0,T],\mathcal{M}_+)$ and let $\varphi\in C^\infty(\mathbb{T}^d)$. Then almost surely for all $t\in[0,T]$*

$$\bar{\pi}_t(\varphi)=\bar{\pi}_0(\varphi)+\int_0^t\frac{1}{2}\bar{\pi}_r(\Delta_C\varphi)\mathrm{d}r. \tag{5.10}$$

Remark 5.21. Note the order of the quantifiers here: The null set where (5.10) fails may depend on φ. With some more work we could show that the null set is independent of φ (for example by considering a countable dense sequence of functions such that $(\Delta_C f_k)_k$ is dense in $(\Delta_C\varphi)_{\varphi\in C^\infty}$), but we will not need this.

Proof. Since the function

$$\nu\mapsto\sup_{t\in[0,T]}\left|\nu_t(\varphi)-\nu_0(\varphi)-\int_0^t\frac{1}{2}\nu_r(\Delta_C\varphi)\mathrm{d}r\right|\wedge1$$

is continuous on $D([0,T],\mathcal{M}_+)$ and bounded, we have

$$\begin{aligned}&\mathbb{E}_{\mu_\rho^N}\left[\sup_{t\in[0,T]}\left|\bar{\pi}_t(\varphi)-\bar{\pi}_0(\varphi)-\int_0^t\frac{1}{2}\bar{\pi}_r(\Delta_C\varphi)\mathrm{d}r\right|\wedge1\right]\\ =&\lim_{k\to\infty}\mathbb{E}_{\mu_\rho^N}\left[\sup_{t\in[0,T]}\left|\bar{\pi}_t^{N(k)}(\varphi)-\bar{\pi}_0^{N(k)}(\varphi)-\int_0^t\frac{1}{2}\bar{\pi}_r^{N(k)}(\Delta_C\varphi)\mathrm{d}r\right|\wedge1\right]=0,\end{aligned}$$

where the last step follows from Lemma 5.19. □

Corollary 5.20 identifies the dynamics of any limiting point $\bar{\pi}$ with a (very) weak solution to the heat equation. To derive a uniqueness result for the limit points from this, we still need to show that all limit points have the same initial condition $\bar{\pi}_0$.

Lemma 5.22. *Let $\rho \in C(\mathbb{T}^d, [0,1])$ and let $\varphi \in C(\mathbb{T}^d)$. Then*

$$\lim_{N\to\infty} \mathbb{E}_{\mu_\rho^N}\left[\left|\bar{\pi}_0^N(\varphi) - \int_{\mathbb{T}^d}\varphi(u)\rho(u)\mathrm{d}u\right|^2\right] = 0.$$

Proof. Under μ_ρ^N we have $\mathbb{E}_{\mu_\rho^N}[\eta_0^N(x)] = \rho(x/N)$, and therefore

$$\mathbb{E}_{\mu_\rho^N}\left[\left|\bar{\pi}_0^N(\varphi) - \int_{\mathbb{T}^d}\varphi(x)\rho(x)\mathrm{d}x\right|^2\right] \lesssim \mathbb{E}_{\mu_\rho^N}\left[\left|\frac{1}{N^d}\sum_{x\in\mathbb{Z}_N^d}\varphi\Big(\frac{x}{N}\Big)\Big(\eta_0^N(x) - \rho\Big(\frac{x}{N}\Big)\Big)\right|^2\right] + \left|\frac{1}{N^d}\sum_{x\in\mathbb{Z}_N^d}\varphi\Big(\frac{x}{N}\Big)\rho\Big(\frac{x}{N}\Big) - \int_{\mathbb{T}^d}\varphi(u)\rho(u)\mathrm{d}u\right|^2.$$

Since under μ_ρ^N the coordinates $\eta_0^N(x)$ and $\eta_0^N(y)$ are independent for $x \neq y$ and since $\mathbb{E}_{\mu_\rho^N}[\eta_0^N(x)] = \rho(x/N)$, the expectation on the right hand side equals

$$\mathbb{E}_{\mu_\rho^N}\left[\left|\frac{1}{N^d}\sum_{x\in\mathbb{Z}_N^d}\varphi\Big(\frac{x}{N}\Big)\Big(\eta_0^N(x) - \rho\Big(\frac{x}{N}\Big)\Big)\right|^2\right] = \frac{1}{N^{2d}}\sum_{x\in\mathbb{Z}_N^d}\varphi^2\Big(\frac{x}{N}\Big)\mathrm{var}(\eta_0^N(x)) \lesssim N^{-d}.$$

On the other hand, $\frac{1}{N^d}\sum_{x\in\mathbb{Z}_N^d}\varphi\big(\frac{x}{N}\big)\rho\big(\frac{x}{N}\big)$ is a Riemann sum approximation of $\int_{\mathbb{T}^d}\varphi(u)\rho(u)\mathrm{d}u$, and therefore the claim follows. □

So to summarize:

Lemma 5.23. *Let $\bar{\pi} \in C([0,T], \mathcal{M}_+)$ be a limit point of $(\bar{\pi}^N)_N$. Then for every $\varphi \in C^\infty(\mathbb{T}^d)$ almost surely*

$$\bar{\pi}_t(\varphi) = \rho(\varphi) + \int_0^t \frac{1}{2}\bar{\pi}_r(\Delta_C\varphi)\mathrm{d}r,$$

where we wrote $\rho(\varphi) := \int_{\mathbb{T}^d}\varphi(x)\rho(x)\mathrm{d}x$.

5.5 Step 3: Uniqueness of the limit

Here we need to show that there exists at most one process in $C([0,T], \mathcal{M}_+)$ which satisfies the characterization in Lemma 5.23. For that purpose let

$$e_m(u) := e^{-2\pi i m\cdot u}, \qquad u \in \mathbb{T}^d,$$

and note that, since $(e_m)_{m\in\mathbb{Z}^d}$ is a countable family of functions, and since

$$(\Delta_C e_m)(u) = \sum_{j,k=1}^d C_{jk}(2\pi i m_j)(2\pi i m_k)e_m(u) = (-4\pi^2 m^T C m)\, e_m(u),$$

we get almost surely for all $m \in \mathbb{Z}^d$ and $t \in [0,T]$

$$\bar{\pi}_t(e_m) = \rho(e_m) + \int_0^t (-2\pi^2 m^T C m)\bar{\pi}_r(e_m)\mathrm{d}r,$$

and therefore

$$\bar{\pi}_t(e_m) = \rho(e_m) \times e^{-t(-2\pi^2 m^T C m)}.$$

In other words, any limit point is uniquely determined on the test functions $(e_m)_{m\in\mathbb{Z}^d}$. Now $\mathbb{T}^d$ is a compact space, and $\mathrm{span}\{(e_m)_{m\in\mathbb{Z}^d}\}$ is an algebra that contains the constant functions and that separates points: for $u\neq v$ we have

$$\frac{e_m(u)}{e_m(v)} = e^{2\pi i m\cdot(u-v)},$$

and if $u-v\notin\mathbb{Z}^d$, then there exists $m\in\mathbb{Z}^d$ with $m\cdot(u-v)\notin\mathbb{Z}^d$. Moreover, $\mathrm{span}\{(e_m)_{m\in\mathbb{Z}^d}\}$ is closed under convex conjugation. So by the Stone-Weierstraß theorem $\mathrm{span}\{(e_m)_{m\in\mathbb{Z}^d}\}$ is dense in $(C(\mathbb{T}^d,\mathbb{C}),\|\cdot\|_\infty)$, and therefore $(\bar\pi_t)_{t\in[0,T]}$ is uniquely determined by $(\bar\pi_t(e_m))_{t\in[0,T],m\in\mathbb{Z}^d}$, which proves the uniqueness of the limit point and identifies it with the unique weak solution to the heat equation

$$\partial_t\bar\pi = \frac{1}{2}\Delta_C\bar\pi, \qquad \bar\pi_0 = \rho.$$

Theorem 5.24. *Let $\rho_0\in C(\mathbb{T}^d,[0,1])$ and let $\eta_0^N\sim\mu_{\rho_0}^N$ for all N. Assume that p is symmetric. Then the sequence of processes $(\bar\pi_t^N)_{t\in[0,T]}\in D([0,T],\mathcal{M}_+)$ converges in probability in $D([0,T],\mathcal{M}_+)$ to the unique weak solution ρ of the heat equation*

$$\partial_t\rho = \frac{1}{2}\Delta_C\rho, \qquad \rho(0,\cdot) = \rho_0.$$

Proof. The only part which we did not prove yet is that the convergence even holds in probability and not only in distribution. For that purpose note that $\pi\mapsto d_S(\pi,\rho)\wedge 1$ is a continuous bounded function on $D([0,T],\mathcal{M}_+)$, and therefore

$$\lim_{N\to\infty}\mathbb{E}_{\mu_\rho^N}[d_S(\bar\pi^N,\rho)\wedge 1] = \mathbb{E}[d_S(\rho,\rho)\wedge 1] = 0,$$

from where an application of Markov's inequality easily gives the convergence in probability. □

6 Equilibrium fluctuations for symmetric exclusion

The results of the previous section give us a precise picture how the density profile of η^N looks on spatial scales of order N and time scales of order N^2. But of course η^N is actually random, and for fixed N the process $\bar\pi^N$ will not agree with the solution to the heat equation ρ, but will have random fluctuations. Here we want to understand these fluctuations better – we will derive a type of central limit theorem, associated to the hydrodynamic limit result Theorem 5.24 which can be understood as a law of large numbers. And while the hydrodynamic limit solves a PDE, we will see that the limit of the fluctuations solves a *stochastic* PDE (SPDE).

To simplify our life, we only consider the situation where $\rho_0\equiv\rho\in(0,1)$, so that $\rho(t,u)\equiv\rho$ for all $(t,u)\in[0,T]\times\mathbb{T}^d$ (the cases $\rho\in\{0,1\}$ are boring because then η^N is deterministic and constant in time). Since then μ_ρ^N is a reversible (and in particular invariant) measure for η^N, we are in *equilibrium*. By Theorem 5.24 we know that

$$\lim_{N\to\infty}|\bar\pi_t^N(\varphi) - \langle\rho,\varphi\rangle_{L^2}| = 0$$

for all $t\in[0,T]$. To see the fluctuations, we should therefore *blow up* the difference and set

$$\mathcal{Y}_t^N(\varphi) := N^\alpha(\bar\pi_t^N(\varphi) - \langle\rho,\varphi\rangle_{L^2}) = N^\alpha(\bar\pi_t^N(\varphi) - \mathbb{E}_{\mu_\rho^N}[\bar\pi_t^N(\varphi)])$$

for an $\alpha>0$ to be determined. To figure out which α to take let us first consider $t=0$, for which we get $\mathbb{E}[\mathcal{Y}_0^N(\varphi)]=0$ and

$$\mathrm{var}(\mathcal{Y}_0^N(\varphi)) = N^{2\alpha}\frac{1}{N^{2d}}\sum_{x\in\mathbb{Z}_N^d}\varphi\Big(\frac{x}{N}\Big)^2\mathrm{var}(\eta_0^N(x)) = N^{2\alpha-2d}\sum_{x\in\mathbb{Z}_N^d}\varphi\Big(\frac{x}{N}\Big)^2\rho(1-\rho).$$

Now the right hand side is a Riemann sum approximation of $\rho(1-\rho)\int_{\mathbb{T}^d}\varphi^2(u)\mathrm{d}u$ if and only if $2\alpha-2d=-d$, i.e. if $\alpha=d/2$. In that case at least $\mathrm{var}(\mathcal{Y}_0^N(\varphi))$ is bounded, and for $\alpha>d/2$ it would diverge to $+\infty$ while for $\alpha<d/2$ it would vanish in the limit. Therefore, we fix $\alpha=d/2$ from now on.

Unfortunately, unlike for the hydrodynamic limit the fluctuations will not be given by a measure valued process, but only by a distribution valued process: We will see that already for $t=0$ the process $\mathcal{Y}_0^N$ converges in distribution to a white noise, which in $d=1$ can be identified with the derivative of the Brownian motion, and since we know from stochastic analysis that the Brownian motion is not of finite variation, the white noise cannot be a measure. To remain in the setting of Subsection 5.2, let us introduce suitable Polish spaces of distributions:

Definition 6.1. *We write for $m\in\mathbb{Z}^d$*

$$e_m(u):=e^{-2\pi i m\cdot u}.$$

For $\alpha\in\mathbb{R}$ we define the (L^2)-Sobolev space H^α as the completion of $C^\infty(\mathbb{T}^d,\mathbb{C})$ with respect to

$$\|\varphi\|_{H^\alpha}^2:=\sum_{m\in\mathbb{Z}^d}(1+|m|^2)^\alpha|\langle\varphi,e_m\rangle_{L^2}|^2.$$

This space can be identified with sequence $(\hat{\mathcal{Z}}(m))_{m\in\mathbb{Z}^d}$ and it is a Hilbert space with inner product

$$\langle\mathcal{Z},\mathcal{Z}'\rangle_{H^\alpha}=\sum_{m\in\mathbb{Z}^d}(1+|m|^2)^\alpha\hat{\mathcal{Z}}(m)\overline{\hat{\mathcal{Z}}'(m)},$$

where $\overline{(\cdot)}$ denotes the complex conjugate, and it is separable because the sequences with finitely many non-zero entries are easily seen to be dense (and then we can consider finite sequences with coefficients of the form $a+ib$ with $a,b\in\mathbb{Q}$, which gives us a countable set).

Exercise 6.1. Show that $\mathrm{span}\{e_m\}$ is dense in H^α for any $\alpha\in\mathbb{R}$.

Solution. *Let $\mathcal{Y}\in H^\alpha$, which we can identify with a sequence $(\hat{\mathcal{Y}}(m))_{m\in\mathbb{Z}^d}$, where formally $\hat{\mathcal{Y}}(m)=\langle\mathcal{Y},e_m\rangle_{L^2}$ (but of course in general $\mathcal{Y}\notin L^2$). Set*

$$\mathcal{Y}_N:=\sum_{|m|\leqslant N}\hat{\mathcal{Y}}(m)\overline{e_m}.$$

Since $\langle e_k,\overline{e_m}\rangle_{L^2}=\delta_{k,m}$, we have $\widehat{\mathcal{Y}_N}(m)=\hat{\mathcal{Y}}(m)$ for all $|m|\leqslant N$ and $\hat{\mathcal{Y}}(m)=0$ for all $m>N$, and therefore

$$\|\mathcal{Y}-\mathcal{Y}_N\|_{H^\alpha}^2=\sum_{m\in\mathbb{Z}^d}(1+|m|^2)^\alpha\big|\hat{\mathcal{Y}}_m-\widehat{\mathcal{Y}_N}(m)\big|^2=\sum_{|m|>N}(1+|m|^2)^\alpha\big|\hat{\mathcal{Y}}_m\big|^2,$$

which, since $\mathcal{Y}\in H^\alpha$ and therefore the sum over all m is finite, converges to zero as $N\to\infty$.

Lemma 6.2. *Any $\mathcal{Y}\in H^\alpha$ defines a linear map on $C^\infty(\mathbb{T}^d)$ via*

$$\mathcal{Y}(\varphi):=\sum_{m\in\mathbb{Z}^d}\hat{\mathcal{Y}}(m)\langle\varphi,\overline{e_m}\rangle_{L^2}=\sum_{m\in\mathbb{Z}^d}\hat{\mathcal{Y}}(m)\overline{\hat{\varphi}(m)}. \tag{6.1}$$

Moreover, we have

$$|\mathcal{Y}(\varphi)|\leqslant\|\mathcal{Y}\|_{H^\alpha}\|\varphi\|_{H^{-\alpha}}$$

and $\|\varphi\|_{H^{-\alpha}}<\infty$ for all $\varphi\in C^\infty(\mathbb{T}^d)$.

Proof. The expression for $\mathcal{Y}(\varphi)$ is derived by formally writing $\mathcal{Y}=\sum_{m\in\mathbb{Z}^d}\hat{\mathcal{Y}}(m)\overline{e_m}$. Since integration by parts gives for any multi-index $\gamma\in\mathbb{N}_0^d$ (with $m^\gamma=m_1^{\gamma_1}\times\ldots\times m_d^{\gamma_d}$ and $\partial_u^\gamma=\partial_{u_1}^{\gamma_1}\ldots\partial_{u_d}^{\gamma_d}$ and $|\gamma|=\gamma_1+\cdots+\gamma_d$)

$$\begin{aligned}(2\pi i m)^\gamma\hat{\varphi}(m)&=\int_{\mathbb{T}^d}\varphi(u)(2\pi i m)^\gamma e^{-2\pi i m\cdot u}\mathrm{d}u=\int_{\mathbb{T}^d}\varphi(u)(-1)^{|\gamma|}\partial_u^\gamma e^{-2\pi i m\cdot u}\mathrm{d}u\\&=\int_{\mathbb{T}^d}\partial_u^\gamma\varphi(u)e^{-2\pi i m\cdot u}\mathrm{d}u=\widehat{\partial_u^\gamma\varphi}(m),\end{aligned}$$

we get $|(2\pi i m)^\gamma \hat{\varphi}(m)| \leqslant \|\partial_u^\gamma \varphi\|_{L^\infty}$, uniformly in $m \in \mathbb{Z}^d$. In particular $\hat{\varphi}(m)$ decays faster than any rational function in m, and the expression in (6.1) is well defined and $\|\varphi\|_{H^{-\alpha}} < \infty$ for all $\varphi \in C^\infty(\mathbb{T}^d)$. Finally, the Cauchy-Schwarz inequality gives

$$\left|\sum_{m\in\mathbb{Z}^d} \hat{\mathcal{Y}}(m)\overline{\hat{\varphi}(m)}\right| = \left|\sum_{m\in\mathbb{Z}^d} (1+|m|^2)^{\alpha/2}\hat{\mathcal{Y}}(m)(1+|m|^2)^{-\alpha/2}\overline{\hat{\varphi}(m)}\right|$$
$$\leqslant \left(\sum_{m\in\mathbb{Z}^d} (1+|m|^2)^{\alpha}|\hat{\mathcal{Y}}(m)|^2\right)^{1/2}\left(\sum_{m\in\mathbb{Z}^d} (1+|m|^2)^{-\alpha}|\overline{\hat{\varphi}(m)}|^2\right)^{1/2} = \|\mathcal{Y}\|_{H^\alpha}\|\varphi\|_{H^{-\alpha}}.$$

□

Lemma 6.3. *Let $\alpha < -d/2$ and set*

$$\mathcal{Y}_t^N(\varphi) := N^{d/2}(\bar{\pi}_t^N(\varphi) - \langle \rho, \varphi\rangle_{L^2}).$$

Then $\mathcal{Y}^N \in D([0,T], H^\alpha)$ for all $t \in [0,T]$ and $N \in \mathbb{N}$.

Proof. Since $\mathcal{Y}^N$ only makes finitely many jumps, it suffices to bound $\|\mathcal{Y}_t^N\|_{H^\alpha}$ for all $t \in [0,T]$. For that purpose fix $m \in \mathbb{Z}^d$ and note that

$$\left|\widehat{\mathcal{Y}_t^N}(m)\right| = \left|N^{-d/2}\sum_{x\in\mathbb{Z}_N^d} e_m\left(\frac{x}{N}\right)(\eta_{N^2t}^N(x) - \rho)\right| \leqslant N^{d/2},$$

and therefore

$$\|\mathcal{Y}_t^N\|_{H^\alpha}^2 \leqslant \sum_{m\in\mathbb{Z}^d} (1+|m|^2)^\alpha N^d < \infty$$

for $\alpha < -d/2$ (where we used that $\sum_{m\in\mathbb{Z}^d\setminus\{0\}} |m|^\gamma < \infty$ if and only if $\gamma < -d$, which can be shown by relating the sum to the corresponding integral and using polar coordinates). □

To prove the convergence of $(\mathcal{Y}^N)_{N\in\mathbb{N}}$ in $D([0,T], H^\alpha)$ for a suitable α (to be determined) we follow the same three steps as for the hydrodynamic limit.

6.1 Step 1: Tightness of the fluctuation field

Recall from Remark 5.14 that to prove the tightness of $(\mathcal{Y}^N)_{N\in\mathbb{N}}$ in $D([0,T], H^\alpha)$ it suffices to verify two conditions:

i. For all $t \in [0,T]$ the sequence $(\mathcal{Y}_t^N)_{N\in\mathbb{N}}$ is tight in H^α;

ii. for all $\varepsilon > 0$ we have

$$\inf_{r>0}\limsup_{N\to\infty} \mathbb{P}_{\mu_\rho^N}(w(\mathcal{Y}^N, r) \geqslant \varepsilon) = 0.$$

To carry out step i., we wirst need to understand compact sets in H^α:

Lemma 6.4. *Let $\alpha \in \mathbb{R}$ and $\beta > \alpha$ and let $A \subset H^\alpha$ be such that*

$$\sup_{\mathcal{Z}\in A} \|\mathcal{Z}\|_{H^\beta} < \infty.$$

Then A is relatively compact.

Proof. Since H^α is a metric space, we have to show that any sequence $(\mathcal{Z}_n)_{n\in\mathbb{N}} \subset A$ has a convergent subsequence. This follows from a diagonal sequence argument: First observe that for $m \in \mathbb{Z}^d$

$$|\widehat{\mathcal{Z}_n}(m)| \leqslant (1+|m|^2)^{-\beta/2}\left(\sum_{\ell\in\mathbb{Z}^d} (1+|\ell|^2)^\beta |\widehat{\mathcal{Z}_n}(\ell)|^2\right)^{1/2} \leqslant (1+|m|^2)^{-\beta/2} \sup_{\mathcal{Z}\in A} \|\mathcal{Z}\|_{H^\beta}$$

is uniformly bounded in n. So let $m_1, m_2, \ldots$ be an enumeration of $\mathbb{Z}^d$ and start with a subsequence $(n_1(k))_k$ such that $\widehat{\mathcal{Z}_{n_1(k)}}(m_1)$ converges to some $\hat{\mathcal{Z}}(m_1)$ for $k \to \infty$. Next let $(n_2(k))_k$ be a subsequence of $(n_1(k))_k$ such that $\widehat{\mathcal{Z}_{n_2(k)}}(m_2)$ converges to some $\hat{\mathcal{Z}}(m_2)$ for $k \to \infty$, and continue by selecting subsequences in that manner. Then for all $i \in \mathbb{N}$ we have convergence along the diagonal sequence,

$$\lim_{k\to\infty} \widehat{\mathcal{Z}_{n_k(k)}}(m_i) = \hat{\mathcal{Z}}(m_i).$$

It remains to show that $\hat{\mathcal{Z}}(m_i) \in H^\alpha$ and that the convergence holds with respect to the H^α norm. By Fatou's lemma we even have

$$\sum_{m\in\mathbb{Z}^d} (1+|m|^2)^\beta |\hat{\mathcal{Z}}(m)|^2 \leqslant \liminf_{k\to\infty} \sum_{m\in\mathbb{Z}^d} (1+|m|^2)^\beta |\hat{\mathcal{Z}}_{n_k(k)}(m)|^2 \leqslant \sup_{\mathcal{Y}\in A} \|\mathcal{Y}\|_{H^\beta}^2 < \infty.$$

Moreover, we have for all $C > 0$

$$\begin{aligned}
\lim_{k\to\infty} \|\mathcal{Z} - \mathcal{Z}_{n_k(k)}\|_{H^\alpha}^2 &= \lim_{k\to\infty} \sum_{m\in\mathbb{Z}^d} (1+|m|^2)^\alpha |\hat{\mathcal{Z}}(m) - \widehat{\mathcal{Z}_{n_k(k)}}(m)|^2 \\
\leqslant \lim_{k\to\infty} \Bigg(\sum_{|m|\leqslant C} (1+|m|^2)^\alpha |\hat{\mathcal{Z}}(m) - \widehat{\mathcal{Z}_{n_k(k)}}(m)|^2 &+ \sum_{|m|>C} (1+|m|^2)^\alpha |\hat{\mathcal{Z}}(m) - \widehat{\mathcal{Z}_{n_k(k)}}(m)|^2 \Bigg) \\
&\lesssim \lim_{k\to\infty} \sum_{|m|\leqslant C} (1+|m|^2)^\alpha |\hat{\mathcal{Z}}(m) - \widehat{\mathcal{Z}_{n_k(k)}}(m)|^2 \\
&\quad + \lim_{k\to\infty} (1+C^2)^{\alpha-\beta} \sum_{|m|>C} (1+|m|^2)^\beta \Big(|\hat{\mathcal{Z}}(m)|^2 + |\widehat{\mathcal{Z}_{n_k(k)}}(m)|^2 \Big) \\
&\lesssim \lim_{k\to\infty} \sum_{|m|\leqslant C} (1+|m|^2)^\alpha |\hat{\mathcal{Z}}(m) - \widehat{\mathcal{Z}_{n_k(k)}}(m)|^2 + (1+C^2)^{\alpha-\beta} \sup_{\mathcal{Z}\in A} \|\mathcal{Z}\|_{H^\beta}^2 \\
&\lesssim (1+C^2)^{\alpha-\beta}.
\end{aligned}$$

Since $\alpha - \beta < 0$ and $C > 0$ is arbitrary, we see that $\mathcal{Z}_{n_k(k)}$ converges to $\mathcal{Z}$ in H^α. □

Lemma 6.5. *For all $\alpha < -d/2$ and $t \in [0,T]$ the sequence $(\mathcal{Y}_t^N)_{N\in\mathbb{N}}$ is tight in H^α and*

$$\sup_{N\in\mathbb{N}, t\in[0,T]} \mathbb{E}_{\mu_\rho^N}[\|\mathcal{Y}_t^N\|_{H^\alpha}^2] < \infty.$$

Proof. Let $\beta \in (\alpha, -d/2)$. Since bounded sets in H^β are relatively compact in H^α it suffices to bound $\mathbb{E}[\|\mathcal{Y}_t^N\|_{H^\beta}^2]$ uniformly in N and t. But

$$\begin{aligned}
\mathbb{E}_{\mu_\rho^N}[\|\mathcal{Y}_t^N\|_{H^\beta}^2] &= \sum_{m\in\mathbb{Z}^d} (1+|m|^2)^\beta \mathbb{E}_{\mu_\rho^N}[|\mathcal{Y}_t^N(e_m)|^2] \\
= \sum_{m\in\mathbb{Z}^d} (1+|m|^2)^\beta N^{-d} \mathbb{E}_{\mu_\rho^N}&\left[\left(\sum_{x\in\mathbb{Z}_N^d} e_m\Big(\frac{x}{N}\Big)(\eta_{N^2 t}^N(x) - \rho) \right) \overline{\left(\sum_{x\in\mathbb{Z}_N^d} e_m\Big(\frac{x}{N}\Big)(\eta_{N^2 t}^N(x) - \rho) \right)} \right] \\
= \sum_{m\in\mathbb{Z}^d} (1+|m|^2)^\beta N^{-d} &\sum_{x\in\mathbb{Z}_N^d} \left| e_m\Big(\frac{x}{N}\Big) \right|^2 \rho(1-\rho) \leqslant \sum_{m\in\mathbb{Z}^d} (1+|m|^2)^\beta \rho(1-\rho),
\end{aligned}$$

where we used that under μ_ρ^N the variables $\eta_{N^2 t}^N(x)$ and $\eta_{N^2 t}^N(y)$ for $x \neq y$ are independent and that the Bernoulli(ρ) distribution has variance $\rho(1-\rho)$. As before, the right hand side is finite and this concludes the proof. □

To prove the dynamic tightness, note that $\sum_{x\in\mathbb{Z}_N^d} \rho \Delta_{x,y}^N \varphi = 0$ because it is a telescope sum with periodic boundary conditions, and therefore we obtain as in (5.8)

$$\mathcal{Y}_t^N(\varphi) = \mathcal{Y}_0^N(\varphi) + \mathcal{A}_t^N(\varphi) + \mathcal{M}_t^N(\varphi), \tag{6.2}$$

where

$$\mathcal{A}_t^N(\varphi) := \int_0^t \frac{1}{2} \sum_{y\in\mathbb{Z}_N^d} p^N(y) \mathcal{Y}_r^N(\Delta_{N\cdot,y}^N \varphi) \mathrm{d}r, \qquad \mathcal{M}_t^N(\varphi) := N^{d/2} \bar{M}_t^{N,\varphi}$$

for the martingale $\bar{M}_t^{N,\varphi}$ of (5.8).

Lemma 6.6. *For all* $\alpha < -d/2 - 2$ *and* $\kappa \in [0, T]$ *we have*

$$\limsup_{N \to \infty} \mathbb{E}_{\mu_\rho^N}[w_{H^\alpha}(\mathcal{A}^N, \kappa)^2] \lesssim \kappa,$$

where the modulus of continuity w_{H^α} *is taken with respect to the* H^α *norm.*

Proof. To control the drift, note that for $[s, t] \subset [0, T]$ by the Cauchy-Schwarz inequality

$$\begin{aligned} \|\mathcal{A}_t^N - \mathcal{A}_s^N\|_{H^\alpha}^2 = & \left\| \int_s^t \frac{1}{2} \sum_{y \in \mathbb{Z}_N^d} p^N(y) \mathcal{Y}_r^N(\Delta_{N \cdot, y}^N \cdot) \mathrm{d}r \right\|_{H^\alpha}^2 \leqslant \left(\int_s^t \sum_{y \in \mathbb{Z}_N^d} p^N(y) \|\mathcal{Y}_r^N(\Delta_{N \cdot, y}^N \cdot)\|_{H^\alpha} \mathrm{d}r \right)^2 \\ \leqslant & |t - s| \int_s^t \left(\sum_{y \in \mathbb{Z}_N^d} p^N(y) \|\mathcal{Y}_r^N(\Delta_{N \cdot, y}^N \cdot)\|_{H^\alpha} \right)^2 \mathrm{d}r \\ \leqslant & |t - s| \int_0^T \left(\sum_{y \in \mathbb{Z}_N^d} p^N(y) |1|^2 \right) \left(\sum_{y \in \mathbb{Z}_N^d} p^N(y) \|\mathcal{Y}_r^N(\Delta_{N \cdot, y}^N \cdot)\|_{H^\alpha}^2 \right) \mathrm{d}r \\ \lesssim & |t - s| \int_0^T \sum_{y \in \mathbb{Z}_N^d} p^N(y) \|\mathcal{Y}_r^N(\Delta_{N \cdot, y}^N \cdot)\|_{H^\alpha}^2 \mathrm{d}r, \end{aligned}$$

and since $\|\Delta_{N \cdot, y}^N e_m\|_\infty \lesssim |y|^2 \|\mathrm{D}^2 e_m\|_\infty \lesssim |y|^2 |m|^2$, we show as in Lemma 6.5 that for all $\alpha < -d/2 - 2$

$$\sup_{N \in \mathbb{N}, r \in [0, T]} \mathbb{E}[\|\mathcal{Y}_r^N(\Delta_{N \cdot, y}^N \cdot)\|_{H^\alpha}^2] \lesssim |y|^2,$$

which leads to

$$\begin{aligned} & \inf_{\kappa > 0} \limsup_{N \to \infty} \mathbb{E}_{\mu_\rho^N} \left(w \left(\int_0^\cdot \frac{1}{2} \sum_{y \in \mathbb{Z}_N^d} p^N(y) \mathcal{Y}_r^N(\Delta_{N \cdot, y}^N \cdot) \mathrm{d}r, \kappa \right)^2 \right) \\ = & \inf_{\kappa > 0} \limsup_{N \to \infty} \mathbb{E}_{\mu_\rho^N} \left[\sup_{|t - s| \leqslant \kappa} \left\| \int_s^t \frac{1}{2} \sum_{y \in \mathbb{Z}_N^d} p^N(y) \mathcal{Y}_r^N(\Delta_{N \cdot, y}^N \cdot) \mathrm{d}r \right\|_{H^\alpha}^2 \right] \lesssim \inf_{\kappa > 0} \kappa = 0. \end{aligned}$$

□

The martingale $\mathcal{M}^N$ is significantly more difficult to control: Since we multiplied $\bar{M}^{N,\varphi}$ with $N^{d/2}$, even for fixed φ the martingale contribution $\mathcal{M}^N(\varphi)$ will not vanish as $N \to \infty$ and the crude arguments based on Lemma 5.17 no longer work. The reason why Lemma 5.17 is not sufficient is that on the right hand side we only have a factor $|t - s|$. If we would know that $\mathbb{E}[|X_t - X_s|^p] \leqslant C|t - s|^{1 + \kappa}$, then we would get

$$\mathbb{E}[w(X, r)^p] \lesssim C r^\kappa.$$

This is reminiscent of Kolmogorov's continuity criterion, where a control of the second moments of the Brownian motion does not suffice to show that it is continuous, but we need to go to higher moments. To bound higher moments of our martingale, we rely on the following martingale inequality:

Lemma 6.7. *Let* $(M_t)_{t \in [0, T]}$ *be a martingale with bounded jumps. Then for all* $0 \leqslant s < t \leqslant T$

$$\mathbb{E}\left[\sup_{r \in [s, t]} |M_r - M_s|^4 \right] \lesssim \sup_{r \in [s, t], \omega \in \Omega} |\Delta M_r(\omega)|^2 \mathbb{E}[|\langle M \rangle_t - \langle M \rangle_s|^2]^{1/2} + \mathbb{E}[|\langle M \rangle_t - \langle M \rangle_s|^2],$$

where $\langle M \rangle$ *denotes the predictable quadratic variation.*

Proof. See [Jacod, Shiryaev, Lemma VII.3.34]. □

Lemma 6.8. *For all $\alpha < -d/2 - 1$ and $\kappa \in [0, T]$ we have*

$$\limsup_{N \to \infty} \mathbb{E}_{\mu_\rho^N}[w_{H^\alpha}(\mathcal{M}^N, \kappa)^4] \lesssim \kappa,$$

where the modulus of continuity w_{H^α} is taken with respect to the H^α norm.

Proof. Note that for all $\lambda > 0$ by the Cauchy-Schwarz inequality

$$\sup_{r \in [s,t]} \|\mathcal{M}_r^N - \mathcal{M}_s^N\|_{H^\alpha}^4 = \sup_{r \in [s,t]} \left| \sum_{m \in \mathbb{Z}^d} (1 + |m|^2)^\alpha |\mathcal{M}_r^N(e_m) - \mathcal{M}_s^N(e_m)|^2 \right|^2$$
$$\leqslant \sup_{r \in [s,t]} \left(\sum_{m \in \mathbb{Z}^d} (1 + |m|^2)^{-d/2-\lambda} \sum_{m \in \mathbb{Z}^d} (1 + |m|^2)^{2\alpha + d/2 + \lambda} |\mathcal{M}_r^N(e_m) - \mathcal{M}_s^N(e_m)|^4 \right)$$
$$\lesssim \sum_{m \in \mathbb{Z}^d} (1 + |m|^2)^{2\alpha + d/2 + \lambda} \sup_{r \in [s,t]} |\mathcal{M}_r^N(e_m) - \mathcal{M}_s^N(e_m)|^4.$$

To apply Lemma 6.7, we would have to split $\mathcal{M}^N(e_m)$ into its real and imaginary parts and argue for both separately. For simplicity we skip this step and pretend that e_m is real valued. Note that by definition of $\mathcal{M}^N$

$$|\Delta \mathcal{M}_r^N(e_m)| = |\Delta \mathcal{Y}_r^N(e_m)| = \left| N^{-d/2} \sum_x e_m\left(\frac{x}{N}\right) \Delta(\eta_{N^2 \cdot}^N(x))_r \right| \leqslant 2N^{-d/2},$$

where we used that at any given time at most one particle can jump and therefore $\eta_{N^2 \cdot}^N(x)$ can only change for two values of x. Moreover, Lemma 5.4 gives

$$|\langle \mathcal{M}^N(e_m) \rangle_t - \langle \mathcal{M}^N(e_m) \rangle_s|^2 = \left| \int_s^t \frac{1}{N^d} \sum_{x,y} \eta_{N^2 r}^N(x)(1 - \eta_{N^2 r}^N(x + y)) p^N(y) [\nabla_{x,y}^N e_m]^2 \mathrm{d}r \right|^2$$
$$\leqslant |t - s| N^{-2d} \int_s^t \left| \sum_{x,y} p^N(y) [\nabla_{x,y}^N e_m]^2 \right|^2 \mathrm{d}r \leqslant |t - s| N^{-2d} |t - s| \times (N^d \|\nabla e_m\|_{L^\infty}^2)^2 \lesssim |t - s|^2 |m|^4.$$

Combining all of the above with Lemma 6.7, we get for all $[s, t] \subset [0, T]$ and for all $\alpha < -d/2 - \lambda/2 - 1$

$$\mathbb{E}_{\mu_\rho^N}\left[\sup_{r \in [s,t]} \|\mathcal{M}_r^N - \mathcal{M}_s^N\|_{H^\alpha}^4\right] \lesssim \sum_{m \in \mathbb{Z}^d} (1 + |m|^2)^{2\alpha + d/2 + \lambda} \mathbb{E}_{\mu_\rho^N}\left[\sup_{r \in [s,t]} |\mathcal{M}_r^N(e_m) - \mathcal{M}_s^N(e_m)|^4\right]$$
$$\lesssim \sum_{m \in \mathbb{Z}^d} (1 + |m|^2)^{2\alpha + d/2 + \lambda} [N^{-d} |t - s| |m|^2 + |t - s|^2 |m|^4]$$
$$\lesssim N^{-d} |t - s| + |t - s|^2,$$

where in the last step we used that $2\alpha + d/2 + \lambda + 2 < -d/2$ if and only if $\alpha < -d/2 - \lambda/2 - 1$. From here Lemma 5.17, which extends without difficulty from real valued to Banach space valued processes, gives

$$\limsup_{N \to \infty} \mathbb{E}_{\mu_\rho^N}[w_{H^\alpha}(\mathcal{M}^N, \kappa)^4] \lesssim \limsup_{N \to \infty} (N^{-d} + \kappa) = \kappa.$$

□

Lemma 6.9. *Let $\alpha < -d/2 - 2$. Then the triple of processes $(\mathcal{Y}^N, \mathcal{A}^N, \mathcal{M}^N)$ is tight in $D([0, T], (H^\alpha)^3)$, and any limit point is supported on $C([0, T], (H^\alpha)^3)$*

Proof. Lemma 6.5, Lemma 6.6 and Lemma 6.8 in combination with Remark 5.14 show that $\mathcal{Y}^N$ is tight in $C([0, T], H^\alpha)$ (simply send $\kappa \to 0$ in Lemma 6.6 and Lemma 6.8). Moreover, since $\mathcal{A}_0^N = 0$ and $\mathcal{M}_0^N = 0$ we have

$$\|\mathcal{A}_t^N\|_{H^\alpha} + \|\mathcal{M}_t^N\|_{H^\alpha} \leqslant w_{H^\alpha}(\mathcal{A}^N, t) + w_{H^\alpha}(\mathcal{M}^N, t),$$

which shows that

$$\sup_{N \in \mathbb{N}, t \in [0,T]} \mathbb{E}[\|\mathcal{A}_t^N\|_{H^\alpha} + \|\mathcal{M}_t^N\|_{H^\alpha}] < \infty,$$

from where we get the relative compactness at fixed times as in Lemma 6.5. Thus, the triple $(\mathcal{Y}^N, \mathcal{A}^N, \mathcal{M}^N)$ is tight. □

6.2 Step 2: Description of the possible limit points

According to Lemma 6.9 we can now find subsequences $(N(k))_{k\in\mathbb{N}}$ such that $(\mathcal{Y}^{N(k)}, \mathcal{A}^{N(k)}, \mathcal{M}^{N(k)})_{k\in\mathbb{N}}$ converges weakly in $D([0,T],(H^\alpha)^3)$ to a limit $(\mathcal{Y}, \mathcal{A}, \mathcal{M})$ that is almost surely in $C([0,T],(H^\alpha)^3)$. Recall that any element of H^α can be identified with a map on $C^\infty(\mathbb{T}^d)$. So to describe the law of $\mathcal{Y}$ we will first study $\mathcal{Y}(\varphi)$ for fixed test functions $\varphi\in C^\infty(\mathbb{T}^d)$.

Lemma 6.10. *We have almost surely for all $\varphi\in C^\infty(\mathbb{T}^d)$ and for all $t\in[0,T]$*

$$\mathcal{A}_t(\varphi) = \int_0^t \frac{1}{2}\mathcal{Y}_s(\Delta_C\varphi)\mathrm{d}s,$$

and

$$\mathcal{Y}_t(\varphi) = \mathcal{Y}_0(\varphi) + \int_0^t \frac{1}{2}\mathcal{Y}_s(\Delta_C\varphi)\mathrm{d}s + \mathcal{M}_t(\varphi).$$

Proof. We only have to show that $\mathcal{A}_t(\varphi) = \int_0^t \frac{1}{2}\mathcal{Y}_s(\Delta_C\varphi)\mathrm{d}s$. For that purpose note that

$$\left\| \sum_{y\in\mathbb{Z}_N^d} p^N(y)\Delta^N_{N\cdot,y} e_m - \Delta_C e_m \right\|_{L^\infty} \lesssim \|\mathrm{D}^3 e_m\|_{L^\infty} \lesssim |m|^3 N^{-1},$$

while also

$$\left\| \sum_{y\in\mathbb{Z}_N^d} p^N(y)\Delta^N_{N\cdot,y} e_m - \Delta_C e_m \right\|_{L^\infty} \leqslant \left\| \sum_{y\in\mathbb{Z}_N^d} p^N(y)\Delta^N_{N\cdot,y} e_m \right\|_{L^\infty} + \|\Delta_C e_m\|_{L^\infty} \lesssim |m|^2.$$

Now let $\lambda\in[0,1]$ and interpolate these two inequalities:

$$\begin{aligned} &\left\| \sum_{y\in\mathbb{Z}_N^d} p^N(y)\Delta^N_{N\cdot,y} e_m - \Delta_C e_m \right\|_{L^\infty} \\ =& \left\| \sum_{y\in\mathbb{Z}_N^d} p^N(y)\Delta^N_{N\cdot,y} e_m - \Delta_C e_m \right\|_{L^\infty}^\lambda \left\| \sum_{y\in\mathbb{Z}_N^d} p^N(y)\Delta^N_{N\cdot,y} e_m - \Delta_C e_m \right\|_{L^\infty}^{1-\lambda} \\ &\lesssim (|m|^3 N^{-1})^\lambda (|m|^2)^{1-\lambda} = |m|^2 |m|^\lambda N^{-\lambda}. \end{aligned}$$

We pick $\lambda>0$ such that $\alpha < -d/2 - 1 - \lambda$ and then we argue as in Lemma 5.19 (also using some of the steps in Lemma 6.6) to obtain

$$\limsup_{N\to\infty} \mathbb{E}_{\mu_\rho^N}\left[\sup_{t\in[0,T]} \left\| \mathcal{A}_t^N - \int_0^t \frac{1}{2}\mathcal{Y}_r^N(\Delta_C\cdot)\mathrm{d}r \right\|_{H^\alpha}^2\right] = 0.$$

Now

$$D([0,T],H^\alpha) \ni \mathcal{Z} \mapsto \int_0^\cdot \frac{1}{2}\mathcal{Z}_r(\Delta_C\cdot)\mathrm{d}r \in C([0,T],H^{\alpha-2})$$

is a continuous functional (which is not difficult to check), so by the continuous mapping theorem $\int_0^\cdot \frac{1}{2}\mathcal{Y}_r^N(\Delta_C\cdot)\mathrm{d}r$ converges weakly in $C([0,T],H^{\alpha-2})$ to $\int_0^\cdot \frac{1}{2}\mathcal{Y}_r(\Delta_C\cdot)\mathrm{d}r$, which concludes the proof. □

Next, let's examine the limiting initial condition $\mathcal{Y}_0$. Recall that if $(X,\mathcal{B},\nu)$ is a measure space, then a *white noise* on X is the centered Gaussian process $(\xi(\varphi))_{\varphi\in L^2(X,\nu)}$ that has covariance

$$\mathbb{E}[\xi(\varphi)\xi(\psi)] = \langle\varphi,\psi\rangle_{L^2}.$$

Lemma 6.11. *The process $(\mathcal{Y}_0(\varphi))_{\varphi\in C^\infty(\mathbb{T}^d)}$ is centered and Gaussian, and has covariance*

$$\mathbb{E}[\mathcal{Y}_0(\varphi)\mathcal{Y}_0(\psi)] = \rho(1-\rho)\langle\varphi,\psi\rangle_{L^2}$$

In other words, $\mathcal{Y}_0$ is a multiple of the white noise on $\mathbb{T}^d$ (restricted from $L^2(\mathbb{T}^d)$ to $C^\infty(\mathbb{T}^d)$).

Proof. To show that $\mathcal{Y}_0$ is centered Gaussian, we have to show that for all $n \in \mathbb{N}$ and all $\varphi_1, \ldots, \varphi_n \in C^\infty(\mathbb{T}^d)$ and all $\kappa_1, \ldots, \kappa_n \in \mathbb{R}$ the random variable $\sum_{k=1}^n \kappa_k \mathcal{Y}_0(\varphi_k)$ is centered Gaussian. But since $\mathcal{Y}_0$ is linear we have

$$\sum_{k=1}^n \kappa_k \mathcal{Y}_0(\varphi_k) = \mathcal{Y}_0\left(\sum_{k=1}^n \kappa_k \varphi_k\right),$$

and therefore it suffices to show for fixed $\varphi \in C^\infty(\mathbb{T}^d)$ that $\mathcal{Y}_0(\varphi)$ is centered Gaussian. Recall that

$$\mathcal{Y}_0^N(\varphi) = \frac{1}{N^{d/2}} \sum_{x \in \mathbb{Z}_N^d} \varphi\left(\frac{x}{N}\right)(\eta_0^N(x) - \rho),$$

where $(\eta_0^N(x))_{x \in \mathbb{Z}_N^d}$ is an i.i.d. family of Bernoulli variables with parameter ρ, so setting $X_k^{N^d} := \varphi\left(\frac{x_k}{N}\right)(\eta_0^N(x_k) - \rho)$ for an enumeration $x_1, \ldots, x_{N^d}$ of $\mathbb{Z}_N^d$ we are in the setting of Proposition 6.12 below (note that the third moment of the Bernoulli(ρ) distribution is of course finite and that φ is bounded). The variance for fixed N is

$$\mathbb{E}[\mathcal{Y}_0^N(\varphi)^2] = \frac{1}{N^d} \sum_{x \in \mathbb{Z}_N^d} \varphi\left(\frac{x}{N}\right)^2 \rho(1-\rho) \longrightarrow \rho(1-\rho) \int_{\mathbb{T}^d} \varphi(x)^2 \mathrm{d}x,$$

and therefore $\mathcal{Y}_0(\varphi) \sim \mathcal{N}(0, \rho(1-\rho)\|\varphi\|_{L^2}^2)$ by Proposition 6.12. From here polarization yields

$$\begin{aligned}\operatorname{cov}(\mathcal{Y}_0(\varphi), \mathcal{Y}_0(\psi)) &= \frac{1}{2}(\operatorname{var}(\mathcal{Y}_0(\varphi+\psi)) - \operatorname{var}(\mathcal{Y}_0(\varphi)) - \operatorname{var}(\mathcal{Y}_0(\psi))) \\ &= \frac{1}{2}(\|\varphi+\psi\|_{L^2}^2 - \|\varphi\|_{L^2}^2 - \|\psi\|_{L^2}^2) = \langle \varphi, \psi \rangle_{L^2},\end{aligned}$$

which concludes the proof. □

Proposition 6.12. *(A variant of the central limit theorem)*

Consider a double sequence $(X_k^N)_{N \in \mathbb{N}, k=1,\ldots,N}$ of independent random variables, such that X_k^N is centered for all N, k and such that

$$\sup_{N \in \mathbb{N}, k=1,\ldots,N} \mathbb{E}[|X_k^N|^3] < \infty.$$

Assume furthermore that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^N \mathbb{E}[|X_k^N|^2] = \sigma^2$$

for some $\sigma^2 \in \mathbb{R}_+$. Then

$$\frac{1}{\sqrt{N}} \sum_{k=1}^N X_k^N$$

converges in distribution to a $\mathcal{N}(0, \sigma^2)$ variable.

Proof. Since in Exercise 4.1 we already gave a characteristic function proof of a central limit theorem, we now use a different method: The "swapping trick" of Lindeberg. For that purpose let $F \in C_b^3(\mathbb{R})$, i.e. F and its derivatives up to order 3 are bounded. By approximation it suffices to show that

$$\lim_{N \to \infty} \mathbb{E}\left[F\left(\frac{1}{\sqrt{N}} \sum_{k=1}^N X_k^N\right)\right] = \int F(x) \mathcal{N}(0, \sigma^2)(\mathrm{d}x).$$

Note that if the (X_k^N) were Gaussian, then $\frac{1}{\sqrt{N}} \sum_{k=1}^N X_k^N$ would be centered Gaussian as a sum of independent centered Gaussians, and since its variance converges to σ^2 by assumption, the claimed convergence would be immediate. Of course, the (X_k^N) are not Gaussian in general and therefore the idea is to replace them successively by independent Gaussian variables $(Z_k^N)_{N \in \mathbb{N}, k \in \mathbb{N}}$, also independent of the (X_k^N), and such that

$$\mathbb{E}[Z_k^N] = 0, \qquad \mathbb{E}[|Z_k^N|^2] = \mathbb{E}[|X_k^N|^2]$$

for all N,k. Let us write

$$S_{k,\ell}^N(X) := \frac{1}{\sqrt{N}}\sum_{j=k}^{\ell} X_j^N, \qquad S_{k,\ell}^N(Z) := \frac{1}{\sqrt{N}}\sum_{j=k}^{\ell} Z_j^N.$$

The proof is complete once we show $\mathbb{E}[F(S_{1,N}^N(X))] - \mathbb{E}[F(S_{1,N}^N(Z))] \to 0$. For that purpose we use a telescope sum that successively swaps the X variables into Z variables:

$$\begin{aligned}F(S_{1,N}^N(X)) - F(S_{1,N}^N(Z)) = &[F(S_{1,N}^N(X)) - F(S_{2,N}^N(X) + S_{1,1}^N(Z))]\\ &+[F(S_{2,N}^N(X) + S_{1,1}^N(Z)) - F(S_{3,N}^N(X) + S_{1,2}^N(Z))]\\ &+[F(S_{3,N}^N(X) + S_{1,2}^N(Z)) - F(S_{4,N}^N(X) + S_{1,3}^N(Z))]\\ &+\ldots\\ &+[F(S_{N,N}^N(X) + S_{1,N-1}^N(Z)) - F(S_{1,N}^N(Z))].\end{aligned}$$

For fixed k we get by a Taylor expansion

$$\begin{aligned}&F(S_{k,N}^N(X) + S_{1,k-1}^N(Z)) - F(S_{k+1,N}^N(X) + S_{1,k}^N(Z))\\ =&F(S_{k+1,N}^N(X) + S_{1,k-1}^N(Z) + N^{-1/2}X_k^N) - F(S_{k+1,N}^N(X) + S_{1,k-1}^N(Z))\\ &+F(S_{k+1,N}^N(X) + S_{1,k-1}^N(Z)) - F(S_{k+1,N}^N(X) + S_{1,k-1}^N(Z) + N^{-1/2}Z_k^N)\\ =&F'(S_{k+1,N}^N(X) + S_{1,k-1}^N(Z))N^{-1/2}(X_k^N - Z_k^N)\\ &+\frac{1}{2}F''(S_{k+1,N}^N(X) + S_{1,k-1}^N(Z))[(N^{-1/2}X_k^N)^2 - (N^{-1/2}Z_k^N)^2]\\ &+O(N^{-3/2})(X_k^N)^3 + O(N^{-3/2})(Z_k^N)^3,\end{aligned}$$

and since X_k^N and Z_k^N are independent of $S_{k+1,N}^N(X) + S_{1,k-1}^N(Z)$ and they have the same first and second moments and bounded third moments (recall that for any centered Gaussian variable Y we have $\mathbb{E}[|Y|^p] = c_p\mathbb{E}[|Y|^2]^{p/2}$ for all $p \in \mathbb{N}$), we get

$$|\mathbb{E}[F(S_{k,N}^N(X) + S_{1,k-1}^N(Z)) - F(S_{k+1,N}^N(X) + S_{1,k}^N(Z))]| = O(N^{-3/2}).$$

Since we have to add N values of k, we obtain

$$|\mathbb{E}[F(S_{1,N}^N(X))] - \mathbb{E}[F(S_{1,N}^N(Z))]| = O(N^{-1/2}),$$

which is what we set out to show. □

Exercise 6.2. Review the proof of the previous proposition. Which property of the normal distribution is the main reason that the proof works? Could we also swap the (X_k^N) with (Y_k^N) that are not normally distributed but that match the first two moments of the (X_k^N) and that have bounded third moments? Which parts of the proof would still work in that case?

Solution. *The key property that we used in the proof is that the sum of two independent Gaussians is still Gaussian. This is not true for other distributions. Indeed, you can use the (classical) central limit theorem to show that if ν is a probability distribution on $\mathbb{R}$ with unit variance, such that whenever $X, Y \sim \nu$ are independent also $\alpha X + \beta Y \sim \nu$ for all $\alpha, \beta > 0$ with $\alpha^2 + \beta^2 = 1$, then $\nu = \mathcal{N}(0,1)$.*

Exercise 6.3. There is actually no need to normalize the sum by $N^{-1/2}$, we could pull this factor into X_k^N (as $\tilde{X}_k^N := N^{-1/2}X_k^N$, although in the following we omit the $\tilde{\cdot}$) and instead assume that

$$\sum_{k=1}^{N} \mathbb{E}[|X_k^N|^2] \longrightarrow \sigma^2. \tag{6.3}$$

Find further conditions on the (X_k^N) under which the convergence

$$\sum_{k=1}^{N} X_k^N \longrightarrow \mathcal{N}(0,\sigma^2) \tag{6.4}$$

can be shown with minor modifications to the arguments in the proof of Proposition 6.12.

Hint: Clearly any family with $X_k^N=0$ for $k=2,\dots,N$ and $\mathbb{E}[|X_1^N|^2]=\sigma^2$ satisfies (6.3) but in general not (6.4).

Solution. *The same proof as for Proposition 6.12 works if we assume*

$$\lim_{N\to\infty}\sum_{k=1}^N \mathbb{E}[|X_k^N|^3]=0.$$

Indeed, in that case we also get for the Gaussians (Z_k^N) that have the same variance as the (X_k^N)

$$\lim_{N\to\infty}\sum_{k=1}^N \mathbb{E}[|Z_k^N|^3]=\lim_{N\to\infty} c_3\sum_{k=1}^N \mathbb{E}[|Z_k^N|^2]^{3/2}=\lim_{N\to\infty} c_3\sum_{k=1}^N \mathbb{E}[|X_k^N|^2]^{3/2}\leqslant \lim_{N\to\infty} c_3\sum_{k=1}^N \mathbb{E}[|X_k^N|^3]=0,$$

and therefore the sum over k of the remainder terms in the Taylor expansion of

$$F(S_{k,N}^N(X)+S_{1,k-1}^N(Z))-F(S_{k+1,N}^N(X)+S_{1,k}^N(Z))$$

does not contribute.

If we do not perform the Taylor expansion to third order but instead take the remainder as

$$|X_k^N|^2\int_0^1(1-\lambda)[F''(S_{k+1,N}^N(X)+S_{1,k-1}^N(Z)+\lambda X_k^N)-F''(S_{k+1,N}^N(X)+S_{1,k-1}^N(Z))]\mathrm{d}\lambda$$

and similarly for (Z_k^N), then we can also show that it suffices to assume the Lindeberg condition

$$\lim_{N\to\infty}\sum_{k=1}^N \mathbb{E}[|X_k^N|^2\mathbb{I}_{|X_k^N|>\delta}]=0$$

for all $\delta>0$. Indeed, given $\varepsilon>0$ we can find $\delta>0$ such that $|F''(x)-F''(y)|\leqslant\varepsilon$ whenever $|x-y|\leqslant\delta$ (note that F'' is uniformly continuous because F''' is bounded). Distinguishing the two alternatives $|X_k^N|>\delta$ and $|X_k^N|\leqslant\delta$, we thus have

$$\left|\mathbb{E}\left[|X_k^N|^2\int_0^1(1-\lambda)[F''(S_{k+1,N}^N(X)+S_{1,k-1}^N(Z)+\lambda X_k^N)-F''(S_{k+1,N}^N(X)+S_{1,k-1}^N(Z))]\mathrm{d}\lambda\right]\right|$$
$$\leqslant 2\|F''\|_\infty\mathbb{E}[|X_k^N|^2\mathbb{I}_{|X_k^N|>\delta}]+\mathbb{E}[|X_k^N|^2]\frac{\varepsilon}{2},$$

and the sum over k is in the limit $\lesssim\varepsilon$ by assumption, so since $\varepsilon>0$ was arbitrary it does not contribute. For the Gaussians we can even control the third moments as before: We have

$$\lim_{N\to\infty}\max_{k=1,\dots,N}\mathbb{E}[|X_k^N|^2]\leqslant\lim_{N\to\infty}\max_{k=1,\dots,N}\mathbb{E}[|X_k^N|^2\mathbb{I}_{|X_k^N|>\delta}]+\delta^2\leqslant\delta^2$$

for all $\delta>0$, so that the maximal variance converges to zero. Thus, we get

$$\sum_{k=1}^N \mathbb{E}[|Z_k^N|^3]\simeq\sum_{k=1}^N \mathbb{E}[|Z_k^N|^2]^{3/2}=\sum_{k=1}^N \mathbb{E}[|X_k^N|^2]^{3/2}\leqslant\max_{k=1,\dots,N}\mathbb{E}[|X_k^N|^2]^{1/2}\sum_{k=1}^N \mathbb{E}[|X_k^N|^2]\longrightarrow 0,$$

and then we can argue as above (taking the Taylor expansion to third order) to show that the remainder coming from the (Z_k^N) vanishes in the limit.

So far we identified the limiting drift $\mathcal{A}=\lim_{k\to\infty}\mathcal{A}^{N(k)}$ and the limiting initial condition $\mathcal{Y}_0=\lim_{k\to\infty}\mathcal{Y}_0^{N(k)}$. It remains to describe $\mathcal{M}=\lim_{k\to\infty}\mathcal{M}^{N(k)}$. We would like to show that $\mathcal{M}(\varphi)$ is a continuous martingale for all $\varphi\in C^\infty(\mathbb{T}^d)$ (in a suitable filtration), and we would like to identify its quadratic variation. For that purpose we need some auxiliary results:

Lemma 6.13. *(Fatou's lemma for convergence in distribution)*

Let $(X_N)_{N\in\mathbb{N}}$ be a family of random variables with values in a Polish space E that converges in distribution to a random variable X. Let $F\colon E\to\mathbb{R}$ be continuous and bounded from below. Then

$$\mathbb{E}[F(X)]\leqslant\liminf_{N\to\infty}\mathbb{E}[F(X_N)].$$

Proof. By the monotone convergence theorem

$$\mathbb{E}[F(X)] = \sup_{C>0} \mathbb{E}[F(X) \wedge C],$$

and for $C > 0$ the function $F \wedge C$ is continuous and bounded and therefore

$$\mathbb{E}[F(X) \wedge C] = \lim_{N\to\infty} \mathbb{E}[F(X_N) \wedge C] = \liminf_{N\to\infty} \mathbb{E}[F(X_N) \wedge C] \leqslant \liminf_{N\to\infty} \mathbb{E}[F(X_N)].$$

Since the right hand side does not depend on C we can take the supremum on the left hand side and obtain the claim. □

Lemma 6.14. *Let $(X_N)_{N\in\mathbb{N}}$ be a family of random variables with values in $\mathbb{R}$ that converges in distribution to a random variable X, and assume that (X_N) is uniformly integrable, i.e.*

$$\lim_{C\to\infty} \sup_{N\in\mathbb{N}} \mathbb{E}[|X_N| \mathbb{I}_{|X_N|>C}] = 0.$$

Then $X \in L^1$ and

$$\mathbb{E}[X] = \lim_{N\to\infty} \mathbb{E}[X_N].$$

Proof. By Lemma 6.13 we know that $\mathbb{E}[|X|] \leqslant \liminf_N \mathbb{E}[|X_N|]$, which is finite because (X_N) is uniformly integrable. Therefore, $X \in L^1$, and with $\varphi_C(x) := (x \wedge C) \vee (-C)$ for $C > 0$ we have

$$\begin{aligned} |\mathbb{E}[X_N] - \mathbb{E}[X]| &\leqslant |\mathbb{E}[X_N] - \mathbb{E}[\varphi_C(X_N)]| + |\mathbb{E}[\varphi_C(X_N)] - \mathbb{E}[\varphi_C(X)]| + |\mathbb{E}[\varphi_C(X)] - \mathbb{E}[X]| \\ &\lesssim \mathbb{E}[|X_N| \mathbb{I}_{|X_N|>C}] + |\mathbb{E}[\varphi_C(X_N)] - \mathbb{E}[\varphi_C(X)]| + \mathbb{E}[|X| \mathbb{I}_{|X|>C}]. \end{aligned}$$

Let now $\varepsilon > 0$, then by assumption there exists $C > 0$ such that

$$\sup_{N\in\mathbb{N}} \mathbb{E}[|X_N| \mathbb{I}_{|X_N|>C}] + \mathbb{E}[|X| \mathbb{I}_{|X|>C}] \leqslant \varepsilon,$$

and using that φ_C is a continuous and bounded function we obtain

$$\lim_{N\to\infty} |\mathbb{E}[X_N] - \mathbb{E}[X]| \lesssim \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, the claim follows. □

To speak about martingales, we first have to introduce a filtration: Let

$$\mathcal{F}_t := \sigma(\mathcal{Y}_s(\varphi) : s \in [0, t], \varphi \in C^\infty(\mathbb{T}^d))$$

Lemma 6.15. *For $\varphi \in C^\infty(\mathbb{T}^d)$ the process $\mathcal{M}(\varphi)$ is a continuous martingale in the filtration $(\mathcal{F}_t)_{t\in[0,T]}$, with quadratic variation*

$$\langle \mathcal{M}(\varphi) \rangle_t = t\rho(1-\rho) \int_{\mathbb{T}^d} \nabla\varphi(u) \cdot (C\nabla\varphi(u)) \mathrm{d}u.$$

Proof. To show that $\mathcal{M}(\varphi)$ is a martingale, we first need to show that it is adapted, which is true because

$$\mathcal{M}_t(\varphi) = \mathcal{Y}_t(\varphi) - \mathcal{Y}_0(\varphi) - \int_0^t \frac{1}{2} \mathcal{Y}_s(\Delta_C \varphi) \mathrm{d}s,$$

and that $\mathcal{M}(\varphi)$ is integrable, which is true because by Lemma 6.13

$$\begin{aligned} \mathbb{E}[|\mathcal{M}_t(\varphi)|^4] &\leqslant \liminf_{N\to\infty} \mathbb{E}[|\mathcal{M}_t^N(\varphi)|^4] \leqslant \liminf_{N\to\infty} \mathbb{E}[\|\mathcal{M}_t^N\|_{H^\alpha}^4] \|\varphi\|_{H^{-\alpha}}^4 \\ &\leqslant \liminf_{N\to\infty} \mathbb{E}[w_{H^\alpha}(\mathcal{M}^N, T)^4] \|\varphi\|_{H^{-\alpha}}^4 < \infty, \end{aligned}$$

where the last step follows from Lemma 6.8. To show the martingale property, we have to show for all $0 \leqslant s \leqslant t \leqslant T$ and for all bounded and $\mathcal{F}_s$-measurable random variables Φ

$$\mathbb{E}[(\mathcal{M}_t(\varphi) - \mathcal{M}_s(\varphi))\Phi] = 0.$$

But by the monotone class theorem, see Lemma 6.16 below, it suffices to consider Φ of the form

$$\Phi(\omega) = F(\mathcal{Y}_{s_1}(\psi_1)(\omega), \ldots, \mathcal{Y}_{s_m}(\psi_m)(\omega))$$

for $0 \leqslant s_1 \leqslant \ldots \leqslant s_m \leqslant s$ and $\psi_1, \ldots, \psi_m \in C^\infty(\mathbb{T}^d)$ and $F: \mathbb{R}^m \to \mathbb{R}$ continuous and bounded. Then the continuous mapping theorem shows that

$$(\mathcal{M}_t(\varphi) - \mathcal{M}_s(\varphi)) F(\mathcal{Y}_{s_1}(\psi_1), \ldots, \mathcal{Y}_{s_m}(\psi_m)) = \lim_{N \to \infty} (\mathcal{M}_t^N(\varphi) - \mathcal{M}_s^N(\varphi)) F(\mathcal{Y}_{s_1}^N(\psi_1), \ldots, \mathcal{Y}_{s_m}^N(\psi_m)),$$

where the convergence is in distribution. We just saw that $\mathcal{M}_t^N(\varphi) - \mathcal{M}_s^N(\varphi)$ is bounded in L^4, and in particular uniformly integrable. Since F is bounded, we can thus apply Lemma 6.14 to obtain

$$\begin{aligned}&\mathbb{E}[(\mathcal{M}_t(\varphi) - \mathcal{M}_s(\varphi)) F(\mathcal{Y}_{s_1}(\psi_1), \ldots, \mathcal{Y}_{s_m}(\psi_m))] \\ &= \lim_{N \to \infty} \mathbb{E}_{\mu_\rho^N}[(\mathcal{M}_t^N(\varphi) - \mathcal{M}_s^N(\varphi)) F(\mathcal{Y}_{s_1}^N(\psi_1), \ldots, \mathcal{Y}_{s_m}^N(\psi_m))] = 0,\end{aligned}$$

which proves that $\mathcal{M}(\varphi)$ is a martingale – and since $|\mathcal{M}_t(\varphi) - \mathcal{M}_s(\varphi)| \leqslant \|\mathcal{M}_t - \mathcal{M}_s\|_{H^\alpha} \|\varphi\|_{H^{-\alpha}}$ and we already know that $\mathcal{M}$ is almost surely continuous in H^α, also $\mathcal{M}(\varphi)$ is almost surely continuous.

To compute its quadratic variation it suffices again to show that

$$\mathbb{E}\left[\left(\mathcal{M}_t^2(\varphi) - \mathcal{M}_s^2(\varphi) - (t-s)\rho(1-\rho) \int_{\mathbb{T}^d} \nabla\varphi(u) \cdot (C \nabla\varphi(u)) \mathrm{d}u\right) F(\mathcal{Y}_{s_1}(\psi_1), \ldots, \mathcal{Y}_{s_m}(\psi_m))\right] = 0.$$

Since in (the proof of) Lemma 6.8 we derived moment bounds on $|\mathcal{M}_t^N(\varphi)|^4$ and $\langle \mathcal{M}^N(\varphi) \rangle_t^2$ that allow us to apply Lemma 6.14 to obtain the convergence of moments, it only remains to show that

$$\lim_{N \to \infty} \langle \mathcal{M}^N(\varphi) \rangle_t = t\rho(1-\rho) \int_{\mathbb{T}^d} \nabla\varphi(u) \cdot (C \nabla\varphi(u)) \mathrm{d}u.$$

Recall that by Lemma 5.4 for all large N (p has finite range and therefore $p^N = p$ for all large N)

$$\begin{aligned}\langle \mathcal{M}^N(\varphi) \rangle_t &= \int_0^t \frac{1}{N^d} \sum_{x,y} \eta_{N^2 r}^N(x)(1 - \eta_{N^2 r}^N(x+y)) p^N(y) [\nabla_{x,y}^N \varphi]^2 \mathrm{d}r \\ &= \sum_y p(y) \int_0^t \frac{1}{N^d} \sum_x \eta_{N^2 r}^N(x)(1 - \eta_{N^2 r}^N(x+y)) [\nabla_{x,y}^N \varphi]^2 \mathrm{d}r.\end{aligned}$$

The right hand side is a finite sum in y, so we can study each addend separately. Note also that the expression for $y = 0$ vanishes (e.g. because $p(y) = 0$ or because $\nabla_{x,0}^N \varphi = 0$), so fix $y \neq 0$ and note that

$$\begin{aligned}&\mathbb{E}_{\mu_\rho^N}\left[\left|\int_0^t \frac{1}{N^d} \sum_x \{\eta_{N^2 r}^N(x)(1 - \eta_{N^2 r}^N(x+y)) - \rho(1-\rho)\} [\nabla_{x,y}^N \varphi]^2 \mathrm{d}r\right|^2\right] \\ &\leqslant t \int_0^t \frac{1}{N^{2d}} \mathbb{E}_{\mu_\rho^N}\left[\left|\sum_x \{\eta_{N^2 r}^N(x)(1 - \eta_{N^2 r}^N(x+y)) - \rho(1-\rho)\}\right|^2\right] [\nabla_{x,y}^N \varphi]^4 \mathrm{d}r \\ &\leqslant t^2 N^{-2d} \sum_{x,x'} \mathbb{E}_{\mu_\rho^N}[\{\eta_0^N(x)(1 - \eta_0^N(x+y)) - \rho(1-\rho)\}\{\eta_0^N(x')(1 - \eta_0^N(x'+y)) - \rho(1-\rho)\}] \\ &\qquad \times [\nabla_{x,y}^N \varphi]^2 [\nabla_{x',y}^N \varphi]^2\end{aligned}$$

where we used that η^N is stationary under μ_ρ^N and therefore we can replace the expetaction at time $N^2 r$ by that at time 0. If $x' \notin \{x, x+y, x-y\}$, then all the random variables in the expectation are independent and the expectation factorizes and thus vanishes. Therefore, we get

$$\mathbb{E}_{\mu_\rho^N}\left[\left|\int_0^t \frac{1}{N^d} \sum_x \{\eta_{N^2 r}^N(x)(1 - \eta_{N^2 r}^N(x+y)) - \rho(1-\rho)\} [\nabla_{x,y}^N \varphi]^2 \mathrm{d}r\right|^2\right] \lesssim N^{-2d} N^d \lesssim N^{-d},$$

and thus

$$\mathbb{E}_{\mu_\rho^N}\left[\left|\langle \mathcal{M}^N(\varphi) \rangle_t - t \frac{1}{N^d} \sum_{x,y} p(y)\rho(1-\rho) [\nabla_{x,y}^N \varphi]^2\right|^2\right] = 0.$$

Moreover, a simple computation shows that

$$\left|\frac{1}{N^d} \sum_{x,y} p(y)\rho(1-\rho) \left([\nabla_{x,y}^N \varphi]^2 - \left|\nabla\varphi\left(\frac{x}{N}\right) \cdot y\right|^2\right)\right| \lesssim N^{-1} \|\mathrm{D}^2 \varphi\|_\infty \|\nabla\varphi\|_\infty,$$

and therefore

$$\begin{aligned}\lim_{N\to\infty}\langle\mathcal{M}^N(\varphi)\rangle_t &= \lim_{N\to\infty} t\frac{1}{N^d}\sum_{x,y} p(y)\rho(1-\rho)\Big|\nabla\varphi\Big(\frac{x}{N}\Big)\cdot y\Big|^2\\ &= \lim_{N\to\infty} t\frac{1}{N^d}\sum_{x,y} p(y)\rho(1-\rho)\sum_{i,j=1}^d \partial_{u_i}\varphi\Big(\frac{x}{N}\Big)y_i\partial_{u_j}\varphi\Big(\frac{x}{N}\Big)y_j\\ &= \lim_{N\to\infty} t\rho(1-\rho)\sum_{i,j=1}^d C_{ij}\frac{1}{N^d}\sum_x \partial_{u_i}\varphi\Big(\frac{x}{N}\Big)\partial_{u_j}\varphi\Big(\frac{x}{N}\Big)\\ &= t\rho(1-\rho)\sum_{i,j=1}^d C_{ij}\int_{\mathbb{T}^d}\partial_{u_i}\varphi(u)\partial_{u_j}\varphi(u)\mathrm{d}u\\ &= t\rho(1-\rho)\int_{\mathbb{T}^d}\nabla\varphi(u)\cdot(C\nabla\varphi(u))\mathrm{d}u,\end{aligned}$$

which concludes the proof. □

In the proof we used the monotone class theorem:

Lemma 6.16. *Let H be a linear space of bounded functions on Ω such that*

- *H contains all constant functions;*
- *if $(h_n)_{n\in\mathbb{N}}\subset H$ and there exists $h:\Omega\to\mathbb{R}$ with $\sup_{\omega\in\Omega}|h_n(\omega)-h(\omega)|\to 0$, then $h\in H$;*
- *if $(h_n)_{n\in\mathbb{N}}\subset H$ is such that $-C\leqslant h_1\leqslant h_2\leqslant\ldots\leqslant C$ for some $C\in\mathbb{R}$ and such that there exists $h:\Omega\to\mathbb{R}$ with $h_n(\omega)\to h(\omega)$ for all $\omega\in\Omega$, then $h\in H$.*

Let $H_0\subset H$ be closed under multiplication (that is $fg\in H_0$ whenever $f,g\in H_0$). Then H contains all $\sigma(H_0)$-measurable functions.

Proof. See [Ethier-Kurtz, Appendix, Corollary 4.4]. □

Note that by Lévy's characterization $\mathcal{M}(\varphi)$ is a multiple of an $(\mathcal{F}_t)$-Brownian motion because its quadratic variation is a deterministic multiple of t. In particular, $\mathcal{M}(\varphi)$ is independent of $\mathcal{Y}_0$ for all $\varphi\in C^\infty(\mathbb{T}^d)$, and it is a Gaussian process.

Remark 6.17. As discussed above, a space-time white noise on $[0,T]\times\mathbb{T}^d$ is a centered Gaussian process $(\xi(\psi))_{\psi\in L^2([0,T]\times\mathbb{T}^d)}$ with covariance $\mathbb{E}[\xi(\psi)\xi(\tilde{\psi})]=\int_{[0,T]\times\mathbb{T}^d}\psi(t,u)\tilde{\psi}(t,u)\mathrm{d}t\mathrm{d}u$. Consider now d independent space-time white noises ξ^1, ..., ξ^d, and let $C^{1/2}$ be a (symmetric) square root of the matrix C. Then for all $\varphi\in C^\infty(\mathbb{T}^d)$ the process

$$\operatorname{div}(C^{1/2}\xi)(\mathbb{I}_{[0,t]}\varphi):=\sum_{j=1}^d\partial_{u_j}(C^{1/2}\xi)_j(\mathbb{I}_{[0,t]}\varphi):=-\sum_{j=1}^d(C^{1/2}\xi)_j(\mathbb{I}_{[0,t]}\partial_{u_j}\varphi),\qquad t\in[0,T],$$

is centered Gaussian with covariance

$$\begin{aligned}&\mathbb{E}[(\operatorname{div}(C^{1/2}\xi)(\mathbb{I}_{[0,t]}\varphi))(\operatorname{div}(C^{1/2}\xi)(\mathbb{I}_{[0,s]}\psi))]\\ &= \sum_{j_1,j_2=1}^d\sum_{k_1,k_2=1}^d(-1)^2\mathbb{E}\big[C^{1/2}_{j_1k_1}\xi^{k_1}(\mathbb{I}_{[0,t]}\partial_{u_{j_1}}\varphi)C^{1/2}_{j_2k_2}\xi^{k_2}(\mathbb{I}_{[0,s]}\partial_{u_{j_2}}\psi)\big]\\ &= \sum_{j_1,j_2=1}^d\sum_{k_1,k_2=1}^d C^{1/2}_{j_1k_1}C^{1/2}_{j_2k_2}\delta_{k_1,k_2}\int_{[0,T]\times\mathbb{T}^d}\mathbb{I}_{[0,t]}(r)\partial_{u_{j_1}}\varphi(u)\mathbb{I}_{[0,s]}(r)\partial_{u_{j_2}}\psi(u)\mathrm{d}r\mathrm{d}u\\ &\overset{C^{1/2}\text{ symm}}{=\!=\!=} (s\wedge t)\sum_{j_1,j_2=1}^d\sum_{k=1}^d C^{1/2}_{j_1k}C^{1/2}_{kj_2}\int_{\mathbb{T}^d}\partial_{u_{j_1}}\varphi(u)\partial_{u_{j_2}}\psi(u)\mathrm{d}u\\ &= (s\wedge t)\int_{\mathbb{T}^d}\sum_{j_1,j_2=1}^d C_{j_1j_2}\partial_{u_{j_1}}\varphi(u)\partial_{u_{j_2}}\psi(u)\mathrm{d}u\\ &= (s\wedge t)\int_{\mathbb{T}^d}\nabla\varphi(u)\cdot(C\psi)(u)\mathrm{d}u,\end{aligned}$$

which is (up to the factor $\rho(1-\rho)$) exactly the covariance of the Gaussian process $(t,\varphi)\mapsto\mathcal{M}_t(\varphi)$.

Definition 6.18. *Let η be a white noise on $\mathbb{T}^d$ (also called a space white noise) and let $\xi^1,\ldots,\xi^d$ be d independent space-time white noises. We say that a process $(\mathcal{Z}_t)_{t\in[0,T]}$ with values in $C([0,T],H^\alpha)$ for $\alpha<-d/2-2$ is a martingale solution to the stochastic PDE (SPDE)*

$$\partial_t\mathcal{Z}_t=\frac{1}{2}\Delta_C\mathcal{Z}_t+\sqrt{\rho(1-\rho)}\operatorname{div}(C^{1/2}\xi),\qquad \mathcal{Z}_0=\sqrt{\rho(1-\rho)}\,\eta, \tag{6.5}$$

if $(\rho(1-\rho))^{-1/2}\mathcal{Z}_0$ is a space white noise and for all $\varphi\in C^\infty(\mathbb{T}^d)$ the process

$$\mathcal{N}_t(\varphi):=\mathcal{Z}_t(\varphi)-\mathcal{Z}_0(\varphi)-\int_0^t\frac{1}{2}\mathcal{Z}_s(\Delta_C\varphi)\mathrm{d}s$$

is a continuous martingale in the filtration $\mathcal{F}_t=\sigma(\mathcal{Z}_s(\varphi)\colon s\in[0,t],\varphi\in C^\infty(\mathbb{T}^d))$, $t\in[0,T]$, with quadratic variation

$$\langle\mathcal{N}(\varphi)\rangle_t=t\rho(1-\rho)\int_{\mathbb{T}^d}\nabla\varphi(u)\cdot(C\psi)(u)\mathrm{d}u.$$

Let us summarize our observations:

Lemma 6.19. *Let $\alpha<-d/2-2$ and let $(\mathcal{Y}_t)_{t\in[0,T]}\in C([0,T],H^\alpha)$ be a limit point of $(\mathcal{Y}^N)_{N\in\mathbb{N}}$. Then $\mathcal{Y}$ is a martingale solution to (6.5).*

6.3 Step 3: Uniqueness of the limit

To conclude the convergence proof, we have to show that the martingale solution to (6.5) is unique in distribution. For that purpose we prove uniqueness of martingale solutions to a much wider class of linear SPDEs, so that the uniqueness for (6.5) becomes a special case.

Definition 6.20. *Let $P(x)=\sum_{|\gamma|\leqslant n}a_\gamma x^\gamma$ be a polynomial, where $n\in\mathbb{N}$ and $\gamma\in\mathbb{N}_0^d$ are multi-indices and $a_\gamma\in\mathbb{R}$ for all γ. Then we write*

$$P(\partial):=\sum_{|\gamma|\leqslant n}a_\gamma\partial_u^\gamma,\qquad P^*(\partial):=\sum_{|\gamma|\leqslant n}(-1)^{|\gamma|}a_\gamma\partial_u^\gamma.$$

Note that

$$\int_{\mathbb{T}^d}(P(\partial)\varphi(u))\psi(u)\mathrm{d}u=\int_{\mathbb{T}^d}\varphi(u)P^*(\partial)\psi(u)\mathrm{d}u$$

for all $\varphi,\psi\in C^\infty(\mathbb{T}^d)$.

Definition 6.21. *Let $(\eta(\varphi))_{\varphi\in C^\infty(\mathbb{T}^d)}$ be a Gaussian process and let $\xi^1,\ldots,\xi^n$ be independent space-time white noises on $[0,T]\times\mathbb{T}^d$. Let P and $Q_1,\ldots,Q_n$ be polynomials and let $\sigma\in\mathbb{R}^{n\times n}$. A stochastic process $(\mathcal{Z}_t)_{t\in[0,T]}$ with values in $C([0,T],H^\alpha)$ for $\alpha\in\mathbb{R}$ is a martingale solution to the SPDE*

$$\partial_t\mathcal{Z}_t=P(\partial)\mathcal{Z}_t+\sum_{j=1}^n Q_j(\partial)(\sigma\xi)^j,\qquad\mathcal{Z}_0=\eta, \tag{6.6}$$

if $\mathrm{law}(\mathcal{Z}_0)=\mathrm{law}(\eta)$ and for all $\varphi\in C^\infty(\mathbb{T}^d)$ the process

$$\mathcal{N}_t(\varphi):=\mathcal{Z}_t(\varphi)-\mathcal{Z}_0(\varphi)-\int_0^t\mathcal{Z}_s(P^*(\partial)\varphi)\mathrm{d}s$$

is a continuous martingale in the filtration $\mathcal{F}_t=\sigma(\mathcal{Z}_s(\varphi)\colon s\in[0,t],\varphi\in C^\infty(\mathbb{T}^d))$, $t\in[0,T]$, with quadratic variation

$$\langle\mathcal{N}(\varphi)\rangle_t=t\sum_{j_1,j_2=1}^n\int_{\mathbb{T}^d}Q_{j_1}^*(\partial)\varphi(u)(\sigma\sigma^*)_{j_1j_2}Q_{j_2}^*(\partial)\varphi(u)\mathrm{d}u=:t\mathcal{B}(\varphi,\varphi).$$

To shorten notation we will also write $\mathcal{B}(\varphi):=\mathcal{B}(\varphi,\varphi)$.

Remark 6.22. In (6.5) we have $n=d$ and $\sigma=C^{1/2}$ and $Q_j=\partial_j$ for all j, and $P(x)=\frac{1}{2}x\cdot(Cx)$.

Our convergence proof is complete once we show that the distribution of the martingale solution to (6.5) is unique. This is a special case of the following theorem.

Theorem 6.23. *In the setting of Definition 6.21 assume that there exists C with*

$$\mathrm{Re}(P(ix))\leqslant 0, \qquad \forall|x|\geqslant C. \tag{6.7}$$

Then any two martingale solutions $\mathcal{Z}$ and $\tilde{\mathcal{Z}}$ to (6.6) have the same distribution, and they are Gaussian processes with explicitly computable (but slightly complicated) expectation and covariance.

The proof follows Appendix C of the recent paper [Mourrat, Weber. Convergence of the two-dimensional Kac-Ising model to Φ^4_2. Comm. Pure Appl. Math. 70, 2017] and it is based on some auxiliary results.

Lemma 6.24. *Let $\mathcal{Z}$ be a martingale solution to (6.6) and let $\varphi\in C^\infty([0,T]\times\mathbb{T}^d)$. Then*

$$\mathcal{R}_t(\varphi):=\mathcal{Z}_t(\varphi(t,\cdot))-\mathcal{Z}_0(\varphi(0,\cdot))-\int_0^t\mathcal{Z}_s((\partial_s+P^*(\partial))\varphi(s,\cdot))\mathrm{d}s, \qquad t\in[0,T],$$

is a continuous martingale with quadratic variation

$$\langle\mathcal{R}(\varphi)\rangle_t=\int_0^t\mathcal{B}(\varphi(s,\cdot))\mathrm{d}s.$$

In particular, $(\mathcal{R}_t(\varphi))_{t\in[0,T],\varphi\in C^\infty([0,T]\times\mathbb{T}^d)}$ is a centered Gaussian process and independent of $\mathcal{F}_0$.

Proof. We decompose

$$\begin{aligned}
\mathcal{Z}_t(\varphi(t,\cdot))-\mathcal{Z}_0(\varphi(0,\cdot)) &= [\mathcal{Z}_t(\varphi(t,\cdot))-\mathcal{Z}_0(\varphi(t,\cdot))]+[\mathcal{Z}_0(\varphi(t,\cdot))-\mathcal{Z}_0(\varphi(0,\cdot))]\\
&=\left[\mathcal{N}_t(\varphi(t,\cdot))+\int_0^t\mathcal{Z}_s(P^*(\partial)\varphi(t,\cdot))\mathrm{d}s\right]+\left[\int_0^t\mathcal{Z}_0(\partial_s\varphi(s,\cdot))\mathrm{d}s\right]\\
&=\left[\mathcal{N}_t(\varphi(t,\cdot))+\int_0^t\mathcal{Z}_s(P^*(\partial)\varphi(s,\cdot))\mathrm{d}s+\int_0^t(\mathcal{Z}_s(P^*(\partial)\varphi(t,\cdot))-\mathcal{Z}_s(P^*(\partial)\varphi(s,\cdot)))\mathrm{d}s\right]\\
&\quad+\left[\int_0^t\mathcal{Z}_s(\partial_s\varphi(s,\cdot))\mathrm{d}s-\int_0^t(\mathcal{Z}_s-\mathcal{Z}_0)(\partial_s\varphi(s,\cdot))\mathrm{d}s\right]\\
&=\left[\mathcal{N}_t(\varphi(t,\cdot))+\int_0^t\mathcal{Z}_s(P^*(\partial)\varphi(s,\cdot))\mathrm{d}s+\int_0^t\mathrm{d}s\int_s^t\mathrm{d}r\mathcal{Z}_s(\partial_rP^*(\partial)\varphi(r,\cdot))\right]\\
&\quad+\left[\int_0^t\mathcal{Z}_s(\partial_s\varphi(s,\cdot))\mathrm{d}s-\int_0^t\mathrm{d}s\left(\int_0^s\mathrm{d}r\mathcal{Z}_r(P^*(\partial)\partial_s\varphi(s,\cdot))+\mathcal{N}_s(\partial_s\varphi(s,\cdot))\right)\right].
\end{aligned}$$

Now compare the two double integrals on the right hand side:

$$\begin{aligned}
&\int_0^t\mathrm{d}s\int_s^t\mathrm{d}r\mathcal{Z}_s(\partial_rP^*(\partial)\varphi(r,\cdot))-\int_0^t\mathrm{d}s\int_0^s\mathrm{d}r\mathcal{Z}_r(P^*(\partial)\partial_s\varphi(s,\cdot))\\
&=\int_0^t\mathrm{d}r\int_0^r\mathrm{d}s\mathcal{Z}_s(\partial_rP^*(\partial)\varphi(r,\cdot))-\int_0^t\mathrm{d}s\int_0^s\mathrm{d}r\mathcal{Z}_r(P^*(\partial)\partial_s\varphi(s,\cdot))=0,
\end{aligned}$$

and therefore

$$\mathcal{Z}_t(\varphi(t,\cdot))-\mathcal{Z}_0(\varphi(0,\cdot))-\int_0^t\mathcal{Z}_s((\partial_s+P^*(\partial))\varphi(s,\cdot))\mathrm{d}s=\mathcal{N}_t(\varphi(t,\cdot))-\int_0^t\mathrm{d}s\mathcal{N}_s(\partial_s\varphi(s,\cdot)).$$

The right hand side is a martingale: it is easily seen to be integrable and adapted, and for $r\leqslant t$ we have

$$\begin{aligned}
\mathbb{E}\left[\mathcal{N}_t(\varphi(t,\cdot))-\int_0^t\mathrm{d}s\mathcal{N}_s(\partial_s\varphi(s,\cdot))\middle|\mathcal{F}_r\right]&=\mathcal{N}_r(\varphi(t,\cdot))-\int_0^r\mathcal{N}_s(\partial_s\varphi(s,\cdot))\mathrm{d}s-\int_r^t\mathcal{N}_r(\partial_s\varphi(s,\cdot))\mathrm{d}s\\
&=\mathcal{N}_r(\varphi(t,\cdot))-\int_0^r\mathcal{N}_s(\partial_s\varphi(s,\cdot))\mathrm{d}s-\mathcal{N}_r(\varphi(t,\cdot)-\varphi(r,\cdot))=\mathcal{N}_r(\varphi(r,\cdot))-\int_0^r\mathcal{N}_s(\partial_s\varphi(s,\cdot))\mathrm{d}s.
\end{aligned}$$

To compute the quadratic variation, note that $t\mapsto\int_0^t\mathcal{N}_s(\partial_s\varphi(s,\cdot))\mathrm{d}s$ is of finite variation and therefore it does not contribute to the quadratic variation. For $t_k^n=tk/n$ we have

$$\begin{aligned}&\sum_{k=0}^{n-1}|\mathcal{N}_{t_{k+1}^n}(\varphi(t_{k+1}^n,\cdot))-\mathcal{N}_{t_k^n}(\varphi(t_k^n,\cdot))|^2\\=&\sum_{k=0}^{n-1}|[\mathcal{N}_{t_{k+1}^n}(\varphi(t_k^n,\cdot))-\mathcal{N}_{t_k^n}(\varphi(t_k^n,\cdot))]+[\mathcal{N}_{t_k^n}(\varphi(t_{k+1}^n,\cdot)-\varphi(t_k^n,\cdot))]|^2\\=&\sum_{k=0}^{n-1}|\mathcal{N}_{t_{k+1}^n}(\varphi(t_k^n,\cdot))-\mathcal{N}_{t_k^n}(\varphi(t_k^n,\cdot))|^2\\&+\sum_{k=0}^{n-1}\mathcal{N}_{t_k^n}(\varphi(t_{k+1}^n,\cdot)-\varphi(t_k^n,\cdot))[2(\mathcal{N}_{t_{k+1}^n}(\varphi(t_k^n,\cdot))-\mathcal{N}_{t_k^n}(\varphi(t_k^n,\cdot)))+\mathcal{N}_{t_k^n}(\varphi(t_{k+1}^n,\cdot)-\varphi(t_k^n,\cdot))].\end{aligned}$$

The first term on the right hand side is a sum of squares of independent Gaussian random variables (recall that $\mathcal{N}(\psi)$ is a multiple of a Brownian motion for every $\psi\in C^\infty(\mathbb{T}^d)$), and the k-th addend has expectation

$$\mathbb{E}[|\mathcal{N}_{t_{k+1}^n}(\varphi(t_k^n,\cdot))-\mathcal{N}_{t_k^n}(\varphi(t_k^n,\cdot))|^2]=\frac{t}{n}\mathcal{B}(\varphi(t_k^n,\cdot)).$$

Since $s\mapsto\mathcal{B}(\varphi(s,\cdot))$ is continuous, it follows from the law of large numbers that

$$\lim_{n\to\infty}\sum_{k=0}^{n-1}|\mathcal{N}_{t_{k+1}^n}(\varphi(t_k^n,\cdot))-\mathcal{N}_{t_k^n}(\varphi(t_k^n,\cdot))|^2=\lim_{n\to\infty}\sum_{k=0}^{n-1}\frac{t}{n}\mathcal{B}(\varphi(t_k^n,\cdot))=\int_0^t\mathcal{B}(\varphi(s,\cdot))\mathrm{d}s.$$

The remainder terms can be estimated by

$$\begin{aligned}&\left|\sum_{k=0}^{n-1}\mathcal{N}_{t_k^n}(\varphi(t_{k+1}^n,\cdot)-\varphi(t_k^n,\cdot))[2(\mathcal{N}_{t_{k+1}^n}(\varphi(t_k^n,\cdot))-\mathcal{N}_{t_k^n}(\varphi(t_k^n,\cdot)))+\mathcal{N}_{t_k^n}(\varphi(t_{k+1}^n,\cdot)-\varphi(t_k^n,\cdot))]\right|^2\\\leqslant&\sum_{k=0}^{n-1}|\mathcal{N}_{t_k^n}(\varphi(t_{k+1}^n,\cdot)-\varphi(t_k^n,\cdot))|^2\sum_{k=0}^{n-1}|2(\mathcal{N}_{t_{k+1}^n}(\varphi(t_k^n,\cdot))-\mathcal{N}_{t_k^n}(\varphi(t_k^n,\cdot)))+\mathcal{N}_{t_k^n}(\varphi(t_{k+1}^n,\cdot)-\varphi(t_k^n,\cdot))|^2,\end{aligned}$$

and the first factor has expectation

$$\mathbb{E}\left[\sum_{k=0}^{n-1}|\mathcal{N}_{t_k^n}(\varphi(t_{k+1}^n,\cdot)-\varphi(t_k^n,\cdot))|^2\right]=\sum_{k=0}^{n-1}t_k^n\mathcal{B}(\varphi(t_{k+1}^n,\cdot)-\varphi(t_k^n,\cdot))\lesssim\sum_{k=0}^{n-1}t|t_{k+1}^n-t_k^n|^2\lesssim n^{-1},$$

where we used that $\mathcal{B}$ is a continuous bilinear map from $C^\infty([0,T]\times\mathbb{T}^d)$ to $\mathbb{R}$. Since the second factor of the remainder is easily seen to be bounded in n, it follows that

$$\langle\mathcal{R}(\varphi)\rangle_t=\int_0^t\mathcal{B}(\varphi(s,\cdot))\mathrm{d}s.$$

To see that $\mathcal{R}$ is a Gaussian process, let $\kappa_1,\kappa_2\in\mathbb{R}$ and $\varphi_1,\varphi_2\in C^\infty([0,T]\times\mathbb{T}^d)$ and $0\leqslant t_1\leqslant t_2\leqslant T$. Then

$$\begin{aligned}\kappa_1\mathcal{R}_{t_1}(\varphi_1)+\kappa_2\mathcal{R}_{t_2}(\varphi_2)&=\kappa_1\mathcal{R}_{t_1}(\varphi_1)+\kappa_2\mathcal{R}_{t_1}(\varphi_2)+\kappa_2(\mathcal{R}_{t_2}(\varphi_2)-\mathcal{R}_{t_1}(\varphi_2))\\&=\mathcal{R}_{t_1}(\kappa_1\varphi_1+\kappa_2\varphi_2)+(\mathcal{R}_{t_2}(\kappa_2\varphi_2)-\mathcal{R}_{t_1}(\kappa_2\varphi_2)).\end{aligned}$$

For fixed φ the process $\mathcal{R}(\varphi)$ is a continuous martingale with deterministic quadratic variation, and therefore it is Gaussian and its increments after time t are independent of $\mathcal{F}_t$. Thus we have a sum of independent Gaussians, which is Gaussian. The same argument extends to finite linear combinations $\kappa_1\mathcal{R}_{t_1}(\varphi_1)+\cdots+\kappa_m\mathcal{R}_{t_m}(\varphi_m)$, and therefore $\mathcal{R}$ is Gaussian. $\mathcal{R}$ is independent of $\mathcal{F}_0$ since $\mathcal{R}_0=0$ and $\mathcal{R}$ has increments that are independent of $\mathcal{F}_0$. □

Proposition 6.25. *Let P be a polynomial satisfying (6.7) and let $\varphi\in C^\infty(\mathbb{T}^d)$. Then there exist unique solutions $\psi,\bar{\psi}\in C^\infty([0,T]\times\mathbb{T}^d)$ to the equations*

$$\partial_t\psi=P(\partial)\psi,\qquad\psi(0)=\varphi\tag{6.8}$$

and

$$\partial_t \bar{\psi} = -P(\partial)\bar{\psi}, \qquad \bar{\psi}(T) = \varphi. \tag{6.9}$$

Moreover, $\bar{\psi}(t) = \psi(T-t)$ *for all* $t \in [0,T]$.

Proof. We will argue similarly as in Section 5.5. Let us assume that we showed existence and uniqueness of the solution ψ to (6.8). Then $\bar{\psi}(t) := \psi(T-t)$ solves (6.9), and if $\tilde{\psi}$ is another solution to (6.9), then $\tilde{\psi}(T-\cdot)$ solves (6.8) and thus $\tilde{\psi}(T-t) = \psi(t) = \bar{\psi}(T-t)$ for all $t \in [0,T]$.

So it suffices to consider the "forward-in-time" case (6.8). If ψ is a solution, then we get from integration by parts for all $m \in \mathbb{Z}^d$

$$\begin{aligned}\partial_t\left(\int_{\mathbb{T}^d}\psi(t,u)e_m(u)\mathrm{d}u\right) &= \int_{\mathbb{T}^d}(P(\partial)\psi(t,u))e_m(u)\mathrm{d}u = \int_{\mathbb{T}^d}\psi(t,u)P^*(\partial)e_m(u)\mathrm{d}u\\ &= \sum_{|\gamma|\leqslant n}(-1)^{|\gamma|}a_\gamma(-2\pi i m)^\gamma\int_{\mathbb{T}^d}\psi(t,u)e_m(u)\mathrm{d}u\\ &= \sum_{|\gamma|\leqslant n}a_\gamma(i2\pi m)^\gamma\int_{\mathbb{T}^d}\psi(t,u)e_m(u)\mathrm{d}u\\ &= P(i2\pi m)\left(\int_{\mathbb{T}^d}\psi(t,u)e_m(u)\mathrm{d}u\right),\end{aligned}$$

with

$$\int_{\mathbb{T}^d}\psi(0,u)e_m(u)\mathrm{d}u = \hat{\varphi}(m),$$

and therefore

$$\int_{\mathbb{T}^d}\psi(t,u)e_m(u)\mathrm{d}u = e^{tP(i2\pi m)}\hat{\varphi}(m)$$

for all $t \in [0,T]$. In other words, $\int_{\mathbb{T}^d}\psi(t,u)e_m(u)\mathrm{d}u$ is uniquely determined for all $m \in \mathbb{Z}^d$ and $t \in [0,T]$, and as in Section 5.5 we easily get the uniqueness of ψ from that. Next, let us show that

$$\psi(t,u) := \sum_{m\in\mathbb{Z}^d} e^{tP(i2\pi m)}\hat{\varphi}(m)e_{-m}(u) \tag{6.10}$$

defines an element of $C^\infty([0,T]\times\mathbb{T}^d)$ that actually solves the equation. To see that the series converges, first note that

$$\sup_{m\in\mathbb{Z}^d, t\in[0,T]}|e^{tP(i2\pi m)}| = \sup_{m\in\mathbb{Z}^d, t\in[0,T]}|e^{t\mathrm{Re}(P(i2\pi m))}| < \infty$$

because for all $|x| \geqslant C$ we have $\mathrm{Re}(P(ix)) \leqslant 0$. Since also $|\hat{\varphi}(m)| \lesssim (1+|m|)^{-k}$ for all $k \in \mathbb{N}$ the series (6.10) converges absolutely. Moreover, for all $k \in \mathbb{N}_0$ and $\gamma \in \mathbb{N}_0^d$

$$\partial_t^k\partial_u^\gamma\sum_{m\in\mathbb{Z}^d}e^{tP(i2\pi m)}\hat{\varphi}(m)e_{-m}(u) = \sum_{m\in\mathbb{Z}^d}(P(i2\pi m))^k e^{tP(i2\pi m)}\hat{\varphi}(m)(2\pi i m)^\gamma e_{-m}(u),$$

which is also absolutely summable, and therefore $\psi \in C^\infty([0,T]\times\mathbb{T}^d)$. For $t=0$ we obtain

$$\psi(0,u) = \sum_{m\in\mathbb{Z}^d}\hat{\varphi}(m)e_{-m}(u) = \varphi(u)$$

for all $u \in \mathbb{T}^d$, which holds because $(e_m)_{m\in\mathbb{Z}^d}$ is an orthonormal basis and $\overline{e_m} = e_{-m}$. Finally

$$\partial_t\psi(t,u) = \sum_{m\in\mathbb{Z}^d}P(i2\pi m)e^{tP(i2\pi m)}\hat{\varphi}(m)e_{-m}(u) = \sum_{m\in\mathbb{Z}^d}e^{tP(i2\pi m)}\hat{\varphi}(m)P(\partial)e_{-m}(u) = P(\partial)\psi(t,u)$$

for all $(t,u) \in [0,T]\times\mathbb{T}^d$, which concludes the proof. □

Remark 6.26. The condition $\mathrm{Re}(P(ix)) \leqslant 0$ for all large x is necessary. For example it is not possible to solve the heat equation backward in time,

$$\partial_t\psi = \Delta\psi, \qquad \psi(T) = \varphi,$$

because the formal solution

$$\psi(t,u) = \sum_{m\in\mathbb{Z}^d} e^{(T-t)|2\pi m|^2}\hat{\varphi}(m)e_{-m}(u)$$

does not converge in any of the function or distribution spaces one typically considers – unless $\hat{\varphi}(m) = 0$ for all but finitely many m. For example the formal solution is not in H^α for any α, because its coefficients grow faster than polynomially in m.

The (forward) heat equation is smoothing out the solution and converges to the (constant in space) average of its initial condition for large times. So running the heat equation backwards in time would have to make smooth functions rougher and rougher, and there is no easy way of setting this up mathematically correct. (This last point is a poor reason though because there are well posed equations that do have roughening effect).

Remark 6.27. The condition $\mathrm{Re}(P(ix)) \leqslant 0$ for all large x is satisfied if and only if one of the following two conditions holds:

i. P is odd, i.e. $a_\gamma = 0$ for all $\gamma \in \mathbb{N}_0^d$ with even length $|\gamma|$; OR

ii. the highest even homogeneity of P satisfies the same condition, i.e. if $k \in \mathbb{N}_0$ is such that $a_\gamma \neq 0$ for some $\gamma \in \mathbb{N}_0^d$ with $|\gamma| = 2k$ and $a_\gamma = 0$ for all γ with $|\gamma| = 2(k+\ell)$ for $\ell \in \mathbb{N}$, then

$$\sum_{\gamma:|\gamma|=2k} a_\gamma (ix)^\gamma \leqslant 0$$

for all $x \in \mathbb{R}^d$.

In particular, P satisfies (6.7) if and only if P^* satisfies (6.7).

Proof. (Proof of Theorem 6.23)

Let $\mathcal{Z}$ be a martingale solution to (6.6) and let $t \in [0,T]$ and $\varphi \in C^\infty(\mathbb{T}^d)$. Consider the solution $\psi \in C^\infty([0,t] \times \mathbb{T}^d)$ to the equation

$$(\partial_s + P^*(\partial))\psi = 0, \qquad \psi(t,u) = \varphi(u),$$

which exists by Proposition 6.25 and Remark 6.27. Then Lemma 6.24 gives

$$\begin{aligned}\mathcal{Z}_t(\varphi) = \mathcal{Z}_t(\psi(t,\cdot)) &= \mathcal{R}_t(\psi) + \mathcal{Z}_0(\psi(0,\cdot)) + \int_0^t \mathcal{Z}_s((\partial_s + P^*(\partial))\psi(s,\cdot))\mathrm{d}s \\ &= \mathcal{R}_t(\psi) + \mathcal{Z}_0(\psi(0,\cdot)),\end{aligned}$$

and since $\mathcal{R}$ and $\mathcal{Z}_0$ are independent Gaussian processes, also $\mathcal{Z}$ is a Gaussian process. Its expectation is

$$\mathbb{E}[\mathcal{Z}_t(\varphi)] = \mathbb{E}[\mathcal{Z}_0(\psi(0,\cdot))] = \mathbb{E}[\eta(\psi(0,\cdot))]$$

and its covariance is

$$\begin{aligned}\mathrm{cov}(\mathcal{Z}_t(\varphi), \mathcal{Z}_s(\tilde{\varphi})) &= \mathrm{cov}(\mathcal{R}_t(\psi), \mathcal{R}_s(\tilde{\psi})) + \mathrm{cov}(\mathcal{Z}_0(\psi(0,\cdot)), \mathcal{Z}_0(\tilde{\psi}(0,\cdot))) \\ &= \int_0^{s\wedge t} \mathcal{B}(\psi(r,\cdot), \tilde{\psi}(r,\cdot))\mathrm{d}r + \mathrm{cov}(\eta(\psi(0,\cdot)), \eta(\tilde{\psi}(0,\cdot))),\end{aligned}$$

in particular the distribution of $\mathcal{Z}$ is unique. □

Remark 6.28. It is not necessary to assume that $(\eta(\varphi))_{\varphi\in C^\infty(\mathbb{T}^d)}$ is a Gaussian process, the same arguments go through for arbitrary random variables η with values in H^α. In that case the process $\mathcal{Z}$ is of course not Gaussian.

We finally proved the main result of this chapter:

Theorem 6.29. *Let $\rho \in (0,1)$ and let $\eta_0^N \sim \mu_\rho^N$ for all N. Assume that p is symmetric. Then for all $\alpha < -d/2 - 2$ the sequence of processes $(\mathcal{Y}_t^N)_{t\in[0,T]} \in D([0,T], \mathcal{M}_+)$ converges in distribution in $D([0,T], H^\alpha)$ to the unique martingale solution $\mathcal{Y}$ of*

$$\partial_t \mathcal{Y}_t = \frac{1}{2}\Delta_C \mathcal{Y}_t + \sqrt{\rho(1-\rho)}\,\mathrm{div}\,(C^{1/2}\xi), \qquad \mathcal{Y}_0 = \sqrt{\rho(1-\rho)}\,\eta,$$

where η is a white noise on $\mathbb{T}^d$ and $\xi^1, \ldots, \xi^d$ are independent space-time white noises on $[0,T] \times \mathbb{T}^d$.