Università di Pisa

Dipartimento di Matematica

Continuous Time Markov Chains and Interacting Particle Systems

Author
Michel Nassif

Supervisor
Dr. Franco Flandoli

Abstract

Interacting particle systems is a recently developed field in the theory of Markov processes with many applications: particle systems have been used to model phenomena ranging from traffic behaviour to spread of infection and tumour growth. We introduce this field through the study of the simple exclusion process. We will construct the generator of this process and we will give a convergence result of the spatial particle density to the solution of the heat equation.

Contents

1 Markov chains

1.1 Discrete-time Markov chains

We start by reviewing the main results in the theory of discrete-time Markov chains.

Definition and basic properties

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let I be a countable set, called the state-space. We say that $(\lambda_i)_{i\in I}$ is a measure on I if $\forall i \in I,\ 0 \le \lambda_i < \infty$. If in addition the total mass $\sum_{i\in I} \lambda_i = 1$, then we call λ a distribution. We say that a matrix $P = (p_{ij})_{i,j\in I}$ is a stochastic matrix if every row $(p_{ij})_{j\in I}$ is a distribution.

Definition 1.1. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of I-valued random variables. Let λ be a distribution on I and P be a stochastic matrix. We say that (X_n) is a Markov chain with initial distribution λ and transition matrix P (Markov(λ, P) for short) if*

(i) X_0 has distribution λ

(ii) for $n \in \mathbb{N}$, conditional on $X_n = i$, X_{n+1} is independent of $X_0, \ldots, X_n$ and has distribution $(p_{ij})_{j\in I}$. More explicitly, for $n \in \mathbb{N}$ and $i_0, \ldots, i_{n+1} \in I$,

$$\mathbb{P}(X_{n+1} = i_{n+1} | X_0 = i_0, \ldots, X_n = i_n) = p_{i_n i_{n+1}}$$

Proposition 1.2. *A random sequence $(X_n)_{n\in\mathbb{N}}$ is Markov(λ, P) if and only if for all $n \in \mathbb{N}$, $i_0, \ldots, i_n \in I$,*

$$\mathbb{P}(X_0 = i_0, \ldots, X_n = i_n) = \lambda_{i_0} p_{i_0 i_1} \ldots p_{i_{n-1} i_n}$$

Definition 1.3. *Let λ be a measure on I and let P, Q be stochastic matrices. We define a new measure λP on I by*

$$(\lambda P)_j = \sum_{i\in I} \lambda_i p_{ij}$$

We also define a new (stochastic) matrix PQ by

$$(PQ)_{ij} = \sum_{k\in I} p_{ik} q_{kj}$$

We define the n-th power of P by

- $P^0 = Id = (\delta_{ij})_{i,j\in I}$
- $P^{n+1} = P^n P$

We write $p_{ij}^{(n)} = (P^n)_{ij}$ for the (i,j) entry in P^n.

Proposition 1.4. *Let $(X_n)_{n\in\mathbb{N}}$ be Markov(λ, P). Then, for all $n, m \in \mathbb{N}$,*

(i) $\mathbb{P}_i(X_n = j) = p_{ij}^{(n)} = \mathbb{P}(X_{n+m} = j | X_m = i)$

(ii) $\mathbb{P}(X_n = j) = (\lambda P)_j$

Theorem 1.5 (Markov property)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be Markov(λ, P), and let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be the natural filtration of (X_n). Then, conditional on $X_m = i$, $(X_{m+n})_{n\in\mathbb{N}}$ is Markov(δ_i, P) and is independent of $\mathcal{F}_m$. More explicitly, for $A \in \mathcal{F}_m$ and $B \in \mathcal{P}(I)^{\otimes\mathbb{N}}$,*

$$\mathbb{P}\Big((X_{m+n})_{n\in\mathbb{N}} \in B, A\Big|X_m = i\Big) = \mathbb{P}\Big((X_n)_{n\in\mathbb{N}} \in B\Big|X_0 = i\Big) \times \mathbb{P}(A|X_m = i)$$

Definition 1.6. *Let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be a filtration on $(\Omega, \mathcal{F})$. A random variable $T : \Omega \to \mathbb{N} \cup \{\infty\}$ is called a stopping time if $\{T \le n\} \in \mathcal{F}_n$. The σ-algebra $\mathcal{F}_T$ of events determined prior to the stopping time T consists of those events $A \in \mathcal{F}$ for which $A \cap \{T \le n\} \in \mathcal{F}_n$ for every $n \in \mathbb{N}$. If $(X_n)_{n\in\mathbb{N}}$ is a random sequence, we define X_T on $\{T < \infty\}$ by*

$$X_T(\omega) = X_{T(\omega)}(\omega)$$

Theorem 1.7 (Strong Markov property)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be Markov(λ, P) and let T be a stopping time for the natural filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ of $(X_n)_{n\in\mathbb{N}}$. Then, conditional on $T < \infty$ and $X_T = i$, $(X_{T+n})_{n\in\mathbb{N}}$ is Markov(δ_i, P) and is independent of $\mathcal{F}_T$. More explicitly, for $A \in \mathcal{F}_T$ and $B \in \mathcal{P}(I)^{\otimes\mathbb{N}}$,*

$$\mathbb{P}\Big((X_{T+n})_{n\in\mathbb{N}} \in B, A\Big|T < \infty, X_T = i\Big) = \mathbb{P}\Big((X_n)_{n\in\mathbb{N}} \in B\Big|X_0 = i\Big) \times \mathbb{P}(A|T < \infty, X_T = i)$$

Class structure and recurrence

Definition 1.8. *If $i, j \in I$, we say that i leads to j and write $i \to j$ if $\mathbb{P}_i(X_n = j$ for some $n) > 0$. We say that i communicates with j and write $i \leftrightarrow j$ if both $i \to j$ and $j \to i$.*

Proposition 1.9. *For distinct states $i, j \in I$ the following are equivalent:*

(i) $i \to j$

(ii) $p_{i_0 i_1} p_{i_1 i_2} \dots p_{i_{n-1} i_n} > 0$ *for some states* $i_0, i_1, \dots, i_n$ *with* $i_0 = i$ *and* $i_n = j$

(iii) $p_{ij}^{(n)} > 0$ *for some* n

It can be verified that $\leftrightarrow$ is an equivalence relation on I, and thus partitions I into communicating classes. We say that a class C is closed if

$$\big(i \in C,\ i \to j\big) \Rightarrow j \in C$$

If there is only one class, we say that the chain (or the transition matrix P) is irreducible.

Definition 1.10. *Let $(X_n)_{n\in\mathbb{N}}$ be a Markov chain with transition matrix P. We say that a state $i \in I$ is recurrent if*

$$\mathbb{P}_i(X_n = i \text{ for infinitely many } n) = 1$$

We say that i is transient if

$$\mathbb{P}_i(X_n = i \text{ for infinitely many } n) = 0$$

We define N_i the first passage time to state i by

$$N_i = \inf\{n \ge 1,\ X_n = i\}$$

Theorem 1.11. *The following dichotomy holds:*

(i) if $\mathbb{P}_i(N_i < \infty) = 1$*, then* i *is recurrent and* $\sum_{n\in\mathbb{N}} p_{ii}^{(n)} = \infty$

(ii) if $\mathbb{P}_i(N_i < \infty) < 1$*, then* i *is transient and* $\sum_{n\in\mathbb{N}} p_{ii}^{(n)} < \infty$

In particular, every state is either recurrent or transient.

Proposition 1.12. *Let* C *be a communicating class. Then either all states in* C *are recurrent or all are transient.*

Proposition 1.13. *(i) Every recurrent class is closed*

(ii) Every finite closed class is recurrent

Proposition 1.14. *Let* P *be irreducible and recurrent. Then for all* $j \in I$ *we have* $\mathbb{P}(N_j < \infty) = 1$.

Invariant distributions and positive recurrence

Definition 1.15. *Let* λ *be a measure on* I *and let* P *be a stochastic matrix. We say that* λ *is invariant for* P *if* $\lambda P = \lambda$.

Proposition 1.16. *Let* $(X_n)_{n\in\mathbb{N}}$ *be Markov*(λ, P) *and suppose that* λ *is invariant for* P*. Then* $(X_{m+n})_{n\in\mathbb{N}}$ *is also Markov*(λ, P).

For a fixed state k, consider for each i the expected time spent in i betweet visits to k

$$\gamma_i^k = \mathbb{E}_k\Big[\sum_{n=0}^{N_k-1} \mathbb{1}_{X_n=i}\Big]$$

Proposition 1.17. *Let* P *be irreducible and recurrent. Then:*

(i) $\gamma_k^k = 1$

(ii) $\gamma^k = (\gamma_i^k)_{i\in I}$ *satisfies* $\gamma^k P = \gamma^k$

(iii) $0 < \gamma_i^k < \infty$ *for all* $i \in I$

Theorem 1.18. *Let* P *be irreducible and let* λ *be an invariant measure for* P *with* $\lambda_k = 1$*. Then* $\lambda \geq \gamma^k$*. If in addition* P *is recurrent, then* $\lambda = \gamma^k$*. In particular,* P *has an invariant measure which is unique up to scalar multiples.*

Definition 1.19. *Let* $i \in I$ *be a recurrent state. We say that* i *is positive recurrent if* $\mathbb{E}_i[N_i] < \infty$*. Otherwise we say that* i *is null recurrent.*

Theorem 1.20. *Let* P *be irreducible. Then the following are equivalent:*

(i) every state is positive recurrent

(ii) some state is positive recurrent

(iii) P has an invariant distribution

Moreover, when (iii) holds, the invariant distribution is unique and is given by

$$\pi_i = \frac{1}{\mathbb{E}_i[N_i]}$$

Convergence to equilibrium

Definition 1.21. *We call a state $i \in I$ aperiodic if $p_{ii}^{(n)} > 0$ for all sufficiently large n*

It can be shown that i is aperiodic if and only if $\gcd\{n \geq 1,\ p_{ii}^{(n)}\} = 1$. It can also be shown that if P is irreducible and has an aperiodic state i, then all states are aperiodic.

Theorem 1.22 (Convergence to equilibrium)**.** *Let P be irreducible and aperiodic. Let λ be any distribution, and let $(X_n)_{n\in\mathbb{N}}$ be Markov(λ, P). Then, for all $j \in I$,*

$$\mathbb{P}(X_n = j) \xrightarrow[n\to\infty]{} \frac{1}{\mathbb{E}_j[N_j]}$$

In particular, for all $i, j \in I$,

$$p_{ij}^{(n)} \xrightarrow[n\to\infty]{} \frac{1}{\mathbb{E}_j[N_j]}$$

Ergodic theorem

Theorem 1.23 (Ergodic theorem)**.** *Let P be irreducible and let λ be any distribution. If $(X_n)_{n\in\mathbb{N}}$ is Markov(λ, P) then*

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_{X_k = i} \xrightarrow[n\to\infty]{} \frac{1}{\mathbb{E}_i[N_i]} \text{ a.s.}$$

Moreover, in the positive recurrent case, for any bounded function $f\colon I \to \mathbb{R}$ we have

$$\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) \xrightarrow[n\to\infty]{} \int_I f d\pi \text{ a.s.}$$

where π is the unique invariant distribution.

1.2 Continuous-time Markov chains

Definition and construction

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let I be a countable set equipped with the discrete topology. A continuous-time random process $(X_t)_{t\geq 0}$ with values in I is a family of random variables $X_t\colon \Omega \to I$. We say that $(X_t)_{t\geq 0}$ is right-continuous if all paths are right-continuous, that is for all $\omega \in \Omega$, $t \mapsto X_t(\omega)$ is right-continuous. By definition of the discrete topology on I, every path of a right-continuous process must remain constant for a while in each new state. As in the discrete case, given a filtration $(\mathcal{F}_t)_{t\geq 0}$ on $(\Omega, \mathcal{F})$ we say that a random variable $T\colon I \to [0,\infty]$ is a stopping time if for $t \geq 0$, $\{T \leq t\} \in \mathcal{F}_t$. The σ-algebra $\mathcal{F}_T$ of events determined prior to the stopping time T consists of those events $A \in \mathcal{F}$ for which $A \cap \{T \leq t\} \in \mathcal{F}_t$ for every $t \geq 0$. If $(X_t)_{t\geq 0}$ is a random process, we define X_T on $\{T < \infty\}$ by

$$X_T(\omega) = X_{T(\omega)}(\omega)$$

Now let $(X_t)_{t\geq 0}$ be a right-continuous process on I. We define the jump times $J_0, J_1, \dots$ of $(X_t)_{t\geq 0}$ by

$$\begin{aligned} J_0 &= 0 \\ J_{n+1} &= \inf\{t \geq J_n,\ X_t \neq X_{J_n}\} \end{aligned}$$

J_n is a stopping time for $(X_t)_{t\geq 0}$ for $n \in \mathbb{N}$. We define the holding times $S_1, S_2, \dots$ by

$$S_n = \begin{cases} J_n - J_{n-1} & \text{if } J_{n-1} < \infty \\ \infty & \text{otherwise} \end{cases}$$

Note that right-continuity forces $S_n > 0$ for all $n \geq 1$. The explosion time ζ is defined by

$$\zeta = \sup_n J_n = \sum_{n=1}^{\infty} S_n$$

The discrete-time process $(\xi_n)_{n\in\mathbb{N}}$ given by $\xi_n = X_{J_n}$ is called the jump process of $(X_t)_{t\geq 0}$. This is simply the sequence of values taken by $(X_t)_{t\geq 0}$ up to explosion.

We shall not consider what happens to a process after explosion. So it is convenient to adjoin to I a new state, say ∞, and require that $X_t = \infty$ if $t \geq \zeta$. Any process satisfying this requirement is called minimal. Note that a minimal process may be reconstructed from its jumping (or holding) times and jump process

$$X_t = \begin{cases} \xi_n \text{ if } J_n \leq t < J_{n+1} \\ \infty \text{ if } t \geq \zeta \end{cases}$$

Definition 1.24. *We say that $Q = (q_{ij})_{i,j\in I}$ is a Q-matrix on I if it satisfies the following conditions:*

(i) $0 \leq -q_{ii} < \infty$ *fir all $i \in I$*

(ii) $q_{ij} > 0$ *for all $i, j \in I$*

(iii) $\sum_{j\in I} q_{ij} = 0$ *for all i*

We will sometimes find it convenient to write a_i or $a(i)$ as an alternative notation for $-q_{ii}$. If Q is a Q-matrix, we define the jump matrix Π of Q by

$$\text{if } a_i > 0, \ \begin{cases} \pi_{ij} = q_{ij}/a_i & \text{if } j \neq i \\ \pi_{ii} = 0 \end{cases}$$

$$\text{if } a_i = 0, \ \begin{cases} \pi_{ij} = 0 & \text{if } j \neq i \\ \pi_{ii} = 1 \end{cases}$$

Then Π is a stochastic matrix on I.

Now we can give the definition of a continuous-time Markov chain. Let $(X_t)_{t\geq 0}$ be a minimal right-continuous process on I. We say that $(X_t)_{t\geq 0}$ is a Markov chain with initial distribution λ and generator matrix Q if its jump chain $(\xi_n)_{n\in\mathbb{N}}$ is a discrete-time Markov chain with initial distribution λ and transition matrix Π and if conditional on $(\xi_n)_{n\in\mathbb{N}}$, its holding times S_n are independent and distributed according to an exponential law of parameter $a(\xi_n)$. We say that $(X_t)_{t\geq 0}$ is Markov(λ, Q) for short.

Remark 1.25. *Notice that the conditional distribution of $(S_1, S_2, \dots, S_n)$ given $(\xi_k)_{k\in\mathbb{N}}$ depends only on $(\xi_0, \xi_1, \dots, \xi_{n-1})$ and therefore, conditional on $(\xi_0, \xi_1, \dots, \xi_{n-1})$, $S_1, S_2, \dots, S_n$ are still independent exponential random variables of parameters $a(\xi_0), a(\xi_1), \dots, a(\xi_{n-1})$ respectively.*

We can construct such a process as follows: let $(\xi_n)_{n\in\mathbb{N}}$ be discrete-time Markov(λ, Π) and let $(T_n)_{n\geq 1}$ be independent exponential random variables of parameter 1, independent of $(\xi_n)_{n\in\mathbb{N}}$. Set $S_n = T_n/a(\xi_{n-1})$, $J_n = \sum_{k=1}^n S_k$ and

$$X_t = \begin{cases} \xi_n \text{ if } J_n \leq t < J_{n+1} \\ \infty \text{ if } t \geq \zeta \end{cases}$$

Then $(X_t)_{t\geq 0}$ has the required properties. We now turn our attention to the explosion time ζ of a Markov chain $(X_t)_{t\geq 0}$ with Q-matrix Q. We are interested in conditions that guarantee $(X_t)_{t\geq 0}$ is non-explosive (i.e. $\zeta = \infty$ a.s.). We begin with a lemma:

Lemma 1.26. *Let $(S_n)_{n\geq 1}$ be a sequence of independent random variables with $S_n \sim \mathcal{E}(\lambda_n)$ and $0 < \lambda_n < \infty$ for all n.*

(i) if $\sum_{n=1}^\infty \frac{1}{\lambda_n} < \infty$, then $\sum_{n=1}^\infty S_n < \infty$ a.s.

(ii) if $\sum_{n=1}^\infty \frac{1}{\lambda_n} = \infty$, then $\sum_{n=1}^\infty S_n = \infty$ a.s.

Proof. (i) Suppose $\sum_{n=1}^\infty \frac{1}{\lambda_n} < \infty$. Then, by monotone convergence

$$\mathbb{E}\big[\sum_{n=1}^\infty S_n\big] = \sum_{n=1}^\infty \frac{1}{\lambda_n} < \infty$$

so $\sum_{n=1}^\infty S_n < \infty$ a.s.

(ii) Suppose that $\sum_{n=1}^\infty \frac{1}{\lambda_n} = \infty$. Then $\sum_{n=1}^\infty \min(1, \frac{1}{2\lambda_n}) = \infty$, and since $\log(1+x) \geq \min(1, \frac{1}{2x})$ for all $x \geq 0$, $\sum_{n=1}^\infty \log(1 + 1/\lambda_n) = \infty$. So, by monotone convergence and independence,

$$\begin{aligned} \mathbb{E}\Big[\exp\Big(-\sum_{n=1}^\infty S_n\Big)\Big] &= \prod_{n=1}^\infty \mathbb{E}\big[\exp(-S_n)\big] \\ &= \prod_{n=1}^\infty \big(1 + \frac{1}{\lambda_n})\big)^{-1} = \exp\Big(-\sum_{n=1}^\infty \log(1 + 1/\lambda_n)\Big) = 0 \end{aligned}$$

So $\exp\Big(-\sum_{n=1}^\infty S_n\Big) = 0$ and $\sum_{n=1}^\infty S_n = \infty$ a.s.

∎

Theorem 1.27. *Let $(X_t)_{t\geq 0}$ be Markov(λ, Q). Then $(X_t)_{t\geq 0}$ does not explode if any one of the following conditions holds:*

(i) $\sup_{i\in I} a_i < \infty$

(ii) $X_0 = i$, *and i is recurrent for the jump chain.*

Proof. Set $T_n = a(\xi_{n-1})S_n$. Then conditional on $(\xi_n)_{n\in\mathbb{N}}$, $(T_n)_{n\geq 1}$ are independent exponential random variables with parameter 1. The conditional distribution of $(T_n)_{n\geq 1}$ given $(\xi_n)_{n\in\mathbb{N}}$ does not depend on $(\xi_n)_{n\in\mathbb{N}}$, so $(T_n)_{n\geq 1}$ are unconditionally independent exponential random variables with parameter 1.

(i) Set $a = \sup_{i\in I} a_i$. Then

$$a\zeta = \sum_{n=1}^{\infty} aS_n \geq \sum_{n=1}^{\infty} T_n = \infty$$

a.s. using the previous lemma.

(ii) i is recurrent for the jump chain, which means that $(\xi_n)_{n\in\mathbb{N}}$ visits i infinitely often, at times $N_1, N_2, \ldots$, says. Then

$$a_i\zeta = \sum_{n=1}^{\infty} a_iS_n \geq \sum_{m=1}^{\infty} a_iS_{N_m+1} = \sum_{m=1}^{\infty} T_{N_m+1} = \infty$$

a.s.

□

Remark 1.28. *In particular, if I is finite or if the jump chain is irreducible and recurrent, then $(X_t)_{t\geq 0}$ does not explode.*

Strong Markov property

Theorem 1.29 (Strong Markov property)**.** *Let $(X_t)_{t\geq 0}$ be a continuous-time Markov chain with Q-matrix Q, and let T be a stopping time for the natural filtration $(\mathcal{F}_t)_{t\geq 0}$ of $(X_t)_{t\geq 0}$. Then, conditional on $T < \infty$ and $X_T = i$, $(X_{t+T})_{t\geq 0}$ is Markov(δ_i, Q) and is independent of $\mathcal{F}_T$.*

Transition function, Kolmogorov equations

In this paragraph, we will assume that $a = \sup_{i\in I} a_i < \infty$ and that for all $i \in I$, $a_i > 0$ so that $\pi_{ii} = 0$. We introduce the matrices $P(t)$ on I defined by

$$p_{ij}(t) = \mathbb{P}_i(X_t = j)$$

the probability to be at site j at time t for the Markov chain that starts from i. It follows from the Markov property of the process that these matrices form a semigroup, i.e. that $P(t+s) = P(t)P(s)$ or more explicitly

$$p_{ij}(t+s) = \sum_{k\in I} p_{ik}(t)p_{kj}(s)$$

On the other hand, since we assumed that $\sup_i a_i < \infty$, it follows from the previous identity that the function $t \mapsto P(t)$ is continuously differentiable. In fact, we have the following result.

Theorem 1.30. *$t \mapsto P(t)$ satisfy the Chapman-Kolmogorov backward equations*

$$\begin{cases} p_{ij}(0) = \delta_{ij} \\ p'_{ij}(t) = \sum_{k \in I} a_i \pi_{ik}(p_{kj}(t) - p_{ij}(t)) \end{cases}$$

which can be written in the more compact form

$$\begin{cases} P(0) = id \\ P'(t) = QP(t) \end{cases}$$

where Q is the Q-matrix of the Markov chain $(X_t)_{t \geq 0}$. In particular,

$$P'(0) = Q$$

so that as $t \to 0$, for all $i, j \in I$,

$$p_{ij}(t) = \delta_{ij} + q_{ij} t + o(t)$$

Proof. The proof relies on the following estimates. For all $i \neq j$,

$$\Big| \frac{p_{ii}(t) - 1}{t} + a_i \Big| \leq 2a^2 t \tag{a}$$

$$\Big| \frac{p_{ij}(t)}{t} - a_i \pi_{ij} \Big| \leq a^2 t \Big(\pi_{ij} + \sum_{n=2}^{\infty} \pi_{ij}^{(n+2)} \frac{(at)^n}{n!} \Big) \tag{b}$$

Assuming these inequalities hold, and using the semigroup property of $(P(t))_{t \geq 0}$, we have

$$\frac{p_{ij}(t+h) - p_{ij}(t)}{h} - (QP(t))_{ij} = \sum_{k \neq i} \Big(\frac{p_{ik}(h)}{h} - a_i \pi_{ik} \Big) p_{kj}(t) + \Big(\frac{p_{ii}(h) - 1}{h} + a_i \Big) p_{ij}(t)$$

Then

$$\begin{aligned} \Big| \frac{p_{ij}(t+h) - p_{ij}(t)}{h} - (QP(t))_{ij} \Big| &\leq \sum_{k \neq i} \Big| \frac{p_{ik}(h)}{h} - a_i \pi_{ik} \Big| + \Big| \frac{p_{ii}(h) - 1}{h} + a_i \Big| \\ &\leq \sum_{k \neq i} a^2 h \Big(\pi_{ik} + \sum_{n \geq 0} \pi_{ik}^{(n+2)} \frac{(ah)^n}{n!} \Big) + 2a^2 h \quad \text{by (a) and (b)} \\ &\leq \big(1 + \exp(ah)\big) a^2 h + 2a^2 h \xrightarrow[h \to 0]{} 0 \end{aligned}$$

so that $p'_{ij}(t) = (QP(t))_{ij}$. By definition of $P(t)$, it is clear that $p_{ij}(0) = \delta_{ij}$ which proves the theorem. It remains to prove inequalities (a) and (b). We start with (a). To compute the probability to be at i at time t for a Markov chain starting from i, we may decompose the event $\{X_t = i\}$ according to the number of jumps before time t. Thus, since $\sup_i a_i < \infty$, the chain is non-explosive and

$$p_{ii}(t) = \mathbb{P}_i(X_t = i) = \sum_{n=0}^{\infty} \mathbb{P}_i(X_t = i, J_n \leq t < J_{n+1}) = \sum_{n=0}^{\infty} \mathbb{P}_i(\xi_n = i, J_n \leq t < J_{n+1})$$

The first two terms in the sum are

$$\mathbb{P}_i(\xi_0 = i, J_0 \le t < J_1) + \mathbb{P}_i(\xi_1 = i, J_1 \le t < J_2) = \mathbb{P}_i(t < J_1) + 0 = \exp(-a_i t)$$

so that

$$\begin{aligned}\Big|\frac{p_{ii}(t)-1}{t} + a_i\Big| &\le \Big|\frac{1}{t}\big(1-\exp(-a_i t)\big) - a_i\Big| + \frac{1}{t}\sum_{n\ge 2}\mathbb{P}_i(\xi_n = i, J_n \le t < J_{n+1}) \\ &\le t a_i^2 + \frac{1}{t}\sum_{n\ge 2}\mathbb{P}_i(\xi_n = i, J_n \le t < J_{n+1})\end{aligned}$$

by Taylor's theorem. Now

$$\begin{aligned}\sum_{n\ge 2}\mathbb{P}_i(\xi_n = i, J_n \le t < J_{n+1}) &= \mathbb{P}_i\Big(\bigcup_{n\ge 2}\{\xi_n = i, J_n \le t < J_{n+1}\}\Big) \\ &\le \mathbb{P}_i(J_2 \le t) \ \text{ because } (J_n)_n \text{ is increasing} \\ &= \sum_{k\ne i}\mathbb{P}_i(S_1 + S_2 \le t | \xi_1 = k)\mathbb{P}_i(\xi_1 = k) \\ &\le \sum_{k\ne i}\mathbb{P}_i\Big(\frac{a_i}{a}S_1 + \frac{a_k}{a}S_2 \le t\Big|\xi_1 = k\Big)\mathbb{P}_i(\xi_1 = k)\end{aligned}$$

Conditional on $\xi_0 = i$ and $\xi_1 = k$, S_1 and S_2 are independent exponential random variables with parameters a_i and a_k respectively. So $\mathbb{P}_i\Big(\frac{a_i}{a}S_1 + \frac{a_k}{a}S_2 \le t\Big|\xi_1 = k\Big) = \mathbb{P}(Y \le t)$ where $Y \sim \Gamma(2, a)$. In particular,

$$\mathbb{P}_i\Big(\frac{a_i}{a}S_1 + \frac{a_k}{a}S_2 \le t\Big|\xi_1 = k\Big) \le \frac{1}{2}a^2 t^2$$

and finally

$$\begin{aligned}\Big|\frac{p_{ii}(t)-1}{t} + a_i\Big| &\le a^2 t + \frac{1}{t}\sum_{k\ne i}\mathbb{P}(Y \le t)\mathbb{P}_i(\xi_1 = k) \\ &\le a^2 t + \frac{1}{t}\mathbb{P}(Y \le t) \le 2a^2 t\end{aligned}$$

which proves (a). We now turn to the estimate (b). With the same decomposition of the event $\{X_t = i\}$, we have that for $j \ne i$

$$\begin{aligned}p_{ij}(t) = \mathbb{P}_i(X_t = j) &= \sum_{n=0}^{\infty}\mathbb{P}_i(\xi_n = j, J_n \le t < J_{n+1}) \\ &= \mathbb{P}_i(\xi_1 = j, J_1 \le t < J_2) + \sum_{n=2}^{\infty}\mathbb{P}_i(\xi_n = j, J_n \le t < J_{n+1})\end{aligned}$$

and then

$$\Big|\frac{p_{ij}(t)}{t} - a_i\pi_{ij}\Big| \le \Big|\frac{1}{t}\mathbb{P}_i(\xi_1 = j, J_1 \le t < J_2) - a_i\pi_{ij}\Big| + \frac{1}{t}\sum_{n\ge 2}\mathbb{P}_i(\xi_n = j, J_n \le t < J_{n+1})$$

Now for any $n \geq 1$, conditional on $(\xi_0, \dots, \xi_n)$, S_{n+1} is an exponential random variable with parameter $a(\xi_n)$. In particular, conditional on $\mathcal{F}_{J_n}$, S_{n+1} is an exponential random variable with parameter $a(\xi_n)$, and

$$\mathbb{P}_i(J_{n+1} > t | \mathcal{F}_{J_n}) = \mathbb{P}_i(S_{n+1} > t - J_n | \mathcal{F}_{J_n}) = \exp\Big(- a(\xi_n)(t - J_n)\Big)$$

We use this for $n = 1$ to get the equality

$$\mathbb{P}_i(\xi_1 = j, J_1 \leq t < J_2) = \mathbb{E}_i\Big[\mathbb{1}_{\xi_1 = j}\mathbb{1}_{J_1 \leq t} \exp\big(- a_j(t - J_1)\big)\Big]$$

and we can compute the last expression and then bound the term

$$\Big|\frac{1}{t}\mathbb{P}_i(\xi_1 = j, J_1 \leq t < J_2) - a_i\pi_{ij}\Big|$$

from above using Taylor's theorem. Similarly, since ξ_n and J_n are $\mathcal{F}_{J_n}$-measurable, we have

$$\begin{aligned}
\mathbb{P}_i(\xi_n = j, J_n \leq t < J_{n+1}) &= \mathbb{E}_i\Big[\mathbb{1}_{\xi_n}\mathbb{1}_{J_n \leq t}\mathbb{E}_i[\mathbb{1}_{J_{n+1} > t} | \mathcal{F}_{J_n}]\Big] \\
&= \mathbb{E}_i\Big[\mathbb{1}_{\xi_n = j}\mathbb{1}_{J_n \leq t} \exp\big(- a(\xi_n)(t - J_n)\big)\Big] \\
&\leq \mathbb{P}_i(\xi_n = j, J_n \leq t) \\
&= \mathbb{P}_i(J_n \leq t | \xi_n = j)\pi_{ij}^{(n)}
\end{aligned}$$

As before, we want to bound $\mathbb{P}_i(J_n \leq t | \xi_n = j)$ writing $J_n = S_1 + \ldots + S_n$ and using the conditional distribution of $(S_1, \dots, S_n)$ given $(\xi_0, \dots, \xi_{n-1})$.

$$\begin{aligned}
\mathbb{P}_i(J_n \leq t | \xi_n = j) &= \mathbb{P}_i(S_1 + \cdots + S_n \leq t | \xi_n = j) \\
&= \sum_{\mathbf{z} = j_1, \dots, j_{n-1}} \mathbb{P}_i(S_1 + \ldots + S_n \leq t | A_z, \xi_n = j)\mathbb{P}_i(A_z | \xi_n = j) \\
&\leq \sum_{\mathbf{z} = j_1, \dots, j_{n-1}} \mathbb{P}(Y \leq t)\mathbb{P}_i(A_z | \xi_n = j) \\
&= \mathbb{P}(Y \leq t) \leq \frac{(at)^n}{n!}
\end{aligned}$$

where $Y \sim \Gamma(n, a)$ and $A_z = \{\xi_1 = j_1, \dots, \xi_{n-1} = j_{n-1}\}$. Combining everything, we get (b). $\square$

Remark 1.31. *In a similar way, we can show that $t \mapsto P(t)$ satisfy the Chapman-Kolmogorov forward equations*

$$P'(t) = P(t)Q$$

Theorem 1.32. *Let $(X_t)_t$ be a continuous-time Markov chain with generator L and let $(\mathcal{F}_t)_t$ be its natural filtration. Let $F \colon \mathbb{R}_+ \times I \to \mathbb{R}$ be a bounded function which is smooth in the first coordinate uniformly over the second: for each $i \in I$, $F(\cdot, i)$ is twice continuously differentiable and there exists a constant C such that for $j = 1, 2$,*

$$\sup_{(s,i)}\Big|\frac{d^j F}{ds^j}(s, i)\Big| \leq C$$

Define the real processes M_t^F and N_t^F by

$$M_t^F = F(t, X_t) - F(0, X_0) - \int_0^t (\partial_s + L)F(s, X_s)ds$$
$$N_t^F = (M_t^F)^2 - \int_0^t \Big(LF^2(s, X_s) - 2F(s, X_s)LF(s, X_s)\Big)ds$$

Then $(M_t^F)_t$ and $(N_t^F)_t$ are $(\mathcal{F}_t)_t$- martingales.

Class structure and recurrence

From now on, we consider only Markov chains that are minimal (i.e. those that die after explosion). Then the class structure is simply the discrete-time class structure of the jump chain $(\xi_n)_{n\in\mathbb{N}}$.

Definition 1.33. *We say that i leads to j and write $i \to j$ if*

$$\mathbb{P}_i(X_t = j \text{ for some } t \geq 0) > 0$$

We say that i communicates with j and write $i \leftrightarrow j$ if both $i \to j$ and $j \to i$.

The notions of communicating class, closed class and irreducibility are inherited from the jump chain.

Proposition 1.34. *For distinct states $i, j \in I$ the following are equivalent:*

(i) $i \to j$

(ii) $i \to j$ *for the jump chain*

(iii) $q_{i_0 i_1} q_{i_1 i_2} \dots q_{i_{n-1} i_n} > 0$ *for some states $i_0, i_1, \dots, i_n$ with $i_0 = i$ and $i_n = j$*

(iv) $p_{ij}(t) > 0$ *for all $t > 0$*

(v) $p_{ij}(t) > 0$ *for some $t > 0$*

Proof. Implications $(iv) \Rightarrow (v) \Rightarrow (i) \Rightarrow (ii)$ are clear.

$(ii) \Rightarrow (iii)$ By Proposition 1.9, there are states $i_0, i_1, \dots, i_n$ with $i_0 = i$ and $i_n = j$ such that $\pi_{i_0 i_1} \pi_{i_1 i_2} \dots \pi_{i_{n-1} i_n} > 0$, which implies (iii)

$(iii) \Rightarrow (iv)$ If $q_{ij} > 0$, then

$$\begin{aligned} p_{ij}(t) &\geq \mathbb{P}_i(J_1 \leq t, \xi_1 = j, S_2 > t) \\ &= (1 - \exp(-a_i t))\pi_{ij} \exp(-a_j t) > 0 \end{aligned}$$

for all $t > 0$, so if (iii) holds, then

$$p_{ij}(t) \geq p_{i_0 i_1}(t/n) \dots p_{i_{n-1} i_n}(t/n) > 0$$

for all $t > 0$, and (iv) holds.

□

Definition 1.35. *We say that a state $i \in I$ is recurrent if*

$$\mathbb{P}_i\Big(\{t \geq 0,\ X_t = i\} \text{ is unbounded}\Big) = 1$$

We say that i is transient if

$$\mathbb{P}_i\Big(\{t \geq 0,\ X_t = i\} \text{ is unbounded}\Big) = 0$$

We define T_i the first passage time of $(X_t)_{t\geq 0}$ to state i by

$$T_i = \inf\{t \geq J_1,\ X_t = i\}$$

We also denote by N_i the first passage time of $(\xi_n)_{n\in\mathbb{N}}$ to state i.

Note that if $(X_t)_{t\geq 0}$ can explode starting from i, then i is certainly not recurrent.

Theorem 1.36. *(i) if i is recurrent for the jump chain $(\xi_n)_{n\in\mathbb{N}}$ then i is recurrent for $(X_t)_{t\geq 0}$*

(ii) if i is transient for the jump chain, then i is transient for $(X_t)_{t\geq 0}$

(iii) every state is either recurrent or transient

(iv) recurrence and transience are class properties

Proof. (i) Suppose i is recurrent for $(\xi_n)_{n\in\mathbb{N}}$. If $X_0 = i$ then $(X_t)_{t\geq 0}$ does not explode and $J_n \to \infty$. Also $X_{J_n} = \xi_n = i$ infinitely often, so $\{t \geq 0,\ X_t = i\}$ is unbounded, with probability 1.

(ii) Suppose i is transient for $(\xi_n)_{n\in\mathbb{N}}$. If $X_0 = i$, then

$$N = \sup\{n \in \mathbb{N},\ \xi_n = i\} < \infty$$

so $\{t \geq 0,\ X_t = i\}$ is bounded by J_{N+1} which is finite, with probability 1, because $(\xi_n,\ n \leq N)$ cannot include an absorbing state.

□

Theorem 1.37. *The following dichotomy hold:*

(i) if $\big(a_i = 0\big)$ or $\big(a_i > 0$ and $\mathbb{P}_i(T_i < \infty) = 1\big)$, then i is recurrent and $\int_0^\infty p_{ii}(t)dt = \infty$

(ii) if $a_i > 0$ and $\mathbb{P}_i(T_i < \infty) < 1$, then i is transient and $\int_0^\infty p_{ii}(t)dt < \infty$

Proof. If $a_i = 0$, then $(X_t)_{t\geq 0}$ cannot leave i, so i is recurrent, $p_{ii}(t) = 1$ for all t, and $\int_0^\infty p_{ii}(t)dt = \infty$. Suppose then that $a_i > 0$. Then

$$\mathbb{P}_i(N_i < \infty) = \mathbb{P}_i(T_i < \infty)$$

so i is recurrent if and only if $\mathbb{P}_i(T_i < \infty) = 1$ by Theorem 1.36 and the corresponding result for the jump chain. We shall now show that

$$\int_0^\infty p_{ii}(t)dt = \frac{1}{a_i}\sum_{n=0}^\infty \pi_{ii}^{(n)}$$

so that i is recurrent if and only if $\int_0^\infty p_{ii}(t)dt = \infty$, again using the corresponding result for the jump chain. To establish the equality, we use Fubini's theorem

$$\begin{aligned}
\int_0^\infty p_{ii}(t)dt &= \int_0^\infty \mathbb{E}_i[\mathbb{1}_{X_t=i}]dt &&= \mathbb{E}_i\int_0^\infty \mathbb{1}_{X_t=i}dt \\
&= \mathbb{E}_i\int_0^\zeta \mathbb{1}_{X_t=i}dt &&= \mathbb{E}_i\sum_{n=0}^\infty\int_{J_n}^{J_{n+1}} \mathbb{1}_{\xi_n=i}dt \\
&= \mathbb{E}_i\sum_{n=0}^\infty S_{n+1}\mathbb{1}_{\xi_n=i} &&= \sum_{n=0}^\infty \mathbb{E}_i[S_{n+1}|\xi_n=i]\mathbb{P}_i(\xi_n=i) \\
&= \frac{1}{a_i}\sum_{n=0}^\infty \pi_{ii}^{(n)}
\end{aligned}$$

□

Finally, we show that recurrence and transience are determined by any discrete-time sampling of $(X_t)_{t\geq 0}$.

Proposition 1.38. *Let $h > 0$ be given and set $Z_n = X_{nh}$.*

(i) if i is recurrent for $(X_t)_{t\geq 0}$, then i is recurrent for $(Z_n)_{n\in\mathbb{N}}$

(ii) if i is transient for $(X_t)_{t\geq 0}$, then i is transient for $(Z_n)_{n\in\mathbb{N}}$

Proof. To prove (i) we use for $nh \leq t < (n+1)h$ the estimate

$$p_{ii}\big((n+1)h\big) \geq \exp(-a_i h)p_{ii}(t)$$

Then by monotone convergence,

$$\int_0^\infty p_{ii}(t)dt \leq h\exp(a_i h)\sum_{n=1}^\infty p_{ii}(nh)$$

and (i) follows by Theorem 1.11. To prove the estimate above, write

$$\begin{aligned}
p_{ii}\big((n+1)h\big) &= \sum_{j\in I} p_{ij}(t)p_{ji}\big((n+1)h-t\big) \\
&\geq p_{ii}(t)p_{ii}\big((n+1)h-t\big)
\end{aligned}$$

combined with the following inequality

$$\begin{aligned}
p_{ii}\big((n+1)h-t\big) &= \mathbb{P}_i(X_{(n+1)h-t}=i) \\
&\geq \mathbb{P}_i(J_1 > (n+1)h-t) \\
&\geq \mathbb{P}_i(J_1 > h) = \exp(-a_i h)
\end{aligned}$$

(ii) is clear. □

Invariant distribution and positive recurrence

Definition 1.39. *Let λ be a measure on I. We say that λ is invariant for the Q-matrix Q if $\lambda Q = 0$.*

Proposition 1.40. *Let Q be a Q-matrix with jump matrix Π and let λ be a measure on I. The following are equivalent:*

(i) *λ is invariant for Q*

(ii) *$\mu\Pi = \mu$ (i.e. μ is invariant for the stochastic matrix Π) where $\mu_i = a_i\lambda_i$*

Proof. We have $a_i(\pi_{ij} - \delta_{ij}) = q_{ij}$ for all $i, j \in I$ so

$$\big(\mu(\Pi - Id)\big)_j = \sum_{i\in I} \mu_i(\pi_{ij} - \delta_{ij}) = \sum_{i\in I} \lambda_i q_{ij} = (\lambda Q)_j$$

□

This means that we can use the existence and uniqueness results for invariant measures established in the discrete-time case.

Theorem 1.41. *Suppose that Q is irreducible and recurrent. Then Q has an invariant measure λ which is unique up to scalar multiples.*

Definition 1.42. *Let $i \in I$ be a recurrent state. If $a_i = 0$ or if $a_i > 0$ and $\mathbb{E}_i[T_i] < \infty$ then we say i is positive recurrent. Otherwise we say that i is null recurrent.*

As in the discrete-time case, positive recurrence is tied up with the existence of an invariant distribution.

Theorem 1.43. *Let Q be an irreducible Q-matrix. Then the following are equivalent:*

(i) *every state is positive recurrent*

(ii) *some state is positive recurrent*

(iii) *Q is non-explosive and has an invariant distribution λ*

Moreover, when (iii) holds, we have for all $i \in I$

$$\lambda_i = \frac{1}{a_i \mathbb{E}_i[T_i]}$$

Proof. Let us exclude the trivial case $I = \{i\}$. Then irreducibility forces $a_i > 0$ for all $i \in I$. It is obvious that (i) implies (ii). Define $\mu^i = (\mu^i_j)_{j\in I}$ by

$$\mu^i_j = \mathbb{E}_i \int_0^{T_i \wedge \zeta} \mathbb{1}_{X_t = j} dt$$

By monotone convergence,

$$\sum_{j\in I}\mu^i_j = \mathbb{E}_i[T_i\wedge\zeta]$$

By Fubini's theorem, and using the identity $T_i = J_{N_i}$,

$$\begin{aligned}
\mu^i_j &= \mathbb{E}_i\sum_{n=0}^{\infty}\int_0^{T_i}\mathbb{1}_{X_t=j}\mathbb{1}_{J_n\le t<J_{n+1}}dt\\
&= \mathbb{E}_i\sum_{n=0}^{N_i-1}\int_0^{\infty}\mathbb{1}_{\xi_n=j}\mathbb{1}_{J_n\le t<J_{n+1}}dt\\
&= \mathbb{E}_i\sum_{n=0}^{\infty}S_{n+1}\mathbb{1}_{\xi_n=j,n<N_i}\\
&= \sum_{n=0}^{\infty}\mathbb{E}_i[S_{n+1}|\xi_n=j]\mathbb{E}_i[\mathbb{1}_{\xi_n=j,n<N_i}]\\
&= \frac{1}{a_j}\mathbb{E}_i\sum_{n=0}^{\infty}\mathbb{1}_{\xi_n=j,n<N_i}\\
&= \frac{1}{a_j}\mathbb{E}_i\sum_{n=0}^{N_i-1}\mathbb{1}_{\xi_n=j} = \frac{\gamma^i_j}{a_j}
\end{aligned}$$

Now suppose that *(ii)* holds, then i is certainly recurrent, so the jump chain is recurrent, and Q is non-explosive. We know that $\gamma^i\Pi = \gamma^i$ by Proposition 1.17, so $\mu^i Q = 0$ by Proposition 1.40. But μ^i has finite total mass

$$\sum_{j\in I}\mu^i_j = \mathbb{E}_i[T_i] < \infty$$

so we obtain an invariant distribution λ by setting $\lambda_j = \mu^i_j/\mathbb{E}_i[T_i]$

On the other hand, suppose *(iii)* holds. Denote by λ an invariant distribution for Q. Fix $i\in I$ and set $\nu_j = \frac{\lambda_j a_j}{\lambda_i a_i}$, then $\nu_i = 1$ and $\nu\Pi = \nu$ by Proposition 1.40, so $\nu_j \ge \gamma^i_j$ for all $j\in I$ by Theorem 1.18. So

$$\begin{aligned}
\mathbb{E}_i[T_i] &= \sum_{j\in I}\mu^i_j = \sum_{j\in I}\frac{\gamma^i_j}{a_j}\\
&\le \sum_{j\in I}\frac{\nu_j}{a_j} = \sum_{j\in I}\frac{\lambda_j}{\lambda_i a_i} = \frac{1}{\lambda_i a_i} < \infty
\end{aligned}$$

so i is positive recurrent. Since Q is irreducible, it is recurrent, hence Π is recurrent. By Theorem 1.18, $\nu_j = \gamma^i_j$ and the preceding inequality becomes $\mathbb{E}_i[T_i] = 1/(\lambda_i a_i)$. □

Theorem 1.44. *Let Q be irreducible and recurrent, and let λ be a measure. Let $s>0$ be given. The following are equivalent:*

(i) $\lambda Q = 0$

(ii) $\lambda P(s) = \lambda$

Proof. Since Q is recurrent, it is non-explosive, and $P(s)$ is recurrent by Proposition 1.38. Hence any λ satisfying (i) or (ii) is unique up to scalar multiples. From the proof of Theorem 1.43, if we fix $i \in I$ and set

$$\mu_j = \mathbb{E}_i \int_0^{T_i} \mathbb{1}_{X_t=j} dt$$

then $\mu Q = 0$. Thus to prove the equivalence, it suffices to show that $\mu P(s) = \mu$. By the strong Markov property at T_i,

$$\mathbb{E}_i \int_{T_i}^{T_i+s} \mathbb{1}_{X_t=j} dt = \mathbb{E}_i \int_0^s \mathbb{1}_{X_t=j} dt$$

Hence, by using Fubini's theorem,

$$\begin{aligned}
\mu_j &= \mathbb{E}_i \int_s^{s+T_i} \mathbb{1}_{X_t=j} dt \\
&= \int_0^\infty \mathbb{P}_i(X_{s+t} = j, t < T_i) dt \\
&= \int_0^\infty \sum_{k \in I} \mathbb{P}_i(X_{s+t} = j, t < T_i | X_t = k) \mathbb{P}_i(X_t = k) \\
&= \int_0^\infty \sum_{k \in I} \mathbb{P}(X_s = j | X_0 = k) \mathbb{P}_i(t < T_i | X_t = k) \mathbb{P}_i(X_t = k) \text{ by the Markov property at } t \\
&= \int_0^\infty \sum_{k \in I} p_{kj}(s) \mathbb{P}_i(t < T_i, X_t = k) \\
&= \sum_{k \in I} \left(\mathbb{E}_i \int_0^{T_i} \mathbb{1}_{X_t=k} dt \right) p_{kj}(s) = \sum_{k \in I} \mu_k p_{kj}(s)
\end{aligned}$$

□

Corollary 1.45. *Let Q be an irreducible non-explosive Q-matrix having an invariant distribution λ. If $(X_t)_{t\geq 0}$ is Markov(λ, Q) then so is $(X_{s+t})_{t\geq 0}$ for any $s \geq 0$.*

Theorem 1.46. *Let Q be a Q-matrix and λ be a measure on I. Suppose that $\sup_i a_i < \infty$ and $a_i > 0$ for all i. Then the following are equivalent:*

(i) $\lambda Q = 0$

(ii) $\lambda P(t) = \lambda$ *for every* $t \geq 0$

Convergence to equilibrium

We now investigate the limiting behaviour of $p_{ij}(t)$ as $t \to \infty$. The situation is simpler than in the discrete-case as there is no longer any possibility of periodicity.

Theorem 1.47 (Convergence to equilibrium)**.** *Let Q be an irreducible, non-explosive Q-matrix having an invariant distribution λ. Suppose $(X_t)_{t\geq 0}$ is Markov(μ, Q) where μ is any distribution on I. Then for all $j \in I$, we have*

$$\mathbb{P}(X_t = j) \xrightarrow[t\to\infty]{} \lambda_j$$

Proof. Let $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ be independent copies of the chain, with $X_0 \sim \mu$ and $Y_0 \sim \pi$, and define $\mathcal{F}_t = \sigma(X_s, Y_s, s \leq t)$. Set $W_t = (X_t, Y_t)$, then $(W_t)_{t\geq 0}$ is an irreducible, non-explosive Markov chain on $I \times I$ with initial distribution

$$\mathbb{P}(W_0 = (i,j)) = \mu_i \pi_j$$

and Q-matrix

$$\widetilde{q}_{(i,j),(k,l)} = \delta_{ik} q_{jl} + \delta_{jl} q_{ik}$$

Denote by $\widetilde{\pi}$ the distribution on $I \times I$ defined by $\widetilde{\pi}_{(i,j)} = \pi_i \pi_j$. Then, using the invariance of π for Q,

$$(\widetilde{\pi}\widetilde{Q})_{(k,l)} = \sum_{i,j\in I} \big(\pi_i \pi_j \delta_{ik} q_{jl} + \delta_{jl} q_{ik}\big) = \sum_{j\in I} \pi_k \pi_j q_{jl} + \sum_{i\in I} \pi_i \pi_l q_{ik} = 0$$

So $\widetilde{\pi}$ is an invariant distribution for $\widetilde{Q}$. In particular, $\widetilde{Q}$ is recurrent and, using Proposition 1.14, we have

$$\mathbb{P}(X_t = Y_t = i \text{ for some } t \geq 0) = 1$$

where $i \in I$ is fixed. Now set

$$\begin{aligned}
\tau &= \inf\{t \geq 0,\ X_t = Y_t = i\} \\
Z_t &= \begin{cases} X_t & \text{for } t \leq \tau \\ Y_t & \text{for } t > \tau \end{cases} \\
Z'_t &= \begin{cases} Y_t & \text{for } t \leq \tau \\ X_t & \text{for } t > \tau \end{cases} \\
W'_t &= (Z_t, Z'_t)
\end{aligned}$$

Then τ is an $(\mathcal{F}_t)_{t\geq 0}$-stopping time a.s. finite, and $(X_\tau, Y_\tau) = (i,i)$. By the strong Markov property applied to $(W_t)_{t\geq 0}$ at time τ, $(X_{t+\tau}, Y_{t+\tau})_{t\geq 0}$ is Markov$(\delta_{(i,i)}, \widetilde{Q})$ and is independent of $\mathcal{F}_\tau$. By the same argument, $(Y_{t+\tau}, X_{t+\tau})_{t\geq 0}$ is Markov$(\delta_{(i,i)}, \widetilde{Q})$ and is independent of $\mathcal{F}_\tau$. So $(W'_t)_{t\geq 0}$ is Markov with the same initial distribution as $(W_t)_{t\geq 0}$ and the same Q-matrix. In particular, $(Z_t)_{t\geq 0} \sim (X_t)_{t\geq 0}$ is Markov(μ, Q).

$$\begin{aligned}
\big|\mathbb{P}(X_t = j) - \pi_j\big| &= \big|\mathbb{P}(Z_t = j) - \mathbb{P}(Y_t = j)\big| \\
&= \big|\mathbb{P}(t \leq \tau, X_t = j) - \mathbb{P}(t \leq \tau, Y_t = j)\big| \\
&= \big|\mathbb{E}\big[\mathbb{1}_{t\leq\tau}(\mathbb{1}_{X_t=j} - \mathbb{1}_{Y_t=j})\big]\big| \\
&\leq \mathbb{P}(\tau \geq t) \xrightarrow[t\to\infty]{} \infty
\end{aligned}$$

$\square$

Theorem 1.48. *Let Q be an irreducible Q-matrix and let μ be any distribution. Suppose that $(X_t)_{t\geq 0}$ is Markov(μ, Q). Then, for all $j \in I$,*

$$\mathbb{P}(X_t = j) \xrightarrow[t\to\infty]{} \frac{1}{a_i \mathbb{E}_j[T_j]}$$

Ergodic theorem

Theorem 1.49 (Ergodic theorem)**.** *Let Q be irreducible and let μ be any distribution. Suppose $(X_t)_{t\geq 0}$ is Markov(μ, Q). Then, for $i \in I$,*

$$\frac{1}{t}\int_0^t \mathbb{1}_{X_s=i} ds \xrightarrow[t\to\infty]{} \frac{1}{a_i \mathbb{E}_i[T_i]} \ a.s.$$

Moreover, in the positive recurrent case, for any bounded function $f\colon I \to \mathbb{R}$ we have

$$\frac{1}{t}\int_0^t f(X_s) ds \xrightarrow[t\to\infty]{} \int f d\lambda \ a.s.$$

where λ is the unique invariant distribution.

2 Simple exclusion process

2.1 Definition and basic properties

For a positive integer N, denote by $\mathbb{T}_N = \mathbb{Z}/N\mathbb{Z}$ the torus with N points and let $\mathbb{T}_N^d = (\mathbb{T}_N)^d$. The points of $\mathbb{T}_N^d$ are called sites and are denoted by the letters x, y and z. An interacting particle system is a continuous-time Markov chain $(\eta_t)_t$ with state space $\mathbb{N}^{\mathbb{T}_N^d}$ describing the collective behaviour of stochastically interacting "particles" evolving in $\mathbb{T}_N^d$. Elements of the state space $\mathbb{N}^{\mathbb{T}_N^d}$ are called configurations and are denoted by the letters η, ξ and ζ. For $\eta \in \mathbb{N}^{\mathbb{T}_N^d}$ and $x \in \mathbb{T}_N^d$, we think of $\eta(x) \in \mathbb{N}$ as the number of particles at site x for configuration η. The simplest example of an interaction particle system is the independent random walks: particles evolve as independent continuous-time random walks on the torus. A more sophisticated interacting particle system and the one we will study is the simple exclusion process: instead of just superposing independent random walks, we add a constraint allowing at most one particle per site. The state space is therefore $\{0,1\}^{\mathbb{T}_N^d}$. The particles move on $\mathbb{T}_N^d$ according to the following rules:

- a particle at $x \in \mathbb{T}_N^d$ waits an exponential time with parameter 1, independently of other particles
- at the end of that time, it chooses a neighbouring site y with uniform probability
- if y is vacant, it goes to y, otherwise it stays at x.

Thus an exclusion interaction is superimposed on otherwise independent continuous-time random walks. We are only interested in the nearest-neighbour simple exclusion process, which means that particles only jump to adjacent sites. The generator of this process is

$$L_N f(\eta) = \sum_{x\in\mathbb{T}_N^d} \sum_{|z|=1} (2d)^{-1}\eta(x)(1-\eta(x+z))\big(f(\eta^{x,x+z}) - f(\eta)\big)$$

for $f\colon \{0,1\}^{\mathbb{T}_N^d} \to \mathbb{R}$ and $\eta \in \{0,1\}^{\mathbb{T}_N^d}$, where $\eta^{x,x+z}$ is the configuration obtained from η by letting a particle jump from x to $x+z$ when this is possible. More explicitly, if $\eta(x) = 1$ and

$\eta(x+z)=0$, then

$$\eta^{x,x+z}(y)=\begin{cases}\eta(y) & \text{if } y\notin\{x,x+z\}\\ \eta(x)-1 & \text{if } y=x\\ \eta(x+z)+1 & \text{if } y=x+z\end{cases}$$

Otherwise, $\eta^{x,x+z}=\eta$. Here is a heuristic argument: we know that the generator may be written as

$$L_N f(\eta)=\sum_{\tilde\eta}\lambda(\eta)p(\eta,\tilde\eta)\big(f(\tilde\eta)-f(\eta)\big)$$

Assume that the process is starting from configuration η. Notice that all accessible states are of the form $\eta^{x,x+z}$ since only one particle can move at a time. For a particle at a site x, denote by $\tau(\eta,x)$ the exponential clock with parameter $\lambda(\eta,x)=1$ associated to the configuration. Now the clock of the system is $\tau(\eta)=\min\big\{\tau(\eta,x),\ x\in\mathbb{T}_N^d \text{ such that } \eta(x)=1\big\}$, which is to say that the configuration changes exactly when one particle moves. Then $\tau(\eta)$ is an exponential random variable with parameter $\lambda(\eta)=\sum_{x,\ \eta(x)=1}\lambda(\eta,x)=\sum_{x\in\mathbb{T}_N^d}\eta(x)$. Now

$$p(\eta,\eta^{x,x+z})=\mathbb{P}\big(\tau(\eta,x)=\tau(\eta),x\xrightarrow{\eta}x+z\big)=\mathbb{P}\big(\tau(\eta,x)=\tau(\eta)\big)\mathbb{P}\big(x\xrightarrow{\eta}x+z|\tau(\eta,x)=\tau(\eta)\big)$$

where $\{x\xrightarrow{\eta}x+z\}$ is the event "A particle at site x in configuration η moves to site $x+z$". By a simple computation, $\mathbb{P}\big(\tau(\eta,x)=\tau(\eta)\big)=\lambda(\eta,x)/\lambda(\eta)=1/\lambda(\eta)$. On the other hand, $\mathbb{P}\big(x\xrightarrow{\eta}x+z|\tau(\eta,x)=\tau(\eta)\big)=(2d)^{-1}\eta(x)\mathbb{1}_{|z|=1}$. Thus the generator becomes

$$L_N f(\eta)=\sum_{x\in\mathbb{T}_N^d}\sum_{|z|=1}(2d)^{-1}\eta(x)\big(f(\eta^{x,x+z})-f(\eta)\big)$$

which coincides with the generator given earlier since $\eta^{x,x+z}=\eta$ when $\eta(x+z)=1$.

For $\alpha\in[0,1]$, we denote by $\nu_\alpha=\nu_\alpha^N$ the Bernoulli product measure of parameter α on $\{0,1\}^{\mathbb{T}_N^d}$. More explicitly, under ν_α, the events $\{\eta,\ \eta(x)=1\}$ are independent for $x\in\mathbb{T}_N^d$ and

$$\nu_\alpha(\{\eta,\ \eta(x)=1\})=\alpha=1-\nu_\alpha(\{\eta,\ \eta(x)=0\})$$

In particular, for any configuration η,

$$\nu_\alpha(\{\eta\})=\alpha^{\lambda(\eta)}(1-\alpha)^{N^d-\lambda(\eta)}$$

where $\lambda(\eta)=\sum_x\eta(x)$ is the number of particles.

Theorem 2.1. *The Bernoulli measures $\{\nu_\alpha,\ \alpha\in[0,1]\}$ are invariant for the simple exclusion process.*

Proof. Let Q be the Q-matrix of the Markov chain. Fix a configuration η.

$$\begin{aligned}\nu_\alpha Q(\eta)&=\sum_{\tilde\eta}\nu_\alpha(\tilde\eta)Q(\tilde\eta,\eta)\\&=\sum_{\substack{x\in\mathbb{T}_N^d\\ \eta(x)=1}}\sum_{\substack{|z|=1\\ \eta(x+z)=0}}\nu_\alpha(\eta^{x,x+z})Q(\eta^{x,x+z},\eta)+\nu_\alpha(\eta)Q(\eta,\eta)\end{aligned}$$

where the second equality makes use of the fact that only one particle can move at a time and it can only jump to adjacent sites. On the one hand, as seen earlier, $\nu_\alpha(\eta)$ only depends on η through the total number of particles $\lambda(\eta)$. In particular, $\nu_\alpha(\eta^{x,x+z}) = \nu_\alpha(\eta)$ since both configurations have the same number of particles. On the other hand, if $\eta(x) = 1$ and $\eta(x+z) = 0$, then

$$\begin{aligned} Q(\eta^{x,x+z}, \eta) &= (2d)^{-1}\eta^{x,x+z}(x+z)(1-\eta^{x,x+z}(x)) \\ &= (2d)^{-1} = Q(\eta, \eta^{x,x+z}) \end{aligned}$$

So

$$\nu_\alpha Q(\eta) = \nu_\alpha(\eta)\Big(\sum_{\substack{x\in\mathbb{T}_N^d \\ \eta(x)=1}} \sum_{\substack{|z|=1 \\ \eta(x+z)=0}} Q(\eta,\eta^{x,x+z}) + Q(\eta,\eta)\Big) = 0$$

using the zero-sum property of a Q-matrix, which concludes the proof. $\square$

2.2 Hydrodynamic equation for simple exclusion process

A hydrodynamic equation describes the macroscopic behaviour of a system when microscopic interactions occur between the particles. In this section we prove the hydrodynamic behaviour of nearest neighbour symmetric simple exclusion processes and show that the hydrodynamic equation is the heat equation. Let $(\eta_t^N)_{t\geq 0}$ be the symmetric simple exclusion process on $\mathbb{T}_N^d$. $(\eta_t^N)_{t\geq 0}$ is a random element taking values in $D\big([0,\infty), \{0,1\}^{\mathbb{T}_N^d}\big)$. In order to work in a fixed space as N increases, we consider the empirical measure associated to the particle system defined by

$$\pi_t^N = \frac{1}{N^d}\sum_{x\in\mathbb{T}_N^d} \eta_t^N(x)\delta_{x/N}$$

We shall prove that π_t^N converges, in a way to be specified later, to a measure absolutely continuous with respect to the Lebesgue measure and satisfying the heat equation in a weak sense.

We briefly present the strategy of the proof. Fix $T > 0$. The time trajectory of the empirical measure $(\pi_t^N)_{0\leq t\leq T}$ is a random element taking values in $D\big([0,T], \mathcal{M}_+(\mathbb{T}^d)\big)$ where $\mathbb{T}^d$ is the d-dimensional torus and $\mathcal{M}_+(\mathbb{T}^d)$ is the set of positive measures on $\mathbb{T}^d$ with mass bounded by 1. Fix a profile $\rho_0 \colon \mathbb{T}^d \to [0,1]$ and denote by $(\mu^N)_N$ a sequence of probability measures on $\{0,1\}^{\mathbb{T}_N^d}$ associated to ρ_0 in a sense to be specified later. Define

$$\begin{aligned} \sigma \colon \{0,1\}^{\mathbb{T}_N^d} &\longrightarrow \mathcal{M}_+(\mathbb{T}^d) \\ \eta &\longmapsto \frac{1}{N^d}\sum_{x\in\mathbb{T}_N^d} \eta(x)\delta_{x/N} \end{aligned}$$

Now let Q^N be the probability measure on $D\big([0,T], \mathcal{M}_+(\mathbb{T}^d)\big)$ corresponding to the process $(\pi_t^N)_{0\leq t\leq T}$ speeded up by N^2 and starting from $\mu^N\sigma^{-1}$. In other words, Q^N is the distribution of the process $(\pi_t^N)_{0\leq t\leq T}$ with initial distribution $\mu^N\sigma^{-1}$. Our goal is to prove that, for each fixed time t, the empirical measure π_t^N converges in probability to $\rho(t,u)du$ where $\rho(t,u)$ is the solution of the heat equation with initial condition ρ_0.

We shall proceed in two steps. We first prove that the process $(\pi_t^N)_{0\leq t\leq T}$ converges in distribution to the deterministic path $(\rho(t,u)du)_{0\leq t\leq T}$ and then argue that convergence in distribution to a deterministic weakly continuous trajectory implies convergence in probability at any fixed time $0 \leq t \leq T$. To prove the first step, it suffices to show the weak convergence of the sequence of probability measures (Q^N) to the Dirac measure concentrated on the solution of the heat equation.

Weak convergence in $D\big([0,T],\mathcal{M}_+(\mathbb{T}^d)\big)$

We work on $\mathcal{M}_+(\mathbb{T}^d)$, the set of positive measures on $\mathbb{T}^d$ with mass bounded by 1 endowed with the weak topology. Denote by $\mathcal{C}(\mathbb{T}^d)$ the space of continuous functions on $\mathbb{T}^d$ with values in $\mathbb{R}$. Since $\mathbb{T}^d$ is compact, $\mathcal{C}(\mathbb{T}^d)$ endowed with the topology of uniform convergence is a separable Banach space and by the Riesz-Markov theorem, its (topological) dual space is $\mathcal{M}(\mathbb{T}^d)$, the space of signed measures on $\mathbb{T}^d$. By Banach-Alaoglu theorem, the unit ball of $\mathcal{M}(\mathbb{T}^d)$ is weakly (sequentially) compact. The dual norm on $\mathcal{M}(\mathbb{T}^d)$ being the total variation norm, $\mathcal{M}_+(\mathbb{T}^d)$ is a subset of the unit ball of $\mathcal{M}(\mathbb{T}^d)$. Furthermore, $\mathcal{M}_+(\mathbb{T}^d)$ is weakly (sequentially) closed in $\mathcal{M}(\mathbb{T}^d)$. Hence, $\mathcal{M}_+(\mathbb{T}^d)$ is weakly (sequentially) compact.

We may define a metric δ on $\mathcal{M}_+(\mathbb{T}^d)$ by introducing a dense countable family $(g_k)_{f\geq 1}$ in $\mathcal{C}(\mathbb{T}^d)$ and by setting

$$\delta(\mu,\nu)=\sum_{k=1}^{\infty}\frac{1}{2^k}\min\Big\{1,\big|\langle\mu,g_k\rangle-\langle\nu,g_k\rangle\big|\Big\}$$

where $\langle\mu,g\rangle=\int_{\mathbb{T}^d} g d\mu$ for $\mu\in\mathcal{M}_+(\mathbb{T}^d)$ and $g\in\mathcal{C}(\mathbb{T}^d)$. It can be shown that the above metric induces the weak topology on $\mathcal{M}_+(\mathbb{T}^d)$. Then $(\mathcal{M}_+(\mathbb{T}^d),\delta)$ is a compact metric space. In particular, it is separable and complete.

We now turn our attention to the space $D\big([0,T],\mathcal{M}_+(\mathbb{T}^d)\big)$. Since $\mathcal{M}_+(\mathbb{T}^d)$ is a compact metric space, we have the following characterisation of relative compactness for sequences of probability measures on $D\big([0,T],\mathcal{M}_+(\mathbb{T}^d)\big)$ (cf. Appendix B).

Theorem 2.2. *Let (P^N) be a sequence of probability measures on $D\big([0,T],\mathcal{M}_+(\mathbb{T}^d)\big)$. The sequence is relatively compact if and only if*

$$\forall\epsilon>0,\ \lim_{\gamma\to 0}\limsup_{N\to\infty}P^N\big(\mu,w'_\mu(\gamma)>\epsilon\big)=0$$

To prove relative compactness of a sequence (Q^N) of probability measures on $D\big([0,T],\mathcal{M}_+(\mathbb{T}^d)\big)$, the condition stated above is often difficult to verify. Instead, the following corollary states that it is enough to check relative compactness for the sequences of probability measures obtained by projecting (Q^N) onto a dense countable subset of $\mathcal{C}(\mathbb{T}^d)$.

Corollary 2.3. *Let (Q^N) be a sequence of probability measures on $D\big([0,T],\mathcal{M}_+(\mathbb{T}^d)\big)$ and let $\{g_k,\ k\geq 1\}$ be a dense family of $\mathcal{C}(\mathbb{T}^d)$. (Q^N) is relatively compact if and only if for every $k\geq 1$, the sequence $(Q^N T_k^{-1})$ of probability measures on $D\big([0,T],\mathbb{R}\big)$ is relatively compact, where*

$$\begin{aligned}T_k\colon D\big([0,T],\mathcal{M}_+(\mathbb{T}^d)\big)&\longrightarrow D\big([0,T],\mathbb{R}\big)\\(\mu_t)_t&\longmapsto\big(\langle\mu_t,g_k\rangle\big)_t\end{aligned}$$

Proof. For the necessity, we show that T_k is continuous. Let (μ^n) be a sequence in $D\big([0,T],\mathcal{M}_+(\mathbb{T}^d)\big)$ converging to μ in the Skorokhod topology. There exists a sequence (λ_n) of strictly increasing continuous mappings of $[0,T]$ onto itself such that $\|\lambda_n-id\|\to 0$ and

$$\sup_t\delta\big(\mu^n\circ\lambda_n(t),\mu(t)\big)\to 0$$

In particular,

$$\sup_t\min\Big\{1,\big|\langle\mu^n\circ\lambda_n(t),g_k\rangle-\langle\mu(t),g_k\rangle\big|\Big\}\leq 2^k\sup_t\delta\big(\mu^n\circ\lambda_n(t),\mu(t)\big)\to 0$$

so that $\sup_t \left|\langle \mu^n \circ \lambda_n(t), g_k\rangle - \langle \mu(t), g_k\rangle\right| \to 0$ which proves that $\langle \mu^n, g_k\rangle \to \langle \mu, g_k\rangle$ in the Skorokhod topology, so that T_k is continuous and the result follows from the continuous mapping theorem. For the sufficiency, fix $\epsilon > 0$ and $\beta > 0$. Choose k_ϵ such that $2^{1-k_\epsilon} \le \epsilon$. Fix $\gamma > 0$ and a partition $\{t_i\}$ such that $\min_i |t_i - t_{i-1}| > \gamma$. For all i and all $t_{i-1} \le s < t < t_i$, we have

$$\begin{aligned}
\delta(\mu(t), \mu(s)) &= \sum_{k=1}^{\infty} 2^{-k} \min\Big\{1, \big|\langle \mu(t), g_k\rangle - \langle \mu(s), g_k\rangle\big|\Big\} \\
&\le \sum_{k=1}^{k_\epsilon} 2^{-k} \big|\langle \mu(t), g_k\rangle - \langle \mu(s), g_k\rangle\big| + \frac{\epsilon}{2} \\
&\le \sum_{k=1}^{k_\epsilon} 2^{-k} w_{\langle \mu, g_k\rangle}[t_{i-1}, t_i) + \frac{\epsilon}{2}
\end{aligned}$$

So

$$w_\mu[t_{i-1}, t_i) = \sup_{t_{i-1} \le s < t < t_i} \delta(\mu(t), \mu(s)) \le \sum_{k=1}^{k_\epsilon} 2^{-k} w_{\langle \mu, g_k\rangle}[t_{i-1}, t_i) + \frac{\epsilon}{2}$$

Now, taking the maximum over i and then the infimum over all γ-sparse sets $\{t_i\}$, we have

$$w'_\mu(\gamma) \le \sum_{k=1}^{k_\epsilon} 2^{-k} w'_{\langle \mu, g_k\rangle}(\gamma) + \frac{\epsilon}{2}$$

By assumption, there exists γ_0 such that for all $k \le k_\epsilon$, $\gamma \le \gamma_0$ and $N \ge 1$,

$$Q^N\Big(\mu,\ w'_{\langle \mu, g_k\rangle}(\gamma) > \epsilon/2\Big) \le \beta 2^{-k}$$

Therefore, for all $\gamma \le \gamma_0$ and $N \ge 1$,

$$\begin{aligned}
Q^N\Big(\mu,\ \sum_{k=1}^{k_\epsilon} 2^{-k} w'_{\langle \mu, g_k\rangle}(\gamma) > \epsilon/2\Big) &\le Q^N\Big(\bigcup_{k=1}^{k_\epsilon} \{\mu,\ w'_{\langle \mu, g_k\rangle}(\gamma) > \epsilon/2\}\Big) \\
&\le \sum_{k=1}^{k_\epsilon} Q^N\Big(\mu,\ w'_{\langle \mu, g_k\rangle}(\gamma) > \epsilon/2\Big) \le \beta
\end{aligned}$$

Finally

$$Q^N\Big(\mu,\ w'_\mu(\gamma) > \epsilon\Big) \le Q^N\Big(\mu,\ \sum_{k=1}^{k_\epsilon} 2^{-k} w'_{\langle \mu, g_k\rangle}(\gamma) > \epsilon/2\Big) \le \beta$$

□

The hydrodynamic equation

Definition 2.4. *Let $\rho_0 \colon \mathbb{T}^d \to [0,1]$ be an initial density profile. We say that a sequence (μ^N) of probability measures on $\{0,1\}^{\mathbb{T}_N^d}$ is associated to ρ_0 if for every continuous function $G \colon \mathbb{T}^d \to \mathbb{R}$ and for every $\epsilon > 0$, we have*

$$\lim_{N \to \infty} \mu^N\Big(\eta,\ \Big|N^{-d} \sum_{x \in \mathbb{T}_N^d} G(x/N)\eta(x) - \int_{\mathbb{T}^d} G(u)\rho_0(u)du\Big| > \epsilon\Big) = 0$$

or equivalently

$$\lim_{N\to\infty} \mu^N\sigma^{-1}\Big(\pi \in \mathcal{M}_+(\mathbb{T}^d),\ \Big|\langle \pi, G\rangle - \langle \rho_0, G\rangle\Big| > \epsilon\Big) = 0$$

Theorem 2.5. *Let $\rho_0\colon \mathbb{T}^d \to [0,1]$ be an initial density profile and let (μ^N) be a sequence of probability measures associated to ρ_0. Let $(\eta_t^N)_t$ be the simple exclusion process on $\mathbb{T}_N^d$ with initial distribution $\mu^N\sigma^{-1}$ and let π_t^N be the empirical measures associated to the configuration η_t^N. Then, for every $t > 0$, the sequence $(\pi_{N^2t}^N)_N$ converges in probability to the absolutely continuous measure $\rho(t,u)du$ where $\rho(t,u)$ is the unique solution to the heat equation with initial condition ρ_0*

$$\begin{cases}\partial_t\rho(t,u) = (2d)^{-1}\Delta\rho(t,u),\ t \geq 0,\ u \in \mathbb{T}^d\\ \rho(0,u) = \rho_0(u), \qquad u \in \mathbb{T}^d\end{cases}$$

Proof. For the sake of simplicity we will prove the theorem for $d = 1$, the proof being valid in higher dimensions. We start by fixing a time $T > 0$ and letting Q^N be the probability measure on $D\big([0,T], \mathcal{M}_+(\mathbb{T})\big)$ corresponding to the Markov process $(\pi_{N^2t}^N,\ t \in [0,T])$ with initial distribution $\mu^N\sigma^{-1}$. To prove that the sequence Q^N is weakly convergent, we shall show that it is relatively compact and that all converging subsequences converge to the same limit which is the Dirac measure (denoted hereafter by Q^*) concentrated on the solution of the heat equation (cf. Corollary A.6).

First step: relative compactness. Denote by $\mathcal{C}^2(\mathbb{T})$ the space of twice continuously differentiable functions on $\mathbb{T}$. $\mathcal{C}^2(\mathbb{T})$ is separable and dense in $\mathcal{C}(\mathbb{T})$ for the uniform topology. Applying Corollary 2.3. it suffices to check that for every $G \in \mathcal{C}^2(\mathbb{T})$ the sequence $Q^NT_G^{-1}$ is relatively compact where

$$\begin{aligned} T_G\colon D\big([0,T], \mathcal{M}_+(\mathbb{T})\big) &\longrightarrow D\big([0,T], \mathbb{R}\big)\\ (\mu_t)_t &\longmapsto \big(\langle \mu_t, G\rangle\big)_t \end{aligned}$$

Fix a function $G \in \mathcal{C}^2(\mathbb{T})$. Notice that $Q^NT_G^{-1}$ is the distribution of the process $\big(\langle \pi_{N^2t}^N, G\rangle,\ t \in [0,T]\big)$ which takes values in $D\big([0,T], \mathbb{R}\big)$. We shall therefore apply Theorem B.5. For the first condition,

$$\begin{aligned} Q^NT_G^{-1}\big(x,\ \|x\| \geq a\big) &= Q^N\big(\pi,\ \sup_t |\langle \pi_t, G\rangle| \geq a\big)\\ &= \mathbb{P}\big(\sup_t |\langle \pi_{N^2t}^N, G\rangle| \geq a\big) \end{aligned}$$

Since the total mass of the empirical measure $\pi_{N^2t}^N$ is bounded by 1, $|\langle \pi_{N^2t}^N, G\rangle| \leq \|G\|$ and it follows that $\lim_{a\to\infty}\limsup_N Q^NT_G^{-1}(x,\ \|x\| \geq a) = 0$. It remains to prove the second condition, or, as we shall do, the one of Remark B.6. Notice that

$$Q^NT_G^{-1}(x,\ w_x(\gamma) > \epsilon) = \mathbb{P}(w_{\langle \pi_{N^2\cdot}^N, G\rangle}(\gamma) > \epsilon)$$

So we shall prove that

$$\lim_{a\to\infty}\limsup_N \mathbb{P}(w_{\langle \pi_{N^2\cdot}^N, G\rangle}(\gamma) > \epsilon) = 0 \tag{a}$$

Now applying Theorem 1.32 with the Markov chain $(\eta_{N^2t}^N)$ whose generator is N^2L_N and the function

$$\begin{aligned} F\colon \mathbb{R}_+ \times \{0,1\}^{\mathbb{T}_N} &\longrightarrow \mathbb{R}\\ (t,\eta) &\longmapsto \langle \sigma(\eta), G\rangle = N^{-1}\sum_x \eta(x)G(x/N) \end{aligned}$$

we have the identity:

$$\langle \pi^N_{N^2t}, G\rangle = \langle \pi^N_0, G\rangle + \int_0^t N^2 L_N F(s, \eta^N_{N^2s})ds + M_t^{G,N} \tag{b}$$

where $(M_t^{G,N})$ is a martingale. But for sites x, z and configuration η such that $\eta(x) = 1$ and $\eta(x+z) = 0$,

$$F(s, \eta^{x,x+z}) - F(s,\eta) = N^{-1}\sum_y \big(\eta^{x,x+z}(y) - \eta(y)\big)G(y/N) = N^{-1}\big\{G((x+z)/N) - G(x/N)\big\}$$

So by a change of variables and a summation by parts

$$\begin{aligned}
L_N F(s,\eta) &= (2N)^{-1}\sum_x \sum_{|z|=1} \eta(x)(1-\eta(x+z))\big\{G((x+z)/N) - G(x/N)\big\} \\
&= -(2N)^{-1}\sum_x \sum_{|z|=1} \eta(x+z)(1-\eta(x))\big\{G((x+z)/N) - G(x/N)\big\} \\
&= (4N)^{-1}\sum_x \sum_{|z|=1} \big(\eta(x) - \eta(x+z)\big)\big\{G((x+z)/N) - G(x/N)\big\} \\
&= (4N)^{-1}\sum_x \sum_{|z|=1} \eta(x)\big\{G((x+z)/N) + G((x-z)/N) - 2G(x/N)\big\} \\
&= (2N^3)^{-1}\sum_x \eta(x)\Delta_N G(x/N) = (2N^2)^{-1}\langle \sigma(\eta), \Delta_N G\rangle
\end{aligned}$$

where $\Delta_N G(x/N) = N^2\big\{G((x+1)/N) + G((x-1)/N) - 2G(x/N)\big\}$ is the discrete Laplacian. The identity (b) becomes

$$\langle \pi^N_{N^2t}, G\rangle = \langle \pi^N_0, G\rangle + (1/2)\int_0^t \langle \pi^N_{N^2s}, \Delta_N G\rangle ds + M_t^{G,N} \tag{c}$$

Now to prove (a), it suffices to prove the same condition for each of the terms of the identity (c). The initial term $\langle \pi^N_0, G\rangle$ does not contribute in this respect. For the second term, the condition follows from the estimate

$$\Big|\int_s^t \langle \pi^N_{N^2s}, \Delta_N G\rangle ds\Big| \le \gamma\|\Delta_N G\| \le \gamma\|G''\|$$

for all $\gamma > 0$ and $|t-s| \le \gamma$. The second inequality is a consequence of Taylor's theorem. As for the third term, we first recall Doob's maximal inequality for a right-continuous, square-integrable martingale (M_t):

$$\mathbb{P}\Big(\sup_{t\in[a,b]}\big|M_t - M_a\big| > \epsilon\Big) \le \frac{1}{\epsilon^2}\mathbb{E}\Big[\sup_{t\in[a,b]}\big|M_t - M_a\big|\Big] \le \frac{4}{\epsilon^2}\mathbb{E}\Big[\big|M_b - M_a\big|^2\Big]$$

Now we fix $\gamma > 0$ and we choose a γ-sparse set $\{t_i\}$. Notice that for any pair $t, s \in [0,T]$ satisfying $|t-s| \le \gamma$, either both t, s lie in the same interval $(t_{i-1}, t_i]$ or they lie in adjacent intervals. Notice also that identity (b) clearly shows $(M_t^{G,N})$ is right-continuous, so we can apply

Doob's inequality. We have

$$\begin{aligned}
\mathbb{P}\big(w_{M_\cdot^{G,N}}(\gamma) > \epsilon\big) &= \mathbb{P}\Big(\sup_{|t-s|\le\gamma} \big|M_t^{G,N} - M_s^{G,N}\big| > \epsilon\Big) \\
&\le \mathbb{P}\Big(\sup_i \sup_{t\in(t_{i-1},t_i]} \big|M_t^{G,N} - M_{t_{i-1}}^{G,N}\big| > \epsilon/3\Big) \\
&\le \sum_i \mathbb{P}\Big(\sup_{t\in(t_{i-1},t_i]} \big|M_t^{G,N} - M_{t_{i-1}}^{G,N}\big| > \epsilon/3\Big) \\
&\le \frac{36}{\epsilon^2}\sum_i \mathbb{E}\Big[\big|M_{t_i}^{G,N} - M_{t_{i-1}}^{G,N}\big|^2\Big] \qquad \text{(d)}
\end{aligned}$$

Define $N_t^{G,N}$ by

$$N_t^{G,N} = (M_t^{G,N})^2 - N^2\int_0^t \Big((L_N F^2)(s,\eta_{N^2s}^N) - F(s,\eta_{N^2s}^N)L_N F(s,\eta_{N^2s}^N)\Big)ds$$

By Theorem 1.32, $(N_t^{G,N})$ is a martingale. A straightforward computation shows that

$$(L_N F^2)(s,\eta) - F(s,\eta)L_N F(s,\eta) = N^{-2}\sum_x\sum_{|z|=1}\eta(x)(1-\eta(x+z))\big(G((x+z)/N) - G(x/N)\big)^2$$

Fix $t, s \in [0,T]$ with $s < t$.

$$\begin{aligned}
\mathbb{E}\Big[\big|M_t^{G,N} - M_s^{G,N}\big|^2\Big] &= \mathbb{E}\Big[(M_t^{G,N})^2 - (M_s^{G,N})^2\Big] \\
&= \mathbb{E}\Big[\int_s^t \sum_x\sum_{|z|=1}\eta_{N^2u}^N(x)(1-\eta_{N^2u}^N(x+z))\big(G((x+z)/N) - G(x/N)\big)^2 du\Big] \\
&\le \frac{\|G'\|^2}{N^2}\mathbb{E}\Big[\int_s^t \sum_x\sum_{|z|=1}\eta_{N^2u}^N(x)(1-\eta_{N^2u}^N(x+z))du\Big] \\
&\le 2\frac{\|G'\|^2}{N^2}\mathbb{E}\Big[\int_s^t \sum_x \eta_{N^2u}^N(x)du\Big] \le 2\frac{\|G'\|^2}{N}|t-s|
\end{aligned}$$

Applying the last inequality with $s = t_{i-1}$ and $t = t_i$, (d) becomes

$$\begin{aligned}
\mathbb{P}\big(w_{M_\cdot^{G,N}}(\gamma) > \epsilon\big) &\le \frac{72}{N\epsilon^2}\|G'\|^2\sum_i |t_i - t_{i-1}| \\
&= \frac{72\,T}{N\epsilon^2}\|G'\|^2 \xrightarrow[N\to\infty]{} 0
\end{aligned}$$

which proves that Q^N is relatively compact.

Second step: uniqueness of limit points. Now that we know that the sequence Q^N is (weakly) relatively compact, it remains to characterise all limit points of Q^N. Let Q be a limit point of Q^N. There exists a sequence (N_i) such that $Q^{N_i} \Rightarrow Q$. We start by proving that Q is concentrated on absolutely continuous measures with respect to the Lebesgue measure. Let $(g_k)_{k\ge1}$ be a dense family in $\mathcal{C}(\mathbb{T})$ and define

$$\begin{aligned}
\phi_k\colon D\big([0,T],\mathcal{M}_+(\mathbb{T})\big) &\longrightarrow \mathbb{R} \\
(\pi_t)_t &\longmapsto \sup_t |\langle \pi_t, g_k\rangle|
\end{aligned}$$

Notice that ϕ_k is continuous: let (μ_n) be a sequence in $D\big([0,T],\mathcal{M}_+(\mathbb{T})\big)$ converging to μ in the Skorokhod topology. There exists a sequence (λ_n) of strictly increasing continuous mappings of $[0,T]$ onto itself such that $\|\lambda_n - id\| \to 0$ and

$$\sup_t \delta\big(\mu^n \circ \lambda_n(t), \mu(t)\big) \to 0$$

Fix $\epsilon > 0$, $\exists n_0$, $\forall n \geq n_0$, $\forall t$, $\delta(\mu^n \circ \lambda_n(t), \mu(t)) \leq \epsilon/2^k$. By definition of δ, we have

$$\frac{1}{2^k} \min\Big\{1, \big|\langle \mu^n \circ \lambda_n(t), g_k\rangle - \langle \mu(t), g_k\rangle\big|\Big\} \leq \frac{\epsilon}{2^k}$$

and

$$\big|\langle \mu^n \circ \lambda_n(t), g_k\rangle - \langle \mu(t), g_k\rangle\big| \leq \epsilon$$

For $n \geq n_0$,

$$\begin{aligned}\big|\langle \mu^n \circ \lambda_n(t), g_k\rangle\big| &\leq \big|\langle \mu^n \circ \lambda_n(t), g_k\rangle - \langle \mu(t), g_k\rangle\big| + \big|\langle \mu(t), g_k\rangle\big| \\ &\leq \epsilon + \sup_t \big|\langle \mu(t), g_k\rangle\big|\end{aligned}$$

Taking the supremum over $t \in [0,T]$ in the last inequality, we have

$$\sup_t \big|\langle \mu^n(t), g_k\rangle\big| = \sup_t \big|\langle \mu^n \circ \lambda_n(t), g_k\rangle\big| \leq \epsilon + \sup_t \big|\langle \mu(t), g_k\rangle\big|$$

Similarly,

$$\sup_t \big|\langle \mu(t), g_k\rangle\big| \leq \epsilon + \sup_t \big|\langle \mu^n(t), g_k\rangle\big|$$

which proves that ϕ_k is continuous. By the continuous mapping theorem, $Q^{N_i}\phi_k^{-1} \Rightarrow Q\phi_k^{-1}$. Notice that

$$\begin{aligned}Q^N\phi_k^{-1}\Big[0, N^{-1}\sum_x |g_k(x/N)|\Big] &= Q^N\Big(\pi,\ \sup_t |\langle \pi_t, g_k\rangle| \leq N^{-1}\sum_x |g_k(x/N)|\Big) \\ &= \mathbb{P}\Big(\sup_t |\langle \pi^N_{N^2 t}, g_k\rangle| \leq N^{-1}\sum_x |g_k(x/N)|\Big) = 1\end{aligned}$$

since there is at most one particle per site. Since g_k is continuous,

$$\frac{1}{N}\sum_{x\in\mathbb{T}_N} |g_k(x/N)| \to \int_{\mathbb{T}} |g_k(u)|du$$

By lemma A.7, we have

$$Q\Big(\pi,\ \sup_t |\langle \pi_t, g_k\rangle| \leq \int_{\mathbb{T}} |g_k|\Big) = Q\phi_k^{-1}\Big[0, \int_{\mathbb{T}} |g_k|\Big] = 1$$

for all $k \geq 1$. Therefore

$$Q\Big(\pi,\ \forall k \geq 1,\ \sup_t |\langle \pi_t, g_k\rangle| \leq \int_{\mathbb{T}} |g_k|\Big) = 1$$

Using the density of (g_k) in $\mathcal{C}(\mathbb{T})$,

$$Q\Big(\pi,\ \forall G \in \mathcal{C}(\mathbb{T}),\ \sup_t |\langle \pi_t, G\rangle| \leq \int_{\mathbb{T}} |G|\Big) = 1$$

But any measure μ on $\mathbb{T}$ satisfying

$$\forall G \in \mathcal{C}(\mathbb{T}),\ |\langle \mu, G\rangle| \leq \int_{\mathbb{T}} |G|$$

is absolutely continuous with respect to the Lebesgue measure. This proves that Q is concentrated on absolutely continuous trajectories with respect to the Lebesgue measure:

$$Q\Big(\pi,\ \forall t,\ \pi_t(du) = \rho(t,u)du\Big) = 1$$

Moreover, Q is concentrated on trajectories that at time 0 are equal to $\rho_0(u)du$. To see this, fix $G \in \mathcal{C}(\mathbb{T})$ and $\epsilon > 0$. By Portmanteau theorem,

$$\begin{aligned}
&Q\Big(\pi \in D\big([0,T], \mathcal{M}_+(\mathbb{T}_N)\big),\ \Big|\langle \pi_0, G\rangle - \langle \rho_0, G\rangle\Big| > \epsilon\Big) \\
&\leq \liminf_i Q^{N_i}\Big(\pi \in D\big([0,T], \mathcal{M}_+(\mathbb{T}_N)\big),\ \Big|\langle \pi_0, G\rangle - \langle \rho_0, G\rangle\Big| > \epsilon\Big) \\
&= \liminf_i \mathbb{P}\Big(\Big|\langle \pi_0^{N_i}, G\rangle - \langle \rho_0, G\rangle\Big| > \epsilon\Big) \\
&= \liminf_i \mu^{N_i}\sigma^{-1}\Big(\pi \in \mathcal{M}_+(\mathbb{T}),\ \Big|\langle \pi, G\rangle - \langle \rho_0, G\rangle\Big| > \epsilon\Big) = 0
\end{aligned}$$

Letting $\epsilon \to 0$, we get

$$Q\Big(\pi,\ \langle \pi_0, G\rangle = \langle \rho_0, G\rangle\Big) = 1$$

for all $G \in \mathcal{C}(\mathbb{T})$. Using the separability of $\mathcal{C}(\mathbb{T})$, this becomes

$$Q\Big(\pi,\ \forall G \in \mathcal{C}(\mathbb{T}),\ \langle \pi_0, G\rangle = \langle \rho_0, G\rangle\Big) = 1$$

which proves the result. Finally, we prove that Q is concentrated on trajectories that are weak solutions for the heat equation. Fix $G \colon \mathbb{T} \to \mathbb{R}$ of class $\mathcal{C}^3$ and $\epsilon > 0$. Apply Theorem 1.32 with the Markov chain $(\eta^N_{N^2 t})$ and the function

$$\begin{aligned}
F \colon \mathbb{R}_+ \times \{0,1\}^{\mathbb{T}_N} &\longrightarrow \mathbb{R} \\
(t, \eta) &\longmapsto \langle \sigma(\eta), G\rangle
\end{aligned}$$

Then the processes defined by

$$\begin{aligned}
M_t^{G,N} &= \langle \pi^N_{N^2 t}, G\rangle - \langle \pi_0, G\rangle - \int_0^t N^2 L_N F(s, \eta^N_{N^2 s}) ds \\
N_t^{G,N} &= (M_t^{G,N})^2 - N^2 \int_0^t \Big((L_N F^2)(s, \eta^N_{N^2 s}) - F(s, \eta^N_{N^2 s}) L_N F(s, \eta^N_{N^2 s})\Big) ds
\end{aligned}$$

are both right-continuous martingales. In particular, by Doob's maximal inequality,

$$\mathbb{P}\Big(\sup_t \big|M_t^{G,N}\big| > \epsilon\Big) \leq \frac{4}{\epsilon^2}\mathbb{E}\Big[(M_T^{G,N})^2\Big]$$

But since $(N_t^{G,N})$ is a martingale, $\mathbb{E}[N_T^{G,N}] = \mathbb{E}[N_0^{G,N}] = 0$ and

$$
\begin{aligned}
\mathbb{E}\Big[(M_T^{G,N})^2\Big] &= N^2 \mathbb{E} \int_0^T \Big((L_N F^2)(s, \eta^N_{N^2 s}) - F(s, \eta^N_{N^2 s}) L_N F(s, \eta^N_{N^2 s})\Big) ds \\
&= \mathbb{E} \int_0^T \sum_x \sum_{|z|=1} \eta^N_{N^2 s}(x)(1 - \eta^N_{N^2 s}(x+z))\big(G((x+z)/N) - G(x/N)\big)^2 ds \\
&\le N^{-2} \sup_u |G'(u)|^2 \, \mathbb{E} \int_0^T \sum_x \sum_{|z|=1} \eta^N_{N^2 s}(x)(1 - \eta^N_{N^2 s}(x+z)) ds \\
&\le 2\, T N^{-1} \sup_u |G'(u)|^2 \xrightarrow[N\to\infty]{} 0
\end{aligned}
$$

So

$$
\lim_N \mathbb{P}\bigg(\sup_t \big|M_t^{G,N}\big| > \epsilon\bigg) = 0 \tag{e}
$$

A similar computation to the one we made in the first step shows that

$$
M_t^{G,N} = \langle \pi^N_{N^2 t}, G\rangle - \langle \pi_0, G\rangle - (1/2) \int_0^t \langle \pi^N_{N^2 s}, \Delta_N G\rangle ds
$$

and we can rewrite (e) as

$$
\begin{aligned}
&\lim_N Q^N\bigg(\pi,\ \sup_t \Big|\langle \pi_t, G\rangle - \langle \pi_0, G\rangle - (1/2)\int_0^t \langle \pi_s, \Delta_N G\rangle ds\Big| > \epsilon\bigg) \\
&= \lim_N \mathbb{P}\bigg(\sup_t \Big|\langle \pi^N_{N^2 t}, G\rangle - \langle \pi_0, G\rangle - (1/2)\int_0^t \langle \pi^N_{N^2 s}, \Delta_N G\rangle ds\Big| > \epsilon\bigg) = 0
\end{aligned} \tag{f}
$$

On the other hand, for $\pi \in D\big([0,T], \mathcal{M}_+(\mathbb{T})\big)$ and $t \in [0,T]$,

$$
\begin{aligned}
\Big|\int_0^t \langle \pi_s, \Delta_N G - \Delta G\rangle ds\Big| &\le \int_0^t \sup_{u\in\mathbb{T}} \big|\Delta_N G(u) - \Delta G(u)\big| ds \\
&\le T \sup_u \big|\Delta_N G(u) - \Delta G(u)\big| \\
&\le \frac{1}{3} T N^{-1} \big\|G^{(3)}\big\| \xrightarrow[N\to\infty]{} 0
\end{aligned}
$$

using Taylor's theorem. In particular,

$$
\lim_N Q^N\bigg(\pi,\ \sup_t \Big|(1/2)\int_0^t \langle \pi_s, \Delta G - \Delta_N G\rangle ds\Big| > \epsilon/2\bigg) = 0 \tag{g}
$$

Now, since the application

$$
\begin{aligned}
D\big([0,T], \mathcal{M}_+(\mathbb{T})\big) &\longrightarrow \mathbb{R} \\
(\pi_t) &\longmapsto \sup_t \Big|\langle \pi_t, G_t\rangle - \langle \pi_0, G_0\rangle - \int_0^t \langle \pi_s, \partial_s G_s + (1/2)\Delta G_s\rangle ds\Big|
\end{aligned}
$$

is continuous, $\{\pi,\ \sup_t \left|\langle \pi_t, G\rangle - \langle \pi_0, G\rangle - (1/2)\int_0^t \langle \pi_s, \Delta G_s\rangle ds\right| > \epsilon\}$ is an open subset of $D\big([0,T],\mathcal{M}_+(\mathbb{T})\big)$, and by Portmanteau theorem,

$$
\begin{aligned}
&Q\bigg(\pi,\ \sup_t \Big|\langle \pi_t, G\rangle - \langle \pi_0, G\rangle - (1/2)\int_0^t \langle \pi_s, \Delta G\rangle ds\Big| > \epsilon\bigg)\\
&\le \liminf_i Q^{N_i}\bigg(\pi,\ \sup_t \Big|\langle \pi_t, G\rangle - \langle \pi_0, G\rangle - (1/2)\int_0^t \langle \pi_s, \Delta G\rangle ds\Big| > \epsilon\bigg)\\
&\le \liminf_i \bigg\{Q^{N_i}\bigg(\pi,\ \sup_t \Big|\langle \pi_t, G\rangle - \langle \pi_0, G\rangle - (1/2)\int_0^t \langle \pi_s, \Delta_N G\rangle ds\Big| > \epsilon/2\bigg)\\
&+ Q^{N_i}\bigg(\pi,\ \sup_t \Big|(1/2)\int_0^t \langle \pi_s, \Delta G - \Delta_N G\rangle ds\Big| > \epsilon/2\bigg)\bigg\} = 0
\end{aligned}
$$

combining (f) and (g). We just proved that for all $G \in \mathcal{C}^3(\mathbb{T})$,

$$
Q\bigg(\pi,\ \langle \pi_t, G\rangle - \langle \pi_0, G\rangle - (1/2)\int_0^t \langle \pi_s, \Delta G\rangle ds = 0\bigg) = 1
$$

By separability and density of $\mathcal{C}^3(\mathbb{T})$ in $\mathcal{C}^2(\mathbb{T})$, Q is concentrated on paths $\{\pi_t,\ 0 \le t \le T\}$ such that for all $G \in \mathcal{C}^2$,

$$
\langle \pi_t, G\rangle - \langle \pi_0, G\rangle - (1/2)\int_0^t \langle \pi_s, \Delta G\rangle ds = 0 \tag{h}
$$

The previous results show that every limit point is concentrated on absolutely continuous trajectories whose density is a weak solution in the sense of (h) for the heat equation and whose density at time 0 is ρ_0. However, in order to prove a uniqueness result of weak solutions for the heat equation, we need to prove the above relation for time dependent functions G. The proof is quite similar to the one we just gave, so we shall skip it. In conclusion, all limit points are concentrated on absolutely continuous trajectories that are weak solutions of the heat equation in the following sense

$$
\langle \pi_t, G_t\rangle - \langle \pi_0, G_0\rangle - \int_0^t \langle \pi_s, \partial_s G_s + (1/2)\Delta G\rangle ds = 0 \tag{i}
$$

for all $G\colon [0,T]\times\mathbb{T} \to \mathbb{R}$ of class $\mathcal{C}^{1,2}$, and whose density at time 0 is ρ_0. It turns out that there exists only one weak solution to equation (i) which therefore coincides with the strong solution. For a proof of this result, we cite [**8**]. This means that every limit point is concentrated on the solution of the heat equation. In other words, all converging subsequences of the sequence Q^N converge to the same limit, which is the Dirac measure concentrated on the solution of the heat equation Q^*.

Third step: convergence in probability at fixed time. In general it is false that the application from $D\big([0,T],\mathcal{M}_+(\mathbb{T})\big)$ to $\mathcal{M}_+(\mathbb{T})$ obtained by taking the value at time $0 < t < T$ of the process is continuous (see Theorem B.7). Therefore to prove convergence at fixed time, we will use the general form of the mapping theorem (see Theorem A.2). Fix $0 < t < T$ and denote by p_t the projection at time t. Let D_t be the set of discontinuities of p_t. Adapting Theorem B.7 in the case of a general complete separable metric space S instead of $\mathbb{R}$ (here we take $S = \mathcal{M}_+(\mathbb{T})$), we have that D_t consists of those $\mu \in D\big([0,T],\mathcal{M}_+(\mathbb{T})\big)$ which are discontinuous at t. But we know that the limiting probability measure of Q^N is Q^*, the dirac measure concentrated on the solution of the heat equation. We also know that the solution of the heat equation is continuous and therefore it does not lie in D_t. In other words D_t has measure 0 with respect to Q^*, and by the mapping theorem

$$Q^N p_t^{-1} \Rightarrow Q^* p_t^{-1}$$

Thus $\pi^N_{N^2 t}$ converges in distribution to the deterministic measure $\rho(t,u)du$. Since convergence in distribution to a deterministic variable implies convergence in probability, the theorem is proved. $\square$

A Weak convergence in metric spaces

Let (S, ρ) be a complete separable metric space, and denote by $\mathcal{S}$ the Borel σ-algebra. Let P_n and P be probability measures on $(S, \mathcal{S})$. We say that P_n converges weakly to P and write $P_n \Rightarrow P$ if for every bounded, continuous real function f on S,

$$\int_S f dP_n \xrightarrow[n\to\infty]{} \int_S f dP$$

Here is a simple condition for weak convergence:

Theorem A.1. *A necessary and sufficient condition for $P_n \Rightarrow P$ is that each subsequence $(P_{\phi(n)})$ contain a further subsequence $(P_{\phi\circ\psi(n)})$ converging weakly to P*

Proof. The necessity is easy. As for sufficiency, if $P_n \not\Rightarrow P$, then $\int_S f dP_n \nrightarrow \int_S f dP$ for some bounded, continuous f. But then, for some positive ϵ and some subsequence $(P_{\phi(n)})$,

$$\Big| \int_S f dP_{\phi(n)} - \int_S f dP \Big| > \epsilon$$

for all n, and no further subsequence can converge weakly to P. $\square$

Let (S', ρ') be a metric space and let P_n and P be probability measures on $(S, \mathcal{S})$. One can easily check that if h is a continuous mapping from S to S', then $P_n \Rightarrow P$ implies $P_n h^{-1} \Rightarrow Ph^{-1}$ where Ph^{-1} is the pushforward measure. However, the continuity assumption can be weakened. Assume only that $h \colon S \to S'$ is measurable with respect to the Borel σ-algebras on S and S', and let D_h be the set of its discontinuities.

Theorem A.2 (Mapping theorem)**.** *If $P_n \Rightarrow P$ and $P(D_h) = 0$, then $P_n h^{-1} \Rightarrow Ph^{-1}$.*

The following notion of tightness plays a fundamental role in the theory of weak convergence.

Definition A.3. *Let Π be a family of probability measures on $(S, \mathcal{S})$. We say that Π is tight if*

$$\forall \epsilon > 0,\ \exists K \subset S \text{ compact}, \forall P \in \Pi,\ P(K) > 1 - \epsilon$$

Theorem A.4. *Each probability measure on $(S, \mathcal{S})$ is tight.*

Proof. Let P be a probability measure on $(S, \mathcal{S})$, and fix $\epsilon > 0$. Denote by $(x_n)_{n\geq 1} \subset S$ a dense subset. For each $k \geq 1$,

$$S = \bigcup_{n=1}^{\infty} B(x_n, 1/k)$$

In particular, there exists N_k such that

$$P\Big(\bigcup_{n=1}^{N_k} B(x_n, 1/k) \Big) > 1 - \frac{\epsilon}{2^k}$$

Let K be the closure of $\bigcap_{k\geq 1}\bigcup_{n\leq N_k} B(x_n, 1/k)$. Then

$$\begin{aligned}
P(K) &\geq P\bigg(\bigcap_{k\geq 1}\bigcup_{n\leq N_k} B(x_n,1/k)\bigg)\\
&= 1 - P\bigg(\bigcup_{k\geq 1}\bigcap_{n\leq N_k} B(x_n,1/k)^c\bigg)\\
&\geq 1 - \sum_{k\geq 1} P\bigg(\bigcap_{n\leq N_k} B(x_n,1/k)^c\bigg)\\
&> 1-\epsilon
\end{aligned}$$

It remains to prove that K is compact. The set $\bigcap_{k\geq 1}\bigcup_{n\leq N_k} B(x_n, 1/k)$ is precompact. So, by the completeness hypothesis, its closure K is compact. $\square$

Let Π a family of probability measures on $(S, \mathcal{S})$. We call Π (weakly) relatively compact if every sequence of elements of Π contains a weakly convergent subsequence; that is if for every sequence (P_n) in Π there exist a subsequence $(P_{\phi(n)})$ and a probability measure Q on $(S, \mathcal{S})$ such that $P_{\phi(n)} \Rightarrow Q$. The following theorem relates tightness of probability measures to relative compactness.

Theorem A.5 (Prokhorov's theorem)**.** *A family Π of probability measures on $(S, \mathcal{S})$ is tight if and only if it is relatively compact.*

Although the converse puts things in perspective, the direct half is what is essential to the applications. The following corollary is useful for proving weak convergence.

Corollary A.6. *If (P_n) is tight, and if each subsequence that converges weakly at all in fact converges weakly to P, then the entire sequence converges weakly to P.*

We conclude this section with a useful lemma.

Lemma A.7. *Let (P_n) be a sequence of probability measures on $\mathbb{R}$ that converges weakly to P, and let $(t_n) \in \mathbb{R}_+^{\mathbb{N}}$ be a sequence converging to $t \geq 0$. Suppose that $P_n([0, t_n]) = 1$ for all n. Then*

$$P([0,t]) = 1$$

Proof. Fix $\epsilon > 0$. There exists $n_0 \in \mathbb{N}$ such that $t_n \leq t + \epsilon$ for $n \geq n_0$. Now

$$\begin{aligned}
P([0,t+\epsilon]) &\geq \limsup_n P_n([0,t+\epsilon])\\
&\geq \limsup_n P_n([0,t_n]) = 1
\end{aligned}$$

using Portmanteau theorem and the monotonicity of probability measures. So for all $\epsilon > 0$, $P([0, t+\epsilon]) = 1$. Now taking $\epsilon = 1/k$ and letting $k \to \infty$, we have $P([0, t]) = 1$ by continuity of P from above. $\square$

B The space D

Let $C = C[0,1]$ be the space of continuous real functions, and let $D = D[0,1]$ be the space of real functions x on $[0,1]$ that are right-continuous and have left-hand limits (cadlag). With very little change, the theory can be extended to functions on $[0,1]$ taking values in separable and complete metric spaces other than $\mathbb{R}$. For simplicity reasons, we develop the theory only in the case of $\mathbb{R}$ and along the way we indicate the changes needed when working with a complete separable metric space (S, ρ) with Borel σ-algebra $\mathcal{S}$.

One can show that a function $x \in D$ has at most countably many discontinuities, that it is bounded and Borel measurable. For $x \in D$ and $T \subset [0,1]$, put

$$w_x(T) = w(x,T) = \sup_{s,t \in T} |x(s) - x(t)|$$

The modulus of continuity of x is defined by

$$w_x(\delta) = w(x,\delta) = \sup_{|s-t| \leq \delta} |x(s) - x(t)|$$

for $\delta \in (0,1]$. It can also be written as

$$w_x(\delta) = \sup_{0 \leq t \leq 1-\delta} w_x[t, t+\delta]$$

Call a set $\{t_i\}$ δ-sparse if it satisfies $0 = t_0 < t_1 < \cdots < t_v = 1$ and $\min_i (t_i - t_{i-1}) > \delta$. Now define a modified modulus of continuity more suitable for the space D by

$$w'_x(\delta) = w'(x,\delta) = \inf_{\{t_i\}} \max_i w_x[t_{i-1}, t_i)$$

for $\delta \in (0,1)$, where the infimum extends over all δ-sparse sets. These definitions make sense even if x does not lie in D, but one can show that $w'_x(\delta) \to 0$ when $\delta \to 0$ if and only if $x \in D$. Similarly, $w_x(\delta) \to 0$ when $\delta \to 0$ if and only if $x \in C$.

The Skorokhod topology

Two functions x and y are near one another in the uniform topology used for C if the graph of $x(t)$ can be carried onto the graph of $y(t)$ by a uniformly small perturbation of the ordinates, with the abscissas kept fixed. In D, we want to allow also a uniformly small deformation of the time scale. This means that a small time change does not modify too much the proximity of two functions. The Skorokhod topology embodies this idea.

Let Λ denote the class of strictly increasing, continuous mappings of $[0,1]$ onto itself. Notice that if $\lambda \in \Lambda$, then $\lambda(0) = 0$ and $\lambda(1) = 1$. Define a distance d on D by setting

$$d(x,y) = \inf_{\lambda \in \Lambda} \min\{\|\lambda - id\|, \|x - y \circ \lambda\|\}$$

for $x, y \in D$, where id is the identity map on $[0,1]$. In this definition, λ represents the uniformly small deformation of the time scale.

A sequence x_n in D converges to x in the Skorokhod topology if and only if there exists a sequence λ_n in Λ such that $\|\lambda_n - id\| \to 0$ and $\|x_n \circ \lambda_n - x\| \to 0$. In particular, uniform convergence implies convergence in the Skorokhod topology. On the other hand, since

$$|x_n(t) - x(t)| \leq |x_n(t) - x \circ \lambda_n(t)| + |x \circ \lambda_n(t) - x(t)|$$

convergence in the Skorokhod topology implies that $x_n(t) \to x(t)$ holds for continuity points t of x. Moreover, if x is continuous on $[0, 1]$, then convergence in the Skorokhod topology implies uniform convergence. Therefore, the Skorokhod topology restricted to C coincides with the uniform topology.

One can show that D is separable under the metric d, but not complete (Consider the sequence $x_n = \mathbb{1}_{[0,2^{-n})}$). We can overcome this issue by defining another metric d° that is equivalent to d (meaning they induce the same topology) but under which D is complete. Start by setting

$$\|\lambda\|^\circ = \sup_{s<t} \left| \log \frac{\lambda t - \lambda s}{t - s} \right|$$

for $\lambda \in \Lambda$. Now for $x, y \in D$, let

$$d^\circ(x, y) = \inf_{\lambda \in \Lambda} \min\{\|\lambda\|^\circ, \|x - y \circ \lambda\|\}$$

It can be shown that this defines a metric on D. Moreover, we have the following theorems.

Theorem B.1. *The metrics d and d° are (topologically) equivalent.*

Theorem B.2. *The space D is separable under d and d° and is complete under d°.*

We now turn our attention to the characterisation of compact sets in D. Using the modulus $w'_x(\delta)$ we can prove an analogue of the Arzelà-Ascoli theorem:

Theorem B.3. *Let $A \subset D$. Then A is relatively compact in the Skorokhod topology if and only if the following conditions hold:*

$$\sup_{x \in A} \|x\| < \infty$$
$$\lim_{\delta \to 0} \sup_{x \in A} w'_x(\delta) = 0$$

Remark B.4. *For the general space S, write $\sup_t \rho(x(t), y(t))$ in place of $\|x - y\|$ in all the definitions and arguments. The first condition in the above theorem should be replaced by the assumption that $\{x(t),\ x \in A,\ t \in [0, 1]\}$ has compact closure in S. The rest needs no change.*

Weak convergence in D

Theorem B.5. *The sequence (P_n) is tight if and only if the following conditions hold:*

$$\lim_{a \to \infty} \limsup_{n} P_n\big(x,\ \|x\| \geq a\big) = 0 \tag{1}$$
$$\forall \epsilon > 0,\ \lim_{\delta \to 0} \limsup_{n} P_n\big(x,\ w'_x(\delta) \geq \epsilon\big) = 0 \tag{2}$$

Remark B.6. *We have the following inequality*

$$w'_x(\delta) \le w_x(2\delta)$$

for $\delta < 1/2$. Consequently, if instead of condition (B.6) of the previous theorem, we prove:

$$\forall \epsilon > 0,\ \lim_{\delta \to 0} \limsup_n P_n\big(x,\ w_x(\delta) \ge \epsilon\big) = 0 \tag{3}$$

then (P_n) is tight.

Finally we take a look at the projections from D to $\mathbb{R}^k$ which will be useful in the proof of the hydrodynamic limit. For $0 \le t_1 < \cdots < t_k \le 1$ define the natural projection $p_{t_1 \cdots t_k}$ as usual:

$$\begin{aligned} p_{t_1 \cdots t_k} : D &\longrightarrow \mathbb{R}^k \\ x &\longmapsto \big(x(t_1), \ldots, x(t_k)\big) \end{aligned}$$

Since each function in Λ fixes 0 and 1, p_0 and p_1 are continuous in the Skorokhod topology. Suppose that $0 < t < 1$. If points x_n converge to x in the Skorokhod topology and x is continuous at t, then as mentioned earlier $x_n(t) \to x(t)$. Suppose, on the other hand, that x is discontinuous at t. If λ_n is the element of Λ that carries t to $t - 1/n$ and is linear on $[0, t]$ and $[t, 1]$ and if $x_n(s) = x \circ \lambda_n(s)$, then x_n converges to x in the Skorokhod topology but $x_n(t)$ does not converge to $x(t)$. Therefore, for $0 < t < 1$, p_t is continuous at x if and only if x is continuous at t.

We must prove that $p_{t_1 \cdots t_k}$ is measurable with respect to the Borel σ-algebra $\mathcal{D}$. We need consider only a single time point t since a mapping into $\mathbb{R}^k$ is measurable if each component mapping is, and we may assume $t < 1$. Let

$$h_\epsilon(x) = \epsilon^{-1} \int_t^{t+\epsilon} x(s)ds$$

If $x_n \to x$ in the Skorokhod topology, then $x_n(s) \to x(s)$ for continuity points s of x as mentioned earlier. But we know that x has at most countably many discontinuities. Therefore $x_n(s) \to x(s)$ holds for points s outside a set of Lebesgue measure 0. Since on the other hand, the x_n are uniformly bounded, we have

$$h_\epsilon(x_n) \xrightarrow[n \to \infty]{} h_\epsilon(x)$$

using Lebesgue's dominated convergence theorem. Thus h_ϵ is continuous in the Skorokhod topology. By right-continuity, $h_{m^{-1}}(x) \to x(t) = p_t(x)$ for each $x \in D$ as $m \to \infty$. Therefore each p_t is measurable. In summary, we have the following theorem:

Theorem B.7. *(i) The projections p_0 and p_1 are continuous.*

(ii) For $0 < t < 1$, p_t is continuous at x if and only if x is continuous at t.

(iii) Each p_t is measurable $\mathcal{D}/\mathcal{R}$, and each $p_{t_1 \cdots t_k}$ is measurable $\mathcal{D}/\mathcal{R}^k$.

References

[1] J. Norris, *Markov Chains*, 1997

[2] T. Liggett, *Continuous Time Markov Processes*, 2010

[3] J.-F. Delmas, B. Jourdain, *Modèles aléatoires*, 2006

[4] T. Liggett, *Interacting Particle Systems*, 2005

[5] C. Kipnis, C. Landim, *Scaling Limits of Interacting Particle Systems*, 1999

[6] C. Bernardin, P. Gonçalves, *From Particle Systems to Partial Differential Equations*, 2012

[7] P. Billingsley, *Convergence of Probability Measures*, 1999

[8] H. Brézis, M. Crandall, *Uniqueness of solutions of the initial value problem for* $u_t - \Delta\phi(u) = 0$, J. Math. Pures Appl. 58, p. 153-163, 1979