Journal of Statistical Physics, Vol. 35, Nos. 5/6, 1984

Metastable Behavior of Stochastic Dynamics: A Pathwise Approach

Marzio Cassandro,[1] Antonio Galves,[2,5] Enzo Olivieri,[3] and Maria Eulália Vares[4]

Received October 19, 1983; revision received January 3, 1984

In this paper a new approach to metastability for stochastic dynamics is proposed. The basic idea is to study the statistics of each path, performing time averages along the evolution. Metastability would be characterized by the fact that the process of these time averages converges, under a suitable rescaling, to a measure valued Markov jump process. Here this convergence is shown for the Curie–Weiss mean field dynamics and also for a model with spatial structure: Harris contact process.

KEY WORDS: Metastability; stochastic dynamics; mean field theory; contact process.

1. INTRODUCTION

A large class of thermodynamical systems undergoing a phase transition exhibits the phenomenon of metastability. Let us consider, for example, a ferromagnetic system below the critical temperature. If we start from an equilibrium state when the external magnetic field is $h = 0^+$ (one can think to prepare this state by slowly switching off a positive magnetic field) and if we let it evolve, after having introduced a small negative magnetic field, then we observe that the initial situation—which is characterized by a

[1] Istituto di Fisica dell'Università di Roma, Roma, Italy.
[2] Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, Brazil.
[3] Istituto di Matematica dell'Università di Roma, Roma, Italy.
[4] Instituto de Matemática Pura e Applicade, IMPA, Rio de Janeiro, Brazil.
[5] Partially supported by CNPq, Grant No. 301301/79.

positive magnetization—persists for a long (macroscopic) time. In other words, the system instead of undergoing the right phase transition, remains for a long time in an apparently stationary situation until some external perturbation or some spontaneous large fluctuation will "nucleate" the new phase, starting an irreversible process leading the system to the true equilibrium phase, with negative magnetization.

In this paper we consider the metastability as a dynamical phenomenon that occurs in a very general situation. We can describe this phenomenon in the following way: we have a stochastic process with a unique stationary probability measure, but if the initial conditions are suitably chosen then the time to get to the asymptotic state becomes very large, and during this time the system behaves as if it were described by another stationary measure; finally, and abruptly, it goes to the true equilibrium.

One possible point of view on metastability is that of the "evolution of the ensembles." It consists in giving prescriptions to construct a state, that is, a probability measure on the configuration space, which should describe the statistical properties of systems in a metastable situation. Then, to study the dynamics of metastability one can consider the time evolution of the probability distribution on the configurations, starting with the "metastable state." In this framework one gives the following dynamical justification for the choice of a given probability μ as the correct one to describe metastability: the average values of the physical observables with respect to the probability distribution at time t are very slowly varying functions of t for suitable values of the thermodynamical parameters. Also, starting from an initial distribution different from μ, the time evolution of the average values is expected to exhibit a rapid variation toward the quasistationary (slowly varying) situation. (See Refs. 2, 10, and 11.)

In this paper we take another, different, point of view, which could be called the "pathwise approach." We consider single typical trajectories and study their statistics in the most natural way, i.e., we perform time averages and try to detect the behavior just described; namely, we try to see if the time averages over almost all trajectories are practically stationary for a very long period, until some large fluctuation will lead the system to another, completely different situation.

One important remark must be made: typically the time until the large fluctuation occurs has a very wide distribution. Thus if we look at the time evolution of the averages with respect to all trajectories, what we usually see is a smooth behavior, even though every single trajectory has a very sharp behavior. The reason why the averages over all trajectories approach the equilibrium very slowly is related to the fact that the time to the large fluctuation is very large.

Then, it is clear that by looking only at the time evolution of the probability distribution, it is very hard to distinguish between the kind of behavior of a single trajectory as described above and another, completely different behavior: a smooth, very slow evolution toward the stable situation. This last one has nothing to do with metastability. Nevertheless, we stress again that in the approach based on the evolution of the ensembles, it is difficult to catch this difference.

To get an intuitive picture of what happens in the class of models we study here, let us consider a diffusion of a particle in a random medium (Brownian noise) with a drift given by $-da(x)/dx$, where $a(x)$ is given in Fig. 1. (This is in fact a continuous version of the first model we consider.) For an initial condition $x < x_0$ the particle is driven to x_1, and under the action of the medium it starts oscillating around it. This is so until a large fluctuation lets it jump beyond the barrier at x_0. Then, the drift leads it to the absolute minimum x_2, where it will start oscillating again until a new fluctuation brings it back to the first region. The smaller is the diffusion coefficient, the larger the time spent before the first jump.

There are two crucial quantities: the exit time from the region of the local minimum T_a, and the "time of thermalization" τ, i.e., the typical time between two consecutive passages through x_1. For very small diffusion coefficient we have $\tau \ll T_a$, as it is well known[12] and as we shall explicitly show in our examples. The main feature of these processes is the relevance of the large fluctuations: the actual transition from one oscillating center to another is not due to oscillations that gradually get larger and larger, but to an abrupt motion against the drift.

To make this picture more precise we will describe these kinds of phenomena in some limit case, where the transition from one region to the other can be made particularly sharp. We shall show in our examples that letting $T_a \to +\infty$ together with the length of the intervals over which we take time averages, and by properly rescaling the time these measure-valued processes converge to a measure-valued Markov jump process. This also allows us to single out the two main regimes.

It is important to stress that the Markovian property of the limit process, i.e., the exponential distribution of its jump time, reveals the impredictability of the transition from one regime to the other. In fact, this impredictability together with the statistical stability of the system before and after the jump may be taken as the basic features of the metastability phenomenon.

What we have described in the last paragraph is in fact a new proposal to describe metastability in stochastic models. In this paper we verify this kind of behavior in two different models: the first is the Curie–Weiss model which, in spite of its inconsistency as a statistical mechanical model, is traditionally considered the first test for new ideas in this area. The second is the basic contact process of Harris (on $\mathbb{Z}$), that has an extra interesting point, which is its spatial structure.

2. METASTABLE BEHAVIOR OF THE CURIE-WEISS MODEL

Let us consider a system of N Ising spins $\sigma_i \in \{-1, +1\}$ described by the Hamiltonian

$$H_N(\boldsymbol{\sigma}_N) = -\frac{J}{N}\sum_{i<j}\sigma_i\sigma_j - h\sum_i \sigma_i - \frac{J}{2} \tag{2.1}$$

where $\boldsymbol{\sigma}_N = (\sigma_1, \ldots, \sigma_N)$. The Gibbs distribution for such a system is

$$P(\boldsymbol{\sigma}_N) = \frac{1}{Z_N} e^{-\beta H_N(\boldsymbol{\sigma}_N)} \tag{2.2}$$

where

$$Z_N = \sum_{\boldsymbol{\sigma}_N \in \{-1,+1\}^N} e^{-\beta H_N(\boldsymbol{\sigma}_N)} \tag{2.3}$$

and β is the inverse temperature. For simplicity we shall choose $J = 1$. If n is the number of positive spins in the configuration σ_N, we have

$$H_N(\boldsymbol{\sigma}_N) = U(n) = -\frac{(2n-N)^2}{2N} - h(2n-N) \tag{2.4}$$

Thus, under the Gibbs measure, the probability of having n positive spins, according to (2.2), is given by

$$P_N(n) = \frac{1}{Z_N}\binom{N}{n} e^{-\beta U(n)} \tag{2.5}$$

As in Griffiths, Weng and Langer[7] we consider a time evolution for our system in terms of a Markov chain in discrete time. The transition probabilities $\{Q_{m,n}\}$ will be given by

$$\begin{aligned} Q_{n,n+1} &= \frac{\alpha}{N}\exp\left\{-\frac{N}{2}\left[\bar{a}_N(n+1) - \bar{a}_N(n)\right]\right\} \\ Q_{n,n-1} &= \frac{\alpha}{N}\exp\left\{-\frac{N}{2}\left[\bar{a}_N(n-1) - \bar{a}_N(n)\right]\right\} \\ Q_{n,n} &= 1 - Q_{nn+1} - Q_{n,n-1} \end{aligned} \tag{2.6}$$

for $n = 0, \ldots, N$ except that $Q_{0,-1} = Q_{N,N+1} = 0$, where

$$\bar{a}_N(n) = \frac{1}{N}\left[\beta U(n) - \log\binom{N}{n}\right] \tag{2.7}$$

It is immediately seen that P_N [given by (2.5)] is the unique invariant probability for this chain.

Now, introducing the "magnetization" variable

$$x = \frac{1}{N} \sum_{i=1}^{N} \sigma_i = \frac{2n - N}{N} \tag{2.8}$$

we define, for $x \in [-1, 1]$, $a_N(x) = \bar{a}_N[(Nx + N)/2]$.

It is easy to check that

$$\lim_{N\to\infty} a_N(x) = a(x) = -\beta\left(\frac{x^2}{2} + hx\right) + \frac{1+x}{2}\log\left(\frac{1+x}{2}\right) + \left(\frac{1-x}{2}\right)\log\left(\frac{1-x}{2}\right) \tag{2.9}$$

and that

$$a_N(x) = a(x) + \frac{v(x)}{N} + O\left(\frac{1}{N^2}\right) \tag{2.10}$$

where v is a continuous function on $[-1, 1]$. The graph of the function $a(\cdot) = a(\beta, h, \cdot)$ for $\beta > 1$, with h positive and small is given in Figure 1, where x_1, x_0, x_2 are the solutions of the equation

$$x = \tanh[\beta(x + h)] \tag{2.11}$$

From now on we shall assume that $\beta > 1$ and $0 < h < [(\beta - 1)/\beta]^{1/2} + (1/\beta)\log[\beta^{1/2} - (\beta - 1)^{1/2}]$ so that we have the picture as given in Fig. 1.

From (2.6) and (2.10) it is clear that for large N, the time evolution of our system is that of a one-dimensional random walk between two reflecting barriers whose drift is given, roughly speaking, by $-(d/dx)a(x)$.

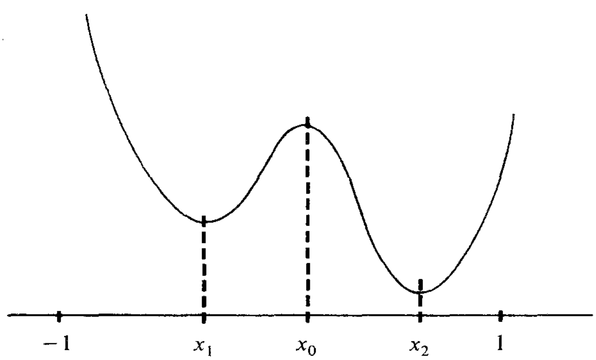

Fig. 1.

Before stating our results we give some definitions. If N is an integer $\geqslant 2$ let

$$\Omega_N = \left\{ -1, -1 + \frac{2}{N}, \ldots, 1 - \frac{2}{N}, 1 \right\}$$

Let $(\xi_N(t))_{t \in N}$ be a Markov chain (M.C.) with values on Ω_N, with transitions

$$P\left(\xi_N(t+1) = -1 + \frac{2m}{N} \,\middle|\, \xi_N(t) = -1 + \frac{2l}{N} \right) = Q_{l,m}$$

where the $Q_{l,m}$ are given in (2.6). Usually we put an upper index to denote the starting point, i.e., $(\xi_N^n(t))_{t \geqslant 0}$ denotes the previous M.C. with $\xi_N^n(0) = -1 + 2n/N$, $n = 0, \ldots, N$. For convenience we may take these processes defined on $(\Omega_N)^{\mathbb{N}}$, the space of all trajectories.

The unique stationary probability measure will be denoted by μ_N.

We denote by $i_0(N)$, $i_1(N)$, $i_2(N)$ and $i_*(N)$, respectively, the integers such that

$$-1 + \frac{2i_j(N)}{N} \leqslant x_j < -1 + \frac{2(i_j(N)+1)}{N} \qquad \text{for} \quad j = 0, 1, 2, \quad \text{and} \tag{2.12}$$

$$-1 + \frac{2i_*(N)}{N} \leqslant x_0 + \frac{1}{N^{1/4}} < -1 + \frac{2(i_*(N)+1)}{N} \tag{2.13}$$

where x_1, x_2, and x_0 are, respectively, the local minimum, the absolute minimum, and the local maximum of the function $a(\cdot)$. Also, let us call $(\bar{\xi}_N^n(t))_{t \geqslant 0}$ the M.C. on $\{-1, -1 + 2/N, \ldots, -1 + 2i_*(N)/N\}$ such that $\bar{\xi}_N^n(0) = -1 + 2n/N$ $(n \leqslant i_*(N))$, and transitions

$$P\left(\bar{\xi}_N^n(t+1) = -1 + \frac{2m}{N} \,\middle|\, \bar{\xi}_N^n(t) = -1 + \frac{2l}{N} \right) = \bar{Q}_{l,m} \tag{2.14}$$

where

$$\bar{Q}_{l,m} = Q_{l,m} \qquad \text{if} \quad l \leqslant i_*(N) - 1$$

$$\bar{Q}_{i_*(N), i_*(N)-1} = Q_{i_*(N), i_*(N)-1}$$

$$\bar{Q}_{i_*(N), i_*(N)} = 1 - Q_{i_*(N), i_*(N)-1}$$

In other words, the process $\bar{\xi}$ has the same transition probabilities as ξ to the left of $-1 + 2i_*(N)/N$ and we have introduced a reflecting barrier at $-1 + 2i_*(N)/N$. Obviously, all these processes may be constructed in $(\Omega_N)^{\mathbb{N}}$ in such a way that if we start at the left of $-1 + 2i_*(N)/N$

$$\bar{\xi}_N^n(t) = \xi_N^n(t) \qquad \forall t \leqslant T_{n, i_*(N)}^N \tag{2.15}$$

where

$$T_{i,j}^N = \inf\left\{ t \geqslant 0 \mid \xi_N^i(t) = -1 + \frac{2j}{N} \right\} \tag{2.16}$$

Let us call $\bar{\mu}_N$ the invariant probability measure for $\bar{\xi}$.

Notations: (1) T_a^N denotes $T_{i_1(N), i_*(N)}^N$.

(2) $\mathfrak{M}_1([-1, +1])$ denotes the space of probability measures on $[-1, +1]$ with w^* topology.

(3) If S is a metric space $D([0, +\infty), S)$ denotes the usual Skorohod space.[1]

(4) $\bar{T}_{i,j}^N$ is defined as in (2.16), but for the process $\bar{\xi}$.

The Results

Given integers $R \geqslant 1$ and $N \geqslant 1$, we define the following empirical process of time averages:

$$A_R^N(s, \cdot) = \frac{1}{R} \sum_{k=[s]+1}^{[s+R]} \delta_{\xi_N^{i_1(N)}(k)}(\cdot), \qquad s \geqslant 0 \tag{2.17}$$

where $\delta_x(\cdot)$ denotes the unit mass at x.

We can state the following:

Theorem 2.1. There exists a sequence of numbers $R_N \nearrow +\infty$ such that the processes $(A_{R_N}^N(sET_a^N, \cdot))_{s \geqslant 0}$ converge in law on $D([0, +\infty), \mathfrak{M}_1[-1, 1])$, as $N \to \infty$, to a jump process $(A(s, \cdot))_{s \geqslant 0}$ given by

$$A(s, \cdot) = \begin{cases} \delta_{x_1} & \text{if } s < T \\ \delta_{x_2} & \text{if } s \geqslant T \end{cases}$$

where T is exponentially distributed, with mean 1.

Theorem 2.1 will be a consequence of the next proposition together with Theorems 2.2 and 2.3.

Proposition 2.1. As $N \to +\infty$ the random variables $T_{i_1(N), i_*(N)}^N / ET_{i_1(N), i_*(N)}^N$ converge in distribution to an exponential distribution with mean 1.

Theorem 2.2. Letting $R_N = E(T_a^N)/N$, and $l_N^* = \max\{ l \in \mathbb{N} : lR_N \leqslant T_a^N \}$, and defining, for all $\epsilon > 0$, all interval $J = [a, b] \subseteq [-1, 1]$ such that $x_1 \in (a, b)$, the event Γ_N as

$$\Gamma_N = [l_N^* = 0] \cup \left[l_N^* \geqslant 1, \sup_{l < l_N^*} |A_{R_N}^N(lR_N, J) - 1| \leqslant \epsilon \right]$$

Then

$$P(\Gamma_N) \to 1 \qquad \text{as} \quad N \to +\infty$$

Remarks. (1) Proposition 2.1 implies, in particular, that $P[l_N^* = 0] \to 0$, so that $P[l_N^* \geqslant 1, \sup_{l<l_N^*} |A_{R_N}^N(lR_N, J) - 1| < \epsilon]$ converge to 1 as $N \to +\infty$ in Theorem 2.2.

(2) Defining $R'_N = r_N R_N$ with R_N as in Theorem 2.2, and $r_N \geqslant 1$ integer such that, as $N \to +\infty$: (a) $r_N \to +\infty$; (b) $R'_N / ET_a^N \to 0$ [and so $P(T_a^N < R'_N) \to 0$, by Proposition 2.1], we can say that for J as above;

$$P\left[T_N > R'_N, \sup_{\substack{l<l'_N \\ l \in \mathbb{N}}} |A_{R'_N}^N(lR_N, J) - 1| \leqslant \epsilon\right] \xrightarrow[N \to +\infty]{} 1$$

where $l'_N = \inf\{l \in \mathbb{N} : (l + r_N - 1)R_N < T_a^N < (l + r_N)R_N\}$. Thus, since $R_N / R'_N \to 0$ we get

$$P\left[T_N > R'_N, \sup_{s<[T_N^a - R'_N]} |A_{R'_N}^N(s, J) - 1| \leqslant \epsilon\right] \xrightarrow[N \to +\infty]{} 1$$

Clearly, such a sequence can be found. Actually, since Theorem 2.2 works equally well for $R_N = ET_a^N / N^2$ we may take $R'_N = ET_a^N / N$, by previous considerations (i.e., the same choice made in Theorem 2.2). Same remark applies to what we do in Theorem 2.3 and 3.3 and their applications to the proof of Theorems 2.1 and 3.1.

In the proof of the previous theorem we shall need the following:

Lemma 1. Let us consider the process ξ_N^n defined by (2.14). Then, for all $J = [a,b] \subseteq [-1,1]$ such that $-1 + [2i_1(N)]/N \in (a,b)$, for all l, for all $\epsilon > 0$, and $T \geqslant (4/\epsilon \vee \bar{\tau}_N/2)^4$, the following inequality holds:

$$P\left[\left|\frac{1}{T}\sum_{t=0}^{T} I_J(\xi_N^l(t)) - \bar{\mu}_N(J)\right| > \epsilon\right]$$

$$\leqslant \text{const}\left[\frac{E(\overline{T}_{l,i_1(N)}) + 4\bar{\tau}_N}{\epsilon T} + \frac{\bar{\sigma}_N}{\bar{\tau}_N^3 T^{1/2}} + \frac{\bar{\sigma}_N}{\bar{\tau}_N}\frac{1}{\epsilon^2 T^{1/4}}\right] \tag{2.18}$$

where $\bar{\mu}_N(J) = \bar{\mu}_N(J \cap \Omega^N)$, $\bar{\tau}_N = E(\overline{T}_{i_1(N),i_1(N)})$ is the mean recurrence time at $i_1(N)$, $\bar{\sigma}_N = E(\overline{T}_{i_1(N),i_1(N)})^2$ and $I_J(\cdot)$ is the characteristic function of the set J.

Remarks. (1) Applying the results of Lemma A1 to the inequality (2.18) we obtain

$$P\left[\left|\frac{1}{T}\sum_{t=0}^{T} I_J\left(\bar{\xi}_N^l(t)\right) - \bar{\mu}_N(J)\right| > \epsilon\right] \leqslant C_1 e^{C_2 N^{3/4}}\left(\frac{1}{\epsilon T} + \frac{1}{\sqrt{T}} + \frac{1}{\epsilon^2 T^{1/4}}\right) \tag{2.19}$$

(2) Sometimes we omit the lower index N, and we write $\bar{\mu}$, $\bar{\tau}$, $\bar{\sigma}$, etc . . . .

Proof of Lemma 1. Let us consider the trajectory $\bar{\xi}_N^l(t)$ for $0 \leqslant t \leqslant T$. Let n_T be the number of returns to $-1 + [2i_1(N)]/N$ and let ρ_i be the time of the ith return with $\rho_0 = T_{l,i_1(N)}^N$. For all trajectories such that $\rho_0 \leqslant T$ we can write

$$\sum_{t=0}^{T} I_J\left(\bar{\xi}_N^l(t)\right) = \sum_{t=0}^{\rho_0} I_J\left(\bar{\xi}_N^l(t)\right) + \sum_{i=1}^{n_T}\sum_{t=\rho_{i-1}+1}^{\rho_i} I_J\left(\bar{\xi}_N^l(t)\right) + \sum_{t=\rho_{n_T}+1}^{T} I_J\left(\bar{\xi}_N^l(t)\right) \tag{2.20}$$

Thus, we have this sum where n_T and the ρ_i's are random variables (if $k = 0$ $\sum_{i=1}^{k} = 0$ by convention), and we get

$$P\left[\left|\frac{1}{T}\sum_{t=1}^{T} I_J\left(\bar{\xi}_N^l(t)\right) - \bar{\mu}_N(J)\right| > \epsilon\right] \leqslant P\left[\frac{\rho_0}{T} > \frac{\epsilon}{4}\right] + P\left[\frac{T - \rho_{n_T}}{T} > \frac{\epsilon}{4}\right] + P\left[\left|\frac{1}{T}\sum_{i=1}^{n_T}\sum_{t=\rho_{i-1}+1}^{\rho_i} I_J\left(\bar{\xi}_N^l(t)\right) - \bar{\mu}_N(J)\right| > \epsilon/2\right] \tag{2.21}$$

The last term on the right-hand side of Eq. (2.21) is bounded by

$$P\left[\left|n_T - T/\bar{\tau}\right| > T^{3/4}\right] + \sum_{n=[T/\bar{\tau} - T^{3/4}]}^{[T/\bar{\tau} + T^{3/4}]} P\left[\left|\frac{1}{n\bar{\tau}}\sum_{i=1}^{n}\sum_{t=\rho_{i-1}+1}^{\rho_i} I_J\left(\bar{\xi}_N^l(t)\right) - \frac{T}{n\bar{\tau}}\,\bar{\mu}_N(J)\right| > \frac{\epsilon T}{2n\bar{\tau}}\,,\, n_T = n\right] \tag{2.22}$$

Then, since $n\bar{\tau} - T^{3/4} \leqslant T \leqslant n\bar{\tau} + T^{3/4}$, if $T > (4/\epsilon)^4$ we have that for all n in this sum, the corresponding summand is bounded above by

$$P\left[\left|\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{1}{\tau}\sum_{t=\rho_{i-1}+1}^{\rho_i} I_J\left(\bar{\xi}_N^l(t)\right) - \bar{\mu}_N(J)\right\}\right| > \epsilon/8\right] \tag{2.23}$$

where we have also relaxed the condition $n_T = n$, so that ρ_n is not constrained to be less or equal to T anymore. Since $[\bar{\xi}_N(t) = -1 + 2i_1(N)/N]$ is a recurrent event which is persistent we have

$$E(n_T) = \frac{T}{\bar{\tau}}\,\mathrm{Var}(n_T) = T\bar{\sigma}/\bar{\tau}^3$$

Moreover, as it is shown in Lemma A2:

$$E\sum_{t=1}^{T^N_{i_1(N),i_1(N)}} I_J\left(\bar{\xi}_N^{i_1(N)}\right) = \bar{\tau}\bar{\mu}_N(J) \tag{2.24}$$

Therefore the following inequalities hold:

$$P\left[\rho_0/T > \epsilon/4\right] \leqslant 4ET^N_{l,i_1(N)}/T\epsilon$$

$$P\left[(T-\rho n_T)/T > \epsilon/4\right] \leqslant 4\bar{\tau}/T\epsilon$$

$$P\left[|n_T - T/\bar{\tau}| > T^{3/4}\right] \leqslant \bar{\sigma}/\bar{\tau}^3 T^{1/2}$$

and

$$P\left[\left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\bar{\tau}}\sum_{t=\rho_{i-1}+1}^{\rho_i} I_J\left(\bar{\xi}_N^{i_1(N)}\right) - \bar{\mu}_N(J)\right| > \epsilon/8\right] \leqslant 64\bar{\sigma}/n\bar{\tau}^2\epsilon^2$$

Thus, for $T > (4/\epsilon \vee \bar{\tau}/2)^4$, we get

$$P\left[\left|\frac{1}{T}\sum_{t=0}^{T} I_J\left(\bar{\xi}_N^l(t)\right) - \bar{\mu}_N(J)\right| > \epsilon\right]$$
$$\leqslant C\left[\frac{ET^N_{l,i_1(N)}}{T\epsilon} + \frac{\bar{\tau}}{T\epsilon} + \frac{\bar{\sigma}}{\bar{\tau}^3 T^{1/2}} + \frac{\bar{\sigma}}{\bar{\tau}\epsilon^2}\frac{1}{T^{1/4}}\right] \tag{2.25}$$

for some constant C. Also, by Lemma A1 it is easily seen that

$$E\bar{T}^N_{l,i_1(N)} \leqslant C_1\exp\left\{N\left[a\left(-1+\frac{2i_0(N)}{N}\right) - a\left(-1+\frac{2i_*(N)}{N}\right)\right]\right\}$$
$$\leqslant C_1\exp\left(C_2N^{3/4}\right)$$

where C_1, C_2 are positive constants. Moreover, by Lemma A3 we get analogous estimates for $\bar{\tau}$ and $\bar{\sigma}$ so that we may also write

$$P\left[\left|\frac{1}{T}\sum_{t=0}^{T} I_J\left(\bar{\xi}_N^l(t)\right) - \bar{\mu}_N(J)\right| > \epsilon\right] \leqslant C_1 \exp\left(C_2 N^{3/4}\right)\left(\frac{1}{\epsilon T} + \frac{1}{\sqrt{T}} + \frac{1}{\epsilon^2 T^{1/4}}\right)$$

for some positive constants C_1, C_2, as claimed in (2.19). ■

Proof of Theorem 2.2. Given ϵ, J, N define the event

$$B_l = \left[\left|\frac{1}{R_N}\sum_{t=lR_N}^{(l+1)R_N} I_J\left(\xi_N^{i_l(N)}(t)\right) - 1\right| < \epsilon\right], \qquad l = 0, 1, \ldots \tag{2.26}$$

For any positive integer K we have

$$\begin{aligned} P(\Gamma_N) &= P[l_N^* = 0] + \sum_{L=1}^{\infty} P(l_N^* = L) P\left(\bigcap_{0 \leqslant l \leqslant L-1} B_l \mid l_N^* = L\right) \\ &\geqslant P[l_N^* = 0] + \sum_{L=1}^{K} P(l_N^* = L)\left(1 - L \sup_{l \leqslant L-1} P\left(\bar{B}_l^c \mid l_N^* = L\right)\right) \end{aligned}$$

where

$$\bar{B}_l^c = \left[\left|\frac{1}{R_N}\sum_{t=lR_N}^{(l+1)R_N} I_J\left(\bar{\xi}_N^{i_l(N)}(t)\right) - 1\right| \geqslant \epsilon\right]$$

and we have used (2.15). Thus, to prove the theorem we have to show that for the R_N given, there exists K_N such that the following conditions are satisfied:

(i) $$K_N R_N / ET_a^{(N)} \xrightarrow[N \to +\infty]{} +\infty$$

(ii) $$K_N^2 \sup_{0 \leqslant l \leqslant K_N} P\left(\bar{B}_l^c\right) \xrightarrow[N \to +\infty]{} 0$$

and this can be easily done, choosing $K_N = N^2$. In fact with this choice (i) is trivially satisfied and for N sufficiently large:

(a) $$|\bar{\mu}_N(J) - 1| < \epsilon/2 \qquad \text{if} \quad x_1 \in (a, b)$$

as we can easily see, and

(b) $$\begin{aligned} P&\left[\left|\frac{1}{R_N}\sum_{t=lR_N}^{(l+1)R_N} I_J\left(\xi_N^{i_l(N)}(t)\right) - \mu_N(J)\right| > \epsilon/2\right] \\ &\leqslant C_1 \exp\left(C_2 N^{3/4}\right) \cdot \frac{N}{\epsilon^2} \exp\left\{-\frac{N}{4}\left[a(x_0) - a(x_1)\right]\right\} \end{aligned}$$

where we have used Lemma 1 for $T = R_N$, and we have expressed ET_a^N by means of the leading term given by Lemma A1 [see (2.26)]. ■

In order to describe the behavior of our system after the first passage through the point $-1+2i^*(N)/N$ we notice that

$$P\left(T^N_{i^*(N),i_2(N)} < T_{i^*(N),i_0(N)}\right) \xrightarrow[N\to\infty]{} 1 \tag{2.27}$$

as a consequence of Eq. (A1′).

Then, using again Eq. (A1) and Chebychev's inequality, it is very easy to see that

$$P\left(T^N_{i^*(N),i_2(N)} > N^2 \mid T^N_{i^*(N),i_2(N)} < T_{i^*(N),i_0(N)}\right) \xrightarrow[N\to\infty]{} 0 \tag{2.28}$$

Now for a given $\bar{i}(N): i_0(N) < \bar{i}(N) < i_2(N)$, that will be chosen later, we introduce a new Markov chain $\tilde{\xi}^n_N(t)$ on $\{-1+2\bar{i}(N)/N, \ldots, 1-2/N, 1\}$ with $\tilde{\xi}^n_N(0) = -1+2n/N$ and transition probability given by

$$P\left(\tilde{\xi}^n_N(t+1) = -1+\frac{2m}{N} \,\middle|\, \tilde{\xi}^n = -1+\frac{2l}{N}\right) = \tilde{Q}_{lm}$$

with

$$\tilde{Q}_{lm} = Q_{lm} \qquad \text{if} \quad l \geqslant \bar{i}(N)+1$$

$$\tilde{Q}_{\bar{i}(N),\bar{i}(N)+1} = Q_{\bar{i}(N),\bar{i}(N)+1}$$

and

$$\tilde{Q}_{\bar{i}(N),\bar{i}(N)-1} = 0$$

$$\tilde{Q}_{\bar{i}(N),\bar{i}(N)} = 1 - \tilde{Q}_{\bar{i}(N),\bar{i}(N)+1}$$

That is, we introduce a reflecting barrier at $\bar{i}(N)$.

Now we specify the choice of the auxiliary point $\bar{i}(N)$. It has to satisfy the following requirements:

(1) The time averages on the time interval $R_N = ET^N_a/N$ are stable during the typical times to reach $\bar{i}(N)$ starting from $i_2(N)$.

(2)

$$\frac{ET_{i_2(N),\bar{i}(N)}}{ET^a_N} \xrightarrow[N\to\infty]{} \infty$$

It turns out that a possible choice for $\bar{i}(N)$ is the following:

$$\bar{i}(N): \quad NET^a_N \leqslant ET_{i_2(N),\bar{i}(N)} \leqslant N^2 ET^a_N$$

Theorem 2.3. $\forall \epsilon > 0$, $\forall J = [a,b]$ such that $(a,b) \ni 1 - 2i_2(N)/N$ if we define for $R_N = ET^N_a/N$

$$l^{**} = \inf\left\{ l \geqslant 0 : lR_N \leqslant T^N_{i_2(N),\bar{i}(N)} < (l+1)R_N \right\}$$

and

$$\bar{\Gamma}(N) = \{l^{**} = 0\} \cup \left\{ l^{**} \geqslant 1,\ \sup_{l < l^{**}} \left| \frac{1}{R_N} \sum_{t = lR_N}^{(l+1)R_N} 1_j\big(\xi^N_{i_2(N)}(t)\big) - 1 \right| < \epsilon \right\}$$

Then

$$P(\tilde{\Gamma}_N) \xrightarrow[N\to\infty]{} 1.$$

Proof. Following the same lines and using the same notations of the proof of Theorem 2.2 we define

$$\tilde{B}_l^c = \left[\left| \frac{1}{R_N} \sum_{t=[lR_N]}^{[(l+1)R_N]} 1_J\big(\tilde{\xi}_N^{i_2(N)}(t)\big) - 1 \right| \geqslant \epsilon \right]$$

and we show that for $K_N = N^4$ the following conditions are satisfied:

(i) $$K_N R_N / ET_{i_2(N),\bar{i}(N)} \xrightarrow[N\to\infty]{} \infty$$

(ii) $$K_N^2 \sup_{l < R_N} P(\tilde{B}_l^c) \xrightarrow[N\to\infty]{} 0 \quad \blacksquare$$

Remark. As in the case of l_N^*, we can see, as a consequence of Proposition 2.1 and Lemma A1, that $P(l_N^{**} = 0) \to 0$ as $N \to +\infty$, for previous choice of R_N.

Proof of Theorem 2.1. For the moment let us assume Proposition 1 is already proved, and let us see how to deduce Theorem 2.1.

From Proposition 2.1, Theorems 2.2 and 2.3, and Eqs. (2.27), (2.28) we deduce quite immediately that for each continuous function f on $[-1, +1]$ the processes $(A^N_{R_N}(sET^N_a, f),\ s \geqslant 0)$ converge in law on $D([0, +\infty), [-\|f\|, \|f\|])$, to a jump process $(X_s(f),\ s \geqslant 0)$, given by

$$\begin{aligned} X_s(f) &= f(x_1) \qquad \text{if} \quad s < T \\ &= f(x_2) \qquad \text{if} \quad s \geqslant T \end{aligned}$$

where T has an exponential distribution with mean 1. [*Remark*: Notice that $ET_{i_2(N),i_0(N)}/ET_{i_1(N),i_*(N)} \to +\infty$ as $N \to +\infty$ and this is why we do not see any return to x_1.]

But the situation is as follows. We have a family $(X_s^N,\ s \geqslant 0)$, $N \geqslant 1$ of real-valued processes with paths in $D([0, +\infty), [-M, M])$ ($M = \sup_x |f(x)|$), verifying the following: there exist random variables τ_N^1 converging weakly to a unit mean exponentially distributed random variable, and random variables $\delta_N \geqslant 0$, $\tau_N^2 \geqslant \tau_N^1$ with $\tau_N^2 \xrightarrow{P} +\infty$ and $\delta_N \xrightarrow{P} 0$, and there exist α, $\beta \in \mathbb{R}$ such that for all $\epsilon > 0$

$$P\left[\sup_{s<\tau_N^1} |X_s^N - \alpha| > \epsilon\right] \to 0$$

$$P\left[\sup_{\delta_N+\tau_N^1 < s < \tau_N^2} |X_s^N - \beta| > \epsilon\right] \to 0$$

as $N \to +\infty$. [Here we are using Theorems 2.2 and 2.3 together with previous remarks that, by Proposition 2.1 we get $P(l_N^* = 0) \to 0$ and $P(l_N^{**} = 0) \to 0$ in those theorems, as well as Remark 2 after Theorem 2.2.]

In fact, in this situation, conditions (i) and (ii) of Billingsley[1] are easily verified, ensuring tightness on $D([0,b],\mathbb{R})$ for each $b > 0$.

Now, we can argue as in Ref. 4 to obtain tightness of the family $(A_{R_N}^N(sET_a^N),\ 0 \leqslant s \leqslant b)$ on $D([0,b],\mathfrak{M}_1[-1,+1])$, for each $b > 0$. Then identification of the limit is very simple. ■

Therefore, what we must do is to prove Proposition 2.1.

Proof of Proposition 2.1. It consists of two steps. Let $S_N = T_a^N / ET_a^N$. Then, (a) if S is the weak limit of some subsequence $(S_{N_k})_{k\geqslant 1}$, it verifies

$$P(S > t + s) = P(S > t)P(S > s) \qquad \text{for all} \quad s,t > 0.$$

(b) The family $(S_N,\ N \geqslant 1)$ is uniformly integrable. ■

Remark. Proposition 1 follows at once from (a) and (b) above.

Now, (b) can be verified through Lemma A1 and analogous calculations for $E(T_a^N)^2$, which can be used to see that $E(T_a^N)^2/(ET_a^N)^2$ remains bounded as $N \to +\infty$.

For the proof of (a) let us recall that for the process $\bar{\xi}_N$ we have

$$\sup_{l<i_*(N)} E\bar{T}_{l,i_1(N)}^N \leqslant C_1 e^{C_2 N^{3/4}}$$

for some constants C_1, C_2. Thus if we take $\alpha_N = e^{N^\gamma}$, where $3/4 < \gamma < 1$, we have

$$\alpha_N / ET_a^N \to 0$$

$$\sup_{l<i_*(N)} E\bar{T}_{l,i_1(N)/\alpha_N}^N \to 0 \qquad \text{as} \quad N \to +\infty$$

Now, if $s,t > 0$, defining

$$t_j^s = \inf\{u \geqslant sET_a^N : \xi_N(u) = j\}$$

if N is large enough so that $\alpha_N < [tET_a^N]$, we get

$$P\left(t_{i_1(N)}^s > sET_a^N + \alpha_N,\, T_a^N > (s+t)ET_a^N\right)$$

$$= P\left[\sum_{u=[sET_a^N]}^{[sET_a^N+\alpha_N]} I_{\{i_1(N)\}}(\xi_N(u)) = 0,\ T_a^N > (s+t)ET_a^N\right]$$

$$\leqslant P\left[\sum_{u=[sET_a^N]}^{[sET_a^N+\alpha_N]} I_{\{i_1(N)\}}\left(\bar{\xi}_N(u)\right) = 0,\ \bar{T}_a^N > (s+t)E\bar{T}_a^N\right]$$

$$\leqslant \sup_{l<i_*(N)} P\left(\bar{T}_{l,i_1(N)}^N > \alpha_N\right)$$

which converges to zero as $N \to +\infty$ (by Chebychev). Thus, if $s, t > 0$,

$$P\left[T_a^N > (s+t)ET_a^N\right] - P\left[T_a^N > (s+t)ET_a^N,\, t_{i_1(N)}^s \leqslant sET_a^N + \alpha_N\right]$$

goes to zero as $N \to +\infty$. But this last term can be written

$$\sum_{k=[sET_a^N]}^{[sET_a^N+\alpha_N]} P\left[T_a^N > k,\, t_{i_1(N)}^s = k\right]$$

$$\times P\left[T_a^N > (s+t)ET_a^N \mid T_a^N > k,\, t_{i_1(N)}^s = k\right]$$

which (using the strong Markov property) is bounded above by

$$P(T_a^N > sET_a^N)P(T_a^N > tET_a^N - \alpha_N)$$

and bounded below by

$$P\left(T_a^N > sET_a^N + \alpha_N,\, t_{i_1(N)}^s \leqslant sET_a^N + \alpha_N\right)P\left(T_a^N > tET_a^N\right)$$

Since $\alpha_N / ET_a^N \to 0$, these facts clearly imply (a) because as above,

$$P\left(T_a^N > sET_a^N + \alpha_N\right) - P\left(T_a^N > sET^N + \alpha_N,\, t_{i_1(N)}^s \leqslant sET_a^N + \alpha_N\right) \to 0$$

3. METASTABLE BEHAVIOR OF THE CONTACT PROCESS

3.1. Definition of the Process

The basic one-dimensional contact process (as defined by Harris) is a continuous time Markov process taking its values on the set $\mathscr{P}(\mathbb{Z})$, of all subsets of $\mathbb{Z}$. It can be described informally as follows: particles are distributed in $\mathbb{Z}$ in such a way that each site is either empty or occupied by at most one particle. $\xi(t)$ denotes the set of occupied sites at time t.

The (stochastic) time evolution can be described as follows. Each particle disappears after waiting a random time exponentially distributed with mean 1, independently of the behavior of the others. Before disappearing each particle keeps trying to create new ones at the neighboring sites, at random times distributed as a Poisson point process with parameter 2λ, where $\lambda > 0$. More precisely, each particle waits a random exponential time with mean $1/2\lambda$, independent of everything else, and then it puts a new particle in one of its nearest neighbors, chosen with probabilities $1/2$, $1/2$, and so on. At each of these creation times, if the chosen site is already occupied nothing happens.

Following Harris[9] we will construct the contact process with the help of a random graph in the "space-time" diagram $\mathbb{Z} \times \mathbb{R}_+$. For each $i \in \mathbb{Z}$ consider three independent Poisson point processes on $\mathbb{R}_+$: $(\vec{\tau}_n^i)_{n\in\mathbb{N}}$, $(\overleftarrow{\tau}_n^i)_{n\in\mathbb{N}}$, $(\overset{+}{\tau}{}_n^i)_{n\in\mathbb{N}}$, with parameters λ, λ, and 1, respectively (i.e., the random variables $\vec{\tau}_1^i, \vec{\tau}_2^i - \vec{\tau}_1^i, \ldots, \vec{\tau}_{n+1}^i - \vec{\tau}_n^i, \ldots$ are independent exponential random variables with mean $1/\lambda$, etc.). We suppose that for i varying in $\mathbb{Z}$ the processes are all independent.

Now, for each $i \in \mathbb{Z}$ we draw arrows from $(i, \vec{\tau}_1^i)$ to $(i+1, \vec{\tau}_1^i)$, from $(i, \vec{\tau}_2^i)$ to $(i+1, \vec{\tau}_2^i)$, etc. Secondly, we draw arrows from $(i, \overleftarrow{\tau}_1^i)$ to $(i-1, \overleftarrow{\tau}_1^i)$, from $(i, \overleftarrow{\tau}_2^i)$ to $(i-1, \overleftarrow{\tau}_2^i)$, etc. Finally we put down $+$ signs at each of the points $(i, \overset{+}{\tau}{}_n^i)$, $n = 1, 2, \ldots$. (For this graphical construction, see also Ref. 5.)

Given two points (i,s) and (j,t) in the space-time $\mathbb{Z} \times \mathbb{R}_+$, with $s < t$, we will say that there is a path from (i,s) to (j,t) if there is a chain of upward vertical segments and arrows in the random graph we have just constructed, leading from (i,s) to (j,t), following (horizontally) the direction of the arrows and without passing (vertically) through a $+$.

Now, given $A \in \mathscr{P}(\mathbb{Z})$, we will define the process $(\xi^A(t),\ t \geqslant 0)$, starting at A, in the following way:

$$\xi^A(0) = A$$

and for $t \geqslant 0$, $\xi^A(t) = \{ j \in \mathbb{Z} : \text{there is a path from } (i,0) \text{ to } (j,t), \text{for some } i \in A \}$.

(A glance at Fig. 2 may help in understanding the definition of the process. For more details see Harris[8] and Griffeath.[4])

Exactly in the same way we can define contact processes taking values on $\mathscr{P}(\{-N, -N+1, \ldots\})$ or on $\mathscr{P}(\{-N, -N+1, \ldots, N-1, N\})$, where N is a positive integer. In the first case it is enough to take Poisson point processes indexed by $i \geqslant -N$ and, for $i = -N$, to consider $(\vec{\tau}_n^{-N})_{n\geqslant 1}$ and $(\overset{+}{\tau}{}_n^{-N})_{n\geqslant 1}$ (i.e., a particle at site $-N$ can only create new particles at site $-N+1$). The contact process with values on $\mathscr{P}(\{-N, -N+1, \ldots\})$ is denoted by $(\xi_{[-N,+\infty)}(t))_{t\geqslant 0}$. Now, in order to define a contact process on $\mathscr{P}(\{-N, -N+1, \ldots, N-1, N\})$ we consider only Poisson point processes indexed by $i = -N, \ldots, N$; for $i = -N$ we only take $(\vec{\tau}_n^{-N})_{n \geqslant 1}$ and $(\overset{+}{\tau}_n^{-N})_{n \geqslant 1}$, and analogously, for $i = N$ we only take $(\overleftarrow{\tau}_n^{N})_{n \geqslant 1}$ and $(\overset{+}{\tau}_n^{N})_{n \geqslant 1}$. The process defined in this way will be denoted by $\xi_N(t)$, $t \geqslant 0$. In all cases the upper index will indicate the initial configuration [i.e., $\xi_N^A(0) = A$].

An enormous amount of results have been produced about the contact process since the pioneer paper by Harris[8] was published. A quite exhaustive review can be found in Griffeath,[6] which contains most of the facts about the process which will appear in the proof of our theorem. Here we just indicate a basic ergodic result which we shall use to enounce this theorem.

It is obvious that the probability measure concentrated at the empty set $\delta_\varnothing$, is invariant for the contact process. Now, there exist λ_* such that if $\lambda > \lambda_*$ then, as $t \to +\infty$, the law of $\xi^{\mathbb{Z}}(t)$ converges weakly to a different probability measure μ, which must be invariant for ξ. Also the same thing happens for $\xi_{[-N,+\infty)}$ and we denote by $\mu_{[-N,+\infty)}$ the corresponding invariant measure. This probability μ (resp. $\mu_{[-N,+\infty)}$) gives mass zero to the set of finite subsets of $\mathbb{Z}$ ($\{-N, -N+1, \ldots\}$ resp.). In particular $\delta_\varnothing$ and μ (as well as $\delta_\varnothing$ and $\mu_{[-N,+\infty]}$) are mutually singular.

Remark. It is known (see Ref. 3) that ξ and $\xi_{[-N,+\infty)}$ have the same critical parameter λ_*.

3.2. The Results

The process $\xi_N(t)$, $t \geqslant 0$ has only one invariant probability measure (which is $\delta_\varnothing$), for any value of λ. (This is a quite elementary fact since $\xi_N(t)$ is a Markov process, with values on a finite set, in which it is always possible to go from any state to any other, except for the empty set, which is a trap.) Nevertheless, if λ is big enough, and if we start the process with a configuration with most of the sites occupied, then for a long time the process will behave as if it were (statistically) in equilibrium with distribution μ restricted to $\{-N, \ldots, N\}$, where μ is the nontrivial invariant distribution corresponding to the contact process on $\mathscr{P}(\mathbb{Z})$ with the same birth rate λ. The point is that μ restricted to $\{-N, \ldots, N\}$ is not invariant for $\xi_N(t)$, $t \geqslant 0$ and what happens is that waiting long enough we will *suddenly* see the system changing completely its statistical behavior, the frequencies becoming stable around the real unique invariant distribution of the finite contact process, i.e., $\delta_\varnothing$. This is what we shall call metastable behavior of the contact process, as discussed in the Introduction.

Now we will enounce precisely our results, which are analogous to those obtained for the Curie–Weiss model, in Section 2. In order to do this, we need to introduce some new notations. First we remember that the upper symbol in the notation of the process indicates its initial value. For instance, $\xi^A_{[-N,+\infty]}$, $t \geqslant 0$ means the process $\xi_{[-N,+\infty]}$ starting with $\xi^A_{[-N,+\infty]}(0) = A$. Let

$$T^A = \inf\{t > 0 : \xi^A(t) = \emptyset\}$$

$$T^A_N = \inf\{t > 0 : \xi^A_N(t) = \emptyset\}$$

We will omit the upper symbol when the process starts with the full set, i.e., $\xi(0) = \mathbb{Z}$, $\xi_N(0) = \{-N, \ldots, N\}$, etc. Also,

$$T_N = \inf\{t > 0 : \xi_N(t) = \emptyset\}, \quad \text{etc.}$$

If $A = \{x\}$ we write ξ^x, T^x and so on.

A cylindrical $f: \mathscr{P}(\mathbb{Z}) \to \mathbb{R}$ is a function such that there exists a finite set $B \subseteq \mathbb{Z}$ so that $f(A) = f(A \cap B)$ for all $A \in \mathscr{P}(\mathbb{Z})$. The smallest such B is called the basis of f.

Given a real number $R > 0$ we define, for $N = 1, 2, \ldots,$ the measure-valued process $(A^N_R(s), s \geqslant 0)$ given by the following empirical distribution:

$$A^N_R(s, \cdot) = \frac{1}{R} \int_s^{s+R} \delta_{\xi_N(t)}(\cdot)\, dt$$

$\mathfrak{M}_1(\mathscr{P}(\mathbb{Z}))$ denotes the space of probability measures on $\mathscr{P}(\mathbb{Z})$, with its weak* topology ($\mathscr{P}(\mathbb{Z}) = \{0, 1\}^{\mathbb{Z}}$ with product topology, and Borel σ field). The processes A^N_R have paths in $D([0, +\infty), \mathfrak{M}_1(\mathscr{P}(\mathbb{Z})))$ and the convergence in the next theorem means convergence of the probability measures in this Skorohod space, with its usual topology and Borel σ field.

In the next proofs we shall use $\vec{\xi}$, $\vec{\xi}_{[-N,+\infty)}$ to denote the corresponding contact processes constructed with arrows pointing to the right only. It is known that $\vec{\xi}$ and $\vec{\xi}_{[-N,+\infty)}$ also present critical phenomena; the critical parameter will be denoted by $\vec{\lambda}_*$. *Remark*: $\vec{\lambda}_* \geqslant \lambda_*$.

In the following theorems we would like to substitute $\vec{\lambda}_*$ by λ_* and we conjecture that this can be done. [It depends on our proof of (3.12) in Theorem 3.2.]

We can now state the following:

Theorem 3.1. If $\lambda > \vec{\lambda}_*$ there exist increasing sequences of real numbers $(R_N)_{N \geqslant 1}$ such that the processes $(A^N_{R_N}(sET_N, \cdot), s \geqslant 0))$ converge in law, when $N \to +\infty$, to a jump Markov process $(A(s, \cdot), s \geqslant 0)$ given by

$$A(s, \cdot) = \begin{cases} \mu & \text{if} \quad s < T \\ \delta_\emptyset & \text{if} \quad s \geqslant T \end{cases}$$

where T is a random time which has an exponential distribution with mean 1.

This theorem will follow, as in the case of the Curie–Weiss model (Theorem 2.1), from the following results, which are stated and proved next.

Theorem 3.2. If $\lambda > \vec{\lambda}_*$, then T_N/ET_N converges in distribution, as $N \to \infty$, to a unit mean exponentially distributed random time.

Theorem 3.3. $\lambda > \vec{\lambda}_*$, there exists an increasing sequence of positive real numbers $(R_N : N \geqslant 1)$ such that (i) $R_N/ET_N \to 0$ as $N \to +\infty$, and (ii) for all $\epsilon > 0$, and all f cylindrical on $\mathscr{P}(\mathbb{Z})$,

$$P\Big(\max_{0<l<K_N} |A^N_{R_N}(lR_N, f) - \mu(f)| > \epsilon, k_N > 0\Big)$$

where $K_N = \max\{l \geqslant 0 : lR_N < T_N\}$.

Proof of Theorem 3.2. We first show (i) for all $s > 0$, $t > 0$

$$\lim_{N\to\infty} P[T_N/\beta_N > s+t] - P[T_N/\beta_N > s]P[T_N/\beta_N > t] = 0 \quad (3.1a)$$

where β_N is defined by $P(T_N > \beta_N) = e^{-1}$, and

$$\text{(ii)} \qquad \lim_{N\to\infty} \frac{\beta_N}{ET_N} \to 1 \qquad (3.1b)$$

To prove (3.1a) let us start considering that

$$\begin{aligned}
&P[T_N/\beta_N > s+t] \\
&\quad = \sum_{\substack{A \subseteq \{-N,\ldots,N\} \\ A \neq \varnothing}} P[T_N/\beta_N > s+t \mid \xi_N(\beta_N s) = A]P[\xi_N(\beta_N s) = A] \\
&\quad = \sum_{\substack{A \subseteq \{-N,\ldots,N\} \\ A \neq \varnothing}} P[T_N/\beta_N > s+t \mid \xi_N(\beta_N s) = A] \\
&\qquad \times P[\xi_N(\beta_N s) = A, T_N > \beta_N s] \\
&\quad = \sum_{\substack{A \subseteq \{-N,\ldots,N\} \\ A \neq \varnothing}} P[T_N^A/\beta_N > t]P[\bar{\zeta}_N(\beta_N s) = A, T_N > \beta_N s] \\
&\quad = P(T_N/\beta_N > t)P(T_N/\beta_N > s) \\
&\qquad + \sum_{\substack{A \subseteq \{-N,\ldots,N\} \\ A \neq \varnothing}} \big[P(T_N^A/\beta_N > t) - P(T_N/\beta_N > t)\big] \\
&\qquad \times P(\xi_N(\beta_N) = A, T_N > \beta_N s)
\end{aligned} \qquad (3.2)$$

Notice that by construction $T_N > T_N^A$ so that

$$\begin{aligned}
0 &\leqslant P(T_N/\beta_N > t)P(T_N/\beta_N > s) - P(T_N/\beta_N > s+t) \\
&= \sum_{\substack{A \subseteq \{-N,\dots,N\} \\ A \neq \varnothing}} \left[P(T_N/\beta_N > t) - P\left(T_N^A/\beta_N > t\right)\right] \\
&\quad \times P(\xi_N(\beta_N s) = A,\, T_N > \beta_N s) \\
&\leqslant \sum_{\substack{A \subseteq \{-N,\dots,N\} \\ A \in E_b}} \left[P(T_N/\beta_N > t) - P\left(T_N^A/\beta_N > t\right)\right] \\
&\quad \times P(\xi_N(\beta_N s) = A,\, T_N > \beta_N s) \\
&\quad + P(\xi_N(\beta_N s) \notin E_b,\, T_N > \beta_N s)
\end{aligned} \tag{3.2$'$}$$

where for $0 < b \leqslant N$,

$$E_b = \left\{ A \subseteq \mathbb{Z} : \frac{|A \cap [-b,b]|}{2b+1} > \frac{\rho}{2} \right\}, \quad \text{and} \quad \rho = \mu\{A : 0 \in A\}.$$

Later on we shall use the fact that $\forall \epsilon > 0$ $\mu(E_b) \geqslant 1 - \epsilon$ if b is large enough, which follows from

$$\mu\left\{ A : \lim_{b \to +\infty} \frac{|A \cap [-b,b]|}{2b+1} = \rho \right\} = 1 \tag{3.3}$$

(See Ref. 3 or 6 for the proof of the ergodicity of μ.)

Thus, to prove (3.1) it is enough to show that for all $\epsilon > 0$ there exist $b = b(\epsilon)$ and $N(\epsilon)$ such that

$$\text{(i)} \qquad P\left[\xi_N(\beta_N s) \notin E_b,\, T_N > \beta_N s\right] < \epsilon \qquad \text{if} \quad N \geqslant N(\epsilon) \tag{3.4}$$

and

$$\begin{aligned}
\text{(ii)} \qquad & \sum_{A \in E_b} P\left[\xi_N(\beta_N s) = A,\, T_N > \beta_N s\right] \\
& \qquad \times \left(P\left[T_N/\beta_N > t\right] - P\left[T_N^A/\beta_N > t\right]\right) < \epsilon
\end{aligned}$$

if $N \geqslant N(\epsilon)$.

Since $b = b(\epsilon)$ is fixed in (i), it is enough to verify that for all $A \in E_b$

$$0 \leqslant P\left[T_N/\beta_N > t\right] - P\left[T_N^A/\beta_N > t\right] < \epsilon \tag{3.5}$$

if $N \geqslant N(\epsilon)$. Notice that the left-hand side in (3.5) is less than or equal to $P[T_N \neq T_N^A]$. We shall show that $P[T_N \neq T_N^A] \leqslant \epsilon$ if $N \geqslant N(\epsilon)$, for all $A \in E_b$.

Since $A \in E_b$ we have $|A \cap [-b,b]| > \rho(2b+1)/2$. On the other side, we know that if $|A| = n$,

$$P[T^A = +\infty] \geqslant P(T^{[1,n]} = +\infty] = \mu\{B : B \cap [1,n] \neq \emptyset\} \nearrow 1$$

as $n \nearrow +\infty$, provided $\lambda > \lambda_*$, where $[1,n]$ really means $[1,n] \cap \mathbb{Z}$.

Given $\epsilon > 0$, let us take $n(\epsilon)$ such that if $n \geqslant n(\epsilon)$ then $\mu\{B : B \cap [1,n] \neq \emptyset\} \geqslant 1 - \epsilon$ see Ref. 5 and take $b = b(\epsilon)$ such that $p(2b+1)/2 \geqslant n(\epsilon)$. Then

$$P[T^{A \cap [-b,b]} = +\infty] \geqslant 1 - \epsilon \tag{3.6}$$

Now, by construction

$$[T^{A \cap [-b,b]} = +\infty] = \bigcup_{x \in A \cap [-b,b]} [T^x = +\infty]$$

If $x \in \{-N, \ldots, N\}$ let

$$U_N^x = \inf\{t : \xi^x(t) \cap \{-N, N\} \neq \emptyset\} \tag{3.7}$$

with the convention that when this set is empty we set $U_N^x(\omega) = +\infty$.

On the other side, by a standard argument for Markov chains

$$P[T^x = +\infty \text{ and } U_N^x = +\infty] = 0 \qquad \text{for each} \quad x \in \{-N, \ldots, N\}$$

Thus, setting

$$\Lambda = \Lambda(A,b,N) \stackrel{\text{def}}{=} [T^x = +\infty \text{ and } U_N^x < +\infty \text{ for some } x \in A \cap [-b,b]]$$

we have

$$P(\Lambda) \geqslant 1 - \epsilon \tag{3.8}$$

for our previous choice of b.

Let us also remark that for each $x \in \{-N, \ldots, N\}$ we have (by construction)

$$\xi^x(t) = \xi_N^x(t) \qquad \text{if} \quad t \leqslant U_N^x \quad \text{and so}$$

$$U_N^x \leqslant T_N^x \leqslant T_N^A \qquad \text{on the set } [T^x = +\infty], \quad \text{if} \quad x \in A.$$

Now,

$$P[T_N \neq T_N^A] \leqslant P(\Lambda^c) + P(\Lambda \cap [T_N \neq T_N^A])$$

with $\Lambda = \Lambda(A,b,N)$, since $P(\Lambda^c) < \epsilon$ for $b = b(\epsilon)$ we need to take care of the last term. We have

$$P(\Lambda \cap [T_N \neq T_N^A]) = \sum_{\emptyset \neq B \subseteq \{-N, \ldots, N\}} P([\bar{\xi}^X(U_N^X) = B] \cap \Lambda)$$

$$\cdot P([T_N \neq T_N^A] \mid [\xi^X(U_N^X) = B] \cap \Lambda) \tag{3.9}$$

where if $\omega \in \Lambda$, $X(\omega)$ is the smallest x in $A \cap [-b, b]$ such that $T^x(\omega) = +\infty$ and $U^x(\omega) < +\infty$. [Outside Λ set $X(\omega) = 0$ for example, so that it is a well-defined random variable.]

It is very simple to verify that

$$\xi^A(t) \cap \left[l^x(t), r^x(t) \right] = \xi(t) \cap \left[l^x(t), r^x(t) \right]$$

if $x \in A$ and $t < T^x$, where

$$l^x(t) = \min \xi^x(t) \quad \text{and} \quad r^x(t) = \max \xi^x(t)$$

As a consequence we have that for all $\omega \in \Lambda$

$$\begin{aligned} \xi^X(t) \cap \left[l^X(t), r^X(t) \right] &= \xi_N^A(t) \cap \left[l^X(t), r^X(t) \right] \\ &= \xi_N(t) \cap \left[l^X(t), r^X(t) \right] \qquad \text{if} \quad t \leqslant U_N^X \end{aligned}$$

It is obvious for $\omega \in \Lambda$ that there is one (and only one) of the next two equalities which holds:

$$l^X\left(U_N^X\right) = -N \quad \text{or} \quad r^X\left(U_N^X\right) = N$$

To fix ideas, let us suppose $l^X(U_N^X) = -N$. Then

$$\xi^A\left(U_N^X\right) \cap \left[-N, r^X\left(U_N^X\right) \right] = \xi\left(U_N^X\right) \cap \left[-N, r^X\left(U_N^X\right) \right]$$

Thus using the spatial symmetry of the process we see that the right-hand side of (3.9) is bounded above by

$$P\Big(\left[|\xi^X\left(U_N^X\right)| \leqslant n(\epsilon) \right] \cap \Lambda\Big)$$
$$+2 \sum_{\substack{-N \in B \subseteq \{-N, \ldots, N\} \\ |B| > n(\epsilon)}} P\Big(\left[\xi^X\left(U_N^X\right) = B \right] \cap \Lambda\Big) P\Big(\left[T_N^B \neq T_N^{B \cup \bar{B}^c} \right]\Big)$$

where $\bar{B} = \{ y \in \mathbb{Z} : \min B \leqslant y \leqslant \max B \}$ and $1 \leqslant n(\epsilon) \leqslant N$. Also, it is known that $\lim_{t \to +\infty} |\bar{\xi}^x(t)|/t = \alpha$, where α is a positive constant for almost all $\omega \in [T^x = +\infty]$. (See Ref. 3, Theorem 9.)

On the other side U_N^x is greater than a sum of $N - b$ independent exponential random variables with mean $1/\lambda$. Thus $U_N^x \to +\infty$ a.s. on $[T^x = +\infty]$; on the other end we have that $\forall m : |\xi^x(t)| > m$ for all sufficiently large times a.s. on $[T_x = +\infty]$.

From this it follows that, given $n(\epsilon)$, there exists $N_1(\epsilon)$ such that if $N \geqslant N_1(\epsilon)$ then $P([|\xi^X(U_N^x)| \leqslant n(\epsilon)] \cap \Lambda) < \epsilon$. So, it is enough to show that we can choose $n(\epsilon)$ large enough so that if $B \subseteq \{-N, \ldots, N\}$ with $-N \in B$ and $|B| \geqslant n(\epsilon)$ then there exists $N_2(\epsilon)$ such that

$$P\Big(\left[T_N^B \neq T_N^{B \cup \bar{B}^c} \right]\Big) < \epsilon \qquad \text{if} \quad N \geqslant N_2(\epsilon) \tag{3.10}$$

Since $\lambda > \vec{\lambda}_*$ we can take $n(\epsilon)$ sufficiently large so that

$$P\left(\left[\vec{T}^{[1,n(\epsilon)]} = +\infty\right]\right) > 1 - \epsilon$$

where $\vec{T}^A = \inf\{t > 0 : \vec{\xi}^A(t) = \emptyset\}$ and $\vec{\xi}(t)$ is the contact process constructed with arrows to the right only.

Thus, if $B \subseteq \{-N, \ldots, N\}$, $-N \in B$ and $|B| > n(\epsilon)$, $P[\vec{T}^B = +\infty] > 1 - \epsilon$. Let

$$V_N^B = \inf\left\{t > 0 : N \in \vec{\xi}^B(t)\right\}$$

with the convention $V_N^B(\omega) = +\infty$ if this set is empty. It is easy to see that for such B, on $[T^B = +\infty]$ we have $V_N^B \leqslant T_N^B$ and also

$$\begin{aligned} &\vec{\xi}^B(t) = \vec{\xi}_N^B(t) \subseteq \xi_N^B(t) \qquad \text{if} \quad t \leqslant V_N^B \quad \text{and} \\ &\xi_N^B(t) = \xi_N^{B \cup \bar{B}^c}(t) \qquad\quad\ \ \text{if} \quad t \geqslant V_N^B \end{aligned} \tag{3.11}$$

Thus

$$P\left(\left[T_N^B \neq T_N^{B \cup \bar{B}^c}\right]\right) \leqslant P\left(\left[\vec{T}^B < +\infty\right]\right) < \epsilon$$

which shows (ii) in (3.4).

Remarks. (1) The reader may check the second line in (3.11) by observing the following: first of all

$$\max \vec{\xi}_N^B(t) \leqslant \max \xi_N^B(t), \qquad \text{if} \quad t \leqslant V_N^B$$

and

$$\begin{aligned} &\left[\min \xi_N^B(t), \max \xi_N^B(t)\right] \cap \xi_N^{B \cup \bar{B}^c}(t) \\ &\qquad = \left[\min \xi_N^B(t), \max \xi_N^B(t)\right] \cap \xi_N^B(t) \end{aligned}$$

if $t \leqslant T_N^B$.

Now, if $-N \in B$, since $\max B < \min \bar{B}^c$ we have

$$\xi_N^B(t) \cap \left[-N, \max \xi_N^B(t)\right] = \xi_N^{B \cup \bar{B}^c}(t) \cap \left[-N, \max \xi_N^B(t)\right] \qquad \text{if} \quad t < T_N^B.$$

But at the time V_N^B we have $\max \xi_N^B(t) = N$ so that $\xi_N^B(V_N^B) = \xi_N^{B \cup \bar{B}^c}(V_N^B)$ and the statement follows.

(2) It is clear that when we write ξ_N^C and $C \not\subseteq \{-N, \ldots, N\}$, what we really mean is that the starting configuration is $C \cap \{-N, \ldots, N\}$.

Now, let us show (i) in (3.4). We have

$$\begin{aligned}
&P\left[\xi_N(\beta_N s) \notin E_b,\, T_N > \beta_N s\right] \\
&\quad \leqslant P\left[\xi_N(\beta_N s) \notin E_b,\, T_N > \beta_N s, \min \xi_N(\beta_N s) < -N + L,\right. \\
&\qquad \left. \max \xi_N(\beta_N s) > N - L\right] \\
&\quad + P\left[\min \xi_N(\beta_N s) \geqslant -N + L,\, T_N > \beta_N s\right] \\
&\quad + P\left[\max \xi_N(\beta_N s) \leqslant N - L,\, T_N > \beta_N s\right]
\end{aligned} \tag{3.12}$$

we are assuming $N > N - L > b$. By construction if $\beta_N s < T_N$

$$\begin{aligned}
&\xi_N(\beta_N s) \cap \left[\min \xi_N(\beta_N s), \max \xi_N(\beta_N s)\right] \\
&\quad = \xi(\beta_N s) \cap \left[\min \xi_N(\beta_N s), \max \xi_N(\beta_N s)\right]
\end{aligned}$$

Thus the first summand in (3.12) may be bounded above by $P[\xi(\beta_N s) \notin E_b]$. And, since $P[\xi(\beta_N s) \notin E_b] \to \mu(E_b^c) < \epsilon$, we can find $N_3(\epsilon) \in \mathbb{N}$ such that

$$P\left[\xi_N(\beta_N s) \notin E_b\right] < 2\epsilon \qquad \text{if} \quad N \geqslant N_3(\epsilon)$$

Finally, we must control $P[\min \xi_N(\beta_N s) > -N + L,\ T_N > \beta_N s]$, the other term in (3.12) being analogous.

It is easy to see that if $t < T_N$ then $\min \xi_N(t) > -N + L$ if and only if $\min \xi_{[-N,+\infty)}(t) > -N + L$. But, from construction $\xi_N(t) \subseteq \xi_{[-N,+\infty)}(t)$ for all $t \geqslant 0$. Thus it suffices to show that if $\min \xi_{[-N,+\infty)}(t) \leqslant -N + L$, then $\min \xi_N(t) \leqslant -N + L$ too, provided $t < T_N$. In effect, if $\min \xi_{[-N,+\infty)}(t) \leqslant -N + L$ then there exists a path from some point $(x, 0) \in \{-N, -N+1, \ldots\} \times \mathbb{R}_+$ to $(y, t) \in \{-N, \ldots, -N + L\} \times \mathbb{R}_+$. If $x \in \{-N, \ldots, N\}$ then $y \in \xi_N(t)$, as we wished. Let us consider then the case $x > N$: since, by hypothesis $t < T_n$, i.e., there exists a path from $(u, 0) \in \{-N, \ldots, N\} \times \mathbb{R}_+$ to $(v, t) \in \{-N, \ldots, N\} \times \mathbb{R}_+$. This path must meet the one which goes from $(x, 0)$ to (y, t), and thus $y \in \xi_N(t)$.

Consequently,

$$\begin{aligned}
&P\left[\min \xi_N(\beta_N s) > -N + L,\, T_N > \beta_N s\right] \\
&\quad \leqslant P\left[\min \xi_{[-N,+\infty)}(\beta_N s) > -N + L\right]
\end{aligned}$$

Let us now take $L = L(\epsilon)$ such that

$$\mu_{[-N,+\infty)}\{A \subseteq [-N, +\infty) \cap \mathbb{Z} : A \cap [-N, -N + L] = \varnothing\} < \epsilon$$

[Such $L(\epsilon)$ exists if $\lambda > \lambda_*$.] Finally, since $\xi_{[-N,+\infty)}(t) \to \mu_{[-N,+\infty)}$ weakly, as $t \to +\infty$, we have

$$P\left(\left[\xi_{[-N,+\infty)}(\beta_N s) > -N + L\right]\right) \leqslant 2\epsilon \qquad \text{if} \quad N \geqslant N_4(\epsilon)$$

thus completing the proof of Eq. (3.1a).

Now, owing to the definition of β_N and the monotonicity in t of $P_N(T_N > \beta_N t)$, it is easy to show that

$$\lim_{N\to\infty} P(T_N/\beta_N > t) = \exp(-t) \tag{3.13}$$

Equation (3.13) together with Eq. (3.1b) clearly implies the result, so that we are left with the proof Eq. (3.1b).

From Eq. (3.2′) and from the definition of β_N we have that

$$\forall \text{ integer } n \qquad P(T_N/\beta_N > n) \leqslant \exp(-n)$$

Now, since

$$ET_N/\beta_N = \int_0^\infty P(T_N/\beta_N > t)\,dt$$

it follows from Lebesgue's theorem that

$$\lim_{N\to\infty} ET_N/\beta_N = \int_0^\infty \lim_{N\to\infty} P(T_N/\beta_N > t)\,dt = \int_0^\infty e^{-t}\,dt = 1$$

Remark. The reader should be aware that in the proof of both theorems we have used the letters A and B to denote at least two things: A denotes a configuration (a subset of $\mathbb{Z}$) and $A_R^N(t)$ denotes a time average (as in Theorem 3.1). B sometimes denotes a configuration and B_k^N denotes an event.

Proof of Theorem 3.3. Since T_N is a.s. finite, for any positive number R_N, K_N is a well-defined and finite random variable with values on $\mathbb{N}$. Moreover, if the R_N verify condition (i) above we know, by Theorem 3.2, that $P(T_N < R_N) \to 0$ as $N \to +\infty$, i.e., $P(K_N = 0) \to 0$. Let us now assume (R_N) is a sequence satisfying (i). For $\epsilon > 0$ and f cylindrical given, let

$$B_k^N = \left[\left|A_{R_N}^N(kR_N, f) - \int f\,d\mu\right| > \epsilon\right]$$

Then, for any integer $m \geqslant 1$

$$\begin{aligned}
P\left[K_N \geqslant 1, \bigcap_{0 \leqslant k < K_N} (B_k^N)^c\right] &= \sum_{j=1}^{+\infty} P[K_N = j]\left(1 - P\left[\bigcup_{k=1}^{j-1} B_k^N \,\middle|\, K_N = j\right]\right) \\
&\geqslant P[1 \leqslant K_N \leqslant m] - \sum_{j=1}^{m} P[K_N = j] \\
&\quad \cdot j \cdot \max_{1 \leqslant k \leqslant j} P[B_k^N \mid K_N = j]
\end{aligned}$$

Now, we can take $L = L(\epsilon)$, positive integer such that the support of f is contained on $[-N + L, N - L]$,

$$\mu_{[-N,+\infty)}\{A : A \cap [-N, -N + L] = \varnothing\} < \frac{\epsilon}{4\|f\|_\infty} \qquad \text{for} \quad N > L$$

On the set

$$\left[\min \xi_N(t) < -N + L, \max \xi_N(t) > N - L\right],$$

we have

$$f(\xi_N(t)) = f(\xi(t))$$

If we let $h(\cdot) = I_{D_{N,L}}(\cdot)$; $D_{N,L} = \{\eta : \eta \cap [-N, -N+L] = \varnothing\}$, then for $t < T_N$

$$h(\xi_N(t)) = h(\xi_{[-N,+\infty)}(t)) \tag{3.14}$$

So, if $k < j$

$$P\left[B_k^N \mid K_N = j\right] \leqslant P\left[\left|\frac{1}{R_N}\int_{kR_N}^{(k+1)R_N} f(\xi(t))\,dt - \int f\,d\mu\right| > \frac{\epsilon}{2}\right]$$
$$+2P\left[2\|f\|_\infty \frac{1}{R_N}\int_{kR_N}^{(k+1)R_N} h(\xi_{[-N,+\infty)}(t))\,dt > \frac{\epsilon}{2}\right]$$

We want to show that we can choose R_N, m_N so that (i) holds and also

$$P\left[K_N \leqslant m_N\right] \to 1$$

$$m_N^2\left\{\max_{k\geqslant 1} P(\Gamma_k^n) + \max_{k \geqslant 1} P\left(\tilde{\Gamma}_k^N\right)\right\} \to 0$$

where

$$\begin{aligned} \Gamma_k^N &= \left[\left|\frac{1}{R_N}\int_{kR_N}^{(k+1)R_N} f(\xi(t))\,dt - \int f\,d\mu\right| > \frac{\epsilon}{2}\right] \\ \tilde{\Gamma}_k^N &= \left[2\|f\|_\infty \frac{1}{R_N}\int_{kR_N}^{(k+1)R_N} h(\xi_{[-N,+\infty)}(t))\,dt > \frac{\epsilon}{2}\right] \end{aligned} \tag{3.15}$$

By construction

$$\xi_{[-N,+\infty)}(kR_N + t) = {}_{kR_N}\xi^{\xi(kR_N)}_{[-N,+\infty)}(t) \subseteq {}_{kR_N}\xi_{[-N,+\infty)}(t)$$

where $\xi_{s[-N,+\infty)}(\cdot)$ is the contact process constructed with the Poisson point process translated in time by s. More precisely, for each A

$$\xi_s^A(t) = \{y \in \mathbb{Z} : \text{there exists a path going up from } (x,s) \text{ to}$$
$$(y, s+t) \text{ for some } x \in A\}.$$

Remark. In formula (3.19) we shall use ξ instead of $\xi_{[-N,+\infty)}$ because the notation would be too heavy.

Since h is decreasing, we have

$$h\left({}_{kR_N}\xi_{[-N,+\infty)}(t)\right) \leqslant h(\xi_{[-N,+\infty)}(kR_N + t))$$

and so

$$\frac{1}{R_N}\int_0^{R_N} h\big({}_{kR_N}\xi(t)\big)\,dt \leqslant \frac{1}{R_N}\int_{kR_N}^{(k+1)R_N} h(\xi(t))\,dt$$
$$= \frac{1}{R_N}\int_0^{R_N} h\big({}_{kR_N}\xi^{\xi(kR_N)}(t)\big)\,dt$$
$$\leqslant \int_{\mu_{[-N,+\infty)}} (dA)\,\frac{1}{R_N}\int_0^{R_N} h\big({}_{kR_N}\xi^A(t)\big)\,dt \qquad (3.19)$$

where in the last step we use that ${}_{kR_N}\xi(t) \supseteq {}_{kR_N}\xi^A(t)$, that h is decreasing and that $\mu_{[-N,+\infty)}$ is an invariant measure for the process.

Now, by Birkhoff theorem the last term on (3.19) can be made arbitrarily close to $\int h\,d\mu_{[-N,+\infty)}$ (a.s.), which can be taken very small by properly choosing L.

We have shown that $\max_{k\geqslant 0} P(\tilde{\Gamma}_k^N) \to 0$ when $R_N \to +\infty$. Let us now look at Γ_k^n.

For this we shall also use a result by Harris and Griffeath[5,6] which tells us that for a cylindrical function f,

$$\frac{1}{s}\int_0^s f(\xi(t))\,dt \to \int f\,d\mu \qquad \text{a.s.}$$

Let us assume for the moment that the cylindrical function f is increasing.

Then, the same argument which gave (3.19)—but for the contact process on $\mathbb{Z}$—allows to write

$$\frac{1}{R_N}\int_0^{R_N} f\big({}_{kR_N}\xi(t)\big)\,dt \geqslant \frac{1}{R_N}\int_0^{R_N} f\big({}_{kR_N}\xi^{\xi(kR_N)}(t)\big)\,dt$$
$$= \frac{1}{R_N}\int_{kR_N}^{(k+1)R_N} f(\xi(t))\,dt$$
$$\geqslant \int \mu(dA)\,\frac{1}{R_N}\int_0^{R_N} f\big({}_{kR_N}\xi^A(t)\big)\,dt$$

Using the ergodic theorem of Birkhoff, the individual ergodic theorem of Harris and Griffeath, and the properties of translation invariance for the Poisson point processes, we have if $k \geqslant 1$

$$P\left[\left|\frac{1}{R_N}\int_{kR_N}^{(k+1)R_N} f(\xi(t))\,dt - \int f\,d\mu\right| > \epsilon\right] < \delta'$$

if $R_N > s(\delta')$, with $s(\delta')$ defined by

$$P\left[\left|\frac{1}{s}\int_0^s f(\xi(t))\,dt - \int f\,d\mu\right| > \epsilon\right] < \frac{\delta'}{2}$$

and

$$P\left[\left|\int \mu(dA)\frac{1}{s}\int_0^s f\big(\xi^A(t)\big)\,dt - \int f\,d\mu\right| > \epsilon\right] < \frac{\delta'}{2}$$

if $s > s(\delta')$.

Let us define

$$\alpha(R_N) = \max_{k \geq 1}\left[P(\Gamma_k^N) + P(\tilde{\Gamma}_k^N)\right]$$

Then $\alpha(\cdot)$ is monotonic decreasing and goes to zero at infinity. We would like to find R_N and m_N satisfying simultaneously

(a) $$P[T_N \leqslant m_N R_N] \to 1$$

(b) $$R_N / ET_N \to 0$$

(c) $$m_N^2 \cdot \alpha(R_N) \to 0 \qquad \text{as} \quad N \to +\infty$$

Let $\psi(N)$ and $\varphi(N)$ be such that $\psi(N) \nearrow +\infty$, $\varphi(N) \nearrow +\infty$. We try to define $R_N = ET_N/\varphi(N)$ and $m_N = \varphi(N)\psi(N)$: with this (i) is immediately verified. To satisfy (ii) it will be enough that

$$\varphi^2(N)\cdot\alpha\left(\frac{ET_N}{\varphi(N)}\right) \to 0$$

Then, it will be enough to take $\psi(\cdot)$ so that $\psi^2(N) \nearrow +\infty$ slower than $1/\alpha(ET_N/\varphi(N))\varphi^2(N)$. Now, let us suppose that for our initial choice $\varphi^2(N)\cdot\alpha(ET_N/\varphi(N)) \not\to 0$. Otherwise the problem is solved). Then, let us take $\hat{\varphi}$ such that

$$\hat{\varphi}(N) \nearrow +\infty$$

$$\hat{\varphi}^2(N)\cdot\alpha\left(\frac{ET_N}{\varphi(N)}\right) \to 0$$

Since α is decreasing $\hat{\varphi}(N)^2\alpha(ET_N/\hat{\varphi}(N)) \to 0$ as $N \to +\infty$, and we define

$$R_N = ET_N/\hat{\varphi}(N) \quad \text{and}$$

$$m_N = \hat{\varphi}(N)\cdot\psi(N)$$

for $\psi(N)$ as discussed above. This completes the proof, if f is increasing. Same argument holds if f is decreasing. The general case follows since any cylindrical function can be written as a finite linear combination of such functions.

Proof of Theorem 3.1. After having proved Theorem 3.2 and Theorem 3.3, this result follows, on the same line of Theorem 2.1. ■

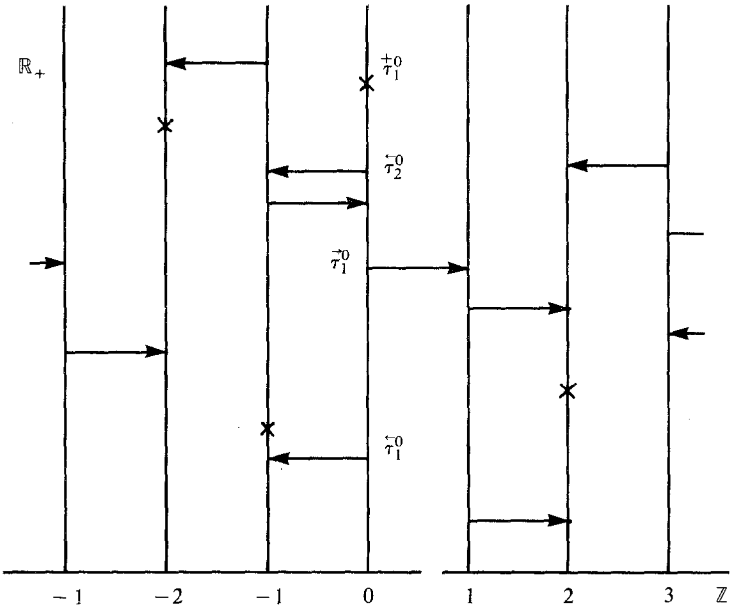

Fig. 2.

ACKNOWLEDGMENT

The discussions we had with Errico Presutti in the very beginning of this work were very important to clarify our basic ideas. We also thank Hermann Rost and François Ledrappier, whose criticism helped us to improve the first version of this text. M.C. and E.O. wish to thank the very kind hospitality at the IME-USP of São Paulo and at the IMPA of Rio de Janeiro where part of this work was done.

APPENDIX

Here we summarize some elementary results which we use.

Lemma A.1. Let $(\xi(t))_{t\in\mathbb{N}}$ be a discrete time Markov chain on $\{0, 1, \ldots, N\}$ with transition probabilities given by $Q_{i,j} = P(\xi(t+1) = j \mid \xi(t) = i)$, where

$$
\begin{aligned}
Q_{i,j} &= p_i && \text{if } j = i+1 \\
&= q_i && \text{if } j = i-1 \\
&= r_i = 1 - (p_i + q_i) && \text{if } j = i
\end{aligned}
$$

with $p_i, q_i > 0$ if $i = 1, \ldots, n-1$, $q_0 = 0$, $P_N = 0$, and $p_i + q_i + r_i = 1$.

Let $T_j = \inf\{t \geqslant 0 : \xi(t) = j\}$. Then for all $n \in \{1, \ldots, N\}$,

$$E_j(T_n) = \sum_{l=j}^{n-1} \left[\frac{1}{p_0} \prod_{i=1}^{l} \frac{q_i}{p_i} + \frac{1}{p_l} + \sum_{i=1}^{l-1} \frac{1}{p_i} \left(\prod_{s=i+1}^{l} \frac{q_s}{p_s} \right) \right] \quad \text{if} \quad j < n \tag{A1}$$

$$P_j(T_n < T_0) = \frac{\sum_{i=1}^{j-1} \prod_{k=1}^{i} \frac{p_k}{q_k}}{\sum_{i=1}^{n-1} \prod_{k=1}^{i} \frac{p_k}{q_k}} \qquad \text{if} \quad 1 \leqslant j < n \tag{A1'}$$

$$E_j(T_n) = \sum_{l=n+1}^{j} \left[\frac{1}{q_N} \left(\prod_{i=l}^{N-1} \frac{p_i}{q_i} \right) + \frac{1}{q_l} + \sum_{i=l+1}^{N-1} \frac{1}{q_i} \left(\prod_{s=l}^{i-1} \frac{p_s}{q_s} \right) \right] \qquad \text{if} \quad j > n \tag{A2}$$

where E_j denotes $E(\cdot \mid \xi(0) = j)$. Similarly for P_j.

Proof. It is obtained by the well-known method of difference equations. ■

Lemma A.2. Let $\xi(t)_{t \in \mathbb{N}}$ be an irreducible discrete time Markov chain, with values on a finite set E. Calling

$$\tau = \inf\{t \geqslant 1 : \xi(t) = x_0\}$$

Then

$$E_{x_0} \sum_{t=1}^{\tau} I_{\{y\}}(\xi(t)) = \mu\{y\} E_{x_0}(\tau) \tag{A3}$$

Proof. We shall show that if $f: E \to \mathbb{R}$

$$\sum_{x \in E} E_x[f(\xi(t))] \cdot E_{x_0}\left(\sum_{s=1}^{\tau} I_{\{x\}} \xi(s) \right)$$
$$= \sum_{x \in E} f(x) E_{x_0}\left[\sum_{s=1}^{\tau} I_{\{x\}}(\xi(s)) \right] \tag{A4}$$

But the right-hand side is equal to

$$f(x_0) + \sum_{k=1}^{\infty} \sum_{s=1}^{k-1} \sum_{x \neq x_0} f(x) P_{x_0}(\xi(k) = x_0\, \xi(s) = x,\, \xi(s') \neq x_0 \text{ if } s' < k)$$
$$= \sum_{s=1}^{\tau} E_{x_0} f(\xi(s))$$

Also, the left-hand side is equal to

$$\sum_{x\in E} E_x f(\xi(t)) E_{x_0} \sum_{s=1}^{\tau} I_{\{x\}}(\xi(s))$$

$$= E_{x_0} \sum_{s=1}^{\tau} E_{\xi(s)} f(\xi(t))$$

$$= E_{x_0} \sum_{s=1}^{\tau} f(\xi(t+s))$$

$$= E_{x_0} \sum_{u=1}^{\tau} \sum_{y} f(y) P_{x_0}(\xi(u)=y) + E_{x_0}\left[\sum_{u=\tau+1}^{\tau+t} f(\xi(u)) - \sum_{u=1}^{t} f(\xi(u))\right]$$

$$= E_{x_0} \sum_{u=1}^{\tau} f(\xi(u))$$

because

$$E_{x_0} \sum_{u=\tau+1}^{\tau+t} f(\xi(u)) = E_{x_0} \sum_{u=1}^{t} f(\xi(u))$$

This proves (A4) and (A3) follows immediately. ■

Lemma A3. Let us consider the process $\bar{\zeta}^N_{i_1(N)}$ of Section 2. Then, for $\bar{T}^N_{i_1(N),i_1(N)}$ as defined in Section 2, we have

$$E\left(\bar{T}^N_{i_1(N),i_1(N)}\right)^k \leqslant e^{c(k+1)N^{3/4}}$$

for all $k \geqslant 1$, and some constant c.

Proof. It can be obtained by the same method used in Lemma A1. ■

REFERENCES

1. P. Billingsley, *Convergence of Probability Measure* (John Wiley and Sons, New York, 1968).
2. D. Capocaccia, M. Cassandro, and E. Olivieri, A study of metastability in the Ising model, *Commun. Math. Physics* **39**:185–205 (1974).
3. R. Durret and D. Griffeath, Supercritical contact processes on $\mathbb{Z}^d$, *Ann. Probability* **11**:1–15 (1983).
4. W. Fleming and M. Viot, Some measure-valued Markov processes in population genetics, *Indiana Univ. Math. J.* **28**:817 (1979).
5. D. Griffeath, *Additive and Cancellative Interacting Particle Systems* (Lecture Notes in Mathematics, No. 724. Springer, New York, 1979).
6. D. Griffeath, The basic contact processes, *Stochastic Processes and Their Applications* **11**:151–185 (1981).

7. R. Griffiths, C. Y. Weng, and J. S. Langer, Relaxation times for metastable states in the mean-field model of a ferromagnet, *Phys. Rev.* **149**:301–305 (1966).
8. T. E. Harris, Contact interactions on a lattice, *Ann. Probability* **2**:969–988 (1974).
9. T. E. Harris, Additive set valued Markov processes and graphical methods, *Ann. Probability* **6**:355–378 (1978).
10. J. L. Lebowitz and O. Penrose, Rigorous treatment of the Van der Walls–Maxwell theory of liquid–vapor transition, *J. Math. Phys.* **7**:98–113 (1966).
11. J. L. Lebowitz and O. Penrose, Rigorous treatment of metastable states in the Van der Walls–Maxwell theory, *J. Stat. Phys.* **3**:211–236 (1971).
12. A. D. Ventsel and M. L. Freidlin, On small random perturbations of dynamical systems, *Russ. Math. Surv.* **25**:1–55 (1970).