Journal of Statistical Physics, Vol. 73, Nos. 1/2, 1993

Renormalization Group for Markov Chains and Application to Metastability

Elisabetta Scoppola[1]

Received December 31, 1993; final May 12, 1993

In this paper we introduce a new renormalization group method for the study of the long-time behavior of Markov chains with finite state space and with transition probabilities exponentially small in an external parameter β. A general approach of metastability problems emerges from this analysis and is discussed in detail in the case of a two-dimensional Ising system at low temperature.

KEY WORDS: Markov chains; renormalization group; metastability; Ising model; Monte Carlo dynamical simulations.

1. INTRODUCTION

In this paper we study the long-time behavior of Markov chains with finite state space S and with transition probabilities satisfying the following condition: for any $x, y \in S$, with $x \neq y$, if $P(x, y) > 0$, then

$$\exp\{-\Delta(x, y)\beta - \gamma\beta\} \leqslant P(x, y) \leqslant \exp\{-\Delta(x, y)\beta + \gamma\beta\} \tag{1.1}$$

where $\Delta(x, y)$ assumes the values $\Delta_0 = 0 < \Delta_1 < \Delta_2 < \cdots < \Delta_n$, and $\gamma = \gamma(\beta) \to 0$ as $\beta \to \infty$. We will consider the case of β very large with respect to the number of states $|S|$.

Stochastic dynamics satisfying this assumption are, for instance, Monte Carlo dynamical simulations of statistical models in the low-temperature regime (as discussed in Section 3), or Markov chains related to small random perturbations of dynamical systems (as explained in the Wentzell and Freidlin theory[1]).

We introduce a new renormalization group method for the study of these Markov chains, inspired by the analysis of metastability for the

---

[1] Dipartimento di Fisica, Università "La Sapienza," 00185 Rome, Italy. E-mail: SCOPPOLA@ROMA1.INFN.IT.

0022-4715/93/1000-0083$07.00/0

two-dimensional Ising model in the low-temperature regime, developed in collaboration with F. Martinelli and E. Olivieri in ref. 8 (see also ref. 12).

Let me consider here a very simple example in order to illustrate the problem and the main ideas of the paper.

Let $S = \{1, 1,..., N\}$ be our state space, and to each state i we associate a nonnegative energy function $H(i)$; the stochastic evolution on S is defined by a Metropolis algorithm: for each $i \neq j$

$$P(i, j) = \begin{cases} 0 & \text{if } \; |i-j| > 1 \\ \frac{1}{2} \exp\{-[H(j) - H(i)]\beta & \text{if } |i-j| = 1 \text{ and } H(j) > H(i) \\ \frac{1}{2} & \text{if } |i-j| = 1 \text{ and } H(j) \leqslant H(i) \end{cases} \tag{1.2}$$

and $P(i, i)$ is fixed by normalization:

$$P(i, i) = 1 - \sum_{j \neq i} P(i, j) \tag{1.3}$$

The parameter β is very large with respect to N. See Fig. 1 for an explicit example. We are considering in this way a kind of random walk on $1,..., N$ with a drift related to the gradient of H.

Suppose we now have to study the long-time behavior of this chain: consider for instance the problem of the determination of the invariant measure of the chain. In the case of the present example the solution is very simple; in fact it is easy to verify that the invariant measure of this chain is given by

$$\nu(i) = \frac{e^{-\beta H(i)}}{\sum_{j=1}^{N} e^{-\beta H(j)}} \tag{1.4}$$

which is mostly concentrated in the states of minimal energy. If we look for a solution of the problem applicable to more general cases, in order words, without using the energy function H, the main difficulty of the problem is then in its multiple tunneling structure. In fact, in the case of large β we cannot use the ergodic theorem, that is, the Monte Carlo method, since the process is trapped in the states corresponding to local minima of the energy for times which are exponentially long in β. We propose here a different approach to the problem based on a multiscale analysis of Markov chains satisfying (1.1) which enables us to control the convergence to equilibrium of the chain.

In order to study the long-time behavior of the chain in a general case we consider a classification of the states in S based on the stability of the states on a sequence of time scales. We will show that such a classification turns out to be different from that suggested in our example by the energy

$H(i)$. For each state i let $E\sigma(i)$ be the mean time necessary to leave the state i. In our example these mean exit times are exponential in the minimal energy differences

$$\min([H(i+1)-H(i)] \vee 0, [H(i-1)-H(i)] \vee 0)$$

where $a \vee b = \max(a, b)$. With such a quantity we can immediately discriminate between the states which are completely unstable and those corresponding to local energy minima. In fact, due to the structure of the transition probability (1.1) there will be states which are completely unstable in the sense that their mean exit time is of order one [in our example the states i such that $H(i) > H(j)$ for $j = i+1$ or $i-1$], and there will be stable states for which the mean exit time is exponentially long in β (in our example the local energy minima). This means that starting from a state x, in a rather short time the process typically (that is, with large probability) reaches a stable state x_1 (or remains in the state x if this is stable and $x = x_1$ in this case) where it spends a very long time, exponential in β; after this time, that is, under the effect of a large fluctuation, the process leaves the state x_1 and reaches in a very short time a stable state x_2 (eventually $x_2 = x_1$, that is, the process goes back to x_1) and so on; unstable states are visited only very briefly during these transitions.

To go on with our classification, we define a first time scale t_1 corresponding to the smaller mean exit time from stable states and we look at our chain on this new time scale. Only the stable states are relevant now, since the ratio of the time spent by the process in unstable states during an interval of time of length t_1, with respect to the time spent in stable states, is exponentially small in β. Among the stable states with mean exit time of order t_1 there will be states which become unstable on this new time scale in the sense that the process typically leaves these states and reaches a different stable state in a time of order $s \cdot t_1$ with s of order one.

In the present paper we realize this time rescaling by defining a new Markov chain, corresponding to the original process observed at time intervals of order t_1, on a new state space containing only the states which were stable on the original time scale. For this new chain we can prove estimates for the transition probabilities of the same kind as (1.1) and thus we can iterate our analysis.

In this way at each step of the iteration we rescale the time by a factor exponentially large in β and we renormalize the chain by considering only the states which are stable on the new time scale.

The goal of this paper is to implement this rescaling in such a way that quantities related to the long-time behavior of the original Markov chain, e.g., the invariant measure in our example, can be estimated in terms of the new renormalized chain. In other words, we develop here a new approach to the study of multiple tunneling problems in the general case, based on a recursive procedure.

We want to note that the time rescaling proposed in this paper is different from the simulated annealing procedure, where the temperature is slowly decreased during the evolution of the process.[5] However, the detailed description of multiple tunneling arising from our analysis can open interesting perspectives from the point of view of applications.

Let me come back to the example with the particular choice of the energy function H shown in Fig. 1. Figure 1a represents all the states $\{1,..., 20\}$ and all the possible transitions of the chain by means of arrows; the numbers associated to each arrow are the value of the quantity $\Delta(i, j)$ (in unity of energy).

Let M be the set of stable states with respect to the definition given above; this set is represented by the points in Fig. 1b. Among these states the less stable ones are 2 and 7, with

$$E\sigma(2) \sim E\sigma(7) \sim e^{\beta} = t_1$$

We will show that the chain represented in Fig. 1a observed at time of order t_1 is a new Markov chain represented in Fig. 1b with transition

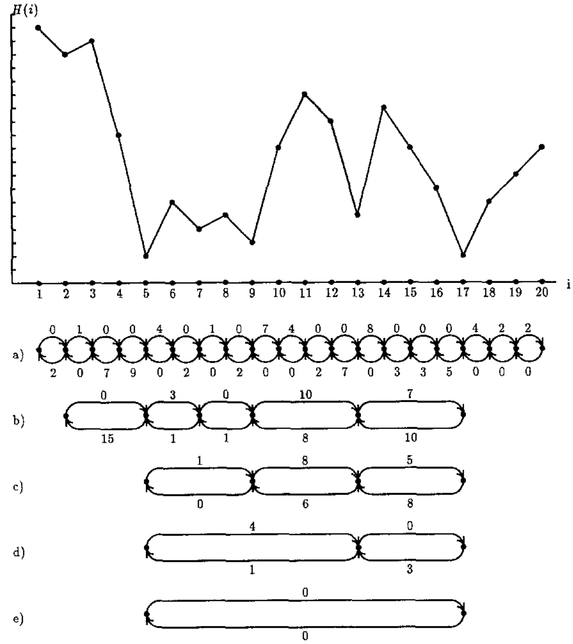

Fig. 1. The renormalization procedure in a simple example.

probabilities $P^{(1)}(x, y)$ different from zero only if x and y are nearest neighbor in M and satisfying (1.1) with

$$\Delta^{(1)}(x, y) = \sum_{i=0}^{n} \Delta(x_i, x_{i+1}) - 1 \tag{1.5}$$

where $x_0 = x$, $x_n = y$, and $|x_i - x_{i+1}| = 1$, with $n = |y - x| - 1$ and where the term -1 represents the entropic factor due to the time rescaling, that is, $1 = (1/\beta) \ln t_1$. The values of $\Delta^{(1)}(x, y)$ are represented in Fig. 1b by the numbers associated to the arrows.

In Figs. 1c and 1d we have the successive steps of the iteration up to Fig. 1e where, under the effect of the rescaling, the initial multiple-well problem is reduced to a double well.

We note that in this construction the states are classified in terms of stability on different time scales and in this classification states of higher energy can be more stable than states of lower energy for a given time scale (see in the example the states 13 and 9).

The main difficulty in the implementation of this rescaling program is that in general, that is, without additional assumptions on the chain, we will not be able to define the new Markov chain corresponding to the rescaled time and satisfying the hypothesis (1.1), on the same probability space by means of a construction path by path, as one might expect, but we have to consider as new state space the set of stable states modulo an equivalence relation which will be precisely defined in the next section. This implies that the new renormalized Markov chain is defined on a different probability space and that the relation between the original chain and the new one is not immediate as one might hope. However, it is possible to state results relating the two chains and which are sufficient to study the long-time behavior of the original process.

In Section 2 the renormalization procedure is described in the general case, while in Section 3 the general scheme is applied to the study of metastable states for the two-dimensional Ising model by considering the Metropolis algorithm, and the results proved by Schonmann and Neves[3,4] in this case are rederived. Similar results may be obtained for different choices of the algorithm as well, e.g., the Kawasaki algorithm, but in this paper, for the sake of brevity, we will consider only the Metropolis case.

In Section 4 we discuss the relation between our approach and some known results on Markov chains.

2. THE RENORMALIZATION PROCEDURE

For the reader's convenience this section is divided into several parts.

2.1. Hypotheses

We consider a Markov chain $\{X_t\}_{t=0,1,2,\dots}$ on a finite state space S with transition probabilities $P(x, y)$ satisfying the ergodicity condition:

$$\exists n \quad \text{such that} \quad \forall x, y \in S \quad P^n(x, y) > 0$$

[where $P^n(x, y)$ is the n-step transition probability] and also satisfying the following additional property:

Property P. There exists a positive parameter β, a function $\Delta(x, y),, x, y \in S$, assuming values $\Delta_0 = 0 < \Delta_1 < \Delta_2 < \cdots < \Delta_n$, for some positive integer n, with $\Delta_n < \infty$, and a positive function $\gamma = \gamma(\beta)$, with $\gamma \to 0$ and $\beta \to \infty$, such that if $x \neq y$ and $P(x, y) > 0$, then

$$\exp\{-\Delta(x, y)\beta - \gamma\beta\} \leqslant P(x, y) \leqslant \exp\{-\Delta(x, y)\beta + \gamma\beta\} \tag{2.1}$$

We will denote by $X_t(x)$ the process starting at x at time 0.

2.2. The Large-Deviations Functional

Following the ideas developed in ref. 1 for the study of diffusion processes in the small-diffusion regime, to each function $\phi: \mathbf{N} \to S$, $\phi = \{\phi_t\}_{t \in \mathbf{N}}$, we associate a functional

$$I_{[0,t]}(\phi) \equiv \sum_{i=0}^{t-1} \Delta(\phi_i, \phi_{i+1}) \tag{2.2}$$

where we define $\Delta(x, x) = 0$ for each $x \in S$ and $\Delta(x, y) = \infty$ if $P(x, y) = 0$. The following large-deviation estimates are easily proved:

Proposition 2.1. Let ϕ be a fixed function starting at x at time 0; then:

(i) We have

$$P(X_s(x) = \phi_s \qquad \forall s \in [0, t]) \leqslant e^{-I_{[0,t]}(\phi)\beta + \gamma t\beta}$$

(ii) If ϕ is such that $\phi_s \neq \phi_{s+1}$ for any $s \in [0, t]$, then we have also a lower bound:

$$P(X_s(x) = \phi_s \qquad \forall s \in [0, t]) \geqslant e^{-I_{[0,t]}(\phi)\beta - \gamma t\beta}$$

(iii) For any constant $r > 0$ and for any $t < e^{\alpha\beta}$ with $\alpha < \Delta_1$

$$\sup_x P(I_{[0,t]}(X_s(x)) \geqslant r) \leqslant e^{-(r \vee \Delta_1)\beta + \varepsilon\beta}$$

where $r \vee \Delta_1 \equiv \max\{r, \Delta_1\}$, with

$$\varepsilon = \frac{3r}{\Delta_1}\left(\gamma + \frac{\ln t}{\beta}\right)$$

We remark that here, and later on, the sup is actually a max (and the inf is a min).

Proof. Points (i) and (ii) are trivial by the definition (2.2) and by Property P.

Point (iii) is also very simple: suppose $r < \Delta_1$; if $I_{[0,t]}(X_s(x)) \geqslant r$, this implies that there exists a time $s \leqslant t$ such that $\Delta(X_s, X_{s+1}) = \Delta_1$ and then

$$\sup_x P(I_{[0,t]}(X_s(x)) \geqslant r) \leqslant t e^{-\Delta_1 \beta + \gamma\beta}$$

In the case $r > \Delta_1$ if $I_{[0,t]}(X_s) \in [jr, (j+1)r), j = 1, 2,...,$ this implies that for each j there exist times $0 \leqslant s_1 < s_2 < \cdots < s_q \leqslant t$ such that $\Delta(X_{s_i}, X_{s_{i+1}}) = r_i$ with $\sum_{i=1}^q r_i = I_{[0,t]}(X_s) \geqslant jr$ and since

$$q \leqslant \frac{I_{[0,t]}(X_s)}{\Delta_1} \leqslant \frac{(j+1)r}{\Delta_1}$$

we obtain

$$\begin{aligned}\sup_x P(I_{[0,t]}(X_s(x)) \geqslant r) &\leqslant \sum_{j=1}^{\infty} P(I_{[0,t]}(X_s(x)) \in [jr, (j+1)r) \\ &\leqslant \sum_{j=1}^{\infty} e^{-jr\beta + (\gamma\beta + \ln t)(j+1)r/\Delta_1} \\ &\leqslant e^{\varepsilon\beta/3} \frac{e^{(-r+\varepsilon/3)\beta}}{1 - e^{(-r+\varepsilon/3)\beta}} \leqslant e^{-r\beta + \varepsilon\beta}\end{aligned}$$

2.3. The Stable States

Following the scheme of ref. 1, we can define an equivalence relation in the state space S by using the functional (2.2). For each couple of states x and y we define

$$V(x, y) \equiv \inf_{t, \phi; \phi_0 = x, \phi_t = y} I_{[0,t]}(\phi) \tag{2.3}$$

and

$$x \sim y \quad \text{iff} \quad V(x, y) = V(y, x) = 0 \tag{2.4}$$

We say that x is a minimum if and only if

$$\text{for any} \quad y \not\to x \quad V(x, y) > 0 \tag{2.5}$$

We will denote by M the set of local minima and we will show that the states in M turn out to be the most probable ones.

For each $x \in S$ we can define the first hitting time to the set M:

$$\tau_M(x) \equiv \inf\{t \geqslant 0; X_t(x) \in M\} \tag{2.6}$$

corresponding to the time spent by the process outside the set M. We have the following:

Proposition 2.2. Let $\delta = 2\gamma\,|S|$; there exist constants $T_0 \in [0, |S|]$ and β_0 such that for any $\beta > \beta_0$:

(i) For any $t > T_0$:

$$\sup_{x \in S} P(\tau_M(x) > t) \leqslant a^{[t/T_0]}$$

with $a = 1 - C^{T_0}$ for some constant $0 < C < 1$ and where $[\cdot]$ denotes the integer part.

(ii) For any $t \geqslant e^{\delta\beta}$

$$\sup_{x \in S} P(\tau_M(x) > t) \leqslant \exp\{-\exp(\delta\beta/2)\}$$

Remark. We wish to note that this proposition is different from the analogous result obtained in the case of a small random perturbation of dynamical systems in ref. 1.

In fact in the present case the attractiveness of the minima is much weaker; this is a consequence of the fact that in ref. 1 the process without random noise corresponds to a deterministic dynamical system for which the set M is the set of ω-limit sets, that is, in the case of zero noise the deterministic solutions have a deterministic and finite exit time from every set not containing any ω-limit set. In the present case this is no longer true and in the case of zero noise ($\beta = \infty$) arbitrarily long trajectories can exist which never visit the set M.

Proof. First of all we want to define the possible classical time T_0 in which the process falls in one of the minima, where classical means in the limit $\beta \to \infty$. For each $x \in S$ let

$$T_0(x) = \inf\{t \geqslant 0; \exists\phi; \phi_0 = x, \phi_t \in M \text{ with } I_{[0,t]}(\phi) = 0\} \tag{2.7}$$

and

$$T_0 = \sup_x T_0(x) \tag{2.8}$$

We first show that these times are well defined and that $T_0 < |S|$. It is sufficient to show that for any x in S there exist $\bar{x} \in M$, a time $T_0(x) < |S|$, and a function ϕ such that $\phi_0 = x$, $\phi_{T_0(x)} = \bar{x}$ and $I_{[0, T_0(x)]}(\phi) = 0$:

if $x \in M$, then $T_0(x) = 0$
if $x \notin M$, then there exists x_1 such that $V(x, x_1) = 0$ and $V(x_1, x) > 0$
if $x_1 \in M$, then $\bar{x} = x_1$, $T_0(x)$ is equal to the time minimizing $V(x, x_1)$
if $x_1 \notin M$, then there exists x_2 such that $V(x_1, x_2) = 0$ and $V(x_2, x_1) > 0$
...
if $x_{n-1} \in M$, then $\bar{x} = x_{n-1}$, $T_0(x)$ is equal to the time minimizing $V(x, x_{n-1}) = 0$
if $x_{n-1} \notin M$, then there exists x_n such that $V(x_{n-1}, x_n) = 0$ and $V(x_n, x_{n-1}) > 0$

The states x_j in this construction must be all different; in fact, if $x_k = x_n$ with $k \leqslant n$, then the following would be true:

$$V(x_{k+1}, x_k) = V(x_{k+1}, x_n) \leqslant V(x_{k+1}, x_{k+2}) + \cdots + V(x_{n-1}, x_n) = 0$$

which contradicts $V(x_{k+1}, x_k) > 0$.

Thus we can conclude that the sequence x_j must be finite, and it stops at $\bar{x} \in M$; by the construction we have $V(x, \bar{x}) = 0$ and then if we choose $T_0(x)$ as the time minimizing $V(x, \bar{x})$ and ϕ the corresponding function, we have proved that $T_0 < |S|$ is well defined. We observe also that this function ϕ, since it is a minimizing function, has the property that $\phi_s \neq \phi_t$, $\forall s \neq t$, s, $t \in [0, T_0(x)]$. We can then apply point (ii) of Proposition 2.1, obtaining

$$\inf_x P(X_s(x) = \phi_s \qquad \forall s \in [0, T_0(x)]) \geqslant e^{-\gamma T_0 \beta}$$

The conclusion of the proof now is very simple: in order to have the estimate (i), we have to choose $C = e^{-\gamma\beta}$. By applying point (i) with $t \geqslant e^{\delta\beta}$ we obtain point (ii):

$$\sup_x P(\tau_M(x) > t) \leqslant \exp\left\{\left[\frac{T}{T_0}\right] \ln(1 - \exp(-\gamma T_0 \beta))\right\} \leqslant \exp\left\{-\exp\frac{\delta\beta}{2}\right\}$$

Corollary 2.1. Let $\alpha \in (\delta, \Delta_1/2 - \gamma)$; for any $t > e^{\alpha\beta} \equiv T$ we have

$$\sup_x P(X_t(x) \notin M) \leqslant e^{-(\Delta_1/2)\beta}$$

Proof. We have

$$\sup_x P(X_t(x) \notin M) \leqslant \sup_x P(I_{[t-T, t]}(X_s(x)) \geqslant \Delta_1)$$
$$+ \sup_x P(X_t(x) \notin M \cap I_{[t-T, t]}(X_s) = 0)$$

The first term is estimated by Proposition 2.1(iii) and the second one by Proposition 2.2(ii).

2.4. The Rescaled Markov Chain $\bar{X}_t$

These results suggest that if we look at the process X_t on a sufficiently large time scale, then it can be described in terms of transitions between states in M; in this way only the behavior of the process on small times is neglected. More precisely, we will construct a new Markov chain with state space M, corresponding to the original process looked at at times sufficiently large. Let us define

$$V_1 \equiv \inf_{x \in M,\, y \in S,\, x \nrightarrow y} V(x, y) \tag{2.9}$$

$$t_1 \equiv e^{V_1 \beta + \delta \beta} \tag{2.10}$$

Let θ be the shift operator for the Markov chain X_n defined by

$$\theta(\{x_0, x_1,..., x_n,...\}) = \{x_1, x_2,..., x_{n+1},...\} \tag{2.11}$$

We will denote by θ_p the p^{th} power pf θ: $\theta_p = \theta \circ \theta \circ \cdots \circ \theta$ p times. We define now recursively a sequence of stopping times:

$$\sigma \equiv \inf\{t > 0; X_t \nrightarrow X_0\} \tag{2.12}$$

$$\tau \equiv \inf\{t \geqslant \sigma; X_t \in M\} \tag{2.13}$$

$$\zeta_1 = \begin{cases} t_1 & \text{if} \quad \sigma > t_1 \\ \tau & \text{if} \quad \sigma \leqslant t_1 \end{cases} \tag{2.14}$$

and for each $n > 1$

$$\zeta_n = \zeta_{n-1} + \zeta_1 \circ \theta_{\zeta_{n-1}} \tag{2.15}$$

These times are stopping times with respect to the σ-algebra associated to the Markov chain. In fact σ and τ are stopping times[11]; we have to prove that ζ_1 is also a stopping time. If we denote by F_t the σ-algebra generated by the process X_t up to time t, then it is easy to show that the event $\{\zeta_1 \leqslant t\}$ belongs to the σ-algebra F_t:

$$\{\zeta_1 \leqslant t\} = \{\sigma > t_1\} \cup [\{\sigma \leqslant t_1\} \cap \{\tau \leqslant t\}]$$

if $t \geqslant t_1$, and all these events belong to F_t; if $t < t_1$, then the event $\{\zeta_1 \leqslant t\} = \{\tau \leqslant t) \in F_t$.

It is easy to prove[11] that if ζ_1 is a stopping time, then ζ_n are stopping times and that the sequence $\bar{X}_n = X_{\zeta_n}$ is an homogeneous Markov chain. For any $x \in M$, we can then consider the new Markov chain $\bar{X}_n$ with $\bar{X}_0(x) = x$.

Before giving an estimate on the transition probabilities of this chain, let me comment on the relation between this chain and the chain which could be considered, in analogy with the work by Wentzell and Freidlin,[1] by looking at the original process at the jumping times between minima, by using only the stopping times σ and τ corresponding to the first exit from the equivalence class of a minimum and to the first subsequent hitting time to the set of the minima. Also with this chain one can study the long-time properties of the process X_t by adapting the techniques and the results given in ref. 1. The crucial difference in the definition of the new Markov chain $\{\bar{X}_n\}$ considered in this paper is that it is strictly related to the time scale t_1 in the sense that it essentially corresponds to the original process viewed on this time scale and this turns out to be crucial in order to classify different minima in terms of their stability, as we will clearify later.

For any couple of states $x, y \in M$ we denote by $\bar{P}(x, y)$ the transition probability of the chain $\bar{X}_n$, that is,

$$\bar{P}(x, y) = P(X_{\zeta_n} = y \mid X_{\zeta_{n-1}} = x) \tag{2.16}$$

We can prove the following:

Proposition 2.3. There exists $\beta_0 > 0$ such that for any $\beta > \beta_0$ and for any $x, y \in M$ with $x \not\sim y$ we have the following:

(a) If for any time $t < |S|$ and any function ϕ such that $\phi_0 = x$, $\phi_t = y$, and $\phi_s \notin M_{x,y}$ for any $s \in (0, t)$, there exists $s' < t$ such that $P(\phi'_s, \phi_{s'+1}) = 0$, then

$$\bar{P}(x, y) = 0$$

where $M_{x,y}$ is the set of the minima with the exception of the states x, y and all the states which are equivalent to x and y.

(b) Otherwise, if the quantity

$$\bar{\Delta}(x, y) = \inf_{t, \phi; \phi_0 = x, \phi_t = y, \phi_s \notin M_{x,y} \forall s \in (0, t)} I_{[0,t]}(\phi) \tag{2.17}$$

is well defined, then

$$t_1 \exp\{-\bar{\Delta}(x, y)\beta - \bar{\gamma}\beta\} \leqslant \bar{P}(x, y) \leqslant t_1 \exp\{-\bar{\Delta}(x, y)\beta + \bar{\gamma}\beta\} \tag{2.18}$$

The quantity $\bar{\Delta}(x, y)$ can assume the values $\bar{\Delta}_1 = V_1 < \bar{\Delta}_2 < \cdots < \Delta_{\bar{n}}$ with $\bar{\Delta}_{\bar{n}} < |S| \Delta_n$ and it is invariant with respect to the equivalence relation in the sense that $\bar{\Delta}(x, y) = \bar{\Delta}(x', y')$ if $x \sim x'$ and $y \sim y'$. The quantity $\bar{\gamma}$ is given by the following:

$$\bar{\gamma} = \left[(|S| + 1)\gamma + \delta + \frac{1 + \ln 2|S|}{\beta} \right] \vee \left[\frac{|S| \Delta_n}{\Delta_1} (\gamma + \delta) + \delta \right]$$

which implies that $\bar{\gamma} \to 0$ as $\beta \to \infty$.

Proof. (a) If there does not exist a function ϕ and a time $t \leqslant |S|$ with $\phi_0 = x$, $\phi_t = y$, and $\phi_s \notin M_{x,y}$ for any $s \in (0, t)$ such that $P(\phi_s, \phi_{s+1}) > 0$, $\forall s \leqslant t$, then obviously

$$\bar{P}(x, y) = P(X_{\zeta_n} = y \mid X_{\zeta_{n-1}} = x) \leqslant \sum_t \sum_\phi P(X_s = \phi) = 0$$

where the second sum is taken over all the functions ϕ such that $\phi_0 = x$, $\phi_t = y$, and $\phi_s \notin M_{x,y}$ for any $s \in (0, t)$.

(b) First of all we prove that the quantity $\bar{\Delta}(x, y)$ defined by (2.17) is invariant with respect to the equivalence relation (2.4), that is, that $\bar{\Delta}(x, y) = \bar{\Delta}(x', y')$ if $x \sim x'$ and $y \sim y'$. Suppose that this is not true and, for instance,

$$\bar{\Delta}(x, y) < \bar{\Delta}(x', y') \tag{2.19}$$

Then this implies that there exists a time t and a function ϕ going from x' to y' in the time t without touching any other minimum in $M_{x,y}$ and such that $I_{[0,t]}(\phi) < \bar{\Delta}(x', y')$, contradicting the definition of $\bar{\Delta}(x', y')$. In fact this function can be constructed by considering the ordered sequence of the following three functions: the function minimizing $V(x', x)$, then the function minimizing $\bar{\Delta}(x, y)$, and finally the function minimizing $V(y, y')$.

Estimate from below: Let $x \not\sim y$ and let $\bar{t}_1 = e^{V_1\beta - \gamma\beta}$; we denote by $\bar{\phi}_s^{x,y}$ the function going from x to y minimizing the quantity $\bar{\Delta}(x, y)$ and by $\bar{t}^{x,y}$ the corresponding time for which we have the trivial and crude estimate $\bar{t}^{x,y} < T \equiv |S|$. We have

$$\bar{P}(x, y) \geqslant \sum_{n=1}^{[\bar{t}_1/T]} \sum_{z \sim x} P_x(\sigma > nT \cap X_{nT} = z \cap X_s = \bar{\phi}_s^{z,y}\ \forall s \in [nT, nT + \bar{t}^{z,y}])$$

$$\geqslant \sum_{n=1}^{[\bar{t}_1/T]} P_x(\sigma > nT) \exp\{-\bar{\Delta}(x, y)\beta - \gamma T\beta\} \tag{2.20}$$

by Proposition 2.1(ii) and by using the fact that $X_{nT}(x) \sim x$ if $\sigma > nT$. On the other hand, by the Markov property we have

$$P_x(\sigma > 1) \geqslant 1 - e^{-V_1\beta + \gamma\beta}$$
$$P_x(\sigma > s) \geqslant P_x(\sigma > s-1)(1 - e^{-V_1\beta + \gamma\beta})$$

that is,

$$P_x(\sigma > s) \geqslant (1 - e^{-V_1\beta + \gamma\beta})^s$$

which implies

$$P_x(\sigma > nT) \geqslant P_x(\sigma > \bar{t}_1) \geqslant (1 - e^{-V_1\beta + \gamma\beta})^{\bar{t}_1} \geqslant \frac{e^{-1}}{2} \tag{2.21}$$

for β sufficiently large. Estimate (2.21) in (2.20) gives

$$\bar{P}(x, y) \geqslant [\bar{t}_1/T](e^{-1}/2) \exp\{-\bar{\Delta}(x, y)\beta - \gamma T\beta\} \geqslant t_1 \exp\{-\bar{\Delta}(x, y)\beta - \bar{\gamma}_1\beta\}$$

with

$$\bar{\gamma}_1 = (T+1)\gamma + \delta + \frac{1 + \ln 2T}{\beta}$$

Estimate from above: We have

$$\bar{P}(x, y) \leqslant P(X_\tau(x) = y \cap \tau - \sigma \leqslant e^{\delta\beta}) + P(\tau - \sigma > e^{\delta\beta}) \tag{2.22}$$

The first term in the rhs of (2.22) is estimated as follows:

$$\begin{aligned}
&\leqslant \sum_{s=1}^{t_1} P(X_u \sim x \; \forall u \leqslant s \cap I_{[s, s+e^{\delta\beta}]}(X_t(x)) \geqslant \bar{\Delta}(x, y)) \\
&\leqslant \sum_{s=1}^{t_1} \sum_{z \sim x} P(I_{[0, e^{\delta\beta}]}(X_t(z)) \geqslant \bar{\Delta}(x, y)) \\
&\leqslant t_1 \exp\{-\bar{\Delta}(x, y)\beta + \varepsilon\beta\}
\end{aligned} \tag{2.23}$$

from Proposition 2.1(iii) with $\varepsilon = \{\bar{\Delta}(x, y)/\Delta_1\}[\gamma + \delta]$. The second term in the rhs of (2.22) by Proposition 2.2(ii) has a superexponential estimate, giving the following global estimate of (2.22):

$$\leqslant t_1 \exp\{-\bar{\Delta}(x, y) + \bar{\gamma}_2\}$$

with $\bar{\gamma}_2 \leqslant \{\bar{\Delta}(x, y)/\Delta_1\}[\gamma + \delta] + \delta$.

The proposition is proved by choosing $\bar{\gamma} = \max(\bar{\gamma}_1, \bar{\gamma}_2)$.

Remarks. 1. The estimates on the time T_0 appearing in Propostion 2.2 and on the times corresponding to functions minimizing the functionals $\bar{\Delta}(x, y)$ in terms of the cardinality of the state space are clearly very crude estimates which can be improved in concrete situations.

2. We observe that the states which are not minima, $S \backslash M$, are involved only in the definition of the quantity $\bar{\Delta}(x, y)$; such quantities are the analog of the quantities $\tilde{V}_D(K_i, K_j)$ defined in ref. 1 (see Chapter 6).

3. We finally remark that the estimates given in Proposition 2.3 are invariant with respect to the equivalence relation (2.4).

2.5. The Renormalized Markov Chain $\tilde{X}_1 = X_t^{(1)}$

From this last remark we are tempted to simplify our problem furthermore by reducing the state space from M to the space of the equivalence classes corresponding to the states in M: $M/_{\sim} \equiv \tilde{M}$, and we expect that it is possible to define a new Markov chain on the space $\tilde{M}$ with transition probabilities satisfying property P and thus we would have the possibility to iterate our argument. However, such a program is not attainable in the most direct way, because of the following remark. Consider our state space $M = m_1 \cup m_2 \cup \cdots \cup m_r$, where m_i are equivalence classes for the relation (2.4), that is,

$$\forall x, y \in m_i \qquad x \sim y$$

and

$$\forall x \in m_i, \quad y \in m_j \qquad i \neq j \qquad x \not\sim y$$

and consider our chain $\bar{X}_t$. If we define a new process

$$Y_t = i \qquad \text{iff} \quad \bar{X}_t \in m_i \tag{2.24}$$

on the state space $\tilde{M} = \{1, 2,..., r\}$, then we can ask if such a process is a Markov chain. Unfortunately I cannot prove this except in the case in which for each m_i

$$\sum_{y \in m_i} \bar{P}(x, y) = \sum_{y \in m_i} \bar{P}(x', y) \qquad \forall x' \sim x \tag{2.25}$$

This relation is "almost true" in the sense that by Proposition 2.3, $\bar{P}(x, y)$ and $\bar{P}(x', y)$ satisfy estimate (2.18) with $\bar{\Delta}(x, y) = \bar{\Delta}(x', y)$, but this is not enough to prove the Markov property for Y_t. An analogous problem is present also in the case of the chain constructed in ref. 1. So a simplification of the process is not allowed in this strong sense of construction of a new Markov chain path by path as suggested by (2.24).

On the state space $\tilde{M}$ we can define a chain $\tilde{X}_t$ with transition probabilities

$$\tilde{P}(i, j) = \frac{1}{\bar{v}(m_i)} \sum_{x \in m_i} \bar{v}(x) \sum_{y \in m_j} \bar{P}(x, y) \tag{2.26}$$

where $\bar{v}$ denotes the invariant measure of the chain $\bar{X}_t$. Here we use the ergodicity condition for the chain X_t in order to have $\bar{v}(m_i) > 0$, $\forall i$.

For such transition probabilities the estimate (2.18) is obviously true, and the invariant measure of the chain $\tilde{X}_t$ is obviously $\{\bar{v}(m_i)\}_{i=1,\dots,r}$. Let $\tilde{W} \subset \tilde{M}$, $\tilde{B} \subset \tilde{W}$, and W and B be the corresponding sets in M ($W = \bigcup_{i \in \tilde{W}} m_i$); we denote by $\bar{q}_W(x, B)$ the probability that the process $\bar{X}_t$ starting at x, at the first entrance time in W hits the set B, and by $\tilde{q}_{\tilde{W}}(i, \tilde{B})$ the analogous quantity for the process $\tilde{X}_t$. The following two lemmas can be easily proved by applying the results of Lemma 3.3 and 3.4 of ref. 1 (Chapter 6, pp. 179–183) to the process $\bar{X}_t$ and to the process $\tilde{X}_t$:

Lemma 2.1. For any $x \in m_i$, $i \in \tilde{M} \backslash \tilde{W}$, $j \in \tilde{W}$,

$$\{\exp(-4r\bar{\gamma}\beta)\}\, \tilde{q}_{\tilde{W}}(i, j) \leqslant \bar{q}_W(x, m_j) \leqslant \{\exp(4r\bar{\gamma}\beta)\}\, \tilde{q}_{\tilde{W}}(i, j)$$

For the Markov chain $\bar{X}_t$ we denote by $\bar{E}_x \bar{\tau}_W$ the mathematical expectation of the number of steps until the first entrance in W, calculated under the assumption that the initial state of the chain is x, and for the chain $\tilde{X}_t$ we denote the corresponding quantity by $\tilde{E}_i \tilde{\tau}_{\tilde{W}}$.

Lemma 2.2. For any $x \in m_i$, $i \in \tilde{M} \backslash \tilde{W}$,

$$\{\exp(-4r\bar{\gamma}\beta)\}\, \tilde{E}_i \tilde{\tau}_{\tilde{W}} \leqslant \bar{E}_x \bar{\tau}_W \leqslant \{\exp(4r\bar{\gamma}\beta)\}\, \tilde{E}_i \tilde{\tau}_{\tilde{W}}$$

In fact the quantities $\tilde{q}_{\tilde{W}}(i, j)$ and $\tilde{E}_i \tilde{\tau}_{\tilde{W}}$ can be explicitly computed in terms of sums and products of the transition probabilities $\tilde{P}(i, j)$ by means of a graph technique. We do not develop here such an aspect, which can be found in ref. 1. These lemmas enable us to obtain results on the long-time behavior of the chain $\bar{X}_t$, and thus of the chain X_t, by using the simpler chain $\tilde{X}_t$ for which Property P holds. We have in fact the following result relating the process X_t to the process $\tilde{X}_t$:

Theorem 2.1. Let $\tilde{W} \subset \tilde{M}$, $\tilde{B} \subset \tilde{W}$, and W and B be the corresponding sets in M; as before we denote by $q_W(x, B)$ the probability that the process X_t starting at x, at the first entrance time in W hits the set B, by $E_x \tau_W$ the mathematical expectation of the n umber of steps until the first entrance in W, calculated under the assumption that the initial state of the chain X_t is x, and by ν the invariant measure of the chain; then for any β sufficiently large and for any $x \in m_i$, $i \in \tilde{M} \setminus \tilde{W}$, $j \in \tilde{W}$:

(i) We have

$$\{\exp(-4r\bar{\gamma}\beta)\}\, \tilde{q}_{\tilde{W}}(i, j) \leqslant q_W(x, m_j) \leqslant \{\exp(4r\bar{\gamma}\beta)\}\, \tilde{q}_{\tilde{W}}(i, j)$$

(ii) There exists a positive η depending on γ with $\eta \to 0$ as $\beta \to \infty$ such that

$$e^{-\eta\beta} t_1 \tilde{E}_i \tilde{\tau}_{\tilde{W}} \leqslant E_x \tau_W \leqslant e^{\eta\beta} t_1 \tilde{E}_i \tilde{\tau}_{\tilde{W}}$$

(iii) There exist constants C and γ', with $\gamma' \to 0$ as $\beta \to \infty$, such that for any $\tilde{B} \subset \tilde{M}$

$$C \cdot t_1 \cdot e^{-\gamma'\beta} \tilde{\nu}(\tilde{B}) \leqslant \nu(B) \leqslant C \cdot t_1 \cdot e^{\gamma'\beta} \tilde{\nu}(\tilde{B})$$

and for any $A \subset S \setminus M$, $\exists B \subset M$ such that:

$$\nu(A) \leqslant e^{-V_1\beta + \gamma'\beta} \nu(B)$$

Proof. Point (i) is an obvious consequence of Lemma 2.1 and the definition of the chain $\bar{X}_t$, from which $q_W(x, B) = \bar{q}_W(x, B)$.

Point (ii) follows from Lemma 2.2 and the following estimates:

Lemma 2.3. For any β sufficiently large there exist a small constant c, $c \to 0$ as $\beta \to \infty$, such that

$$t_1 \bar{E}_x \bar{\tau}_W e^{-c\beta} \leqslant E_x \tau_W \leqslant t_1 \bar{E}_x \bar{\tau}_W e^{c\beta}$$

Proof. By using the strong Markov property we have

$$\begin{aligned}
E_x \tau_W &= E_x \sum_{n=1}^{\infty} \zeta_n \chi(\bar{\tau}_W = n) \\
&= E_x \sum_{n=1}^{\infty} [\zeta_1 + \zeta_1 \circ \theta_{\zeta_1} + \cdots + \zeta_1 \circ \theta_{\zeta_{n-1}}] \\
&\quad \times \chi(X_{\zeta_1} \notin W)\, \chi(X_{\zeta_2} \notin W) \cdots \chi(X_{\zeta_{n-1}} \notin W)\, \chi(X_{\zeta_n} \in W) \\
&= \sum_{n=1}^{\infty} \sum_{m=1}^{n-2} E_x \{\chi(X_{\zeta_1} \notin W)\, \chi(X_{\zeta_2} \notin W) \cdots \chi(X_{\zeta_m} \notin W) \\
&\quad \times E_{X_{\zeta_m}} [\zeta_1 \chi(X_{\zeta_1} \notin W)\, E_{X_{\zeta_1}} \chi(X_{\zeta_2} \notin W) \cdots \chi(X_{\zeta_{n-m-1}} \in W)]\} \\
&\quad + \sum_{n=1}^{\infty} E_x [\chi(\bar{\tau}_W > n-1)\, E_{X_{\zeta_{n-1}}} \zeta_1 \chi(X_{\zeta_1} \in W)] \qquad (2.27)
\end{aligned}$$

Since ζ_1 is given by

$$\zeta_1 = \sum_{s=1}^{t_1} (s + \tau_M(X_s))\, \chi(\sigma = s) + t_1 \chi(\sigma > t_1) \tag{2.28}$$

we have the following estimates for ζ_1:

$$\zeta_1 \leqslant t_1 + K + \sum_{j>K} j\chi(\exists x \in S\backslash M; \tau_M(x) = j) \tag{2.29}$$

$$\zeta_1 \geqslant \tilde{t}_1 - \tilde{t}_1 \chi(\exists x \in M; \sigma(x) \leqslant \tilde{t}_1) \tag{2.30}$$

with $\tilde{t}_1 = e^{V_1\beta - 2\gamma\beta}$, and thus we obtain

$$\begin{aligned}
E_x \tau_W \leqslant{} & (t_1 + K)\, \bar{E}_x \bar{\tau}_W + \sum_{n=1}^{\infty} \mathrm{J}_{m=1}^{n-2} E_x \Big\{ \chi(\bar{\tau}_W > m) \\
& \times E_{X_{\zeta_m}} \Big[\sum_{j>K} j\chi(\exists x \in S\backslash M; \tau_M(x) = j)\, \chi(X_{\zeta_1} \notin W) \\
& \times E_{X_{\zeta_1}} \chi(\bar{\tau}_W = n - m - 1) \Big] \Big\} \\
& + \sum_{n=1}^{\infty} E_x \Big\{ \chi(\bar{\tau}_W > n-1)\, E_{X_{\zeta_{n-1}}} \sum_{j>K} j\chi(\exists x \in S\backslash M; \tau_M(x) = j) \Big\}
\end{aligned} \tag{2.31}$$

$$\begin{aligned}
\leqslant{} & (t_1 + K)\, \bar{E}_x \bar{\tau}_W \\
& + \sum_{m=1}^{\infty} \sum_{n=m+1}^{\infty} \bar{P}_x(\bar{\tau}_W > m) \sup_{y \in M\backslash W} \bar{P}_y(\bar{\tau}_W = n - m - 1) \\
& \times \sup_{z \notin M} \sum_{j>K} jP_z(\tau_M(z) = j) \\
& + \sum_{n=1}^{\infty} \bar{P}_x \chi(\bar{\tau}_W > n-1) \sup_{z \notin M} \sum_{j>K} jP_z(\tau_M(z) = j)
\end{aligned}$$

$$\leqslant (t_1 + K)\, \bar{E}_x \bar{\tau}_W + 2\bar{E}_x \bar{\tau}_W \sup_{z \notin m} \sum_{j>K} jP_z(\tau_M(z) = j) \tag{2.32}$$

and since by Proposition 2.2

$$\sup_{z \notin M} \sum_{j>K} jP_z(\tau_M(z) = j) \leqslant e^{-\alpha K}$$

for some constant α, by chosing $K = e^{\delta'\beta}$ we obtain the estimate

$$E_x \tau_W \leqslant e^{c\beta} t_1 \bar{E}_x \bar{\tau}_W$$

with c exponentially small in β.

As far as the lower bound is concerned, by a similar argument we obtain

$$E_x\tau_W \geqslant \tilde{t}_1 \bar{E}_x \bar{\tau}_W - \tilde{t}_1 \sum_{n=1}^{\infty} \sum_{m=1}^{n-2} E_x\{\chi(\bar{\tau}_W > m)\, E_{X_{\zeta_m}}[\chi(\exists x \in M;\, \sigma(x) \leqslant \tilde{t}_1)$$
$$\times \chi(X_{\zeta_1} \notin W)\, E_{X_{\zeta_1}} \chi(\bar{\tau}_W = n - m - 1)]\}$$
$$+ \sum_{n=1}^{\infty} E_x\{\chi(\bar{\tau}_W > n-1)\, E_{X_{\zeta_{n-1}}} \chi(\exists x \in M;\, \sigma(x) \leqslant \tilde{t}_1)\} \tag{2.33}$$

By an estimate like (2.21)

$$\sup_{x \notin W} \bar{P}_x(\sigma(x) \leqslant \tilde{t}_1)$$
$$\leqslant 1 - \{1 - \exp(-V_1\beta + \gamma\beta)\}^{\tilde{t}_1}$$
$$\leqslant 1 - \exp(-\tilde{t}_1 \tfrac{1}{2} e^{-(V_1\beta + \gamma\beta)}) \leqslant 1 - \exp(-\tfrac{1}{2} e^{-\gamma\beta}) \leqslant \tfrac{1}{2} \exp(-\gamma\beta) \tag{2.34}$$

we obtain

$$E_x\tau_W \geqslant \tilde{t}_1 \bar{E}_x \bar{\tau}_W - \tilde{t}_1 \bar{E}_x \bar{\tau}_W e^{-\gamma\beta} \geqslant t_1 \bar{E}_x \bar{\tau}_W e^{-c\beta}$$

for $c > 2\gamma$.

Point (iii) is an easy consequence of the fact that $\bar{v}$ coincides with $\bar{v}$ and there exists a constant C such that for any $B \subset S$,[2]

$$v(B) = C \sum_{y \in M} \bar{v}(y) E_y \sum_{t=0}^{\zeta_1} \chi_B(X_t)$$

where the constant C is fixed by normalization.

In fact, if $B_1 \subset M$, by the definition of ζ_1,

$$v(B_1) = C \sum_{y \in B_1} \bar{v}(y) E_y \sum_{t=0}^{\zeta_1} \chi_{\{y\}}(X_t)$$

and for any $y \in M$ we have

$$t_1 \geqslant E_y \sum_{t=0}^{\zeta_1} \chi_{\{y\}}(X_t) \geqslant E_y(\sigma \vee \bar{t}_1) \geqslant \bar{t}_1 P(\sigma > \bar{t}_1) \geqslant \frac{e^{-1}}{2} \bar{t}_1$$

where $\bar{t}_1 = e^{V_1\beta - \gamma\beta}$, as before.

For $A \subset S \backslash M$, by using Proposition 2.2, we have

$$v(A) = C \sum_{y \in M} \bar{v}(y) E_y \sum_{t=0}^{\zeta_1} \chi_A(X_t) \leqslant C \cdot K \cdot T_0$$

for a suitable constant K independent of β.

2.6. The Iteration Scheme

From these results it is now clear what the strategy suggested in this paper is in order to study the long-time behavior of the chain X_t: it can be summarized in the following steps:

1. Construct the set M.

2. Consider on M the chain $\bar{X}_t$ and calculate its transition probabilities by means of the quantities $\bar{\Delta}(x, y)$.

3. Consider the new chain $\tilde{X}_t$ on $\tilde{M}$: we can call it $X_t^{(1)}$ on $S^{(1)} = \tilde{M}$; it satisfies Condition P; and then go back to step 1 for this chain.

We obtain in this way a sequence of chains $X_t^{(n)}$ on the spaces $S^{(n)}$ which are smaller and smaller. The iteration scheme is the following:

$S^{(0)} = S$, $\Delta^{(0)}(x, y) = \Delta(x, y)$ defined by (2.1), $\bar{\Delta}^{(0)}(x, y) = \bar{\Delta}(x, y)$, $\forall x, y \in M$, defined by (2.17), $M^{(0)} = M$ defined immediately after by (2.5), V_1 is defined by (2.9); for any $n \geqslant 1$ we define the quantities

$$S^{(n)} = M^{(n-1)}/_{\sim^{(n-1)}} \tag{2.35}$$

$$\Delta^{(n)}(x, y) = \bar{\Delta}^{(n-1)}(x, y) - V_n \qquad \forall x, y \in S^{(n)} \tag{2.36}$$

for any $\phi: \mathbf{N} \to S^{(n)}$:

$$I_{[0,t]}^{(n)}(\phi) = \sum_{i=0}^{t-1} \Delta^{(n)}(\phi_i, \phi_{i+1}) \tag{2.37}$$

$$V^{(n)}(x, y) = \inf_{t, \phi; \phi_0 = x, \phi_t = y} I_{[0,t]}^{(n)}(\phi) \qquad \forall x, y \in S^{(n)} \tag{2.38}$$

$$x \sim^{(n)} y \qquad \text{if and only if} \quad V^{(n)}(x, y) = V^{(n)}(y, x) = 0 \tag{2.39}$$

$$M^{(n)} = \{x \in S^{(n)}; \forall y \in S^{(n)}, y \not\sim^{(n)} x \quad V^{(n)}(x, y) > 0\} \tag{2.40}$$

$$\bar{\Delta}^{(n)}(x, y) = \inf_{t, \phi; \phi_0 = x, \phi_t = y, \phi_s \notin M^{(n)} \text{or } \phi_s \sim^{(n)} x, y, \forall s \in [0, t]} I_{[0,t]}^{(n)}(\phi) \qquad \forall x, y \in M^{(n)} \tag{2.41}$$

$$V_{n+1} = \inf_{x \in M^{(n)}, y \in S^{(n)} x \not\sim^{(n)} y} V^{(n)}(x, y) \tag{2.42}$$

$$t_{n+1} = e^{V_{n+1}\beta + \delta\beta} \tag{2.43}$$

By Theorem 2.1 the results on the quantities $q_W(x, B)$, $E_x \tau_W$, and $v(B)$ can be obtained by looking at the same quantities for the chains $X_t^{(n)}$.

2.7. Results in Some Particular Cases ($\bar{X}_t = \tilde{X}_t$)

We remark that in the cases in which at each step of this procedure there are no equivalent states, then stronger results can be obtained. In this cases in fact there is a correspondence path by path between the chains X_t and $X_t^{(n)}$, living now on the same probability space, and this implies that not only the quantities $q_W(x, B)$ and $E_x \tau_W$ can be compared.

Theorem 2.2. If the chain X_t satisfies the additional condition

$$\forall n \geqslant 0 \qquad \forall x \in S^{(n)}, \qquad \nexists y \in S^{(n)}, \quad x \neq y; \qquad x \sim^{(n)} y \tag{2.44}$$

then for any $t > T_k e^{\delta\beta} \cdot 2^k$, with $T_k = t_1 \cdot t_2 \cdots t^k$, for any $W \subset S^{(k)}$ and for any $x \in S^{(k)}$,

$$P(\tau_W(x) > t) \leqslant P\left(\tau_W^{(k)}(x) > \frac{t}{T_k 2^k}\right) + e^{-c_1 e^{\delta\beta}} \tag{2.45}$$

for some constant c_1, where $\tau_W^{(k)}$ is the hitting time to W for the process $X_t^{(k)}$, moreover, there exists a constant Δ such that

$$P(X_t(x) \notin S^{(k+1)}) \leqslant e^{-\Delta\beta} \tag{2.46}$$

Proof. We have

$$P(\tau_W(x) > t) = P(\tau_W^{(k)}(x) > v^{(k)}(t)) \tag{2.47}$$

where $v^{(k)}(t)$ is the number of transitions of the chain $X^{(k)}$ within the time t. This probability can be estimated in the following way:

$$\leqslant P\left(\tau_W^{(k)}(x) > \frac{t}{T_k 2^k}\right) + P\left(v^{(k)}(y) < \frac{t}{T_k 2^k}\right) \tag{2.48}$$

and this last term has a superexponential estimate. In fact

$$P\left(v^{(1)}(t) < \frac{t}{t_1 2}\right) = P(\zeta_{m_1} > t) \tag{2.49}$$

with $m_1 = t(t_1 2)$ and by a Chebyshev estimate, for any $\lambda > 0$ and by using the Markov property, we obtain

$$\begin{aligned} P(\zeta_{m_1} > t) &\leqslant e^{-\lambda t} \, [\sup_{x \in M} E_x e^{\lambda \zeta_1}]^{m_1} \\ &\leqslant e^{-\lambda t} e^{\lambda t_1 m_1} [\sup_{x \notin M} E_x e^{\lambda \tau_M}]^{m_1} \leqslant e^{-\lambda t} e^{\lambda t_1 m_1} e^{\lambda c m_1} \end{aligned} \tag{2.50 }$$

for some constant c independent of β, if λ is sufficiently small, In conclusion, (2.49) can be estimated by

$$\leqslant e^{-\lambda t/2} e^{\lambda c(t/t_1 2)} \leqslant e^{-c' t} \tag{2.51}$$

By iterating the same argument, we easily obtain that

$$
\begin{aligned}
&P\left(\nu^{(k)}(t)<\frac{t}{T_k 2^k}\right)\\
&\quad\leqslant P\left(\nu^{(k-1)}(t)<\frac{t}{T_{k-1}2^{k-1}}\right)+P\left(\nu^{(k)}(t)<\frac{t}{T_k 2^k}\cap \nu^{(k-1)}(t)>\frac{t}{T_{k-1}2^{k-1}}\right)\\
&\quad\leqslant e^{-c_1 e^{\delta\beta}} \qquad (2.52)
\end{aligned}
$$

since the second term in the rhs of (2.52) is estimated as in (2.51. With such an estimate we can evaluate by iteration also (2.46) in the following way:

$$
\begin{aligned}
P(X_t(x)\notin S^{(k+1)})\leqslant P(X_t(x)\notin S^{(k)})&+P\left(X^{(k)}_{\nu^{(k)}(t)}\notin S^{(k+1)}\cap \nu^{(k)}(t)>\frac{t}{T_k 2^k}\right)\\
&+P\left(\nu^{(k)}(t)<\frac{t}{T_k 2^k}\right)\leqslant e^{-\Delta\beta} \qquad (2.53)
\end{aligned}
$$

for some constant Δ, since the second term in the rhs of (2.53) can be estimated by Corollary 2.1 applied to the chain $X_t^{(k)}$.

2.8. A Final Remark

Finally let me comment on the applicability of such a renormalization procedure.

The crucial point in our construction is the evaluation of the quantities $\bar{\Delta}(x, y)$ in which also the nonequilibrium states in $S\backslash M$ have a role. We note that we can replace the quantities $\bar{\Delta}(x, y)$ with $+\infty$ for all the transitions $x\to y$ such that the function $\bar{\phi}^{(x,y)}$ minimizing the quantity $V(x, y)$ [defined by (2.3)] in a time $\bar{t}^{(x,y)}$ is such that there exists $s\in(0, \bar{t}^{(x,y)})$ such that $\bar{\phi}_s^{(x,y)}=z\in M_{x,y}$. Such a transition $x\to y$ for the chain $\bar{X}_t$ can be neglected [i.e., we can consider $\bar{P}(x, y)=0$] since it will never appear in the functions minimizing $V^{(1)}$ and $\bar{\Delta}^{(1)}$. In fact in this case we have

$$V(x, y)=V(x, z)+V(z, y) \qquad (2.54)$$

and the transitions $x\to z$ and $z\to y$ contribute with a quantity

$$
\begin{aligned}
\Delta^{(1)}(x, z)+\Delta^{(1)}(z, y)&=\bar{\Delta}(x, z)-V_1+\bar{\Delta}(z, y)-V_1\\
&=V(x, z)-V_1+V(z, y)-V_1\\
&=V(x, y)-2V_1<\bar{\Delta}(x, y)-V_1 \qquad (2.55)
\end{aligned}
$$

3. METROPOLIS ALGORITHMS FOR TWO-DIMENSIONAL ISING SYSTEMS: METASTABILITY

In this section we want to apply the renormalization scheme developed in Section 2 to the Markov chain given by the Metropolis algorithm for the two-dimensional Ising system, providing a new proof of the results on metastability obtained by Neves and Schonmann.[3,4]

Consider a ferromagnetic nearest-neighbor Ising model in a finite box $\Lambda \subset \mathbf{Z}^2$ of side L: to each $x \in \Lambda$ we associate a spin variable $\sigma(x) = \pm 1$ and to each spin configuration $\sigma \in \{-1, +1\}^\Lambda = S$ we associate a Hamiltonian

$$H_\Lambda(\sigma) = -\frac{1}{2} \sum_{x, y \in \Lambda, |x-y|=1} \sigma(x)\,\sigma(y) - \frac{h}{2} \sum_{x \in \Lambda} \sigma(x) \tag{3.1}$$

where h is a uniform external magnetic field, with fixed boundary conditions.

In order to compute quantities such as the mean value of an arbitrary observable f with respect to the Gibbs measure

$$\mu_\Lambda(\sigma) = \frac{e^{-\beta H_\Lambda(\sigma)}}{Z_\Lambda} \tag{3.2}$$

here Z_Λ is the partition fuinction, with the Monte Carlo method one considers a Markov chain $\{\sigma_t\}_{t \in \mathbf{N}}$ with state space $S = \{-1, +1\}^\Lambda$, and transition probabilities $P(\sigma, \sigma') = P(\sigma_{t+1} = \sigma' \mid \sigma_t = \sigma)$ satisfying the detailed balance condition:

$$\mu_\Lambda(\sigma)\, P(\sigma, \sigma') = \mu_\Lambda(\sigma')\, P(\sigma', \sigma) \tag{3.3}$$

and the ergodic condition.

An explicit construction of this Markov chain can be given by the Metropolis algorithm, which is defined as follows: given the spin configuration $\sigma = \sigma_t$ at time t, in order to compute the configuration σ_{t+1}, we choose at random with uniform distribution a site x in Λ and we compute

$$\Delta_x H(\sigma) = H_\Lambda(\sigma^x) - H_\Lambda(\sigma) \tag{3.4}$$

with

$$\sigma^x(y) = \begin{cases} \sigma(y) & \text{if } x \neq y \\ -\sigma(y) & \text{if } x = y \end{cases} \tag{3.5}$$

If $\Delta_x H(\sigma) \leq 0$, we flip the spin at x; otherwise we flip it with a probability

$$\exp(-\beta \Delta_x H(\sigma))$$

In terms of transition probabilities this correspond to the following definition: if $\sigma \neq \sigma'$,

$$P(\sigma, \sigma') = \begin{cases} 0 & \text{if } \sigma' \neq \sigma^x \text{ for all } x \in \Lambda \\ (1/|\Lambda|) \exp\{-\beta(\Delta_x H(\sigma) \vee 0)\} & \text{if } \sigma' = \sigma^x \text{ for some } x \end{cases} \tag{3.6}$$

As before, we will denote by $\sigma_t(\sigma)$ the process starting at σ.

This is one of the infinite possible algorithm we can construct satisfying the conditions of reversibility and ergodicity, and even if from a numerical point of view this algorithm is not the best, we will study this case in order to discuss the consequences of our multiscale analysis on metastable situations. Let us consider the Ising model (3.1) in the case of small positive h and with periodic boundary conditions on a sufficiently large box Λ, in the limit of small temperature. Consider the configuration $-\mathbf{1}$ in which all the spins are -1; in the case $h=0$ this state is an equilibrium state of the system; however, for $h>0$ it becomes metastable in the sense that the magnetic field decides the phase even if it is very small, but its effects become relevant only on a scale sufficiently large $[l \geqslant l_c(h) \sim 2/h]$, as only on large scales does the volume energy dominate the surface energy. From a dynamical point of view this means that starting form the configuration $-\mathbf{1}$ locally the system will undergo only small fluctuations around the metastable state $-\mathbf{1}$ for a certain amount of time, very large if β/h is large, until if will "tunnel" to the true equilibriuim $+\mathbf{1}$. The main physical feature of this transition is the existence of a critical value $l_c(h)$ for the size of the droplets: droplets whose side is smaller than $l_c(h)$ tend to shrink, whereas the larger ones tend to grow and there is an "activation energy" which is necessary to create them.

A genuine dynamical argument for the existence of the critical size $l_c(h)$ was introduced for the first time by Neves and Schonmann[3,4] in the framework of a regorous analysis of the Metropolis algorithm in finite volume in the zero-temperature limit and by Martinelli *et al.*[8] (see also ref. 9) for a random cluster algorithm (Swendsen and Wang dynamics) in the thermodynamic limit at low temperature. Nucleation from a metastable state is also studied in ref. 6 for an anisotropic Ising model at very low temperature and in ref. 7 in the case of isotropic nearest-neighbor and next-nearest-neighbor interactions.

The main results proved in ref. 3 and 4 can be summarized as follows.

Let $\tau_\eta(\sigma) \equiv \inf\{t \geqslant 0;\ \sigma_t(\sigma) = \eta\}$ be the first hitting time to the configuration η starting from σ, so that $\tau_{+1}(-1)$ is the nucleation time, that is, the time needed to reach the configuration $+\mathbf{1}$; let R be the set of configurations with all spins -1 except for those in a rectangle $l_1 \times l_2$, which are $+1$, and let $l(\eta) = \min(l_1, l_2)$ for every $\eta \in R$.

Theorem 3.1.[3,3] For any h arbitrarily small, $h<2$, with $2/h$ not integer and Λ sufficiently large:

(a) For all $\eta \in R$: If $l(\eta)<2/h$,

$$\lim_{\beta \to \infty} P(\tau_{-1}(\eta)<\tau_{+1}(\eta))=1$$

If $l(\eta)>2/h$,

$$\lim_{\beta \to \infty} P(\tau_{+1}(\eta)<\tau_{-1}(\eta))=1$$

(b) We have

$$\lim_{\beta \to \infty} P(\tau_G(-\mathbf{1})<\tau_{+1}(-\mathbf{1}))=1$$

where G is the set of configurations in R in which the spins $+1$ forms a square droplet of side l_c.

(c) We have

$$\lim_{\beta \to \infty} \frac{1}{\beta} \log E(\tau_{+1}(-\mathbf{1}))=\Gamma(h)$$

where $\Gamma(h)$ is explicitly computed in terms of the parameter of the Hamiltonian:

$$\Gamma(h)=4l_c-(l_c^2-l_c+1)h \sim \frac{4}{h} \tag{3.7}$$

with $l_c=[2/h]+1$.

(d) We have

$$\lim_{\beta \to \infty} \frac{1}{\beta} \log(\tau_{+1}(-\mathbf{1}))=\Gamma(h) \qquad \text{in probability}$$

(e) $\tau_{+1}(-\mathbf{1})/E(\tau_{+1}(-\mathbf{1}))$ converges in distribution as $\beta \to \infty$ to an exponential random variable of mean one.

All these results are obtained with the volume Λ and h fixed, with $L \gg 2/h$, and β is subsequently taken large enough.

We will give here the strategy of a proof of Theorem 3.1, which is completely different from the original proof given in refs. 3 and 4, based on the multiscale analysis developed in the previous section, for the Markov chain defined by (3.6).

Our proof is by induction. We define, first of all, the step 0 of our induction and we prove it.

Step 0. The starting point is the construction of set set $M^{(0)} = M$ of the minima [see definition (2.5)]. For any configuration σ we denote by $C_+(\sigma)$ the set of plus spins:

$$C_+(\sigma) \equiv \{x \in \Lambda;\ \sigma(x) = +1\} \tag{3.8}$$

Lemma 3.1(0). A configuration σ belongs to M, i.e., it is a minimum, if and only if

$$\forall x \in C_+(\sigma) \qquad \sum_{y;\,|y-x|=1} \sigma(y) \geqslant 0$$

$$\forall x \in C_-(\sigma) \qquad \sum_{y;\,|y-x|=1} \sigma(y) \leqslant -2$$

This means that a configuration σ belongs to M if and only if

$$C_+(\sigma) = \bigcup_i R^i$$

where R^i are rectangles on the lattice of sides $l_1^i \leqslant l_2^i$ with $l_1^i > 1$ for any i, and $\mathrm{dist}(R^i, R^j) > 2$ if $i \neq j$.

This union can be empty, that is, $-\mathbf{1}$ is a minimum.

Proof. Suppose that

$$C_+(\sigma) \neq \bigcup_i R^i$$

or

$$C_+(\sigma) = \bigcup_i R^i \qquad \text{but with} \quad l_1^i = 1 \quad \text{for some } i$$

or

$$\mathrm{dis}(R^i, R^j) \leqslant 2 \qquad \text{for some} \quad i \neq j$$

Then there is an $x \in \Lambda$ such that at least one of the following statements holds:

(a) $\sigma(x) = +1$ and at least three among the four nearest-neighbor sites of x are -1.

(b) $\sigma(x) = -1$ and two nearest neighbors of x are $+1$.

Then there exists a configuration σ' such that $V(\sigma, \sigma') = 0$ and $V(\sigma', \sigma) > 0$, where σ' is given by σ^x [see (3.5)] for x satisfying (a) and (b) and thus σ is not a minimum.

If $C_+(\sigma) = \bigcup_i R^i$ with $l_1^i > 1$ for any i, and $\text{dist}(R^i, R^j) > 2$ if $i \neq j$, then for every $x \in \Lambda$ the configuration σ^x has $H_\Lambda(\sigma^x) > H_\Lambda(\sigma)$ and thus σ is a minimum.

Lemma 3.2.(0). The chain $X_t^{(0)} \equiv X_t$ has nonequivalent configurations and thus $M = S^{(1)}$.

The proof easily follows from the fact that for our chain $\Delta(\sigma, \eta) = 0$ implies $H(\sigma) > H(\eta)$.

We have now to evaluate the transition probabilities for the new chain in M. Let $\sigma \in M$, $\sigma \neq \mathbf{-1}$ be a configuration such that

$$C_+(\sigma) = \bigcup_{i \in I} R^i$$

For each $i \in I$ and $j = 1, 2$ we denote by $\sigma_{(+L_j^i)}$ and $\sigma_{(-L_j^i)}$ the stable configurations given respectively by the creation or the annihilation of a row L_j^i of length l_j^i in the rectagle R^i. This means the following: let η be the configuration defined by

$$C_+(\eta) = R^i \cup L_j^i \bigcup_{k \in I, k \neq i} R^k$$

where $R^i \cup L_1^i$ is a rectangle of sides l_1^i, $l_2^i + 1$, and $R^i \cup L_2^i$ is a rectangle of sides $l_1^i + 1$, l_2^i (see Fig. 2).

If η is stable on scale 0 ($\eta \in M$), then $\sigma_{(+L_j^i)} = \eta$; otherwise the configuration $\sigma_{(+L_j^i)}$ is the stable configuration to which η is connected. In fact, if $\eta \notin M$, this implies that $R^i \cup L_j^i$ has a distance equal to 2 from another rectangle R^k. In this case the configuration $\sigma_{(+L_j^i)}$ is such that its C_+ set contains the smallest rectangle containing $R^i \cup L_j^i \cup R^k$. In general, $\sigma_{(+L_j^i)}$ is the stable configuration such that there exists a sequence of configurations ϕ_i such that $\phi_0 = \eta$, $\phi_t = \sigma_{(+L_j^i)}$, and $I_t(\phi) = 0$.

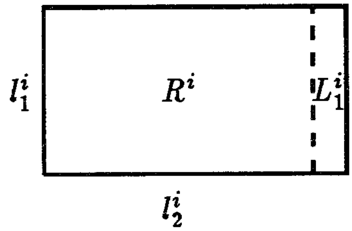

Fig. 2. The sets R^i and $R^i \cup L_1^i$.

An analogous definition can be done for $\sigma_{(-L_j^i)}$: if $(R^i\backslash L_j^i)$ is a rectangle with the smaller side larger than or equal to 2, i.e., if the configuration η such that $C_+(\eta)=(R^i\backslash L_j^i)\bigcup_{k\in I,k\neq i}R^k$ belongs to M, then $\sigma_{(-L_j^i)}=\eta$.

Otherwise

$$C_+(\sigma_{(-L_j^i)})=\bigcup_{k\in I,k\neq i}R^k$$

$\sigma_{(+L_j^i)}$ and $\sigma_{(-L_j^i)}$ are the configuration "near" σ in the following sense:

Lemma 3.3(0). (a) We have the following:

$$\bar{\Delta}(\sigma,\sigma_{(+L_j^i)})=2-h \tag{3.9}$$

$$\bar{\Delta}(\sigma,\sigma_{(-L_j^i)})\begin{cases}=h(l_i^j-1) & \text{if } \ l_i^j<l_c\\ >2-h & \text{if } \ l_i^j\geqslant l_c\end{cases} \tag{3.10}$$

(b) For any $\eta\neq\sigma_{(\pm L_j^i)}$, $\forall i\in I$, $j=1,2$, at least one of the following statements holds:

1. If σ and η are such that $C_+(\eta)\not\subset C_+(\sigma)$, then for any sequence ϕ_s with $\phi_0=\sigma$, $\phi_t=\eta$, and $\phi_{s+1}=(\phi_s)^x$, $\forall s\leqslant t$, for some x, there are $k=k(\phi)\geqslant 1$ times in which a new row of plus spins is created, that is, $i_1,\ldots,i_k$, such that $C_+(\phi_{i_j})=C_+(\phi_{i_j-1})\cup\{x_{i_j}\}$, where, in the configuration ϕ_{i_j-1}, x_{i_j} has at least three minus neighbor spins. Let $k=k(\sigma,\eta)=\inf_{t,\phi;\phi_0=\sigma,\phi_t=\eta}k(\phi)$. In this case

$$\bar{\Delta}(\sigma,\eta)\geqslant(2-h)k \tag{3.11}$$

2. If σ and η are such that $C_+(\sigma)\not\subset C_+(\eta)$, then either

$$\bar{\Delta}(\sigma,\eta)\geqslant 2+h \tag{3.12}$$

or

$$\bar{\Delta}(\sigma,\eta)\geqslant h(C(\sigma,\eta)-r(\sigma,\eta))=\sum_{m\leqslant r(\sigma,\eta)}[h(l^m-1)] \tag{3.13}$$

where $C(\sigma,\eta)=|C_+(\sigma)\backslash(C_+(\eta)\cap C_+(\sigma))|$, $r(\delta,\eta)$ is the number of rows in $C_+(\sigma)\backslash(C_+(\eta)\cap C_+(\sigma))$, and the l^m are the lengths of a sequence of rows L^m such that $\bigcup_{m\leqslant r(\sigma,\eta)}L^m=C_+(\sigma)\backslash(C_+(\eta))\cap C_+(\sigma))$. We stress here that a row L is a set of plus spins $L=\{x_k\}_{k=0}^l$ with $x_k=x_0+k\bar{e}_i$, where $\bar{e}_i$ is a unitary vector of the lattice and $\sigma(x_k+\bar{e}_j)=+1$ for some $\bar{e}_j$ orthogonal to $\bar{e}_i$, $\sigma(x_k-\bar{e}_j)=-1$ and $\sigma(x_0-\bar{e}_i)=-1=\sigma(x_l+\bar{e}_i)$. We note that $C_+(\sigma)\backslash(C_+(\eta)\cap C_+(\sigma))$ can always be written as $\bigcup_{m\leqslant r(\sigma,\eta)} L^m$ with $l^m \geqslant 2$, since $\sigma, \eta \in M$. We also remark that the number $r(\sigma, \eta)$ is independent of the choice of the sequence of L^m and it can be defined in the following way: if $C_+(\sigma) \not\subset C_+(\eta)$ for any sequence ϕ_s with $\phi_0 = \sigma$, $\phi_t = \eta$, and $\phi_{s+1} = (\phi_s)^{x_s}$, $\forall s \leqslant t$, for some x_s, there are $r = r(\phi) \geqslant 1$ times in which x_s in the configuration ϕ_s is a plus spin with three minus neighbor spins.

Let $r = r(\sigma, \eta) = \inf_{t,\phi;\phi_0=\sigma,\phi_t=\eta} r(\phi)$. By definition, $r(\sigma, \eta) = k(\eta, \sigma)$.

(c) Let σ_{Q_l} be a configuration such that $C_+(\sigma_{Q_l}) = Q_l$ with Q_l a square of side l; then

$$\bar{\Delta}(-\mathbf{1}, \sigma_{Q_2}) = 8 - 3h \tag{3.14}$$

and for every $\sigma \in M$, $\sigma \neq \sigma_{Q_2}$,

$$\bar{\Delta}(-\mathbf{1}, \sigma) \geqslant \bar{\Delta}(-\mathbf{1}, \sigma_{Q_2}) + K(2-h) \tag{3.15}$$

where $K = k(\sigma_{Q_2}, \sigma) = \inf_{t,\phi;\phi_0=\sigma_{Q_2},\phi_t=\sigma} k(\phi)$ [see definition in point (b1)].

(d) Moreover,

$$\bar{\Delta}(\sigma, \eta) = h \qquad \text{implies} \quad H(\sigma) > H(\eta) \tag{3.16}$$

Proof. First of all a general remark: for any $\sigma. \eta \in M$, if $C_+(\eta)$ has a row of length 1 which is not in $C_+(\sigma)$, then from any ϕ going from η to σ we have

$$I_t(\phi) \begin{cases} \geqslant h(l-1) & \text{if} \quad l < l_c \\ > 2-h & \text{if} \quad l \geqslant l_c \end{cases} \tag{3.17}$$

and for any ϕ' going from σ to η we have

$$I_t(\phi') \geqslant 2-h \tag{3.18}$$

In fact, either the function ϕ has a move with $\Delta(\phi_i, \phi_{i+1}) \geqslant 2-h$, and in this case (3.17) is verified, or it is a sequence of single-spin-flip moves such that there exists at least $l-1$ times $s_1,..., s_{l-1}$ such that at each s_i a plus spin, with two neighbor plus spins, is flipped to minus. Thus (3.17) is trivially verified if we note that in the case $l \geqslant l_c$ the inequality $(l-1)h > 2-h$ holds true.

In the function ϕ' there must exist at least a time s in which a new row is created, that is, a minus spin with three neighbor minus spins is flipped to plus.

(a) To prove (3.9), we note that

$$\bar{\Delta}(\sigma, \sigma_{(+L_j^i)}) \leqslant I_t(\phi) \tag{3.19}$$

with $t=l_1^i$, $\phi_0=\sigma$, and for $k\leqslant l_1^i$, $C_+(\phi_k)=R^i\cup x_k\bigcup_{l\in I, l\neq i}R^l$, where c_k is a row of sites of length k and $R^i\cup c_k$ is given in Fig. 3.

For such a ϕ we have

$$I_t(\phi)=2-h$$

since only the contribution due to the first transition $\sigma\to\phi_1$ is different from zero.

On the other hand, by (3.18),

$$\bar{\Delta}(\sigma, \sigma_{(+L_j^i)})\geqslant 2-h$$

and thus (3.9) is proved.

In order to obtain (3.10), we follow a similar argument with a function ϕ' such that $\phi'_0=\sigma$ and for $k\leqslant l_1^k$, $C_+(\phi'_k)=(R^i\backslash c_k)\bigcup_{l\in I, l\neq i}R^l$.

(b) Let now $\eta\neq\sigma_{(\pm L_j^i)}$, $\forall i\in I$, $j=1, 2$, and suppose that there exists a point $x\in C_+(\eta)$ such that $x\notin C_+(\sigma)$. This implies that every function ϕ realizing the transition $\sigma\to\eta$ must have at least a time in which a new row is created somewhere. If k is the minimal number of rows which have to be created in order to obtain η, then (3.11) follows from the definition of $\Delta(\sigma, \eta)$.

Suppose now that $C_+(\sigma)\not\subset C_+(\eta)$; then either the configuration η is obtained by σ with at least a move corresponding to a $\Delta(\phi_i, \phi_i+1)\geqslant 2+h$, and then (3.12) holds, or otherwise for each function ϕ going from σ to η, for each time i in which a plus spin is flipped to minus we have $\Delta(\phi_i, \phi_{i-1})\leqslant h$. This means that there are $r>1$ entire rows L^1,..., L^r of plus spins contained in $C_+(\sigma)$ which are annihilated with this procedure with a contribution of $\bar{\Delta}(\sigma, \eta)$ equal to $\sum_{j=1}^r(l^j-1)$. Thus (3.13) is proved.

In order to prove point (c) of the lemma, we consider the function

$$\phi_0=-\mathbf{1},\qquad C_+(\phi_1)=\{x\},\qquad C_+(\phi_2)=\{x, x+e_1\}$$

$$C_+(\phi_3)=\{x, x+e_1, x+e_2\},\qquad C_+(\phi_4)=\{x, x+e_1, x+e_2, x+e_1+e_2\}$$

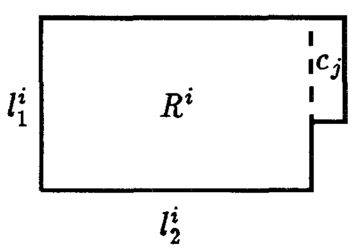

Fig. 3. The set $R^i\cup c_j$.

where e_1 and e_2 are the basic unitary vectors of our lattice, and

$$I_4(\phi) = 4 - h + 2 - h + 2 - h = 8 - 3h$$

Also in this case we remark that each function ϕ' going from $-\mathbf{1}$ to σ_{Q_2} has a functional $I_i(\phi') \geqslant 8 - 3h$, since there must exist at least a time s_1 in which a minus spin four minus neighbor spins is flipped and at least two subsequent times $s_2, s_3 > s_1$ in which a minus spin with three minus neighbor spins is flipped to plus.

In order to prove (3.15), we note first of all that for each $\sigma \in M$ with $C_+(\sigma) = \bigcup_{i \in I} R^i$, in the function ϕ minimizing $\bar{\Delta}(-\mathbf{1}, \sigma)$ there are $|I|$ times in which a minus spin with four minus neighbors is flipped. In fact the mechanism of creation of a new cluster by splitting is not competitive since it requires a larger functional

$$2 - h + 2 - h + 2 - h + h > 4 - h$$

By an argument similar to that of (3.18) we obtain

$$\bar{\Delta}(-\mathbf{1}, \sigma) = |I| (4 - h) + (2 - h) \sum_{i \in I} (l_1^i - 1 + l_2^i - 1)$$

and (3.15) easily follows.

(d) If $\bar{\Delta}(\sigma, \eta) = h$, by Eqs. (3.9)–(3.12) this implies that $C_+(\sigma) = C_+(\eta) \cup L_1^i$, where L_1^i is a row of length 2 and thus $H(\eta) < H(\sigma)$.

Lemma 3.4(0). We have

$$V_1 = h, \qquad t_1 = e^{h\beta + \delta\beta} \tag{3.20}$$

The chain $X_t^{(1)}$ on $M = S^{(1)}$, corresponding to the original chain with a rescaled time by the quantity t_1, has transition probabilities given by the quantities

$$\Delta^{(1)}(\sigma, \eta) = \bar{\Delta}(\sigma, n) - h \tag{3.21}$$

with $\bar{\Delta}(\sigma, \eta)$ evaluated in Lemma 3.2(0).

Proof. The proof of (3.20) easily follows from (3.9)–(3.12). The transition probabilities of the chain $X_t^{(1)}$ are evaluated by Proposition 2.3.

Definition of Step *n*. The Markov chain $X_t^{(n)}$ defined on the state space $S^{(n)}$ has the following properties:

Lemma 3.1(n). A configuration $\sigma \in M^{(n)}$, i.e., is a stable state for the chain $X_t^{(n)}$, if and only if

$$C_+(\sigma) = \bigcup_i R^i$$

where R^i are rectangles of sides $l_1^i \leqslant l_2^i$ with $l_1^i > n+1$ for any i, and $\operatorname{dist}(R^i, R^j) > 2$ if $i \neq j$. This union can be empty, that is, $-\mathbf{1} \in M^{(n)}$.

Lemma 3.2(n). The chain $X_t^{(n)}$ has nonequivalent configurations and thus $M^{(n)} = S^{(n+1)}$.

Let $\sigma_{(+L_j^i)}$ and $\sigma_{(-L_j^i)}$ be defined as in step 0, where now M is replaced by $M^{(n)}$.

Lemma 3.3(n). (a) We have the following:

$$\bar{\Delta}^{(n)}(\sigma, \sigma_{(+L_j^i)}) = 2 - (n+1)h \tag{3.22}$$

$$\bar{\Delta}^{(n)}(\sigma, \sigma_{(-L_j^i)}) = h(l_j^i - (n+1)) \qquad \forall i, j; \quad l_j^i < l_c \tag{3.23}$$

(b) For any $\eta \neq \sigma$ at least one of the following statements holds:

1. If $C_+(\eta) \not\subset C_+(\sigma)$ and $k = k(\sigma, \eta)$ is the minimal number of the new row of plus spins which have to be created [see Lemma 3.3(0)],

$$\bar{\Delta}^{(n)}(\sigma, \eta) \geqslant [2 - (n+1)h]k \tag{3.24}$$

2. If $C_+(\sigma) \not\subset C_+(\eta)$, then either

$$\bar{\Delta}^{(n)}(\sigma, \eta) \geqslant 2 - (n-1)h \tag{3.25}$$

or

$$\bar{\Delta}^{(n)}(\sigma, \eta) \geqslant h[C(\sigma, \eta) - r(\sigma, \eta)(n+1)] = \sum_{m \leqslant r(\sigma, \eta)} [h(l^m - (n+1))] \tag{3.26}$$

where $r(\sigma, \eta)$ and $C(\sigma, \eta)$ are defined as in Lemma 3.3(0).

(c) We have

$$\bar{\Delta}^{(n)}(-\mathbf{1}, \sigma_{Q_{n+2}}) = \bar{\Delta}^{(n-1)}(-\mathbf{1}, \sigma_{Q_{n+1}}) + 4 - 2(n+1)h - 3h \tag{3.27}$$

for any $l > n+2$.

For any $\sigma \in M^{(n)}$, $\sigma \neq \sigma_{Q_{n+2}}$,

$$\bar{\Delta}^{(n)}(-\mathbf{1}, \sigma) \geqslant \bar{\Delta}^{(n)}(-\mathbf{1}, \sigma_{Q_{n+2}}) + K(2 - (n+1)h) \tag{3.28}$$

where

$$K = k(\sigma_{Q_{n+2}}, \sigma) = \inf_{t, \phi; \phi_0 = \sigma_{Q_{n+2}}, \phi_t = \sigma, \phi_{i+1} = (\phi_i)^x} k(\phi)$$

(d) Moreover,

$$\bar{\Delta}^{(n)}(\sigma, n) = h \qquad \text{implies} \quad H(\sigma) > H(\eta) \tag{3.29}$$

Lemma 3.4(***n***). We have

$$V_{n+1} = h, \qquad t_{n+1} = e^{h\beta + \delta\beta} \tag{3.30}$$

The chain $X_t^{(n+1)}$ on $M^{(n)} = S^{(n+1)}$, corresponding to the original chain with a rescaled time by the quantity $T_{n+1} = t_{n+1} t_n \cdots t_1$, has transition probabilities given by the quantities

$$\Delta^{(n+1)}(\sigma, \eta) = \bar{\Delta}^{(n)}(\sigma, \eta) - h \tag{3.31}$$

with $\bar{\Delta}^{(n)}(\sigma, \eta)$ evaluated in Lemma 3.3 (n).

From Step n to Step $n+1$

Here we prove the results stated in the previous subsection by induction. We suppose that they are proved up to steps $n < l_c - 3$.

Proof of Lemma 3.1 $(n+1)$. A configuration $\sigma \in S^{(n+1)} = M^{(n)}$, by Lemma 3($n$), is such that

$$C_+(\sigma) = \bigcup_i R^i$$

where R^i are rectangles of sides $l_1^i \leqslant l_2^i$ with $l_1^i > n+1$ for any i, and $\operatorname{dist}(R^i, R^j) > 2$ if $i \neq j$.

If σ is such that there is i with $l_1^i = n+2$, then by Lemma 3.3(n) we can show that σ is not stable on scale $n+1$; in fact,

$$\begin{aligned}\Delta^{(n+1)}(\sigma, \sigma_{(-L_j^i)}) &= \bar{\Delta}^{(n)}(\sigma, \eta) - h \\ &= h(l_i^1 - (n+1)) - h = h - h = 0\end{aligned}$$

On the other hand, for any configuration $\eta \in S^{(n+1)}$ such that $l_1^i > n+2$ for any i, by Lemma 3.3(n), we have that

$$\Delta^{(n+1)}(\eta, \zeta) > 0 \qquad \forall \zeta \in S^{(n+1)}$$

Proof of Lemma 3.2($n+1$). This result easily follows from (3.32) and Lemma 3.3(n)(d), since

$$\Delta^{(n+1)}(\sigma, \eta) \equiv \bar{\Delta}^{(n)}(\sigma, \eta) - h = 0$$

implies

$$H(\sigma) > H(\eta)$$

Proof of Lemma 3.3(n + 1). (a) Equations (3.22) and (3.23) at step $n+1$ easily follow from Eq. (3.31) and by Lemma 3.3(n) (a) and (b).

In fact,

$$\bar{\Delta}^{(n+1)}(\sigma, \sigma_{(+L_j^i)}) = \inf_{t, \phi \in \Phi(S^{(n+1)}); \phi_0 = \sigma, \phi_t = \sigma_{(+L_j^i)}, \phi_s \notin M^{(n+1)}_{\sigma, \sigma_{(+L_j^i)}}} I_t^{(n+1)}(\phi)$$

$$\Delta^{(n+1)}(\sigma, \sigma_{(+L_j^i)}) = \bar{\Delta}^{(n)}(\sigma, \sigma_{(+L_j^i)}) - h = 2 - (n+2)h$$

and for any $\phi \in \Phi(S^{(n+1)})$ going from σ to $\sigma_{(+L_j^i)}$ in a time $t > 1$ without touching $M^{(n+1)}_{\sigma, \sigma_{(+L_j^i)}}$ there exists at least a time $s \leqslant t$ such that ϕ_{s-1} and ϕ_s verify case (b1) of Lemma 3.3(n) and thus

$$\bar{\Delta}^{(n+1)}(\sigma, \sigma_{(+L_j^i)}) = \Delta^{(n+1)}(\sigma, \sigma_{(+L_j^i)}) = 2 - (n+2)h$$

In the same way

$$\bar{\Delta}^{(n+1)}(\sigma, \sigma_{(-L_j^i)}) = \Delta^{(n+1)}(\sigma, \sigma_{(-L_j^i)}) = h(l_j^i - (n+2))$$

since for any $\phi \in \Phi(S^{(n+1)})$ going from σ to $\sigma_{(-L_j^i)}$ in a time $t > 1$ without touching $M^{(n+1)}_{\sigma, \sigma_{(+L_j^i)}}$ and minimizing $I_t(\phi)$, there exists at least a time $s \leqslant t$ such that $\Delta^{(n+1)}(\phi_s, \phi_{s+1}) \geqslant 2 - nh$ or a time s such that $L_j^i \subset C_+(\phi_s)$ but $L_j^i \not\subset C_+(\phi_{s+1})$ and in this case $\Delta^{(n+1)}(\phi_s, \phi_{s+1}) \geqslant h(l_j^i - (n+2))$.

(b) Equation (3.24) at step $n+1$ follows from the same equation at step n. In fact, if $C_+(\eta) \not\subset C_+(\sigma)$, then for every function in $\phi \in \Phi(S^{(n+1)})$ such that $\phi_0 = \sigma$, $\phi_t = \eta$, let $k(\phi_i, \phi_{i+1})$ be the minimal number of new rows of plus spins which have to be created at step i. By the definition of k we have that

$$k(\phi_0, \phi_1) + k(\phi_1, \phi_2) + \cdots + k(\phi_{t-1}, \phi_t) \geqslant k(\sigma, \eta)$$

In particular, this inequality holds for the function $\bar{\phi} \in \Phi(S^{(n+1)})$, minimizing the quantity $I_t^{(n+1)}$ in the definition of $\bar{\Delta}^{(n)}(\sigma, \eta)$. By applying Lemma 3.3($n$)(b1) and Eq. (3.31) we obtain

$$\begin{aligned}
\bar{\Delta}^{(n)}(\sigma, \eta) &= I_t^{(n+1)}(\bar{\phi}) \\
&\geqslant \sum_{i \leqslant t; k(\bar{\phi}_{i-1}, \bar{\phi}_i) \geqslant 1} \Delta^{(n+1)}(\bar{\phi}_{i-1}, \bar{\phi}_i) \\
&\geqslant \sum_{i \leqslant t; k(\bar{\phi}_{i-1}, \bar{\phi}_i) \geqslant 1} [2 - (n+1)h]\, k(\bar{\phi}_{i-1}, \bar{\phi}_i) - h \\
&\geqslant [2 - (n+2)h] \sum_{i \leqslant t; k(\bar{\phi}_{i-1}, \bar{\phi}_i) \geqslant 1} k(\bar{\phi}_{i-1}, \bar{\phi}_i) \\
&\geqslant [2 - (n+2)h]\, k(\sigma, \eta)
\end{aligned}$$

If $C_+(\sigma) \not\subset C_+(\eta)$, let $\bar{\phi}$ be a function in $\Phi(S^{(n+1)})$ minimizing the quantity $I_t^{(n+1)}$ in the definition of $\bar{\Delta}^{(n)}(\sigma, \eta)$. If $\bar{\phi}$ reaches η in only one step, then the result follows from Lemma 3.3(n)(b). If $\bar{\phi}$ needs more than one step, say t, then we have two possibilities:

(i) There exists $1 < k \leqslant t$ such that $\Delta^{(n+1)}(\bar{\phi}_{k-1}, \bar{\phi}_k) \geqslant 2 - nh$ and thus (3.25) at step $n+1$ holds.

(ii) For any $1 < k \leqslant t$ we have $\Delta^{(n+1)}(\bar{\phi}_{k-1}, \bar{\phi}_k) < 2 - nh$; then by the result at step n we have

$$
\begin{aligned}
\Delta^{(n+1)}(\sigma, \eta) &= I^{(n+1)}(\bar{\phi}) \\
&\geqslant \sum_{i; C_+(\bar{\phi}_{i-1}) \not\subset C_+(\bar{\phi}_i)} \Delta^{(n+1)}(\bar{\phi}_{i-1}, \bar{\phi}_i) \\
&\geqslant \sum_{i; C_+(\bar{\phi}_{i-1}) \not\subset C_+(\bar{\phi}_i)} \sum_{m_i < r_i} h[l^{m_i} - (n+1)] - h \\
&\geqslant \sum_{m \in I(\bar{\phi})} h[l^m - (n+1)] - h \\
&= h[C(\sigma, \eta) - r(\sigma, \eta)(n+2)]
\end{aligned}
$$

where $r_i = r(\bar{\phi}_{i-1}, \bar{\phi}_i)$, $I(\bar{\phi})$ is the set of index of rows which are annihilated along the function $\bar{\phi}$ such that $\bigcup_{m \in I(\bar{\phi})} L^m = C_+(\sigma) \backslash [C_+(\eta) \cap C_+(\sigma)]$, and with $|I(\bar{\phi})| = r(\sigma, \eta)$.

(c) By the definition of $\bar{\Delta}^{(n+1)}$ and by Lemma 3.3(n) we have that

$$
\begin{aligned}
\bar{\Delta}^{(n+1)}(-\mathbf{1}, \sigma_{Q_{n+3}}) &\leqslant \Delta^{(n+1)}(-\mathbf{1}, \sigma_{Q_{n+2}}) \\
&\quad + \Delta^{(n+1)}(\sigma_{Q_{n+2}}, \sigma_{R_{n+2,n+3}}) + \Delta^{(n+1)}(\sigma_{R_{n+2,n+3}}, \sigma_{Q_{n+3}}) \\
&= \bar{\Delta}^{(n)}(-\mathbf{1}, \sigma_{Q_{n+2}}) + 2[2 - (n+2)h] - 3h \qquad (3.32)
\end{aligned}
$$

We have to prove that for every $\phi \in \Phi(S^{(n+1)})$ such that

$$
\phi_0 = -\mathbf{1}, \qquad \phi_t = \sigma_{Q_{n+3}}, \qquad \phi_s \notin M^{(n+1)}_{-\mathbf{1}, \sigma_{Q_{n+3}}}, \quad \forall s < t
$$

then we have

$$
I_t^{(n+1)}(\phi) \geqslant \bar{\Delta}^{(n)}(-\mathbf{1}, \sigma_{Q_{n+2}}) + 2[2 - (n+2)h] - 3h
$$

If $\phi_1 = \sigma_{Q_{n+2}}$, there must exist in ϕ two times i_1, i_2 between 1 and t such that $k(\phi_{i_k-1}, \phi_{i_k}) \geqslant 1$ or a time i in which $k(\phi_{i-1}, \phi_i) = 2$. In both cases by (3.22) and (3.24) the result follows. If $\phi_1 \neq \sigma_{Q_{n+2}}$, by (3.28) we have

$$
\Delta^{(n+1)}(-\mathbf{1}, \phi_1) \geqslant \bar{\Delta}^{(n)}(-\mathbf{1}, \sigma_{Q_{n+2}}) + k(\sigma_{Q_{n+2}}, \phi_1)[2 - (n+1)h] - 3h
$$

with $k \geqslant 1$ and if $k=1$, by (3.22) and (3.24), $\Delta^{(n+1)}(\phi_{t-1}, \phi_t) \geqslant 2-(n+1)h-h$. To prove (3.28) at step $n+1$ we note that for every $\sigma \in M^{(n+1)}$, $\sigma \neq \sigma_{Q_{n+3}}$,

$$\bar{\Delta}^{(n+1)}(-\mathbf{1}, \sigma) \geqslant I_t^{(n+1)}(\bar{\phi})$$

for some function $\bar{\phi} \in \Phi(S^{(n+1)})$ minimizing $\bar{\Delta}^{(n+1)}(-\mathbf{1}, \sigma)$ and thus

$$\begin{aligned}
&\bar{\Delta}^{(n+1)}(-\mathbf{1}, \sigma) \\
&\quad \geqslant \bar{\Delta}^{(n)}(-\mathbf{1}, \sigma_{Q_{n+2}}) + k(\sigma_{Q_{n+2}}, \bar{\phi}_1)[2-(n+1)h]-h \\
&\qquad + \sum_{i;k(\bar{\phi}_{i-1}, \bar{\phi}_i) \geqslant 1} \{[2-(n+1)h]\, k(\bar{\phi}_{i-1}, \bar{\phi}_i) - h\} \\
&\quad \geqslant \bar{\Delta}^{(n)}(-\mathbf{1}, \sigma_{Q_{n+2}}) + [2-(n+1)h][k(\sigma_{Q_{n+2}}, \bar{\phi}_1) \\
&\qquad + \sum_{i;k(\bar{\phi}_{i-1}, \bar{\phi}_i) \geqslant 1} k(\bar{\phi}_{i-1}, \bar{\phi}_i) - h(1+\mathrm{card}\{i; k(\bar{\phi}_{i-1}, \bar{\phi}_i) \geqslant 1\})] \\
&\qquad\qquad \text{(card is the cardinality)} \\
&\quad \geqslant \bar{\Delta}^{(n+1)}(-\mathbf{1}, \sigma_{Q_{n+3}}) + k(\sigma_{Q_{n+3}}, \sigma)[2-(n+2)h]
\end{aligned}$$

Proof of Lemma 3.4(n + 1). By Lemma 3.3($n+1$) we have $V_{n+2} = h$ and thus the lemma follows from Proposition 2.3.

The Step $l_c - 2$

Up to now we have proved that for any $n \leqslant l_c - 3$, Lemmas 3.1–3.4 hold and thus the transition probabilities of the chain $X^{(l_c-2)}$ can be computed by means of the equality $\Delta^{(l_c-2)}(\cdot, \cdot) = \bar{\Delta}^{(l_c-3)}(\cdot, \cdot) - h$.

Lemmas 3.1–3.3 at step $l_c - 2$ hold and the proof is exactly the same as before, but now

$$V_{l_c-1} = \inf_{\sigma, \eta \in M^{(l_c-2)}} \Delta^{(l_c-2)}(\sigma, \eta) = 2-(l_c-1)h < h$$

Thus the chain $X^{(l_c-1)}$ is such that $\Delta^{(l_c-1)}(\sigma, \sigma_{(+L_j^i)}) = 0$ and thus the only configurations belonging to $M^{(l_c-1)}$ are $+\mathbf{1}$ and $-\mathbf{1}$, and

$$V_{l_c} = \bar{\Delta}^{(l_c-1)}(-\mathbf{1}, +\mathbf{1}) = \bar{\Delta}^{(l_c-2)}(-\mathbf{1}, \sigma_{Q_{l_c}}) - 2 + (l_c-1)h \tag{3.33}$$

We can conclude that the chain $X^{(l_c)}$ on the states space $\{-\mathbf{1}, +\mathbf{1}\}$ has

$$\Delta^{(l_c)}(-\mathbf{1}, +\mathbf{1}) = \bar{\Delta}^{(l_c-1)}(-\mathbf{1}, +\mathbf{1}) - V_{l_c} = 0 \tag{3.34}$$

that is $M^{(l_c)} = +\mathbf{1}$.

Proof of Theorem 3.1. With this construction the proof of Theorem 3.1 is quite simple. We observe that with the notation given in Lemma 2.1

$$P(\tau_A(\eta) < \tau_B(\eta)) = q_{\{A \cup B\}}(\eta, A) \tag{3.35}$$

and thus it is sufficient to control such quantities for one of the chains $X_t^{(n)}$. If we look for instance at point (a) of the theorem, if $l(\eta) = l$, then $\eta \in M^{(l-2)}$, but $\eta \notin M^{(l-1)}$. If we look at the chain $X^{(l-1)}$ defined on $M^{(l-2)}$, we have that

$$P(\eta, -\mathbf{1}) \quad \text{is of order one}$$

and

$$P(\eta, M^{(l-1)} \backslash -\mathbf{1}) \quad \text{is exponentially small in } \beta \text{ for each} \quad \eta \in M^{(l-2)} \backslash M^{(l-1)} \tag{3.36}$$

We can estimate $P(\tau_{-\mathbf{1}}(\eta) > \tau_{+\mathbf{1}}(\eta))$ as follows:

$$\leqslant P(\tau_{M^{(l-1)}}(\eta) > e^{\delta\beta}) + P(\tau_{M^{(l-1)} \backslash -\mathbf{1}} < e^{\delta\beta}) \tag{3.37}$$

The first term is estimated by Proposition 2.2 and turns out to be superexponentially small and the second one tends to zero as β tends to infinity by (3.36). The proof of point (b) is completely similar.

Point (c) follows from Theorem 2.1(ii) and the evaluation of $E^{(l_c)}(\tau_{+\mathbf{1}}(-\mathbf{1}))$ for the chain $X_t^{(l_c)}$. In fact, by (3.34),

$$\frac{1}{\beta} \log E^{(l_c)}(\tau_{+\mathbf{1}}(-\mathbf{1})) = 0$$

and thus by Theorem 2.1(ii)

$$\frac{1}{\beta} \log E(\tau_{+\mathbf{1}}(-\mathbf{1})) = V_1 + V_2 + \cdots + V_{l_c} = (l_c - 2)h + \bar{\Delta}^{(l_c-2)}(-\mathbf{1}, Q_{l_c}) \tag{3.38}$$

By Eq. (3.27) we have

$$\begin{aligned} \bar{\Delta}^{(l_c-2)}(-\mathbf{1}, Q_{l_c}) &= 8 + 4(l_c - 2) - h[3 + 3(l_c - 2) + (l_c - 1)(l_c - 2)] \\ &= 4l_c - h[l_c^2 - 1] \end{aligned} \tag{3.39}$$

and by (3.38)

$$\frac{1}{\beta} \log E(\tau_{+\mathbf{1}}(-\mathbf{1})) = (l_c - 2)h + 4l_c - h[l_c^2 - 1] = 4l_c - h[l_c^2 - l_c + 1] \tag{3.40}$$

Points (e) easily follow from Theorem 2.2 and standard arguments (see, e.g., ref. 10):

If we define

$$f(\sigma, t) = P(\tau_{+1}(\sigma) > tE(\tau_{+1}(-\mathbf{1}))) \tag{3.41}$$

we have to prove that

$$f(-1, t+s) = f(-1, t)\, f(-1, s) + o(\beta) \tag{3.42}$$

with $o(\beta) \to 0$ as $\beta \to \infty$. This is an easy consequence of Theorem 2.2; in fact, by using the Markov property and the fact that $S^{(l_c)} = \{+\mathbf{1}, -\mathbf{1}\}$ we obtain

$$f(-1, t+s) = f(-1, t)\, f(-1, s) + P(X_{tE(\tau_{+1}(-\mathbf{1}))} \notin S^{(l_c)}) \tag{3.43}$$

Point (d) easily follows from points (c) and (e).

4. SOME REMARKS ON THE RENORMALIZATION PROCEDURE

As is well known in the theory of Markov chains (see, e.g., ref. 1), the invariant measure of the chain X_t with finite state space S and transition probabilities $P(i, j)$ can be computed in terms of a graph technique which is summatized here.

Let W be a subset of the state space S; a W-graph is defined as a graph consisting of arrows $m \to n$, $m \in S \setminus W$, $n \in S$, $n \neq m$, satisfying the following conditions:

1. Precisely one arrow comes from any $m \in S \setminus W$.
2. There are no closed cycles in the graph.

The set of all W-graphs is denoted by $G(W)$ and for any graph $g \in G(W)$ we define $\pi(g) = \prod_{m \to n \in g} P(m, n)$. With these notations it is very easy to verify that the invariant measure ν of the chain X_t is given by

$$\nu(i) = \frac{Q_i}{\sum_{j=1}^{|S|} Q_j} \tag{4.1}$$

where

$$Q_i = \sum_{g \in G\{i\}} \pi(g) \tag{4.2}$$

by showing that Q_i satisfy the invariant equation

$$Q_i = \sum_{j=1}^{|S|} Q_j P(j, i) \tag{4.3}$$

If we suppose that the chain X_t is ergodic and satisfies Hypothesis P [where we consider $\Delta(x, y) = \infty$ if $P(x, y) = 0$], then the quantities $\pi(g)$ can be estimated by

$$e^{-\Delta(g)\beta - |S|\gamma\beta} \leqslant \pi(g) \leqslant e^{-\Delta(g)\beta + |S|\gamma\beta} \tag{4.4}$$

where

$$\Delta(g) = \sum_{m \to b \in g} \Delta(m, n) \tag{4.5}$$

and if we define

$$V\{i\} = \min_{g \in G\{i\}} \Delta(g) \tag{4.6}$$

then the invariant measure of the chain can be easily estimated in terms of $e^{-V\{i\}\beta}$.

Our renormalization procedure in this context corresponds to breaking up the problem into finding the graph g minimizing $V\{i\}$ in simpler minimization problems on a sequence of independent scales. In fact, we can note that if we consider the first chain constructed in our analysis, that is, the chain $\bar{X}_t$ on the stable states, and if we suppose to know the graph $\bar{g}_i^*$ minimizing $\bar{V}\{i\}$, defined as (4.6) for the chain $\bar{X}_t$, we can then construct a graph g_i^* minimizing $V\{i\}$. In fact, to each $m \to n \in \bar{g}_i^*$, m, $n \in M$, we associate the sequence of arrows defined by the function ϕ which minimizes the quantity $\bar{\Delta}(m, n)$ and we add to the set of arrows constructed in this way all the sequences of arrows starting from each unstable state not yet touched by the previous sequences of arrows, and going to a stable state and such that their contribution to $\Delta(g)$ is zero (see the first part of the proof of Proposition 2.2 for a proof that this is always possible). In this way we have defined a set of arrows containing an $\{i\}$-graph g_i^* such that

$$\Delta(g_i^*) \leqslant \bar{V}\{i\} + (|M| - 1) V_1 \tag{4.7}$$

On the other hand, we have the inequality

$$V\{i\} \geqslant \bar{V}\{i\} + (|M| - 1) V_1 \tag{4.8}$$

In fact, if g is a graph minimizing $V\{i\}$, then for each $k \in M$ it contains a sequence of arrows going from k to another point in M, say j, without touching other stable states and contributing to $V\{i\}$ with a quantity greater than or equal to $\bar{\Delta}(k, j)$. The set of these transitions between stable states provides an $\{i\}$-graph in the space M.

From (4.7) and (4.8) we can conclude that the graph g_i^* constructed above minimizes $V\{i\}$. By repeating this argument on each scale we obtain the solution of the minimization problem.

A similar discussion can be developed for the problem of the exit of the Markov process from a given domain, which can be analyzed in terms of graphs.

ACKNOWLEDGMENT

Special thanks are due to R. Schonmann for his accurate and constructive criticism and for many suggestions during the preparation of this paper. I also thank L. Bertini, N. Cancrini, M. Cassandro, G. Jona-Lasinio, R. Kotecky, F. Martinelli, W. Metzner, G. Nappo, E. Olivieri, C. Parrinello, and M. Virasoro for many important discussions and constructive advice. This research has been partially supported by SC1-CT91-0695 of the Commission of European Communities.

REFERENCES

1. M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems* (Springer-Verlag, Berlin, 1984).
2. R. Z. Has'minskii, *Stochastic Stability of Differential Equations* (Sijthoff-Noordhoff, 1980).
3. E. J. Neves and R. H. Schonmann, Behaviour of droplets for a class of Glauber dynamics at very low temperatures, *Commun. Math. Phys.* **137**:209 (1991).
4. E. J. Neves and R. H. Schonmann, Critical droplets and metastability for a Glauber dynamics at very low temperatures, *Prob. Theory Related Fields* **91**:331 (1992).
5. S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Optimization by simulated annealing, *Science* **220**:671 (1983).
6. R. Kotecky and E. Olivieri, Droplet dynamics for asymmetric Ising model, Preprint (1992).
7. R. Kotecky and E. Olivieri, Shapes of growing droplets—A model of escape from a metastable phase, in preparation.
8. F. Martinelli, E. Oliveri, and E. Scoppola, On the Swendsen and Wang dynamics II: Critical droplets and homogeneous nucleation at low temperature for the 2 dimensional Ising model, *J. Stat. Phys.* **62**:1135 (1991).
9. F. Martinelli, E. Olivieri, and E. Scoppola, Metastability and exponential approach to equilibrium for low temperature stochastic Ising models, *J. Stat. Phys.* **61**:1105 (1990).
10. F. Martinelli, E. Olivieri, and E. Scoppola, Small random perturbations of finite and infinite dimensional dynamical systems: Unpredictability of exit times, *J. Stat. Phys.* **55**:477 (1989).
11. D. Revuz, *Markov Chains* (North-Holland, Amsterdam, 1975), Chapter 1, Section 3.
12. E. Scoppola, Metastability and nucleation for 2-dimensional Ising systems, *Physica A* (Proceedings Statphys 18—Berlin 1992), to appear.