METASTABILITY FOR MARKOV CHAINS: A GENERAL PROCEDURE BASED ON RENORMALIZATION GROUP IDEAS

ELISABETTA SCOPPOLA*
Dipartimento di Fisica
Università 'La Sapienza'
Piazzale A. Moro 2
00185 Roma
Italy
e-mail: scoppola@roma1.infn.it

Abstract. The paper is a report on results on the long time behavior of Markov chains with finite state spaces and with transition probabilities exponentially small in an external parameter β. A general approach based on renormalization group ideas is presented and discussed in the simple case of reversible Markov chains. Applications are also discussed.

Key words: Markov chains, renormalization group, metastability, reversibility, invariant measure, first hitting time, Metropolis algorithm.

1. Introduction

In this note I review some results on the long time behavior of Markov chains characterized by the following property: the state space S is discrete and finite and the transition probabilities $P(x,y)$ can be estimated from above and from below exponentially in a large parameter β: if $P(x,y) > 0$ then

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

where $\Delta(x,y)$ has non-negative values (0 is a possible value), β is sufficiently large and γ tends to zero as β tends to infinity.

Markov chains of this kind arise for instance in Monte Carlo simulations of statistical mechanics models at low temperature (see, e.g., [S2]), or in the Freidlin–Wentzel analysis of diffusion processes given by small random perturbations of dynamical systems [FW].

Let us consider, as an example, the Markov chain defined by the Metropolis algorithm for the two dimensional ferromagnetic Ising model in a finite box $\Lambda \subset \mathbb{Z}^2$ with external magnetic field $h > 0$.

To each $i \in \Lambda$ we associate a spin variable $\sigma(i) = \pm 1$ and to each spin configuration $\sigma \in \{-1,+1\}^\Lambda = S$ we associate the Hamiltonian:

$$H_\Lambda(\sigma) = -\frac{1}{2} \sum_{i,j\in\Lambda,\ |i-j|=1} \sigma(i)\sigma(j) - \frac{h}{2}\sum_{i\in\Lambda}\sigma(i) \tag{1.2}$$

*Partially supported by grant SC1-CT91-0695 of the Commission of European Communities.

where h is a uniform positive external magnetic field, and we can consider for instance periodic boundary conditions.

Mean values of arbitrary observables with respect to the Gibbs measure

$$\mu_\Lambda(\sigma) = \frac{e^{-\beta H_\Lambda(\sigma)}}{Z_\Lambda} \tag{1.3}$$

(where Z_Λ is the partition function), can be computed by using the Monte Carlo method, by defining a Markov chain $\{\sigma_t\}_{t\in\mathbb{N}}$ with state space $S = \{-1, +1\}^\Lambda$, and transition probabilities $P(\sigma, \sigma') = P(\sigma_{t+1} = \sigma' \,|\, \sigma_t = \sigma)$ satisfying the detailed balance condition:

$$\mu_\Lambda(\sigma)P(\sigma, \sigma') = \mu_\Lambda(\sigma')P(\sigma', \sigma) \tag{1.4}$$

and the ergodicity condition:

$$\exists n \in \mathbb{N} \text{ such that, } \forall \sigma, \eta,\ P^n(\sigma, \eta) > 0 \tag{1.5}$$

where $P^n(.,.)$ is the n-step transition probability.

An explicit construction of this Markov chain can be given by the Metropolis algorithm which is defined as follows: for any $\sigma \in S$ and any i in Λ let

$$\Delta_i H(\sigma) \;=\; H_\Lambda(\sigma^i) - H_\Lambda(\sigma) \tag{1.6}$$

with

$$\sigma^i(j) = \begin{cases} \sigma(j) & \text{if } i \neq j \\ -\sigma(j) & \text{if } i = j. \end{cases} \tag{1.7}$$

We consider the following transition probabilities: if $\sigma \neq \sigma'$

$$P(\sigma, \sigma') = \begin{cases} 0 & \text{if} \quad \sigma' \neq \sigma^i \text{ for all } i \in \Lambda \\ |\Lambda|^{-1} \exp\{-\beta(\Delta_i H(\sigma) \vee 0)\} & \text{if} \quad \sigma' = \sigma^i \text{ for some } i. \end{cases} \tag{1.8}$$

$P(\sigma, \sigma)$ is obtained by normalization.

We will denote by $\sigma_t(\sigma)$ the process starting at σ.

The detailed balance condition (1.4) ensures that the invariant measure of the chain σ_t is the Gibbs measure and thus it is concentrated on the configuration $+\mathbf{1}$ in which all the spins are plus. The configuration $-\mathbf{1}$ is only a local minimum of the Hamiltonian H (remember $h > 0$) and if we define $E(l) \equiv H(\sigma_{(l)}) - H(-\mathbf{1})$, where $\sigma_{(l)}$ is a configuration in which the plus spins form a square of side l, it is very simple to compute $E(l) = 4l - hl^2$ which is maximum for $l = 2/h$. This means that the magnetic field determines the phase even if it is very small, but its effects become relevant only on a scale sufficiently large ($l \geq l_c(h) \sim 2/h$), as only on large scales does the volume energy dominate the surface energy. Neves and Schonmann have studied in [NSch1, 2] this metastable behavior of the state $-\mathbf{1}$ from a dynamical point of view by showing that the Markov chain σ_t, defined above, starting from the configuration $-\mathbf{1}$, locally undergoes only small fluctuations around the metastable state $-\mathbf{1}$ for a certain amount of time, very large if β/h is large, until it will 'tunnel' to the true equilibrium $+\mathbf{1}$. 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 sides are 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.

These results have been obtained in [NSch1, 2] with h fixed, the side L of the box Λ sufficiently large (i.e., $L \gg 2/h$) and β large enough.

More precisely we can summarize these results in the following theorem. Let $\tau_\eta(\sigma) \equiv \inf\{t \geq 0; \sigma_t(\sigma) = \eta\}$ be the first hitting time to the configuration η starting from σ, so that $\tau_{+\mathbf{1}}(-1)$ 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 1.1 [NSch1, 2]. *For any h arbitrarily small, $h < 2$, with $2/h$ not an integer and Λ sufficiently large,*

(a) *for all $\eta \in R$,*

$$\lim_{\beta\to\infty} P(\tau_{-\mathbf{1}}(\eta) < \tau_{+\mathbf{1}}(\eta)) = 1 \quad \textit{if } l(\eta) < 2/h,$$
$$\lim_{\beta\to\infty} P(\tau_{+\mathbf{1}}(\eta) < \tau_{-\mathbf{1}}(\eta)) = 1 \quad \textit{if } l(\eta) > 2/h;$$

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

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

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

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

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

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

(d)
$$\lim_{\beta\to\infty} \frac{1}{\beta} \log(\tau_{+\mathbf{1}}(-\mathbf{1})) = \Gamma(h) \quad \textit{in probability};$$

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

Similar results have been obtained by Martinelli, Olivieri and Scoppola [MOS1] for a random cluster algorithm (Swendsen–Wang dynamics) in the thermodynamic limit at low temperature. Nucleation from a metastable state is also studied in [KO1] for an anisotropic Ising model and in [KO2] in the case of isotropic nearest neighbours and next nearest neighbours interactions.

Let us note here that metastability results, like the previous theorem, can be used to prove the exponential convergence to equilibrium of the chain *uniformly* in the volume [MOS2, 3]. More precisely for every local observable f if we denote by $\mu_\Lambda(f)$ its mean value with respect to the Gibbs measure (i.e., $\mu_\Lambda(f) = \sum_\sigma f(\sigma)\mu_\Lambda(\sigma)$), then

$$\sup_{\sigma\in S} |\mu_\Lambda(f) - Ef(\sigma_t(\sigma))| \le C_f e^{-mt}$$

for any $t > t_0(\beta, h)$, where E is the expectation over the process σ_t, C_f is a constant depending only on f and m is independent of Λ.

It is not difficult to show (see, e.g., [S1]) that results like Theorem 1.1 can be easily obtained if one controls the following quantities characterizing the long time behavior of the chain σ_t:

$$\nu(D), \quad \forall D \in S, \tag{1.9}$$

$$E\tau_D(\sigma), \quad \forall D \in S,\ \sigma \in S, \tag{1.10}$$

$$P(\tau_D > t), \quad t \in \mathbb{N}, \tag{1.11}$$

$$P(\sigma_{\tau_D}(\sigma) = \eta), \quad \forall D \in S,\ \sigma \in S,\ \eta \in D, \tag{1.12}$$

where $\nu(.)$ denotes the invariant measure of the chain, E and P are the expectation and the probability and τ_D is the first hitting time to the set D.

I will present in this note results recently obtained in [S1] on the control of these quantities, in a general case, that is for Markov chains satisfying condition (1.1), by means of a general procedure, model-independent, based on renormalization group ideas. Even if the results obtained in [S1] hold for the general class of Markov chains with exponential behavior of transition probabilities, I will discuss here a smaller class of Markov chains obtained by imposing two additional assumptions, reversibility and non-degeneracy, which simplify the construction and the proofs.

Let us conclude this introduction with a short discussion of the main idea behind this renormalization approach. I will denote by $X_t(x)$ an arbitrary Markov chain on the states space S starting from the state $x \in S$, and with transition probability exponentially small in β (the chain σ_t defined above by the Metropolis algorithm is a particular example). I will prove that it is possible to control, with estimates from above and from below, the quantities (1.9)–(1.12) for the process X_t by means of an iterative argument: I will introduce a classification of the states in terms of their stability $S \supseteq S^{(1)} \supseteq S^{(2)} \supseteq \cdots \supseteq S^{(n)}$. This classification enables us to define a sequence of Markov chains $X_t^{(k)}$ defined over the sequence of states spaces $S^{(k)}$ and corresponding to the initial chain X_t viewed on a sequence of times T_1, T_2, T_3, ... This means that the chain $X_t^{(k)}$ is a coarse grain version of the chain X_t in the sense that passing from the chain X_t to the chain $X_t^{(k)}$ we give a less detailed description of the process but we lose information only about events which occur in a typical time less than or equal to T_k.

At each step of such an iteration the quantities (1.9)–(1.12) are estimated in terms of the same quantities for the chain of the next step. Since, in the construction $S^{(k)} \subseteq S^{(k-1)}$ and $S^{(k+1)} \subset S^{(k-1)}$, the idea of the method is to iterate the argument up to an n sufficiently large, such that the space $S^{(n)}$ is sufficiently small and the quantities (1.9)–(1.12) are easily evaluable at this level.

The paper is organized as follows: in Section 2 we will give the precise assumptions on the Markov chain considered, and we will state the main result. In Section

3 we will prove the result. In Section 4 we will consider some applications. In particular we briefly discuss there the problem of the exit of the process by a domain D containing several stable states (i.e., states with an exponentially long mean exit time).

2. Hypotheses and the Main Theorem

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 following conditions:

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

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

(2) Reversibility: Let $\mathcal{F} \subset S^2$ be the space of pairs of states (x,y) such that $P(x,y) > 0$. Then there exists a function H defined on the space $S \cup \mathcal{F}$ with values in $\mathbb{R}$ such that

$$H(x,y) = H(y,x), \tag{2.2}$$

$$H(x,y) \ge H(x) \vee H(y), \tag{2.3}$$

$$\Delta(x,y) = H(x,y) - H(x). \tag{2.4}$$

This implies that the transition probabilities satisfy the detailed balance condition with respect to the measure $\mu(x) = {Z_S}^{-1} \exp\{-\beta H(x)\}$ in the limit $\beta \to \infty$.

(3) Non-Degeneracy: We suppose that

$$H(x) \neq H(y) \quad \forall x \neq y,\ x,y \in S. \tag{2.5}$$

Remarks. Hypothesis (1) is the exact statement of (1.1). Hypothesis (2) is the reversibility property in a general form. The case of the Metropolis algorithm, discussed in the introduction, corresponds to the choice

$$H(x,y) = H(x) \vee H(y).$$

Hypothesis (3) enables us to simplify the exposition; it is obviously satisfied if $|S|$ is finite with an arbitrarily small change in the function H. We note here that hypothesis (3) is not verified by the Metropolis algorithm (not in this strong form). However we want to stress that we use this hypothesis only to simplify the exposition, and the only crucial assumption is property P.

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

Main Theorem. *Let X_t be a Markov chain satisfying the previous conditions then it is possible to define a finite sequence of Markov chains $X_t^{(1)}, X_t^{(2)}, \dots, X_t^{(n)}$ on state spaces $S \supseteq S^{(1)} \supseteq S^{(2)} \supseteq \cdots \supseteq S^{(n)}$ and with $S^{(n)} = x_m$ (where x_m is the* *state of absolute minimum for the function H: $H(x) > H(x_m)$ $\forall x \neq x_m$) such that each Markov chain $X_t^{(k)}$ satisfies hypotheses (1)–(3) with new functions $\Delta^{(k)}$ and $H^{(k)}$. The processes $X_t^{(k)}$ correspond to the chain X_t on a sufficiently large time scale T_k, exponentially long in β, in the following sense. Let W be a subset of S; we denote by ν the invariant measure of the chain, by τ_W the first hitting time to W and by $E_x\tau_W$ its mathematical expectation calculated under the assumption that the initial state of the chain X_t is x; analogous quantities can be defined for the process $X_t^{(k)}$, $k = 1, 2, \ldots$, and we will use the same notation with the superscript ${}^{(k)}$. Then for any β sufficiently large and for any $k = 1, \ldots, n$ let $W \subset S^{(k)}$, $x \in S^{(k)} \backslash W$, $y \in W$:*

(i)

$$P(X_{\tau_W}(x) = y) = P\left(X^{(k)}_{\tau^{(k)}_W}(x) = y\right); \tag{2.6}$$

(ii) *there exists a positive η depending on γ and k, with $\eta \to 0$ as $\beta \to \infty$, such that*

$$e^{-\eta\beta} T_k E_x \tau_W^{(k)} \leq E_x \tau_W \leq e^{\eta\beta} T_k E_x \tau_W^{(k)}; \tag{2.7}$$

(iii) *there exist constants C and γ', with $\gamma' \to 0$ as $\beta \to \infty$, such that for any $B \subset S^{(k)}$:*

$$C.T_k.e^{-\gamma'\beta}\nu^{(k)}(B) \leq \nu(B) \leq C.T_k.e^{\gamma'\beta}\nu^{(k)}(B); \tag{2.8}$$

(iv) *for any $t > T_k e^{\delta\beta}.2^k$, for any $W \subset S^{(k)}$ and for any $x \in S^{(k)}$:*

$$P(\tau_W(x) > t) \leq P\left(\tau_W^{(k)}(x) > \frac{t}{T_k 2^k}\right) + \exp\left\{-c_1 e^{\delta\beta}\right\}; \tag{2.9}$$

for some constant c_1; moreover there exists a constant Δ such that

$$P(X_t(x) \notin S^{(k+1)}) \leq e^{-\Delta\beta}. \tag{2.10}$$

The previous construction is explicit and the quantities $\Delta^{(k)}(x, y)$, T_k and the state spaces $S^{(k)}$ are explicitly defined in terms of the quantities $\{\Delta(u, v)\}_{u,v\in S}$.

Remarks. Point (i) is a consequence of the fact that, if hypothesis (3) holds, then the definition of the chains $X_t^{(k)}$ is a path by path construction. In the general case in which only hypothesis (1) is verified, one can obtain estimates from above and from below of the probabilities appearing in (2.6).

Point (ii) is the exact statement of the time rescaling.

With point (iii) we have a relation among the invariant measures of the chains at different steps of the iteration. We want to remark here that the classification of states induced by the construction of the sets $S^{(k)}$ does not correspond to the classification of states given by the invariant measure ν, i.e., given by the Hamiltonian in our reversible case, even if $x_m \in S^{(k)}$ for any $k \leq n$. This means that some state in $S^{(k)}$ could have an invariant measure exponentially smaller than the invariant measure of some state in $S \backslash S^{(k)}$.

Estimate (2.9) is a quite crude bound based on the Chebyshev estimate. It is sufficient to prove that the classification of states in the sets $S^{(k)}$ is strictly related to the time scales T_k. More precisely, with probability exponentially near to one, the process is in a state in $S^{(k)}$ after a time of order T_k (see (2.10)). In other words, the classification of states considered here is based on the stability of the states.

3. Proof of the Main Theorem

We will construct in details the first chain $X_t^{(1)}$ (see Subsections 3.1 and 3.2) and we will prove that it satisfies hypotheses (1)–(3). For this chain we will verify points (i)–(iv) of the theorem (see 3.3). The proof of the theorem will follow by induction (Subsections 3.4 and 3.5).

3.1. The Stable States

We define the state x in S to be stable if and only if it is a local minimum of the function H, i.e., $H(x) < H(x,y)$ for any $x \neq y$.

We will denote by M the set of local minima (i.e., stable states). For each $x \in S$ we can define the first hitting time to the set M:

$$\tau_M(x) \equiv \min\{t \geq 0;\ X_t(x) \in M\} \tag{3.1}$$

corresponding to the time spent by the process outside the set M. In order to obtain estimates on the time τ_M, following the ideas developed in [FW], to each function $\phi : \mathbb{N} \to S$, $\phi = \{\phi_t\}_{t\in\mathbb{N}}$, we associate a functional

$$I_{[0,t]}(\phi) \equiv \sum_{i=0}^{t-1} \Delta(\phi_i, \phi_{i+1}) \tag{3.2}$$

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 very easily proved (see [S1]).

Proposition 3.1. *Let ϕ be a fixed function starting at x at time 0. Then*

(i)
$$P\big(X_s(x) = \phi_s,\ \forall s \in [0,t]\big) \leq \exp\left\{-I_{[0,t]}(\phi)\beta + \gamma t\beta\right\};$$

(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:*

$$P\big(X_s(x) = \phi_s,\ \forall s \in [0,t]\big) \geq \exp\left\{-I_{[0,t]}(\phi)\beta - \gamma t\beta\right\};$$

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

$$\sup_x P(I_{[0,t]}(X_s(x)) \geq r) \leq e^{-r\beta+\epsilon\beta}$$

where
$$\epsilon = \frac{3r}{\Delta_1}\left[\gamma + \frac{\ln t}{\beta}\right].$$

An immediate consequence of Proposition 3.1 is the following.

Proposition 3.2. *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 $t > T_0$,*
$$\sup_{x\in S} P(\tau_M(x) > t) \leq a^{[t/T_0]}$$

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

(ii) *for any $t \geq e^{\delta\beta}$*

$$\sup_{x \in S} P(\tau_M(x) > t) \leq \exp\left\{-e^{\delta\beta/2}\right\};$$

(iii) *let $\alpha \in (\delta, \frac{1}{2}\Delta_1 - \gamma)$, for any $t > e^{\alpha\beta}$ we have*

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

The idea of the proof of this proposition is very simple: if $x \notin M$ this means that there exists a sequence of unstable states $x_0 = x, x_1, x_2, \ldots, x_n$ such that $H(x_i) > H(x_{i+1})$ and $H(x_i, x_{i+1}) = H(x_i)$ for any $i = 0, \ldots, n$. Such a sequence is finite and thus it ends in a stable state x_{n+1}. By using Proposition 3.1 it is easy to prove that with large probability the process follows such a sequence of states and thus it reaches M in a time independent of β (see [S1] for a detailed proof).

3.2. The Rescaled Markov Chain $X_t^{(1)}$

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 $X_t^{(1)}$ with state space $S^{(1)} \equiv M$, corresponding to the original process looked at times sufficiently large. Let us define:

$$V_1 \equiv \min_{x \in M,\, y \in S,\, x \neq y} \left\{H(x,y) - H(x)\right\}, \tag{3.3}$$

$$t_1 \equiv e^{V_1\beta + \delta\beta}. \tag{3.4}$$

We define a sequence of stopping times

$$\sigma_1 \equiv \min\{t > 0;\, X_t \neq X_0\}, \tag{3.5}$$

$$\tau_1 \equiv \min\{t \geq \sigma;\, X_t \in M\}, \tag{3.6}$$

$$\zeta_1 = \begin{cases} t_1 & \text{if } \sigma_1 > t_1, \\ \tau_1 & \text{if } \sigma_1 \leq t_1, \end{cases} \tag{3.7}$$

and for each $n > 1$,

$$\sigma_n \equiv \min\{t > \zeta_{n-1};\, X_t \neq X_{\zeta_{n-1}}\}, \tag{3.8}$$

$$\tau_n \equiv \min\{t \geq \sigma_n;\, X_t \in M\}, \tag{3.9}$$

$$\zeta_n = \begin{cases} \zeta_{n-1} + t_1 & \text{if } \sigma_n - \zeta_{n-1} > t_1, \\ \tau_n & \text{if } \sigma_n - \zeta_{n-1} \leq t_1. \end{cases} \tag{3.10}$$

It is simple to prove that these times are stopping times with respect to the σ-algebra associated with the Markov chain [S1], and thus the sequence $X_n^{(1)} = X_{\zeta_n}$ is an homogeneous Markov chain. For any $x \in M$, we can then consider the new Markov chain $X_t^{(1)}$ with $X_0^{(1)}(x) = x$ on the state space $M \equiv S^{(1)}$.

Let us note that this new Markov chain is strictly related to the time scale t_1 (see (3.4)) in the sense that, as we will show later on,

$$P\Big(\zeta_{n+1} - \zeta_n \in [t_1 e^{-\gamma\beta}, t_1 e^{\gamma\beta}]\Big) \sim 1.$$

For any pair of states $x, y \in M$ we denote by $P^{(1)}(x,y)$ the transition probability of the chain $X_n^{(1)}$, that is

$$P^{(1)}(x,y) = P(X_{\zeta_n} = y \mid X_{\zeta_{n-1}} = x). \tag{3.11}$$

We will prove the following.

Proposition 3.3. *There exists $\beta_0 > 0$ such that for any $\beta > \beta_0$ and for any $x, y \in M$ with $x \neq y$ 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*

$$P^{(1)}(x,y) = 0,$$

where $M_{x,y}$ is the set of the minima with the exception of the states x, y.

(b) *Otherwise, if the quantity*

$$\overline{\Delta}(x,y) = \inf\big\{I_{[0,t]}(\phi) : t, \phi \text{ such that } \phi_0 = x,\ \phi_t = y,\ \phi_s \notin M_{x,y},\ \forall s \in (0,t)\big\} \tag{3.12}$$

is well defined, then

$$t_1\, e^{-\overline{\Delta}(x,y)\beta - \overline{\gamma}\beta} \ \leq\ P^{(1)}(x,y)\ \leq\ t_1\, e^{-\overline{\Delta}(x,y)\beta + \overline{\gamma}\beta}. \tag{3.13}$$

The quantity $\overline{\Delta}(x,y)$ can assume the values $\overline{\Delta}_1 = V_1 < \overline{\Delta}_2 < \cdots < \overline{\Delta}_{\overline{m}}$ with $\overline{\Delta}_{\overline{m}} < |S|\Delta_m$. The quantity $\overline{\gamma}$ is given by the following:

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

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

Proof.

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

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

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

(b) *Estimate from below*: Let $x \neq y$ and let $\overline{t}_1 = e^{V_1\beta - \gamma\beta}$; we denote by $\overline{\phi}_s^{x,y}$ the function going from x to y minimizing the quantity $\overline{\Delta}(x,y)$ and by $\overline{t}^{x,y}$ the corresponding time for which we have the trivial and crude estimate $\overline{t}^{x,y} < T \equiv |S|$. We have that

$$P^{(1)}(x,y) \geq \sum_{n=1}^{[\overline{t}_1/T]} P_x\Big(\{\sigma_1 > nT\} \cap \{X_s = \overline{\phi}_s^{z,y},\ \forall s \in [nT, nT + \overline{t}^{z,y}]\}\Big)$$
$$\geq \sum_{n=1}^{[\overline{t}_1/T]} P_x(\sigma_1 > nT) e^{-\overline{\Delta}(x,y)\beta - \gamma T\beta} \tag{3.14}$$

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

$$P_x(\sigma_1 > 1) \geq 1 - e^{-V_1\beta + \gamma\beta},$$

and by the Markov property:

$$P_x(\sigma_1 > s) \geq (1 - e^{-V_1\beta + \gamma\beta})^s,$$

which implies that

$$P_x(\sigma_1 > nT) \geq P_x(\sigma_1 > \overline{t}_1) \geq (1 - e^{-V_1\beta + \gamma\beta})^{\overline{t}_1} \geq \tfrac{1}{2} e^{-1} \tag{3.15}$$

for β sufficiently large. Estimate (3.15) in (3.14) gives

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

with

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

Estimate from above:

$$P^{(1)}(x,y) \leq P\big(\{X_{\tau_1}(x) = y\} \cap \{\tau_1 - \sigma_1 \leq e^{\delta\beta}\}\big) + P(\tau_1 - \sigma_1 > e^{\delta\beta}). \tag{3.16}$$

The first term on the right-hand side of (3.19) is bounded above by

$$\sum_{s=1}^{t_1} P\Big(\{X_u = x\ \forall u \leq s\} \cap \{I_{[s, s+e^{\delta\beta}]}(X_t(x)) \geq \overline{\Delta}(x,y)\}\Big)$$
$$\leq \sum_{s=1}^{t_1} P\big(I_{[0, e^{\delta\beta}]}(X_t(x)) \geq \overline{\Delta}(x,y)\big) \leq t_1 e^{-\overline{\Delta}(x,y)\beta + \epsilon\beta} \tag{3.17}$$

from Proposition 2.1(iii) with

$$\epsilon = \frac{\overline{\Delta}(x,y)}{\Delta_1}[\gamma + \delta].$$

The second term on the right-hand side of (3.16) has, by Proposition 2.2(ii), a super-exponential estimate giving the following upper bound for (3.17):

$$t_1 e^{-\overline{\Delta}(x,y)+\overline{\gamma}_2}$$

with

$$\overline{\gamma}_2 \leq \frac{\overline{\Delta}(x,y)}{\Delta_1}[\gamma+\delta]+\delta.$$

The proposition is proved by choosing $\overline{\gamma} = \max\{\overline{\gamma}_1, \overline{\gamma}_2\}$. □

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

Proposition 3.4. *For any $x, y \in M \equiv S^{(1)}$ with $x \neq y$, let*

$$\Delta^{(1)}(x,y) \equiv \overline{\Delta}(x,y) - V_1 \tag{3.18}$$

and let $\mathcal{F}^{(1)} \subset (S^{(1)})^2$ be the space of pairs of states (x,y) such that $P^{(1)}(x,y) > 0$. Then there exists a function $H^{(1)}$ on the space $S^{(1)} \cup \mathcal{F}^{(1)}$ with values in $\mathbb{R}$ such that

$$H^{(1)}(x,y) = H^{(1)}(y,x), \tag{3.19}$$

$$H^{(1)}(x,y) \geq H^{(1)}(x) \vee H^{(1)}(y), \tag{3.20}$$

$$\Delta^{(1)}(x,y) = H^{(1)}(x,y) - H^{(1)}(x), \tag{3.21}$$

$$H^{(1)}(x) \neq H^{(1)}(y), \ \forall x \neq y, \ x, y \in S^{(1)}. \tag{3.22}$$

Proof. We define

$$H^{(1)}(x) \equiv H(x), \ \forall x \in S^{(1)}, \tag{3.23}$$

and

$$H^{(1)}(x,y) \equiv \overline{\Delta}(x,y) - V_1 + H(x). \tag{3.24}$$

Then equation (3.22) immediately follows by hypothesis (3) on the initial function H, and equation (3.21) is a consequence of (3.18). To prove equation (3.19) let us introduce the following definition: for any function $\phi : \mathbb{N} \to S$ starting at x and such that $\phi_t = y$ let $(\phi)^{\mathrm{tr}}$ be the time reversed function going from y to x that is

$$(\phi)^{\mathrm{tr}}_s = \phi_{t-s}.$$

By the definition of $\Delta(.,.)$ and by the symmetry of $H(x,y)$ it is trivial to verify that

$$I_{[0,t]}(\phi) + H(x) = I_{[0,t]}((\phi)^{\mathrm{tr}}) + H(y). \tag{3.25}$$

This identity implies that if we denote by $\overline{\phi}^{x,y}$ the function going from x to y and minimizing the functional I in the definition of $\overline{\Delta}(x,y)$, then

$$(\overline{\phi}^{x,y})^{\mathrm{tr}} = \overline{\phi}^{y,x}$$

and thus, using again (3.25) and the definition of $H^{(1)}(x,y)$, we prove immediately its symmetry property (3.19).

Equation (3.20) easily follows from definition (3.24), equation (3.19) and the fact that the quantity $\overline{\Delta}(x,y) - V_1$ by definition is non-negative. □

3.3. Proof of (i)–(iii) for $X_t^{(1)}$

We have now to prove points (i)–(iii) of our main theorem for the chain $X_t^{(1)}$.

Let W be a subset of $S^{(1)}$, and τ_W its first hitting time. Then for any $x \in S^{(1)} \setminus W$, $y \in W$,

$$P\left(X^{(1)}_{\tau^{(1)}_W}(x) = y\right) = P\left(X_{\zeta_{\tau^{(1)}_W}}(x) = y\right) = P(X_{\tau_W}(x) = y)$$

since by definition $\zeta_{\tau^{(1)}_W} = \tau_W$.

The proof of point (ii) is based on the application of the strong Markov property and on the fact that the Markov time ζ_1 can be estimated from above and from below with times of order t_1. In fact

$$\begin{aligned}
E_x \tau_W &= E_x \sum_{n=1}^{\infty} \zeta_n \chi(\tau^{(1)}_W = n) \\
&= E_x \sum_{n=1}^{\infty} [\zeta_1 + \zeta_2 - \zeta_1 + \cdots + \zeta_n - \zeta_{n-1}] \\
&\qquad\qquad \times \chi(X_{\zeta_1} \notin W)\chi(X_{\zeta_2} \notin W)\ldots\chi(X_{\zeta_{n-1}} \notin W)\chi(X_{\zeta_n} \in W) \\
&= \sum_{n=1}^{\infty}\sum_{m=1}^{n-2} E_x\Big\{\chi(X_{\zeta_1} \notin W)\chi(X_{\zeta_2} \notin W)\ldots\chi(X_{\zeta_m} \notin W) \\
&\qquad\qquad \times E_{X_{\zeta_m}}[\zeta_1\chi(X_{\zeta_1} \notin W)E_{X_{\zeta_1}}\chi(X_{\zeta_2} \notin W)\ldots\chi(X_{\zeta_{n-m-1}} \in W)]\Big\} \\
&\qquad\qquad + \sum_{n=1}^{\infty} E_x[\chi(\overline{\tau}_W > n-1)E_{X_{\zeta_{n-1}}}\zeta_1\chi(X_{\zeta_1} \in W)].
\end{aligned} \tag{3.26}$$

Since ζ_1 is given by

$$\zeta_1 = \sum_{s=1}^{t_1}(s + \tau_M(X_s))\chi(\sigma_1 = s) + t_1\chi(\sigma_1 > t_1) \tag{3.27}$$

we have the following estimates for ζ_1:

$$\begin{aligned}
&\zeta_1 \le t_1 + K + \sum_{j>K} j\chi\big(\exists x \in S\setminus M;\ \tau_M(x) = j\big), \\
&\zeta_1 \ge \widetilde{t}_1 - \widetilde{t}_1\chi\big(\exists x \in M;\ \sigma_1(x) \le \widetilde{t}_1\big),
\end{aligned}$$

with $\widetilde{t}_1 = e^{V_1\beta - 2\gamma\beta}$. By applying these estimates in (3.26) it is simple to show (see [S1]) that (2.7) holds.

Point (iii) is an easy consequence of the fact that there exists a constant C such that for any $B \subset S$ (see [H])

$$\nu(B) = C\sum_{y\in M}\nu^{(1)}(y)E_y\sum_{t=0}^{\zeta_1}\chi_B(X_t) \tag{3.28}$$

where the constant C is fixed by normalization.

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

$$\nu(B_1) = C \sum_{y \in B_1} \nu^{(1)}(y) E_y \sum_{t=0}^{\zeta_1} \chi_{\{y\}}(X_t)$$

and for any $y \in M$ we have:

$$t_1 \geq E_y \sum_{t=0}^{\zeta_1} \chi_{\{y\}}(X_t) \geq E_y(\sigma_1 \vee \overline{t}_1) \geq \overline{t}_1 P(\sigma_1 > \overline{t}_1) \geq \tfrac{1}{2} e^{-1} \overline{t}_1,$$

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

We remark here that since the chain $X_t^{(1)}$ satisfies the reversibility condition with respect to the function $H^{(1)}$ and $H^{(1)}(x) = H(x)$ for any $x \in M$, then point (iii) of the main theorem can be proved also by a direct computation and by using Lemma 3.2 of [FW, Chap. 6] concerning the invariant measure of chains satisfying estimates like (2.1).

3.4. The Iteration Scheme

The iteration scheme now is the following. For any $k \geq 1$ we define the following quantities: for any $\phi : \mathbb{N} \to S^{(k)}$,

$$I_{[0,t]}^{(k)}(\phi) = \sum_{i=0}^{t-1} \Delta^{(k)}(\phi_i, \phi_{i+1}), \tag{3.29}$$

$$M^{(k)} = \{x \in S^{(k)}; \forall y \in S^{(k)},\ y \neq x,\ H^{(k)}(x) < H^{(k)}(x,y)\}, \tag{3.30}$$

$$\overline{\Delta}^{(k)}(x,y) = \inf\Big\{ I_{[0,t]}^{(k)}(\phi) : t, \phi, \phi_0 = x, \phi_t = y, \phi_s \notin M^{(k)},\ \forall s \in [0,t] \Big\}, \quad \forall x, y \in M^{(n)}, \tag{3.31}$$

$$S^{(k+1)} = M^{(k)}, \tag{3.32}$$

$$V_{k+1} = \inf\Big\{ H^{(k)}(x,y) - H^{(k)}(x) : x \in M^{(k)}, y \in S^{(k)}, x \neq y \Big\}, \tag{3.33}$$

$$t_{k+1} = e^{V_{k+1}\beta + \delta\beta}, \tag{3.34}$$

$$T_1 = t_1,$$

$$T_{k+1} = t_1 t_2 \dots t_k t_{k+1}, \tag{3.35}$$

$$\Delta^{(k+1)}(x,y) = \overline{\Delta}^{(k)}(x,y) - V_{k+1}, \quad \forall x, y \in S^{(k+1)}, \tag{3.36}$$

$$H^{(k+1)}(x) \equiv H(x), \quad \forall x \in S^{(k)}, \tag{3.37}$$

$$H^{(k+1)}(x,y) \equiv \Delta^{(k+1)}(x,y) + H(x). \tag{3.38}$$

In order to prove that the sequence of state spaces $S^{(1)} \supseteq S^{(2)} \supseteq \cdots \supseteq S^{(n)}$ is finite we note that $S \supset S^{(2)}$. In fact if S contains unstable states then $S \supset$

$M \equiv S^{(1)}$. If S does not contain unstable states then in $S^{(1)}$ there is at least an unstable state since, by the definition of V_1, there exist $x, y \in S = M$ such that $\Delta^{(1)}(x,y) \equiv \overline{\Delta}(x,y) - V_1 = 0$.

By equation (3.37) and by using the property (3.20) at step k, it is trivial to verify that the state x_m, corresponding to the absolute minimum of the function H, is contained in every set $S^{(k)}$.

Points (i)–(iii) of the main theorem are thus proved by iteration, since for each k the chain $X_t^{(k)}$ satisfies hypotheses (1)–(3).

3.5. Proof of (iv)

We are finally going to prove the last point (iv) of the theorem,

$$P(\tau_W(x) > t) = P(\tau_W^{(k)}(x) > \mu^{(k)}(t)) \tag{3.39}$$

where $\mu^{(k)}(t)$ is the number of transitions of the chain $X^{(k)}$ within time t. This probability can be bounded above by

$$P\left(\tau_W^{(k)}(x) > \frac{t}{T_k 2^k}\right) + P\left(\mu^{(k)}(t) < \frac{t}{T_k 2^k}\right), \tag{3.40}$$

and this last term has a super-exponential estimate. In fact,

$$P\left(\mu^{(1)}(t) < \frac{t}{2t_1}\right) = P(\zeta_{m_1} > t) \tag{3.41}$$

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

$$\begin{aligned} P(\zeta_{m_1} > t) &\le e^{-\lambda t}\left[\sup_{x\in M} E_x e^{\lambda\zeta_1}\right]^{m_1} \\ &\le e^{-\lambda t} e^{\lambda t_1 m_1}\left[\sup_{x\notin M} E_x e^{\lambda\tau_M}\right]^{m_1} \le e^{-\lambda t} e^{\lambda t_1 m_1} e^{\lambda c m_1}, \end{aligned} \tag{3.42}$$

for some constant c independent of β, if λ is sufficiently small. In conclusion, (3.41) can be bounded above by

$$e^{-\lambda t/2} e^{\lambda c t/(2t_1)} \le e^{-c' t}. \tag{3.43}$$

By iterating the same argument we easily obtain that

$$\begin{aligned} P\left(\mu^{(k)}(t) < \frac{t}{T_k 2^k}\right) &\le P\left(\mu^{(k-1)}(t) < \frac{t}{T_{k-1}2^{k-1}}\right) \\ &\quad + P\left(\left\{\mu^{(k)}(t) < \frac{t}{T_k 2^k}\right\} \cap \left\{\mu^{(k-1)}(t) > \frac{t}{T_{k-1}2^{k-1}}\right\}\right) \\ &\le \exp\left\{-c_1 e^{\delta\beta}\right\}, \end{aligned} \tag{3.44}$$

since the second term on the right-hand side of (3.44) is estimated as in (3.43). With such an estimate we can evaluate by iteration also (2.10) in the following way:

$$\begin{aligned} P(X_t(x) \notin S^{(k+1)}) \leq{} & P(X_t(x) \notin S^{(k)}) \\ & + P\left(\left\{X^{(k)}_{\nu^{(k)}(t)} \notin S^{(k+1)}\right\} \cap \left\{\mu^{(k)}(t) > \frac{t}{T_k 2^k}\right\}\right) \\ & + P\left(\mu^{(k)}(t) < \frac{t}{T_k 2^k}\right) \\ \leq{} & e^{-\Delta\beta}, \end{aligned} \tag{3.45}$$

for some constant Δ, since the second term on the right-hand side of (3.45) can be estimated by Proposition 3.2(iii) applied to the chain $X_t^{(k)}$. This concludes the proof of the main theorem.

3.6. Remarks

In the figure we give a simple example of the iterative construction. We define $S = \{1, 2, \dots, N\}$, with $N = 21$, and $\mathcal{F} = \{(i, i+1)\}_{i=1}^N$ with periodic boundary conditions: $N + 1 = 1$. The Markov chain X_t is characterised by the function $H(x)$, $x \in \mathbb{R}$, where $H(i)$ is the value assumed by H in the integer coordinate i, and $H(i, i+1)$ is defined as $\sup_{x \in [i,i+1]} H(x)$.

We note that at each step of the iteration the function H is 'smoothed'.

We conclude this section with a final remark on the applicability of such a procedure. We note that the procedure described in this paper fixes a stategy for the study of the long time behavior of general Markov chains in a model independent way. However its efficiency turns out to be strictly related to the model considered. This means that the computation of the quantities $\overline{\Delta}^{(k)}(x, y)$, which is necessary to the definition of the new step $k + 1$ of the iterative procedure, could be difficult in some cases.

However a simplification comes from the following remark. In the construction of our first renormalized chain $X_t^{(1)}$ all transitions between states (x, y) such that

$$\Delta(x, y) \geq \Delta(x, z) + \Delta(z, y)$$

for some state z, were not present in the function minimizing the functional $I_{[0,t]}(\phi)$ in the definition of $\overline{\Delta}(.,.)$. The same remark applied at each step of the iteration implies that it will not be necessary to compute exactly the values of $\Delta^{(k)}(.,.)$ for each pair of states in $S^{(k)}$, but for some transitions an estimate will be sufficient. This remark greatly simplifies, for instance, the analysis of the chain given by the Metropolis algorithm for the two dimensional Ising model discussed in [S1].

4. Applications

The renormalization procedure discussed in the previous section can be usefully applied to the case of the Metropolis algorithm defined in the introduction, thus

Fig. 1. An example of the iterative construction.

providing a proof of the metastability theorem stated in the introduction completely different from that given in [NSch1, 2]. A detailed analysis of this case can be found in [S1].

We want here to mention another problem which can be easily discussed by using the sequence of Markov chains $X_t^{(k)}$ defined in the previous section. Let us consider a chain X_t satisfying hypotheses (1)–(3) of Section 2, and we denote by $M \subseteq S$ the set of stable states corresponding to local minima of the function H. Let $D \subset S$ be a set containing several stable states, and we consider the problem of the first exit of the process $X_t(x)$, $x \in D$, from the domain D.

In [FW], Freidlin and Wentzel approached such a problem by means of a graphical technique providing estimates from above and from below of the probability that the process exits in a given state and estimating the mean exit time σ_D. More precisely, for any set of states $W \subset S$ define a W-graph as a graph consisting of arrows $m \to n$ with $m \in S \backslash W$ and $n \in W$, $m \neq n$ and satisfying the following properties:

(1) every state $m \in S \backslash W$ is the initial point of exactly one arrow,
(2) there are no closed cycles in the graph.

Condition (2) can be replaced by

(2′) for any state $m \in S \backslash W$ there exists a sequence of arrows leading from it to some point $n \in W$.

The set of W-graphs is denoted by $G(W)$, for any $i \in S \backslash W$ and $j \in W$ we denote by $G_{ij}(W)$ the set of W-graphs in which the sequence of arrows leading from i into W ends at the point j.

Let $G(i \not\to W)$ be the set of graphs containing $|S \backslash W| - 1$ arrows $m \to n$, $m \in S \backslash W$, $n \in S$, $m \neq n$, and not containing chains of arrows leading from i into W.

With this notation, we summarize the Freidlin–Wentzel results in the following way.

Theorem 4.1 [FW]. *For any $x \in D$ let*

$$\sigma_D \equiv \min\{t > 0;\, X_t(x) \notin D\};$$

then

$$\lim_{\beta \to \infty} \frac{1}{\beta} \ln E\sigma_D(x) = W_D - M_D(x). \tag{4.1}$$

For any β sufficiently large there exists $\delta > 0$ with $\delta \to 0$ as $\beta \to \infty$, such that for any $x \in D$ and $y \in \partial D$,

$$\exp\bigl\{-\beta(W_D(x,y) - W_D - \delta)\bigr\} \le P(X_{\sigma_D}(x) = y) \le \exp\bigl\{-\beta(W_D(x,y) - W_D + \delta)\bigr\}, \tag{4.2}$$

where

$$W_D = \min_{g \in G(D^c)} \sum_{m \to n \in g} \Delta(m,n), \tag{4.3}$$

$$M_D(x) = \min_{g \in G(x \not\to D^c)} \sum_{m \to n \in g} \Delta(m,n), \tag{4.4}$$

$$W_D(x,y) = \min_{g \in G_{xy}(D^c)} \sum_{m \to n \in g} \Delta(m,n). \tag{4.5}$$

Let us now suppose that the domain D contains a unique stable state x_0 and that there exists a unique function $\overline{\phi}^{x_0,\partial D}$ starting at x_0, exiting for the first time from D in a time t, minimizing the functional $I_{[0,t]}(\phi)$ and such that $\overline{\phi}_i^{x_0,\partial D} \neq \overline{\phi}_j^{x_0,\partial D}$ for any $i \neq j$. Then if we denote by θ the last time at which the process visits the stable state x_0 before it leaves the set D,

$$\theta \equiv \max\{t < \sigma_D;\, X_t(x_0) = x_0\}, \tag{4.6}$$

then

$$\lim_{\beta \to \infty} P\left(X_s(x_0) = \overline{\phi}_{s-\theta}^{x_0,\partial D},\ \forall s \in (\theta, \sigma_D)\right) = 1. \tag{4.7}$$

The natural question at this point is the following. Is it possible to have a more detailed description of the large deviation producing the exit? In other words: is it possible to control not only the most probable point of the boundary touched at the exit but also the most probable exit trajectory (or tube of trajectories), as in the case of a domain attracted by a single stable state?

The difficult point is the following: even if there exists a unique function $\overline{\phi}^{x,\partial D}$ starting at x, exiting from D at time t and minimizing the functional $I_{[0,t]}(\phi)$, if this function visits a stable state in D, the process, with large probability, will spend a random exponential time in that stable state. This means that if D contains several stable states we have no hope to fix an exit trajectory as in (4.7) which is followed with large probability by the process.

In [OS] we solve this problem by showing that the construction of the chains $X_t^{(k)}$ provides a useful tool to study the behavior of the process near each stable state. More precisely we can define a tube of probable exit trajectories without fixing the times spent near each stable state.

Let us suppose for simplicity that the absolute minimum of the function H is not contained in D. (If this is not the case we can consider a new chain with this property by changing the function H outside the set D without affecting the property of the process up to the first exit time from D.)

Let N be the first integer such that $X_t^{(N)}$ has stable states only outside D:

$$N = \min\{k;\, S^{(k+1)} \subset D^c\}.$$

Then the process $X_t^{(N)}$ with large probability exits from D by following a trajectory $\overline{\phi}^{(N)}$ such that $I^{(N)}(\overline{\phi}^{(N)}) = 0$ in a mean time of order 1. This result is completely trivial for the chain $X_t^{(N)}$, but, by using our main theorem, it gives immediately the results (4.1) and (4.2).

By this result we have thus obtained a first approximation of the probable tube of trajectories exiting from D: all the trajectories of the process X_t corresponding at level N to the trajectory $\overline{\phi}^{(N)}$. By using the explicit construction of the renormalized chains at each step of the procedure, it is not difficult to prove that each transition of the chain $X^{(N)}$ corresponds with large probability to well defined sequences of transitions for the chain $X^{(N-1)}$. By iterating this idea in [OS] a complete characterization of the exit from a domain containing several stable states is obtained.

References

[FW] Freidlin, M. I. and Wentzell, A. D. (1984). *Random Perturbations of Dynamical Systems.* Springer-Verlag, Berlin.

[H] Has'minskii, R. Z. (1980). *Stochastic Stability of Differential Equations.* Sijthoff-Noordhoff.

[KO1] Kotecky, R. and Olivieri, E. (1992). Droplet dynamics for asymmetric Ising model (to appear).

[KO2] Kotecky, R. and Olivieri, E. Shapes of growing droplets — a model of escape from a metastable phase (to appear).

[MOS1] Martinelli, F., Olivieri, E., and Scoppola, E. (1991). On the Swendsen and Wang dynamics II: Critical droplets and homogeneous nucleation at low temperature for the 2 dimensional Ising model. *Journal of Statistical Physics* **62**, 135.

[MOS2] Martinelli, F., Olivieri, E., and Scoppola, E. (1990). Metastability and exponential approach to equilibrium for low temperature stochastic Ising models. *Journal of Statistical Physics* **61**, 1105.

[MOS3] Martinelli, F., Olivieri, E., and Scoppola E. (1991). On the Swendsen and Wang dynamics I: Exponential convergence to equilibrium. *Journal of Statistical Physics* **62**, 117.

[NSch1] Neves, E. J. and Schonmann R. H. (1991). Behaviour of droplets for a class of Glauber dynamics at very low temperatures. *Communications in Mathematical Physics* **137**, 209.

[NSch2] Neves, E. J. and Schonmann, R. H. (1992). Critical droplets and metastability for a class of Glauber dynamics at very low temperatures. *Probability Theory and Related Fields* **91**, 331.

[OS] Olivieri, E. and Scoppola, E. (to appear).

[S1] Scoppola, E.. Renormalization group for Markov chains and application to metastability. *Journal of Statistical Physics* (to appear).

[S2] Scoppola, E.. Metastability and nucleation for 2-dimensional Ising systems. *Physica A* (to appear).