MARKOV CHAINS WITH EXPONENTIALLY SMALL TRANSITION PROBABILITIES: FIRST EXIT PROBLEM FROM A GENERAL DOMAIN

I. THE REVERSIBLE CASE.

E.Olivieri(1), E.Scoppola(2)

(1) Dipartimento di Matematica - II Università di Roma - Tor Vergata
Via della Ricerca Scientifica - 00173 ROMA - Italy
E-mail: OLIVIERI@MAT.UTOVRM.IT

(2) Dipartimento di Fisica - Università di Bari. V. Amendola 173 - 70126 BARI - Italy
E-mail: SCOPPOLA@ROMA1.INFN.IT

Dedicated to the memory of Claude Kipnis

Abstract

We consider general ergodic aperiodic Markov chains with finite state space whose transition probabilities between pairs of different communicating states are exponentially small in a large parameter β.

We extend previous results by Freidlin and Wentzell ([FW]) on the first exit problem from a general domain Q.

In the present paper we analyze the case of *reversible* Markov chains. The general case will be studied in a forthcoming paper.

We prove, in a purely probabilistic way and without using F-W graphical technique, some results on the first exit problem from a general domain Q containing many attractors. In particular we analyze the properties of special domains called *cycles* and, by using the new concept of *temporal entropy*, we obtain new results leading to a complete description of the typical tube of trajectories during the first excursion outside Q.

Keywords : Markov chains, first exit problem, large deviations, reversibility.

Section 1. Introduction.

In this paper we consider ergodic, aperiodic Markov chains, with finite state space S and with transition probabilities $P(x,y)$ satisfying the following:

Property $\mathcal{P}$

If x and y are communicating states, i.e. $x \neq y$ and $P(x,y) > 0$, then

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

where $\Delta(.,.)$ is a non-negative function on the set of pairs of communicating states and $\gamma \to 0$ as $\beta \to \infty$.

Freidlin and Wentzell introduced this kind of Markov chains as auxiliary stuctures in their study of the asymptotic properties of diffusion processes describing small random perturbation of dynamical systems.

Another very interesting application, which actually is our main motivation, comes from non-equilibrium statistical mechanics: stochastic dynamics for interacting particle systems at very low temperature, like Glauber dynamics for Ising-like models, in a finite volume, satisfy property $\mathcal{P}$ (in this case β is the inverse temperature). (See e.g. [NSch1], [NSch2], [Sch], [KO1], [KO2], [MOS]).

We will mainly study the problem of the first exit from a domain $Q \subset S$ containing many attracting equilibrium states for the dynamics at $\beta = \infty$. Many results on this subject are already known: in particular Freidlin and Wentzel [FW] proved estimates for the average exit time and the typical point, on the boundary of Q, reached during the first excursion outside Q. They also describe the tube of typical trajectories exiting from Q, when this set contains a unique attracting state.

The study of the typical exiting trajectories is of fundamental importance and, in a certain sense, it is the central problem in the description of nucleation phenomenon, in the framework of general stochastic Ising models (see [Sch], [KO1], [KO2]). In that case we are interested in the analysis of the typical sequence of growing droplets and in particular in their shapes. Indeed the growth of the so called critical nucleus can be seen as a particular case of the first exit from a non-completely attracted domain.

It turns out, by looking at several particular models, that a crucial ingredient in the description of the growth is given by the *resistance times* inside some subsets of Q. These can be considered as a sort of *temporal entropy* related to fluctuations, taking place during suitable random times which grow exponentially fast in β. This temporal entropy turns out to be necessary to give rise to an efficient escape mechanism. Neglecting these random fluctuations, during the escape, would lead to a mechanism extremely depressed in probability.

In other words the last escape from Q, occurs in very different way in the two cases of one or several attracting point in Q. In the completely attracted case, the typical trajectories, during the first excursion outside Q, spend a finite time, independent of β, to bring the process out of Q without any "hesitation". In the general case the last escape takes place by visiting a suitable sequence of more or less stable attractors $z_1, ..., z_n$ and spending some suitable random times inside certain domain $A_1, ..., A_n$ which can be considered as a sort of generalized basins of attraction of $z_1, ..., z_n$ respectively.

We can say that the formulation itself of the problem of the characterization of the tube of typical exiting trajectories, in the general case, requires new concepts with respect to what has been done in the completely attracted case.

On one hand our work can be considerd as a completion and a generalization of the results contained in chapter 6 of the Freidlin and Wentzell's book [FW]; on the other hand we formulate a general set-up useful to treat, in a unified way, a large class of stochstic dynamics.

Our results are general and we are able to reduce the solution of the above mentioned typical large deviation problem, connected to the escape from a general domain, to the solution of a well defined sequence of variational problems. These variational problems constitute the model-dependent work to be done. In other words we state the results concerning the general behaviour of the class of Markov chains satisfying property $\mathcal{P}$, by specifying their common features and by reducing the model-dependent work to the solution of some well specified problems whose formulation can be given in general.

In the present note we will concentrate ourselves on the *reversible* case (see hypotheses 1 and 2 in section 3) where the unique invariant measure μ of the chain has the Gibbsian form: $\mu = \exp(-\beta H)/Z$ with a given energy function H on S.

For the general case we just give here the formulation of the problem, the complete treatement being the object of a forthcoming paper.

The discussion of the general case will require some generalization of the graphical technique introduced by Freidlin and Wentzel (see [FW] pg.177) and, more important, the use of the approach introduced by one of the authors in [S1]. This approach is based on the introduction of *renormalized chains*, obtained by a time rescaling related to the degree of stability of different attracting equilibrium states.

The reversible case is much easier. The crucial point of our approach to that case is to base our discussion on the analysis of the "energy landscape".

We will provide new probabilistic proofs of results obtained by Freidlin and Wentzell with their graphical technique. Moreover we will prove new results on the characterization of the tube of the typical exiting trajectories.

The paper is organized as follows: in section 2 we state the problem for the general case and we sketch the strategy for the solution of the general problem, in section 3 we consider the reversible case and we give the definition and the main properties of the cycles; finally in section 4 we define the typical exiting tube of trajectories of the first escape in the reversible case.

Section 2. The exit problem and the renormalization procedure.

Let X_t be a Markov chain satisfying property $\mathcal{P}$ above; given any set of states $Q \subset S$, we will denote by τ_Q the first hitting time to Q:

$$\tau_Q \equiv \min\{t > 0;\, X_t \in Q\}$$

We define the (outer) boundary ∂Q of Q as the set:

$$\partial Q = \{x \notin Q : \exists\, x' \in Q : P(x', x) > 0\}$$

A first description of the exit of the chain X_t from the set Q can be given by means of the following two quantities: the expectation of the first exit time from Q:

$$E_x \tau_{\partial Q} \tag{2.1}$$

and the spatial distribution of the first exit:

$$P_x(X_{\tau_{\partial Q}} = y) \tag{2.2}$$

with $x \in Q$, $y \in \partial Q$ (we denote by P_x the probability distribution on the process starting from x at $t = 0$; E_x denotes the corresponding expectation).

Estimates of these quantities, from above and from below, are given by Freidlin and Wentzell in the framework of their study of diffusion processes describing small random perturbations of dynamical systems [FW]. In fact they proved that these continuous processes can be approximated by discrete Markov chains satisfying property $\mathcal{P}$. Moreover they show that the quantities (2.1) and (2.2) can be expressed in terms of sums of products of transition probabilities of the chain, and these products can be defined by means of graphs of arrows. Then the estimates of these quantities can be reduced to a problem of minimization of a suitable cost function associated to each graph (see [FW] ch 6, sect. 3, pg. 176).

The theory of large deviations, developed in [FW] for the continuous case, can be very easily adapted to our discrete Markov chain. To each path, i.e. to each function $\phi : \mathbf{N} \to S, \quad \phi = \{\phi_t\}_{t\in\mathbf{N}}$, we can associate a functional

$$I_{[0,t]}(\phi) \equiv \sum_{i=0}^{t-1} \Delta(\phi_i, \phi_{i+1}) \tag{2.3}$$

where the fonction $\Delta(x, y)$ is defined in (1.1) and we set $\Delta(x, x) = 0$ for each $x \in S$ and $\Delta(x, y) = \infty$ if $P(x, y) = 0$. This functional can be interpreted as the cost function of each path ϕ, and the following large deviation estimates can be easily proved (see [S1]):

Lemma 2.1

Let ϕ be a fixed function starting at x at time 0, then

i)

$$P_x(X_s = \phi_s \quad \forall s \in [0, t]) \le e^{-I_{[0,t]}(\phi)\beta+\gamma t\beta}$$

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

$$P_x(X_s = \phi_s \quad \forall s \in [0, t]) \ge e^{-I_{[0,t]}(\phi)\beta-\gamma t\beta}$$

iii) for any constant $I_0 > 0$, for any sufficiently small $\alpha > 0$, for any $t < e^{\alpha\beta}$ and for any sufficiently large β

$$\sup_x P_x(I_{[0,t]}(X_s) \ge I_0) \le e^{-I_0\beta+\epsilon\beta}$$

where $\epsilon \to 0$ as $\beta \to \infty$.

Following Freidlin and Wentzell, an equivalence relation in the state space S can be defined by means of the functional $I_{[0,t]}(\phi)$: for each couple of states x, y we define

$$V(x, y) \equiv \inf_{t,\phi;\,\phi_0=x,\,\phi_t=y} I_{[0,t]}(\phi) \tag{2.4}$$

and

$$x \sim y \quad \text{iff} \quad V(x,y) = V(y,x) = 0 \tag{2.5}$$

We say that x is a stable state if and only if

$$\text{for any} \quad y \not\sim x \quad V(x,y) > 0 \tag{2.6}$$

i.e. if each path leaving from x has a positive cost. We will denote by M the set of stable states.

It is immediate to see that if the set M contains a state x then it contains the whole equivalence class of x, namely $M \supset (x)_\sim$, where $(x)_\sim \equiv \{y \in S;\, y \sim x\}$

An immediate consequence of lemma 2.1 is the following:

Lemma 2.2

There exist constants $T_0 \in [0, |S|]$ and β_0 such that for any $\beta > \beta_0$:

i) for any $t > T_0$:

$$\sup_{x \in S} P_x(\tau_M > t) \le a^{[\frac{t}{T_0}]}$$

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

ii) for any $\eta > 0$ and for any $t \ge e^{\eta\beta}$ and β sufficiently large we have:

$$\sup_{x \in S} P_x(\tau_M > t) \le \exp\{-e^{\frac{\eta\beta}{2}}\}$$

Let us now come back to the problem of the exit of our chain X_t from a domain Q and let us suppose that this set contains a unique stable state x_0 completely attracting this set, i.e. for each $y \in Q$ there exists a path $y_0 = y, y_1, ..., y_n = x_0$ such that $\Delta(y_1, y_{i+1}) = 0 \quad \forall i < n$, while $\Delta(y,z) > 0$ for each $y \in Q, z \in \partial Q$. Then in this case Freidlin and Wentzell can describe in complete detail the exit from Q.

First of all in this case the quantities (2.1) and (2.2) can be easily estimated as follows (see Th. 2.1 and Th.4.1 ch 4 [FW]): for all $x \in Q$

$$\lim_{\beta \to \infty} \frac{1}{\beta} \ln E_x \tau_{\partial Q} = \min_{y \in \partial Q} V(x_0, y)$$

and if there exists a unique state $y_0 \in \partial Q$ such that $V(x_0, y_0) = \min_{y \in \partial Q} V(x_0, y)$ then

$$\lim_{\beta\to\infty} P_x(X_{\tau_{\partial Q}} = y_0) = 1$$

We notice, at this point, that in the case of reversible chains, i.e. when the function $\Delta(x,y)$ defined in (1.1) can be expressed in terms of a unique energy function H on S, then the quantities (2.1) and (2.2) can be estimated more easily in terms of the function H, since the quantity $\min_{y\in\partial Q} V(x_0,y)$ can be easily expressed in terms of H.

Moreover in this case of a domain Q containing a unique stable state, the last escape can be described quite precisely and completely.

We state now the result in our discrete case of Markov chains (see [FW] Th.2.3, ch 4 for the continuous version of this result).

We want to notice here that in the continuous case of diffusion processes discussed in [FW], the dynamics corresponding to zero random noise, was given by a dynamical system, that is the unperturbed system was completely deterministic and for each starting point there was a unique deterministic path emerging from it. The tube of typical exiting trajectories was given, in that case, as a neighborhood, in the uniform topology, of such a deterministic path.

Here, in the discrete case of Markov chains, the situation is different and even for $\beta = \infty$ the system can still be random. This means that there is not a unique deterministic path but several possible paths emerging from the same starting point. Moreover, we do not have to consider a neighborhood since the space is discrete. So the typical exiting tube, in this case, is a finite set of individual paths.

Proposition 2.1

Let Q be a set of states containing a unique stable state x_0 and for each given α define

$$\Phi \equiv \{\{\phi_s\}_{s\in\mathbf{N}};\ \phi_0 = x_0,\ \phi_{T_\phi} \in \partial Q,\ \phi_s \in Q,\ \forall s < T_\phi, \quad \textit{with} \quad T_\phi < e^{\alpha\beta}\} \tag{2.7}$$

and

$$\bar\Phi \equiv \{\phi_s \in \Phi;\ I_{[0,T_\phi]}(\phi) = \min_{y\in\partial Q} V(x_0,y)\} \tag{2.8}$$

then there exists α_0 (see lemma 2.1) such that if $\alpha < \alpha_0$ we have:

$$\lim_{\beta\to\infty} P_x(X_{\theta_{x_0}+t} = \phi_t \quad \forall t = 0,...,\tau_{\partial Q} - \theta_{x_0} \quad \textit{for some} \quad \phi \in \bar\Phi) = 1 \tag{2.9}$$

where $\theta_{x_0} \equiv \max\{t < \tau_{\partial Q};\ X_t = x_0\}$.

A proof of this proposition can be easily obtained by applying lemmas 2.1 and 2.2. Indeed by lemma 2.2 we are able to exclude that $\tau_{\partial Q} - \theta_{x_0}$ is exponentially too large in β, and then we can apply the large deviation estimates. More precisely the probability of the complementary event can be estimated from above as follows:

$$P_x(\{X_{\theta_{x_0}+t}\}_{t=0}^{\tau_{\partial Q}-\theta_{x_0}} \not\in \bar\Phi) = \sum_{s=0}^{\infty} P_x([\{X_{s+t}\}_{t=0}^{\tau_{\partial Q}-s} \not\in \bar\Phi] \bigcap [\theta_{x_0} = s]) =$$

$$= \sum_{s=0}^{\infty} P_x([\{X_{s+t}\}_{t=0}^{\tau_{\partial Q}-s} \not\in \bar\Phi] \bigcap [X_{s'} \in Q \quad \forall s' < s,\ X_s = x_0,\ X_{s+1} \neq x_0] \bigcap$$

$$\bigcap [X_{\tau^{>s}_{x_0 \cup \partial Q}} \in \partial Q]) \tag{2.10}$$

where

$$\tau^{>s}_{x_0 \cup \partial Q} \equiv \min\{t > s;\ X_t \in \{x_0 \cup \partial Q\}\}$$

The r.h.s. of (2.10) can be estimated, by using the Markov property, as follows:

$$\leq \sum_{s=0}^{\infty} P_x(\tau_{\partial Q} > s) P_{x_0}([\{X_t\}_{t=0}^{\tau_{\partial Q}} \not\in \bar\Phi] \bigcap [X_1 \neq x_0,\ X_{\tau_{x_0 \cup \partial Q}} \in \partial Q]) \leq$$

$$\leq E_x \tau_{\partial Q} \,.\, \{P_{x_0}([\{X_t\}_{t=0}^{\tau_{\partial Q}} \not\in \Phi] \bigcap [X_1 \neq x_0,\ X_{\tau_{x_0 \cup \partial Q}} \in \partial Q]) + P_{x_0}(\{X_t\}_{t=0}^{\tau_{\partial Q}} \in \Phi \backslash \bar\Phi)\} \leq$$

$$\leq e^{\min_{y \in \partial Q} V(x_0,y)\beta + \delta\beta} \,.\, \{\exp\{-e^{\frac{\alpha\beta}{2}}\} + P(I_{[0,e^{\alpha\beta}]}(X_s(x_0)) \geq \min_{y \in \partial Q} V(x_0,y) + d)\}$$

for some positive constant d. The theorem follows by applying lemma 2.2 and by using that δ goes to 0 as $\beta \to \infty$ while d is fixed.

We remark that in this theorem the hypothesis of the uniqueness of the stable state in Q is crucial; in fact the large deviation estimate can be applied, here as in the continuous case, only on intervals of time which do not grow too fast in β (they have to be bounded by $e^{\alpha\beta}$ with α sufficiently small). However, if the set Q contains several stable states, then the functions in $\bar{\Phi}$ can (and we will see that they will do) visit other stable states before leaving Q where the process is likely to spend exponentially long times. This means that in this case, due to the above mentioned resistance times, the time $\tau_{\partial Q} - \theta_{x_0}$ is exponentially large with large probability. Thus an extension of proposition 2.1 to a general domain Q would require a control on large deviation estimates over exponentially long intervals of time. This is the crucial point to be solved. Really new ideas and techniques are necessary; new concepts will be needed to define the tube.

To this end we recall here some recent results on Markov chains satisfying property $\mathcal{P}$, representing a completly different approach to the study of long time behavior of Markov chains, based on a renormalization procedure [S1].

The main idea of this *renormalization* procedure can be summarized as follows.

The behaviour of the chain X_t involves a sequence of different time scales T_1, T_2, T_3, ..., exponentially large in β, related to the stability of the different states. For each time scale T_i it is possible to define a *renormalized chain* $X_t^{(i)}$, corresponding to the original chain X_t on this time scale; this is a coarse grained version of the chain X_t in the sense that it gives a less detailed description of the process, but the loss of information concerns only events which occur in a typical time less or equal to T_i.

In order to explain more clearly how this renormalization works, we briefly recall the construction of the first chain $X_t^{(1)}$. (see [S1] for more details)

A first simple classification of states is that given in terms of stable and unstable states (see (2.6)).

If we denote by M the set of stable states, then, as shown in lemma 2.2, the process spends, with large probability, almost all the time in M. This result suggests 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 behaviour of the process on small times is neglected.

In fact we can consider the less stable states in M and we can define a time scale T_1 corresponding to this smallest stability:

$$T_1 \equiv e^{V_1\beta+\delta\beta} \tag{2.11}$$

where

$$V_1 \equiv \min_{x\in M,\, y\in S,\, x\not\sim y} V(x,y) \tag{2.12}$$

and δ is a small constant going to zero as β goes to infinity.

We can then construct a new Markov chain $X_t^{(1)}$ with state space M, corresponding to the original process with a rescaling time T_1, by defining a sequence of stopping times $\zeta_1, \zeta_2, ..., \zeta_n, ...$ such that $\zeta_{n+1} - \zeta_n$ is of order T_1 with large probability and X_{ζ_n} belongs to M.

More precisely we define the sequence of stopping times:

$$\sigma_1 \equiv \min\{t > 0;\; X_t \not\sim X_0\}$$

$$\tau_1 \equiv \min\{t \geq \sigma_1;\; X_t \in M\}$$

$$\zeta_1 = \begin{cases} t_1 & \text{if } \sigma_1 > t_1 \\ \tau_1 & \text{if } \sigma_1 \leq t_1 \end{cases} \tag{2.13}$$

and for each $n > 1$:

$$\sigma_n \equiv \min\{t > \zeta_{n-1};\; X_t \not\sim X_{\zeta_{n-1}}\}$$

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

$$\zeta_n = \begin{cases} \zeta_{n-1} + t_1 & \text{if } \sigma_n - \zeta_{n-1} > t_1 \\ \tau_n & \text{if } \sigma_n - \zeta_{n-1} \leq t_1 \end{cases} \tag{2.14}$$

It is simple to prove that the sequence $X_n^{(1)} = X_{\zeta_n}$ is a homogeneous Markov chain. For any couple of states $x, y \in M$ we denote by $P^{(1)}(x,y)$ the transition probability of the chain $X_n^{(1)}$, it is possible to prove that these transition probabilities satisfy the same assumption (property $\mathcal{P}$) verified by the original chain X_t, provided we identify states which are equivalent with respect to the relation (2.5). More precisely for any $x, y \in M, \quad x \not\sim y$:

$$\exp\{-\Delta^{(1)}(x,y)\beta - \gamma'\beta\} \leq P^{(1)}(x,y) \equiv P(X_{\zeta_n} = y |\, X_{\zeta_{n-1}} = x) \leq \exp\{-\Delta^{(1)}(x,y)\beta + \gamma'\beta\} \tag{2.15}$$

The quantities $\Delta^{(1)}(x,y)$ are defined by:

$$\Delta^{(1)}(x,y) = \inf_{\substack{t,\phi;\; \phi_0 = x,\; \phi_t = y, \\ \phi_s \notin M \setminus [(x)_\sim \cup (y)_\sim]}} I_{[0,t]}(\phi) - V_1 \tag{2.16}$$

and $\gamma' \to 0$ as $\beta \to \infty$.

Thus we have a new chain $X_t^{(1)}$ on the state space $S^{(1)} = M$ (modulo the equivalence relation (2.5))$\equiv M/_{\sim})$, to which we can apply again the same analysis, by defining new stable states, a time scale T_2, a corresponding chain $X_t^{(2)}$ and so on.

The iteration scheme is the following one (see [S1]): for any $k \geq 1$ we define the following quantities:

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

$$I_{[o,t]}^{(k)}(\phi) = \sum_{i=0}^{t-1} \Delta^{(k)}(\phi_i, \phi_{i+1}) \tag{2.17}$$

$$V^{(k)}(x,y) \equiv \min_{t,\phi;\, \phi_0 = x,\, \phi_t = y} I_{[o,t]}^{(k)}(\phi) \quad \forall x, y \in S^{(k)} \tag{2.18}$$

$$x \sim^{(k)} y \quad \text{if and only if} \quad V^{(k)}(x,y) = V^{(k)}(y,x) = 0 \tag{2.19}$$

$$M^{(k)} = \{x \in S^{(k)}; \forall y \in S^{(k)},\, y \not\sim^{(k)} x \quad V^{(k)}(x,y) > 0\} \tag{2.20}$$

$$V_{k+1} = \min_{x \in M^{(k)}, y \in S^{(k)}\ x \not\sim^{(k)} y} V^{(k)}(x,y) \tag{2.21}$$

$$t_{k+1} = e^{V_{k+1}\beta + \delta\beta} \tag{2.22}$$

$$T_1 = t_1$$

$$T_{k+1} = t_1 t_2 ......t_k t_{k+1} \tag{2.23}$$

$$S^{(k+1)} = M^{(k)}/_{\sim^{(k)}} \tag{2.24}$$

$$\Delta^{(k+1)}(x,y) = \min_{\substack{t,\phi;\phi_0 = x, \phi_t = y, \\ \phi_s \not\in M^{(k)} \setminus [(x)_{\sim^{(k)}} \cup (y)_{\sim^{(k)}}]\ \forall s \in [0,t]}} I_{[0,t]}^{(k)}(\phi) - V_{k+1} \quad \forall x, y \in S^{(k+1)} \tag{2.25}$$

The main results proved in [S1] can be summarized as follows: we can estimate from above and from below the following quantities for the original chain X_t:

i) the invariant measure of the chain: $\mu(.)$

ii) the expected value of the first hitting time of the process starting at x to a set $B \subset S$: $E\tau_B(x) \quad B \in S,\, x \in S$

iii) the spatial distribution of the first entry: $P(X_{\tau_B}(x) = y) \quad B \in S,\, x \in S,\, y \in B$

in terms of the same quantities for the rescaled chains $X^{(i)}$.

Since $|S^{(i)}| \leq |S^{(i-1)}|$ (actually one can prove that $|S^{(i+1)}| < |S^{(i-1)}|$), the above results provide a useful tool for the evaluation of these quantities when $|S|$ is large, in fact one can consider a time rescaling T_n so large that the corresponding state space $S^{(n)}$ is so small that explicit computations are easy at this level.

In [S2] the easier case of reversible Markov chains is analyzed. The construction of the renormalized chains is more explicit providing, in that case, a sequence of "smoothed" hamiltonian functions $H^{(k)}$.

It is possible to establish a connection between these results and the graphical approach by Freidlin and Wentzell. The minimization problem involved in the graph approach to estimate (2.1) and (2.2), turns out to be broken by this renormalization procedure in a sequence of easier minimization problems. (See [S3]).

We propose here a new application of this renormalization procedure. We claim that we can use the renormalized chains $X_t^{(k)}$ to control the large deviation phenomena, for the Markov chain X_t, taking place during exponentially long times T_k.

In fact the control on the chain $X_t^{(k)}$ over a small time $t < e^{\alpha\beta}$ given by the large deviation estimates (lemma 2.1) applied to this chain, implies a control on the original chain X_t in an interval of time $t.T_k$ exponentially long in β. In other words we are able to characterize the typical behaviour of the chain X_t in each interval of time having a length of order T_k corresponding to a single jump $X_s^{(k)} \to X_{s+1}^{(k)}$ at level (k).

This is exactly the new ingredient necessary to the construction of the exiting tube, as discussed at the end of proposition 2.1.

A complete development of this idea will be given in [OS]. Here we only outline the strategy for applying the renormalization approach to the study of the exit problem in the general case of a set Q containing several stable states.

In order to describe the first exit from the domain Q of a Markov chain X_t we can consider a new chain X_t^Q on the state space $Q \cup \partial Q$ which is equivalent to X_t up to its first exit from Q, but with almost absorbing states in ∂Q. More precisely we define the following transition probabilities:

$$P^Q(x,y) = P(x,y) \quad \text{for any } x \in Q \tag{2.26}$$

$$P^Q(x,y) = e^{-\beta\Delta(\partial Q)} P(x,y) \quad \text{for } x \in \partial Q,\, y \neq x \tag{2.27}$$

where $\Delta(\partial Q) >> \sum_{y,z\in Q} \Delta(y,z)$.

To this chain we apply the previously discussed analysis by evaluating the sequence of renormalized chains $X_t^{(1)}, ..., X_t^{(i)}, ...$ and the corresponding sequence of state spaces $S^{(1)}, ..., S^{(i)}, ...$ We warn the reader of an abuse of notation: we will omit from now on the superscript Q.

Since $\Delta(\partial Q) >> \sum_{y,z\in Q} \Delta(y,z)$ it is immediate to show that there exists a step N of the iteration such that in Q there are only unstable states. More precisely let $N = N(Q) = \inf\{n;\, S^{(n+1)} \subset \partial Q\}$ then $X^{(N)}$ has stable states only on the boundary of Q and all the states in ∂Q are stable. Thus the exit from Q for the chain $X_t^{(N)}$ is a trivial problem since it is a downhill exiting. This means that for each $x \in S^{(N)} \cap Q$ there exists at least a time k and a sequence $x_0^{(N)}, ...x_k^{(N)}$ of states in $S^{(N)}$ such that $x_0^{(N)} = x, \quad x_i^{(N)} \in Q \quad \forall i < k, \quad x_k^{(N)} \in \partial Q$ and $\Delta^{(N)}(x_i^{(N)}, x_{i+1}^{(N)}) = 0$ for each $i < k$. The sequence $x_0^{(N)}, ...x_k^{(N)}$ is not necessarily unique and it represents a probable exit path from Q for the chain $X^{(N)}$. Let us first suppose that this exiting path is unique. We can then give a first approximation of the tube of trajectories that the stochastic process follows with probability almost one when it exits from the domain Q. Indeed let

$$\Phi(S) = \{\{\phi_i\}_{i\in\mathbf{N}},\, \phi \in S\} \tag{2.28}$$

be the space of all the trajectories of the Markov chain X_t .

On each path $\phi \in \Phi(S)$ we evaluate the sequence of stopping times ζ_n and a path $\phi^{(1)} \in \Phi(S^{(1)})$ given by $\phi_i^{(1)} = \phi_{\zeta_i}$.

Using the same construction we can thus define a sequence of trajectories $\{\phi_i^{(n)}\}_{i\in\mathbf{N}}$ in the spaces $\Phi(S^{(n)})$ with n=2,3,...

For any fixed sequence of states in $S^{(n)} : \psi = \psi_1, \psi_2, ..., \psi_k$ we can thus define a tube of trajectories in $\Phi(S)$ as

$$\Theta(n,\psi) = \{\phi \in \Phi(S);\, \phi_i^{(n)} = \psi_i \quad \forall i \le k\} \tag{2.29}$$

In the case of uniqueness of the exiting path for the chain $X_t^{(N)}$ we can conclude that

$$P_x(X_t) \in \Theta(N, \{x_0^{(N)}, ...x_k^{(N)}\})) \tag{2.30}$$

tends to one as $\beta \to \infty$.

Indeed

$$\{X_t \in \Theta(N, \{x_0^{(N)}, ...x_k^{(N)}\})\} = \{X_i^{(N)} = x_i^{(N)} \quad \forall i = 0, ..., k\} \tag{2.31}$$

and the probability of the second event tends to one as $\beta \to \infty$ if $\{x_0^{(N)}, ...x_k^{(N)}\}$ is the unique deterministic path exiting from $Q \cap S^{(N)}$ for the chain $X_t^{(N)}$.

In the case in which there are several exiting paths for the process $X_t^{(N)}$ we have to consider the union of this tubes.

As a consequence of this fact we can easily find the results by Freidlin and Wentzell on the mean exit time from Q and on the best exit point of the process X_t. Indeed by the results proved in [S1], we can evaluate the quantities $E\sigma_Q$ and $P(X_{\sigma_Q} = y)$ for the renormalized chain $X_t^{(N)}$.

Coming back to our problem of the determination of the minimal tube of trajectories that the stochastic process follows with large probability, we want now to define a tube of exiting trajectories smaller than $\Theta(N, \{x_0^{(N)}, ...x_k^{(N)}\})$. A first remark is the following one: if we look at one transition $x_i^{(N)} \to x_{i+1}^{(N)}$ of the chain $X^{(N)}$ it easy to show that with large probability it corresponds to a sequence of transitions of the chain $X^{(N-1)}$ minimizing the functional appearing in the definition of $\Delta^{(N)}(x_i^{(N)}, x_{i+1}^{(N)})$. Indeed for each $x, y \in S^{(n)}$ let $\omega(x,y)_i$, $i \le t$ be the sequence (that here we suppose unique) of states in $S^{(n-1)}$ minimizing the quantity $\Delta^{(n)}(x,y)$. It is easy to prove that

$$P_x(X_{s+\sigma}^{(n-1)} = \omega(x,y)_s,\, \forall s \le t | X_1^{(n)} = y) \to 1 \qquad \text{as } \beta \to \infty \tag{2.32}$$

where σ is the exit time from the state x of the process $X^{(n-1)}$. One can iterate this argument and look at the most probable sequence of transitions of the chain $X^{(N-2)}$ corresponding to each transition of $X^{(N-1)}$ and so on.

In order to complete the analysis we have to look for a characterization of the behaviour of the original chain X_t when a given renormalized chain $X^{(n)}$ remains in the same state for a time σ; namely we want to analyze in more detail the transition $x \to x$ for the chain $X^{(n)}$.

In the second paper of this series [OS] we will answer the above question; on this ground we will perform a complete analysis of the first exit tube.

In the rest of this paper we will consider the exit problem for reversible chains. We will not follow the strategy proposed in this section but we will perform explicit computations which are easy in the reversible case. We will relate these computations to the general strategy in the forcoming general paper in which this reversible case will represent a useful example to clarify general statements and definitions.

Section 3. The cycles and their properties

We consider an ergodic aperiodic Markov chain with a finite space state S and with transition probabilities $P(x,y)$ satisfying property $\mathcal{P}$ and the following assumption :

Hypothesis H1

There exists a function $H : S \to \mathbf{R}^+$ *such that ,* $\forall\, x, y \in S$ *with* $P(x,y) > 0$*,*

$$\Delta(x,y) \;=\; [H(y) - H(x)]_+ \tag{3.1}$$

where $(a)_+$ *is the positive part* $(\equiv a \vee 0)$ *of the real number* a*.*

The above choice corresponds to a *quasi-reversible Metropolis Markov chain* (here "Metropolis " is used in a slightly more general meaning than the common Metropolis algorithm).

The most general quasi-reversible Markov chain can be defined in the following way: we suppose given a function $H : \Gamma \cup S \to \mathbf{R}^+$, Γ being the subset of $S \times S$ of *communicating* pairs of states, namely pairs $\{x,y\}$ where $P(x,y)$ is strictly positive, with

$$H(\{x,y\}) \geq H(x) \vee H(y)$$

and

$$\Delta(x,y) = H(\{x,y\}) - H(x) \tag{3.2}$$

It is not difficult to convince oneself that the paricular Metropolis choice, corresponding to $H(\{x,y\}) = H(x) \vee H(y)$, does not represent a real loss of generality : one can always enlarge the space S to include a new point for each "bond" $\{x,y\} \in \Gamma$ between communicating states assigning to it the energy $H(\{x,y\})$.

Moreover the Metropolis form is used in many applications, like the usual stochastic Ising models.

We will always assume that $P(x,y)>0$ implies $P(y,x)>0$ for any $x,y \in S$.

For most of our result we will assume not only the hypothesis H1 of quasi-reversibility, but the stronger hypothesis of *reversibility*:

Hypothesis H2 (Detailed balance condition)

There exists a strictly positive probability measure μ on S such that:

$$\forall\, x,x' \, \in S \;:\; \mu(x)P(x,x') \;=\; \mu(x')P(x',x) \tag{3.3}$$

¿From (3.3) one immediately deduce that μ is the unique invariant measure of the chain.

A *path* ω is a sequence $\omega \equiv x_1,\dots,x_N,\; N \in \mathbf{N}$, with $x_j,x_{j+1},\; j=1,\dots,N-1$, communicating states (i.e. $P(x_j,x_{j+1}) \;>\; 0$).

A set $Q \subset S$ is *connected* if $\forall\, x,x' \in Q$ there esists a path $\omega : x \to x'$ all contained in Q (we use the notation $\omega : x \to x'$ to denote a path ω joining x to x').

Given $Q \subset S$ we denote by $U = U(Q)$ the set of all minima of the energy on the boundary ∂Q of Q :

$$U(Q) = \{z \in \partial Q : \min_{x \in \partial Q} H(x) = H(z)\} \tag{3.4}$$

Given $Q \subset S$ we denote by $F = F(Q)$ the set of all minima of the energy on Q :

$$F(Q) = \{y \in Q : \min_{x \in Q} H(x) = H(y)\} \tag{3.5}$$

A connected set of equal energy states $\bar{P} \subset S$ is called *plateau.* It is easy to convince oneself that, in the framework of our asymptotic estimates, exponential in the parameter β, we can identify these plateaux with single points. In other words, states which are equivalent with respect to the relation (2.5) of the previous section, can be identified.

Moreover it follows from (3.3) the property: $P(x,y)>0$ *implies* $P(y,x)>0$.

It is immediate to verify that in the quasi reversible case a state x is stable, in the sense of definition (2.6), if either it is a local minimum of the function H or it belongs to a plateau equivalent to a local minimum.

Definition 3.1

A connected set A which satisfies :

$$\max_{x\in A} H(x) = \bar{H} < \min_{z\in\partial A} H(z) \equiv H(U(A))$$

is called cycle.

It is easy to see that, under the quasi-reversibility hypothesis for our Markov chain, the above definition is equivalent to the one given by Freidlin-Wentzell (see [FW] pag. 198).

In the following we will give some propositions (beside other definitions). Most of them are intended to clarify the structural properties of the cycles. For some of them the proof is immediate and we omit it. The first non-immediate (though in a sense elementary) statement is contained in Proposition 3.7 for which we provide a proof.

Proposition 3.1

Given a state $\bar{x} \in S$ and a real number c the set of all x's connected to $\bar{x}$ by paths with energy always below a given c either coincides with S or it is a cycle A with

$$H(U(A)) \geq c$$

Proposition 3.2

Given two cycles $A_1,\ A_2$, either
1) $A_1 \cap A_2 = \emptyset$
or
2) $A_1 \subset A_2$ or, viceversa, $A_2 \subset A_1$

Proof.

Let $A_1 \cap A_2 \neq \emptyset$.
It is immediate to see that one cannot have that, at the same time,

$$\exists x_1 \in \partial A_1 \cap A_2, \qquad \exists x_2 \in \partial A_2 \cap A_1,$$

otherwise one would have, at the same time:

$$H(x_1) < H(x_2)$$

and

$$H(x_2) < H(x_1)$$

Then, either $A_1 \subset A_2$ or, $A_2 \subset A_1$

Definition 3.2

A cycle A such that $\forall z \in U(A)$ one has that for all x's communicating with z for which:

$$H(x) < H(z)$$

belong to A is called stable or attractive cycle.

Definition 3.3

A cycle A for which there exists $y^* \in U(A)$ downhill connected to some point x in A^c (namely $\exists\ x \notin A$, communicating with y^*, with $H(x) < H(y^*) \equiv H(U(A))$), is called transient; points like y^* are called (minimal) saddles.

Definition 3.4

A transient cycle A such that $\exists\ \bar{x} \notin A$ with $H(\bar{x}) < H(F(A))$, there exists $y^* \in \mathcal{S}(A)$ and a path $\omega : y^* \to \bar{x}$ below y^* (namely $\forall x \in \omega : H(x) < H(y^*)$), is called metastable.

Remark

It can happen that there exist several points $z \in U(A)$ communicating with each other and with the same energy: as far as asymptotic estimates, exponential in β, are concerned we can consider them as a unique point.

Definition 3.5

For each couple of states $x, y \in S$ we define their minimal saddle $\mathcal{S}(x, y)$ as the set of states corresponding to the solution of the following minimax problem:

$$\min_{\omega: x \to y} \max_{z \in \omega} H(z) \ = \ H(\mathcal{S}(x,y))$$

where $\omega : x \to y$ is a generic path joining x to y. For any $z, z' \in \mathcal{S}(x,y)$ we have $H(z) = H(z')$ and we set $H(\mathcal{S}(x,y)) \equiv H(z) \quad \forall z \in \mathcal{S}(x,y)$.

¿From our assumptions on the chain it immediately follows that $\mathcal{S}(x,y) = \mathcal{S}(y,x)$ $\forall x, y \in S$.

Remark

We will call *trivial* the saddles corresponding to a trivial minimax solution, i.e. if $\mathcal{S}(x,y) \ni x$ or y. We want to note that saddles between stable states, not equivalent with respect to the relation $\sim$ introduced in (2.5), are not trivial, since, by the definition of stable states, the energy of the saddle must be strictly larger than the energy of both the stable states considered. The saddles between stable states will be called *natural saddles* .

Proposition 3.3

If A is a cycle then:

i) for each $x, y, z \in A$ and $w \notin A$:

$$H(\mathcal{S}(x,y)) < H(\mathcal{S}(z,w))$$

ii) for any $x \in A$ the set

$$\mathcal{S}(x,A) \ = \ \{z \in \mathcal{S}(x,w) \ \textit{for some} \ w \in \partial A \, ; \quad \min_{w \in \partial A} H(\mathcal{S}(x,w)) = H(z)\}$$

does not depend on x and we denote it by $\mathcal{S}(A)$. Moreover

$$\mathcal{S}(A) = U(A)$$

Proof

By the definition of cycle for any $x, y \in A$ there exists a path $\omega : x \to y$ contained in A such that for any $x_i \in \omega$ $H(x_i) < H(U(A))$. On the other hand any path going from $z \in A$ to $w \in \partial A$ maximizes H in ∂A.

Thus we have

$$\min_{w \in \partial A} H(\mathcal{S}(x,w)) = H(U(A))$$

Definition 3.6

Given a stable state i.e. a local minimum for the energy $x \in M$, we define the following basins for x:
i) $\hat{B}(x) =$ the wide basin of attraction of x the (necessarily nonempty) set given by :

$$\hat{B}(x) = \{z : \exists \text{ downhill path } \omega : z \to x\} \tag{3.6}$$

(a path $\omega \equiv x_1, \ldots, x_N$ is downhill iff $H(x_{j+i}) < H(x_j),\ j = 1, \ldots, N-1$)
ii) $\bar{B}(x) =$ the (not necessarily non-empty) basin of attraction of x given by:

$$\bar{B}(x) = \{z : \text{every downhill path starting from } z \text{ ends in } x\} \tag{3.7}$$

$\bar{B}(x)$ can be seen as the usual basin of attraction of x with respect to the $\beta = \infty$ dynamics.
iii) $B(x) =$ the strict basin or attraction of x given by :
if S is reduced to $\bar{B}(x)$:

$$B(x) = \bar{B}(x) = S$$

Otherwise:

$$B(x) = \{z \in \bar{B}(x) \ : \ H(z) < H(U(\bar{B}(x)))\} \tag{3.8}$$

Remark

1) Every local minimum for the energy is (a trivial) cycle.
2) Any cycle A contains at least a stable state belonging to $F(A)$
3) The above definitions of different basins of attraction can immediately be extended to the case in which, in place of a local minimum x, there is a generic cycle A. In this way one gets $\hat{B}(A), \bar{B}(A), B(A)$. Of course one can even consider the basins of attraction of more general sets (not necessarily cycles).
4) The case of a Markov chain for which the space of the states S reduces itself to the basin of attraction of the absolute minimum $\bar{x}$ for the energy H is, in a sense, almost trivial; we shall refer to this case as to the *one-well* case.

Proposition 3.4

Suppose S not one-well. Let G be a subset of S and suppose that there exist $x \in G, x' \notin G$ with the following properties:
i) G is connected
ii) there exists a path $\omega : x \to y^*$ for some $y^* \in U(G)$, with

$$H(z) < H(y^*)\ \forall z \in \omega, z \neq y$$

iii) there exists a path $\omega : y^* \to x'$ with

$$H(z) < H(y^*)\ \forall z \in \omega, z \neq y$$

(obviously it has to be $H(x') < H(y^*)$)
Then if $A = \{z\ :\ \exists\ \ \omega : z \to x;\ \forall y \in \omega, \quad H(y) < H(y^*)\} \ \equiv$ maximal connected set containing x with energy less than $H(y^*)$:
1) $A \subset G$
2) A is a cycle with $U(A) \ni y^*$
3) y^* is a minimax between x and x' namely $y^* \in \mathcal{S}(x, x')$.

The above proposition provides a constructive criterion to find the minimax between different points.

The notion of local minimum and of its basins of attraction will be useful to provide a decomposition of the cycles. More precisely, given a cycle A containing several local minima, we will consider some smaller cycles A_i contained in A, by looking at the sequence of internal saddles between minima in A. Let us start with a trivial corollary of proposition 3.2:

Proposition 3.5

If a cycle A contains an internal natural saddle y, namely a non trivial minimax between two stable states $z, z' \in A \cap M$ with energy $H(y) < H(U(A))$, then it contains all the cycles A_j having y in the sets $U(A_j)$'s.

Proposition 3.6

Let A be a cycle which does not reduce itself to a (proper or not) subset of a strict basin of attraction of some local minimum x. Let $y_1, \ldots, y_m$ be the (supposed non-empty) set of *its internal natural saddles with maximal energy; namely if*

$$C(A) = \{y \in A : \exists x, x' \in A \cap M : y \in \mathcal{S}(x, x')\} \tag{3.9}$$

is the set of internal natural saddles in A, then

$$\max_{z \in C(A)} H(z) = H(x) \textit{ if and only if } x \in \{y_1, \dots, y_m\} \tag{3.10}$$

so that, in particular,

$$H(y_i) \;=\; H(y_j) \;\; \forall\, i,\; j \;= 1, \dots, m$$

The number m of internal saddles with maximal energy of a cycle A is denoted by $N(A)$ (if A reduces itself to a subset of a strict basin of attraction of some local minimum x one has $N(A) = 0$).

Then the following is true :

i) the following decomposition holds:

$$A = A_1 \cup \ldots \cup A_k \cup V \tag{3.11}$$

where $A_1, \dots, A_k$ are transient cycles with $H(U(A)) = H(y_i)$ and $\forall x_i \in A_i, x_j \in A_j$: $\mathcal{S}(x_i, x_j) \subset \{y_1, \dots, y_m\}$

ii) At least one of the A_j's must contain points in $F(A)$

iii)

$$V = \{y_1\} \cup \ldots \{y_m\} \cup \tilde{V} \tag{3.12}$$

$\tilde{V}$, if it is non-empty, is made by points with energy strictly larger than $H(y_i)$ and it is completely attracted by the union

$$\tilde{A} = A_1 \cup \ldots \cup A_k \tag{3.13}$$

namely

$$\tilde{V} \cap M = \emptyset \quad \tilde{V} \subset \bar{B}(\tilde{A})$$

iv) for all y_k there exists $\omega : y_k \to F(A)$:

$$H(x) \le H(y_k)\; \forall x \in \omega$$

Remark

1) In the case where A is a subset of the basin of attraction $\bar{B}(x)$, for some x, the corresponding $\tilde{A}$ can be assimilated to x and $V \equiv \tilde{V}$ to $A \setminus \{x\}$.

2) The point iv) above says that a y_k can even be only in common to cycles A_j not containing points in $F(A)$ but, since there exists a path ω from y_k to $F(A)$ never overpassing in energy $H(y_k)$, at least one of these A_j's must have a minimal saddle $y_h,\ h \neq k$, in its boundary, which is again connected to $F(A)$ by a path never overpassing in energy $H(y_k) = H(y_h)$.

The simplest non-trivial example is the one in which $k = 2$, the number of internal saddles of maximal energy is one:

$$A = A_1 \cup A_2 \cup V,$$

$$V \ = \ \tilde{V} \cup \{y_1\},$$

and

$$|F(A)| = 1 \quad : F(A) \equiv \{x\}.$$

Given a subset $Q \subset S$ we use as before τ_Q to denote the first hitting time to Q:

Proof.

Since A is a cycle then $H(U(A)) > H(y_i)$. By propositions 3.1 and 3.5 the maximal connected components $A_1, \dots, A_k \subset A$ of states in A with energy strictly less than $H(y_i)$ are cycles contained in A with $H(U(A_i)) \ = \ H(y_i)$.

Since, by proposition 3.3 $\mathcal{S}(x_i, x_j)$ does not depend on x_i and x_j, then $\forall x_i \in A_i,\ x_j \in A_j\ :\ \mathcal{S}(x_i, x_j) \ = \ \mathcal{S}(F(A_i), F(A_j)) \subset \{y_1, \dots y_m\}$.

The only remaining, not completely immediate point is iv): it easily follows from the fact that

$$\forall\, x \ \in \ A \cap M, \quad H(\mathcal{S}(x, F(A)) \ \leq \ H(y_i)$$

since $\{y_1, \dots, y_m\}$ are the internal natural saddles with maximal energy.

The results contained in the following Proposition 3.7 are less immediate than the previous ones. As a matter of fact they are already known even in a more general situation (see [FW] for a proof based on the F-W graphical technique). We provide here, in the reversible case, a new, purely probabilistic, proof which, in our opinion, is much more transparent since it is based on a simple intuitive argument in which the *resistance times* play an important role. As we already remarked in the introduction and it will appear clear in the following section this is a crucial mechanism that, apparently, escaped to previous researches on the subject.

Proposition 3.7

Suppose hypotheses H1 and H2 are satisfied. Given a cycle A

i) for all $\epsilon > 0$ *there exist* $\beta_0 > 0$ *and* $k > 0$ *such that for any* $\beta > \beta_0$ *and* $\forall x \in A$

$$P_x(\ \tau_{\partial A} \ < \ \exp(\beta\ [\ H(\mathcal{S}(A)) - H(F(A))\ + \varepsilon\ ])\ )\ \geq\ 1 - e^{-k\beta}$$

ii) there exist $\delta > 0$, $\beta_0 > 0$ *and* $k' > 0$ *such that for all* $\beta > \beta_0$ *and* $\forall x, x' \in A$ *:*

$$P_x(\tau_{x'}\ <\ \tau_{\partial A}\ ;\ \tau_{x'}\ <\ \exp(\beta[H(\mathcal{S}(A)) - H(F(A)) - \delta])\ \geq\ 1 - e^{-k'\beta}$$

iii) $\forall x \ \in A$, $\forall \epsilon > 0$ $\hat{y} \ \in \ \partial A$ *and* β *sufficiently large:*

$$P_x(\ X_{\tau_{\partial A}} = \hat{y}) \geq \exp(-\varepsilon\beta) \exp(-\beta[H(\hat{y}) - H(\mathcal{S}(A))]\ )$$

Proof

The proof uses induction on the total number of the internal natural saddles $|C(A)|$. We first assume that proprties i) ,ii), iii), are verified for all cycles A with $|C(A)|$ less or equal to an integer $n \geq 0$ and we prove them for all the cycles A with a number of internal natural saddles $|C(A)|$ less or equal to $n+1$; then we prove these properties for the case $n = 0$, the basis of the induction, which corresponds to A being the strict domain of attraction of a "plateau" $F(A)$ of communicating, equal energy, points (in particular $F(A)$ can reduce to a single local minimum x).

Consider a generic cycle A with a number of internal *maximal* saddles $N(A) = m$. We can use the decomposition given by Proposition 3.6, namely :

$$A \ = \ \{y_1, \ldots, y_m\} \cup \tilde{V} \cup \tilde{A}$$

where $\tilde{A}$, defined in (3.13), is a union of cycles A_j which, beyond satisfying the properties specified in Proposition 3.6, have, $\forall j = 1, \ldots, m$, a number of internal saddles $|C(A_j)|$ less or equal to n and then satisfy the recursive hypotheses i), ii), iii).

Let us start proving i) for our cycle A. Given any sufficiently small $\varepsilon > 0$ let

$$T_1 = T_1(\varepsilon) \ \equiv \exp(\beta[H(y_j) - H(F(A)) + \varepsilon/2]) \tag{3.14}$$

and

$$T_2 = T_2(\varepsilon) \ \equiv \exp(\beta[H(\mathcal{S}(A)) - H(F(A)) + \varepsilon]). \tag{3.15}$$

Then the argument is the following one : we will construct, for every state $x \in A$, an event E_{x,T_1} starting at x at time $t = 0$, taking place in the interval of time $[0, T_1]$ and satisfying the following conditions:

1) if E_{x,T_1} takes place, our Markov chain X_t hits ∂A before T_1 and

2)

$$\inf_{x\in A} P(E_{x,T_1}) \ \geq \ \alpha_{T_1} \ > \ 0 \quad \text{with} \quad \lim_{\beta\to\infty} (1-\alpha_{T_1})^{T_2/T_1} = 0 \tag{3.16}$$

In particular we will take

$$\alpha_{T_1} = \exp(-\beta[H(\mathcal{S}(A)) - H(y_j) + \varepsilon/4]) \tag{3.17}$$

Let us now divide the interval $[0 \ T_2]$ into $q = [T_2/T_1]$ (here [ ] means integer part) intervals of length T_1 ; by properties 1) and 2) above of E_{x,T_1}, we easily get the proof of our Proposition 3.7. For, if $\tau_{\partial A} > T_2$, necessarily, by property 1), in none of the q intervals of length T_1 the (translated of) event E_{x,T_1} can have taken place; by 2) and the strong Markov property , part i) of our Proposition directly follows. Then we are reduced to the construction of such an event E_{x,T_1}.

Let us first give a rapid description, in words, of E_{x,T_1}.

Let y^* be a state in $U(A)$, by definition there exists a downhill trajectory from y^* to the set $\tilde{A}$:

$$\bar{x}_0 = y^*, \bar{x}_1, \bar{x}_2, \ldots., \bar{x}_k \in \tilde{A}; \qquad \bar{x}_1, \ldots., \bar{x}_{k-1} \in V$$

with $H(\bar{x}_{i+1}) \leq H(\bar{x}_i)$. We put $k = 1$ if y^* is communicating with $\tilde{A}$. Let A_{j^*} be the particular component of $\tilde{A}$ hitted by this trajectory (i.e. $\bar{x}_k \in A_{j^*}$). The event E_{x,T_1} is then defined by requiring that the process hits the set A_{j^*} in a time much shorter than T_1, and then follows the trajectory obtained by $\bar{x}_0, \ldots., \bar{x}_k$ by time reversal.

More precisely let $\epsilon' < \epsilon/2$ and let

$$\tau_{\partial A_{j^*}}^{(>\tau_{A_{j^*}})} \equiv \min\{t > \tau_{A_{j^*}};\, X_t \notin A_{j^*}\}$$

where, as before, $\tau_{A_{j^*}}$ is the first hitting time to the set A_{j^*}. Then:

$$E_{x,T_1} \equiv \{\tau_{A_{j^*}} < T_1 e^{-\epsilon'\beta}\} \cap \{\tau_{\partial A_{j^*}}^{(>\tau_{A_{j^*}})} < T_1 e^{-\epsilon'\beta}\} \cap \{X_{\tau_{\partial A_{j^*}}^{(>\tau_{A_{j^*}})}} = \bar{x}_{k-1}\} \cap$$

$$\{\cap X_{\tau_{\partial A_{j^*}}^{(>\tau_{A_{j^*}})}+s} = \bar{x}_{k-1-s} \qquad \forall s = 1, ..., k-1\}$$

By using the strong Markov property we have:

$$P(E_{x,T_1}) = P_x(\tau_{A_{j^*}} < T_1 e^{-\epsilon'\beta}) P_{X_{\tau_{A_{j^*}}}}(\{\tau_{\partial A_{j^*}} < T_1 e^{-\epsilon'\beta}\} \cap \{X_{\tau_{\partial A_{j^*}}} = \bar{x}_{k-1}\}).$$

$$.P(\bar{x}_{k-1}, \bar{x}_{k-2}).P(\bar{x}_{k-2}, \bar{x}_{k-3}) ...... P(\bar{x}_1, y^*) \tag{3.18}$$

We start estimating the first term in the r.h.s. of (3.18). For any path ω going from x to A_{j^*}, we define the following times, depending on ω:

$$s_0 = 0$$

$$t_k = \min\{t \geq s_{k-1}; \quad \omega_t \in \tilde{A}\}$$

$$s_k = \min\{t > t_k; \quad \omega_t \notin A_{j_k}\}$$

where $j_k = j_k(\omega)$ is such that $\omega_{t_k} \in A_{j_k}$ and $k = 1, ..., l(\omega)$ so that $j_{l(\omega)} = j^*$.
Let

$$\bar{A} \equiv \tilde{A} \cup \{x \in A :\, H(x) = H(y_i)\}$$

¿From Proposition 3.6 it is easy to see that there exists a set of paths $\bar{\Omega}$ going from x to A_{j^*} with the following characteristics:

Given any $\bar{\omega} \in \bar{\Omega}$, if $\bar{\omega}_1 = x$ with $x \notin \bar{A}$, $\bar{\omega}$ first enters $\bar{A}$, following a downhill sequence, and then it does not leave $\bar{A}$ anymore. Subsequently $\bar{\omega}$ follows a well specified sequence of cycles $A_1, ..., A_{j^*}$ and saddles $y_1, ..., y_l$, spending a certain time in each A_j (typically of order of $\exp\{\beta[H(y_j) - H(F(A_j))]\}$, exiting from A_j through the saddles $y_j \in \partial A_j$. Moreover the path $\bar\omega$ does not visit more than once the same saddle y_j and it is downhill for each $t \in (s_k, t_k]$ for each $k = 1, ..., l(\bar\omega)$, $l(\bar\omega) \le m$ and $H(\bar\omega_{s_k}) = H(\mathcal{S}(A_{j_k})) = H(y_i)$. The existence of such an $\bar\Omega$ suggests the way to estimate the first factor in the r.h.s. of (3.18). In order to simplify the notation we set:

$$A_{j_1(\bar\omega)}, ..., A_{j_{l(\bar\omega)}(\bar\omega)} \equiv A_1, ..., A_{j^*}$$

$$s_1(\bar\omega), ..., s_{l(\bar\omega)}(\bar\omega) \equiv s_1, ..., s_l$$

We have:

$$P_x(\tau_{A_{j^*}} < T_1 e^{-\epsilon'\beta}) \ge$$

$$\ge P_x(X_t = \bar\omega_t \quad \forall t \le t_1) \min_{x\in A_1} P_x(\{\tau_{\partial A_1} < \frac{T_1}{l(\bar\omega)} e^{-\epsilon'\beta}\} \cap \{X_{\tau_{\partial A_1}} = \bar\omega_{s_1}\}).....$$

$$..... \min_{x\in A_{j^*-1}} P_x(\{\tau_{\partial A_{j^*-1}} < \frac{T_1}{l(\bar\omega)} e^{-\epsilon'\beta}\} \cap \{X_{\tau_{\partial A_{j^*-1}}} = \bar\omega_{s_{l-1}}\}).$$

$$.P_{\bar\omega_{s_{l-1}}}(X_t = \bar\omega_{t+s_{l-1}} \quad \forall t \in (o, t_l - s_{l-1}))$$

Now, since $\bar\omega$ is downhill for any $t \in (s_{k-1}, t_k)$, $\forall k = 1, ..., l$, by using the following obvious inequality:

$$P_{\bar\omega_{s_n}}(X_{t-s_n} = \bar\omega_t \quad \forall t \in (s_n, t_{n+1})) \ge e^{-\alpha'\beta}$$

for some $\alpha' \to 0$ as $\beta \to \infty$ and by using the iterative hypotheses which imply:

$$\max_{j=1,..,j^*} P(\{\tau_{\partial A_j} < \frac{T_1}{l(\bar\omega)} e^{-\epsilon'\beta}\} \cap \{X_{\tau_{\partial A_j}} = \bar\omega_{s_j}\}) > e^{-\alpha''\beta}$$

with $\alpha'' \to 0$ as $\beta \to \infty$, we have:

$$P_x(\tau_{A_{j^*}} < T_1 e^{-\epsilon'\beta}) \ge e^{-\alpha\beta}$$

with $\alpha \to 0$ as $\beta \to \infty$.

Notice that we have used $\bar\omega$ only to deduce, from it, a sequence $A_1, ...A_{j^*}$ of cycles and a sequence of downhill paths emerging from saddles. Moreover, in the above argument it is essential to give the process the freedom to spend in each cycle A_j a suitable random time.

Again by the iterative hypotheses the second term in r.h.s. of (3.18) is estimated by:

$$P_{X_{\tau_{A_{j^*}}}}(\{\tau_{\partial A_{j^*}} < T_1 e^{-\epsilon'\beta}\} \cap \{X_{\tau_{\partial A_{j^*}}} = \bar{x}_{k-1}\}) \geq e^{-\epsilon\beta} e^{-\beta[H(\bar{X}_{k-1})-H(U(A_{j^*}))]}$$

and in conclusion we obtain the following estimate:

$$P(E_{x,T_1}) \geq e^{-\alpha'\beta} e^{-\beta[H(y^*)-H(U(A_{j^*}))]} = e^{-\alpha'\beta} e^{-\beta[H(\mathcal{S}(A))-H(y_i)]}$$

This concludes the proof of property i) for A

Let us now go to the proof of ii). We proceed similarly to what we did to prove i) using recurrence in $\tilde{A}$ and strong Markov property. We first prove that, for some positive δ :

$$\lim_{\beta\to\infty} P_x(\tau_{x'} \ < \ \exp(\beta[H(\mathcal{S}(A)) - H(F(A)) - \delta]) \ = \ 1 \tag{3.19}$$

Again for any pair $x, x' \in A$ appearing in ii) we construct an event E'_{x,x',T_1} similar to E_{x,T_1}; $E'_{x,x,'T_1}$ contains trajectories of our process starting at x ending at x' during a time interval at most T_1. Notice that $T_1(\varepsilon)$, for ε sufficiently small, is of the form $\exp(\beta[H(\mathcal{S}(A)) - H(F(A)) - \delta])$ (appearing in ii)) with, say, $\delta = [H(\mathcal{S}(A) - H(y_j)]/2$.

We describe E'_{x,x',T_1} in words leaving to the reader the easy task of a precise definition, along the same lines previously seen for E_{x,T_1} .

If E'_{x,x',T_1} takes place then, starting from x, we first descend in a time much shorter than $T_1(\varepsilon)/3$ to $\tilde{A}$ (if we still were not there) and this happens with a probability almost one for β large; then in a time of order $T_1(\varepsilon)/3$, if ε is chosen sufficiently small : 1) we never get out of the set $\bar{A} \ \equiv \ \{y_1, \dots, y_m\} \cup \tilde{A}$; this happens with probability approaching one for $\beta \to \infty$, as it easily follows from reversibility, since, to get out of $\bar{A}$ it will typically take a time of order $\exp(\beta[H(\mathcal{S}(A) - H(F(A))])$, much larger than $T_1(\varepsilon)$ for ε sufficiently small since, of course, $H(y_j) \ < \ H(\mathcal{S}(A))$; (see, for instance Lemma 1 in [KO1] ; 2) we enter the cycle A_j say A', containing x'. It follows from the recursive hypothesis ii) valid for A' that,$\forall \varepsilon > 0$ and β large, with probability almost one, before leaving A', and in a time typically much shorter that $T_1(\varepsilon)/3$ we touch x'. This concludes the proof of (3.19).

By using reversibility, we get the full condition ii) since with probability tending to one as $\beta \to \infty$ for every $x \in A$, $\varepsilon > 0$ one has:

$$\tau_{\partial A} > \ \exp(\beta[H(\mathcal{S}(A)) - H(F(A)) - \varepsilon) \tag{3.20}$$

and choosing ε sufficiently small :

$$\exp(\beta[H(\mathcal{S}(A)) - H(F(A)) - \delta]) \; < \; \exp(\beta[H(\mathcal{S}(A)) - H(F(A)) - \varepsilon)$$

Let us now prove point iii). Given $y \in F(A)$ we can estimate from below the probability $P_x(X_{\tau_{\partial A}} = \hat{y})$ by imposing to the process to visit the state y before the exit time $\tau_{\partial A}$ as follows:

$$P_x(X_{\tau_{\partial A}} = \hat{y}) \geq \sum_{z \in A} \sum_{s=0}^{\infty} P_x(X_{s'} \in A \quad \forall s' \leq s, \, X_s = z) P_z(\tau_y < \tau_{\partial A}).$$

$$\cdot \sum_{t=1}^{\infty} \sum_{\bar{x}_1,\dots \bar{x}_{t-1} \in A \setminus y} P_y(X_1 = \bar{x}_1, \dots.., X_{t-1} = \bar{x}_{t-1}, \, X_t = \hat{y}) \tag{3.21}$$

By using reversibility and the already proved point ii), valid now for the whole cycle A, the last term in the r.h.s. of (3.21) can be estimated as follows:

$$\sum_{t=1}^{\infty} \sum_{\bar{x}_1,\dots \bar{x}_{t-1} \in A \setminus y} P_y(X_1 = \bar{x}_1, \dots.., X_{t-1} = \bar{x}_{t-1}, \, X_t = \hat{y}) =$$

$$e^{-\beta[H(\hat{y}) - H(y)]} . \sum_{t=1}^{\infty} \sum_{\bar{x}_1,\dots \bar{x}_{t-1} \in A \setminus y} P_{\hat{y}}(X_1 = \bar{x}_{t-1}, \dots.., X_{t-1} = \bar{x}_1, \, X_t = y) =$$

$$= e^{-\beta[H(\hat{y}) - H(y)]} . [P(\hat{y}, y) +$$

$$+ \sum_{t=2}^{\infty} \sum_{\bar{x}_{t-1} \in A \setminus y} P(\hat{y}, \bar{x}_{t-1}) \sum_{\bar{x}_1,\dots \bar{x}_{t-2} \in A \setminus y} P_{\bar{x}_{t-1}}(X_1 = \bar{x}_{t-2}, \dots.., X_{t-1} = y)] \geq$$

$$\geq e^{-\beta[H(\hat{y}) - H(y) + \epsilon']} \min_{z \in A \setminus y} P_z(\tau_y < \tau_{\partial A}) \geq e^{-\beta[H(\hat{y}) - H(y) + \epsilon]} .$$

Putting this estimate in (3.21) and using again point ii) to estimate from below the quantity $P_z(\tau_y < \tau_{\partial A})$ we obtain, for β large enough,

$$P_x(X_{\tau_{\partial A}} = \hat{y}) \geq e^{-\beta[H(\hat{y}) - H(y) + \epsilon']} \sum_{s=0}^{N} P_x(\tau_{\partial A} \geq s) \geq$$

$$\geq e^{-\beta[H(\hat{y})-H(y)+\epsilon']}.N.P_x(\tau_{\partial A} \geq N)$$

It is easy to prove that there exists ζ going to zero as $\beta \to \infty$ such that if we chose $N = e^{\beta[H(\mathcal{S}(A))-H(F(A))-\zeta]}$ we have $P_x(\tau_{\partial A} \geq N) > 1/2$.

This conclude the proof of iii).

To conclude the proof of our proposition we have to show that properties i), ii), iii) are true for A such that the number $N(A)$ of internal saddles is zero, namely when A is part of (or coincides with) the strict basin of attraction of $F(A)$, $F(A)$ being a plateau in the previously specified sense.

Let we are given such an A. Property i) easily follows by the same argument used before: we costruct, for any $x \in A, \varepsilon > 0$ an event $E_{x,T}$ with $T = T(\varepsilon) = \exp(\beta\varepsilon/2)$ which consists in descending from x to $F(A)$ in a time at most $T/2$ following a downhill path ω from x to $F(A)$; then in following an uphill path ω' from $F(A)$ up to $U(A)$ in a time shorter that $T/2$. This path ω' is the time-reversed of a path going downhill from $U(A)$ to $F(A)$. The paths ω and ω' certainly exist as A is the strict basin of attraction of $F(A)$.

With T_2 given by (3.15) we easily verify, in the present case, (3.16) and (3.17) with $T_1 \equiv T(\varepsilon)$ since : 1) for every $\varepsilon > 0$ the descent to $F(A)$, along a downhill path, takes place in a suitable finite time, much smaller than $T(\varepsilon)$, with a probability approaching one as $\beta \to \infty$ and : 2) the ascent from $F(A)$ to $U(A)$, along an uphill path, in a suitable finite time , much smaller than $T(\varepsilon)$ $\forall \varepsilon > 0$ and β sufficiently large takes place with a probability larger than $\exp(\beta[H(U(A)) - H(F(A)) + \varepsilon/4])$. Then property i) easily follow.

Combining the previously discussed methods, that we used before to prove ii) and iii) starting from the inductive hypothesis, with the idea leading to the construction of the above event $E_{x,T}$ we easily get, in our present case of $A = \bar{B}(F(A))$, ii) and iii) (we leave the details to the reader).

This concludes the proof of Proposition 3.7.

We will analyze now, in more detail, the first exit from a cycle A. In particular, following the ideas developed in the framework of the so called "pathwise approach to metastability", (see [CGOV]), we will prove asymptotic exponentiality of the properly renormalized first exit time from any cycle in the limit $\beta \to \infty$. Then we will deduce the asymptotic behaviour of the expectation of this exit time; notice that the methods developed in the previous sections naturally lead only to estimates in probability of the typical exit times but, as we will see, we can even get a good control on the tails of the distribution of these random variables and this will be important to get the asymptotics of the averages.

Suppose given, once for ever, a cycle A. Given a point $x \in F(A)$ let the time $T_\beta = T_\beta(x)$ be defined by

$$P_x(\tau_{\partial A} > T_\beta) \ = \ e^{-1}. \tag{3.22}$$

The above definition is interesting since, as it easily follows from Proposition 3.7, the asymptotic distribution, (in the sense of the most probable behaviour), of the first exit time from A , does not depend on $x \ \in \ A$ in the sense of logarithmic equivalence; thus, as we will see, also T_β will not depend on $x \ \in \ A$, always in the sense of logarithmic equivalence.

Proposition 3.7, i) and the usual estimate, (see Lemma 1 in [KO1]) based on reversibility, give, $\forall\, \varepsilon > 0, \ \forall\, x \ \in \ A$,

$$\lim_{\beta\to\infty} P_x(\exp(\beta\,[H(\mathcal{S}(A)) - H(F(A)) - \varepsilon]) < \tau_{\partial A} < \exp(\beta[H(\mathcal{S}(A)) - H(F(A)) + \varepsilon])$$

$$= 1 \tag{3.23}$$

¿From (3.22), (3.23) it easily follows that, $\forall\, \varepsilon > 0, \ \forall\, x \ \in \ A$,

$$\exp(\beta\,[H(\mathcal{S}(A)) - H(F(A)) - \varepsilon]) < T_\beta < \exp(\beta[H(\mathcal{S}(A)) - H(F(A)) + \varepsilon]$$

Proposition 3.8

$\forall\, x \ \in \ A, \forall\, s \ \in \ \mathbf{R}^+$:

$$\lim_{\beta\to\infty} P_x(\frac{\tau_{\partial A}}{T_\beta} \ > \ s\ ) = \ e^{-s}. \tag{3.24}$$

Proof.

Given $s,\ t \ \in \ \mathbf{R}^+$ we write:

$$P_x(\tau_{\partial A} \ > \ (t+s)\ T_\beta) = \sum_{\bar{x}\in A} P_x(\tau_{\partial A} \ > \ (t+s)\ T_\beta\ ;\ X_t = \bar{x})$$

$$= \sum_{\bar{x}\in A} P_x(\tau_{\partial A} \ > \ t\ T_\beta\ ;\ X_t = \bar{x}) \ \ P_{\bar{x}}(\tau_{\partial A} \ > \ s\ T_\beta) \tag{3.25}$$

We can write, $\forall T < sT_\beta$:

$$P_{\bar{x}}(\tau_{\partial A} \ > \ s\,T_\beta\,) = P_{\bar{x}}(\tau_{\partial A} \ > \ s\,T_\beta\,;\tau_x < T) + P_{\bar{x}}(\tau_{\partial A} \ > \ s\,T_\beta\,;\tau_x \geq T). \tag{3.26}$$

We have, from Proposition 3.7 ii), that there exists a $\delta > 0$ such that for the time $\bar{T}_1$, defined as

$$\bar{T}_1 = \exp(\beta[H(\mathcal{S}(A)) - H(F(A)) - \delta] \tag{3.27}$$

moreover, denoting by $\mathrm{o}(\beta)$ an infinitesimal quantity as $\beta \to \infty$,

$$P_x(\tau_{\partial A} \ > \ s\,T_\beta) \ - \mathrm{o}(\beta) \leq P_{\bar{x}}(\tau_{\partial A} \ > \ s\,T_\beta\,;\tau_x < \bar{T}_1) \ \leq$$

$$\leq P_x(\tau_{\partial A} \ > \ s\,(T_\beta - \bar{T}_1/s)) \qquad (3.28).$$

since, for β sufficiently large, $\bar{T}_1 < sT_\beta$. On the other hand, from (3.26) and again from Proposition 3.7 ii), we get:

$$P_{\bar{x}}(\tau_{\partial A} \ > \ s\,T_\beta\,) - \mathrm{o}(\beta) \leq P_{\bar{x}}(\tau_{\partial A} \ > \ s\,T_\beta\,;\tau_x < T) \leq P_{\bar{x}}(\tau_{\partial A} \ > \ s\,T_\beta\,) \qquad (3.29).$$

¿From (3.22), (3.25), (3.28), (3.29), we get the result; for, (3.25) together with (3.28), (3.29), since $\lim_{\beta\to\infty} \bar{T}_1/T_\beta = 0$, imply the convergence of $P_x(\frac{\tau_{\partial A}}{T_\beta} \ > \ s\ )$ to a (possibly degenerate) exponential; (3.22) fixes the average ($\equiv 1$) of the limiting exponential.

Proposition 3.9

For every $x \in A$ and $\varepsilon > 0$, if $\mathbf{E}_x$ denotes average over the trajectories of the process starting, at $t = 0$, from x, we have:

$$\exp(\beta\,[H(\mathcal{S}(A)) - H(F(A)) - \varepsilon]) < \mathbf{E}_x(\tau_{\partial A}) < \exp(\beta[H(\mathcal{S}(A)) - H(F(A)) + \varepsilon] \tag{3.30}$$

Proof.

For every integer valued random variable ξ we have:

$$\mathbf{E}(\xi) \ = \ \sum_{m=1}^{\infty} \mathbf{P}(\xi \geq m) \tag{3.31}$$

Now, following the argument of proof of Proposition 3.7 i), based on the introduction of the set E_{x,T_1}, using estimate (3.15) with nT_2 in place of $T_2 = T_2(\varepsilon)$ and the strong Markov property, it is easy to get the following estimate :

$$P_x(\tau_{\partial A} \ > \ nT_2) \ < \ \exp(-nc) \tag{3.32}$$

valid for every $x \ \in \ A$, $\varepsilon > 0$, for a suitable positive constant c independent of β ; we recall that

$$T_2 = T_2(\varepsilon) \ \equiv \exp(\beta[H(\mathcal{S}(A)) - H(F(A)) + \varepsilon])$$

Applying formula (3.31) we get:

$$\mathbf{E}^{(\beta)}{}_x(\tau_{\partial A}) \ = \ \sum_{m=1}^{\infty} P_x^{(\beta)}(\tau_{\partial A} \geq m) \tag{3.33}$$

where we have put in evidence, with a superscript, the dependence on β of the distribution of our process.

Now the result of Proposition 3.9 follows, via (3.32), from standard arguments (see, for instance, [CGOV]).

Section 4. The exit tube.

In this section we will analyze the typical trajectories of the first excursion outside a cycle A.

We will start analyzing the first descent from any point y_0 in A to $F(A)$. For every $y_0 \ \in A$ we will define a tube of trajectories $\mathcal{T}_{y_0}$ (it will be called " standard"); then we will prove that with high probability, during the first excursion from y_0 to $F(A)$, we will follow this tube. As we will see this tube will be , in general, very widely defined. This means that both possible multiple bifurcations and resistance times will be considered in its definition.

Roughly speaking $\mathcal{T}_{y_0}$ will be characterized by a set of sequences (cascades) of minimaxes $y_1, \dots, y_n$ towards $F(A)$, decreasing in energy, intercalated by sequences of downhill paths $\omega_1, \dots, \omega_n$ and sets $Q_1, \dots, Q_n$ which are a sort of generalized cycles.

Given any point y_0 in A and a downhill path ω_1 starting from y_0, we will define a set $Q_1 \ = \ Q_1(y_0, \omega_1)$. This set Q_1 is a union of cycles having common minimal saddles of the same height. It will have the meaning of the first set where our process, during its first excursion to $F(A)$, is captured if it follows the path ω_1; after entering into Q_1 it will spend some time inside it before leaving it to enter, after another downhill path ω_2 , into another similar set Q_2 and so on untill it enters a cycle containing part of $F(A)$.

Finally we will analyze the problem of the typical tube of trajectories during the first excursion outside a cycle A. As already observed by R. Schonmann for the case of stochastic Ising models (see [S 2]), it will turn out, using reversibility, that this tube is simply related, via a time reversal transformation, to the typical tube followed by the process during the first descent to the bottom $F(A)$ of A.

The construction of Q_1.

Given y_0 let us consider a downhill path ω_1 starting from y_0. We stress, once more, that this path is not in general unique. This means that the whole construction we are defining must be repeated for each path.

Let x_1 be the first stable state in ω_1 (see fig.1 as an example). If such a point x_1 is in $F(A)$ then y_0 belongs to the wide basin of attraction of a connected component $G = G(A, y_0, \omega_1)$ of $F(A)$, i.e. $y_0 \in \bar{B}(G)$.

In this degenerate case we set $Q_1 \equiv G(A, y_0, \omega_1)$ and the *cascade of saddles* $y_0, y_1, \ldots, y_n$ reduces to y_0.

Let us now suppose that $x_1 \notin F(A)$. Let H_1 be the energy of the saddle (not necessarily unique) between x_1 and $F(A)$:

$$H(\mathcal{S}(x_1, F(A))) = H_1$$

We denote by $A^{(1)}$ the cycle containing x_1 with energy less than H_1. By definition of H_1 we have that $A^{(1)} \cap F(A) = \emptyset$. We define

$$\mathcal{S}^{(1)} \equiv \mathcal{S}(A^{(1)})$$

and

$$\tilde{\mathcal{S}}^{(1)} \equiv \cap_{z \in \mathcal{S}^{(1)}} \bar{P}(z)$$

Let us now consider the cycles: $A_1^{(2)}, \ldots, A_{k_2}^{(2)}$ with energy less than H_1 not coinciding with $A^{(1)}$, with which $\tilde{\mathcal{S}}^{(1)}$ is downhill communicating and such that $A_j^{(2)} \cap F(A) = \emptyset \ \forall\ j = 1, \ldots, k_2$. If there are no such cycles we define $N = 1$ otherwise we continue, by iteration, the construction as follows.

We call $A^{(2)}$ the union of all the cycles $A_j^{(2)}$'s (which do not contain points in $F(A)$),

$$A^{(2)} = \cup_j A_j^{(2)}$$

and $\tilde{\mathcal{S}}^{(2)}$ the union of all the plateaux containing minimal saddles of the $A_j^{(2)}$'s which are not contained in $\tilde{\mathcal{S}}^{(1)}$. Now consider, similarly to before, the cycles $A_1^{(3)}, \dots, A_{k_3}^{(3)}$ with energy less than H_1, not coinciding with any of the previous $A_j^{(2)}$'s, with which $\tilde{\mathcal{S}}^{(2)}$ is downhill communicating and such that $A_j^{(3)} \cap F(A) \;=\; \emptyset \;\forall\, j = 1, \dots, k_3$.

If there are no such cycles we define $N = 2$ otherwise we iterate the construction.

This procedure stops at a given finite index $N = N(x_1, A)$.

It easy to convince oneself that

$$\forall\, j = 2, \dots, N : A^{(j)} \cap A^{(l)} = \emptyset \;\forall\, l = 1, \dots, j-1$$

and

$$\forall\, j = 2, \dots, N-1 : \mathcal{S}^{(j)} \cap \mathcal{S}^{(l)} = \emptyset \;\forall\, l = 1, \dots, j-1$$

We set

$$Q_1 = Q_1(y_0, \omega_1) \equiv (\cup_{j=1}^N A^{(j)}) \cup (\cup_{j=1}^N \cup_{z \in U(A^{(j)})} \bar{P}(z)\;)$$

.

It is easily seen that Q_1 is the maximal connected set, containing x_1, of points x such that

$$H(\mathcal{S}(x, F(A))) = H_1$$

The boundary of Q_1 is given by

$$\partial Q_1 = \partial^u Q_1 \cup \partial^d Q_1 \quad \partial^u Q_1 \cap \partial^d Q_1 = \emptyset$$

where $\partial^u Q_1$ is made of points z with energy larger than H_1 and so, trivially, with

$$H(\mathcal{S}(z, F(A))) > H_1$$

whereas $\partial^d Q_1$ is made of points z with energy smaller than H_1 , belonging to some cycle containing points of $F(A)$ and then such that

$$H(\mathcal{S}(x, F(A))) < H_1.$$

Let us call $\bar{\mathcal{S}}_1$ the subset of $\tilde{\mathcal{S}}^{(1)} \cup \ldots \cup \tilde{\mathcal{S}}^{(N-1)}$ which is downhill communicating with $\partial^d Q_1$. Chose a point y_1 in $\bar{\mathcal{S}}_1$ and a downhill path ω_2 starting from y_1 not belonging to Q_1. Start again from y_1, ω_2 a hierarchical construction totally analogous to the previous one. Denote by x_2 the first stable state in ω_2 and by H_2 the energy of the saddle between x_2 and $F(A)$. By definition $H_2 < H_1$. As before we construct new sets Q_2 (depending on the choice of y_1 in $\bar{\mathcal{S}}_1$ and of the path ω_2 starting from y_1).
If we continue in this way and recursively construct $y_2, y_3, \ldots$, $\omega_3, \omega_4, \ldots$, $Q_3, Q_4, \ldots$ and $\mathcal{S}_3, \mathcal{S}_4, \ldots$ we necessarily end up with a y_{M-1}, ω_M, for some finite M, with x_M belonging to a connected component G^* (plateau) of $F(A)$.

We set $Q_M \equiv G^*$.

We are now ready to define our tube of trajectories $\mathcal{T}_{y_0}$.

Suppose first that $y_0, \omega_1, y_1, \omega_2, \ldots, y_{M-1}, \omega_M$ are given.

Then , automatically, $x_2, \ldots, x_{M-1}, x_M, Q_1, \ldots, Q_{M-1}, Q_M$ are given.

Let

$$\mathcal{T}(y_0, \omega_1, y_1, \omega_2, \ldots, y_{N-1}, \omega_M) \ = \ y_0 \cup \omega_1 \cup Q_1 \cup \omega_2, \ldots, Q_{N-1} \cup \omega_M \cup Q_M, \tag{4.1}$$

Any sequence like $\mathcal{T}(y_0, \omega_1, y_1, \omega_2, \ldots, y_{N-1}, \omega_M)$, obtained via the above construction, will be called *standard cascade*; it can be visualized as a sequence of falls and communicating lakes.

We set :

$$\mathcal{T}_{y_0} = \bigcup_{\omega_1, y_1, \ldots, y_{M-1}, \omega_M} \ = \ \mathcal{T}(y_0, \omega_1, y_1, \omega_2, \ldots, y_{N-1}, \omega_M) \tag{4.2}$$

We call $\mathcal{T}_{y_0}$ *standard tube* starting from y_0. Notice that, by construction:

$$H(y_{j-1}) \ < \ H(y_j), \quad \max_{F(Q_j)} H \ < \ H(y_{j-1}), \ j = 1, \ldots M$$

and $F(Q_1), \ldots, F(Q_{M-1})$ are strictly higher, in energy, than $F(A)$

A particularly simple case is when each Q_j reduces itself to a single cycle A_j^*.

Now we can state our main result.

Theorem 1

Given a cycle A, for every $y_0 \in A$,

i)

$$\exists\, \delta > 0 \ \ \textit{such that} \ \ \lim_{\beta\to\infty} P_{y_0}(\tau_{F(A)} < \exp(\beta[H(y_1) - H(F(A)) - \delta]) = 1$$

ii)

$$\lim_{\beta\to\infty} P_{y_0}(x_t \ \in \ \mathcal{T}_{y_0} \ \forall\, t \le \tau_{F(A)}) \ = \ 1$$

moreover :

iii)

$$\lim_{\beta\to\infty} P_{y_0}(\forall\, t \le \tau_{F(A)} \ : \ x_t \ \in \ \mathcal{T}(y_0, \omega_1, y_1, \omega_2, \dots, y_{N-1}, \omega_M)$$

$$\textit{for some}\, y_0, \omega_1, y_1, \omega_2, \dots, y_{N-1}, \omega_M) \ = \ 1,$$

more precisely :

iv)

with a probability tending to one as $\beta \to \infty$*, there exists a sequence* $y_0, \omega_1, y_1, \omega_2, \dots, y_{N-1}, \omega_M$ *such that our process starting at* $t = 0$ *from* y_0*, between* $t = 0$ *and* $t = \tau_{F(A)}$*, after having followed the initial downhill path* ω_1*, visits, sequentially, the sets* Q_1, Q_2, Q_{M-1} *exiting from* Q_j *through* y_j *and then following the path* ω_{j+1} *before entering* Q_{j+1}*.*

For every $\varepsilon > 0$ *with a probability tending to one as* $\beta \to \infty$ *the process spends inside each* Q_j *a time less than* $\exp(\beta[H(y_j) - H(F(Q_j)) + \varepsilon])$

Finally: before exiting from Q_j *it can perform an arbitrary sequence of passages through the cycles* $A^{(j)}$ *belonging to* Q_j *. Each passage is made through a minimal saddle* z_j *in the boundary of* $A^{(j)}$ *; for every* $\varepsilon > 0$ *with a probability tending to one as* $\beta \to \infty$*, once the system enters into a particular* $A^{(j)}$*, it spends there, a time* T *:* $\exp(\beta[H(z_j) - H(F(A^{(j)})) - \varepsilon]) \ < \ T \ < \ \exp(\beta[H(z_j) - H(F(A^{(j)})) + \varepsilon])$

Proof

By construction, using Proposition 3.7 applied either directly to our original cycle A if $y_0 \in \bar{B}(F(A))$ or to the cycles in Q_1 otherwise. The rest of the theorem also easily follows from Proposition 3.7 applied to the cycles contained in the Q_j's. We leave the details to the reader.

Now, given any cycle A, we want to describe, in the maximal possible detail, the first excursion from $F(A)$ to ∂A.

Following Schonmann ([Sch2]), we first give some simple general definitions.

Given $x, y \in S$, we denote by $\Omega^*(x,y)$ the set of all paths ω starting in x, visiting y at some finite time t and never visiting x and y in between :

$$\Omega^*(x,y) \equiv \{\omega = x_1, \dots, x_t \text{ for some } t \ : \ x_1 = x, \ x_t = y, \ ; \ x_2, \dots, x_{t-1} \neq x, y\} \tag{4.3}$$

We denote by R the time reversal operator defined on finite paths:

$$\forall \ \omega \equiv (x_1, \dots, x_t) \ : \ R\omega \equiv \bar{\omega} \equiv (x_t, \dots, x_1) \tag{4.4}$$

We naturally define, for every set of paths Δ,

$$R\Delta = \{\bar{\omega} = R\omega \ ; \ \omega \ \in \ \Delta\}.$$

Let us call $\bar{\tau}_x$ the last time our process visits the state x before touching, for the first time, y, namely :

$$\bar{\tau}_x \ \equiv \ \max\{t < \tau_y : X_t = x\} \tag{4.5}$$

Given a finite path $\tilde{\omega} = \tilde{x}_1, \dots, \tilde{x}_t$, we say that our process $\{X_t\}_{t>0}$ *starts as* $\tilde{\omega}$ if $X_1 = \tilde{x}_1, \dots, X_t = \tilde{x}_t$.

For any $x, y \in S$ we define a measure on the (infinite) paths $\omega = x_1, \dots, x_t, ...$ starting at x $(x_1 \equiv x)$, as follows:

if $\omega \notin \Omega^*(x,y)$ then $\rho(\omega) = 0$

if $\omega \in \Omega^*(x,y)$ we set $\tilde{\omega}_i = \omega_i$ for all $i < \inf\{t > 0; \omega_t = y\}$, that is $\tilde{\omega}$ is the finite path given by the first segment of ω before hitting y. Then

$$\rho(\omega) = P_x(\ \{X_t\}_{t\geq 0} \text{ starts as } \tilde{\omega} \ | \omega \in \Omega^*(x,y))$$

and by $\bar{\rho}(\omega)$ the measure on the paths $\omega = x_1, \dots, x_t, ...$ given by :

$$\bar{\rho}(\omega) = P_y(\ \{X_t\}_{t\geq 0} \text{ starts as } R\tilde{\omega} \ | \omega \in \ \Omega^*(y,x) \ ) \quad \text{, if } \ \omega \ \in \ \Omega^*(y,x)$$

$$\bar{\rho}(\omega) = 0 \quad \text{otherwise}$$

In [Sch2] it is proven that, for every $x, y \in S$, every $\Lambda \in \Omega^*(x,y)$,

$$P_x(X_t \ \in \ \Lambda \ \forall t \ \in [\bar{\tau}_x, \tau_y]) = \rho(\Lambda) = \bar{\rho}(R\Lambda) = P_y(X_t \ \in \ R\Lambda \ \forall t \ \in [\bar{\tau}_y, \tau_x]) \tag{4.6}$$

Let us now denote by $\partial^- F(A)$ the set of all $\hat{x} \ \in \ F(A)$, uphill communicating with $A \setminus F(A)$. Given a point $\hat{x}$ in $\partial^- F(A)$, consider the set $V(\hat{x})$ of all the points $x \ \in \ A$

uphill communicating with some point $\hat{y} \in U(A)$ and such that there exists a standard cascade $\mathcal{T}(y_0 = \hat{y}, \omega_1, y_1, \omega_2, \dots, y_{N-1}, \omega_M)$ starting from $\hat{y}$ and ending in $\hat{x}$: $\hat{x}$ will belong to some component G^* of $F(A)$; $G^* \equiv Q_M$ and ω_M will end entering into Q_M at $\hat{x}$.

Now we are able to state our main result about the typical trajectories realizing the escape from a cycle A.

Theorem 2

Let

$$\bar{\tau}_{F(A)} = \max\{t < \tau_{S\setminus A} : X_t \in F(A).\}$$

Call $\bar{\partial}^- A$ the set of all the points $x \in A$ uphill connected to $U(A)$ and

$$\mathcal{T} = \cup_{y_0 \in \bar{\partial}^- A} \mathcal{T}_{y_0} \tag{4.7}$$

the set of all possible standard tubes starting from points in $\bar{\partial}^- A$ and ending in $F(A)$. Then:

i)

$$P_{F(A)}(X_{\tau_{S\setminus A}} \in U(A) \; ; \; X_t \in R\mathcal{T} \; \forall \, t \in [\bar{\tau}_{F(A)}, \tau_{S\setminus A} - 1]) \to 1 \text{ as } \beta \to \infty \tag{4.8}$$

ii)

Given $\hat{x} \in \partial^- F(A)$ and any $x \in A$,

$$P_x(\exists \, \hat{y} \in V(\hat{x}) \; : \; X_t \in R\mathcal{T}(y_0 = \hat{y}, \omega_1, y_1, \omega_2, \dots, y_{N-1}, \omega_M) \forall t \in [\bar{\tau}_{\hat{x}}, \tau_{\hat{y}} - 1] | X_{\bar{\tau}_{F(A)}} = \hat{x})$$

$$\to 1 \text{ as } \beta \to \infty \tag{4.9}$$

iii)

During the first excursion from $F(A)$ to $S \setminus A$, conditioning to $X_{\bar{\tau}_{F(A)}} = \hat{x}$ (for some $\hat{x} \in \partial^- F(A)$), to $X_{\tau_{S\setminus A}-1} = \hat{y}$ (for some $\hat{y} \in V(\hat{x})$) and to follow a particular " anticascade" $R\mathcal{T}(y_0 = \hat{y}, \omega_1, y_1, \omega_2, \dots, y_{N-1}, \omega_M)$ between $\bar{\tau}_{\hat{x}}$ and $\tau_{\hat{y}}$, all the " time reversed " of the properties specified in point iv) of Theorem 1 hold true ; namely $\forall \, \varepsilon > 0$, with probability tending to one as $\beta \to \infty$ our process , during the above mentioned first excursion, visits the time reversed of the sequence specified in point iv) of Theorem 1 spending in each set the same typical times given there.

Remark

In the particular case (relevant for the applications to stochastic Ising models) where the sets Q_i always coincide with a single cycle A_i, it immediately follows from Theorem 2, that the typical tube of trajectories during the first excursion from $F(A)$ to $S \setminus A$ is an anticascade starting from $\hat{x} \in \partial^- F(A)$ and ending in some $y^* \in U(A)$ given by a sequence: $\bar{A}_1, \bar{\omega}_1, \bar{y}_1, \bar{A}_2, \bar{\omega}_2, \bar{y}_2, \dots, \bar{A}_M, \bar{\omega}_M, y^*$ with the properties:

i)

$H(\bar{y}_i) < H(\bar{y}_{i+1}),\ 1 = 1, \dots, M-1$

ii)

$\bar{y}_i \in S(\bar{A}_{i+1})$.

Some ω_i can be empty; in that case : $\bar{y}_i$ is also a saddle point in $\partial \bar{A}_i$

Aknowledgements.

We want to thank Roman Kotecky, Fabio Martinelli and Roberto Schonmann for very interesting and stimulating discussion.

We want to express thanks for the kind hospitality to the Isaac Newton Institute in Cambridge where part of this work has been done.

After the completion of the present paper we became aware that concepts and arguments, somehow related to the ones developed in the present note, have been considered by O. Catoni [Ca] in the framework of the theory of simulated annealing.

We want to thank G. Ben Arous who gave us the information about that.

This work has been partially supported by the grant SC1-CT91-0695 of the Commission of European Communities

References

[CGOV] M.Cassandro, A.Galves, E.Olivieri, M.E.Vares, " Metastable behaviour of stochastic dynamics: A pathwise approach", J.Stat.Phys. **35**, 603-634 (1984)

[Ca] O.Catoni, "Sharp large deviations estimates for simulated annealing algorithms ", Ann. Inst. H. Poincaré **27**, 291-383 (1991)

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

[KO1] R.Kotecky, E.Olivieri "Droplet dynamics for asymmetric Ising model" Preprint (1992).

[KO2] R.Kotecky, E.Olivieri "Shapes of growing droplets- a model of escape from a matastable phase" in preparation.

[MOS] F. Martinelli, E. Olivieri, E. Scoppola, "Metastability and exponential approach to equilibrium for low temperature stochastic Ising models" Journal of Stat. Phys. **61**, N.5/6 1105 1990.

[NSch1] E.J. Neves, R.H. Schonmann, "Behaviour of droplets for a class of Glauber dynamics at very low temperatures" Comm. Math.Phys. **137** 209 (1991).

[NSch2 E.J. Neves, R.H. Schonmann, "Critical Droplets and Metastability for a Glauber Dynamics at Very Low Temperatures", Prob.Theor.Rel.Fields **91**, 331 (1992)

[OS] E. Olivieri, E. Scoppola, " Markov chains with exponentially small transition probabilities: First exit problem from a general domain -II. The general case" In preparation.

[Sch] R. H. Schonmann, " The pattern of escape from metastability of a stochastic Ising model ", Comm. Math. Phys. **147** (1992) 231-240.

[S1] E.Scoppola " Renormalization group for Markov chains and application to metastability", Jour. Stat. Phys. **73**, 83 (1993)

[S2] E.Scoppola " Metastability for Markov chains: a general procedure based on renormalization group ideas", in "Probability and Phase Transition", Ed. G.Grimmett, NATO ASI Series, Kluwer Ac. Publ.

[S3] E.Scoppola " Renormalization and graph methods for Markov chains", Proceedings of the Conference "Advances in Dynamical Systems and Quantum Physics" - Capri 1993. In press