METASTABILITY AND TYPICAL EXIT PATHS IN STOCHASTIC DYNAMICS

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

(1) Dipartimento di Matematica - II Università di Roma - Tor Vergata
Via della Ricerca Scientifica - 00133 ROMA - Italy
E-mail: OLIVIERI@MAT.UTOVRM.IT

(2) Dipartimento di Matematica - Università di Roma Tre.
V. Segre 2 - 00146 ROMA - Italy
E-mail: SCOPPOLA@ROMA1.INFN.IT, SCOPPOLA@MAT.UNIROMA3.IT

Abstract

In this paper we review and discuss results on metastability. We consider ergodic aperiodic Markov chains with exponentially small transition probabilities and we give a complete description of the typical tube of trajectories during the first excursion outside a general domain Q.

Section 1. Introduction.

In this review we want to describe some recent results in the theory of metastability from the point of view of mathematical physics and probability theory.

Mathematically we will deal with the study of some large deviation problems for families of Markov chains with finite state space and transition probabilities decaying exponentially fast in a large external parameter β: the so called Freidlin-Wentzell regime.

From physical point of view we analyze a central problem arising in the dynamical description of phase transitions: the decay from metastable to stable equilibrium.

Phase transitions (like, for instance, liquid-vapour for a fluid or positive-negative magnetized phase for ferromagnetic systems) are, no doubt, the most important phenomena in statistical mechanics. As it is well known all equilibrium aspects are satisfactorily taken into account in the framework of the " Gibbs paradigm" (see, for instance [Ru], [Si], [Ge]). On the other hand a general theory of non-equilibrium statistical mechanics is still lacking and this is one of the main challenges for contemporary fundamental physics.

Certainly metastability and in particular its peculiar dynamical aspects, play a fundamental role in non-equilibrium statistical mechanics.

Let us rapidly introduce the subject.

Metastability is a relevant phenomenon for thermodynamic systems close to a first order phase transition.

Suppose to start from a given pure equilibrium phase in a suitable region of the phase diagram and change the thermodynamic parameters to values corresponding to a different equilibrium phase; then, in particular experimental situations, the system, instead of undergoing a phase transition, can still remain in a "wrong" equilibrium, far from the "true" one but actually close to what the equilibrium would be at the other side of the transition. This apparent equilibrium, often called "metastable state", persists untill an external perturbation or some spontaneous fluctuation leads the system to the stable equilibrium.

Examples of metastable states are supersaturated vapours and solutions, supercooled liquids and ferromagnets with magnetization opposite to the field. We will concentrate on this last example. In this case a metastable state can be obtatined starting from a ferromagnet in presence of a small negative external magnetic field $h < 0$ by suddenly switching this field to a small positive value: our system remains with negative magnetization (opposite to the actual value of h) for a time which is long if the external field is small.

We will not provide here a complete list of references on metastability; we want only to mention the paper by O. Penrose and J.L.Lebowitz [PL1], (see also [PL2]), where for the first time the theory of the metastable states has a rigorous formulation, and the paper by M.Cassandro, A.Galves, E.Olivieri and M.E.Vares [CGOV] where the so called "pathwise approach to metastability" has been introduced. This approach turns out to be the starting point of the recent developments on the theory of metastability.

We summarize, in what follows, the mathematical description of metastability, based on the pathwise approach, for the simple case of ferromagnetic spin systems on a lattice.

We will give a simple description of our systems by means of the standard Ising model. Consider a spin variable $\sigma(i)$, assuming values $+1$ or -1, associated to each site i in a box $\Lambda \subset \mathbf{Z}^d$.

To any configuration $\sigma \in \{-1,+1\}^\Lambda$ we associate the energy:

$$H_\Lambda(\sigma) = -\frac{1}{2} \sum_{i,j\in\Lambda, |i-j|=1} \sigma(i)\sigma(j) - \frac{h}{2} \sum_{i\in\Lambda} \sigma(i) \tag{1.1}$$

with $h > 0$ and we assume, for instance, periodic boundary conditions on the boundary of Λ.

The Gibbs measure associated to this Hamiltonian, describing the equilibrium properties of our system, is given by

$$\mu_\Lambda(\sigma) = \frac{e^{-\beta H_\Lambda(\sigma)}}{Z_\Lambda} \tag{1.2}$$

where Z_Λ is the normalization factor called partition function:

$$Z_\Lambda = \sum_{\sigma\in\{-1,+1\}^\Lambda} e^{-\beta H_\Lambda(\sigma)}$$

and $\beta > 0$ represents the reciprocal of the absolute temperature.

It is possible to show (see [Ru], [Si]) that, for $h > 0$ ($h < 0$), in the thermodynamic limit $\Lambda \to \mathbf{Z}^d$ we get , for every $\beta > 0$, a unique Gibbs state with positive (negative) magnetization.

We will be interested to the regime in which $h > 0$ sufficiently small is given, Λ sufficiently large but finite is given and β tends to infinity. In these conditions the Gibbs measure will be concentrated on the unique ground state for the energy H.

This configuration of minimal energy is $+\mathbf{1}$, in which $\sigma(i) = +1$ for all i. The configuration $-\mathbf{1}$ ($\sigma(i) = -1$ for all i) is only a local minimum for the Hamiltonian H_Λ (remember $h > 0$) and it is naturally related to metastability.

Let us now define $E(l) \equiv H_\Lambda(\sigma_{(l)}) - H_\Lambda(-\mathbf{1})$, where $\sigma_{(l)}$ is the configuration in which the plus spins are exactly the sites internal to a square of side l drawn in the dual lattice $\mathbf{Z}^d + (\frac{1}{2}, \frac{1}{2})$.

We have: $E(l) = 4l - hl^2$ which is maximum for $l = \frac{2}{h}$. This means that even though an arbitrarily small non-vanishing magnetic field determines the phase, its effects become relevant only on a sufficiently large scale ($l \geq l_c(h) \sim \frac{2}{h}$), since only on these large scales the volume energy dominates the surface energy.

A stochastic dynamics with discrete time is introduced in our system by means of a Markov chain on $S = \{-1, +1\}^\Lambda$ with transition probabilities $P(\sigma, \sigma')$ satisfying the detailed balance condition:

$$\mu_\Lambda(\sigma)P(\sigma, \sigma') = \mu_\Lambda(\sigma')P(\sigma', \sigma) \tag{1.3}$$

and the ergodicity condition

$$\forall \sigma, \eta \quad \exists t \in \mathbf{N} \quad \text{such that} \quad P^t(\sigma, \eta) > 0 \tag{1.4}$$

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

We will denote by σ_t the process at time t starting at σ at time 0.

By using the ergodic theorem for Markov chains it is easy to prove that the t-step transition probabilities $P^t(\sigma, \sigma')$ converge, as t goes to infinity, to the unique invariant measure of the chain that, due to the detailed balance condition (1.3), turns out to be the Gibbs measure $\mu_\Lambda(\sigma')$. This means that the mean value of each observable f with respect to the Gibbs measure μ_Λ, $\mu_\Lambda(f) = \sum_\sigma f(\sigma)\mu_\Lambda(\sigma)$, can approximately be evaluated as the mean of the observable on the process, $Ef(\sigma_t)$, at large t.

Markov chains satisfying conditions (1.3),(1.4) are actually used for Monte Carlo simulations.

An explicit construction of a Markov chain satisfying the previous hypotheses can be given, for instance, by the Glauber-Metropolis algorithm which is defined as follows: for any $\sigma \in S$ and any i in Λ let

$$\Delta_i H(\sigma) \ = \ H_\Lambda(\sigma^i) - H_\Lambda(\sigma) \tag{1.5}$$

with

$$\sigma^i(j) = \begin{cases} \sigma(j) & \text{if } i \neq j \\ -\sigma(j) & \text{if } i = j \end{cases} \tag{1.6}$$

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

$$P(\sigma,\sigma') = \begin{cases} 0 & \text{if} \quad \sigma' \neq \sigma^i \text{ for all } i \in \Lambda \\ \frac{1}{|\Lambda|} \exp\{-\beta(\Delta_i H(\sigma) \vee 0)\} & \text{if} \quad \sigma' = \sigma^i \text{ for some } i \end{cases} \tag{1.7}$$

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

We remark that the dynamics given by (1.7) is not conservative in the sense that the total magnetization is not a conserved quantity. Conservative dynamics (like e.g. Kawasaki dynamics) are not considered in the present paper; the study of metastability in the conservative case is more complicated and still open.

This kind of dynamics are often called stochastic Ising models. Suppose now to analyze the typical paths of our Markov chain defined by (1.7) starting from $-\mathbf{1}$ in the above described Freidlin-Wentzell regime (Λ large, h small fixed, β very large). For instance we can use a computer simulation and perform a large number of independent runs (see [ToMi]). What we see is that, during the evolution, in the sea of minus spins small droplets of pluses appear but shrinking and disappearing before they are able to become large.

Only after a very long time, under the effect of a large fluctuation, a cluster sufficiently large will appear in the sea. Its behavior is then completely different and it grows without hesitations.

In order to understand this behavior of typical trajectories, let us compare the shrinking and the growing probabilities for a connected cluster of spins $+1$ in the sea of -1's in the two dimensional case.

First of all each cluster of spins $+1$ becomes rectangular in a finite time, independent of β, with a probability of order one following a sequence of transitions with $\Delta_i H(\sigma) < 0$. Indeed the rectangle is the only shape of a cluster in which there are no spins with a number of nearest neighbour opposite spins larger than 2 and there are no spins -1 for which this number is equal to 2.

Starting from a rectangular cluster of $+1$'s, the cut of a corner has a cost $\Delta_i H(\sigma) = h$ and thus the lost of a row of the rectangle of length l has a cost $h(l-1)$ (the last $+1$ in the row is lost with cost 0); similarly for a column. On the other hand the cost of the addition of a new line to the rectangular cluster is $\Delta_i H(\sigma) = 2 - h$.

This means that if the minimal size l of the cluster is such that $h(l-1) > 2-h$, i.e. $l \geq l_c = [\frac{2}{h}] + 1$, the cluster tends to grow while if $l < l_c$ the cluster tends to shrink.

This heuristic argument has been developed in a rigorous way and metastability has been studied in detail for the 2-d standard stochastic Ising model by R.Schonmann and J.Neves in [NSch1], [NSch2] and [Sch1] (see also [Sch2], [Sch3]).

In [MOS1] metastability has been analyzed in the case of a cluster stochastic dynamics (Swendsen-Wang dynamics).

R.Kotecky and E.Olivieri in [KO1], [KO2] considered different hamiltonians. In particular they studied anisotropic 2-d Ising model and isotropic next nearest neighbour interactions 2-d Ising model and they proved that decay from the metastable to the stable situations typically follows a non-equilibrium path given by a sequence of "non-Wulff" configurations. Other 2-d lattice models have been studied in [CiO], [NaO]

Recently metastability in the three dimensional Ising model has been studied by G.Ben Arous and R.Cerf in [BACe].

We want also to mention the papers by R.Schonmann [Sch2], [Sch3] where the more difficult problem of the limit $h \to 0$ for fixed temperature is discussed.

Dynamics like the one we have described above (in a continuous time version) can also be condidered in the infinite volume situation. This gives rise to one of the most relevant examples of stochastic system of interacting particles (see [Li]). For these stochastic Ising models in infinite volume we have non-ergodicity of the process on the coexistence region: $h = 0, \beta > \beta_c$, ($\beta_c \equiv$ the reciprocal of the critical temperature). On the other hand in the complementary, uniqueness region: $h \neq 0$ and/or $\beta < \beta_c$, one expects good ergodic properties. Very intesting and deep results have been obtained in this field in particular by Holley (see, for instance, the celebrated papers [Ho1,2]). The methods that have been developed allow to handle the region far away from the coexistence line; near this coexistence line it turns out to be quite difficult to prove, for instance, exponential approach to equilibrium. Consider, in particular, the region $\beta > \beta_c, h$ small but non-vanishing. Nucleation phenomena and the behaviour of the system on a scale $l_c(h) = \frac{2}{|h|}$ (going to infinity as h tends to zero) are certainly to be taken into account. In other words the local relaxation from metastability to stability has implications in the global aproach to equilibrium. In [MOS2,3], by using results on the behaviour of the typical trajectories in a finite volume regime, the authors where able to prove strong exponential ergodicity for $\beta \gg \beta_c$ and $|h| \neq 0$ arbitrarily small in infinite volume. This paper represented a breakthrough in the understanding of the uniqueness regime that now is almost complete ([MO1], [MO2], [MOSch], [SchSh], [CesM])

Problems somehow related to metastability arise in many different areas of science like biology ([EGJLa]), sociology ([GlHu]), linguistics ([CG], [NiBe]), paleontology ([NewCoKi])...(see also [GSchm]).

In the present paper we want to collect and discuss results on metastability in the limit

$\beta \to \infty$ from a general point of view.

This means that we will consider ergodic aperiodic Markov chains X_t with finite state space S and with transition probabilities satisfying:

Property $\mathcal{P}$

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

$$\exp(-\Delta(x,y)\beta - \gamma\beta) \leq P(x,y) \leq \exp(-\Delta(x,y)\beta + \gamma\beta) \tag{1.8}$$

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

This kind of Markov chains appear in many models, not only related to physical problems.

As it has been shown by Freidlin and Wentzell [FW], this kind of Markov chains are also useful to study the asymptotic properties of diffusion processes describing small random perturbation of dynamical systems.

Similar Markov chains in a non-stationary situation are also considered in the theory of simulated annealing and optimization ([Ca1,2], [T2], [Ce]).

Let us now start the discussion of the problem of the first exit from a general domain $Q \subset S$.

We will denote by P_x the probability distribution of the process X_t starting from the state x at $t = 0$ and by E_x the corresponding expectation. Moreover, given any set of states $Q \subset S$, we will denote by τ_Q the first hitting time to Q:

$$\tau_Q \equiv \min\{t > 0 : \, X_t \in Q\} \tag{1.9}$$

For any set $Q \subset S$ we will denote by $Q^c \equiv S \backslash Q$ the complement of Q.

Many results on the first exit problem have been obtained by Freidlin and Wentzel [FW]. In particular they found estimates for the average exit time $E_x \tau_{Q^c}$ and the typical point, on the boundary of Q, reached during the first excursion outside Q, i.e. estimates on $P_x(X_{\tau_{Q^c}} = y)$. Our aim is to provide a complete description of the typical behavior of the Markov chain X_t up to the time τ_{Q^c}, for any set $Q \subset S$, and for sufficiently large β.

We are looking for the minimal set of exiting trajectories such that the sum of the corresponding probabilities is of order one as β goes to infinity.

A state x such that for any $y \neq x$ the quantity $\Delta(x, y)$ appearing in (1.8) is strictly positive is called a *stable state*. For instance in the case of the above defined Metropolis dynamics any strict local minimum for the energy function H is a stable state.

Consider the case of a domain Q completely attracted by a unique stable state $z_0 \in Q$; namely every "null cost" trajectory starting in Q ends in z_0 where the cost of a trajectory $x_0, x_1, ..., x_n$ is given by $\Delta(x_0, x_1) + \Delta(x_1, x_2) + ... + \Delta(x_{n-1}, x_n)$. This situation has been analyzed in detail by Freidlin and Wentzel (see [FW]). It turns out that the typical trajectories during the first exit from Q i.e. in the interval of time $[\theta_{\{z_0\}}, \tau_{Q^c}]$ where, for $D \subset Q$

$$\theta_D = \max\{t < \tau_{Q^c};\ X_t \in D\} \tag{1.10}$$

are suitable sequences "against the drift" emerging from z_0 and leading outside Q without hesitations in a finite time independent of β.

The case of a general Q containing several stable states is much more complicated. In this general case the last escape takes place by visiting a suitable sequence of stable states $z_1, ..., z_n$ and spending suitable random times inside certain domains $Q_1, ..., Q_n$ (*permanence sets*) containing $z_1, ..., z_n$ respectively. This means that in order to define our typical exiting tube we have to consider *resistance times* associated to stable states. These can be considered as a sort of *temporal entropy* that turns out to be necessary to give rise to an efficient escape mechanism. Neglecting these random times, during the escape, would lead to a mechanism extremely depressed in probability.

These permanence sets are union of suitable sets called *cycles* (see Definition 2.2 below). The characteristic property of a cycle A is that the process spends exponentially long time visiting all states in A before leaving it.

Cycles have been introduced by Freidlin and Wentzell, under some particular hypotheses, and generalization of this definition can be found in [OS2] and in [T1], [ChiCho1] where the cycle decomposition is used in the framework of simulated annealing.

If we consider reversible Markov chains, i.e. if we add to property $\mathcal{P}$ the hypothesis of detailed balance condition (1.3), the problem of the definition of cycles and of the determination of the typical tube is easier and can be reduced to the discussion of the energy landscape. In particular the sequence of sets defining the exiting tube can be obtained by solving a well defined sequence of variational problems (minimax problems) relative to the energy function H. It turns out that these variational problems constitute the only model-dependent work to be done to study the first exit from a general domain in the reversible case (see [OS1]).

In the general non reversible case, when only property $\mathcal{P}$ is verified, the cycles can be defined, following Freidlin and Wentzell, by means of a graphical technique.

As far as the problem of the determination of the typical exiting tube is concerned, one can adopt two different strategies.

Following Catoni and Cerf [CaCe] (see also [ChiCho2]), given a subset $D \subset Q$ and the cycle decomposition of the set $Q \setminus D$, one can associate to each path of cycles the probability that the process visits exactly this sequence of cycles in the time interval $[\theta_D, \tau_{Q^c}]$. The determination of the typical exiting tube can thus be reduced to the determination of the cycle paths of large probability.

A different approach to the problem is introduced in [OS2], where the analysis of the typical exiting paths is obtained by means of the renormalization procedure introduced in [S1].

In that paper the long time behavior of the chain X_t was studied by means of the construction of a sequence of *renormalized Markov chains* $X_t^{(1)}, X_t^{(2)}, \dots X_t^{(j)} \dots$ whose state spaces $S^{(1)}, S^{(2)}, \dots, S^{(j)} \dots$ were composed by stable states of increasing stability. These chains provide a rougher and rougher description of our stochastic evolution adapted to the analysis of phenomena taking place in increasing scales of time (exponential in β).

If $S \equiv S^{(0)}$ is the original state space then $S^{(1)}$ is just the set of stable states for the original Markov chain X_t; $X_t^{(1)}$ is suitably defined on $S^{(1)}$. $S^{(2)}$ is the set of stable states for $X_t^{(1)}$ and so on. In [S1] with this construction of renormalized chains, some results on the typical long time behavior of the original chain X_t were easily obtained. In fact to each exponentially long path of the chain X_t, a short path of a chain $X_t^{(N)}$ was associated, with a suitable renormalization index N.

Given $Q \subset S$ we define its boundary:

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

In order to simplify the description of the first exit from the domain Q of a Markov chain X_t we can consider a new auxiliary 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:

$$P^Q(x,y) = P(x,y) \quad \text{for any } x \in Q \tag{1.11}$$

$$P^Q(x,y) = e^{-\beta\Delta(\partial Q)} P(x,y) \quad \text{for } x \in \partial Q,\ y \neq x \tag{1.12}$$

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

To this chain is applied in [OS2] the renormalization procedure by introducing 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: in what follows we will only consider the chain P^Q but, to semplify notation, we will omit 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 renormalization iteration such that in Q there are only unstable states with respect to the corresponding chain $X_t^{(N)}$. More precisely let $N = N(Q)$ be the level such that the $(N+1)$-the renormalized Markov chain does not contain states inside Q:

$$N(Q) = \min\{n : \; S^{(n+1)} \cap Q = \emptyset\} \tag{1.13}$$

This means that the first excursion outside Q for the chain $X_t^{(N)}$ is a sort of "descent" along the drift.

A first rough approximation of the typical tube of escape from Q is thus given by the set of typical trajectories followed during the first excursion outside Q by the chain $X_t^{(N)}$.

It remains the question of "reading" the result in terms of the paths followed on the original scale of time by our original chain X_t.

This problem of analyzing the set of trajectories of the original chain X_t corresponding to a given trajectory of the chain $X_t^{(N)}$, is solved in [OS2] by associating to any state x^N of $S^{(N)}$ a suitable set $Q_{x^N} \subset S$, a sort of generalized cycle, representing the set where the original process X_t typically remains in the interval of time corresponding to a jump of the chain $X_t^{(N)}$.

By iterating this idea the exiting tube of paths can be easily determined.

In the remaining part of the present paper we will summarize the results proved in [OS1],[OS2] under some additional assumption that greatly simplify the exposition.

The paper is organized as follows. In section 2 we recall the construction of the exiting tube in the reversible case and we state the main results proved in [OS1]. In sections 3 and 4 we will consider non reversible chains but under a very strong additional hypothesis that allows us to give a simple presentation of the results on the first exit.

Section 2. The exit problem in the reversible case

In this section we will consider ergodic aperiodic Markov chains with transition probabilities satisfying the following:

Hypothesis M

There exists a function $H : S \to \mathbf{R}^+$ such that

$$P(x,y) \ = \ q(x,y) \exp(-\beta[H(y) - H(x)]_+) \tag{2.1}$$

where $q(x,y) \ = \ q(y,x)$ and $(a)_+$ is the positive part $(:= a \vee 0)$ of the real number a.

The above choice corresponds to a *Metropolis Markov chain* (see (1.7)) which is *reversible* in the sense of (1.3) with

$$\mu(x) \propto \exp(-\beta H(x)) \tag{2.2}$$

Since the symmetric matrix q is independent of β we immediately get, from (2.1), that Propery $\mathcal{P}$ is satisfied.

Hypothesis ND

$$H(x) \neq H(y) \quad \forall x \neq y \tag{2.3}$$

This hypothesis of *non-degeneracy* of the energy function $H(x)$ is introduced only to simplify the exposition. It is not at all needed and more general reversible Markov chains satisfying only the detailed balance condition (1.3) can be considered (see [OS1]) without any additional difficulty.

We observe that we can obtain the validity of (2.3) starting from any Metropolis Markov chain with an arbitrarily small change in the energy landscape.

Following the ideas presented in the introduction, we first define the cycles and we discuss their properties.

In order to construct the set of typical trajectories exiting from a cycle A (see Definition 2.2 below) we will consider in section 2.2 the first descent from any point y_0 in A to the "bottom" $F(A)$ of A. As already observed by R. Schonmann (see [Sch1]), the problem of the typical tube of trajectories during the first excursion outside a cycle A turns out to be simply related, via a time reversal transformation, to the typical tube followed by the process during the first descent to $F(A)$. Using this idea in section 2.3 we will define the typical tube of trajectories exiting from A.

2.1 - Cycles and their properties

A *path* ω is a sequence $\omega := x_1, \dots, x_N,\ N \in \mathbf{N}$, with $x_j, x_{j+1},\ j = 1, \dots, N-1$, communicating states (i.e. $P(x_j, x_{j+1}) > 0$). We often write $\omega : x \to y$ to denote a path joining x to y.

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 denote by M the set of "stable" states namely the local minima of the energy: $x \in M$ iff $H(x) \le \min_{y \in \partial\{x\}} H(y)$.

We say that a state x is *downhill connected* to a state y if there exists a *downhill path* $\omega = (x_0 = x, x_1, \dots, x_k = y)$ with $H(x_{i+1}) \le H(x_i),\ i = 0, , \dots, k-1$.

It is useful to conventionally extend every down-hill path up to a final stable state $y \in M$.

On the other side a state x is *uphill connected* to a state y if there exists a path $\omega = (x_0 = x, x_1, \dots, x_k = y)$ with $H(x_{i+1}) \ge H(x_i),\ i = 0, , \dots, k-1$.

Notice that the validity of Hypothesis ND does not prevent, in general, the existence of many different down-hill (or up-hill) paths emerging from the same point y.

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

$$F(Q) := \{y \in Q : \min_{x \in Q} H(x) = H(y)\} \tag{2.4}$$

and by $U = U(Q)$ the minimum set of the energy on its boundary:

$$U(Q) \equiv F(\partial Q) := \{z \in \partial Q : \min_{x \in \partial Q} H(x) = H(z)\} \tag{2.5}$$

We notice that by Hypothesis ND both $U(Q)$ and $F(Q)$ contain only one state.

Definition 2.1

For each pair of states $x, y \in S$ we define their minimal saddle point $\mathcal{S}(x, y)$ as follows: Let

$$\bar{H}_{x,y} := \min_{\omega : x \to y} \max_{z \in \omega} H(z).$$

Then $\mathcal{S}(x, y)$ is uniquely defined by

$$H(\mathcal{S}(x, y)) = \bar{H}_{x,y} \tag{2.6}$$

This definition is well-posed because of the validity of Hypothesis ND.

By definition we have that that $\mathcal{S}(x,y) = \mathcal{S}(y,x)\ \forall x, y \in S$. Notice that it may happen that $x = \mathcal{S}(y,x)$ or $y = \mathcal{S}(y,x)$.

Given $x \in S,\ Q \subset S,\ x \notin Q$,we define $\mathcal{S}(x,Q)$ by:

$$\min_{w \in Q} H(\mathcal{S}(x,w)) = H(\mathcal{S}(x,Q)) \tag{2.7}$$

Definition 2.2

A connected set A which satisfies :

$$\max_{x \in A} H(x) = \bar{H} < H(U(A))$$

is called cycle.

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

In the following we will recall the main properties of cycles (for proofs see [OS1]).

Proposition 2.1

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

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

Moreover, given two cycles A_1, A_2, either

1) $A_1 \cap A_2 = \emptyset$

or

2) $A_1 \subset A_2$ or, viceversa, $A_2 \subset A_1$

Proposition 2.2

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 $\mathcal{S}(x, A^c)$ does not depend on x and coincides with $U(A)$

Proposition 2.3
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 $\forall x \in A$

$$P_x(\exp(\beta\, [\, H(U(A)) - H(F(A)) \; -\varepsilon\,]) \; < \; \tau_{\partial A} \; < \; \exp(\beta\, [\, H(U(A)) - H(F(A)) \; +\varepsilon\,])\;) \; \geq$$

$$\geq \; 1 - e^{-k\beta}$$

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

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

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

$$P_x(\; X_{\tau_{\partial A}} = y) \geq \exp(-\beta[H(y) - H(U(A)) + \varepsilon]\;)$$

The property described by point ii) is usually called the *recurrence property* of cycles. It is the main property characterizing cycles and it means that with large probability each state in the cycle is visited by the process before the exit.

A consequence of this property is the asymptotically exponential distribution of the first exit time. More precisely let A be a given cycle. Given a point $x \in F(A)$ let the time $T_\beta = T_\beta(x)$ be defined by

$$P_x(\tau_{\partial A} > T_\beta(x)) \; = \; e^{-1}. \tag{2.8}$$

One has that T_β does not depend on $x \in A$, in the sense of logarithmic equivalence; namely we have $\forall x, y \in A$:

$$\lim_{\beta\to\infty} \frac{1}{\beta} \log\left[\frac{T_\beta(x)}{T_\beta(y)}\right] = 0 \qquad \lim_{\beta\to\infty} \frac{1}{\beta} \log[T_\beta(x)] \; = H(U(A)) - H(F(A)) \tag{2.9}$$

Moreover 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. Namely, $\forall x, y \in A$, $\varepsilon > 0$:

$$\lim_{\beta\to\infty} P_x(T_\beta(y)e^{-\varepsilon\beta} \; < \tau_{\partial A} \; < \; T_\beta(y)e^{+\varepsilon\beta}) \; = \; 1 \tag{2.10}$$

Proposition 2.4

Let $T^*_\beta = T_\beta(x^*)$ where x^* is a particular point in A chosen once for all. Then $\forall\ x\ \in A,\ \forall\ s\ \in\ \mathbf{R}^+$:

$$\lim_{\beta\to\infty} P_x\big(\frac{\tau_{\partial A}}{T^*_\beta} >\ s\ \big) =\ e^{-s}. \tag{2.11}$$

Moreover, 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:

$$\exp(\beta\ [H(U(A)) - H(F(A)) - \varepsilon]) < \mathbf{E}_x(\tau_{\partial A}) < \exp(\beta[H(U(A)) - H(F(A)) + \varepsilon] \tag{2.12}$$

2.2 - The permanence sets Q_i and the standard cascade

We want to construct now the set of typical trajectories starting from an arbitrary state $y_0 \in A$ and descending to $F(A)$. 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 stable state in ω_1. If such a point x_1 is $F(A)$ (y_0 belongs to the "basin of attraction" of $F(A)$) the *cascade* is simply given by $y_0, \omega_1, F(A)$

Let us now suppose that $x_1 \neq F(A)$. Let H_1 be the energy of the saddle between x_1 and $F(A)$:

$$H(\mathcal{S}(x_1, F(A))) = H_1 \tag{2.13}$$

and let $C(x_1, H_1)$ be the cycle containing x_1 with states of energy stricly less than H_1.

The first *permanence set* Q_1 is the maximal connected set, containing x_1, of points x such that

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

The permanence set Q_1 turns out to be the union of the cycle $C(x_1, H_1)$, the saddle $\mathcal{S}(x_1, F(A))$ and possibly other cycles contained in A and having $\mathcal{S}(x_1, F(A))$ as common minimal saddle in their boundaries. The above definition of permanence set can be used also in the case in which Hypothesis ND does not hold.

Let y_1 be a point in ∂Q_1 such that $\mathcal{S}(y_1, F(A)) < H_1$. We note that there is at least a point in ∂Q_1 with this property since, otherwise, by definition of Q_1, (2.13) does not hold.

We can iterate the procedure by defining a downhill path ω_2 starting from y_1 and lying outside Q_1. Denote by x_2 the final 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 and of the path ω_2 starting from y_1).
If we continue in this way and recursively construct a sequence of the form: $y_2, \omega_3, Q_3 y_3, \ldots$, we necessarily end up with a y_{N-1}, ω_N, for some finite N, with $x_N = F(A)$.

We set $Q_N := F(A)$.

Suppose first that a particular choice is made of $y_0, \omega_1, y_1, \omega_2, \ldots, y_{N-1}, \omega_N$, compatible with the above construction.

Then , automatically, $x_2, \ldots, x_{N-1}, x_N, Q_1, \ldots, Q_{N-1}, Q_N$ are given.

Let

$$\mathcal{T}(y_0, \omega_1, y_1, \omega_2, \ldots, y_{N-1}, \omega_N) \;=\; y_0 \cup \omega_1 \cup Q_1 \cup \omega_2, \ldots, Q_{N-1} \cup \omega_N \cup Q_N, \qquad (2.14)$$

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

Notice that, by construction:

$$H(y_{j-1}) \;>\; H(y_j), \quad H(F(Q_j)) \;\le\; H(y_{j-1}),\; j = 1, \ldots N$$

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

With the above definition of standard cascade we can state the result on typical paths for the first descent to the bottom of the cycle A.

Theorem 2.1

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

i)

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

ii)

$$\lim_{\beta\to\infty} P_{y_0}(\forall\, t \le \tau_{F(A)} \;:\; x_t \;\in\; \mathcal{T}(y_0, \omega_1, y_1, \omega_2, \ldots, y_{N-1}, \omega_N)$$
$$\text{for some } y_0, \omega_1, y_1, \omega_2, \ldots, y_{N-1}, \omega_N) \;=\; 1,$$

iii) Moreover, with probability $\to\; 1$ as $\beta \to \infty$, there exists a sequence $y_0, \omega_1, y_1, \omega_2, \ldots, y_{N-1}, \omega_N$ defining a standard cascade, 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_{N-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 which can be estimated from above and from below by $\exp(\beta[H(U(Q_j)) - H(F(Q_j)) \pm \varepsilon]), j = 1, \dots, N-1$.

Finally: before exiting from Q_j it can perform an arbitrary sequence of passages through the cycles belonging to Q_j with the typical behavior described in Proposition 2.3.

2.3 - By reversibility from the typical descent to the typical exit

Following Schonmann ([Sch1]), 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 from x, visiting y at some finite time t and never visiting x and y in between :

$$\Omega^*(x,y) := \{\omega = x_1, \dots, x_t \text{ for some } t \; : \; x_1 = x, \; x_t = y, \; ; \; x_2, \dots, x_{t-1} \neq x, y\} \qquad (2.15)$$

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

$$\forall \, \omega := (x_1, \dots, x_t) \; : \; R\omega := \bar{\omega} := (x_t, \dots, x_1) \qquad (2.16)$$

We naturally define, for every set of paths Δ,

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

Let us call τ_{xy} the last time our process visits the state x before touching, for the first time, y, namely :

$$\tau_{xy} \; := \; \max\{t < \tau_y : X_t = x\} \qquad (2.17)$$

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 ($x_1 := x$), 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 $\omega = x_1, \dots, x_t, \dots$ 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 [Sch1] it is proven that, for every $x, y\ \in\ S$, every $\Lambda\ \subset\ \Omega^*(x,y)$,

$$P_x(\{X_t\}_{t\in[\tau_{xy},\tau_y]}\ \in\ \Lambda) = \rho(\Lambda) = \bar{\rho}(R\Lambda) = P_y(\{X_t\}_{t\ \in[\tau_{yx},\tau_x]}\ \in\ R\Lambda) \tag{2.18}$$

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

Theorem 2.2
Let

$$\theta_{F(A)} = \max\{t < \tau_{A^c} : X_t\ = F(A).\}$$

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

$$P_{F(A)}(X_{\tau_{A^c}}\ = U(A)) \to\ 1\ as\quad \beta\ \to\ \ \infty \tag{2.19}$$

ii)

$$P_{F(A)}(\exists y_0 \in \bar{\partial}^- A, \mathcal{T}(y_0,\omega_1,y_1,\omega_2,\dots,y_{N-1},\omega_N)\ :$$

$$X_t \in R\mathcal{T}(y_0,\omega_1,y_1,\omega_2,\dots,y_{N-1},\omega_N) \forall t \in [\theta_{F(A)}, \tau_{A^c} - 1]) \to 1 as\ \beta \to \infty$$

iii)

During the first excursion from $F(A)$ to A^c, conditioning to $X_{\tau_{A^c}-1}\ =\ y$ (for some $y\ \in \bar{\partial}^- A$) and to follow a particular " anticascade" $R\mathcal{T}(y_0 = y,\omega_1,y_1,\omega_2,\dots,y_{N-1},\omega_N)$, with probability tending to one as $\beta\ \to\ \infty$ the process , between $\theta_{F(A)}$ and τ_y visits the sequence $Q_{N-1},...,Q_1$, spending in each set a time which can be estimated from above and from below by $\exp(\beta[H(U(Q_j)) - H(F(Q_j)) \pm \varepsilon]), j = 1,\dots,N-1$.

Section 3. The renormalization procedure and the cycle decomposition.

In the remaining part of this paper we will consider generic Markov chains (i.e. non necessarily reversible) but we will add to property $\mathcal{P}$ the following very particular hypothesis which greately simplifies the exposition of results and proofs. The reader can find in [OS2] the general results.

Hypothesis RI

There is at most a couple of states $x, y;\ x \neq y$ *such that* $\Delta(x, y) = 0$ *and moreover every non trivial linear combination with integer coefficients of the not vanishing quantities* $\Delta(x, y)$ *is different from zero.*

Remark

The rational independence condition in RI is even more strict than the analogous condition of non-degeneracy ND of the reversible case. The condition that for a unique pair of states x, y one has $\Delta(x, y) = 0$ is included in RI because it is conserved (and in any case produced) by the renormalization procedure.

Again we notice that with an arbitrarily small change of the $\Delta(x, y)$, any Markov chain satisfying property $\mathcal{P}$ can be modified to satisfy Hypothesis RI.

Given a chain X_t on S let $\Phi(S) \equiv \{\{\phi_i\}_{i\in\mathbf{N}} : \phi_i \in S\}$ be the set of paths. Following the theory of large deviations, developed in [FW], we define, for each $t \in \mathbf{N}$, a functional $I_{[0,t]}$ on $\Phi(S)$ associating to each path $\phi \in \Phi(S)$ the value

$$I_{[0,t]}(\phi) \equiv \sum_{i=0}^{t-1} \Delta(\phi_i, \phi_{i+1}) \tag{3.1}$$

where for $x \neq y,\ P(x, y) > 0$, the function $\Delta(x, y)$ has been defined in (1.8) and we set $\Delta(x, x) = 0$ for each $x \in S$ and $\Delta(x, y) = \infty$ if $P(x, y) = 0$. This functional is the *cost function* of each path ϕ in the sense of the following:

Lemma 3.1

Let ϕ *be a given path starting from* x *at time 0; then, for* $t \in \mathbf{N}$

i)

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

where γ *is the quantity introduced in (1.8)*

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

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

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

$$\sup_x P_x(I_{[0,t]}(X_\cdot) \geq I_0) \leq e^{-I_0\beta + \epsilon\beta}$$

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

For each pair of states x, y we define

$$V(x,y) \equiv \inf_{t,\phi:\, \phi_0 = x,\, \phi_t = y} I_{[0,t]}(\phi) \tag{3.2}$$

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

$$\text{for any} \quad y \neq x \quad V(x,y) > 0 \tag{3.3}$$

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

We note that by Hypothesis RI we have that there is at most an unstable state, i.e.$|S \backslash M| \leq 1$. Moreover again Hypothesis RI allows us to exclude the existence of equivalent states, i.e. of states x, y such that $V(x,y) = V(y,x) = 0$, and thus we do not have to discuss the problem in terms of equivalence classes as it has been done in [OS2].

3.1 - Construction of the renormalized chain

It is immediate to prove, using lemma 3.1, that the process spends, with large probability, relatively very short times in the unstable state. This 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.

Indeed we can consider the less stable states in M and we can define a time scale t_1 corresponding to this smallest stability:

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

where

$$V_1 \equiv \min_{x\in M,\, y\in S,\, x\neq y} V(x,y) \tag{3.5}$$

and $\delta = \delta(\beta) := 2\gamma(\beta)|S|$ (see (1.8)), goes to zero as β tends 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:

$$\zeta_0 \equiv \min\{t \geq 0 :\ X_t \in M\}$$

and for each $n \geq 1$:

$$\sigma_n \equiv \min\{t > \zeta_{n-1} :\ X_t \neq X_{\zeta_{n-1}}\}$$

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

$$\zeta_n = \begin{cases} \zeta_{n-1} + t_1 & \text{if } \sigma_n - \zeta_{n-1} > t_1 \\ \tau_n & \text{if } \sigma_n - \zeta_{n-1} \leq t_1 \end{cases} \tag{3.6}$$

It is easy to see that the sequence $X_n^{(1)} = X_{\zeta_n}$ is a homogeneous Markov chain on the state space $S^{(1)} = M$. For any pair of states $x, y \in M$ we denote by $P^{(1)}(x,y)$ the transition probabilities of the chain $X_n^{(1)}$; it is possible to prove that these transition probabilities satisfy the same assumptions (property $\mathcal{P}$ and Hypothesis RI) verified by the original chain X_t.

More precisely for any $x, y \in M, \quad x \neq y$:

$$\exp\{-\Delta^{(1)}(x,y)\beta - \gamma'\beta\} \leq P^{(1)}(x,y) \equiv P(X_{\zeta_n} = y |\, X_{\zeta_{n-1}} = x) \leq \exp\{-\Delta^{(1)}(x,y)\beta + \gamma'\beta\} \tag{3.7}$$

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

$$\Delta^{(1)}(x,y) = \bar{\Delta}(x,y) - V_1 \tag{3.8}$$

where $\bar{\Delta}(x,y) = [\Delta(x,y) \wedge (\Delta(x,\hat{y}^0) + \Delta(\hat{y}^0,y))]$ if $\hat{y}^0 := S\backslash S^{(1)} \neq \emptyset$ and $\bar{\Delta}(x,y) = \Delta(x,y)$ if $\hat{y}^0 = \emptyset$, and $\gamma' \to 0$ as $\beta \to \infty$ (see [S1]).

Hypothesis RI is also satisfied by the chain $X_t^{(1)}$ since the new quantities $\Delta^{(1)}(x,y)$ turn out to be linear combination of the old quantities $\Delta(.,.)$ with integer coefficients and there exists at most one unstable state x for the new chain $X_t^{(1)}$.

Thus we have a new chain $X_t^{(1)}$ on the state space $S^{(1)}$, 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.

This procedure of renormalization can be accomplished for more general Markov chains satisfying only property $\mathcal{P}$. In our case, under Hypothesis RI, at most one unstable state can be removed going from $S^{(n)}$ to $S^{(n+1)}$, i.e. $|S^{(n)}\backslash S^{(n+1)}| \leq 1$. It is also easy to prove that $|S^{(n+2)}| < |S^{(n)}|$.

We will denote by $\hat{y}^n$ the unstable state of $S^{(n)}$, if any, i.e. $\hat{y}^n = S^{(n)}\backslash S^{(n+1)}$. Thus we have: $S^{(n)} = \cup_{i\geq n} \hat{y}^i$.

We recall now here the iteration scheme introduced in [S1] (see also [S2]).

Notation: The superscript $^{(k)}$ will denote the various quantities referring to the k-th chain $X_t^{(k)}$; e.g. $\tau_Q^{(k)} = \min\{t:\ X_t^{(k)} \in Q\}, \quad Q \subset S^{(k)}$.

For any $k \geq 1$ we define the following quantities:
for any $\phi : \mathbf{N} \to S^{(k)}$:

$$I_{[o,t]}^{(k)}(\phi) = \sum_{i=0}^{t-1} \Delta^{(k)}(\phi_i, \phi_{i+1}) \tag{3.9}$$

$$V^{(k)}(x,y) \equiv \min_{t,\phi:\ \phi_0=x,\ \phi_t=y} I_{[o,t]}^{(k)}(\phi) \quad \forall x,y \in S^{(k)} \tag{3.10}$$

$$M^{(k)} = \{x \in S^{(k)} : \forall y \in S^{(k)},\ y \neq x \quad V^{(k)}(x,y) > 0\} \tag{3.11}$$

$$V_{k+1} = \min_{x\in M^{(k)}, y \in S^{(k)}\ x\neq y} V^{(k)}(x,y) \tag{3.12}$$

$$t_{k+1} = e^{V_{k+1}\beta + \delta\beta} \tag{3.13}$$

$$T_1 = t_1$$

$$T_{k+1} = t_1 t_2 ......t_k t_{k+1} \tag{3.14}$$

$$S^{(k+1)} = M^{(k)} \tag{3.15}$$

$$\Delta^{(k+1)}(x,y) = \bar{\Delta}^{(k)}(x,y) - V_{k+1} \quad \forall x,y \in S^{(k+1)} \tag{3.16}$$

where $\bar{\Delta}^{(k)}(x,y) = [\Delta^{(k)}(x,y) \wedge (\Delta^{(k)}(x,\hat{y}^k) + \Delta^{(k)}(\hat{y}^k,y))]$ if $\hat{y}^k \neq \emptyset$ and $\bar{\Delta}^{(k)}(x,y) = \Delta^{(k)}(x,y)$ if $\hat{y}^k = \emptyset$.

The main results proved in [S1] can be summarized as follows:

Theorem 3.1

Let $W \subset S^{(1)}$ and $B \subset W$. Then for any sufficiently large β and for any $x \in S^{(1)} \backslash W$, $y \in W$ there is a positive γ' depending on γ and tending to zero as $\beta \to \infty$ such that:

$$\exp\{-\gamma'\beta\} P_x^{(1)}(X_{\tau_W}^{(1)} = y) \leq P_x(X_{\tau_W} = y) \leq \exp\{\gamma'\beta\} P_x^{(1)}(X_{\tau_W}^{(1)} = y)$$

$$\exp\{-\gamma'\beta\} t_1 E_x^{(1)} \tau_W^{(1)} \leq E_x \tau_W \leq \exp\{\gamma'\beta\} t_1 E_x^{(1)} \tau_W^{(1)}$$

$$e^{-\gamma'\beta} \mu^{(1)}(B) \leq \mu(B) \leq e^{\gamma'\beta} \mu^{(1)}(B)$$

for any $B \subset S^{(1)}$ where μ and $\mu^{(1)}$ are the invariant measures of the chains X_t and $X_t^{(1)}$ respectively. Moreover for any $A \in S \backslash M$

$$\mu(A) \leq e^{-V_1\beta + \gamma'\beta}$$

3.2 - Tube of paths corresponding to a renormalized path

By using the previous construction of renormalized chains we can immediately define a correspondence between trajectories of chains at different levels of the iteration.

More precisely, for any integer n, to any path ϕ of the Markov chain X_t, $\phi \in \Phi(S)$, we can associate a path $\phi^{(n)}$ of the Markov chain $X_t^{(n)}$, $\phi^{(n)} \in \Phi(S^{(n)})$, which is in a sense a projection of the path ϕ in the smaller space $\Phi(S^{(n)})$. On the other hand to any path $\phi^{(n)} \in \Phi(S^{(n)})$ we want to associate a set, say a tube, of paths $\phi \in \Phi(S)$ having projection $\phi^{(n)}$.

Definition 3.1

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

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

On the other side, to any given sequence of states $\psi^{(n)}$ in $S^{(n)}$ of length $T:\ \psi^{(n)}_i$, $i \le T$ we can associate a tube of trajectories in $\Phi(S)$ as

$$\mathcal{T}(\psi^{(n)}) = \{\phi \in \Phi(S):\ \phi^{(n)}_i = \psi^{(n)}_i \quad \forall i \le T\} \tag{3.17}$$

By construction this application between trajectories in $\Phi(S)$ and trajectories in $\Phi(S^{(n)})$ is such that

$$P^{(n)}(X^{(n)}_s = \phi^{(n)}_s,\, \forall s \le T) \asymp P(X_t \in \mathcal{T}(\psi^{(n)})) \tag{3.18}$$

where we say that two quantities A and B are *logarithmically equivalent* and we write $A \asymp B$ if they have the same exponential asymptotic behavior in β namely

$$A \asymp B \quad \text{if and only if} \quad \lim_{\beta\to\infty} \frac{1}{\beta}\log A = \lim_{\beta\to\infty}\frac{1}{\beta}\log B \tag{3.19}$$

With this application we can also define a sequence of random times Z^n_k corresponding, on the original time scale, to the times k on the time scale of the chain $X^{(n)}$. More precisely, given a path $\phi \in \Phi(S)$, we have defined a sequence of times $\zeta_0, ..., \zeta_k, ...$ and a path $\phi^{(1)} \in \Phi(S^{(1)})$ associated to it, and, iterating, sequences of times $\zeta^{(i)}_0, ..., \zeta^{(i)}_k, ...,$ and paths $\phi^{(i+1)} \in \Phi(S^{(i+1)})$ for each $i = 0, 1, ...$

Definition 3.2

$$Z^n_k \equiv \zeta_{\zeta^{(1)}_{\zeta^{(2)}_{\ddots_{\zeta^{(n-1)}_k}}}} \qquad n = 1, 2, ... \tag{3.20}$$

This is a random time with respect to the process X_t, corresponding to the time k for the chain $X^{(n)}$.

These times Z^n_k with large probability are of order T_n in the following sense:

Proposition 3.1

For any $N \ge 1$ and for any sufficiently small positive constant α there exists a positive constant $k = k(\alpha, N)$ such that for any $x \in S^{(N)}$

$$P_x(T_N e^{-\alpha\beta} \le Z_1^N \le T_N e^{\alpha\beta}) \ge 1 - e^{-k\beta}$$

for any sufficiently large β.

Similar results are obtained in [OS2] in the more general case in which Hypothesis RI is not assumed.

In order to complete the characterization of the typical tube of paths in $\Phi(S)$ corresponding to a given path $\phi^{(n)}$, we need to define the cycles. In fact it is possible to associate to each state $x^n \in S^{(n)}$ a cycle Q_{x^n} for the original chain X_t representing the region of the state space S corresponding to the state $x^n \in S^{(n)} \subset S$ under the time rescaling T_n.

3.3 - Cycles associated to renormalized states

Let us first recall some definition introduced by Freidlin and Wentzell, in order to discuss the cycle decomposition.

Definition 3.3 (B-graphs) : for any set of states $B \subset S$, a B-graph is a graph consisting of arrows $m \to n$ with $m \in B$ and $n \in S$, $m \neq n$ and satisfying the following properties:

1) every state $m \in B$ 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 $m \in B$ there exists a sequence of arrows leading from it to some point $n \in B^c$.

In other words a graph is a forest of trees with roots in B^c and with branches given by arrows directed to the root (i.e. the set B^c is the target of the sequences of arrows).

The set of B-graphs is denoted by $G(B)$; for any graph $g \in G(B)$ we define $\pi(g) = \prod_{m\to n \in g} P(m,n)$.

Warning:

Our notation is opposite with respect to that used in [FW], where a W-graph was a graph with target W , i.e. a W^c-graph in our notation.

By using this graphic formalism Freidlin and Wentzel in [FW] prove the following lemmas.

Lemma 3.2 (FW)
The invariant measure of the Markov chain X_t : $\mu(i)$, $i \in S$ is given by :

$$\mu(i) = \frac{q_i}{\sum_{j \in S} q_j}$$

where

$$q_i = \sum_{g \in G\{S \setminus i\}} \pi(g) \tag{3.21}$$

Lemma 3.3 (FW)
For any $B \subset S$ let τ_B be the first hitting time to B, then for any $j \in B$

$$P_i(X_{\tau_B} = j) = \frac{\sum_{g \in G_{ij}(B^c)} \pi(g)}{\sum_{g \in G(B^c)} \pi(g)} \tag{3.22}$$

where for any $i \in B^c$ and $j \in B$ we denote by $G_{ij}(B^c)$ the set of B^c-graphs in which the sequence of arrows leading from i into B ends at the point j (i.e. i belongs to the tree with root j).

Lemma 3.4(FW)

$$E_i \tau_B = \frac{\sum_{g \in G(B^c \setminus \{i\})} \pi(g) + \sum_{j \in B^c,\, j \neq i} \sum_{g \in G_{ij}(B \setminus \{j\})} \pi(g)}{\sum_{g \in G(B^c)} \pi(g)} =$$

$$= \frac{\sum_{g \in G(i \not\to B)} \pi(g)}{\sum_{g \in G(B^c)} \pi(g)} \tag{3.23}$$

where $G(i \not\to B)$ is the set of graphs (without closed loops) containing $|B^c| - 1$ arrows $m \to n$ each one emerging from a different point $m \in B^c$ and with $n \in S$, $m \neq n$ and not containing chains of arrows leading from i into B.

Lemma 3.2 is easily proved by showing that the quantities q_i satisfy the stationarity equation, Lemmas 3.3 and 3.4 can be proved by induction over the number of states contained in B. (See [FW], pg.179,182).

By using property $\mathcal{P}$ these results can be reformulated as follows (see [FW]). Consider the problem of the first exit of our chain X_t from a domain $Q \subset S$ and let τ_{Q^c} be the first exit time from the domain Q.

Proposition 3.2 (FW)
For any $\delta > 0$ there exists $\beta_o > 0$ such that for any $\beta > \beta_0$:

$$e^{-\beta(W_Q(x,y)-W_Q+\delta)} \leq P_x(X_{\tau_{Q^c}} = y) \leq e^{-\beta(W_Q(x,y)-W_Q-\delta)} \tag{3.24}$$

for any $x \in Q$ and $y \in Q^c$, with

$$W_Q(x,y) = \min_{g \in G_{xy}(Q)} W(g) \tag{3.25}$$

$$W_Q = \min_{g \in G(Q)} W(g) \tag{3.26}$$

$$W(g) \equiv \sum_{m \to n \in g} \Delta(m,n) \tag{3.27}$$

Moreover for any $x \in Q$:

$$\lim_{\beta \to \infty} \frac{1}{\beta} \log E_x \tau_{Q^c} = W_Q - M_Q(x) \tag{3.28}$$

where

$$M_Q(x) = \min_{g \in G(x \not\to Q^c)} W(g) \tag{3.29}$$

For each $Q \subset S$ let $g^*_Q \in G\,(Q)$ be a Q-graph minimizing the quantity W_Q defined in (3.26).

We note that by Hypothesis RI for any set $Q \subset S$ there exists a unique graph g^*_Q. Indeed if there were two different minimizing graphs $(g^*_Q)_1$ and $(g^*_Q)_2$, denoting by $(g^*_Q)'_1$ the arrows contained in $(g^*_Q)_1$ not contained in $(g^*_Q)_2$ (and the analog for $(g^*_Q)'_2$) it would be

$$0 = W((g^*_Q)_1) - W(g^*_Q)_2)) = W((g^*_Q)'_1) - W((g^*_Q)'_2)$$

against Hypothesis RI.

For any set Q and $x \in Q$ let

$$R_Q(x) = \{y \in Q^c : \exists x \to ... \to y \in g^*_Q\}$$

Definition of cycles of rank zero

Single states in S are called cycles of rank zero or 0-cycles.

Definition of cycles of rank one

Consider a set $C = \{x_1, ..., x_n\}$ contained in S with $n \geq 2$ such that $x_{i+1} = R_{\{x_i\}}(x_i)$ for each $i = 1, ..., n-1$ and $x_1 = R_{\{x_n\}}(x_n)$. Such a set is called 1-cycle or cycle of rank one.

Every single state not belonging to one of these 1-cycles constitutes by itself a 1-cycle.

Thus we can decompose the state space in 1-cycles: $S = \cup_i C_i$.

If C is a 1-cycle we have that

$$|R_C(x)| = 1 \quad \text{and} \quad R_C(x) = R_C(y) = R_C \quad \forall x, y \in C \tag{3.30}$$

This property is immediately verified under our Hypothesis RI. This means that graphs minimizing the exit from a cycle are made by a unique tree.

We can then iterate our construction by defining:

Definition of cycles of rank k

Consider a set $C = \{C_1, ..., C_n\}$ contained in S where C_i are $k-1$ cycles and $i = 1..., n$; $n \geq 2$ such that $R_{C_i} \in C_{i+1}$ for each $i = 1, ..., n-1$ and $R_{C_n} \in C_1$. Such a set is called k-cycle or cycle of rank k.

Every $(k-1)$-cycle not belonging to one of these k-cycles constitutes by itself a k-cycle.

Thus we can decompose the state space in k-cycles.

If C is a k-cycle we can iteratively verify that

$$|R_C(x)| = 1 \quad \text{and} \quad R_C(x) = R_C(y) = R_C \quad \forall x, y \in C \tag{3.31}$$

The main properties of cycles are summarized in the following.

Proposition 3.3

For any set C with $|C| > 1$, the following are equivalent:

i) C is a cycle (of some rank)

ii) there exists $K > 0$ such that for every $x \in C$, $B \subset C$, $x \notin B$ and β sufficiently large we have:

$$P_x(X_{\tau_{C^c \cup B}} \notin B) < e^{-K\beta} \tag{3.32}$$

iii) $E_x \tau_{C^c}$ is independent of $x \in C$ in the sense of logarithmic equivalence and for any $B \subset C$ and for any $x \in B$ and $z \in C$:

$$\lim_{\beta\to\infty} \frac{1}{\beta} \log E_x \tau_{B^c} < \lim_{\beta\to\infty} \frac{1}{\beta} \log E_z \tau_{C^c} \tag{3.33}$$

As in the reversible case, a consequence of the recurrence property of cycles (3.32) is the exponential distribution of their first exit times.

Let us now define the cycle associated to a renormalized state. For any $x^n \in S^{(n)}$ let Q_{x^n} be the maximum cycle such that

$$Q_{x^n} \cap S^{(n-1)} = x^n$$

We will show that this cycle is the permanence set associated to x^n i.e. the set of states visited by the chain X_t in the interval of time corresponding to a transition $x^n \to y^n$ of the chain $X_t^{(n)}$. More precisely let Z_t^{n-1} be defined by (3.20), let

$$\sigma^{(n-1)} := \tau^{(n-1)}_{S^{(n-1)} \setminus x^n}$$

and for any x^n, $y^n \in S^{(n)}$ let $\phi^{(n-1)}(x^n, y^n)$ be the path of states in $S^{(n-1)}$ going from x^n to y^n without visiting other states in $S^{(n)}$ and minimizing the quantity $\bar{\Delta}^{(n-1)}(x^n, y^n)$.

By our hypothesis RI we know that this path is given by:

$$\phi^{(n-1)}(x^n, y^n) = \begin{cases} x^n, y^n & \text{if } \bar{\Delta}^{(n-1)}(x^n, y^n) = \Delta^{(n-1)}(x^n, y^n) \\ x^n, \hat{y}^{n-1}, y^n & \text{otherwise} \end{cases} \tag{3.34}$$

Theorem 3.2

For any sufficiently small $a > 0$ (see proposition 3.1), we define:

$$A_1 := \{Z^{n-1}_{\sigma^{(n-1)}-1} \in [T_n e^{-a\beta}, T_n e^{a\beta}]\}$$

$$A_2 := \{Z^{n-1}_{\sigma(n-1)-1} < \tau_{(Q_{x^n})^c} \leq Z^{n-1}_{\sigma(n-1)}\}$$

$$A_3 := \{X_t \in \mathcal{T}(\phi^{(n-1)}(x^n, y^n)) \quad \forall t > Z^{n-1}_{\sigma(n-1)-1}\}$$

$$G := \{X_0^{(n)} = x^n, \quad X_1^{(n)} = y^n\}$$

There exists a positive constant K, depending on a but independent of β, such that for any sufficiently large β we have :

$$P_{x^n}(A_1 \cap A_2 \cap A_3 | G) \geq 1 - e^{-K\beta}$$

Moreover if $x^n = \hat{y}^n$ and if we define

$$A_4 := \{\forall y \in Q_{x^n} \quad \exists t < Z^{n-1}_{\sigma(n-1)-1} \text{ such that } X_t = y\}$$

then we have, for any sufficiently large β

$$P_{x^n}(A_1 \cap A_2 \cap A_3 \cap A_4 | G) \geq 1 - e^{-K\beta}$$

Sketch of the proof

It is sufficient to show that $P_{x^n}(A_i|G)$ are exponentially small for $i = 1, ..., 4$.

Suppose first that $P_{x^n}(G) \asymp 1$.

The estimate of $P_{x^n}(A_1|G)$ follows from Proposition 3.1.

To estimate $P_{x^n}(A_2|G)$ we first note that, since $Q_{x^n} \cap S^{(n-1)} = x^n$, we have $\tau_{(Q_{x^n})^c} \leq Z^{n-1}_{\sigma(n-1)}$. On the other hand

$$P_{x^n}(\tau_{Q^c_{x^n}} < Z^{n-1}_{\sigma(n-1)-1}) \leq P_{x^n}(X_{\tau_{Q^c_{x^n}}} \notin R_{Q_{x^n}}) + P_{R_{Q_{x^n}}}(X_{\tau_{S^{(n-1)}}} = x^n) \qquad (3.35)$$

The first term in the r.h.s. of (3.35) is estimated by proposition 3.2. The second one, can be estimated by considering the minimum cycle containing $R_{Q_{x^n}}$ with non empty intersection with $S^{(n-1)}$. This cycle does not contain x^n, by the maximality of Q_{x^n}, and thus by proposition 3.3 also this second term is exponentially small.

The estimate of $P_{x^n}(A_3|G)$ can be obtained by using (3.18) and by applying lemma 3.1 to the chain $X_t^{(n-1)}$; the last probabilty $P_{x^n}(A_4|G)$ is estimated by proposition 3.3.

The case $P_{x^n}(G)$ exponentially small is more interesting and more difficult to treat. The reader can find a detailed proof in [OS2].

Section 4. The exit problem in the general case

Let us come now to the problem of the determination of the tube of exit from an arbitrary set $Q \subset S$. We will use the renormalization procedure and the results of the previous section to define the tube of typical exiting paths in terms of a sequence of permanence sets.

As discussed in the introduction, we can apply the renormalization procedure to the auxiliary chain X_t^Q with almost absorbing states in ∂Q (see (1.11),(1.12)) up to the step $N = N(Q)$ of the iteration, such that in Q there is only the unstable state $\hat{y}^N$.

Due to our hypothesis RI there exists a unique transition $\hat{y}^N, z^N$ for the chain $X^{(N)}$ exiting from $\hat{y}^N$ with $\Delta^{(N)}(\hat{y}^N, z^N) = 0$.

By lemma 3.1 we have the following inequality:

$$P_{\hat{y}^N}(X_1^{(N)} = z^N) > 1 - e^{-k\beta} \tag{4.1}$$

This transition $\hat{y}^N, z^N$ can be analyzed in terms of paths of the chain $X^{(N-1)}$.

We can proceed as follows: by applying Theorem 3.2 we can associate to this transition a cycle $Q_{\hat{y}^N}$ and the path $\phi^{(N-1)}(\hat{y}^N, z^N)$ in $\Phi(S^{(N-1)})$ defined by (3.34)

To each transition of this path at level $N-1$ we can associate, again by Theorem 3.2, a cycle $Q_{\phi_j^{N-1}}$ and a new path.

By iterating this argument we have that the transition $\hat{y}^N, z^N$ is described by a sequences of cycles:

$$Q_{\hat{y}^{n_0}},\, Q_{\hat{y}^{n_1}}, ..., Q_{\hat{y}^{n_L}}.$$

Let us analyze the sequence $\hat{y}^{n_0}, ..., \hat{y}^{n_L}$ arising from the above iteration.

We start with the sequence $\hat{y}^N, z^N$. At the first step we consider transitions at level $N-1$ by substituting to the previous sequence the sequence $\phi^{(N-1)}(\hat{y}^N, z^N)$. At the k-th step of the iteration we start by a sequence of transitions of the chain at level $N-k+1$, $\mathbf{\hat{y}^{N-k+1}} = \{\hat{y}^{n_i}\}_i$ with $n_i \geq N-k+1$ and we substitute to each couple of consecutive states $\hat{y}^{n_i}$, $\hat{y}^{n_{i+1}}$ the sequence of transitions between states belonging to $S^{(N-k)}$ given by $\phi^{(N-k)}(\hat{y}^{n_i}, \hat{y}^{n_{i+1}})$. We obtain in this way a new sequence $\hat{\mathbf{y}}^{(\mathbf{N-k})}$ that we call the *refinement* of $\hat{\mathbf{y}}^{\mathbf{N-k+1}}$. This sequence corresponds to the "typical exit path" for the chain $X_t^{(N-k)}$.

The final sequence $\hat{\mathbf{y}}^{\mathbf{0}} = \hat{y}^{n_0}, ..., \hat{y}^{n_L}$, with $n_0 = N$, arising from the above iteration will be the refinement of the refinement.....of the refinement of $\hat{y}^N, z^N$.

In order to characterize the typical exit path in the time interval $(\theta_{\hat{y}^N}, \tau_{Q^c}]$ we cancel the first term in $\hat{\mathbf{y}}^{\mathbf{0}}$ corresponding to the first interval of time $[0, \theta_{\hat{y}^N}]$, and we obtain:

$$\hat{\mathbf{y}} := \hat{y}^{n_1}, ..., \hat{y}^{n_L}$$

To this path we can associate the corresponding sequence of permanence sets:

$$\mathcal{T}_{\hat{\mathbf{y}}} := Q_{\hat{y}^{n_1}}, Q_{\hat{y}^{n_1}}, ..., Q_{\hat{y}^{n_L}}$$

The result on the typical exiting tube can thus be stated as follows:

Theorem 4.1

For each positive a there exists a positive constant K such that for each sufficiently large β we have:

$$P_{\hat{y}^N}(\{X_t\}_{t \in (\theta_{\hat{y}^N}, \tau_{Q^c}]} \in \mathcal{T}_{\hat{\mathbf{y}}}) \geq 1 - e^{-K\beta}$$

Moreover with probability going to one as $\beta \to \infty$ the process visits sequentially the sequence of sets given by $\mathcal{T}_{\hat{\mathbf{y}}}$ spending inside each cycle $Q_{\hat{y}^{n_i}}$ a time which can be estimated from above and from below by $T_{n_i} e^{\pm a\beta}$ and visiting all the states inside the cycle before the exit from it.

The proof of this theorem can be obtained by using Theorem 3.2.

Let us now make some remarks on theorem 4.1.

Given the sequence $\hat{\mathbf{y}}$ consider each element $\hat{y}^{n_j}$ belonging to a portion of the sequence where n_j is monotonocally decreasing, i.e. such that $n_j < n_{j-1}$ and $Q_{\hat{y}^{n_j}} \subset Q_{\hat{y}^{n_{j-1}}}$. This means that the information given by the event:

$$\tau_{Q^c_{\hat{y}^{n_j}}} \asymp T_{n_j} \tag{4.2}$$

can be considered as a "negligible correction" to the statement:

$$\tau_{Q^c_{\hat{y}^{n_{j-1}}}} \asymp T_{n_{j-1}}$$

Then from any sequence $\hat{\mathbf{y}}$ we can extract a new sequence $\hat{\mathbf{y}}'$ in which this kind of monotonically decreasing subsequences have been removed.

More precisely:

Definition 4.1

Let $\{\hat{\mathbf{y}}''\}_i$ be a maximal segment of $\hat{\mathbf{y}}$

$$\{\hat{\mathbf{y}}''\}_i := \{\hat{y}^{n_{l_i}},\, \hat{y}^{n_{l_i+1}},\, ...,\, \hat{y}^{n_{m_i}}\}$$

such that for any $j \in [l_i, m_i]$ we have $Q_{\hat{y}^{n_j}} \subseteq Q_{\hat{y}^{n_{k_i}}}$ for some $k_i \in [l_i, m_i]$.

Let $\hat{\mathbf{y}}'$ be the sequence obtained by $\hat{\mathbf{y}}$ by replacing each segment $\{\hat{\mathbf{y}}''\}_i$ with the unique term $\hat{y}^{n_{k_i}}$.

This means that by considering the sequence $\hat{\mathbf{y}}'$ instead of $\hat{\mathbf{y}}$, we have removed all "neglegible corrections" like (4.2).

With theorem 4.1 we have described in full detail the exit; with the refinement procedure we have obtained at the same time two different kinds of information. First of all, we have obtained what we call *primary information* (or *level one approximation of the tube of exit*): the family of ordered sequences of different permanence sets visited by the process (namely the set of standard descents $Q_{\hat{\mathbf{y}}'}$) and the typical times spent by our process inside them. The *secondary information* (or level 2 approximation of the tube of exit) concerns further details about the history of the process inside each cycle that it visits when it stays inside some Q_i in $Q_{\mathbf{y}'}$ corresponding to the segments $\{\hat{\mathbf{y}}''\}_i$. It will describe the first descent to the bottom of the cycles and the first excursion outside them.

Up to now we have considered the starting point $\hat{y}^N$. As a final remark we want to note that a general starting point can be considered by distinguishing a first interval of time in which the process falls to the bottom $\hat{y}^N$, and a second, much longer interval of time in which the process exits from Q following the description obtained above, since we can now consider $\hat{y}^N$ as a starting point.

The first transient consisting in the descent to the bottom of Q can be analyzed with the same procedure, by considering refining paths.(See [OS2] for more details).

Aknowledgements

This work has been partially supported by the grant CHRX-CT93-0411 of the Commission of European Communities.

References

[BACe] G. Ben Arous, R.Cerf "Metastability of the three dimensional Ising model on a torus at very low temperature", Preprint 1995

[CG] M.Cassandro, A.Galves "Language acquisition and change in generalized GW model", Preprint MOL4 1995

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

[Ca1] O.Catoni, "Rough large deviations estimates for simulated annealing. Application to exponential schedules", Ann. Prob. **20**, 1109-1146 (1992)

[Ca2] O.Catoni, "Sharp large deviations estimates for simulated annealing algorithms", Ann. Inst. H. Poincare **27 no.3**, 291-382 (1991)

[CaCe] O.Catoni, R.Cerf "The exit path of a Markov chain with rare transitions", Preprint 1995

[Ce] R.Cerf, "Une théorie asymptotique des algorithmes génétiques" These Universite Montpellier II (1994)

[CesM] F.Cesi, F.Martinelli, "On the layering transition of an SOS surface interacting with a wall. I. Equilibrium results" J. Stat. Phys. **82**, 823 (1996)

[ChiCho1] T.S.Chiang, Y.Chow "A limit theorem for a class of inhomogeneous Markov processes", Ann. Prob. **17**, 1483-1502 (1989)

[ChiCho2] T.S.Chiang, Y.Chow "On the Exit Problem from a Cycle of Simulated Annealing Processes with application - A Backward Equation Approach", Tech. Rept. Inst. of Math. Academia Sinica, 1995

[CiO] E.Cirillo, E.Olivieri,"Metastability and nucleation for the Blume -Capel model: different mechanisms of transition" Journ. Stat. Phys. **83**, 473-571 (1996)

[EGJLa] V.R.Eston, A.Galves, C.M.Jacobi, R.Langevin, "Dominance switch between two interacting species and metastability" Atas do II Simposio dos Ecosistemas da costa sul Brasileira, CACIESP, Sau Paulo (1988)

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

[GSchm] A.Galves, B.Schmitt, "Occurrence of rare events for mixing dynamical systems" Ann. Inst. H. Poincare Phys. Theor. **52**, 267 (1990)

[Ge] H.O.Georgii, Gibbs measures and Phase Transitions, Walter de Gruyter (de Gruyter Studies in Mathematics Vol. 9), Berlin, New York (1988)

[GlHu] S.Glance, B.Huberman, "The dynamics of social dilemmas", Scient. Amer. March 76 (1994)

[Ho1] R.Holley, "Possible rates of convergence in finite range, attractive spin systems", Contemp. Math. **41**, 215 (1985)

[Ho2] R.Holley, "Rapid convergence to equilibrium in one dimensional stochastic Ising models", Ann. Prob. **13**, 72-89 (1985)

[KO1] R.Kotecky, E.Olivieri, "Droplet dynamics for asymmetric Ising model" Journ. Stat. Phys. **70** , 1121-1148, (1993).

[KO2] R.Kotecky, E.Olivieri "Shapes of growing droplets- a model of escape from a matastable phase" Journ. Stat. Phys. **75** , 409-507, (1994).

[Li] T.M.Liggett, Interacting particle systems, Springer Verlag (1985)

[MO1] F.Martinelli, E.Olivieri, "Approach to equilibrium of Glauber dynamics in the one phase region. I: the attractive case", Comm. Math. Phys. **161**, 447-486 (1994)

[MO2] F.Martinelli, E.Olivieri, "Approach to equilibrium of Glauber dynamics in the one phase region II: the general case", Comm. Math. Phys. **161**, 487-514 (1994)

[MOSch] F.Martinelli, E.Olivieri, R.H.Schonmann, "For 2-d lattice spin systems weak mixing implies strong mixing", Comm. Math. Phys. **165**, 33-47 (1994)

[MOS1] F.Martinelli, E.Olivieri, E.Scoppola, "On the Swendsen and Wang dynamics II: Critical droplets and homogeneous nucleation at low temperature for the two dimensional Ising model", J.Stat.Phys. **62**, 135-159 (1991)

[MOS2] F.Martinelli, E.Olivieri, E.Scoppola, "Metastability and exponential approach to equilibrium for low temperature stochastic Ising models", J.Stat.Phys **61**, 1105 (1990)

[MOS3] F.Martinelli, E.Olivieri, E.Scoppola, "On the Swendsen and Wang dynamics I: Exponential convergence to equilibrium", J. Stat. Phys. **62**, 117-134 (1991)

[NaO] F.R.Nardi, E.Olivieri,"Low temperature stochastic dynamics for an Ising model with alternating field."Markov Proc. Rel. Field **2**, 117-166 (1996)

[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)

[NewCoKi] C.Newman, J.E.Cohen, C.Kipnis, "Neo-darwinian evolution implies punctuated equilibria", Nature **315**, 400 (1985)

[NiBe] P.Niyogi, R.C.Berwick, "Formalizing Triggers: a Learning model for finite spaces", Preprint MIT 1993

[OS1] E. Olivieri, E. Scoppola, " Markov chains with exponentially small transition probabilities: First exit problem from a general domain -I. The reversible case" JSP **79**, 613 (1995)

[OS2] E. Olivieri, E. Scoppola, " Markov chains with exponentially small transition probabilities: First exit problem from a general domain -II The general case". JSP **84**, 987-1041 (1996)

[PL1] O.Penrose, J.L.Lebowitz "Rigorous treatment of metastable states in van der Waals theory", J.Stat.Phys. **3**, 211-236 (1971)

[PL2] O.Penrose, J.L.Lebowitz "Towards a rigorous molecular theory of metastability" In Fluctuation Phenomena (second ed.) E. W.Montroll, J.L.Lebowitz ed., North Holland Publ. (1987)

[Ru] D.Ruelle, Statistical Mechanics, Benjamin Inc. (1969)

[Sch1] R.Schonmann, "The pattern of escape from metastability of a stochastic Ising model", Comm. Math. Phys. **147**, 231-240 (1992)

[Sch2] R.Schonmann, "Slow droplet driven relaxation of stochastic Ising models in the vicinity of the phase coexistence region", Comm. Math. Phys. **161**, 1-49 (1994)

[Sch3] R.Schonmann, "Theorems and conjectures on the droplet driven relaxation of stochastic Ising model" in Probability and Phase Transition, Ed. G.Grimmett, NATO ASI Series, Kluwer Ac. Publ.pg. 265-301 (1994)

[SchSh] R.Schonmann, S. Shlosman, "Complete Analiticity for 2D Ising completed", Comm. Math. Phys. **170**, 453 (1995)

[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. (1994)

[Si] Ya.G.Sinai, Theory of Phase Transition: Rigorous Results, Akademiai Kiado, Budapest (1982)

[ToMi] H.Tomita, S.Miyashita "Statistical properties of the relaxation process of metastable states in the kinetic Ising model", Phys.Rev.B. Condensed Matter **46**, 8886-8893 (1992)

[T1] A.Touvé "Cycle Decompositions and Simulated Annealing". SIAM J. Control Optim. To appear

[T2] A.Trouvé "Rough large deviation estimates for the optimal convergence speed exponent of generalized simulated annealing algorithms", Ann. Inst. H. Poicaré To appear