THE EXIT PATH OF A MARKOV CHAIN WITH RARE TRANSITIONS

Olivier CATONI – Raphaël CERF

CNRS, Ecole Normale Supérieure – CNRS, Université Paris Sud

October 1995

Abstract. We study the exit path from a general domain after the last visit to a set of a Markov chain with rare transitions. We prove several large deviation principles for the law of the succession of the cycles visited by the process (the cycle path), the succession of the saddle points gone through to jump from cycle to cycle on the cycle path (the saddle path) and the succession of all the points gone through (the exit path). We estimate the time the process spends in each cycle of the cycle path and how it decomposes into the time spent in each point of the exit path. We describe a systematic method to find the most likely saddle paths. We apply these results to the reversible case of the Metropolis dynamics.

1. Introduction

Markov chains with rare transitions appear in a variety of contexts ranging from statistical mechanics to optimization and reliability theory. This discrete model of a dynamical system with small random perturbations has been studied for itself for the first time by Freidlin and Wentzell in connection with diffusion models [10]. Its main feature is that the transition rates are supposed to obey some large deviations principle with parameter β (called the inverse temperature). In other words the transition rate between states x and y is assumed to be of order $\exp -\beta V(x,y)$ for some rate function V. Exponential rates come from the discrete approximation of more complex events, such as the jumps of a diffusion process with small perturbations from one attractor to another, or the rate of failure of

1991 *Mathematics Subject Classification.* 60J10 60F10 82C44.

Key words and phrases. Freidlin–Wentzell theory, large deviations, exit, metastability.

a machine. They may also be used to simulate on a computer the distribution at thermal equilibrium of a system of particules or magnetic spins, called the Gibbs ensemble by physicists. They may even be used as a rough model for the microscopic dynamics of such a system of statistical mechanics (whether at equilibrium or not).

The main problem in the study of Markov chains with rare transitions is to estimate their rate of convergence towards equilibrium. Some global approaches lead to an estimation of the second eigenvalue of the transition matrix [8,9]. However it is often interesting to have a more precise description of the evolution towards equilibrium than what can be deduced from the knowledge of this second eigenvalue. Namely one is often interested in the occurrence of a partial equilibrium in a subdomain of the state space, as well as in the lifetimes of these partial equilibria. The global relaxation time of the system is the time needed for partial equilibria to melt into a global one. In statistical mechanics this phenomenon is known as the transition from a metastable state to a stable one. For an optimization algorithm, it is described as the problem of escape from non–global minima.

In the context of stochastic optimization, a quite precise description of the trajectories followed by the system to escape from any subdomain of the state space is given in Catoni [7] for simulated annealing. The proofs are inspired by the Freidlin–Wentzell theory which is applied inductively on subsets of increasing cardinality. They are aimed at solving the non–stationary case, that is the case when the rare transition probabilities are decreased with time as is usually done in stochastic optimization algorithms. All the tools needed to generalize this approach to general non reversible Markov chains are contained in [6,20,22], although the study of metastability is not the explicit purpose of these papers.

Another group of people interested in statistical mechanics studied the stationary case for its own sake [16,18]. Although this work was posterior to the studies of the non stationary case, it was done independently. Moreover, only some part of the results established for the non stationary case were rediscovered in the special case of stationary processes.

Nevertheless it is clear that although the stationary case is contained in the non stationary one, it would be useful to provide simpler proofs for it. This is the aim of the present paper, where we use the graphs techniques introduced by Freidlin and Wentzell to prove various large deviations estimates related to the exit path of a Markov chain with rare transitions. Special proofs for the stationary case are useful for optimization theory, as shown in [5] where piecewise stationary strategies are used to obtain almost optimal rates of convergence.

We could have started with refined estimates from which we would have deduced coarser and coarser ones using the contraction principle of large deviations theory. We have instead chosen to proceed step by step from simple estimates to more refined ones. Although this is not the most economical approach from the mathematical point of view, we think it gives an easier understanding of the behaviour of the process. We start with a rough description of the exit path, where we study only the succession of the maximal cycles the process goes through. Then we look for the entrance and exit points of the visited cycles, thus describing the exit saddle path. We give precise estimates for the time the process spends in each cycle of the cycle path. We describe eventually more precisely the succession of arrows the trajectory is likely to go through and we give estimates for the time spent in each point. We put forward an efficient method to find the exit saddle paths the system is likely to take. We also simplify some results in the reversible case of a Metropolis dynamics. This method is applied to study the metastability of the three dimensional Ising model on a torus at very low temperatures in [2], the model dependent variational problems being solved in [1]. This study of a huge and intricate energy landscape would not have been possible without the use of a systematic way to find the exit path and illustrates the efficiency of the method.

2. The main problem

Let E be a finite space. We consider a family of time homogeneous Markov chains (X, P_β) on E indexed by a positive parameter β (the inverse temperature). More precisely, we consider the coordinate process $X = (X_n)_{n\in\mathbb{N}}$ on the space $E^{\mathbb{N}}$ defined by $X_n : (\omega_0, \dots) \mapsto \omega_n$ together with a family of probabilities (P_β) indexed by β; under each of them the coordinate process is a Markov chain. We suppose that these Markov chains are in the Freidlin–Wentzell regime, that is their transition mechanisms satisfy

$$a(\beta)\,\exp -\beta\, V(x,y) \;\leq\; P_\beta(X_{n+1}=y/X_n=x) \;\leq\; \frac{1}{a(\beta)}\,\exp -\beta\, V(x,y)$$

for all x, y in E, where $\beta \mapsto a(\beta)$ is a positive function such that $\lim_{\beta\to\infty} \ln a(\beta)/\beta = 0$ and $V : E\times E \to \mathbb{R}_+ \cup \{\infty\}$ is an irreducible cost function i.e.

$$\forall x, y \in E\times E \qquad \exists i_0, i_1, \dots, i_r \qquad i_0 = x, \quad i_r = y, \quad V(i_0, i_1) + \cdots + V(i_{r-1}, i_r) < \infty.$$

For C an arbitrary subset of E we define the time $\tau(C,m)$ of exit from C after time m

$$\tau(C,m) \;=\; \min\{\, n\geq m : X_n \notin C\,\}$$

(we make the convention that $\tau(C) = \tau(C,0)$).
We define also the time $\theta(C,m)$ of the last visit to the set C before time m

$$\theta(C,m) \;=\; \max\{\, n\leq m : X_n \in C\,\}$$

(if the chain has not visited C before m, we take $\theta(C,m)=0$).
Remark that τ is a stopping time when m is deterministic whereas θ isn't.

Let G and D be two subsets of E such that $G\subset D$ and let x be a starting point in G. Our aim is to describe the behaviour of the chain (X_n) after its last visit to the set G before it escapes from the set D. More precisely we will study the asymptotic behaviour as β goes to infinity of the law of $(X_k,\, \theta(G, \tau(D)) \leq k \leq \tau(D))$. We will identify the points of $D\setminus G$ the Markov chain (X_n) is likely to visit on its exit path as well as the typical times it spends in the subsets of $D\setminus G$ it crosses. The key idea to achieve this study, introduced by Catoni [7], is to decompose $D\setminus G$ into its maximal cycles and to focus on the jumps of (X_n) between these cycles, which may be seen as abstract states.

3. Freidlin and Wentzell lemmas on Markov chains

These lemmas give useful formulas for the invariant measure and for the laws of the exit time and exit point for an arbitrary subset of E. These formulas are rational fractions of the coefficients of the transition matrix whose numerators and denominators are most conveniently written as sums over particular types of graphs.

Definition 3.1. (the graphs $G(W)$)
Let W be an arbitrary non–empty subset of E.
An oriented graph on E is called a W–graph if and only if

- there is no arrow starting from a point of W
- each point of W^c is the initial point of exactly one arrow
- for each point x in W^c, there exists a path in the graph leading from x to W.

The set of all W–graphs is denoted by $G(W)$.

Remark. The third condition above is equivalent to

- there is no cycle in the graph.

Definition 3.2. (the graphs $G_{x,y}(W)$)
Let W be an arbitrary non–empty subset of E, let x belong to E and y to W.
If x belongs to W^c, the set $G_{x,y}(W)$ is the set of all oriented graphs on E such that

- there is no arrow starting from a point of W
- each point of W^c is the initial point of exactly one arrow
- for each point z in W^c, there exists a path in the graph leading from z to W
- there exists a path in the graph leading from x to y.

More concisely, they are the graphs of $G(W)$ which contain a path leading from x to y.
If x belongs to W, the set $G_{x,y}(W)$ is empty if $x \neq y$ and is equal to $G(W)$ if $x = y$.

Remark. The graphs in $G_{x,y}(W)$ have no cycles. For any x in E and y in W, the set $G_{x,y}(W)$ is included in $G(W)$.

Definition 3.3. (the graphs $G(x \not\to W)$)
Let W be an arbitrary non–empty subset of E and let x be a point of E.
If x belongs to W the set $G(x \not\to W)$ is empty.
If x belongs to W^c the set $G(x \not\to W)$ is the set of all oriented graphs on E such that

- there is no arrow starting from a point of W
- each point of W^c except one, say y, is the initial point of exactly one arrow
- there is no cycle in the graph
- there is no path in the graph leading from x to W.

The third condition (no cycle) is equivalent to

- for each z in $W^c \setminus \{y\}$, there is a path in the graph leading from z to $W \cup \{y\}$.

Lemma 3.4. *Let W be an arbitrary non–empty subset of E and let x be a point of E. The set $G(x \not\rightarrow W)$ is the union of all the sets $G_{x,y}(W \cup \{y\})$, $y \in W^c$.*

Remark. In the case $x \in W^c, y \in W$, the definitions of $G_{x,y}(W)$ and $G(x \not\rightarrow W)$ are those given by Wentzell and Freidlin [10]. We have extended these definitions to cover all possible values of x. With our choice for the definition of the time of exit $\tau(W^c)$ (the first time greater than or equal to zero when the chain is outside W^c), the formulas for the law of $X_{\tau(W^c)}$ and for the expectation of $\tau(W^c)$ will remain valid in all cases.

Let g be a graph on E, we define its probability $p_\beta(g)$ by

$$p_\beta(g) = \prod_{(x \to y) \in g} P_\beta(X_{n+1} = y / X_n = x).$$

Lemma 3.5. *(expected number of visits before exit)*
For any non–empty subset W of E, y in W^c and x in E,

$$\sum_{n=0}^{\infty} P_\beta(X_n = y,\, \tau(W^c) > n / X_0 = x) = \frac{\displaystyle\sum_{g \in G_{x,y}(W \cup \{y\})} p_\beta(g)}{\displaystyle\sum_{g \in G(W)} p_\beta(g)}.$$

This lemma, due to Catoni [5], may be used to prove the following three lemmas of Wentzell and Freidlin.

Lemma 3.6. *(stationary measure)*
The stationary measure μ_β of the Markov chain $((X_n)_{n \in \mathbb{N}}, P_\beta)$ is

$$\forall x \in E \qquad \mu_\beta(x) = \frac{\displaystyle\sum_{g \in G(x)} p_\beta(g)}{\displaystyle\sum_{y \in E} \sum_{g \in G(y)} p_\beta(g)}.$$

Lemma 3.7. *(exit point)*
For any non–empty subset W of E, any y in W and x in E,

$$P_\beta(X_{\tau(W^c)} = y / X_0 = x) = \frac{\displaystyle\sum_{g \in G_{x,y}(W)} p_\beta(g)}{\displaystyle\sum_{g \in G(W)} p_\beta(g)}.$$

Lemma 3.8. *(exit time)*
For any subset W of E and x in E,

$$E_\beta(\tau(W^c)/X_0 = x) \;=\; \frac{\displaystyle\sum_{y\in W^c}\;\sum_{g\in G_{x,y}(W\cup\{y\})} p_\beta(g)}{\displaystyle\sum_{g\in G(W)} p_\beta(g)} \;=\; \frac{\displaystyle\sum_{g\in G(x\not\to W)} p_\beta(g)}{\displaystyle\sum_{g\in G(W)} p_\beta(g)}.$$

4. The cycle decomposition

We now recall briefly some basic facts and definitions of quantities concerning the decomposition of E into cycles. For a detailed exposition of this question, we refer the reader to Catoni [5,6] and Trouvé [20,21].

Definition 4.1. (cost of a graph)
We define the cost of a graph g over E by

$$V(g) = \sum_{(x\to y)\in g} V(x,y).$$

Clearly, we have $\lim_{\beta\to\infty} \ln p_\beta(g)/\beta = -V(g)$.

We next define a delicate but extremely useful tool.

Definition 4.2. (renormalized communication cost)
Lemma 3.7. implies the existence of the limits for any subset D of E

$$\begin{aligned}
\forall x\in D \quad \forall y\notin D \qquad & \lim_{\beta\to\infty} -\frac{1}{\beta}\ln P_\beta(X_{\tau(D)} = y/X_0 = x) \;=\; C_D(x,y)\\
&\qquad = \;\min\{\, V(g) : g\in G_{x,y}(D^c)\,\} \;-\; \min\{\, V(g) : g\in G(D^c)\,\},\\
\forall x\notin D \quad \forall y\notin D \qquad & \lim_{\beta\to\infty} -\frac{1}{\beta}\ln P_\beta(X_{\tau(D,1)} = y/X_0 = x) \;=\; C_D(x,y)\\
= \;\min\{\, V(x,z) + V(g) &: z\in D\cup\{y\},\, g\in G_{z,y}(D^c)\,\} \;-\; \min\{\, V(g) : g\in G(D^c)\,\}.
\end{aligned}$$

For x in E and y in D we set $C_D(x,y) = C_{D\setminus\{y\}}(x,y)$ so that C_D is defined over $E\times E$.
The quantity $C_D(x,y)$ is called the renormalized communication cost in D.
For a set A, we define also $C_D(x,A) = \min\{\, C_D(x,y) : y\in A\,\}$.

Remark. The first formula for the renormalized communication cost (corresponding to the case $x\in D$, $y\notin D$) is the formula given by Wentzell and Freidlin.

Remark. For any set D and any x in E, we have $C_D(x,x) = 0$.

Lemma 4.3. *For $x \notin D$ and $y \notin D$,*

$$C_D(x,y) \;=\; \min\{\, V(x,z) + C_D(z,y) : z \in D \cup \{y\}\,\}.$$

For $x \in D$, $y \in D$ and $z \notin D$,

$$C_D(x,z) \;\leq\; C_D(x,y) + C_D(y,z).$$

Proof. The first equation is a direct consequence of the expression of $C_D(x,y)$ and $C_D(z,y)$ in terms of graphs. A probabilistic proof is also possible: let x, y belong to D^c; we have

$$P_\beta(X_{\tau(D,1)} = y / X_0 = x) \;=\; \sum_{z \in D \cup \{y\}} P_\beta(X_{\tau(D)} = y / X_0 = z) P_\beta(X_1 = z / X_0 = x)$$

and the first equation follows letting β tend to infinity.
Now let $x \in D$, $y \in D$ and $z \in D^c$. We can decompose $P_\beta(X_{\tau(D)} = z / X_0 = x)$ as

$$\begin{aligned} P_\beta(X_{\tau(D \backslash \{y\})} = y,\, X_{\tau(D)} = z / X_0 = x) \;+\; P_\beta(X_{\tau(D\backslash\{y\})} = z / X_0 = x) \\ \geq\; P_\beta(X_{\tau(D\backslash\{y\})} = y,\, X_{\tau(D)} = z / X_0 = x) \\ =\; P_\beta(X_{\tau(D\backslash\{y\})} = y / X_0 = x)\, P_\beta(X_{\tau(D)} = z / X_0 = y). \end{aligned}$$

Letting β tend to infinity we obtain

$$C_D(x,z) \;\leq\; C_{D\backslash\{y\}}(x,y) \;+\; C_D(y,z).$$

Moreover by convention $C_{D\backslash\{y\}}(x,y) = C_D(x,y)$. □

For another construction of the renormalized communication cost, see Trouvé [20,21].

Proposition 4.4. *(properties of the renormalized communication cost)*
For any subset D of E and for any point x in E, the cost $C_D(x, D^c)$ is null.
For any x, y, the set function $A \mapsto C_A(x,y)$ is non–increasing i.e.

$$A \subset B \qquad \Longrightarrow \qquad C_A(x,y) \;\geq\; C_B(x,y).$$

Proof. These properties are easy consequences of the probabilistic definition of the renormalized cost. Let us prove the last assertion. First we have $C_A(x,y) = C_{A\backslash\{y\}}(x,y)$, $C_B(x,y) = C_{B\backslash\{y\}}(x,y)$ and we may thus assume that y is in B^c. We have the inclusion $B^c \subset A^c$ so that for $i = 0$ or $i = 1$,

$$\{\, X_{\tau(A,i)} = y \,\} \;\subset\; \{\, X_{\tau(B,i)} = y \,\} \quad \Longrightarrow \quad P_\beta(\, X_{\tau(A,i)} = y \,) \;\leq\; P_\beta(\, X_{\tau(B,i)} = y \,)$$

and finally $C_A(x,y) \geq C_B(x,y)$. □

Definition 4.5. (virtual energy)
Lemma 3.6. implies that for any x in E,

$$\begin{aligned}\lim_{\beta\to\infty} -\frac{1}{\beta}\ln\mu_\beta(x) &= W(x)\\ &= \min\{\,V(g) : g \in G(\{x\})\,\} - \min\{\,V(g) : g\in G(\{y\}),\, y\in E\,\}.\end{aligned}$$

The quantity $W(x)$ is called the virtual energy of x.
For an arbitrary set D, we define its virtual energy $W(D)$ by

$$W(D) = \min\{\,W(x) : x \in D\,\}.$$

The bottom $F(D)$ of D is the set of points of D with virtual energy $W(D)$ i.e.

$$F(D) = \{\,x\in D : W(x) = W(D)\,\}.$$

Definition 4.6. (localized virtual energy)
For any subset D of E, any point x of E and any y in D, lemma 3.5. implies the existence of the limit

$$\lim_{\beta\to\infty} -\frac{1}{\beta}\ln\sum_{n=0}^{\infty} P_\beta(X_n = y,\, \tau(D,1) > n / X_0 = x) = W_D(x,y).$$

The quantity $W_D(x,y)$ is the logarithmic rate of the potential of the Markov chain starting from x and killed outside D.
The expression of $W_D(x,y)$ in terms of graphs is for any x in D and y in D,

$$W_D(x,y) = \min\{\,V(g) : g\in G_{x,y}(D^c\cup\{y\})\,\} - \min\{\,V(g) : g \in G(D^c)\,\}$$

and for any x in D^c and y in D,

$$\begin{aligned}W_D(x,y) &= \min\{\,V(x,z)+V(g) : z\in D,\, g\in G_{z,y}(D^c\cup\{y\})\,\} - \min\{\,V(g) : g\in G(D^c)\,\}\\ &= \min\{\,V(x,z) + W_D(z,y) : z \in D\,\}.\end{aligned}$$

Lemma 4.7. *(link between the localized virtual energy and the renormalized cost)*
Let D be an arbitrary subset of E, x in E and y in D. Then

$$W_D(x,y) + C_{D\setminus\{y\}}(y, D^c) = C_{D\setminus\{y\}}(x,y).$$

Proof. Let us assume first that x is in D. We have

$$W_D(x,y) \;=\; \min\{\, V(g) : g \in G_{x,y}(D^c \cup \{y\})\,\} \;-\; \min\{\, V(g) : g \in G(D^c)\,\}$$

and

$$C_{D\setminus\{y\}}(y,D^c) \;=\; \min\{\, V(g) : g \in G(D^c)\,\} \;-\; \min\{\, V(g) : g \in G(D^c \cup \{y\})\,\}.$$

Thus

$$\begin{aligned} W_D(x,y) \,+\, C_{D\setminus\{y\}}(y,D^c) \;&=\; \min\{\, V(g) : g \in G_{x,y}(D^c \cup \{y\})\,\} \\ &\qquad\qquad - \min\{\, V(g) : g \in G(D^c \cup \{y\})\,\} \\ &=\; C_{D\setminus\{y\}}(x,y). \end{aligned}$$

When x is in D^c we can write

$$\begin{aligned} W_D(x,y) \,+\, C_{D\setminus\{y\}}(y,D^c) \;&=\; \min\{\, V(x,z) + W_D(z,y) : z \in D\,\} \,+\, C_{D\setminus\{y\}}(y,D^c) \\ &=\; \min\{\, V(x,z) + C_{D\setminus\{y\}}(z,y) : z \in D\,\} \\ &=\; C_{D\setminus\{y\}}(x,y). \quad \square \end{aligned}$$

Remark. A probabilistic proof is also possible, starting from the identity

$$\Big(\sum_{n=0}^{\infty} P_\beta(X_n = y,\, \tau(D) > n / X_0 = x)\Big)\, P_\beta(X_{\tau(D\setminus\{y\})} \notin D / X_0 = y) \;=\; P_\beta(X_{\tau(D\setminus\{y\})} = y / X_0 = x).$$

Definition 4.8. (height of a set)
The height $H(D)$ of the set D is defined by

$$H(D) \;=\; \max_{x\in D} \quad \lim_{\beta\to\infty} \frac{1}{\beta} \ln E_\beta(\tau(D)/X_0 = x).$$

By lemma 3.8. this limit exists and is equal to

$$H(D) \;=\; -\min\{\, V(g) : x \in D,\, g \in G(x \not\to D^c\,\} + \min\{\, V(g) : g \in G(D^c)\,\}.$$

Definition 4.9. (boundary)
For a subset D of E we define its boundary $B(D)$

$$B(D) \;=\; \{\, y : y \notin D,\, \exists\, x \in D \quad V(x,y) < \infty \,\}$$

and its principal boundary $\widetilde{B}(D)$

$$\widetilde{B}(D) \;=\; \{\, y \in B(D) : \exists x \in D \quad \lim_{\beta\to\infty} -\frac{1}{\beta} \ln P_\beta(X_{\tau(D)} = y / X_0 = x) \;=\; 0 \,\}.$$

Remark. The principal boundary of a set D may equivalently be defined by

$$\widetilde{B}(D) \;=\; \{\, y \in B(D) : \exists x \in D \quad C_D(x,y) = 0 \,\}.$$

Definition 4.10. (cycle)
A cycle π is a subset of E which is either reduced to a point or satisfy

$$\forall x, y \in \pi, \quad x \neq y, \qquad \lim_{\beta\to\infty} -\frac{1}{\beta} \ln P_\beta(X_{\tau(\pi\setminus\{y\})} \neq y / X_0 = x) \;>\; 0.$$

The set of cycles is a tree for the set inclusion, two cycles being either disjoint or comparable for the inclusion relation.

Proposition 4.11. *(mixing properties of cycles)*
Let π be a cycle. For any points x, y in π, the cost $C_\pi(x,y)$ is null.
For any subset D of E containing π and for any x, y in π, we have for any z in E

$$W_D(z,x) - W_D(z,y) \;=\; W(x) - W(y)$$
$$C_D(x,z) \;=\; C_D(y,z), \quad C_D(z,x) \;=\; C_D(z,y)$$

and for any z in D, we have $W_D(x,z) \;=\; W_D(y,z)$.

As a consequence, we can define without ambiguity the renormalized costs $C_D(\pi, z)$, $C_D(z, \pi)$ for any cycle π included in D and any point z in E, as well as $C_D(\pi_1, \pi_2)$ for any cycles π_1, π_2 included in D.

Definition 4.12. (maximal partition)
Let D be a subset of E. The partition of D into its maximal subcycles is denoted by $\mathcal{M}(D)$. For x in D we denote by $\pi(x, D)$ the unique cycle of $\mathcal{M}(D)$ containing x (for x in D^c we make the convention that $\pi(x, D) = \{x\}$).

Remark. In case D is a cycle, we have $\mathcal{M}(D) = \{D\}$.
The relevance of the cycles of $\mathcal{M}(D)$ for the study of the behaviour of the Markov chain in D will already appear in the next lemma.

Notation. Let g be a graph and A a subset of E, we denote by $g|_A$ the restriction of g to A, obtained by deleting from g all the arrows starting outside from A.

Lemma 4.13. *Let D be an arbitrary subset of E. We have*

$$\min\{\,V(g) : g \in G(D^c)\,\} \;=\; \sum_{\pi \in \mathcal{M}(D)} \min\{\,V(g) : g \in G(\pi^c)\,\}.$$

Moreover for any graph $\widehat{g}$ in $G(D^c)$ realizing the minimum $\min\{\,V(g) : g \in G(D^c)\,\}$ and any cycle π of $\mathcal{M}(D)$, we have $V(\widehat{g}|_\pi) = \min\{\,V(g) : g \in G(\pi^c)\,\}$ and $\widehat{g}|_\pi$ has only one arrow with endpoint outside of π.

Proof. Let us prove first that any $\widehat{g}$ realizing the minimum $\min\{\,V(g) : g \in G(\pi^c)\,\}$ has only one arrow with endpoint outside of π. For this purpose, let us assume that $(x \to y) \in \widehat{g}$, $(x' \to y') \in \widehat{g}$ with $x, x' \in \pi$, $y, y' \in \pi^c$ and $y \neq y'$. The graph $\widehat{g} \setminus \{(x' \to y')\}$ would belong to $G_{x,y}(\pi^c \cup \{x'\})$ and would satisfy

$$V(\widehat{g} \setminus \{(x' \to y')\}) \;=\; \min\{\,V(g) : g \in G(\pi^c \cup \{x'\})\,\},$$

therefore we would have that y belong to $\widetilde{B}(\pi \setminus \{x'\}) \cap \pi^c$ which is in contradiction with the fact that π is a cycle.

Now let $\widehat{g}$ be any graph of $G(D^c)$. For any π in $\mathcal{M}(D)$ let $\widehat{g}|_\pi$ denote its restriction to π (obtained by removing arrows with starting point outside π). Clearly the graph $\widehat{g}|_\pi$ belongs to $G(\pi^c)$. Therefore

$$V(\widehat{g}) \;=\; \sum_{\pi \in \mathcal{M}(D)} V(\widehat{g}|_\pi) \;\geq\; \sum_{\pi \in \mathcal{M}(D)} \min\{\,V(g) : g \in G(\pi^c)\,\} \tag{1}$$

and there is equality if and only if for each π in $\mathcal{M}(D)$ we have $V(\widehat{g}|_\pi) = \min\{\,V(g) : g \in G(\pi^c)\,\}$. Thus to end the proof of the lemma, all we have to do is to build a graph $\widehat{g}$ in $G(D^c)$ such that for any π in $\mathcal{M}(D)$, $V(\widehat{g}|_\pi) = \min\{\,V(g) : g \in G(\pi^c)\,\}$. For this purpose, let us consider the graph $\mathcal{G}$ over the set $\mathcal{M}(D) \cup D^c$ (the points of D^c being identified with one point cycles) defined by

$$(\pi_1 \to \pi_2) \in \mathcal{G} \quad\Longleftrightarrow\quad \widetilde{B}(\pi_1) \cap \pi_2 \neq \emptyset \quad\Longleftrightarrow\quad C_{\pi_1}(\pi_1, \pi_2) = 0.$$

Let $\mathcal{G}|_{\mathcal{M}(D)}$ be its restriction to the cycles of $\mathcal{M}(D)$ (that is we only keep the arrows of $\mathcal{G}$ whose starting point is a cycle of $\mathcal{M}(D)$). This restriction has no stable irreducible component in $\mathcal{M}(D)$. Indeed such a component could not be reduced to one cycle, because no cycle (and in fact no set) has an empty principal boundary, and on the other hand, if it were made of more than one cycle, the union of these cycles would be a cycle of D, and this would contradict the maximality of the elements of $\mathcal{M}(D)$. Consequently, we can extract from $\mathcal{G}|_{\mathcal{M}(D)}$ a spanning collection of oriented trees $\mathcal{H}$ belonging to $G(\mathcal{M}(D)^c)$, that is a graph on $\mathcal{M}(D) \cup D^c$ such that

- each cycle of $\mathcal{M}(D)$ is the starting point of exactly one arrow.
- there is no arrow starting from D^c.
- there is no loops in $\mathcal{H}$.

Now for each π in $\mathcal{M}(D)$, considering π' such that $(\pi \to \pi')$ is in $\mathcal{H}$, we can choose y in $\widetilde{B}(\pi) \cap \pi'$ (because $\mathcal{H} \subset \mathcal{G}$) and find a graph $\widehat{g}_\pi$ in $G_{x,y}(\pi^c)$, where x is some point of π, such that $V(\widehat{g}_\pi) = \min\{\, V(g) : g \in G(\pi^c)\,\}$ (because y is in $\widetilde{B}(\pi)$). Then we know that y is the only endpoint of the arrows of $\widehat{g}_\pi$ lying outside of π. Let $\widehat{g}$ be the union of all the graphs $\widehat{g}_\pi$, $\pi \in \mathcal{M}(D)$. The graph $\widehat{g}$ has no loop, because any loop in $\widehat{g}$ would correspond to a loop in $\mathcal{H}$. Therefore $\widehat{g}$ belongs to $G(D^c)$ and answers the question. $\square$

The formula of lemma 4.13. is very useful to perform several computations of renormalized costs and heights of sets. As an application, we compute the cost appearing in lemma 4.7. above.

Proposition 4.14. *For any subset D of E and s in D we have*

$$C_{D\setminus\{s\}}(s, D^c) \;=\; H(\pi(s,D)) \;+\; W(\pi(s,D)) \;-\; W(s)$$

(where $\pi(s,D)$ is the greatest cycle included in D and containing s.)

Proof. As we already noticed in the proof of lemma 4.7., we have

$$C_{D\setminus\{s\}}(s, D^c) \;=\; \min\{\, V(g) : g \in G(D^c)\,\} \;-\; \min\{\, V(g) : g \in G(D^c \cup \{s\})\,\}.$$

Using lemma 4.13. we obtain

$$\begin{aligned} C_{D\setminus\{s\}}(s, D^c) \;&=\; \sum_{\pi \in \mathcal{M}(D)} \min\{\, V(g) : g \in G(\pi^c)\,\} \;-\; \sum_{\pi \in \mathcal{M}(D\setminus\{s\})} \min\{\, V(g) : g \in G(\pi^c)\,\} \\ &=\; \min\{\, V(g) : g \in G(\pi(s,D)^c)\} \;-\; \sum_{\pi \in \mathcal{M}(\pi(s,D)\setminus\{s\})} \min\{\, V(g) : g \in G(\pi^c)\,\} \\ &=\; \min\{\, V(g) : g \in G(\pi(s,D)^c)\} \;-\; \min\{\, V(g) : g \in G(\pi(s,D)^c \cup \{s\})\}. \end{aligned}$$

Let e belong to $F(\pi(s,D))$. We have also (from the first equation of proposition 4.11.)

$$\min\{\, V(g) : g \in G(\pi(s,D)^c \cup \{s\})\} - \min\{\, V(g) : g \in G(\pi(s,D)^c \cup \{e\})\} \;=\; W(s) - W(e)$$

(i.e. the computation of a difference of virtual energies can be done within a cycle containing the points) so that finally

$$\begin{aligned} C_{D\setminus\{s\}}(s, D^c) \;=\; &\min\{\, V(g) : g \in G(\pi(s,D)^c)\,\} \;-\; \min\{\, V(g) : g \in G(\pi(s,D)^c \cup \{e\})\,\} \\ &+ W(\pi(s,D)) \;-\; W(s) \;=\; H(\pi(s,D)) \;+\; W(\pi(s,D)) \;-\; W(s). \quad \square \end{aligned}$$

Corollary 4.15. $\forall D \subset E \quad \forall s \in D \qquad C_{D\setminus\{s\}}(s, D^c) \;=\; C_{\pi(s,D)\setminus\{s\}}(s, \pi(s,D)^c).$

Proof. We apply proposition 4.14. to the set $D = \pi(s,D)$. Since $\pi(s,\pi(s,D)) = \pi(s,D)$ (see the remark after definition 4.12.), we see that the cost $C_{\pi(s,D)\setminus\{s\}}(s,\pi(s,D)^c)$ is also equal to $H(\pi(s,D)) + W(\pi(s,D)) - W(s)$. $\square$

5. The exit cycle path

We will study how the Markov chain $(X_n)_{n\in\mathbb{N}}$ jumps between the cycles of $\mathcal{M}(D\setminus G)$, the partition of $D\setminus G$ into its maximal subcycles, after its last visit to G.

Definition 5.1. (cycle path)
We define recursively a sequence of random times and cycles:

$$\begin{array}{ll}
\tau_{-1} = \theta(G,\tau(D)), & \pi_{-1} = \{X_{\tau_{-1}}\},\\
\tau_0 = \tau_{-1}+1, & \pi_0 = \pi(X_{\tau_0}, D\setminus G),\\
\tau_1 = \tau(\pi_0,\tau_0), & \pi_1 = \pi(X_{\tau_1}, D\setminus G),\\
\quad\vdots & \quad\vdots\\
\tau_k = \tau(\pi_{k-1},\tau_{k-1}), & \pi_k = \pi(X_{\tau_k}, D\setminus G),\\
\quad\vdots & \quad\vdots\\
\tau_r = \tau(D), & \pi_r = \{X_{\tau_r}\}.
\end{array}$$

Notice that the length r is itself random, since r is defined by the equality $\tau_r = \tau(D)$. The sequence $(\pi_{-1},\pi_0,\dots,\pi_{r-1},\pi_r)$ is called the cycle path of (X_n) relative to D,G and denoted by $\pi(X,D,G)$. It is a random variable with values in the cycle path space

$$\Psi(D,G) \;=\; \big\{\,\{y\} : y\in G\,\big\} \times \bigcup_{r=0}^{\infty} \big\{\,(\pi_0,\dots,\pi_{r-1}) \in \mathcal{M}(D\setminus G) : \pi_k \neq \pi_{k-1}, 1\le k<r\,\big\}$$
$$\times \big\{\,\{z\} : z\in D^c\,\big\}$$

of finite sequences of cycles starting in G, traveling through $\mathcal{M}(D\setminus G)$ and ending in D^c. We define a cost function $V_x(D,G)$ on the space $\Psi(D,G)$ by

$$V_x(D,G)(\{y\},\pi_0,\dots,\pi_{r-1},\{z\}) \;=\;$$
$$W_D(x,y) \;+\; V(y,\pi_0) \;+\; \sum_{k=1}^{r-1} C_{\pi_{k-1}}(\pi_{k-1},\pi_k) \;+\; C_{\pi_{r-1}}(\pi_{r-1},z)$$

where we recall that by definition,

$$V(y,\pi_0) = \min_{u\in\pi_0} V(y,u), \qquad C_{\pi_{k-1}}(\pi_{k-1},\pi_k) = \min_{v\in\pi_k} C_{\pi_{k-1}}(\pi_{k-1},v).$$

Since the sets D and G will be fixed in the sequel, we will drop them in the notations whenever no confusion is possible: for instance we will write $V_x, \Psi, \pi(X)$ instead of $V_x(D,G), \Psi(D,G), \pi(X,D,G)$.

Theorem 5.2. *(estimation of the probability of a cycle path)*
There exists a positive constant K_1 (depending only on the cardinality of $D \setminus G$) such that for any exit cycle path $(\{y\}, \pi_0, \dots, \pi_{r-1}, \{z\})$ in $\Psi(D,G)$ and any β we have

$$\begin{aligned}(K_1^{-1} a(\beta)^{K_1})^{r+1} \exp -\beta V_x(\{y\}, \pi_0, \dots, \pi_{r-1}, \{z\})& \\ \leq\ P_\beta\big(\pi(D,G)\ =\ (\{y\}, \pi_0, \dots, \pi_{r-1}, \{z\})/X_0 = x\big)\ \leq& \\ (K_1 a(\beta)^{-K_1})^{r+1} \exp -\beta V_x(\{y\}, \pi_0, \dots, \pi_{r-1}, \{z\}).&\end{aligned}$$

Proof. Conditioning by the last visit of the chain to the set G and applying the Markov property we get

$$\begin{aligned}&P_\beta\big(\pi(D,G)\ =\ (\{y\}, \pi_0, \dots, \pi_{r-1}, \{z\})/X_0 = x\big) \qquad (2)\\ &\quad = \sum_{n=0}^{\infty} P_\beta(X_n = y, \tau(D) > n/X_0 = x)\ \prod_{k=0}^{r} P_\beta(X_{\tau_k} \in \pi_k / X_{\tau_l} \in \pi_l, -1 \leq l < k).\end{aligned}$$

By lemma 3.5., there exist positive constants C_1, C_1' such that

$$\begin{aligned}C_1' a(\beta)^{C_1} \exp -\beta W_D(x,y) \leq \sum_{n=0}^{\infty} P_\beta(X_n = y, \tau(D) > n/X_0 = x)&\\ \leq C_1'^{-1} a(\beta)^{-C_1} \exp -\beta W_D(x,y).&\end{aligned}$$

Moreover the first term of the product in formula (2) satisfies

$$a(\beta) \exp -\beta V(y, \pi_0)\ \leq\ P_\beta(X_{\tau_0} \in \pi_0 / X_{\tau_{-1}} = y)\ \leq\ a(\beta)^{-1} |\pi_0| \exp -\beta V(y, \pi_0)$$

and for $k \geq 1$ we have

$$\begin{aligned}&P_\beta(X_{\tau_k} \in \pi_k / X_{\tau_l} \in \pi_l, -1 \leq l < k)\ = \qquad (3)\\ &\quad \sum_{u \in \pi_{k-1}} P_\beta(X_{\tau_k} \in \pi_k / X_{\tau_{k-1}} = u)\, P_\beta(X_{\tau_{k-1}} = u / X_{\tau_l} \in \pi_l, -1 \leq l < k).\end{aligned}$$

Yet there exists a positive constant C_2 such that for any u in π_{k-1},

$$\begin{aligned}C_2^{-1} a(\beta)^{C_2} \exp -\beta C_{\pi_{k-1}}(\pi_{k-1}, \pi_k)&\\ \leq\ P_\beta(X_{\tau_k} \in \pi_k / X_{\tau_{k-1}} = u)\ \leq&\\ C_2 a(\beta)^{-C_2} \exp -\beta C_{\pi_{k-1}}(\pi_{k-1}, \pi_k)&\end{aligned}$$

whence, substituting this inequality in the previous equation and summing over $u \in \pi_{k-1}$, we get

$$C_2^{-1}a(\beta)^{C_2} \exp -\beta C_{\pi_{k-1}}(\pi_{k-1},\pi_k)$$
$$\leq\ P_\beta(X_{\tau_k} \in \pi_k / X_{\tau_l} \in \pi_l, -1 \leq l < k)\ \leq$$
$$C_2 a(\beta)^{-C_2} \exp -\beta C_{\pi_{k-1}}(\pi_{k-1},\pi_k)\,.$$

and formula (2) yields the result. □

The cost of a cycle path includes the sum of the costs of each of its arrows $(\pi_{k-1} \to \pi_k)$. Thus any path of bounded cost has a bounded number of arrows of positive cost. However it may have an arbitrary large number of arrows of null cost if there exist loops of null cost.

Definition 5.3. The cycle path $(\{y\},\pi_0,\dots,\pi_{r-1},\{z\})$ is optimal knowing that $X_0 = x$ if its cost $V_x(\{y\},\pi_0,\dots,\pi_{r-1},\{z\})$ is equal to the infimum

$$\inf\{\, V_x(\psi_{-1},\dots,\psi_{s-1},\{z\}) : s \in \mathbb{N}, (\psi_{-1},\dots,\psi_{s-1},\{z\}) \in \Psi(D,G)\,\}.$$

Let $\mathcal{G}_x$ be the graph containing the arrows of all optimal paths knowing that $X_0 = x$.

Proposition 5.4. *The cycle path $(\pi_{-1},\dots,\pi_r)$ is optimal knowing that $X_0 = x$ if and only if it is contained in the optimal graph $\mathcal{G}_x$ (that is, all its arrows do appear in the optimal graph).*

Remark. That any optimal path belongs to $\mathcal{G}_x$ is obvious. The interesting point is that each path contained in the optimal graph $\mathcal{G}_x$ yields also an optimal path knowing $X_0 = x$.

Proof. Let us define the cost of a beginning path $(\pi_{-1},\pi_0,\dots,\pi_k)$ where the π_k's belong to $\mathcal{M}(D \setminus G)$ by

$$V_x(\pi_{-1},\pi_0,\dots,\pi_k)\ =\ W_D(x,\pi_{-1})\ +\ V(\pi_{-1},\pi_0)\ +\ \sum_{l=1}^{k} C_{\pi_{l-1}}(\pi_{l-1},\pi_l).$$

Then for any optimal path $(\pi_{-1},\dots,\pi_r)$ and any $k < r$, $(\pi_{-1},\dots,\pi_k)$ is optimal among the paths ending in π_k. Conversely, assume that $(\pi_{-1},\dots,\pi_r)$ belongs to $\mathcal{G}_x$ i.e. that all the arrows of the path do appear in the graph $\mathcal{G}_x$. We prove by induction that $(\pi_{-1},\dots,\pi_k)$ is optimal among the paths ending in π_k. Suppose the result is true at rank $k-1$. By the very definition of the optimal graph $\mathcal{G}_x$, there exists an optimal path γ such that the arrow (π_{k-1},π_k) is in γ. Let γ_{k-1} be the path γ truncated at π_{k-1} and γ_k be the path γ truncated at π_k. By the induction hypothesis $V_x(\gamma_{k-1}) = V_x(\pi_{-1},\dots,\pi_{k-1})$ so that $V_x(\gamma_k) = V_x(\pi_{-1},\dots,\pi_k)$. Since γ_k is optimal, $(\pi_{-1},\dots,\pi_k)$ is also optimal. □

Remark. The same kind of proof appears in the study of the well known dynamic programming algorithm.

We study now the link between the cycle path cost and the renormalized communication cost.

Theorem 5.5. *Let D be a domain, let x belong to D and let y be a point of E. Let π_x and π_y be the cycles of $\mathcal{M}(D)$ containing x and y respectively (if y is not in D, we put $\pi_y = \{y\}$). The cost $C_D(x,y)$ is equal to the infimum*

$$\inf\left\{ \sum_{k=1}^{r} C_{\pi_{k-1}}(\pi_{k-1},\pi_k) : r\in\mathbb{N},\ \pi_0=\pi_x,\ (\pi_1,\dots,\pi_{r-1})\in\mathcal{M}(D)^{r-1},\ \pi_r=\pi_y \right\}.$$

Remark. Proposition 4.11. shows that $C_D(x,y) = C_D(\pi_x,\pi_y)$ is independent of the pair (x,y) chosen in $\pi_x\times\pi_y$.

Remark. If we had not used maximal cycles in theorem 5.5. as well as in the construction of the exit cycle path, the renormalized communication cost would have been strictly inferior to the infimum of the cycle paths costs and all the exit cycle paths of finite length would have had an exponentially vanishing probability to be taken during the last excursion. This shows that the maximal cycles of $\mathcal{M}(D)$ represent the right granularity at which the last excursion should be described.

Proof. In case $\pi_x=\pi_y$, the cost is zero, as well as the infimum. We assume now that π_x and π_y are distinct. By definition of the renormalized cost, we have $C_D(x,y) = C_{D\setminus\{y\}}(x,y)$ and since x belongs to $D\setminus\{y\}$, then

$$C_{D\setminus\{y\}}(x,y) \;=\; \min\{\,V(g) : g\in G_{x,y}(D^c\cup\{y\})\,\} \;-\; \min\{\,V(g) : g\in G(D^c\cup\{y\})\,\}.$$

Let $\overline{g}$ be a graph realizing the first minimum i.e. $\overline{g}$ belongs to $G_{x,y}(D^c\cup\{y\})$ and

$$V(\overline{g}) \;=\; \min\{\,V(g) : g\in G_{x,y}(D^c\cup\{y\})\,\}.$$

There exists a path of arrows in $\overline{g}$ leading from x to y. Let $\pi_0\to\cdots\to\pi_r$ be the sequence of the successive cycles of $\mathcal{M}(D)$ visited by this path. Since the path of points starts at x and ends at y we have $\pi_0=\pi_x$ and $\pi_r=\pi_y$. In addition we remove the loops in this cycle path, in order to obtain a path of distinct cycles. Now

$$C_D(x,y) \;=\; V(\overline{g}) \;-\; \min\{\,V(g) : g\in G(D^c\cup\{y\})\,\}.$$

Decomposing $\overline{g}$ on the cycles of $\mathcal{M}(D)$ and applying lemma 4.13. we get

$$\begin{aligned}
C_D(x,y) \;&=\; \sum_{\pi\in\mathcal{M}(D)} V(\overline{g}|_\pi) \;-\; \sum_{\pi\in\mathcal{M}(D\setminus\{y\})} \min\{\,V(g) : g\in G(\pi^c)\,\}\\
&=\; \sum_{\pi\in\mathcal{M}(D),\pi\neq\pi_y} \Big(V(\overline{g}|_\pi) - \min\{\,V(g) : g\in G(\pi^c)\,\}\Big) \;+\; V(\overline{g}|_{\pi_y}) -\\
&\qquad\qquad \sum_{\pi\in\mathcal{M}(D\setminus\{y\}),\pi\subset\pi_y} \min\{\,V(g) : g\in G(\pi^c)\,\}.
\end{aligned}$$

However the set of cycles $\{\, \pi : \pi \in \mathcal{M}(D \setminus \{y\}),\ \pi \subset \pi_y \,\}$ is exactly the maximal partition $\mathcal{M}(\pi_y \setminus \{y\})$ whence by applying once more lemma 4.13., the cost $C_D(x,y)$ is equal to

$$\sum_{\pi \in \mathcal{M}(D), \pi \neq \pi_y} \Big(V(\overline{g}|_\pi) - \min\{\, V(g) : g \in G(\pi^c) \,\} \Big) + V(\overline{g}|_{\pi_y}) - \min\{\, V(g) : g \in G(\pi_y^c \cup \{y\}) \,\}$$

Since $\overline{g}$ belongs to $G_{x,y}(D^c \cup \{y\})$, for any π in $\mathcal{M}(D) \setminus \pi_y$, its restriction $\overline{g}|_\pi$ belongs to $G(\pi^c)$. Moreover, $\overline{g}|_{\pi_y}$ belongs to $G(\pi_y^c \cup \{y\})$. As a consequence,

$$C_D(x,y) \;\geq\; \sum_{k=0}^{r-1} \Big(V(\overline{g}|_{\pi_k}) - \min\{\, V(g) : g \in G(\pi_k^c) \,\} \Big).$$

(where $\pi_0, \ldots, \pi_r$ is the sequence of cycles previously described).
By construction of the π_k's, there are x_k in π_k and y_k in π_{k+1} such that $\overline{g}|_{\pi_k}$ is in $G_{x_k,y_k}(\pi_k^c)$, thus $V(\overline{g}|_{\pi_k}) - \min\{\, V(g) : g \in G(\pi_k^c) \} \;\geq\; C_{\pi_k}(x_k, y_k)$. We obtain finally the desired inequality

$$C_D(x,y) \;\geq\; \sum_{k=0}^{r-1} C_{\pi_k}(\pi_k, \pi_{k+1}).$$

Conversely, let $\pi_0 \to \cdots \to \pi_r$ be a cycle path in $\mathcal{M}(D)$ starting at $\pi_0 = \pi_x$, ending at $\pi_r = \pi_y$ and realizing the infimum of the theorem. Let us introduce the stopping times

$$\nu_{-1} = 0, \quad \nu_k = \tau(\pi_k, \nu_{k-1}) \text{ for } 0 \leq k < r, \quad \nu_r = \tau(\pi_r \setminus \{y\}, \nu_{r-1}).$$

We have, putting $\pi_{r+1} = \{y\}$,

$$\begin{aligned} P_\beta(X_{\tau(D\setminus\{y\})} &= y / X_0 = x) \\ &\geq \prod_{k=0}^{r} P_\beta(X_{\nu_k} \in \pi_{k+1} / X_{\nu_s} \in \pi_{s+1}, s < k, X_0 = x) \\ &\geq \prod_{k=0}^{r-1} \inf_{u \in \pi_k} P_\beta(X_{\tau(\pi_k)} \in \pi_{k+1} / X_0 = u) \times \inf_{u \in \pi_r} P_\beta(X_{\tau(\pi_r \setminus \{y\})} = y / X_0 = u). \end{aligned}$$

Using the definition of the communication cost and letting β tend to infinity we get that

$$C_D(x,y) \;\leq\; \sum_{k=0}^{r-1} C_{\pi_k}(\pi_k, \pi_{k+1}). \quad \square$$

We first obtain a slight improvement of the second equation of lemma 4.3..

Corollary 5.6. *For x in D and y in D^c, the cost $C_D(x,y)$ is equal to*

$$C_D(x,y) \;=\; \min\{\, C_D(x,z) + C_{\pi(z,D)}(z,y) : z \in D \,\}.$$

We apply now theorem 5.5. to the set $D \setminus G$.

Corollary 5.7. *For any y in G and z in D^c, the cost $C_{D\setminus G}(y,z)$ is equal to the infimum*

$$\inf \left\{ V(y,\pi_0) \,+\, \sum_{k=1}^{r-1} C_{\pi_k}(\pi_k,\pi_{k+1}) : (\{y\},\pi_0,\dots,\pi_{r-1},\{z\}) \in \Psi(D,G) \right\}.$$

Proof. This is a consequence of the fact that

$$C_{D\setminus G}(y,z) \;=\; \min\,\left\{\, V(y,u) \,+\, C_{D\setminus G}(u,z) : u \in D\setminus G \cup \{z\} \,\right\}$$

(lemma 4.3.) and of theorem 5.5.. $\square$

Corollary 5.8. *The function L_x from $\mathcal{M}(D\setminus G) \cup \{\, \{z\} : z \in D^c \,\}$ to $\mathbb{R}^+$ defined by*

$$L_x(\pi) \;=\; \min_{y\in G} \big(W_D(x,y) \,+\, C_{D\setminus G}(y,\pi) \big)$$

is non–decreasing along the optimal cycle paths knowing that $X_0 = x$.

6. The pruned cycle path

The estimates in theorem 5.2. are not very satisfactory because their precision depends on the length of the path and also because paths of bounded cost do not necessarily have a bounded length. In this section we will eliminate the loops in the graph of transitions of null cost between cycles. We define a relation $\rightarrow$ on the cycles of $\mathcal{M}(D\setminus G)$ by

$$\pi_1 \rightarrow \pi_2 \qquad \Longleftrightarrow \qquad \widetilde{B}(\pi_1)\cap\pi_2 \neq \emptyset \qquad \Longleftrightarrow \qquad C_{\pi_1}(\pi_1,\pi_2) = 0.$$

The associated equivalence relation $\equiv$ is

$$\pi \equiv \pi' \qquad \Longleftrightarrow \qquad \exists \pi_1,\dots,\pi_r,\pi'_1,\dots,\pi'_h \in \mathcal{M}(D\setminus G)$$
$$\pi \rightarrow \pi_1 \rightarrow \cdots \rightarrow \pi_r \rightarrow \pi' \rightarrow \pi'_1 \rightarrow \cdots \rightarrow \pi'_s \rightarrow \pi.$$

We denote by $\overline{\mathcal{M}}(D\setminus G)$ the partition of $\mathcal{M}(D\setminus G)$ determined by this equivalence relation. Formally, the elements of $\overline{\mathcal{M}}(D\setminus G)$ are sets of cycles, but we rather consider them as subsets of E by identifying $\overline{\pi}$ in $\mathcal{M}(D\setminus G)$ with the set $\{\, x : \exists \pi \in \overline{\pi},\ x \in \pi \,\}$. For a point x in $D\setminus G$, we note $\overline{\pi}(x, D\setminus G)$ the unique element of $\overline{\mathcal{M}}(D\setminus G)$ containing x.
For any $\overline{\pi}$ in $\overline{\mathcal{M}}(D\setminus G)$ and for any y in E the cost $C_{\overline{\pi}}(x,y)$ is independent of the point x in $\overline{\pi}$ (by theorem 5.5.). We can thus define the cost $C_{\overline{\pi}}(\overline{\pi},y)$. Once more, if A is a subset of E, we set $C_{\overline{\pi}}(\overline{\pi},A) = \min\{\, C_{\overline{\pi}}(\overline{\pi},y) : y \in A \,\}$.

Definition 6.1. (pruned cycle path)
We define recursively a sequence of random times and cycles:

$$\begin{array}{ll}
\overline{\tau}_{-1} = \theta(G, \tau(D)), & \overline{\pi}_{-1} = \{X_{\overline{\tau}_{-1}}\}, \\
\overline{\tau}_0 = \overline{\tau}_{-1} + 1, & \overline{\pi}_0 = \overline{\pi}(X_{\overline{\tau}_0}, D \setminus G), \\
\overline{\tau}_1 = \tau(\overline{\pi}_0, \overline{\tau}_0), & \overline{\pi}_1 = \overline{\pi}(X_{\overline{\tau}_1}, D \setminus G), \\
\quad\vdots & \quad\vdots \\
\overline{\tau}_k = \tau(\overline{\pi}_{k-1}, \overline{\tau}_{k-1}), & \overline{\pi}_k = \overline{\pi}(X_{\overline{\tau}_k}, D \setminus G), \\
\quad\vdots & \quad\vdots \\
\overline{\tau}_r = \tau(D), & \overline{\pi}_r = \{X_{\overline{\tau}_r}\}.
\end{array}$$

The sequence $(\overline{\pi}_{-1}, \overline{\pi}_0, \dots, \overline{\pi}_{r-1}, \overline{\pi}_r)$ is called the pruned cycle path of (X_n) relative to D, G and denoted by $\overline{\pi}(X, D, G)$. It is a random variable with values in the pruned cycle path space

$$\overline{\Psi}(D,G) \;=\; \big\{ \{y\} : y \in G \big\} \times \bigcup_{r=0}^{\infty} \big\{ (\overline{\pi}_0, \dots, \overline{\pi}_{r-1}) \in \overline{\mathcal{M}}(D \setminus G) : \overline{\pi}_k \neq \overline{\pi}_{k-1}, 1 \leq k < r \big\} \\ \times \big\{ \{z\} : z \in D^c \big\}$$

of finite sequences of sets of equivalent cycles starting in G, traveling through $\overline{\mathcal{M}}(D \setminus G)$ and ending in D^c. We define a cost function $\overline{V}_x(D,G)$ on the space $\overline{\Psi}(D,G)$ by

$$\overline{V}_x(D,G)(\{y\}, \overline{\pi}_0, \dots, \overline{\pi}_{r-1}, \{z\}) \;= \\ W_D(x,y) \;+\; V(y, \overline{\pi}_0) \;+\; \sum_{k=1}^{r-1} C_{\overline{\pi}_{k-1}}(\overline{\pi}_{k-1}, \overline{\pi}_k) \;+\; C_{\overline{\pi}_{r-1}}(\overline{\pi}_{r-1}, z).$$

Notice that the pruned cycle path of (X_n) is a partition of the cycle path of (X_n): it is obtained by regrouping equivalent cycles in $\mathcal{M}(D \setminus G)$. Theorem 5.5. shows that the V_x–cost of a cycle path and the $\overline{V}_x$–cost of the corresponding pruned cycle path are equal i.e. we have $V_x(\pi(X,D,G)) \;=\; \overline{V}_x(\overline{\pi}(X,D,G))$.

Theorem 6.2. *(estimation of the probability of a pruned cycle path)*
There exists a positive constant K_2 (depending only on the cardinality of $D \setminus G$) such that for any pruned cycle path $(\{y\}, \overline{\pi}_0, \dots, \overline{\pi}_{r-1}, \{z\})$ in $\overline{\Psi}(D,G)$ and any β we have

$$(K_2^{-1} a(\beta)^{K_2})^{r+1} \exp -\beta \overline{V}_x(\{y\}, \overline{\pi}_0, \dots, \overline{\pi}_{r-1}, \{z\}) \\ \leq\; P_\beta\big(\overline{\pi}(D,G) \;=\; (\{y\}, \overline{\pi}_0, \dots, \overline{\pi}_{r-1}, \{z\}) / X_0 = x\big) \;\leq \\ (K_2 a(\beta)^{-K_2})^{r+1} \exp -\beta \overline{V}_x(\{y\}, \overline{\pi}_0, \dots, \overline{\pi}_{r-1}, \{z\}).$$

Proof. This is the same kind of proof as for theorem 5.2.. The only difference is that we now use the estimation (for some constant C_4), for all u in $\overline{\pi}_{k-1}$,

$$C_4^{-1} a(\beta)^{C_4} \exp -\beta C_{\overline{\pi}_{k-1}}(\overline{\pi}_{k-1}, \overline{\pi}_k) \\ \leq \; P_\beta(X_{\overline{\tau}_k} \in \overline{\pi}_k / X_{\overline{\tau}_{k-1}} = u) \; \leq \\ C_4 a(\beta)^{-C_4} \exp -\beta C_{\overline{\pi}_{k-1}}(\overline{\pi}_{k-1}, \overline{\pi}_k). \quad \square$$

The key fact that makes $\overline{V}_x$ a good cost function to work with is the following

Lemma 6.3. *Let* $(\overline{\pi}_0, \ldots, \overline{\pi}_r)$ *be a sequence such that* $\overline{\pi}_k \in \overline{\mathcal{M}}(D \setminus G)$, $\overline{\pi}_k \neq \overline{\pi}_{k-1}$, $1 \leq k \leq r$. *There exists a positive constant* δ *independent of* $(\overline{\pi}_0, \ldots, \overline{\pi}_r)$ *such that, if* r *is greater than the cardinality of* $\overline{\mathcal{M}}(D \setminus G)$, *then*

$$\sum_{k=1}^{r} C_{\overline{\pi}_{k-1}}(\overline{\pi}_{k-1}, \overline{\pi}_k) \;\geq\; \delta \;>\; 0.$$

Remark. The graph of the jumps of null cost over $\overline{\mathcal{M}}(D \setminus G)$ has no loop. We could have replaced the cardinality of $\overline{\mathcal{M}}(D \setminus G)$ by the diameter of this graph, that is the number of vertices of its longest paths.

Proof. Since r is greater than or equal to $|\overline{\mathcal{M}}(D \setminus G)|$, two elements $\overline{\pi}_i$ and $\overline{\pi}_j$, $0 \leq i < j \leq r$ of the sequence $\overline{\pi}_0, \ldots, \overline{\pi}_r$ have to be equal. Since $\overline{\pi}_i \neq \overline{\pi}_{i+1}$, then $j > i+1$ and $\overline{\pi}_i, \overline{\pi}_{i+1}, \ldots, \overline{\pi}_j = \overline{\pi}_i$ is a loop. Therefore one of its jumps at least has a positive cost and the sum of the lemma is not smaller than

$$\delta \;=\; \min\{C_{\overline{\pi}}(\overline{\pi}, y) : \overline{\pi} \in \overline{\mathcal{M}}(D \setminus G),\; y \in B(\overline{\pi}),\; C_{\overline{\pi}}(\overline{\pi}, y) > 0\}. \quad \square$$

Corollary 6.4. *The cost* $\overline{V}_x$ *gives a control on the length i.e.*
$\forall \lambda > 0 \quad \exists\, R(\lambda) \quad \forall r \geq R(\lambda) \quad \forall (\overline{\pi}_0, \ldots, \overline{\pi}_r) \in \overline{\Psi}(D \setminus G) \qquad \overline{V}_x(\overline{\pi}_0, \ldots, \overline{\pi}_r) \geq \lambda.$

Corollary 6.5. *The cost function* $\overline{V}_x$ *is a good rate function on* $\overline{\Psi}$ *i.e. its level sets are compact (finite in our situation).*

Proof. Corollary 6.4. shows that the sets $\{ (\overline{\pi}_0, \ldots, \overline{\pi}_r) : \overline{V}_x(\overline{\pi}_0, \ldots, \overline{\pi}_r) \leq \lambda \}$ contain only sequences of bounded length and are thus finite. $\square$

The length of the cycle path $\pi(X)$ (i.e. the number of jumps it involves) will be denoted by $|\pi(X)|$ in the sequel. For instance $|(\pi_0, \ldots, \pi_r)| = r$.

Proposition 6.6. *There exist a positive constant K_3 and a function $h:\mathbb{N}\to\mathbb{R}_+$ such that $\lim_{+\infty} h=+\infty$ and*

$$P_\beta(|\overline{\pi}(X)|\geq R/X_0=x)\ \leq\ (K_3 a(\beta)^{-K_3})^R \exp -\beta h(R).$$

Proof.

$$\begin{aligned}
&P_\beta(|\overline{\pi}(X)|\geq R/X_0=x)\\
&\quad=\sum_{y,\overline{\pi}_0,\dots,\overline{\pi}_R} P_\beta(\overline{\pi}_{-1}(X)=\{y\},\overline{\pi}_0(X)=\overline{\pi}_0,\dots,\ \overline{\pi}_R(X)=\overline{\pi}_R/X_0=x)\\
&\quad\leq\sum_{y,\overline{\pi}_0,\dots,\overline{\pi}_R}(K_2 a(\beta)^{-K_2})^{R+1}\exp -\beta\Big(W_D(x,y)+V(y,\overline{\pi}_0)+\sum_{k=1}^{R} C_{\overline{\pi}_{k-1}}(\overline{\pi}_{k-1},\overline{\pi}_k)\Big)\\
&\quad\leq\ |G|\big(|\overline{\mathcal{M}}(D\setminus G)|\,K_2\,a(\beta)^{-K_2}\big)^{R+1}\exp -\beta(\lfloor R/|\overline{\mathcal{M}}(D\setminus G)|\rfloor\delta - C_3)
\end{aligned}$$

where δ is as in lemma 6.3. and

$$C_3\ =\ -\inf\{\,W_D(x,y)+V(y,\overline{\pi}_0):y\in G,\,\overline{\pi}_0\in\overline{\mathcal{M}}(D\setminus G)\,\}.\quad\square$$

Theorem 6.7. *(large deviations upper bound)*
For any positive λ, there exists a positive constant K_4 such that for any β

$$P_\beta(V_x(\pi(X))\geq\lambda/X_0=x)\ \leq\ (K_4 a(\beta))^{-K_4}\exp -\beta\lambda.$$

Proof. We decompose the event $\{\,V_x(\pi(X))\geq\lambda\,\}$ according to the length of the associated pruned cycle path. Let R be such that $h(R)>\lambda$. We have

$$\begin{aligned}
&P_\beta(V_x(\pi(X))\geq\lambda/X_0=x)\ =\ P_\beta(\overline{V}_x(\overline{\pi}(X))\geq\lambda/X_0=x)\\
&\quad=\ P_\beta(\overline{V}_x(\overline{\pi}(X))\geq\lambda,\,|\overline{\pi}(X)|\geq R/X_0=x)\ +\ P_\beta(\overline{V}_x(\overline{\pi}(X))\geq\lambda,\,|\overline{\pi}(X)|<R/X_0=x).
\end{aligned}$$

The first term in the sum is controlled by proposition 6.6.. The second term is equal to

$$\sum P_\beta(\overline{\pi}(X)=(\overline{\pi}_0,\dots,\overline{\pi}_r)/X_0=x)$$

where the sum runs over all pruned cycle paths $(\overline{\pi}_0,\dots,\overline{\pi}_r)$ satisfying the conditions $r<R$, $\overline{V}_x(\overline{\pi}_0,\dots,\overline{\pi}_r)\geq\lambda$. For such a path we have by theorem 6.2.

$$P_\beta(\overline{\pi}(X)=(\overline{\pi}_0,\dots,\overline{\pi}_r)/X_0=x)\ \leq\ K_2^{r+1}a(\beta)^{-K_2(r+1)}\exp -\beta\lambda$$

and there is only a finite number of such terms (the length of the paths being bounded). The desired upper bound follows easily. $\square$

7. The exit saddle path

In this section, we will study the entrance and exit points of the cycles of the exit cycle path.

Definition 7.1. (saddle path)
We define recursively a sequence of random times and points:

$$
\begin{array}{lll}
\tau_0 = \theta(G, \tau(D)) + 1, & s_0 = X_{\tau_0 - 1}, & s_1 = X_{\tau_0}, \\
\tau_1 = \tau(\pi(s_1, D \setminus G), \tau_0), & s_2 = X_{\tau_1 - 1}, & s_3 = X_{\tau_1}, \\
\quad\vdots & \quad\vdots & \quad\vdots \\
\tau_k = \tau(\pi(s_{2k-1}, D \setminus G), \tau_{k-1}), & s_{2k} = X_{\tau_k - 1}, & s_{2k+1} = X_{\tau_k}, \\
\quad\vdots & \quad\vdots & \quad\vdots \\
\tau_r = \tau(D), & s_{2r} = X_{\tau_r - 1}, & s_{2r+1} = X_{\tau_r}.
\end{array}
$$

Remark. The times τ_i are the same as those used in the definition of the cycle path. We recall also that the length r is random, since r is defined by the equality $\tau_r = \tau(D)$.

The sequence $(s_0, \dots, s_{2r+1})$ is called the saddle path of (X_n) relative to D, G and denoted by $\mathcal{S}(X, D, G)$. The saddle path space relative to D, G starting at x is the set

$$
\begin{aligned}
\mathcal{S}(D,G) \;=\; \big\{ (s_0, \dots, s_{2r+1}) : s_0 \in G,\ s_{2r+1} \in D^c,\ \forall k \in \{1, \dots, r\} \quad s_{2k-1} \in D \setminus G, \\
s_{2k} \in \pi(s_{2k-1}, D \setminus G),\ V(s_{2k}, s_{2k+1}) < \infty,\ s_{2k+1} \in B(\pi(s_{2k-1}, D \setminus G)) \big\}.
\end{aligned}
$$

We define a cost function $V_x(D, G)$ on the space $\mathcal{S}(D, G)$ by

$$
\begin{aligned}
&V_x(D,G)(s_0, \dots, s_{2r+1}) \;=\\
&W_D(x, s_0) + V(s_0, s_1) + \sum_{k=1}^{r} \Big(W(s_{2k}) + V(s_{2k}, s_{2k+1}) - H(\pi(s_{2k}, D \setminus G)) - W(\pi(s_{2k}, D \setminus G)) \Big).
\end{aligned}
$$

Since the sets D and G will be fixed in the sequel, we will drop them in the notations whenever no confusion is possible: for instance we will write $V_x, \mathcal{S}, \mathcal{S}(X)$ instead of $V_x(D,G), \mathcal{S}(D,G), \mathcal{S}(X,D,G)$.
The next lemma gives the probability of one fixed jump of a saddle path.

Lemma 7.2. *(last point and exit point of a cycle)*
There exists a positive constant K_5 such that for any cycle π, $x, y \in \pi$ and $z \in B(\pi)$,

$$
\begin{aligned}
K_5^{-1} a(\beta)^{K_5} \exp -\beta\big(W(y) + V(y,z) - H(\pi) - W(\pi)\big) \\
\le\; P_\beta(X_{\tau(\pi)} = z, X_{\tau(\pi)-1} = y / X_0 = x) \;\le \\
K_5 a(\beta)^{-K_5} \exp -\beta\big(W(y) + V(y,z) - H(\pi) - W(\pi)\big).
\end{aligned}
$$

Proof of lemma. By summing over the possible values of the last instant in π, we obtain

$$P_\beta(X_{\tau(\pi)} = z, X_{\tau(\pi)-1} = y/X_0 = x) \;=\; \sum_{k=0}^{\infty} P_\beta(\tau(\pi) > k,\, X_k = y, X_{k+1} = z/X_0 = x)$$
$$= \Big(\sum_{k=0}^{\infty} P_\beta(\tau(\pi) > k,\, X_k = y/X_0 = x)\Big) P_\beta(X_1 = z/X_0 = y).$$

Lemma 3.5. yields the existence of a positive constant C_5 such that

$$C_5^{-1} a(\beta)^{C_5} \exp -\beta W_\pi(x,y) \le \sum_{k=0}^{\infty} P_\beta(\tau(\pi) > k,\, X_k = y/X_0 = x)$$
$$\le C_5 a(\beta)^{-C_5} \exp -\beta W_\pi(x,y)$$

By lemma 4.7., we know that $W_\pi(x,y) = C_{\pi\setminus\{y\}}(x,y) - C_{\pi\setminus\{y\}}(y,\pi^c)$. Since π is a cycle, the cost $C_{\pi\setminus\{y\}}(x,y)$ is null. Moreover proposition 4.14. gives

$$-W_\pi(x,y) \;=\; C_{\pi\setminus\{y\}}(y,\pi^c) \;=H\,(\pi) + W(\pi) - W(y).$$

The required estimation follows easily. $\square$

Corollary 7.3. *For any cycle π, any point x in π and any point y in $B(\pi)$, we have*

$$C_\pi(x,y) \;=\; \min\{\, W(z) + V(z,y) : z \in \pi\} - H(\pi) - W(\pi).$$

Theorem 7.4. *(estimation of the probability of a saddle path)*
There exists a positive constant K_6 (depending only on the cardinality of $D \setminus G$) such that for any saddle path $(s_0, \dots, s_{2r+1})$ in $\mathcal{S}(D \setminus G)$ and any β we have

$$(K_6^{-1} a(\beta)^{K_6})^{r+1} \exp -\beta V_x(s_0, \dots, s_{2r+1})$$
$$\le\; P_\beta(\mathcal{S}(X) = (s_0, \dots, s_{2r+1})/X_0 = x) \;\le$$
$$(K_6 a(\beta)^{-K_6})^{r+1} \exp -\beta V_x(s_0, \dots, s_{2r+1}).$$

Proof. If $S(X) = (s_0, \dots, s_{2r+1})$, the last visit of the chain (X_n) in G occurred at s_0. Conditioning according to the time of this last visit and then applying the Markov property, we obtain

$$
\begin{aligned}
&P_\beta(\mathcal{S}(X)=(s_0,\dots,s_{2r+1})/X_0=x)\\
&=\sum_{k=0}^{\infty} P_\beta(\mathcal{S}(X)=(s_0,\dots,s_{2r+1}),\,X_k=s_0,\,\forall l\quad k<l\le\tau(D),\,X_l\notin G/X_0=x)\\
&=\sum_{k=0}^{\infty} P_\beta(\mathcal{S}(X)=(s_0,\dots,s_{2r+1}),\,X_k=s_0,\,\forall l\quad k<l\le\tau(D),\,X_l\notin G/X_k=s_0)\times\\
&\hfill P_\beta(X_k=s_0,\,k<\tau(D)/X_0=x)\\
&=\Big(\sum_{k=0}^{\infty} P_\beta(X_k=s_0,\,k<\tau(D)/X_0=x)\Big)\times \qquad (4)\\
&\qquad P_\beta(\mathcal{S}(X)=(s_0,\dots,s_{2r+1}),\,X_0=s_0,\,\forall l\quad 0<l\le\tau(D),\,X_l\notin G/X_0=s_0)
\end{aligned}
$$

Lemma 3.5. yields the existence of a positive constant C_6 such that

$$
\begin{aligned}
C_6^{-1}a(\beta)^{C_6}\exp-\beta W_D(x,s_0)\ &\le\ \sum_{k=0}^{\infty} P_\beta(X_k=s_0,\,k<\tau(D)/X_0=x) \qquad (5)\\
&\le\ C_6 a(\beta)^{-C_6}\exp-\beta W_D(x,s_0)
\end{aligned}
$$

We now deal with the second term in formula (4). By repeated conditionings and applications of the strong Markov property we get

$$
\begin{aligned}
&P_\beta(\mathcal{S}(X)=(s_0,\dots,s_{2r+1}),\,\forall l\quad 0<l\le\tau(D),\,X_l\notin G/X_0=s_0)\ =\\
&P_\beta(X_{\tau_r}=s_{2r+1},X_{\tau_r-1}=s_{2r}/X_{\tau_{r-1}}=s_{2r-1})\times\dots\times P_\beta(X_{\tau_1}=s_3,X_{\tau_1-1}=s_2/X_{\tau_0}=s_1)\\
&\hfill\times\ P_\beta(X_{\tau_0}=s_1,X_{\tau_0-1}=s_0,\tau_0=1/X_0=s_0)\ =\\
&P_\beta(X_1=s_1/X_0=s_0)\prod_{k=1}^{r} P_\beta(X_\tau=s_{2k+1},X_{\tau-1}=s_{2k},\tau=\tau(\pi(s_{2k-1},D\backslash G))/X_0=s_{2k-1})
\end{aligned}
$$

Several applications of lemma 7.2. to the above expression, together with formulas (4) and (5), yield the desired estimation. $\square$

We would like to show that the law of the saddle path of (X_n) concentrates on the set of sequences in $\mathcal{S}$ having a null V_x–cost as β goes to infinity.

Definition 7.5. (optimal saddle)
Let π be a cycle. A pair (y,z) is an optimal saddle escaping from π if

$$y\in\pi,\quad z\in\widetilde{B}(\pi),\quad W(y)+V(y,z)=W(\pi)+H(\pi).$$

Let $(s_0,\dots,s_{2r+1})$ belong to $\mathcal{S}$ and suppose there exists a sequence $(t_0,\dots,t_{2h+1})$ in $D\setminus G$ such that

$$t_0\in\pi(s_1,D\setminus G),\quad t_1\in\widetilde{B}(\pi(s_1,D\setminus G)),\quad t_{2h+1}\in\pi(s_1,D\setminus G),$$

and (t_{2k},t_{2k+1}) is an optimal saddle escaping from $\pi(t_{2k-1},D\setminus G)$ for all k in $\{1,\dots,h\}$. Then $(t_0,\dots,t_{2h+1})$ may be seen as a portion of a saddle path of $\mathcal{S}$. Since (t_{2k},t_{2k+1}) is an optimal saddle escaping from $\pi(t_{2k-1},D\setminus G)$ we have

$$\sum_{k=0}^{h}\big(W(t_{2k})+V(t_{2k},t_{2k+1})-H(\pi(t_{2k},D\setminus G))-W(\pi(t_{2k},D\setminus G))\big)=0.$$

Now $(s_0,s_1,t_0,\dots,t_{2h+1},s_2,\dots,s_{2r+1})$ is in the saddle path set $\mathcal{S}$ and has the same V_x–cost as the initial saddle path $(s_0,\dots,s_{2r+1})$. As a consequence, V_x is not in general a good rate function (its level sets are not compact) and with V_x we won't be able to control the length of the saddle path. What we have to do is to get rid of sequences like $(t_0,\dots,t_{2h+1})$ by considering only the relevant jumps between cycles of $\mathcal{M}(D\setminus G)$.

8. The pruned saddle path

Definition 8.1. (pruned saddle path)
We define recursively a sequence of random times and points:

$$\begin{array}{lll}
\overline{\tau}_0=\theta(G,\tau(D))+1, & \overline{s}_0=X_{\overline{\tau}_0-1}, & \overline{s}_1=X_{\overline{\tau}_0},\\
\overline{\tau}_1=\tau(\overline{\pi}(\overline{s}_1,D\setminus G),\overline{\tau}_0), & \overline{s}_2=X_{\overline{\tau}_1-1}, & \overline{s}_3=X_{\overline{\tau}_1},\\
\vdots & \vdots & \vdots\\
\overline{\tau}_k=\tau(\overline{\pi}(\overline{s}_{2k-1},D\setminus G),\overline{\tau}_{k-1}), & \overline{s}_{2k}=X_{\overline{\tau}_k-1}, & \overline{s}_{2k+1}=X_{\overline{\tau}_k},\\
\vdots & \vdots & \vdots\\
\overline{\tau}_r=\tau(D), & \overline{s}_{2r}=X_{\overline{\tau}_r-1}, & \overline{s}_{2r+1}=X_{\overline{\tau}_r}.
\end{array}$$

The sequence $(\overline{s}_0,\dots,\overline{s}_{2r+1})$ is called the pruned saddle path of (X_n) relative to D,G and is denoted by $\overline{\mathcal{S}}(X,D,G)$. The pruned saddle path space relative to D,G starting at x is the set

$$\begin{aligned}\overline{\mathcal{S}}(D,G)\;=\;\big\{(\overline{s}_0,\dots,\overline{s}_{2r+1}):\overline{s}_0\in G,\ \overline{s}_{2r+1}\in D^c,\ \forall k\in\{1,\dots,r\}\quad \overline{s}_{2k-1}\in D\setminus G,\\ \overline{s}_{2k}\in\overline{\pi}(\overline{s}_{2k-1},D\setminus G),\ V(\overline{s}_{2k},\overline{s}_{2k+1})<\infty,\ \overline{s}_{2k+1}\in B(\overline{\pi}(\overline{s}_{2k-1},D\setminus G))\big\}.\end{aligned}$$

We define a cost function $\overline{V}_x(D,G)$ on the space $\mathcal{S}(D,G)$ by

$$\begin{aligned}\overline{V}_x(D,G)(\overline{s}_0,\dots,\overline{s}_{2r+1})\;=\;&W_D(x,\overline{s}_0)\;+\;V(\overline{s}_0,\overline{s}_1)\\ &+\;\sum_{k=1}^{r}\Big(W(\overline{s}_{2k})+V(\overline{s}_{2k},\overline{s}_{2k+1})-H(\overline{\pi}(\overline{s}_{2k},D\setminus G))-W(\overline{\pi}(\overline{s}_{2k},D\setminus G))\Big).\end{aligned}$$

Notice that the pruned saddle path of (X_n) is a subsequence of the saddle path of (X_n): it is obtained by deleting the jumps between equivalent cycles of $\mathcal{M}(D\setminus G)$. These removed jumps have a null cost so that the V_x–cost of a saddle path and the $\overline{V}_x$–cost of the corresponding pruned saddle path are equal i.e. we have $V_x(\mathcal{S}(X)) = \overline{V}_x(\overline{\mathcal{S}}(X))$.

Theorem 8.2. *(estimation of the probability of a pruned saddle path)*
There exists a positive constant K_7 (depending only on the cardinality of $D\setminus G$) such that for any pruned saddle path $(\overline{s}_0,\dots,\overline{s}_{2r+1})$ in $\overline{\mathcal{S}}(D\setminus G)$ and any β we have

$$(K_7^{-1}a(\beta)^{K_7})^{r+1}\exp-\beta\overline{V}_x(\overline{s}_0,\dots,\overline{s}_{2r+1})$$
$$\leq\ P_\beta(\overline{\mathcal{S}}(X)=(\overline{s}_0,\dots,\overline{s}_{2r+1})/X_0=x)\ \leq$$
$$(K_7a(\beta)^{-K_7})^{r+1}\exp-\beta\overline{V}_x(\overline{s}_0,\dots,\overline{s}_{2r+1}).$$

The proof is of the same kind as for theorem 7.4.. The only difference lies in the equivalent form of lemma 7.2., that we state now.

Lemma 8.3. *(inner and outer exit points from a set of equivalent cycles)*
There exists a positive constant K_8 such that for any cycle $\overline{\pi}$, $x,y\in\overline{\pi}$ and $z\in B(\overline{\pi})$,

$$K_8^{-1}a(\beta)^{K_8}\exp-\beta\big(W(y)+V(y,z)-H(\overline{\pi})-W(\overline{\pi})\big)$$
$$\leq\ P_\beta(X_{\tau(\overline{\pi})}=z,X_{\tau(\overline{\pi})-1}=y/X_0=x)\ \leq$$
$$K_8a(\beta)^{-K_8}\exp-\beta\big(W(y)+V(y,z)-H(\overline{\pi})-W(\overline{\pi})\big).$$

Proof. The proof follows the line of the proof of lemma 7.2.. Here the quantity characterizing the rate of decrease of the event under consideration (i.e. the exit from the set $\overline{\pi}$ through the arrow (y,z)) is $W_{\overline{\pi}}(x,y)+V(y,z)$. Application of lemma 4.7. yields $W_{\overline{\pi}}(x,y)=C_{\overline{\pi}\setminus\{y\}}(x,y)-C_{\overline{\pi}\setminus\{y\}}(y,\overline{\pi}^c)$. Since x and y are in $\overline{\pi}$, we have by the definition of $\overline{\pi}$ combined with theorem 5.5. that the cost $C_{\overline{\pi}\setminus\{y\}}(x,y)$ is null. Next, proposition 4.14. yields

$$C_{\overline{\pi}\setminus\{y\}}(y,\overline{\pi}^c)=H(\pi(y,\overline{\pi}))+W(\pi(y,\overline{\pi}))-W(y)$$

(where we recall that $\pi(y,\overline{\pi})$ is the greatest cycle included in $\overline{\pi}$ which contains y). Anyway the quantity $H(\pi)+W(\pi)$ is the same for all the maximal subcycles of $\overline{\pi}$ and coincides with $H(\overline{\pi})+W(\overline{\pi})$. Thus $W_{\overline{\pi}}(x,y)+V(y,z)=W(y)+V(y,z)-H(\overline{\pi})-W(\overline{\pi})$. $\square$

Let $\overline{\Theta}$ be the map from $\overline{\mathcal{S}}(D,G)$ to $\overline{\Psi}(D,G)$ defined by

$$\overline{\Theta}(\overline{s}_0,\dots,\overline{s}_{2r+1})\ =\ (\{\overline{s}_0\},\overline{\pi}(\overline{s}_1,D\setminus G),\dots,\overline{\pi}(\overline{s}_{2r-1},D\setminus G),\{\overline{s}_{2r+1}\}).$$

Let $(\overline{s}_0,\dots,\overline{s}_{2r+1})$ in $\overline{\mathcal{S}}(D,G)$ and $(\overline{\pi}_{-1},\dots,\overline{\pi}_r)$ in $\overline{\Psi}(D,G)$ be such that

$$\overline{\Theta}(\overline{s}_0,\dots,\overline{s}_{2r+1})\ =\ (\overline{\pi}_{-1},\dots,\overline{\pi}_r).$$

We have the inclusion $\big\{\overline{\mathcal{S}}(X)=(\overline{s}_0,\dots,\overline{s}_{2r+1})\big\}\ \subset\ \big\{\overline{\pi}(X)=(\overline{\pi}_{-1},\dots,\overline{\pi}_r)\big\}$.
Theorems 6.2. and 8.2. then imply that $\overline{V}_x(\overline{s}_0,\dots,\overline{s}_{2r+1})\geq\overline{V}_x(\overline{\pi}_{-1},\dots,\overline{\pi}_r)$.
As a consequence, the conclusions of corollaries 6.4. and 6.5. are still valid for the rate function $\overline{V}_x$ on the space $\overline{\mathcal{S}}$. We restate them in the next two corollaries.

Corollary 8.4. *The cost $\overline{V}_x$ gives a control on the length i.e.*
$\forall \lambda > 0 \quad \exists\, R(\lambda) \quad \forall r \geq R(\lambda) \quad \forall (\overline{s}_0, \dots, \overline{s}_{2r+1}) \in \overline{\mathcal{S}}(D \setminus G) \qquad \overline{V}_x(\overline{s}_0, \dots, \overline{s}_{2r+1}) \geq \lambda.$

Corollary 8.5. *The cost function $\overline{V}_x$ is a good rate function on $\overline{\mathcal{S}}$ i.e. its level sets are compact (finite in our situation).*

The length of the saddle path $\mathcal{S}(X)$ (i.e. the number of jumps it involves) will be denoted by $|\mathcal{S}(X)|$ in the sequel. For instance $|(s_0, \dots, s_{2r+1})| = r$.

Corollary 8.6. *There exist two positive constants K_9, K_{10} and a function $h : \mathbb{N} \to \mathbb{R}_+$ such that $\lim_{+\infty} h = +\infty$ and*

$$P_\beta(|\overline{\mathcal{S}}(X)| \geq R / X_0 = x) \;\leq\; (K_9 a(\beta)^{-K_{10}})^R \exp -\beta h(R).$$

Proof. This a straightforward consequence of the equality $|\overline{\mathcal{S}}(X)| = |\overline{\pi}(X)|$ and of proposition 6.6.. $\square$

Theorem 8.7. *(large deviations upper bound)*
For any positive λ, there exist two positive constants K_{11}, K_{12} such that for all β

$$P_\beta(V_x(\mathcal{S}(X)) \geq \lambda / X_0 = x) \;\leq\; K_{12} a(\beta)^{-K_{11}} \exp -\beta\lambda.$$

Proof. We decompose the event $\{\, V_x(\mathcal{S}(X)) \geq \lambda \,\}$ according to the length of the associated pruned saddle path. Let R be such that $h(R) > \lambda$. We have

$$\begin{aligned} &P_\beta(V_x(\mathcal{S}(X)) \geq \lambda / X_0 = x) \;=\; P_\beta(\overline{V}_x(\overline{\mathcal{S}}(X)) \geq \lambda / X_0 = x) \\ &\quad = \; P_\beta(\overline{V}_x(\overline{\mathcal{S}}(X)) \geq \lambda,\, |\overline{\mathcal{S}}(X)| \geq R / X_0 = x) \,+\, P_\beta(\overline{V}_x(\overline{\mathcal{S}}(X)) \geq \lambda,\, |\overline{\mathcal{S}}(X)| < R / X_0 = x). \end{aligned}$$

The first term in the sum is controlled by corollary 8.6.. The second term is equal to

$$\sum P_\beta(\overline{\mathcal{S}}(X) = (\overline{s}_0, \dots, \overline{s}_{2r+1}) / X_0 = x)$$

where the sum runs over all pruned saddle paths $(\overline{s}_0, \dots, \overline{s}_{2r+1})$ satisfying the conditions $r < R$, $\overline{V}_x(\overline{s}_0, \dots, \overline{s}_{2r+1}) \geq \lambda$. For such a path we have by theorem 8.2.

$$P_\beta(\overline{\mathcal{S}}(X) = (\overline{s}_0, \dots, \overline{s}_{2r+1}) / X_0 = x) \;\leq\; (K_7 a(\beta)^{-K_7})^{r+1} \exp -\beta\lambda$$

and there is only a finite number of such terms (the length of the paths being bounded). The desired upper bound follows easily. $\square$

We have just proved that the law of the pruned saddle path satisfies a large deviation principle with the good rate function $\overline{V}_x$. As a consequence, as β goes to infinity, this law concentrates on the paths of null V_x–cost.

Proposition 8.8. *(concentration of the law of the saddle path)*

$$\lim_{\beta\to\infty} P_\beta(V_x(\mathcal{S}(X)) = 0/X_0 = x) \;=\; 1.$$

We finally give a criterion to decide whether a saddle path has a positive asymptotic probability to occur.

Proposition 8.9. *Assume that* $\liminf\limits_{\beta\to\infty} a(\beta) > 0$. *Then*

$$\liminf_{\beta\to\infty} P_\beta(\mathcal{S}(X) = (s_0,\dots,s_{2r+1})/X_0 = x) > 0 \quad\Longleftrightarrow\quad V_x(s_0,\dots,s_{2r+1}) \;=\; 0.$$

9. The saddle paths of null V_x–cost

This section is devoted to the investigation of the structure of the saddle paths having a null V_x–cost.

Theorem 9.1. *(characterization of the saddle paths of null V_x–cost)*
Let $s = (s_0,\dots,s_{2r+1})$ be a saddle path in $\mathcal{S}(D,G)$. Let $r_* = \min\{k : s_{2k+1} \notin \pi(s_0,D)\}$. The saddle path s has a null V_x–cost if and only if

- $C_{D\setminus s_0}(x,s_0) \;=\; 0,$
- $$V(s_0,s_1) + \sum_{1\le k\le r_*}\Big(W(s_{2k}) + V(s_{2k},s_{2k+1}) - H(\pi(s_{2k},D\setminus G)) - W(\pi(s_{2k},D\setminus G))\Big) = H(\pi(s_0,D)) + W(\pi(s_0,D)) - W(s_0),$$
- $\forall k,\quad r_* < k \le r,\quad (s_{2k},s_{2k+1})$ *is an optimal saddle escaping from* $\pi(s_{2k-1},D\setminus G)$.

Remark. The second property is equivalent to saying that s_{2r_*+1} is in the principal boundary of $\pi(s_0,D)$ and that the sequence $(s_0,\dots,s_{2r_*+1})$ is a saddle path of null V_{s_0}-cost of $\Psi(\pi(s_0,D),G)$.

Remark. We recall that the third property says that (see definition 7.5.)

$$\begin{gathered}\forall k\in\{r_*+1,\dots,r\},\quad s_{2k}\in\pi(s_{2k-1},D\setminus G),\quad s_{2k+1}\in\widetilde{B}(\pi(s_{2k-1},D\setminus G)),\\ W(s_{2k}) + V(s_{2k},s_{2k+1}) = W(\pi(s_{2k-1},D\setminus G)) + H(\pi(s_{2k-1},D\setminus G)).\end{gathered}$$

Proof. We have the inequalities

$$
\begin{aligned}
V_x(s_0,\dots,s_{2r+1}) \;&\geq\; W_D(x,s_0) \,+\, V(s_0,s_1) \,+ \sum_{1\leq k\leq r_*}\Big(W(s_{2k})+V(s_{2k},s_{2k+1})\\
\text{(truncation of the sum)}\qquad &\qquad\qquad -H(\pi(s_{2k},D\setminus G))-W(\pi(s_{2k},D\setminus G))\Big)\\
\text{(Corollary 7.3.)}\qquad &\geq\; W_D(x,s_0) \,+ \sum_{0\leq k\leq r_*} C_{\pi(s_{2k},D\setminus G)}(s_{2k},s_{2k+1})\\
\text{(Proposition 4.4.)}\; &\geq\; W_D(x,s_0) \,+\, C_{\pi(s_0,D)\setminus G}(s_0,s_{2r_*+1})\\
\text{(Lemma 4.7.)}\qquad &\geq\; C_{D\setminus\{s_0\}}(x,s_0)-C_{D\setminus\{s_0\}}(s_0,D^c)+C_{\pi(s_0,D)\setminus G}(s_0,s_{2r_*+1})\\
\text{(Corollary 4.15.)}\; &\geq\; C_{D\setminus\{s_0\}}(x,s_0)-C_{\pi(s_0,D)\setminus\{s_0\}}(s_0,\pi(s_0,D)^c)+C_{\pi(s_0,D)\setminus G}(s_0,s_{2r_*+1})\\
&\geq\; C_{D\setminus\{s_0\}}(x,s_0)\;\geq\;0.
\end{aligned}
$$

Thus, for the saddle path to have a null cost, all these successive inequalities have to be equalities. That is, we must have $C_{D\setminus\{s_0\}}(x,s_0)=0$, the quantity

$$
V(s_0,s_1)\,+\sum_{1\leq k\leq r_*}\Big(W(s_{2k})+V(s_{2k},s_{2k+1})-H(\pi(s_{2k},D\setminus G))-W(\pi(s_{2k},D\setminus G))\Big)
$$

must be equal to $H(\pi(s_0,D))\,+\,W(\pi(s_0,D))\,-\,W(s_0)$ and for all k in $\{r_*+1,\dots,r\}$, we must have $W(s_{2k})+V(s_{2k},s_{2k+1})-H(\pi(s_{2k},D\setminus G))-W(\pi(s_{2k},D\setminus G))\;=\;0$, this last condition being equivalent to saying that (s_{2k},s_{2k+1}) is an optimal saddle escaping from the cycle $\pi(s_{2k},D\setminus G)$.
Conversely, if these three conditions are satisfied, we check directly that the V_x–cost of the saddle path is null. $\square$

Definition 9.2. Let A and B be two disjoint subsets of E. We define $\mathcal{V}(A,B)$, the set of points from which A is visited before B, by

$$
\mathcal{V}(A,B)\;=\;\{\,x\in E: C_{A^c\cap B^c}(x,B)>0\,\}.
$$

We define $\mathcal{R}(A,B)$, the set of ridge points between A and B, by

$$
\mathcal{R}(A,B)\;=\;\{\,x\in E: C_{A^c\cap B^c}(x,B)=C_{A^c\cap B^c}(x,A)=0\,\}.
$$

The probabilistic counterpart of this definition is the following:
a point x is in $\mathcal{V}(A,B)$ if and only if

$$
\lim_{\beta\to\infty}-\frac{1}{\beta}\ln P_\beta(\tau(B^c)<\tau(A^c)/X_0=x)\;>\;0\,,
$$

a point x is in $\mathcal{R}(A,B)$ if and only if

$$\lim_{\beta\to\infty} -\frac{1}{\beta}\ln P_\beta(\tau(B^c)<\tau(A^c)/X_0=x) \;=\; 0\,,$$
$$\lim_{\beta\to\infty} -\frac{1}{\beta}\ln P_\beta(\tau(B^c)>\tau(A^c)/X_0=x) \;=\; 0\,.$$

Remark. Notice that for any point of E, we have $C_{A^c\cap B^c}(x,A\cup B)=0$ so that at least one of the two costs $C_{A^c\cap B^c}(x,A)$, $C_{A^c\cap B^c}(x,B)$ is null. Henceforth,

$$C_{A^c\cap B^c}(x,A)>0 \quad\Longrightarrow\quad C_{A^c\cap B^c}(x,B)=0\,,$$
$$C_{A^c\cap B^c}(x,B)>0 \quad\Longrightarrow\quad C_{A^c\cap B^c}(x,A)=0\,.$$

As a consequence, the three sets $\mathcal{V}(A,B),\mathcal{R}(A,B),\mathcal{V}(B,A)$ form a partition of E.

Proposition 9.3. *Let $s=(s_0,\dots,s_{2r+1})$ be a saddle path in $\mathcal{S}(D,G)$ of null V_x–cost. This saddle path visits successively the three sets $\mathcal{V}(G,D^c)$, $\mathcal{R}(G,D^c)$, $\mathcal{V}(D^c,G)$: if we let $r_*=\min\{k: s_{2k+1}\notin\mathcal{V}(G,D^c)\}$ and $r^*=\min\{k\ge r_*: s_{2k+1}\notin\mathcal{R}(G,D^c)\}$, then*

$$\begin{aligned}
\forall k,\quad 0\le k\le 2r_*, &\qquad s_k\in\mathcal{V}(G,D^c),\\
\forall k,\quad 2r_*<k\le 2r^*, &\qquad s_k\in\mathcal{R}(G,D^c),\\
\forall k,\quad 2r^*<k\le 2r+1, &\qquad s_k\in\mathcal{V}(D^c,G).
\end{aligned}$$

In addition the index r_* coincides with $\min\{k: s_{2k+1}\notin\pi(s_0,D)\}$. Thus a saddle path of null V_x–cost may be decomposed into three parts: an ascending part $(s_k,\,0\le k\le 2r_*)$, a ridge part $(s_k,\,2r_*<k\le 2r^*)$ and a descending part $(s_k,\,2r^*<k\le 2r+1)$.

Proof. We first prove that the two indices $r_*^1=\min\{k: s_{2k+1}\notin\pi(s_0,D)\}$ and $r_*^2=\min\{k: s_{2k+1}\notin\mathcal{V}(G,D^c)\}$ are equal. We have $\pi(s_0,D)\subset\mathcal{V}(G,D^c)$ so that $r_*^1\le r_*^2$. Theorem 9.1. shows that $C_{D\setminus G}(s_{2r_*^1+1},D^c)=0$ so that $s_{2r_*^1+1}$ is not in $\mathcal{V}(G,D^c)$ and $r_*^1\ge r_*^2$. Thus $r_*=r_*^1=r_*^2$.
That the points s_k, $0\le k\le 2r_*$ are in $\mathcal{V}(G,D^c)$ is straightforward; in fact they all belong to $\pi(s_0,D)$ (remember that $s_{2k}\in\pi(s_{2k-1},D\setminus G)$). Theorem 9.1. shows that all the subsequent points of the saddle path satisfy $C_{D\setminus G}(s_k,D^c)=0$ so that they are not in $\mathcal{V}(G,D^c)$. The definition of r^* implies that $s_k\in\mathcal{R}(G,D^c)$ for $2r_*<k\le 2r^*$. Finally, theorem 9.1. shows that for $k>2r^*$, we have $C_{D\setminus G}(s_{2r^*+1},s_k)=0$. Each point z such that $C_{D\setminus G}(s_{2r^*+1},z)=0$ is such that $C_{D\setminus G}(z,G)>0$ and is thus in $\mathcal{V}(D^c,G)$ (otherwise we would have $C_{D\setminus G}(s_{2r^*+1},G)=0$). Therefore the points $(s_k,\,2r^*<k\le 2r+1)$ are in $\mathcal{V}(D^c,G)$. $\square$

Definition 9.4. (global saddles)
Let $s = (s_0, \dots, s_{2r+1})$ be a saddle path in $\mathcal{S}(D, G)$ of null V_{s_0}–cost. The global saddles of this saddle path are the saddles $((s_{2k}, s_{2k+1}), r_* \le k \le r^*)$. The set of the global saddles between G and D^c is the union of the global saddles of all the saddle paths of $\mathcal{S}(D, G)$ of null cost.

Definition 9.5. (global saddle points)
A point y is a global saddle point between G and D^c if there exists a global saddle between G and D^c entering the cycle $\pi(y, D \setminus G)$ and a global saddle between G and D^c escaping from the cycle $\pi(y, D \setminus G)$.

Remark. Equivalently, the global saddle points are the ridge points between G and D^c which may be visited with a non exponentially vanishing probability during the last excursion from G before leaving D.

The task of determining the set of all the saddle paths of null V_x–cost can be performed in the following way:

$i)$ first find the points s of G such that $C_D(x, s) = 0$;

$ii)$ for each such point s, find the points s' of the principal boundary of the cycle $\pi(s, D)$;

$iii)$ for each such point s' in $\widetilde{B}(\pi(s, D))$, determine all the sequences of cycles $\pi_0, \dots, \pi_r$ in $\mathcal{M}(D \setminus G)$ such that $s' \in \pi_0$, $\widetilde{B}(\pi_k) \cap \pi_{k+1} \neq \emptyset$ for k in $\{0, \dots, r-1\}$, $\widetilde{B}(\pi_r) \cap D^c \neq \emptyset$;

$iv)$ for each such cycle path $\pi_0, \dots, \pi_r$ determine all the saddle paths $(s_0, \dots, s_{2r+1})$ such that $s_0 \in \pi_0$, $s_{2r+1} \in D^c$, $(s_{2k-1}, s_{2k}) \in \pi_k$ for k in $\{1, \dots, r\}$, and (s_{2k}, s_{2k+1}) is an optimal saddle escaping from π_k for all k in $\{0, \dots, r\}$;

$v)$ finally determine all the saddle paths $(s'_0, \dots, s'_{2r'+1})$ in $\Psi(\pi(s, D), G)$ starting at $s'_0 \in \pi(s, D) \cap G$, ending at $s'_{2r'+1} = s'$ and such that the cost

$$V(s'_0, s'_1) + \sum_{1 \le k \le r'} \Big(W(s'_{2k}) + V(s'_{2k}, s'_{2k+1}) - H(\pi(s'_{2k}, D \setminus G)) - W(\pi(s'_{2k}, D \setminus G)) \Big)$$

is equal to $H(\pi(s'_0, D)) + W(\pi(s'_0, D)) - W(s'_0)$;

$vi)$ the set of all the saddle paths of null V_x–cost passing trough s' is the set of the saddle paths $(s'_0, \dots, s'_{2r'+1}, s_0, s_1, \dots, s_{2r+1})$ obtained by glueing together two saddle paths obtained at step $iv)$ and $v)$.

Remark that step $iii)$ may fail: it may happen that there is no cycle path of null cost in $\mathcal{M}(D \setminus G)$ starting at s' and escaping from D. In that case, there is no saddle path of null cost through s' and it is not necessary to perform step $v)$ (this is the reason why we propose to do step $v)$ at the end). Such a point s' leads to a dead–end. What we do at step $iii)$ is to look at all the cycles we can reach from s' in the graph over $\mathcal{M}(D \setminus G)$ defined by the relation

$$\pi_1 \to \pi_2 \qquad \Longleftrightarrow \qquad \widetilde{B}(\pi_1) \cap \pi_2 \neq \emptyset \qquad \Longleftrightarrow \qquad C_{\pi_1}(\pi_1, \pi_2) = 0$$

i.e. we determine the orbit of the cycle $\pi(s', D\setminus G)$ in this graph (the minimal stable subgraph containing $\pi(s', D\setminus G)$). The good cycle paths are the cycle paths included in this graph which end at a cycle whose principal boundary contains points of D^c.
Depending on the situation, one might also determine first the pruned cycle paths of null cost and then search for the whole cycle paths.
Saddle paths corresponding to step v) always exist: in fact, we have $\pi(s,D)\cap G\neq\emptyset$ and thus for all s_0 in $\pi(s,D)\cap G$, the cost $C_D(x,s_0)$ is null. We consider for instance a saddle path of null $V_s(\pi(s,D),\{s\})$–cost starting at s and escaping from $\pi(s,D)$ at s' (such a saddle path exists since s' is in the principal boundary of $\pi(s,D)$). We choose for s'_0 the last point of the saddle path belonging to G. The portion of the saddle path after s'_0 satisfies the requirements.

10. Time spent in each state of the exit path

We now study the times $\tau_1-\tau_0,\dots,\tau_{r+1}-\tau_r$ between the jumps along the saddle path as well as the times the process spends in the points of the cycle path, knowing that the Markov chain follows a fixed saddle path. The key lies in the following Wentzell–Freidlin type formulas.

Lemma 10.1. *(conditional expected number of visits before exit)*
For an arbitrary set W, two points x,z in W^c and y in W, we have

$$\sum_{n=0}^{\infty}P_\beta\big(X_n=z,\tau(W^c)>n/X_0=x,X_{\tau(W^c)}=y\big)=\frac{\displaystyle\sum_{g_1\in G_{x,z}(W\cup\{z\})}p_\beta(g_1)\sum_{g_2\in G_{z,y}(W)}p_\beta(g_2)}{\displaystyle\sum_{g_1\in G(W)}p_\beta(g_1)\sum_{g_2\in G_{x,y}(W)}p_\beta(g_2)}$$

Proof. We just write (where $1(A)$ denotes the characteristic function of the event A)

$$\begin{aligned}\sum_{n=0}^{\infty}&P_\beta(X_n=z,\tau(W^c)>n/X_0=x,X_{\tau(W^c)}=y)\ =\\ &=\sum_{k=0}^{\infty}\frac{P_\beta(\tau(W^c)>k,\,X_{\tau(W^c)}=y,\,X_k=z/X_0=x)}{P_\beta(X_{\tau(W^c)}=y/X_0=x)}\\ &=\sum_{k=0}^{\infty}\frac{P_\beta(\tau(W^c)>k,\,X_k=z/X_0=x)P_\beta(X_{\tau(W^c)}=y/X_0=z)}{P_\beta(X_{\tau(W^c)}=y/X_0=x)}\end{aligned}$$

and the result follows by applying lemmas 3.5. and 3.7.. $\square$

Lemma 10.2. *(conditional expected exit time)*
For an arbitrary set W, x in W^c and y in W, we have

$$E_\beta(\tau(W^c)/X_0=x,X_{\tau(W^c)}=y) \;=\; \frac{\displaystyle\sum_{z\in W^c}\;\sum_{g_1\in G_{x,z}(W\cup\{z\})} p_\beta(g_1)\sum_{g_2\in G_{z,y}(W)} p_\beta(g_2)}{\displaystyle\sum_{g_1\in G(W)} p_\beta(g_1)\sum_{g_2\in G_{x,y}(W)} p_\beta(g_2)}$$

Proof. We either do a proof in the spirit of the proof of lemma 10.1. or we simply write

$$E_\beta(\tau(W^c)/X_0=x,\,X_{\tau(W^c)}=y)=\sum_{z\in W^c}\sum_{n=0}^\infty P_\beta(X_n=z,\,\tau(W^c)>n/X_0=x,X_{\tau(W^c)}=y)$$

and apply the formula proved in lemma 10.1.. □

Remark. It is interesting to notice that for these conditional formulas the graph summations are coupled in the numerator and are independent for the denominator.

Corollary 10.3. *(large deviations of the conditional expected number of visits)*
As β goes to infinity,

$$\frac{1}{\beta}\ln\sum_{n=0}^\infty P_\beta\big(X_n=z,\tau(W^c)>n/X_0=x,X_{\tau(W^c)}=y\big)$$

goes to

$$\min\{\,V(g_1):g_1\in G(W)\,\}\;+\min\{\,V(g_2):g_2\in G_{x,y}(W)\,\}-\\ \min\{\,V(g_1)+V(g_2):g_1\in G_{x,z}(W\cup\{z\}),\,g_2\in G_{z,y}(W)\,\}.$$

Corollary 10.4. *(large deviations of the expected conditional exit time)*
As β goes to infinity, $\ln E(\tau(W^c)/X_0=x,X_{\tau(W^c)}=y)/\beta$ goes to

$$\min\{\,V(g_1):g_1\in G(W)\,\}\;+\min\{\,V(g_2):g_2\in G_{x,y}(W)\,\}-\\ \min\{\,V(g_1)+V(g_2):g_1\in G_{x,z}(W\cup\{z\}),\,g_2\in G_{z,y}(W),\,z\in W^c\,\}.$$

In general these quantities do really depend on the conditioning event. However we will only be concerned with the case of a cycle.

Proposition 10.5. *For a cycle π, two points x, z in π and y in $B(\pi)$ we have*

$$\begin{aligned}\lim_{\beta\to\infty}\frac{1}{\beta}\ln\sum_{n=0}^{\infty}P_\beta\big(X_n=z,\tau(\pi)>n/X_0=x,X_{\tau(\pi)}=y\big) &= W(\pi)+H(\pi)-W(z),\\ \lim_{\beta\to\infty}\frac{1}{\beta}\ln E(\tau(\pi)/X_0=x,X_{\tau(\pi)}=y) &= H(\pi)\end{aligned}$$

that is, the expected number of visits to a point and the exit time are of the same order (on the logarithmic scale) as if there were no condition imposed on the exit point.

Proof. Since π is a cycle, for any z in π, the minimum $\min\{\, V(g_2) : g_2 \in G_{z,y}(\pi^c \cup \{y\})\,\}$ is equal to $\min\{\, V(g_2) : g_2 \in G(\pi^c \cup \{y\})\,\}$ (which is independent of z). The quantity appearing in corollary 10.3. is thus reduced to

$$\min\{\, V(g_2) : g_2 \in G_{x,y}(\pi^c)\,\} - \min\{\, V(g_1) : g_1 \in G_{x,z}(\pi^c \cup \{z\})\,\}$$

which is exactly $W(\pi)+H(\pi)-W(z)$ (by proposition 4.14.). This proves the first formula stated in the proposition. Taking the infimum over z in π (which corresponds to the quantity appearing in corollary 10.4.), we obtain $H(\pi)$, proving the second formula. □

Proposition 10.6. *(estimation of the conditioned number of visits to a state before exit)*

For any cycle π, any x, u in π, any y in $B(\pi)$, we have for any positive ϵ

$$\lim_{\beta\to\infty} -\frac{1}{\beta}\ln P_\beta\bigg(\sum_{k=0}^{\tau(\pi)}1(X_k=u) > \exp-\beta(W_\pi(u,u)-\epsilon)/X_0=x,X_{\tau(\pi)}=y\bigg) = +\infty$$

and for any small enough positive ϵ

$$\liminf_{\beta\to\infty} -\frac{1}{\beta}\ln P_\beta\bigg(\sum_{k=0}^{\tau(\pi)}1(X_k=u) < \exp-\beta(W_\pi(u,u)+\epsilon)/X_0=x,X_{\tau(\pi)}=y\bigg) \geq \epsilon.$$

Proof. Let us consider the random variable

$$\phi = \sum_{k=0}^{\tau(\pi)}1(X_k=u)$$

and the sequence of stopping times

$$\begin{aligned}\theta_0 &= \inf\{\, n\geq 0 : X_n = u\,\},\\ \theta_k &= \inf\{\, n > \theta_{k-1} : X_n = u\,\} \quad , \text{ for } k\geq 1.\end{aligned}$$

For any $N \geq M > 0$ we have

$$
\begin{aligned}
P_\beta(\phi > N,\, X_{\tau(\pi)} = y/X_0 = x) \;&=\; P_\beta(\theta_N < \tau(\pi),\, X_{\tau(\pi)} = y/X_0 = x) \\
&=\; P_\beta(\theta_0 < \tau(\pi)/X_0 = x)\, P_\beta(\theta_N < \tau(\pi)/X_0 = u)\, P_\beta(X_{\tau(\pi)} = y/X_0 = u) \\
&\leq\; P_\beta(\theta_0 < \tau(\pi)/X_0 = x)\, P_\beta(X_{\tau(\pi)} = y/X_0 = u)\, P_\beta(\theta_M < \tau(\pi)/X_0 = u)^{\lfloor N/M \rfloor} \\
&\leq\; P_\beta(X_{\tau(\pi)} = y/X_0 = x) \left(\frac{E_\beta(\phi/X_0 = u)}{M} \right)^{\lfloor N/M \rfloor}
\end{aligned}
$$

Let us take $M = \lceil e\, E_\beta(\phi/X_0 = u) \rceil$. We obtain

$$
P_\beta(\phi > N,\, X_{\tau(\pi)} = y/X_0 = x) \;\leq\; P_\beta(X_{\tau(\pi)} = y/X_0 = x)\, \exp - \left\lfloor \frac{N}{e\, E_\beta(\phi/X_0 = u) + 1} \right\rfloor
$$

so that

$$
P_\beta(\phi > N/X_{\tau(\pi)} = y, X_0 = x) \;\leq\; \exp - \left\lfloor \frac{N}{e\, E_\beta(\phi/X_0 = u) + 1} \right\rfloor
$$

from which it is easy to deduce the first equation of proposition 10.6..
We now prove the second inequality. We have

$$
\begin{aligned}
P_\beta(\phi < N,\, X_{\tau(\pi)} = y/X_0 = x) \;&=\; \sum_{k=0}^{N-1} P_\beta(\phi = k,\, X_{\tau(\pi)} = y/X_0 = x) \\
=\; P_\beta(X_{\tau(\pi \setminus \{u\})} = y/X_0 = x) \;&+\; \sum_{k=1}^{N-1} P_\beta(\theta_k < \tau(\pi),\, X_{\tau(\pi)} = y,\, \theta_{k+1} > \tau(\pi)/X_0 = x) \\
=\; P_\beta(X_{\tau(\pi \setminus \{u\})} = y/X_0 = x) \;&+\; \sum_{k=1}^{N-1} P_\beta(\theta_k < \tau(\pi)/X_0 = x)\, P_\beta(X_{\tau(\pi \setminus \{u\})} = y/X_0 = u).
\end{aligned}
$$

Moreover

$$
P_\beta(X_{\tau(\pi)} = y/X_0 = u) \;=\; P_\beta(X_{\tau(\pi \setminus \{u\})} = y/X_0 = u)\, E_\beta(\phi/X_0 = u)
$$

so that

$$
\begin{aligned}
&P_\beta(\phi < N,\, X_{\tau(\pi)} = y/X_0 = x) \;\leq \\
&\qquad P_\beta(X_{\tau(\pi \setminus \{u\})} = y/X_0 = x) \;+\; \frac{N-1}{E_\beta(\phi/X_0 = u)} P_\beta(X_{\tau(\pi)} = y/X_0 = u)
\end{aligned}
$$

and the second inequality stated in proposition 10.6. will follow from the two following equations:

$$(6) \qquad \lim_{\beta\to\infty} \frac{P_\beta(X_{\tau(\pi)} = y / X_0 = u)}{P_\beta(X_{\tau(\pi)} = y / X_0 = x)} = 1\,,$$

$$(7) \qquad \liminf_{\beta\to\infty} -\frac{1}{\beta} \ln \frac{P_\beta(X_{\tau(\pi\setminus\{u\})} = y / X_0 = x)}{P_\beta(X_{\tau(\pi)} = y / X_0 = x)} > 0\,.$$

Formula (6) follows from the three equations:

$$\begin{aligned}
&P_\beta(X_{\tau(\pi)} = y / X_0 = u) \geq P_\beta(X_{\tau(\pi\setminus\{x\})} = x / X_0 = u)\, P_\beta(X_{\tau(\pi)} = y / X_0 = x)\,,\\
&P_\beta(X_{\tau(\pi)} = y / X_0 = x) \geq P_\beta(X_{\tau(\pi\setminus\{u\})} = u / X_0 = x)\, P_\beta(X_{\tau(\pi)} = y / X_0 = u)\,,\\
&\lim_{\beta\to\infty} P_\beta(X_{\tau(\pi\setminus\{u\})} = u / X_0 = x) = \lim_{\beta\to\infty} P_\beta(X_{\tau(\pi\setminus\{x\})} = x / X_0 = u) = 1\,.
\end{aligned}$$

Let us finally prove formula (7). We have

$$\begin{aligned}
\frac{P_\beta(X_{\tau(\pi\setminus\{u\})} = y / X_0 = x)}{P_\beta(X_{\tau(\pi)} = y / X_0 = x)} &= 1 - P_\beta(X_{\tau(\pi\setminus\{u\})} = u / X_0 = x)\, \frac{P_\beta(X_{\tau(\pi)} = y / X_0 = u)}{P_\beta(X_{\tau(\pi)} = y / X_0 = x)}\\
(8) \qquad &\leq 1 - \big(1 - P_\beta(X_{\tau(\pi\setminus\{u\})} \neq u / X_0 = x)\big)\big(1 - P_\beta(X_{\tau(\pi\setminus\{x\})} \neq x / X_0 = u)\big).
\end{aligned}$$

Moreover, we know that

$$\begin{aligned}
&\liminf_{\beta\to\infty} -\frac{1}{\beta} \ln P_\beta(X_{\tau(\pi\setminus\{u\})} \neq u / X_0 = x) > 0\,,\\
&\liminf_{\beta\to\infty} -\frac{1}{\beta} \ln P_\beta(X_{\tau(\pi\setminus\{x\})} \neq x / X_0 = u) > 0\,.
\end{aligned}$$

We can thus make an asymptotic expansion in formula (8), from which formula (7) follows. □

Corollary 10.7. *For any cycle π, any x, u in π, any y in $B(\pi)$, any positive ϵ, we have*

$$\lim_{\beta\to\infty} P_\beta\left(\exp -\beta\epsilon \leq \sum_{k=0}^{\tau(\pi)} 1(X_k = u) \exp \beta W_\pi(u,u) \leq \exp \beta\epsilon \,/\, X_0 = x, X_{\tau(\pi)} = y \right) = 1\,.$$

Remark. It is clear from our proof that we could get stronger inequalities than those of proposition 10.6. if we strengthen the hypothesis on $a(\beta)$ (for instance $\lim a(\beta) = a > 0$). We leave the details to the reader.

Corollary 10.8. *(estimation of the conditioned exit time)*
For each $\epsilon > 0$, each cycle π, x in π, y in $B(\pi)$, we have

$$\lim_{\beta\to\infty} P_\beta\Big(\exp\beta(H(\pi)-\epsilon) \leq \tau(\pi) \leq \exp\beta(H(\pi)+\epsilon) / X_0 = x,\, X_{\tau(\pi)} = y \Big) = 1.$$

Proposition 10.9. *For any saddle path $(s_0, \dots, s_{2r+1})$ we have*

$$\lim_{\beta\to\infty} P_\beta\Big(\forall k \in \{1,\dots,r\} \quad \exp\beta(H(\pi(s_{2k}, D\setminus G)-\epsilon) \leq \tau_{k+1}-\tau_k$$
$$\leq \exp\beta(H(\pi(s_{2k}, D\setminus G)+\epsilon) \Big) = 1$$

conditionally on the event $\big\{\pi(X) = \big(\{s_0\}, \pi(s_1, D\setminus G), \dots, \pi(s_{2r-1}, D\setminus G), \{s_{2r+1}\}\big)\big\}$.

Remark. The conditioning can also be done using the event $\big\{\mathcal{S}(X) = (s_0, \dots, s_{2r+1})\big\}$.

Proof. This result follows from repeated conditionings and applications of the preceding corollary 10.8.. $\square$

Proposition 10.10. *For any cycle path $(\{y\}, \pi_0, \dots, \pi_{r-1}, \{z\})$ we have*

$$\lim_{\beta\to\infty} P_\beta\Big(\forall u \in \bigcup_{k=0}^{r-1} \pi_k\,, \quad \exp-\beta(W_{D\setminus G}(u,u)+\epsilon) \leq \sum_{n=\theta(G,\tau(D))}^{\tau(D)} 1(X_n = u)$$
$$\leq \exp-\beta(W_{D\setminus G}(u,u)-\epsilon) \,/\, \pi(X,D,G) = (\{y\}, \pi_0, \dots, \pi_{r-1}, \{z\}) \Big) = 1\,.$$

More precisely, for any u in $\bigcup_{k=0}^{r-1} \pi_k$, for any positive ϵ,

$$\lim_{\beta\to\infty} -\frac{1}{\beta} \ln P_\beta\Big(\sum_{n=\theta(G,\tau(D))}^{\tau(D)} 1(X_n = u) > \exp-\beta(W_{D\setminus G}(u,u)-\epsilon)$$
$$/\, \pi(X,D,G) = (\{y\}, \pi_0, \dots, \pi_{r-1}, \{z\}) \Big) = +\infty$$

and for any small enough positive ϵ

$$\liminf_{\beta\to\infty} -\frac{1}{\beta} \ln P_\beta\Big(\sum_{n=\theta(G,\tau(D))}^{\tau(D)} 1(X_n = u) < \exp-\beta(W_{D\setminus G}(u,u)+\epsilon)$$
$$/\, \pi(X,D,G) = (\{y\}, \pi_0, \dots, \pi_{r-1}, \{z\}) \Big) \geq \epsilon.$$

Proof. We recall that $-W_{D\setminus G}(u,u) = W(\pi(u, D\setminus G)) + H(\pi(u, D\setminus G)) - W(u)$. These estimations follow from proposition 10.6.. $\square$

Remark. The states of the exit saddle path are not necessarily those in which the exit path spends most of its time. The states in which the exit path spends most of its time are the ground states of the deepest cycles of the exit cycle path.

11. Chronology of the exit path

In this section we will be interested in a more refined large deviation upper–bound which controls all the arrows of the exit path and the order in which they appear. With the same notations as before, we introduce the cost function

$$V_x(D,G)(X) \;=\; W_D(x,X_\theta) \,+ \sum_{n=\theta+1}^{\tau(D)} \big(V(X_{n-1},X_n)+W_{\{X_{\theta+1},\ldots,X_{n-1}\}}(X_{n-1},X_{n-1})\big)^+$$

with the convention that $W_\emptyset(u,u)=0$ in all cases. We consider the sequence of random times $\nu_0=0$,

$$\nu_k \;=\; \inf\big\{n>\nu_{k-1} : V(X_{n-1},X_n)+W_{S_n}(X_{n-1},X_{n-1})>0\,\big\},$$

where we have put $S_n=\{\,X_{\theta+1},\ldots,X_{n-1}\,\}$. The cost function defined above is related to the events

$$\begin{aligned}&\mathcal{E}_r(y,(s_l),(A_k),z) \;=\\ &\quad\big\{\,X_\theta=y,\ X_{\nu_k-1}=s_{2k-1},\ X_{\nu_k}=s_{2k},\ S_{\nu_k}=A_k,\ 1\le k\le r,\ \tau(D)<\nu_{r+1},\ X_{\tau(D)}=z\,\big\}\end{aligned}$$

where the sets A_k are subsets of $D\setminus G$ and the points s_l are in $D\setminus G$. For each of these events we have the following estimates.

Theorem 11.1. *There is a positive constant K_{13} such that for any β*

$$\begin{aligned}&(K_{13}^{-1}a(\beta)^{K_{13}})^{r+1}\exp-\beta V_x(\mathcal{E})\\ &\qquad\le\; P_\beta(\mathcal{E}_r(y,(s_l),(A_k),z)/X_0=x)\;\le\\ &\qquad\qquad(K_{13}a(\beta)^{-K_{13}})^{r+1}\exp-\beta V_x(\mathcal{E})\,.\end{aligned}$$

Proof.

$$\begin{aligned}&P_\beta(\mathcal{E}_r(y,(s_l),(A_k),z)/X_0=x) \;=\\ &\quad\bigg(\sum_{m=0}^{\infty}P_\beta(X_m=y,\ \tau(D)>m/X_0=x)\bigg)P_\beta(\mathcal{E}_r(y,(s_l),(A_k),z),\ X_{\tau(D\setminus G)}\in D^c/X_0=y).\end{aligned}$$

Moreover

$$\begin{aligned}P_\beta(\mathcal{E}_r(y,(s_l),(A_k),z),X_{\tau(D\setminus G)}\in D^c/X_0=y)&=P_\beta(X_{\tau(D\setminus G)}=z,\nu_{r+1}>\tau(D)/X_{\nu_r}=s_{2r})\\ &\times\prod_{k=1}^{r}P_\beta(X_{\nu_k-1}=s_{2k-1},\ X_{\nu_k}=s_{2k},\ S_{\nu_k}=A_k/X_{\nu_{k-1}}=s_{2k-2},\ S_{\nu_{k-1}}=A_{k-1}).\end{aligned}\tag{9}$$

We have to estimate the term

$$P_\beta(X_{\nu_k-1}=s_{2k-1},\,X_{\nu_k}=s_{2k},\,S_{\nu_k}=A_k/X_{\nu_{k-1}}=s_{2k-2},\,S_{\nu_{k-1}}=A_{k-1}). \tag{10}$$

Let $\mathcal{I}_k$ be the collection of all the sequences $(u_0,\dots,u_m)$ in E^{m+1} such that

- $u_0=s_{2k-2}$,
- $u_0,\dots,u_m$ are distinct points,
- $A_{k-1}\cup\{\,u_0,\dots,u_m\,\}=A_k$,
- $C_{A_{k-1}\cup\{\,u_0,\dots,u_{l-1}\,\}}(s_{2k-2},u_l)=0$.

If $\mathcal{I}_k$ is empty, then the term in formula (10) is zero, otherwise it is equal to

$$\begin{aligned}\sum_{(u_0,\dots,u_m)\in\mathcal{I}_k}\prod_{l=1}^m & P_\beta\big(X_{\alpha_l}=u_l,\,\nu_k>\alpha_l/X_{\alpha_{l-1}}=u_{l-1},\,S_{u_{l-1}}=A_{k-1}\cup\{\,u_0,\dots,u_{l-2}\,\}\big)\\ &\times P_\beta\big(X_{\nu_k-1}=s_{2k-1},\,X_{\nu_k}=s_{2k},\,S_{\nu_k}=A_k/X_{\alpha_m}=u_m,\,S_{\alpha_m+1}=A_k\big)\end{aligned} \tag{11}$$

where $\alpha_0=\nu_{k-1}$ and $\alpha_l=\inf\{\,n>\alpha_{l-1}/X_n\notin S_n\,\}$. But

$$P_\beta(X_{\alpha_l}=u_l,\,\nu_k\geq\alpha_l/\cdots)\;=\;P_\beta(X_{\alpha_l}=u_l/\cdots)\;-\;P_\beta(X_{\alpha_l}=u_l,\,\nu_k<\alpha_l/\cdots).$$

In addition, for some positive constants γ,C_7,C_8,

$$\begin{gathered}P_\beta(X_{\alpha_l}=u_l,\,\nu_k<\alpha_l/\cdots)\;\leq\;P_\beta(\nu_k<\alpha_l/\cdots)\;\leq\;C_7a(\beta)^{-C_8}\exp(-\gamma\beta)\,,\\ C_7^{-1}a(\beta)^{C_8}\;\leq P_\beta(X_{\alpha_l}=u_l/\cdots)\;\leq\;C_7a(\beta)^{-C_8}.\end{gathered}$$

Therefore for β large enough and for some positive constants C_9,C_{10},

$$C_9^{-1}a(\beta)^{C_{10}}\;\leq\;P_\beta(X_{\alpha_l}=u_l,\,\nu_k\geq\alpha_l/\cdots)\;\leq\;C_9a(\beta)^{-C_{10}}\,. \tag{12}$$

We also have for some positive constants C_{11},C_{12}

$$\begin{aligned}C_{11}^{-1}a(\beta)^{C_{12}}\,\exp-\beta\big(W_{A_k}(u_m,s_{2k-1})+V(s_{2k-1},s_{2k})\big)\;\leq&\\ P_\beta(X_{\nu_k-1}=s_{2k-1},\,X_{\nu_k}=s_{2k},\,S_{\nu_k}=A_k/X_{\alpha_m}=u_m,\,S_{\alpha_m+1}=A_k)&\\ \leq\;C_{11}a(\beta)^{-C_{12}}\,\exp-\beta\big(W_{A_k}(u_m,s_{2k-1})+V(s_{2k-1},s_{2k})\big)\,.&\end{aligned} \tag{13}$$

Let us remark eventually that $W_{A_k}(u_m,s_{2k-1})=W_{A_k}(s_{2k-1},s_{2k-1})$ because it is possible to go from u_m to s_{2k-1} in A_k at null cost. Putting together (9), (11), (12), (13) and noticing that $P_\beta(X_{\tau(D\setminus G)}=z,\,\nu_{r+1}>\tau(D)/X_{\nu_r}=s_{2r})$ is logarithmically equivalent to zero, we obtain the estimates stated in the theorem. $\square$

Theorem 11.2. *For any positive λ there is a positive constant K_{14} such that, for any β,*

$$P_\beta(V_x(X) > \lambda / X_0 = x) \ \le\ K_{14}\, a(\beta)^{-K_{14}} \exp -\beta\lambda.$$

Proof. Let

$$\alpha \ =\ \min\big(\big\{\, \big(V(x,y) + W_A(x,x)\big)^+ : x, y \in E,\ A \subset E \,\big\} \setminus \{0\}\big).$$

We write

$$P(\nu_k < \tau(D)/X_0 = x) \ =\ \sum \Bigg(\Big(\sum_{m=0}^{\infty} P_\beta(X_m = y,\ \tau(D) > m / X_0 = x) \Big)$$

$$\times \prod_{l=1}^{k} P_\beta(X_{\nu_l - 1} = s_{2l-1},\ X_{\nu_l} = s_{2l},\ S_{\nu_l} = A_l / X_{\nu_{l-1}} = s_{2l-2},\ S_{\nu_{l-1}} = A_{l-1}) \Bigg)$$

where the first summation set is $\{\, y \in G, (s_1, \dots, s_{2k}) \in E^{2k}, (A_1, \dots, A_k) \in \mathcal{P}(E)^k \}$. The same estimates as in theorem 11.1. show that for any k there exist positive constant C_{13}, C_{14} such that

$$P_\beta(\nu_k \le \tau(D)/X_0 = x) \ \le\ C_{13}\, a(\beta)^{-C_{14}} \exp -\beta(k\alpha + \min_{y \in G} W_D(x,y)).$$

We choose k such that $k\alpha + \min_{y\in G} W_D(x,y)) > 2\lambda$ and we write

$$P_\beta(V_x(X) > \lambda/X_0 = x) \ \le\ P_\beta(\nu_k \le \tau(D)/X_0 = x) + P_\beta(V_x(X) > \lambda,\ \nu_k > \tau(D)/X_0 = x).$$

The second term of the righthand side is the probability of a finite union of events of the type described in theorem 11.1.; applying the estimates of theorem 11.1. to each of these events yields the desired inequality. $\square$

12. Applications

Falling to the bottom. Let π be a cycle and x a starting point in π. We take $D = \pi \setminus F(\pi)$ and $G = \pi(x, D)$ i.e. we want to study the way the chain either leaves π or reaches the bottom $F(\pi)$ of the cycle. For any point s_0 in $\pi(x,D)$ we have $C_{D\setminus\{s_0\}}(x, s_0) = 0$ (since D contains the cycle $\pi(x,D)$). Theorem 9.1. shows that we must necessarily have (second point of the characterization of the saddle paths of null cost)

$$V(s_0, s_1) \ =\ H(\pi(s_0, D)) + W(\pi(s_0, D)) - W(s_0).$$

Moreover s_1 does not belong to G and is thus in the boundary of $\pi(s_0, D) = \pi(x, D)$. Since by corollary 7.3.

$$H(\pi(s_0, D)) \ +\ W(\pi(s_0, D)) \ -\ W(s_0) \ \le\ V(s_0, s_1) \ -\ C_{\pi(s_0, D)}(s_0, s_1)$$

a saddle path of null V_x–cost has to satisfy $C_{\pi(s_0,D)}(s_0, s_1) = 0$ so that in fact s_1 belongs to $\widetilde{B}(\pi(s_0, D))$. Necessarily, the point s_1 belongs to π. In addition, $C_{D\setminus G}(s_1, s_{2r+1}) = 0$ whence s_{2r+1} is in the bottom $F(\pi)$ (the exit of D takes place in $F(\pi)$). In this situation, the index r_* is null, the ascending part of the saddle path is reduced to (s_0).

Escaping from the bottom. Let π be a cycle and x a starting point in the bottom $F(\pi)$. We take $D=\pi$ and $G=F(\pi)$ i.e. we want to study the way the chain leaves π after its last visit to $F(\pi)$. For any point s_0 in $F(\pi)$ we have $\pi(s_0,D)=\pi$ and $C_{D\setminus\{s_0\}}(x,s_0)=0$. By proposition 4.14. we have also $C_{D\setminus\{s_0\}}(s_0,D^c)=H(\pi)$. In this situation the indices r_* and r^* are equal to the length r of the saddle path, the ascending part is $(s_0,\dots,s_{2r})$, the ridge part is empty, the descending part is s_{2r+1} and s_{2r+1} belongs to the principal boundary of π. The global saddles are the optimal saddles exiting from π i.e. the saddles (y,z) such that $y\in\pi$, $z\in\widetilde{B}(\pi)$ and $W(y)+V(y,z)=W(\pi)+H(\pi)$.

Recursive applications of the result. Knowing that the Markov chain follows a fixed saddle path, we can recursively apply our result to obtain information on the way it enters and exits from the successive cycles of the associated cycle path: for instance, on the way it falls to the bottom or escape from the bottom of an intermediate cycle.

13. The reversible dynamics of Metropolis

Let $U:E\mapsto\mathbb{R}_+$ be a potential whose global minimum over E is zero. We suppose that

$$\forall x,y\in E\qquad V(x,y)<\infty\quad\Longleftrightarrow\quad V(y,x)<\infty$$

and that $V(x,y)=\big(U(y)-U(x)\big)^+$ whenever $V(x,y)<\infty$. In this situation the virtual energy W coincides with the potential U and each quantity admits a simplified expression with the help of U. For instance, for any cycle π, $H(\pi)=\min\{\,(U(x)-U(\pi))^+ : x\in B(\pi)\,\}$.

Proposition 13.1. *(reversibility of the saddle path)*
Let $\mathcal{R}$ be the reversing operator defined on the set of sequences of points of E of odd length by $\mathcal{R}(s_0,\dots,s_{2r+1})=(s_{2r+1},\dots,s_0)$. The operator $\mathcal{R}$ is a one to one map between the saddle paths of $\mathcal{S}(D,G)$ and $\mathcal{S}(G^c,D^c)$.

Proof. The map $\mathcal{R}$ is clearly one to one since $\mathcal{R}\circ\mathcal{R}$ is the identity map. That $\mathcal{R}(\mathcal{S}(D,G))=\mathcal{S}(G^c,D^c)$ is a consequence of the set equality $D\setminus G=G^c\setminus D^c$ and of the following facts:

$$\begin{aligned}
V(s_{2k},s_{2k+1})<\infty &\quad\Longleftrightarrow\quad V(s_{2k+1},s_{2k})<\infty\,,\\
s_{2k}\in\pi(s_{2k-1},D\setminus G) &\quad\Longleftrightarrow\quad s_{2k-1}\in\pi(s_{2k},G^c\setminus D^c)\,,\\
s_{2k+1}\in B(\pi(s_{2k},D\setminus G)) &\quad\Longleftrightarrow\quad s_{2k}\in B(\pi(s_{2k+1},G^c\setminus D^c))\,.\quad\square
\end{aligned}$$

Lemma 13.2. *Let $(s_0,\dots,s_{2r+1})$ be a saddle path. For all k in $\{1,\dots,r_*\}$ we have*

$$\begin{aligned}
U(s_{2k})\,+\,V(s_{2k},s_{2k+1})\,-\,H(\pi(s_{2k},D\setminus G))\,-\,U(\pi(s_{2k},D\setminus G))&\\
\geq\,\max(U(s_{2k}),U(s_{2k+1}))\,-\,\max(U(s_{2k-2}),U(s_{2k-1})).&
\end{aligned}$$

Proof. Since s_{2k-2} does not belong to $\pi(s_{2k-1}, D \setminus G)$ (which coincides with $\pi(s_{2k}, D \setminus G)$) and $V(s_{2k-2}, s_{2k-1})$ is finite (which implies by reversibility that $V(s_{2k-1}, s_{2k-2})$ is also finite), then s_{2k-2} belongs to the boundary of $\pi(s_{2k}, D \setminus G)$. Either $\pi(s_{2k}, D \setminus G)$ is reduced to $\{s_{2k}\}$ and $H(\pi(s_{2k}, D \setminus G)) = 0$, in which case the inequality of the lemma is satisfied, or $H(\pi(s_{2k}, D \setminus G)) > 0$; in this last case, we have

$$U(s_{2k-2}) \;\geq\; H(\pi(s_{2k}, D \setminus G)) \;+\; U(\pi(s_{2k}, D \setminus G))$$

which gives the desired inequality. $\square$

Theorem 13.3. *(reversibility of the ascending part)*
Let $(s_0, \ldots, s_{2r+1})$ be a saddle path of $\mathcal{S}(D, G)$ of null $V_{s_0}(D, G)$–cost.
For any k in $\{1, \ldots, r_*\}$, the point s_{2k-2} belongs to the principal boundary of $\pi(s_{2k-1}, D \setminus G)$. Equivalently, the ascending part $(s_0, \ldots, s_{2r_*})$ of the saddle path $(s_0, \ldots, s_{2r+1})$ is the descending part of a saddle path of $\mathcal{S}(G^c, \pi(s_0, D)^c)$ of null $V_{s_{2r_*+1}}(G^c, \pi(s_0, D)^c)$–cost.

Remark. For $s \in \pi(s_0, D)$ we have $\pi(s, D \setminus G) = \pi(s, \pi(s_0, D) \setminus G) = \pi(s, G^c \setminus \pi(s_0, D)^c)$.

Proof. Let $(s_0, \ldots, s_{2r+1})$ be a saddle path of null V_{s_0}–cost. We have by lemma 13.2.

$$\begin{aligned} V(s_0, s_1) \;+ \sum_{1 \leq k \leq r_*} &\Big(U(s_{2k}) + V(s_{2k}, s_{2k+1}) - H(\pi(s_{2k}, D \setminus G)) - U(\pi(s_{2k}, D \setminus G)) \Big) \\ &\geq\; V(s_0, s_1) \;+ \sum_{1 \leq k \leq r_*} \max(U(s_{2k}), U(s_{2k+1})) \;-\; \max(U(s_{2k-2}), U(s_{2k-1})) \\ &\geq\; U(s_{2r_*+1}) \;-\; U(s_0). \end{aligned}$$

Now s_{2r_*+1} belongs to the principal boundary of $\pi(s_0, D)$ whence

$$U(s_{2r_*+1}) - U(s_0) = U(\pi(s_0, D)) + H(\pi(s_0, D)) - U(s_0)$$

which is also the value of the initial sum in the above formula (a necessary condition for $(s_0, \ldots, s_{2r+1})$ to be a saddle path of null V_{s_0}–cost, see the second condition given in theorem 9.1.). It follows that all the intermediate inequalities (given by lemma 13.2.) are in fact equalities i.e.

$$\forall k \in \{1, \ldots, r_*\} \quad U(\pi(s_{2k-1}, D \setminus G)) + H(\pi(s_{2k-1}, D \setminus G)) \;=\; \max(U(s_{2k-2}), U(s_{2k-1})).$$

This equality is equivalent to the fact that s_{2k-2} belongs to the principal boundary of $\pi(s_{2k-1}, D \setminus G)$. $\square$

Theorem 13.4. *(global reversibility of saddle paths of null cost between two cycles) Let π_1 and π_2 be two disjoint cycles of E and set $G = \pi_1$, $D = \pi_1^c \cap \pi_2^c$. The reversing operator $\mathcal{R}$ maps the saddle paths of null cost of $\mathcal{S}(D,G)$ onto the saddle paths of null cost of $\mathcal{S}(G^c, D^c)$; it leaves the ridge part invariant and exchanges the ascending and the descending part. More precisely, let $(s_0, \dots, s_{2r+1})$ be a saddle path of $\mathcal{S}(D,G)$ of null V_{s_0}–cost. Then $\mathcal{R}(s_0, \dots, s_{2r+1})$ is a saddle path of $\mathcal{S}(G^c, D^c)$ of null $V_{s_{2r+1}}$–cost. Its ascending part is $(s_{2r+1}, \dots, s_{2r^*+1})$, its ridge part is $(s_{2r^*}, \dots, s_{2r_*+1})$ and its descending part $(s_{2r_*}, \dots, s_0)$.*

Proof. Let $(s_0, \dots, s_{2r+1})$ be a saddle path of null V_{s_0}–cost. Proposition 13.1. shows that $\mathcal{R}(s_0, \dots, s_{2r+1})$ belongs to $\mathcal{S}(G^c, D^c)$. We have

$$\sum_{1 \le k \le r} \Big(U(s_{2k}) + V(s_{2k}, s_{2k+1}) - H(\pi(s_{2k}, D \setminus G)) - U(\pi(s_{2k}, D \setminus G)) \Big)$$
$$= \sum_{1 \le k \le r} \Big(\max\big(U(s_{2k-1}), U(s_{2k-2})\big) - H(\pi(s_{2k-1}, D \setminus G)) - U(\pi(s_{2k-1}, D \setminus G)) \Big)$$
$$- \max(U(s_1), U(s_0)) + \max(U(s_{2r+1}), U(s_{2r}))$$

so that

$$V_{s_0}(s_0, \dots, s_{2r+1}) - V_{s_{2r+1}}(\mathcal{R}(s_0, \dots, s_{2r+1})) = V(s_0, s_1) - V(s_{2r+1}, s_{2r})$$
$$- \max(U(s_1), U(s_0)) + \max(U(s_{2r+1}), U(s_{2r}) + W_D(s_0, s_0) - W_{G^c}(s_{2r+1}, s_{2r+1})$$
$$= U(\pi(s_{2r+1}, G^c)) + H(\pi(s_{2r+1}, G^c)) - U(\pi(s_0, D)) - H(\pi(s_0, D))$$

(where we have used lemma 4.7. and proposition 4.14. to compute W_D and W_{G^c}). However the cycles $\pi(s_0, D)$ and $\pi(s_{2r+1}, G^c)$ do not depend on the points s_0, s_{2r+1} in $\pi_1 \times \pi_2$: these cycles are two disjoint cycles which are maximal proper subcycles of the smallest cycle π containing both π_1 and π_2. As a consequence, the quantities $U(\pi(s_0, D)) + H(\pi(s_0, D))$ and $U(\pi(s_{2r+1}, G^c)) + H(\pi(s_{2r+1}, G^c))$ are equal to the level of π whence

$$V_{s_0}(s_0, \dots, s_{2r+1}) = V_{s_{2r+1}}(\mathcal{R}(s_0, \dots, s_{2r+1})).$$

Thus $\mathcal{R}(s_0, \dots, s_{2r+1})$ has a null $V_{s_{2r+1}}$–cost. The assertion concerning the three parts of $\mathcal{R}(s_0, \dots, s_{2r+1})$ is a straightforward consequence of the definition (proposition 9.3.). $\square$

Corollary 13.5. *Let π_1 and π_2 be two disjoint cycles of E. The set of global saddles (respectively the set of global saddle points) between π_1 and π_2 coincides with the set of global saddles (resp. the set of global saddle points) between π_2 and π_1.*

References

1. L. Alonso–R. Cerf, *The three dimensional polyominoes of minimal area*, preprint (1995).
2. G. Ben Arous–R. Cerf, *Metastability of the three dimensional Ising model on a torus at very low temperatures*, preprint (1995).
3. M. Cassandro–A. Galves–E. Olivieri–M.E. Vares, *Metastable behaviour of stochastic dynamics: a pathwise approach*, Jour. Stat. Phys. **35 nos. 5/6** (1984), 603–634.
4. O. Catoni, *Large deviations for annealing*, PhD Thesis, University Paris XI (1990).
5. O. Catoni, *Notes de cours de DEA* (1995), Université d'Orsay.
6. O. Catoni, *Rough large deviations estimates for simulated annealing. Application to exponential schedules*, The Annals of Probability **20** (1992), 1109–1146.
7. O. Catoni, *Sharp large deviations estimates for simulated annealing algorithms*, Ann. Inst. Henri Poincaré **Vol. 27 no. 3** (1991), 291–383.
8. P. Diaconis–D. Stroock, *Geometric bounds for eigenvalues of Markov chains*, The Annals of Applied Probability **Vol. 1 no. 1** (1991), 36–61.
9. J.D. Deuschel–C. Mazza, *L^2 convergence of time nonhomogeneous Markov processes: I. Spectral Estimates*, The Annals of Applied Probability **Vol. 4 no. 4** (1994), 1012–1056.
10. M.I. Freidlin–A.D. Wentzell, *Random perturbations of dynamical systems*, Springer–Verlag, New York, 1984.
11. C.R. Hwang–S.J. Sheu, *Singular perturbed Markov chains and exact behaviour of simulated annealing processes*, J. Theoret. Prob. **Vol. 5 no. 2** (1992), 223–249.
12. R. Kotecky–E. Olivieri, *Droplet dynamics for asymmetric Ising model*, Jour. Stat. Phys. **70 nos. 5/6** (1993), 1121–1148.
13. R. Kotecky–E. Olivieri, *Shapes of growing droplets – a model of escape from a metastable phase*, Jour. Stat. Phys. **75 nos 3/4** (1994), 409–506.
14. E.J. Neves–R.H. Schonmann, *Critical droplets and metastability for a Glauber dynamics at very low temperatures*, Commun. Math. Phys. **137** (1991), 209–230.
15. E.J. Neves–R.H. Schonmann, *Behaviour of droplets for a class of Glauber dynamics at very low temperatures*, Prob. Th. Related Fields **91** (1992), 331–354.
16. E. Olivieri–E. Scoppola, *Markov chains with exponentially small transition probabilities: first exit problem from a general domain* (1994).
17. R.H. Schonmann, *The pattern of escape from metastability of a stochastic Ising model*, Commun. Math. Phys **147** (1992), 231–240.
18. E. Scoppola, *Renormalization group for Markov chains: a general procedure based on renormalization group ideas*, Jour. Stat. Phys. **73 nos. 1/2** (1993), 83–121.
19. A. Trouvé, *Convergence optimale pour les algorithmes de recuits généralisés*, C. R. Acad. Sci. Paris **t.315 Série I** (1992), 1197–1202.
20. A. Trouvé, *Cycle decompositions and simulated annealing*, SIAM J. Control Optimization, (1995) (to appear).
21. A. Trouvé, *Parallélisation massive du recuit simulé*, PhD Thesis, University Paris XI (1993).
22. A. Trouvé, *Rough large deviation estimates for the optimal convergence speed exponent of generalized simulated annealing algorithms*, Ann. Inst. H. Poincaré, (1995) (to appear).

O. Catoni, DIAM, Laboratoire de Mathématiques de l'Ecole Normale Supérieure, UA 762 du CNRS, 45 rue d'Ulm, 75230 Paris Cedex 05–France

R. Cerf, Université Paris Sud, Mathématique, Bâtiment 425, Modélisation stochastique et statistique, URA D 0743 du CNRS, 91405 Orsay Cedex–France

E-mail address: catoni@dmi.ens.fr – cerf@stats.matups.fr