THÈSE

présentée en vue de l'obtention du

DOCTORAT DE L'ECOLE NORMALE SUPERIEURE DE CACHAN

Spécialité : Probabilités

par

Pierre Tarrès

sous la direction de

Michel Benaïm

Pièges des algorithmes stochastiques et marches aléatoires renforcées par sommets.

Soutenue le 24 Octobre 2001 devant le jury composé des Professeurs :

M.	Michel Benaïm	Université de Cergy-Pontoise	Directeur
Mme.	Marie Duflo	Université de Marne la Vallée	Examinateur
M.	Yves Le Jan	Université Paris-Sud	Examinateur
M.	Yves Meyer	Ecole Normale Supérieure de Cachan	Examinateur

Au vu des rapports des Professeurs :

Mme.	Marie Duflo	Université de Marne la Vallée
M.	Robin Pemantle	Ohio State University

Centre de Mathématiques et Leurs Applications
UMR CNRS 8536, Ecole Normale Supérieure de Cachan

Remerciements

Je tiens à exprimer ici ma grande reconnaissance à Michel Benaïm, qui a accepté de diriger cette thèse de doctorat. Son enseignement, sa disponibilité et la richesse des sujets qu'il m'a proposés m'ont permis de beaucoup progresser au cours de ces trois années. Il m'a conduit à prendre confiance dans les idées que je lui présentais jusqu'à me laisser une très grande autonomie, ce qui a été une grande chance pour moi.

Je veux remercier Marie Duflo et Robin Pemantle de m'avoir fait l'honneur d'accepter d'être rapporteurs de cette thèse. Leurs encouragements chaleureux au sujet de mes résultats ont contribué à l'enthousiasme avec lequel j'ai travaillé. Plus récemment, les conseils de Marie Duflo sur les estimations m'ont été particulièrement utiles pour rédiger une nouvelle version de la preuve de la conjecture de Pemantle et Volkov.

Je suis particulièrement heureux qu'Yves Le Jan ait accepté de faire partie de ce jury. Je le remercie de m'avoir indiqué sur le sujet des marches renforcées des directions nouvelles et d'autres angles d'attaque qui me seront très utiles.

J'ai eu le privilège de faire la connaissance d'Yves Meyer lors de mon mémoire de DEA, qui concernait un problème d'analyse d'image en biologie moléculaire et était donc lié à certains de ses travaux. Je lui suis profondément reconnaissant d'avoir continué depuis à s'intéresser à ma recherche et d'être le président de ce jury, à la mesure de l'admiration que je lui porte.

Je tiens également beaucoup à remercier Gérard Ben Arous de s'être intéressé à mes travaux. J'ai la chance de travailler prochainement sous sa direction, et d'avoir reçu de sa part des perspectives de recherche passionnantes.

J'ai beaucoup bénéficié des conseils de Jean-Michel Morel, qui a bien voulu m'accueillir au CMLA dont il est le directeur. Je le remercie aussi pour les discussions très encourageantes, et pour sa grande disponibilité.

C'est à Robert Azencott que je dois, par la variété des cours du DEA qu'il dirigeait et par les conseils qu'il m'a donnés, d'avoir pu choisir judicieusement le domaine dans lequel j'ai travaillé. Plusieurs entretiens avec lui ont eu une importance déterminante pour moi, et j'en éprouve une profonde gratitude.

Je voudrais enfin remercier toutes les personnes avec lesquelles j'ai pu avoir des contacts souvent très stimulants au cours de cette thèse ; je citerais par exemple Jean-Baptiste Bardet, Sébastien Blachère, Gilles Blanchard, Bruno Blanchet, Odile Brandière, Jean Couellier, Bruno Fabre, Samuel Herrmann, Saïma Khenissy, Grégoire Lecerf, Yannick Le Guilcher, Philippe Marchal, Sébastien Schreiber, Michel Sortais, Françoise Veretout, mais cette liste est loin d'être exhaustive.

Résumé

Nous étudions dans une première partie des algorithmes stochastiques à temps discret, s'écrivant localement comme une approximation des solutions d'une équation différentielle ordinaire. Nous rappelons des résultats de Benaïm et Hirsch, montrant sous des hypothèses faibles que les ensembles limites de ces algorithmes appartiennent à une catégorie d'ensembles caractéristiques du système dynamique associé appelés ensembles intérieurement transitifs par chaînes. Il est naturel de se demander quels ensembles, parmi ces ensembles caractéristiques, sont effectivement atteints asymptotiquement par l'algorithme avec une probabilité non nulle.

Les attracteurs sont effectivement rejoints avec probabilité non nulle. Nous nous sommes intéressés dans cette thèse au cas des ensembles répulsifs, qui contiennent dans chacun de leurs voisinages des trajectoires ne les ayant pas comme ensembles limites. L'hypothèse des points linéairement instables a été beaucoup étudiée, en particulier par Pemantle (1990), Brandière et Duflo (1996), et Benaïm (1997). Nous montrons dans le cas unidimensionnel que tous les types de points instables (c'est à dire tels que la fonction sous-jacente soit positive à droite et négative à gauche) sont évités lorsque le bruit est assez excitant et les pas sont déterministes. Ce résultat nous permet, dans le cas plus général de la dimension finie, d'élargir le cadre de théorèmes montrant sous certaines hypothèses qu'on évite presque-sûrement les ensembles normalement hyperboliques, et en particulier les points linéairement instables et les orbites périodiques hyperboliques linéairement instables.

La condition d'excitation du bruit est importante dans ces résultats : nous obtenons, dans le cas de l'algorithme de Narendra sur le bandit à deux bras, que le point instable de la dynamique associée (qui correspond au choix du mauvais bras) est atteint avec une probabilité non nulle pour des pas $\gamma_n = c/n^\gamma$, $c \in]0,1[$, $\alpha < 1$, ce résultat étant dû à la faiblesse et l'irrégularité du bruit. Par contre cet algorithme converge bien avec probabilité 1 vers le choix du bon bras pour des pas $\gamma_n = c/n$, $c \in]0,1[$.

Nous avons étudié, en collaboration avec Benaïm et Schreiber, une modélisation des dynamiques génétiques par des urnes de Pólya généralisées. Les processus liés au problème sont alors des approximations bruitées de solutions d'équations différentielles. Mais les pas liés à l'évolution du système sont aléatoires et liés

à la taille de la population, et les théorèmes précédents ne peuvent s'appliquer directement. Il est nécessaire de prendre en compte les zones de croissance et de décroissance de la population pour identifier les états limites possibles du système, l'extinction étant envisageable.

La deuxième partie concerne les marches aléatoires renforcées par sommets, qui sont des processus évoluant dans un environnement qu'ils contribuent à faire changer par le fait qu'ils ont une probabilité plus grande de revenir aux endroits déjà visités.

Pemantle et Volkov ont montré que, lorsque le graphe sous-jacent est $\mathbb{Z}$, la marche reste presque-sûrement bloquée dans un ensemble fini de points et, avec une probabilité strictement positive, dans un ensemble de cinq points consécutifs. Ils ont conjecturé que ce deuxième événement se réalisait avec probabilité un.

Nous apportons la preuve de cette conjecture. Le plan général de cette démonstration est le suivant. Nous commençons par étudier la dynamique liée aux mesures d'occupations sur les points qui suivent immédiatement le premier point visité un nombre infini de fois. Cette étude préliminaire conduit à se poser la question de la probabilité d'événements apparaissant comme "instables". Nous ne pouvons pas utiliser les résultats précédents parce que nous n'avons pas de description dynamique des mesures d'occupation sur l'ensemble des points visités un nombre infini de fois par la marche. Cependant le bruit généré par cette marche est assez régulier pour que nous puissions appliquer l'heuristique de résultats sur la non-convergence vers les points instables, en utilisant un ordre partiel sur ce type de marches aléatoires rendu possible par l'aspect unidimensionnel du problème.

Nous donnons enfin une nouvelle preuve des résultats obtenus par Pemantle et Volkov, en utilisant en partie certaines des techniques développées pour la preuve de la conjecture.

Table des matières

Chapitre 1

Introduction

1.1 Algorithmes stochastiques

1.1.1 Introduction et cadre général

Un algorithme stochastique est un processus à temps discret $(x_n)_{n\in\mathbb{N}}$ prenant ses valeurs dans un espace euclidien et évoluant à chaque instant n suivant une formule itérative du type

$$x_{n+1} - x_n = \gamma_{n+1} V_{n+1},$$

où γ_{n+1} est le pas à l'instant $n+1$, et V_{n+1} est une variable aléatoire.

La valeur x_n peut par exemple caractériser l'état d'un système à l'instant n dont l'évolution à l'instant $n+1$ dépend de son propre état et d'un élément extérieur ξ_{n+1}, c'est à dire

$$V_{n+1} = f(x_n, \xi_{n+1}).$$

Les principaux résultats présentés ici sur les algorithmes stochastiques se placent dans l'hypothèse où l'évolution est essentiellement une version bruitée d'une dynamique inhérente au système étudié, de sorte que

$$V_{n+1} = F(x_n) + a_n(\epsilon_{n+1} + r_{n+1}),$$

où ϵ_{n+1} est un bruit, c'est à dire une variable aléatoire d'espérance nulle, et r_{n+1} est petit.

Plus précisément, considérons un espace de probabilités $(\Omega, \mathcal{F}, P)$ muni d'une filtration $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$. Nous dirons ici qu'un processus $(x_n)_{n\in\mathbb{N}}$ adapté à la filtration $\mathbb{F}$ est un algorithme stochastique à valeurs dans $\mathbb{R}^m$ $(m \geq 1)$ si les conditions suivantes **H-1** sont satisfaites :

$$x_{n+1} - x_n = \gamma_{n+1} F(x_n) + c_{n+1}(\epsilon_{n+1} + r_{n+1}) \tag{1.1}$$

où

- $F : \mathbb{R}^m \to \mathbb{R}^m$ est borélienne
- $(\gamma_n)_{n\in\mathbb{N}^*}$, $(c_n)_{n\in\mathbb{N}^*}$ sont des suites déterministes à valeurs dans $\mathbb{R}_+$, telles que (c_n) a une infinité de termes non nuls,
- $(\epsilon_n)_{n\in\mathbb{N}^*}$, $(r_n)_{n\in\mathbb{N}^*}$ sont des suites à valeurs dans $\mathbb{R}^m$ adaptées à la filtration $\mathbb{F}$ et telles que
$$E(\epsilon_{n+1} \mid \mathcal{F}_n) = 0, \ \lim_{n\to\infty} r_n = 0.$$

Nous supposons dans les sections 1.1.1 à 1.1.4 que l'hypothèse **H-1** est vérifiée.

L'étude des algorithmes stochastiques a commencé dans les années 1950 avec les travaux de Robbins et Monro (1951,[42]) et Kiefer et Wolfowitz (1952,[29]), et a été l'objet de nombreux travaux dans le cadre de questions liées au traitement du signal, au contrôle adaptatif ([31],[33],[34]) et aux estimations récursives ([35]). Cette étude est fortement liée à des problèmes d'apprentissage, en particulier à l'étude théorique des réseaux de neurones ([25],[50]) et des algorithmes de recuit simulé ([23]). Plus récemment, ce formalisme a trouvé des applications à des problèmes liés à la théorie des jeux ([10],[27]), à la modélisation des phénomènes d'épidémie ([9]), également à l'étude des marches aléatoires renforcées ([4],[39]) et de la dynamique des populations ([44]). Ces deux dernières applications sont envisagées dans cette thèse.

Donnons deux exemples pour commencer.

Exemple 1.1.1 Bandit à deux bras.

Un bandit à deux bras est une machine à sous dotée de deux bras A et B ; en appuyant sur A (resp. B), on gagne un franc avec probabilité θ_A (resp. θ_B) et rien sinon.

On peut chercher un algorithme qui détermine le meilleur bras à jouer (celui pour lequel la probabilité de gagner est la plus forte) sans trop s'attarder sur le moins bon bras.

L'algorithme de Narendra a pour but de répondre à cette question, et est défini de la manière suivante : à chaque instant $n \in \mathbb{N}^*$, je joue le bras A avec probabilité X_n, et le bras B avec probabilité $1-X_n$. Initialement $X_1 = x_1 \in]0,1[$; à l'instant n, je détermine X_{n+1} de la manière suivante :

$$X_{n+1} = \begin{cases} (1-\gamma_n)X_n + \gamma_n \text{ si j'ai joué } A \text{ et } A \text{ est gagnant} \\ (1-\gamma_n)X_n \text{ si j'ai joué } B \text{ et } B \text{ est gagnant} \\ X_n \text{ sinon,} \end{cases}$$

étant donnée une suite déterministe $(\gamma_n)_{n\in\mathbb{N}^*}$. On obtient en résumé

$$X_{n+1} - X_n = \begin{cases} \gamma_n(1-X_n) \text{ avec probabilité } X_n.\theta_A \\ -\gamma_n X_n \text{ avec probabilité } (1-X_n).\theta_B \\ 0 \text{ avec probabilité } 1-(X_n\theta_A + (1-X_n)\theta_B), \end{cases}$$

ces probabilités étant entendues conditionnellement à $\mathcal{F}_n$.

Une méthode pour étudier ce processus est d'observer que

$$E(X_{n+1}-X_n \mid \mathcal{F}_n) = \gamma_n(1-X_n).X_n\theta_A - \gamma_n X_n.(1-X_n)\theta_B = \gamma_n X_n(1-X_n)(\theta_A-\theta_B).$$

Il semble ainsi naturel de comparer l'évolution du processus $(X_n)_{n\in\mathbb{N}}$ avec les solutions de l'équation différentielle ordinaire

$$\frac{dx}{dt} = x(1-x)(\theta_A-\theta_B),$$

puisque le bruit s'ajoutant à chaque étape à $E(X_{n+1}-X_n \mid \mathcal{F}_n)$ a une influence négligeable sur l'évolution de X_n sur des intervalles de temps longs. Cet argument permet d'affirmer (cf. théorèmes 1.1.1 et 1.1.3 plus bas) que X_n tend vers 0 ou 1 si $\sum_n \gamma_n = +\infty$ et, pour tout $c > 0$, $\sum_n e^{-c/\gamma_n} < +\infty$.

Une telle heuristique pourrait nous amener à penser que, dans la mesure où toutes les solutions de cette équation différentielle ne partant pas de 0 tendent vers 1 si $\theta_A > \theta_B$, la suite X_n tend alors vers 1 sous cette hypothèse, ce qui serait le résultat souhaitable pour cet algorithme.

Les méthodes de comparaison avec les solutions de l'équation différentielle ne permettent en fait pas de déduire un tel résultat. Dans ce cas précis, nous avons montré que cette affirmation est fausse pour des pas $\gamma_n = c/n^\alpha$, $\alpha \in]0,1[$, et vraie pour $\gamma_n = c/n$ si $c \in]0,1[$ (voir chapitre 3).

□

Exemple 1.1.2 Réseaux de neurones.

Le principe des algorithmes d'apprentissage en réseaux de neurones peut se résumer de la manière suivante. On se donne au départ une fonction

$$\begin{array}{rcl} \phi : \mathbb{R}^m \times \mathbb{R}^k & \longrightarrow & \mathbb{R} \\ (\omega, x) & \longmapsto & \phi(\omega, x). \end{array}$$

On cherche alors une méthode qui permette de déterminer à partir d'un ensemble de données $(x_k, y_k)_{k\in E}$ (E fini) un vecteur de poids $\omega \in \mathbb{R}^m$ assurant la prédiction des sorties y à partir des entrées x par $\phi(\omega, x)$ avec le plus de précision possible, autrement dit qui minimise les différents écarts $\|y_k - \phi(\omega, x_k)\|$.

Formellement, considérons une suite de variables aléatoires i.i.d $(x_n, y_n) \in \mathbb{R}^k \times \mathbb{R}$, $n \in \mathbb{N}^*$, de loi μ. Un algorithme supervisé en réseaux de neurones $(\omega_n)_{n\in\mathbb{N}}$ est défini par une règle du type suivant : ω_0 fixé et, pour tout $n \in \mathbb{N}$,

$$\omega_{n+1} - \omega_n = -\gamma_{n+1} \frac{\delta C}{\delta \omega}(\omega_n, x_{n+1}, y_{n+1})$$

où

$$C(\omega, x, y) = (y - \phi(\omega, x))^2,$$

étant donnée une suite déterministe $(\gamma_n)_{n\in\mathbb{N}^*}$. On effectue donc à chaque itération une descente de gradient basée sur la nouvelle information (x_{n+1}, y_{n+1}).

Ceci est rendu possible par la méthode de rétropropagation du gradient, qui permet de calculer directement

$$e(\omega, x, y) = \frac{\delta C}{\delta \omega}(\omega, x, y) = -2(y - \phi(\omega, x)) \frac{\delta \phi}{\delta \omega}(\omega, x).$$

Remarquons que, pour tout $n \in \mathbb{N}$,

$$E[\omega_{n+1} - \omega_n \mid \omega_n = \omega] = -\gamma_{n+1} E[e(\omega, x_1, y_1)] = -\gamma_{n+1} \frac{\delta}{\delta \omega}[E(C(\omega, x_1, y_1))].$$

Ceci justifie de chercher à comparer l'évolution du processus $(\omega_n)_{n\in\mathbb{N}}$ aux solutions de l'équation différentielle ordinaire

$$\frac{d\omega}{dt} = -\frac{\delta}{\delta\omega}E[C(\omega, x_1, y_1)] = -\nabla\overline{C}(\omega)$$

avec

$$\overline{C}(\omega) = E[C(\omega, x_1, y_1)] = E[(y_1 - \phi(\omega, x_1))^2].$$

Concrétement, on travaille sur une base d'apprentissage finie $(x_n, y_n)_{n=1,\ldots,N}$. Si

$$\mu = \frac{1}{N}\sum_{i=1}^{N}\delta_{(x_i,y_i)}$$

où $\delta_{(x,y)}$ désigne la mesure de Dirac en (x, y), alors

$$\overline{C}(\omega) = \frac{1}{N}\sum_{i=1}^{N}[y_i - \phi(\omega, x_i)]^2.$$

Une tel algorithme n'assure pas a priori une convergence vers le meilleur choix de ω, c'est à dire celui qui minimise $\overline{C}(\omega)$. Les résultats qui suivent permettent de mieux en comprendre le comportement asymptotique.

Ainsi, nous pouvons par exemple appliquer les théorèmes 1.1.1 et 1.1.2 (Benaïm et Hirsch) pour montrer que, génériquement, l'algorithme converge vers un point critique de $\overline{C}$. Le théorème 1.1.3 (Benaïm, Duflo) prouve que tous les minimas locaux sont atteints avec probabilité strictement positive. Ce qui implique qu'en général la probabilité de convergence vers le minimum global de $\overline{C}$ est différente de 1. Enfin les résultats de non-convergence vers les pièges obtenus par Pemantle, Brandière et Duflo, Benaïm, et ceux présentés au cours de cette thèse dans les chapitres 2 et 3, permettent de montrer que les maximas locaux et les points cols sont p.s évités.

Nous donnons cet exemple pour montrer la variété des applications des algorithmes stochastiques. Cependant l'étude de ce type d'algorithmes d'apprentissage par réseaux de neurones a donné lieu à un grand nombre de résultats (voir par exemple [23], [25], [50]).

□

1.1.2 Résultats de convergence vers certains ensembles caractéristiques du système dynamique associé

Un algorithme stochastique (1.1) peut être considéré comme un algorithme de Cauchy-Euler donnant une approximation en temps discret des solutions de l'équation différentielle ordinaire (EDO)

$$\frac{dx}{dt} = F(x), \tag{1.2}$$

ce qui permet d'espérer lier son comportement asymptotique à celui de cette EDO (1.2), comme nous l'avons remarqué au cours des exemples précédents. Cet argument appelé méthode de l'EDO a été introduit par Ljung en 1977 ([32]) et a inspiré un grand nombre de travaux (voir par exemple [12],[23],[30] et [31]).

Nous présentons ici le principe de résultats obtenus par Benaïm et Hirsch en 1996 ([3],[5],[6],[8]) montrant que les ensembles limites de $(X_n)_{n\in\mathbb{N}}$ font partie d'un certain type d'ensembles caractéristiques du système dynamique associé à l'EDO, appelés ensembles intérieurement transitifs par chaînes (I.C.T).

L'intérêt de ces résultats a été d'expliquer ce lien entre l'algorithme et son EDO associée sans se limiter à des dynamiques simples pour F, ce qui était essentiellement supposé précédemment (par exemple des dynamiques de descente de gradient, c'est à dire telles que F soit l'opposé du gradient d'une fonction de coût, comme dans le cas de l'exemple 1.1.2).

Nous supposerons dans la suite que F est bornée et lipschitzienne. Introduisons quelques définitions.

Tout d'abord, cette dernière hypothèse implique que le champ de vecteurs F a d'uniques courbes intégrales. Nous définissons alors le flot Φ associé à F

$$\begin{array}{rcl}\Phi : \mathbb{R}\times\mathbb{R}^m & \longrightarrow & \mathbb{R}^m\\ (t,x) & \longmapsto & \Phi(t,x)=\Phi_t(x).\end{array}$$

Remarquons que, pour tous $t,\ s\in\mathbb{R}^2$,

$$\Phi_0 = Id\ ,\ \ \Phi_{t+s}=\Phi_t\circ\Phi_s.$$

Pour toute fonction $Y:\mathbb{R}_+\to\mathbb{R}^m$, nous notons

$$L(Y)=\cap_{t\geq 0}\overline{Y([t,+\infty[)}$$

son ensemble limite.

Une fonction $Y:\mathbb{R}_+\to\mathbb{R}^m$ est appelée pseudo-trajectoire asymptotique (P.T.A) de Φ si et seulement si, pour tout $T>0$,

$$\lim_{t\to+\infty}\sup_{0\leq h\leq T} d(Y(t+h),\Phi_h(Y(t)))=0.$$

Ce qui signifie que, pour tout $T>0$ fixé, la courbe

$$[0,T]\to\mathbb{R}^m:\ \ h\mapsto Y(t+h)$$

suit la solution de l'EDO partant du point $Y(t)$ sur l'intervalle $[0,T]$ d'aussi près que nécessaire pour t assez grand.

Posons par ailleurs

$$\tau_0=0\ ,\ \ \tau_n=\sum_{i=1}^n\gamma_i \text{ pour } n\geq 1,$$

et définissons la fonction "inverse" $m : \mathbb{R}_+ \to \mathbb{N}$ de $n \mapsto \tau_n$, par

$$m(t) = \sup\{k \geq 0 : \ t \geq \tau_k\}.$$

Considérons le processus en temps continu $X(\cdot)$ associé à X_n par

$$X(t) = X_n \text{ pour } \tau_n \leq t < \tau_{n+1}.$$

Nous disons qu'un processus $(\epsilon_n)_{n\in\mathbb{N}}$ est sous-gaussien s'il existe un réel strictement positif Γ tel que, pour tout $\theta \in \mathbb{R}^m$,

$$E(\exp(<\theta, \epsilon_{n+1}>) \mid \mathcal{F}_n) \leq \exp(\frac{\Gamma}{2}\|\theta\|^2).$$

Alors nous avons le résultat suivant :

Théorème 1.1.1 *(Benaïm,[3] et [5]).*
Supposons $\gamma_n = c_n$, $\sum_n \gamma_n = +\infty$ et

[H-2.1] *il existe $q \geq 2$ tel que*

$$(a1) \qquad \sum_n \gamma_n^{1+q/2} < +\infty$$
$$(a2) \qquad \sup_n E(\|\epsilon_{n+1}\|^q) < +\infty,$$

ou

[H-2.2]

$$(b1) \qquad \textit{pour tout } c \geq 0, \ \sum_n e^{-c/\gamma_n} < +\infty$$
$$(b2) \qquad (\epsilon_n)_{n\in\mathbb{N}} \textit{ est sous-gaussien.}$$

Alors le processus en temps continu X associé à $(X_n)_{n\in\mathbb{N}}$ est une pseudo-trajectoire asymptotique (P.T.A) de Φ.

Il nous reste maintenant à comprendre quels ensembles peuvent être ensembles limites de pseudo-trajectoires asymptotiques. Commençons par deux définitions.

Définition 1.1.1 *(Bowen et Conley, [14] et [18]). Soient $\Lambda \in \mathbb{R}^m$ et $a, b \in \Lambda$. Nous écrivons $a \hookrightarrow_\Lambda b$ si et seulement si, pour tous $\delta > 0$, $T > 0$, il existe $n \in \mathbb{N}^*$ et une suite $(x_i)_{0\leq i\leq n} \in \Lambda^{n+1}$ tels que*

$$\begin{aligned} & d(x_0, a) < \delta \\ & \forall i < n, \ \exists t_i > T / \ d(\Phi_{t_i}(x_i), x_{i+1}) < \delta \\ & x_n = b. \end{aligned}$$

Définition 1.1.2 *(Bowen et Conley, [14] et [18]). Un ensemble Λ est dit intérieurement transitif par chaînes (I.C.T) ssi, pour tous a, $b \in \Lambda$, $a \hookrightarrow_\Lambda b$.*

Un ensemble Λ est dit intérieurement récurrent par chaînes (I.C.R) ssi, pour tout $a \in \Lambda$, $a \hookrightarrow_\Lambda a$.

La notion d'ensemble intérieurement transitif par chaînes est bien adaptée à la description des ensembles limites de pseudo-trajectoires asymptotiques, comme le montre le théorème suivant.

Théorème 1.1.2 *(Benaïm et Hirsch, [8]).*

(i) *Soit $Y : \mathbb{R}_+ \to \mathbb{R}^m$ une pseudo-trajectoire asymptotique bornée. Alors $L(Y)$ est intérieurement transitif par chaînes.*

(ii) *Soit $L \subset \mathbb{R}^m$ un ensemble intérieurement transitif par chaînes. Alors il existe une pseudo-trajectoire asymptotique Y telle que $L(Y) = L$.*

Ce résultat conduit à se poser deux questions : en premier lieu, quel type d'ensembles recouvre la notion d'I.C.T ? Enfin, parmi ces ensembles I.C.T, lesquels sont effectivement atteints avec une probabilité strictement positive par l'algorithme ? En effet il y a bien une équivalence entre ensembles I.C.T et ensembles limites de P.T.A, mais les P.T.A possibles du flot associé à F ne sont pas toutes décrites par l'algorithme.

Nous avons besoin des définitions suivantes.

Un ensemble $A \subset \mathbb{R}^m$ est dit invariant ssi $\Phi_t(A) = A$ pour tout $t \in \mathbb{R}$.

Soit $A \subset \mathbb{R}^m$. Notons **(i)** et **(ii)** les conditions suivantes

(i) A est non vide, compact et invariant

(ii) Il existe un voisinage $W \subset \mathbb{R}^m$ de A tel que $\mathsf{dist}(\Phi_t(x), A)$ tend vers 0 quand $t \to \infty$, uniformément par rapport à $x \in W$.

Nous dirons que A est un attracteur si et seulement si **(i)** et **(ii)** sont vérifiées.

Nous dirons que A est un ensemble répulsif si et seulement si **(i)** est vérifiée et **(ii)** n'est pas vérifiée.

Un point $p \in \mathbb{R}^m$ est un équilibre si $F(p) = 0$ (ce qui implique que $\{p\}$ est invariant).

Le résultat suivant, dû à Bowen ([14]) précise la nature des ensembles I.C.T et I.C.R, ce qui est l'objet de la première question.

Théorème 1.1.3 *(Bowen,[14]). Soit $\Lambda \subset \mathbb{R}^m$. Les affirmations suivantes sont équivalentes :*

(i) *Λ est intérieurement transitif par chaînes*

(ii) *Λ est connexe et intérieurement récurrent par chaînes*

(iii) *Λ est un ensemble compact et invariant, et $\Phi|\Lambda$ n'a pas d'attracteur propre.*

Plusieurs résultats répondent partiellement à la deuxième question.

En premier lieu, il est possible de montrer que pour tout attracteur A, la probabilité que l'ensemble limite de l'algorithme soit inclus dans A est strictement positive sous certaines conditions : c'est l'objet de la partie suivante 1.1.3.

En deuxième lieu, certains ensembles répulsifs ne sont presque-sûrement pas atteints lorsque le bruit est assez excitant, le résultat inverse étant possible (convergence avec probabilité non nulle vers l'ensemble) si cette dernière hypothèse n'est pas satisfaite. C'est l'objet de la partie 1.1.4.

1.1.3 Convergence avec probabilité non nulle vers un attracteur

Les résultats de convergence avec probabilité non nulle vers un attracteur présentés ici sont dûs à Benaïm ([5]) et Duflo ([24]). Ils se placent sous l'hypothèse suivante **H-3** : il existe une fonction $w : \mathbb{R}_+^3 \to \mathbb{R}_+$ telle que, pour tous $\delta > 0$ et $T > 0$,

$$P(\sup_{s\geq t}[\sup_{0\leq h\leq T} d(X(s+h), \Phi_h(X(s)))] \geq \delta \mid \mathcal{F}_{\tau_{m(t)}}) \leq w(t,\delta,T)$$

et $\lim_{t\to\infty} w(t,\delta,T) = 0$.

Cette condition **H-3** est satisfaite pour un grand nombre d'algorithmes stochastiques. Une condition suffisante est qu'il existe une fonction $r : \mathbb{R}_+^3 \to \mathbb{R}_+$ telle que

$$P(\sup_{0\leq h\leq T} d(X(t+h), \Phi_h(X(t))) \geq \delta \mid \mathcal{F}_{\tau_{m(t)}}) \leq \int_t^{t+T} r(s,\delta,T)\, ds$$

et

$$\int_0^\infty r(t,\delta,T)\, dt < \infty.$$

Cette dernière condition est toujours vérifiée sous les hypothèses du théorème 1.1.1 (voir [5]).

Définissons l'ensemble $Att(X)$ des points $x \in \mathbb{R}^m$ atteignables par l'algorithme, c'est à dire tels que, à chaque instant $t \geq 0$ et pour tout voisinage U de x,

$$P(\exists s \geq t : \ X(s) \in U) > 0.$$

Alors nous avons le résultat suivant :

Théorème 1.1.4 *(Benaïm, [5], Duflo,[24]) Soit A un attracteur ayant pour bassin d'attraction $B(A)$. Supposons $B(A) \cap Att(X) \neq \emptyset$; alors*

$$P(L(X) \subset A) > 0.$$

Plus précisément, pour tout ouvert $U \subset \mathbb{R}^m$ tel que $\overline{U} \subset B(A)$, il existe δ, $T > 0$ (dépendants de U) tels que

$$P(L(X) \subset A) \geq (1 - w(t,\delta,T))P(\exists s \geq t : \ X(s) \in U).$$

1.1.4 Non-convergence vers certains ensembles répulsifs

Commençons par rappeler certains résultats de non-convergence vers les points linéairement instables (ou pièges) et plus généralement vers les ensembles normalement hyperboliques, obtenus par Pemantle ([38]), Brandière et Duflo ([15],[17]), et Benaïm ([5],[7]).

Introduisons auparavant les définitions suivantes.

Définition 1.1.3 *Soit $p \in \mathbb{R}^m$ un équilibre du système dynamique engendré par F, c'est à dire tel que $F(p) = 0$. Notons $H = DF(p)$. Nous dirons que p est hyperbolique si toutes les valeurs propres de $DF(p)$ ont des parties réelles non nulles. L'équilibre p sera dit linéairement instable si certaines valeurs propres de H ont une partie réelle strictement positive.*

Nous notons $K_+(H)$ (resp. $K_-(H)$) l'espace propre généralisé correspondant aux valeurs propres de H à partie réelle strictement positive (resp. négative ou nulle).

Pour tout $x \in \mathbb{R}^m$, nous notons $x^{(r)}$ la projection de x sur $K_+(H)$ parallèlement à $K_-(H)$.

Définition 1.1.4 *Soit $\alpha \in]0,1]$. Une fonction f sera dite $C^{1+\alpha}$ si f est C^1 et sa dérivée est α-hölderienne. Remarquons que C^{1+1} est plus faible que C^2. Une variété est $C^{1+\alpha}$ si ses fonctions de transition peuvent être choisies $C^{1+\alpha}$.*

Le résultat suivant donne un critère de non-convergence vers les points linéairement instables : lorsque le bruit est assez excitant et régulier, et sous certaines hypothèses les pas, le piège est presque-sûrement évité.

Théorème 1.1.5 *(Brandière et Duflo,[17]). Soit $(x_n)_{n\in\mathbb{N}}$ un algorithme stochastique à valeurs dans $\mathbb{R}^m$. Supposons F C^{1+1}, et soit p un équilibre linéairement instable.*

Nous faisons les hypothèses suivantes sur les perturbations

- $\liminf_{n\to\infty} E(\|\epsilon_{n+1}^{(r)}\| \mid \mathcal{F}_n) > 0,\ \limsup_{n\to\infty} E(\|\epsilon_{n+1}\|^2 \mid \mathcal{F}_n) < \infty,$
- $\sum_{n\in\mathbb{N}} \|r_n\|^2 < \infty,$

et sur les pas

$$\gamma_n = O(c_n),\ \sum_{n\in\mathbb{N}} \gamma_n = \infty,\ \sum_{n\in\mathbb{N}} c_n^2 < \infty.$$

Alors $P(\lim_{n\to\infty} x_n = p) = 0$.

Nous donnons maintenant un résultat de non-convergence vers les ensembles normalement hyperboliques, dû à Benaïm ([5]), utilisant en partie des idées dûes à Pemantle ([39]) pour son aspect probabiliste.

Les ensembles normalement hyperboliques sont des ensembles Γ invariants par Φ et vérifiant les conditions suivantes : $\Gamma \subset S$ où S est une sous-variété de $\mathbb{R}^m$ de dimension $(m-d)$, $d \in \{1, \dots m\}$, localement invariante (il existe un voisinage U de Γ dans $\mathbb{R}^m$ et $t_0 \in \mathbb{R}_*^+$ tels que $\forall t \in \mathbb{R}$, $|t| \leq t_0$, $\Phi_t(U \cap S) \subset S$), enfin $\mathbb{R}^m = T_p S \oplus E_p^u$ pour tout point $p \in \Gamma$ avec

(i) $p \to E_p^u$ continue de Γ dans la variété Grassmannienne $G(d, m)$ des plans de dimension d dans $\mathbb{R}^m$

(ii) $D\Phi_t(p)E_p^u = E_{\Phi_t(p)}^u$ pour tout $t \in \mathbb{R}$, $p \in \Gamma$

(iii) Il existe $\lambda > 0$ et $C > 0$ tels que, pour tout $p \in \Gamma$, $w \in E_p^u$ et $t \geq 0$, $\|D\Phi_t(p)w\| \geq Ce^{\lambda t}\|w\|$.

La notion d'ensemble normalement hyperbolique recouvre celle de point linéairement instable et d'orbite périodique hyperbolique linéairement instable ; nous développons ce point dans les deux exemples suivants.

Exemple 1.1.3 Soit p un équilibre linéairement instable de $F \in C^1$. Alors $\Gamma = \{p\}$ est un ensemble normalement hyperbolique. En effet, reprenons les notations de la définition 1.1.3. Il existe (voir par exemple [43] ou [47]) une sous-variété S localement invariante tangente à $K_-(H)$, qui est C^k si F est C^k. Comme $D\Phi_t(p) = e^{tH(p)}$, il existe $\lambda > 0$ et $c > 0$ tels que $\|D\Phi_t(p)w\| \geq Ce^{\lambda t}\|w\|$, pour tout $w \in K_+(H)$, ce qui permet de conclure avec $E_p^u = K_+(H)$.

□

Exemple 1.1.4 Soit $\Gamma \in \mathbb{R}^m$ une orbite de période $T > 0$ de $F \in C^k$, $k \geq 1$. Pour tout $p \in \Gamma$, $D\Phi_{T+t}(x) = D\Phi_T(\Phi_t(x)) \circ D\Phi_t(x)$ et $D\Phi_{t+T}(x) = D\Phi_t(x) \circ D\Phi_T(x)$, donc

$$D\Phi_T(\Phi_t(x)) = D\Phi_t(x) \circ D\Phi_T(x) \circ (D\Phi_t(x))^{-1}.$$

Par conséquent les valeurs propres $\lambda_1 = e^{\mu_1}$, ..., $\lambda_m = e^{\mu_m}$ (comptées avec leurs multiplicités) de $D\Phi_T(p)$ sont indépendantes de $p \in \Gamma$. Elles sont appelées multiplicateurs de Floquet. L'unité est toujours un multiplicateur de Floquet car, pour tout $p \in \Gamma$, $\Phi_T(\Phi_t(p)) = \Phi_t(p)$ implique $D\Phi_T(p).F(p) = F(p)$ par dérivation en 0, et $F(p)$ est non nul car sinon l'orbite serait réduite à un point.

Γ est dit hyperbolique si 1 est un multiplicateur de multiplicité 1, et les autres multiplicateurs ont des modules différents de 1. Γ est dit linéairement instable si certains multiplicateurs ont un module strictement supérieur à 1.

Supposons que Γ soit une orbite hyperbolique linéairement instable pour le champ de vecteurs F. Alors l'hyperbolicité implique (voir par exemple [47]) qu'il existe C, $\lambda > 0$, et une décomposition de $T_\Gamma\mathbb{R}^m$ comme somme directe de trois faisceaux de vecteurs :

$$T_\Gamma\mathbb{R}^m = E^s(\Gamma) \oplus E^u(\Gamma) \oplus E^\Phi(\Gamma)$$

invariante par $T_\Gamma\Phi$ et telle que pour tout $p \in \Gamma$, $t \geq 0$,

$$\|D\Phi_t(p)|_{E_p^s}\| \leq Ce^{-\lambda t}, \ \|D\Phi_{-t}(p)|_{E_p^u}\| \leq Ce^{-\lambda t}$$

et

$$E_p^\Phi = \text{ Vect } (F(p))$$

où $p \times E_p^u$ désigne la fibre de $E^u(\Gamma)$ en p, et de même pour E_p^s et E_p^Φ.

La dimension de l'espace vectoriel E_p^u (pour tout $p \in \Gamma$) est au moins 1, car Γ est linéairement instable.

La fonction $p \mapsto E_p^u$ est une fonction C^k à valeurs dans la variété Grassmannienne des plans de dimension appropriée.

□

Théorème 1.1.6 *(Benaïm,[5]) Soit $(x_n)_{n\in\mathbb{N}}$ un algorithme stochastique à valeurs dans $\mathbb{R}^m$, tel que pour tout $n \in \mathbb{N}$, $\gamma_n = c_n$ et $r_n = 0$, et soit Γ un ensemble normalement hyperbolique. Supposons*

(i) *$(\epsilon_n)_{n\in\mathbb{N}}$ bornée, c'est à dire $\limsup_{n\to\infty} \|\epsilon_n\| < \infty$*

(ii) *(γ_n) vérifie la condition (b1) du théorème 1.1.1*

(iii) *Il existe un voisinage $\mathcal{N}(\Gamma)$ de Γ et $b > 0$ tels que, pour tout vecteur unité $v \in \mathbb{R}^m$,*

$$E(<\epsilon_{n+1}, v>^+ \mid \mathcal{F}_n) \geq b1_{\{x_n \in \mathcal{N}(\Gamma)\}}.$$

(iv) *Il existe $\alpha \in]1/2, 1]$ tel que*

(a) *F et S sont $C^{1+\alpha}$*

(b)

$$\lim_{n\to\infty} \frac{\gamma_{n+1}^\alpha}{\sqrt{\lim_{m\to\infty} \sum_{i=n+1}^m \gamma_i^2}} = 0.$$

Alors

$$P(\lim_{n\to\infty} d(x_n, \Gamma) = 0) = 0.$$

Nous nous sommes tout d'abord intéressés dans cette thèse au comportement des algorithmes au voisinage de points répulsifs qui ne vérifient pas les conditions précédentes, c'est à dire non linéairement instables, dans le cas de la dimension un. Brandière a obtenu sur ce sujet le résultat que nous énonçons ci-après.

Nous nous plaçons donc dans le cas $m = 1$. Définissons l'hypothèse **H-4.1** sur les perturbations : étant donné $a > 2$,

- $\liminf_{n\to\infty} E(|\epsilon_{n+1}|^2 \mid \mathcal{F}_n) > 0$, $\limsup_{n\to\infty} E(|\epsilon_{n+1}|^a \mid \mathcal{F}_n) < \infty$
- $\sum_{n\in\mathbb{N}} r_n^2 < \infty$

Nous pouvons sans perte de généralité supposer que le point répulsif considéré est en 0.

Théorème 1.1.7 *(Brandière,[16]). Supposons $F(0) = 0$, F de classe C^1 au voisinage de 0, et $F(x) \sim \alpha x^{2k+1}$ avec $\alpha > 0$ et $k \in \mathbb{N}$, $\gamma_n = c_n = g/n$ (avec $g \in \mathbb{R}_*^+$), $r_n = 0$, et* **H-4.1** *avec $a > 2 + k^{-1}$. Alors $P(\lim_{n\to\infty} x_n = 0) = 0$.*

Le théorème que nous avons montré au cours de cette thèse se base sur l'hypothèse suivante **H-4.2** sur la fonction F :

Il existe un voisinage V de 0 tel que F soit positive ou nulle sur $V \cap \mathbb{R}_*^+$ et négative ou nulle sur $V \cap \mathbb{R}_*^-$.

Théorème 1.1.8 *(Chapitre 3) Soit $(x_n)_{n\in\mathbb{N}}$ un algorithme stochastique. Alors, sous les hypothèses* **H-4.1** *et* **H-4.2**, $P(\lim_{n\to\infty} x_n = 0) = 0$.

Nous montrons ce théorème comme une conséquence du résultat suivant :

Théorème 1.1.9 *(Chapitre 3) Soient $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ des variables aléatoires à valeurs dans $\mathbb{R}$ adaptées à la filtration $\mathbb{F}$ et vérifiant*

$$x_{n+1} = y_n + c_{n+1}(\epsilon_{n+1} + r_{n+1}) \ \textit{si} \ X_n \in V,$$

où V est un voisinage de 0 dans $\mathbb{R}$, $(\epsilon_n)_{n\in\mathbb{N}}$, $(r_n)_{n\in\mathbb{N}}$ sont des suites réelles adaptées à la filtration $\mathbb{F}$, $(c_n)_{n\in\mathbb{N}}$ est une suite déterministe à valeurs dans $\mathbb{R}^+$ ayant une infinité de termes non nuls.

Nous faisons l'hypothèse **H-4.1** *et, pour tout $n \in \mathbb{N}$, $|y_n| \geq |x_n|$. Alors $P(\lim_{n\to\infty} x_n = 0) = 0$.*

Le théorème 1.1.9 implique le théorème 1.1.8 en posant, pour tout $n \in \mathbb{N}$,

$$y_n = x_n + \gamma_{n+1} F(x_n).$$

Remarquons que l'hypothèse d'excitation est particulièrement importante : l'algorithme de Narendra sur le bandit à deux bras, présenté dans l'exemple 1.1.1 et détaillé dans le chapitre 4, en donne un contre-exemple. En effet, lorsque $\theta_A > \theta_B$, 0 est un point linéairement instable pour l'algorithme car $F(x) = x(1-x)(\theta_A - \theta_B)$; par contre, pour tout $a > 1$, $E(|\epsilon_{n+1}|^a \mid \mathcal{F}_n) \leq X_n(\theta_A + \theta_B)$ si X_n est assez proche de 0. Donc l'hypothèse **H-4.1** n'est pas vérifiée sur l'événement $\{\lim_{n\to\infty} X_n = 0\}$.

Le principe développé dans les théorèmes 1.1.8 et 1.1.9 nous a permis de prouver le théorème suivant, qui élargit les conditions d'application du théorème 1.1.6.

Notons, pour tout $n \in \mathbb{N}$,

$$V_{n+1} = F(x_n) + \epsilon_{n+1} + r_{n+1}.$$

Théorème 1.1.10 *(Chapitre 2) Soit $(x_n)_{n\in\mathbb{N}}$ un algorithme stochastique à valeurs dans $\mathbb{R}^m$, tels que pour tout $n \in \mathbb{N}$, $\gamma_n = c_n$, et soit Γ un ensemble normalement hyperbolique. Supposons*

(i) *$(\epsilon_n)_{n\in\mathbb{N}}$ bornée, c'est à dire* $\limsup_{n\to\infty} \|\epsilon_n\| < \infty$

(ii) $\sum \|r_n\|^2 < \infty$

(iii) *Il existe un voisinage $\mathcal{N}(\Gamma)$ de Γ et $b > 0$ tel que, pour tout vecteur unité $v \in \mathbb{R}^m$,*

$$E(< V_{n+1}, v >^+ \mid \mathcal{F}_n) \geq b 1_{\{x_n \in \mathcal{N}(\Gamma)\}}.$$

(iv) *Il existe $\alpha \in]1/2, 1]$ tel que F et S soient $C^{1+\alpha}$.*

Alors

$$P(\lim_{n\to\infty} d(x_n, \Gamma) = 0) = 0.$$

Le changement par rapport au théorème 1.1.6 est de ne faire que l'hypothèse $\gamma_n = c_n$ sur les pas, et d'ajouter un terme de perturbation $(r_n)_{n\in\mathbb{N}}$ tel que $\sum \|r_n\|^2 < \infty$.

1.1.5 Modélisation de dynamiques génétiques

Nous avons étudié, en collaboration avec Benaïm et Schreiber, la modélisation de phénomènes d'évolution à partir de modèles d'urnes généralisées : ce travail est présenté dans le chapitre 5. Un processus d'urnes généralisées se définit de la manière suivante. L'urne contient des boules de différentes couleurs, et, à chaque étape, on ajoute ou enlève certaines de ces boules suivant des probabilités ne dépendant que du nombre ou de la distribution des boules dans l'urne à l'instant considéré.

Le lien avec les dynamiques d'évolution s'effectue en associant chaque couleur à un génotype, un phénotype ou une statégie comportementale. Le fait d'ajouter ou d'enlever des boules d'une couleur donnée correspond alors pour les éléments de la population étudiée au fait de se reproduire ou de disparaître.

Nous étudions ainsi une chaîne de Markov homogène

$$(z_n)_{n\in\mathbb{N}} = (z_n^1, \ldots, z_n^k)_{n\in\mathbb{N}}$$

à valeurs dans $\mathbb{N}^k$. Pour tous $i \in \{1, \ldots, k\}$ et $n \in \mathbb{N}$, z_n^i est associé au nombre de boules de couleur i à l'instant n. Nous notons $\Pi : \mathbb{N}^k \times \mathbb{N}^k \mapsto [0, 1]$ le noyau de transition de cette chaîne de Markov ; pour tous z, $z' \in \mathbb{N}^k$,

$$\Pi(z, z') = P[z_{n+1} = z' \mid z_n = z].$$

Posons, pour tout $w = (w^1, \ldots, w^k) \in \mathbb{R}^k$,

$$|w| = |w^1| + \ldots + |w^k| \ , \ \alpha(w) = w^1 + \ldots + w^k,$$

et définissons le $k-1$-simplexe $S_k \subset \mathbb{R}^k$ par

$$S_k = \{x = (x^1, \ldots, x^k) \in \mathbb{R}^k : \ \forall i \in \{1, \ldots, k\}, \ x^i \geq 0, \ \sum_{i=1}^k x^i = 1\}.$$

Nous formulons les hypothèses suivantes **H-5.1** et **H-5.2** sur la chaîne de Markov : à chaque étape, d'une part on n'ajoute ou n'enlève qu'un nombre limité de boules de chaque couleur, et d'autre part les probabilités de transition ne dépendent asymptotiquement (pour un grand nombre de boules) que des proportions des différentes couleurs :

H-5.1 $\exists m \in \mathbb{N} / \ \forall n \in \mathbb{N}, \ |z_{n+1} - z_n| \leq m$

H-5.2 Il existe des fonctions lipschitziennes

$$p_w : \ S_k \to [0,1] : \ w \in \mathbb{Z}^k, \ |w| \leq m$$

et un réel $a > 0$ tels que

$$|p_w(z/|z|) - \Pi(z, z+w)| \leq a/|z|$$

pour tous $z \in \mathbb{N}^k \setminus \{0\}$ et $w \in \mathbb{Z}^k$ tels que $|w| \leq m$.

Dans ce contexte, il est naturel d'associer à z_n le processus x_n défini par

$$x_n = \begin{cases} \frac{z_n}{|z_n|} & \text{si } z_n \neq 0 \\ 0 & \text{si } z_n = 0 \end{cases}$$

Notons que, pour tout $n \in \mathbb{N}$ tel que $z_n \neq 0$, x_n est un élément de S_k.

Exemple 1.1.5 Processus de sélection de fertilité.

Considérons une population monoïque (i.e ayant un seul sexe) d'individus diploïdes, avec k allèles différentes A_1, ..., A_k occupant un seul locus. Pour tout i, $j \in \{1, \ldots, k\}$, nous ne distinguons pas les génotypes A_iA_j et A_jA_i.

Chaque élément de la population choisit son partenaire de façon équiprobable sur l'ensemble des autres individus, se reproduit et meurt immédiatement après.

Plus précisément, nous supposons que l'évolution s'effectue de la manière suivante. Nous fixons $m \in \mathbb{N}$ et considérons, pour chaque paire de génotypes A_iA_j et A_rA_s, où i, j, r, $s \in \{1, \ldots, k\}$, une suite de variables aléatoires i.i.d $G_n(ij, rs)$, $n \in \mathbb{N}$, à valeurs dans $\{0, \ldots, m\}$, représentant le nombre de descendants créés par leur éventuelle reproduction à l'instant n.

A chaque instant, la population se modifie de la manière suivante :

(1) S'il y a moins de deux individus, la population disparaît (il y a extinction)

(2) Nous tirons au hasard, sans replacement, deux individus dans la population ; supposons ici qu'ils aient pour génotypes A_iA_j et A_rA_s.

(3) Nous supprimons ces derniers éléments de la population (ils meurent).

(4) Nous ajoutons $G_n(ij, rs)$ individus à la population. Le génotype A_uA_v étant déterminé par : $u = i$ ou j avec probabilité $1/2$, $v = r$ ou s avec probabilité $1/2$, indépendamment.

Nous définissons, pour tous i, $j \in \{1, \ldots, k\}$ tels que $i \neq j$, z_n^{ij} le nombre d'individus de génotype A_iA_j à l'instant n, et z_n^{ii} deux fois le nombre d'individus de génotype A_iA_i à l'instant n.

Nous étudions alors l'évolution de $z_n = (z_n^{ij}) \in (\mathbb{N}^{k\times k})^{\mathbb{N}}$. Notons que $|z_n| = \sum_{i,j} z_n^{ij}$ vaut deux fois la taille de la population étudiée. Nous montrons facilement que $(z_n)_{n\in\mathbb{N}}$ est une chaîne de Markov homogène, et que les conditions **H-5.1** et **H-5.2** sont vérifiées.

□

Exemple 1.1.6 Processus de sélection de fertilité avec mutations

Pour tenir compte des mutations dans les processus de sélection de fertilité, nous introduisons pour tous i, j, r, $s \in \{1, \ldots, k\}$ des probabilités de mutation $\mu(ij, rs)$ du génotype A_iA_j au génotype A_rA_s. Alors, pour tous i, $j \in \{1, \ldots, k\}$, $\sum_{r\leq s} \mu(ij, rs) = 1$.

Ces mutations s'introduisent dans la définition du processus de sélection de fertilité de l'exemple précédent, en considérant qu'à la fin de chaque étape, les individus de génotype A_uA_v obtenus par la règle (4) deviennent des individus $A_{\tilde{u}}A_{\tilde{v}}$ avec probabilité $\mu(uv, \tilde{u}\tilde{v})$.

Le processus $(z_n)_{n\in\mathbb{N}}$, défini à la fin de l'exemple 1.1.5, est toujours une chaîne de Markov homogène et vérifie également les hypothèses **H-5.1** et **H-5.2**.

□

Dans le cas général, lorsque $|z_n|$ est grand, alors d'une part $x_{n+1} - x_n$ est petit puisque d'après l'hypothèse **H-5.1** $|z_{n+1} - z_n|$ est borné, et d'autre part la quantité $E(x_{n+1} - x_n \mid \mathcal{F}_n)$ peut d'après l'hypothèse **H-5.2** être déterminée précisément à partir des valeurs de $p_w(x_n)$ pour $|w| \leq m$.

Nous pouvons donc espérer comparer x_n aux solutions d'une équation différentielle. Notons $w = x_{n+1} - x_n$, et calculons

$$x_{n+1}-x_n = \frac{z_{n+1}}{|z_{n+1}|}-\frac{z_n}{|z_n|} = \frac{z_{n+1} - z_n}{|z_n|}+\frac{z_{n+1}}{|z_{n+1}|}(\frac{|z_n| - |z_{n+1}|}{|z_n|}) = \frac{1}{|z_n|}[w-x_{n+1}\alpha(w)].$$

Donc $|x_{n+1} - x_n| \leq 2m/|z_n|$, et également

$$x_{n+1} - x_n = \frac{1}{|z_n|}[w - x_n\alpha(w)] + b_{n+1},$$

où

$$|b_{n+1}| = |(x_{n+1} - x_n)\frac{\alpha(w)}{|z_n|}| \leq \frac{2m^2}{|z_n|^2}.$$

Le lemme suivant développe cette remarque.

Lemme 1.1.1 *Considérons une chaîne de Markov homogène $(z_n)_{n\in\mathbb{N}}$ sur $\mathbb{N}^k$ satisfaisant les hypothèses* **H-5.1** *et* **H-5.2**. *Soit $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$ la filtration engendrée* *par $(z_n)_{n\in\mathbb{N}}$. Il existe des suites de variables aléatoires $(U_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ adaptées à la filtration $\mathbb{F}$ telles que, pour tout $n \in \mathbb{N}$,*

(i) Si $z_n \neq 0$, alors

$$x_{n+1} - x_n = \frac{1}{|z_n|}\Big(\sum_{w\in\mathbb{Z}^k} p_w(x_n)(w - x_n\alpha(w)) + U_{n+1} + b_{n+1}\Big) \tag{1.3}$$

(ii) $E[U_{n+1} \mid z_n] = 0$

(iii) $\|U_{n+1}\| \leq 4m$ et $E[\|U_{n+1}\|^2 \mid \mathcal{F}_n] \leq 4m^2$

(iv) $\|b_{n+1}\| \leq \frac{2m^2}{\max\{1,|z_n|\}}$.

Le processus $(x_n)_{n\in\mathbb{N}}$ suit donc un algorithme bruité d'approximation de Cauchy-Euler, de pas $1/|z_n|$ à l'instant n, de l'équation différentielle ordinaire

$$\frac{dx}{dt} = \sum_{w\in\mathbb{Z}^k} p_w(x)(w - x\alpha(w)). \tag{1.4}$$

Remarquons cependant que $(x_n)_{n\in\mathbb{N}}$ n'est pas un algorithme stochastique au sens de l'hypothèse **H-1**, les pas $1/|z_n|$ n'étant pas déterministes. L'évolution de la taille de z_n joue un rôle prépondérant dans le contrôle que nous pouvons avoir sur x_n. Nous avons donc besoin de prendre en compte la dynamique de la taille de la population pour énoncer des résultats comparables aux théorèmes précédents, et les résultats présentés dans les sections 1.1.2 à 1.1.4 ne s'appliquent pas directement.

Il est en revanche possible de décrire précisément les ensembles limites de la distribution génotypique lorsque la population croît assez rapidement, comme l'énonce le théorème suivant, dû à Schreiber (2001, [44]).

Nous définissons le processus continu $X(.)$ associé à $(x_n)_{n\in\mathbb{N}}$ de la même manière qu'en 1.1.1, avec des pas $\gamma_i = 1/|z_i|$.

Nous écrivons dans la suite qu'un processus $(z_n)_{n\in\mathbb{N}}$ est un processus d'urnes généralisées si $(z_n)_{n\in\mathbb{N}}$ est une chaîne de Markov homogène sur $\mathbb{N}^k$ satisfaisant les conditions **H-5.1** et **H-5.2**.

Théorème 1.1.11 *Soit $(z_n)_{n\in\mathbb{N}}$ un processus d'urnes généralisées. Alors, p.s sur l'événement $\{\liminf \frac{|z_n|}{n} > 0\}$,*

(1) le processus $X(.)$ associé à $(x_n)_{n\in\mathbb{N}}$ est p.s une pseudo-trajectoire asymptotique du flot associé à (1.4).

(2) l'ensemble limite $L((x_n)_{n\in\mathbb{N}})$ de $(x_n)_{n\in\mathbb{N}}$ est p.s un ensemble intérieurement récurrent pas chaînes du flot associé à (1.4).

La proposition suivante énonce que les sous-ensembles du simplexe sur lesquels la population a tendance à décroître ne peuvent p.s pas être ensembles limites de l'algorithme.

Remarquons qu'asymptotiquement, la croissance moyenne de la population partant de $x \in S_k$ est

$$\sum_w p_w(x)\alpha(w).$$

Proposition 1.1.1 *Soit $(z_n)_{n\in\mathbb{N}}$ un processus d'urnes généralisées. Si $K \subset S_k$ est un ensemble compact satisfaisant*

$$\sup_{x\in K} \sum_w p_w(x)\alpha(w) < 0,$$

alors

$$P[\{L((x_n)_{n\in\mathbb{N}}) \subset K\} \cap \{\lim_{n\to\infty} |z_n| = \infty\}] = 0.$$

Nous pouvons a contrario montrer qu'un attracteur sur lequel la population a tendance à croître est ensemble limite de l'algorithme avec probabilité non nulle.

Définissons l'ensemble $Att_\infty(X)$ des points $x \in S_k$ tels que, pour tout $M \in \mathbb{N}$ et pour tout voisinage ouvert U de x,

$$P[\exists n \in \mathbb{N}/ \ |z_n| \geq M \text{ et } x_n \in U] > 0.$$

Cet ensemble correspond aux régions du simplexe atteignables par une population de grande taille.

Théorème 1.1.12 *Soit $(z_n)_{n\in\mathbb{N}}$ un processus d'urnes généralisées. Soit $\mathcal{A}$ un attracteur de l'équation différentielle associée (1.4), de bassin d'attraction $\mathcal{B}(\mathcal{A})$. Supposons*

$$a = \inf_{x\in\mathcal{A}} \sum_w p_w(x)\alpha(w) > 0,$$

et définissons

$$\mathcal{C} = \{\liminf_{n\to\infty} \frac{|z_n|}{n} \geq a\} \cap \{L((x_n)_{n\in\mathbb{N}}) \subset \mathcal{A}\}.$$

Si U est un ensemble ouvert dont la fermeture est incluse dans $\mathcal{B}(\mathcal{A})$, alors il existe une constante $K > 0$ telle que, pour tout $M \in \mathbb{N}$,

$$P[\mathcal{C}] \geq (1 - \frac{K}{M})P[\exists n \in \mathbb{N} \ |z_n| \geq M \ \textit{et} \ x_n \in U].$$

En particulier, si

$$\mathcal{B}(\mathcal{A}) \cap Att_\infty(X) \neq \emptyset,$$

alors $P[\mathcal{C}] > 0$.

Nous pouvons bien sûr estimer la vitesse de convergence d'une population dont la distribution génotypique converge vers un ensemble sur lequel la population a tendance à croître, comme l'énonce la proposition suivante.

Proposition 1.1.2 *Soit $(z_n)_{n\in\mathbb{N}}$ un processus d'urnes généralisées. Soit $K \subset S_k$ un ensemble compact. Si*

$$a = \inf_{x\in K} \sum_w p_w(x)\alpha(w) > 0,$$

alors

$$\liminf_{n\to\infty} \frac{|z_n|}{n} \geq a$$

sur l'événement $\{L((x_n)_{n\in\mathbb{N}}) \subset K\} \cap \{\lim_{n\to\infty} |z_n| = \infty\}$.

Il est également possible d'adapter le théorème 1.1.10 sur la non-convergence vers les points ou orbites linéairement instables, mais le résultat reste subordonné au fait que la population croisse suffisamment régulièrement.

Définition 1.1.5 *Soit $(z_n)_{n\in\mathbb{N}}$ un processus d'urnes généralisées. Etant donné un compact $\mathcal{U} \subset int(S_k)$, nous disons que $(z_n)_{n\in\mathbb{N}}$ est non-dégénéré en $\mathcal{U}$ si, pour tout $x \in \mathcal{U}$,*

$$Vect\{w \in \mathbb{Z}^k : \ p_w(x) > 0\} = \mathbb{R}^k$$

où, pour tout sous-ensemble A de $\mathbb{R}^k$, $Vect(A) \subset \mathbb{R}^k$ désigne l'espace vectoriel engendré par les points de A.

Théorème 1.1.13 *Soit $(z_n)_{n\in\mathbb{N}}$ un processus d'urnes généralisées. Soit $\mathcal{U} \subset int(S_k)$ un équilibre linéairement instable ou une orbite hyperbolique linéairement instable du système différentiel associé (1.4). Supposons*

(a) pour tout $w \in \mathbb{Z}^k$ tel que $|w| \leq m$, la fonction p_w est $C^{1+\beta}$ sur un voisinage de $\mathcal{U}$, avec $\beta > 1/2$

(b) $(z_n)_{n\in\mathbb{N}}$ est non dégénéré sur $\mathcal{U}$.

Alors $P[L((x_n)_{n\in\mathbb{N}}) \subset \mathcal{U}$ et $\liminf_{n\to\infty} \frac{|z_n|}{n} > 0] = 0$.

Les résultats qui précèdent impliquent le théorème suivant.

Théorème 1.1.14 *Soit $(z_n)_{n\in\mathbb{N}}$ un processus d'urnes généralisées. Supposons que toutes les fonctions p_w soient $C^{1+\beta}$ avec $\beta > 1/2$, que $(z_n)_{n\in\mathbb{N}}$ soit non dégénéré sur S_k, que $Att_\infty(X) = S_k$ et que les ensembles intérieurement récurrents par chaînes de l'équation différentielle associée (1.4) soient des équilibres hyperboliques q vérifiant $\sum_w p_w(q)\alpha(w) \neq 0$. Alors*

1) *$P[\liminf \frac{|z_n|}{n} > 0] > 0$ si et seulement s'il existe un équilibre linéairement stable q tel que $\sum_w p_w(q)\alpha(w) > 0$*

2) *$P[\{L((x_n)_{n\in\mathbb{N}}) = q\}] > 0$ pour tout équilibre linéairement stable q satisfaisant $\sum_w p_w(q)\alpha(w) > 0$*

3) *$P[\{L((x_n)_{n\in\mathbb{N}}) = q\}] = 0$ pour tout équilibre linéairement instable q.*

Nous donnons un exemple d'application de ce résultat aux processus additifs de sélection de fertilité avec mutations.

Exemple 1.1.7 Processus additifs de sélection de fertilité. Nous reprenons les notations des exemples 1.1.5 et 1.1.6. Nous disons qu'un processus de sélection de fertilité (avec ou sans mutations) est additif ssi le nombre moyen de descendants produits par reproduction entre deux individus A_iA_j et A_rA_s est additif : autrement dit, s'il existe une suite $(\gamma_{ij})_{i,j\in\{1,\dots,k\}}$ telle que, pour tous i, j, r, s $\in \{1,\dots,k\}$,

$$g(ij, rs) = E[G_0(ij, rs)] = \gamma_{ij} + \gamma_{rs}.$$

□

Proposition 1.1.3 *Considérons un processus $(z_n)_{n\in\mathbb{N}}$ additif de sélection de fertilité avec mutations. Supposons*

- *$\forall\ i,\ j,\ r,\ s \in \{1,\dots,k\}$, $\mu(rs, ij)$ est strictement positif, et suffisamment petit lorsque $\{r,s\} \neq \{i,j\}$*
- *$\forall\ i,\ j,\ r,\ s \in \{1,\dots,k\}$, $P[G_0(ij, rs) \geq 3] > 0$.*

Supposons également que les équilibres q de l'équation de sélection de fertilité sans mutations sont hyperboliques et vérifient $\sum_w p_w(q)\alpha(w) \neq 0$.

Alors les conclusions du théorème 1.1.14 sont satisfaites.

1.2 Marches aléatoires renforcées par sommets

1.2.1 Introduction

La deuxième partie de cette thèse est consacrée aux marches aléatoires renforcées par sommets, définies par Pemantle en 1988 ([37]), qui sont des processus évoluant dans un environnement qu'ils contribuent à faire changer par le fait qu'ils ont une probabilité plus grande de revenir aux endroits déjà visités.

Plus précisément, plaçons nous sur un espace de probabilité $(\Omega, \mathcal{F}, P)$ et considérons un graphe G localement fini, $\sim$ sa relation d'adjacence et $S(G)$ l'ensemble de ses sommets.

Soit $(X_n)_{n\in\mathbb{N}}$ un processus à valeurs dans $S(G)$; notons $\mathcal{F}_n = \sigma(X_0,\dots,X_n)$, et $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$. Définissons, pour tout sommet $v \in S(G)$, $Z_n(v)$ le nombre de fois où ce sommet a été visité, c'est à dire

$$Z_n(v) = 1 + \sum_{i=0}^{n} 1_{X_i=v},$$

en supposant par convention qu'à l'instant 0 tous les sommets ont été visités une fois par la marche.

Considérons une application $a : S(G) \times S(G) \longrightarrow \mathbb{R}^+$ telle que, pour tous v, $w \in S(G)$, $a(v,w) > 0 \Longleftrightarrow v \sim w$.

Alors $(X_n)_{n\in\mathbb{N}}$ est une *marche aléatoire renforcée par sommets* de renforcement a et issue de $v_0 \in S(G)$ si $X_0 = v_0$ et, pour tout $n \in \mathbb{N}$,

$$P(X_{n+1} = v \mid \mathcal{F}_n) = 1_{v\sim X_n} \frac{a(X_n, v) Z_n(v)}{\sum_{w\sim X_n} a(X_n, w) Z_n(w)}.$$

Ce type de processus avec mémoire a été étudié en détail sur des graphes finis par Pemantle (1992,[39]) et Benaïm (1997,[3]). Il peut s'appliquer à la modélisation de phénomènes d'auto-organisation, liés à l'apprentissage ou à des problèmes économiques. Une description détaillée de ces applications et des résultats connus sur le sujet est donnée dans [40] (voir aussi [48]).

Cette idée de marches aléatoires avec renforcement trouve son origine dans les travaux de Coppersmith et Diaconis en 1986 ([19]), avec la notion de *marches aléatoires renforcées par arêtes* définies de telle manière que les probabilités d'atteindre un sommet dépendent du nombre de fois où l'arête menant à ce sommet a été visitée (et non du nombre de fois où ce sommet a été visité). Ces processus, dont la définition diffère peu de celle des marches aléatoires renforcées par sommets, donnent néanmoins lieu à des comportements asymptotiques sensiblement différents.

Les marches aléatoires renforcées par sommets ont également un équivalent en temps continu, avec la notion de diffusions auto-attractives introduite par Cranston et Le Jan en 1995 ([20]). Ces diffusions auto-attractives ont été étudiées par Raimond en 1996 ([41]), par Benaïm, Ledoux et Raimond en 2000 ([11]), et également par Herrmann et Roynette en 2000 ([28]).

Les résultats obtenus dans le cadre de cette thèse se placent dans le cas où $G = \mathbb{Z}$ et $a(v, w) = 1_{v\sim w}$. Sous ces hypothèses, Pemantle et Volkov ont obtenu des énoncés très précis sur le comportement asymptotique de X_n ([40]) [1], que nous rappelons ci-dessous.

Définissons

$$R = \{v \in \mathbb{Z} / \ \exists n \in \mathbb{N},\ X_n = v\}\ ,\ R' = \{v \in \mathbb{Z} / \ X_n = v \text{ une infinité de fois}\}.$$

Pour $v \in \mathbb{Z}$ et $\alpha \in]0, 1[$, considérons les six événements :

(i) $R' = \{v - 2, v - 1, v, v + 1, v + 2\}$;

(ii) $\ln Z_n(v - 2)/\ln n \to \alpha$;

(iii) $\ln Z_n(v + 2)/\ln n \to 1 - \alpha$;

(iv) $Z_n(v - 1)/n \to \alpha/2$;

(v) $Z_n(v + 1)/n \to (1 - \alpha)/2$;

(vi) $Z_n(v)/n \to 1/2$.

[1] partiellement généralisés par Volkov ([49]) à une classe assez large de graphes localement finis.

Théorème 1.2.1 *([40], Pemantle et Volkov)* $P(|R| < +\infty) = 1$.

Théorème 1.2.2 *([40], Pemantle et Volkov)* $P(|R'| \leq 4) = 0$.

Théorème 1.2.3 *([40], Pemantle et Volkov) Pour tout ouvert* $I \subset]0,1[$ *et tout* $v \in \mathbb{Z}$, *il existe avec une probabilité strictement positive un réel* $\alpha \in I$ *tel que les événements (i) à (vi) se réalisent.*

Conjecture 1.2.1 *([40], Pemantle et Volkov) Il existe presque-sûrement* $v \in \mathbb{Z}$ *et* $\alpha \in]0,1[$ *tels que les événements (i) à (vi) se réalisent.*

Nous avons démontré cette conjecture au cours de cette thèse. La preuve en est présentée dans le chapitre 8. Auparavant, nous donnons un plan de ses arguments principaux dans la section 1.2.3, et un résumé détaillé dans le chapitre 7.

Nous avons également proposé une nouvelle preuve des théorèmes 1.2.1, 1.2.2 et 1.2.3, obtenue en mettant en oeuvre des techniques utilisées pour la preuve de cette conjecture. Cette démonstration est présentée dans le chapitre 6 ; nous en donnons l'heuristique dans la section 1.2.2.

1.2.2 Un résultat de Pemantle et Volkov sur $\mathbb{Z}$

Pour comprendre les causes du blocage de la marche avec probabilité strictement positive dans cinq points, nous commençons par étudier le comportement d'une marche de même nature évoluant sur trois points C, D et E.

Notons, pour tout $n \in \mathbb{N}$, $C_n = Z_n(C)$, $D_n = Z_n(D)$ et $E_n = Z_n(E)$ les nombres de fois où les points C, D et E ont été visités à l'instant n.

A chaque instant $n \in \mathbb{N}$, si $X_n = C$ ou E, alors $X_{n+1} = D$; si $X_n = D$, alors $X_n = C$ avec une probabilité $C_n/(C_n + E_n)$ égale au nombre de fois où C a été visité sur le nombre total de fois où C ou E ont été visités.

Ce mécanisme est équivalent à celui d'une urne de Pólya à deux couleurs C et E, ayant au départ respectivement C_0 et E_0 boules. A chaque étape (qui se déroule un instant sur deux dans notre exemple) on tire au hasard une boule, que l'on remet en ajoutant une autre boule de la même couleur.

Alors un résultat classique affirme que $C_n/(C_n + E_n)$ tend vers une loi béta de paramètres C_0 et E_0 (voir par exemple [26], volume 2, chapitre VII). Une conséquence immédiate en est que $C_n/(C_n + E_n)$ converge p.s vers un réel strictement compris entre 0 et 1. Donnons une preuve courte de cette dernière affirmation, dûe à Volkov ([49]).

Affirmation 1.2.1 *Presque-sûrement, il existe* $\alpha \in]0,1[$ *tel que* $C_n/(C_n+E_n) \to \alpha$.

PREUVE: Définissons, pour tout $n \in \mathbb{N}$,

$$U_n = \ln(C_n + E_n) - \ln(E_n - 1).$$

Alors $(U_n)_{n\in\mathbb{N}}$ est une sur-martingale positive. En effet, pour tout $n \in \mathbb{N}$, si $X_n = D$, alors

$$E(U_{n+1} - U_n \mid \mathcal{F}_n) = \ln(1 + \frac{1}{C_n + E_n}) + \frac{E_n}{C_n + E_n} \ln(1 - \frac{1}{E_n}) < 0.$$

Donc $(U_n)_{n\in\mathbb{N}}$ converge p.s. Nous pouvons appliquer le même argument en inversant C et E, ce qui permet de conclure.

□

Ajoutons deux points A et B à gauche de C, et notons, pour tout $n \in \mathbb{N}$, $A_n = Z_n(A)$, $B_n = Z_n(B)$, t_n (resp. u_n) l'instant de n-ième visite en C (resp. en D).

Expliquons le principe du processus par lequel la marche reste dans certains cas bloquée à droite de A.

Considérons, pour tout $n \in \mathbb{N}$, les événements suivants :

1) A été visité a fois (où a est une constante),

2) B a été peu visité par rapport à C, c'est à dire $B_n \leq C_n^\xi$ (où $\xi \in]0,1[$ est une constante),

3) C n'a pas été trop visité par rapport à D, c'est à dire $C_n/D_n \leq \zeta$ (où $\zeta < \xi$ est une constante).

Soit $m \in \mathbb{N}$ un instant auquel **1)**, **2)**, **3)** sont réalisés et $X_m = C$. Alors ces propriétés **1)**, **2)** et **3)** auront une grande tendance à rester vraies. En effet,

(i) A chaque instant t_n de visite en C, si **1)**, **2)** et **3)** sont vérifiées, alors la probabilité de visite en A avant l'instant t_{n+1} est

$$P(X_{t_n+2} = A \mid \mathcal{F}_{t_n}) = \frac{B_{t_n}}{B_{t_n} + D_{t_n}} \frac{A_{t_n}}{A_{t_n} + C_{t_n}} \leq \frac{B_{t_n}}{C_{t_n}} \frac{A_{t_n}}{C_{t_n}} \leq \frac{1}{n^{1-\xi}} \frac{a}{n} \frac{\mu}{n^{1-\xi}} \leq \frac{a}{n^{2-\xi}}.$$

Donc la probabilité que **1)** reste vraie tant que **2)** et **3)** restent vraies est supérieure à

$$1 - \sum_{n \geq C_m} \frac{a}{n^{2-\xi}} \geq 1 - \frac{a}{(1-\xi)} \frac{1}{(C_m - 1)^{1-\xi}}.$$

(ii) De même, à chaque instant t_n de visite en C, si **1)**, **2)** et **3)** sont vérifiées, nous pouvons faire les remarques suivantes concernant le nombre de visites en B avant t_{n+1}. D'une part si A n'est pas visité à l'instant $t_n + 2$, B est visité au plus une fois avant t_{n+1}. Et d'autre part

$$P(X_{t_n+1} = B \mid \mathcal{F}_{t_n}) = \frac{B_{t_n}}{B_{t_n} + D_{t_n}} \leq \frac{B_{t_n}}{C_{t_n}} \frac{C_{t_n}}{D_{t_n}} \leq \frac{\zeta B_{t_n}}{n}.$$

Par conséquent

$$E(B_{\kappa_{n+1}} \mid \mathcal{F}_{t_n}) \leq B_{t_n}(1 + \zeta/n) \leq n^\xi(1 + \xi/n) \simeq (n+1)^\xi.$$

Ce qui implique que, en moyenne, **2)** a également tendance à se reproduire tant que **1)** et **3)** restent vraies. Nous obtenons, en prenant en compte les perturbations autour de cette tendance moyenne, le résultat que B_n/C_n^ξ n'augmente que peu avec une grande probabilité.

(iii) A chaque instant u_n de visite en D, la probabilité d'aller en C est $C_{u_n}/(C_{u_n}+E_{u_n})$, qui est inférieur à C_{u_n}/D_{u_n}. Par conséquent

$$E(C_{u_{n+1}}/D_{u_{n+1}} \mid \mathcal{F}_{u_n}) \leq C_{u_n}/D_{u_n}.$$

Il nous reste à évaluer $C_{u_{n+1}} - C_{u_n+1}$, c'est à dire le nombre de visites de B vers C avant le retour en D à l'instant u_{n+1}. Nous pouvons montrer que ces visites, qui contribuent à faire augmenter C_{u_n}/D_{u_n}, ont en moyenne un impact peu important sur son évolution lorsque l'on part d'une situation dans laquelle **1)** et **2)** sont réalisés, puisqu'alors la probabilité d'aller en B partant de C est faible.

En résumé, de la même manière qu'en **(ii)**, nous pouvons montrer qu'en moyenne C_{u_n}/D_{u_n} n'augmente que peu sous cette hypothèse.

Le résultat qui correspond à ces considérations heuristiques peut s'énoncer de la manière suivante. Soient $\mu > 1$, ζ, $\xi \in]0,1[$ tels que $\xi > \zeta$. Définissons, pour tout $n \in \mathbb{N}$, les événements suivants

$\mathcal{E}_1(n)_a : A_n = a$

$\mathcal{E}_2(n)_{\mu,\xi} : B_n \leq \mu C_n^\xi$

$\mathcal{E}_3(n)_\zeta : C_n/D_n \leq \zeta$,

et $\mathcal{E}(n)_{a,\mu,\zeta,\xi} = \mathcal{E}_1(n)_a \cap \mathcal{E}_2(n)_{\mu,\xi} \cap \mathcal{E}_3(n)_\zeta$.

Lemme 1.2.1 *Supposons $a \in \mathbb{N}^*$, $\mu > 1$, ζ, $\xi \in]0,1[$ fixés tels que $\xi > \zeta$. Alors, pour tout $\nu < \min(1/2, 1-\xi)$, il existe $k_0 \in \mathbb{N}$ ne dépendant que de μ, ζ et ξ tel que pour tout $k \geq k_0$,*

$$P(\cap_{n \geq t_k} \mathcal{E}(n)_{a,\mu,\zeta+k^{-\nu},\xi} | \mathcal{F}_{t_k}) \geq c 1_{\mathcal{E}(t_k)_{a,1,\zeta,\xi}}.$$

Nous pouvons déduire les théorèmes 1.2.1 et 1.2.3 de ce lemme 1.2.1, à partir des arguments suivants. Nous donnons ici uniquement le schéma des preuves données dans le chapitre 6.

Commençons par le théorème 1.2.3 : supposons $v \geq v_0$ par exemple, notons $A = v-3$ et B à G les points qui suivent A par ordre alphabétique. Soient ζ_1, ζ_2 $\in]0,1[$ quelconques. Après l'instant initial, il est possible avec une probabilité non nulle de se rendre en D puis de ne faire pendant un certain temps que des allers et retours entre $\{D\}$ et $\{C, E\}$, de sorte qu'il existe un instant $n \in \mathbb{N}$ auquel le rapport $C_n/(C_n+E_n)$ soit dans l'intervalle $]\zeta_1, 1-\zeta_2[$, avec C_n et E_n aussi grands que voulus.

En appliquant le lemme 1.2.1 sur les points A à G d'une part et G à A d'autre part nous obtenons que, partant d'une telle situation, on va avec une grande probabilité ne plus jamais visiter les points A et G, peu visiter les points B et

F, le rapport $C_n/(C_n+E_n)$ restant proche de l'intervalle $]\zeta_1, 1-\zeta_2[$. Ce qui, en appliquant des résultats de convergence sur les martingales liées à ce phénomène, permet de conclure que les événements (i) à (vi) sont bien vérifiés.

La preuve du théorème 1.2.1 repose sur la remarque suivante : à chaque fois qu'un point v à gauche de v_0 est visité pour la deuxième fois (rappelons que par convention, à l'instant initial tous les sommets ont été visités une fois par la marche), tous ses antécédents n'ont été visités qu'une fois. Notons $A = v-2$, et B à G les points suivants.

On peut alors trouver, pour tout $k_0 \in \mathbb{N}$, avec une probabilité ne dépendant que de k_0, un instant $n \in \mathbb{N}$ auquel $A_n = 1$, $B_n = 1$, $C_n = k_0$ et $D_n \geq 2k_0$, ce qui implique en reprenant les notations du lemme 1.2.1 que $\mathcal{E}(t_{k_0})_{1,1,1/2,3/4}$ est vérifiée. Enfin, en appliquant ce lemme, nous pouvons conclure qu'alors le sommet $v-2$ n'est plus jamais visité avec une grande probabilité si k_0 est supposé assez grand.

En résumé, pour tout $v \leq v_0$, la probabilité que $v-2$ ne soit jamais visité conditionnellement au fait que v a été visité est supérieure à une constante. Par conséquent, la probabilité qu'un sommet v_0-2k soit visité au moins une fois tend vers 0 quand k tend vers l'infini, ce qui implique que R a p.s une borne inférieure. On montre de la même manière que R a p.s une borne supérieure.

Il reste maintenant à expliquer le principe de la démonstration que nous proposons du théorème 1.2.2.

Commençons par introduire, pour tout $A \in \mathbb{Z}$, l'événement

$$\Omega_A = \{A = \inf R' - 1\}.$$

Nous savons, d'après le théorème 1.2.1, que

$$P(\cup_{A\in\mathbb{Z}}\Omega_A) = 1.$$

Lorsque nous sommes, pour $A \in \mathbb{Z}$, sur un ensemble Ω_A, il n'y a en tout qu'un nombre fini de visites en A et, a contrario d'un des arguments du lemme 1.2.1, il y a peu de visites de B vers C (si B à G sont les entiers qui suivent A).

Plus précisément, introduisons, pour tous entiers L et M consécutifs ($M = L \pm 1$), l'événement

$$\Upsilon_{L,M} = \{\sum \frac{1_{\{X_n=L, X_{n+1}=M\}}}{M_n} < +\infty\}.$$

Cet événement correspond au fait qu'il y ait peu de visites de L vers M. Nous pouvons par exemple déduire du lemme de Kronecker que, sur $\Upsilon_{L,M}$,

$$\sum_{k=1}^{n} 1_{\{X_n=L, X_{n+1}=M\}} = o(M_n).$$

Le lemme du Borel-Cantelli conditionnel (voir par exemple [21]) nous permet d'affirmer que

$$\Omega_A \subset \Upsilon_{B,A} = \{\sum \frac{1_{\{X_n=B, X_{n+1}=A\}}}{A_n} < +\infty\} = \{\sum \frac{1_{X_n=B}}{A_n+C_n} < +\infty\} = \Upsilon_{B,C}.$$

Nous reprenons d'autre part une idée proche de celle développée dans la partie (**iii**) de l'argument de la preuve du lemme 1.2.1, à savoir que si l'on observe le processus un instant sur deux, $C_n/(C_n + E_n)$ n'augmente pas en moyenne lors des visites venant de D ou F, et est donc essentiellement affecté par l'impact des visites de B vers C. Nous déduisons de ces deux remarques que $C_n/(C_n + E_n)$ converge vers un réel $\alpha \in [0, 1[$. Ce qui est résumé par le résultat suivant, qui constitue une partie de la proposition 8.3.1 de la preuve de la conjecture (Chapitre 8)

Proposition 8.3.1 *(début). Soient Q, R, S, T et U cinq entiers consécutifs. Alors*

$$\textbf{a)}\ Y_n = \ln(R_n + T_n) - \ln(T_n - 1) - \sum_{k \le n} \frac{1}{R_{k-1} + T_{k-1}} 1_{\{X_{k-1}=Q, X_k=R\}}$$

est une sur-martingale.

$$\textbf{b)}\ \Upsilon_{Q,R} \subset \{\exists \alpha \in [0,1[/\ \frac{R_n}{R_n + T_n} \to \alpha\}.$$

L'affirmation **b)** se déduit de **a)** par le fait que sur $\Upsilon_{Q,R}$, Y_n est p.s minorée, et donc converge p.s.

Cette proposition implique, avec $Q := B$, $R := C$ ... que $\lim E_n/(C_n + E_n) > 0$ sur Ω_A, donc en particulier que $\liminf E_n/D_n > 0$. Par ailleurs, à partir du point E, lui-même visité un nombre infini de fois sur Ω_A, la probabilité d'aller en F est

$$\frac{F_n}{D_n + F_n} \ge \frac{1}{D_n + 1}.$$

Ce qui permet d'affirmer que

$$\begin{aligned} \Omega_A \quad &\subset \{\sum \frac{1_{X_n=E} F_n}{D_n + F_n} \ge \sum \frac{1_{X_n=E}}{D_n + 1} = \sum \frac{1_{X_n=E}}{E_n} \frac{E_n}{D_n + 1} = +\infty\} \\ &\subset \{\sum 1_{X_n=E, X_{n+1}=F} = +\infty\} \subset \{F \text{ visité un nombre infini de fois }\}. \end{aligned}$$

Ce qui permet de conclure la preuve du théorème 1.2.2.

1.2.3 Preuve d'une conjecture de Pemantle et Volkov

Nous présentons maintenant le plan général de la preuve de la conjecture 1.2.1 donnée dans le chapitre 8. Un résumé détaillant l'heuristique des résultats de cette preuve est donné dans le chapitre 7. Pour simplifier les notations, nous reprenons la numérotation des résultats de la preuve, chapitre 8.

Tout d'abord, nous utilisons toujours le fait que

$$P(\cup_{A\in\mathbb{Z}}\Omega_A) = 1.$$

Nous fixons donc $A \in \mathbb{Z}$, et nous plaçons sur Ω_A. Notons B, C, ... les points qui suivent A. Nous avons déjà énoncé (proposition 3.1) que

$$\Omega_A \subset \Upsilon_{B,A} = \Upsilon_{B,C} \subset \{\exists\alpha \in]0,1]/\ \frac{C_n}{C_n+E_n} \to \alpha\}.$$

Il est possible d'obtenir une règle du même type pour la convergence du rapport $G_n/(G_n+I_n)$ lorsque $\alpha > 0$, en utilisant un autre résultat énoncé dans la proposition 8.3.1.

Proposition 8.3.1 *(fin). Soient Q, R, S, T et U cinq entiers consécutifs, croissants ou décroissants. Alors*

a)

$$Z_n = \ln(R_n+T_n) - \ln(T_n-1) - \sum_{k=1}^{n} \frac{1}{R_{k-1}+T_{k-1}} 1_{\{X_{k-1}=Q, X_k=R\}} + \sum_{k=1}^{n} \frac{R_{k-1}}{(R_{k-1}+T_{k-1})T_{k-1}} 1_{\{X_{k-1}=U, X_k=T\}}$$

est une sur-martingale.

b) $\Upsilon_{Q,R} \cap \{\alpha > 0\} \subset \Upsilon_{U,T}$.

Nous utilisons ici cette proposition avec $Q := B$, $R := C$, $S := D$, $T := E$, $U := F$, ce qui permet d'obtenir

$$\Upsilon_{B,C} \cap \{\alpha > 0\} \subset \Upsilon_{F,E} = \Upsilon_{F,G} \subset \{\exists\beta \in]0,1]/\ \frac{G_n}{G_n+I_n} \to \beta\}.$$

Nous avons besoin d'étudier plus précisément les cas $\alpha = 0$ et $\beta = 0$. Nous allons en fait montrer que $\alpha = 0$ est p.s impossible, et qu'en revanche $\beta = 0$ est possible mais implique $H_n/(H_n+J_n) \to \beta' > 0$. Ces résultats reposent sur les lemmes 8.6.1, 8.6.2 et 8.2.4.

Lemme 8.6.1 *Soient P, Q, R, S et T cinq entiers consécutifs. Alors*

$$\Upsilon_{P,Q} \cap \{\lim \frac{Q_n}{Q_n+S_n} = 0\} \subset \Upsilon_{Q,R} = \Upsilon_{Q,P}.$$

Lemme 8.6.2 *Soient P, Q, R, S et T cinq entiers consécutifs. Alors*

$$\Upsilon_{P,Q} \cap \{\lim \frac{Q_n}{Q_n+S_n} = 0\} \subset \{\limsup \frac{R_n}{R_n+T_n} < 1\}.$$

Lemme 8.2.4 *Soient P, Q et R trois entiers consécutifs. Alors*

$$\Upsilon_{P,Q} \cap \Upsilon_{R,Q} = \{Q_\infty < \infty\}.$$

Comme nous allons le voir, le lemme 8.6.1 est une conséquence du lemme 8.6.2.

L'argument de la preuve du lemme 8.6.2 est détaillé dans la partie 2 du résumé, chapitre 7. L'idée générale est celle d'une compétition entre les nombres de visites aux points Q et T. Nous voulons montrer que

$$\Gamma_0 = \Upsilon_{P,Q} \cap \{\lim \frac{Q_n}{Q_n + S_n} = 0\} \cap \{\liminf \frac{T_n}{R_n + T_n} = 0\}$$

est un événement de probabilité nulle.

D'une part, si l'on part d'une situation dans laquelle T a été plus visité que Q, alors le nombre de visites en T va avoir tendance à augmenter, et on obtiendra finalement que $\liminf S_n/R_n > 1$ avec une probabilité raisonnable. Nous pouvons ainsi supposer, pour tout $\mu > 1$, que T_n reste inférieur à μQ_n pour n assez grand.

D'autre part, sous cette dernière hypothèse le bruit inhérent au comportement de la marche aléatoire est suffisamment régulier pour que l'on puisse, à partir d'un ordre partiel sur ce type de marches aléatoires, utiliser l'heuristique de résultats de non-convergence vers les points instables et montrer que $\liminf S_n/R_n \neq 1$ avec probabilité 1.

Enfin, le lemme 8.6.1 se déduit du lemme 8.6.2, en utilisant l'inclusion

$$\Upsilon_{P,Q} \cap \{\limsup \frac{R_n}{R_n + T_n} < 1\} \subset \Upsilon_{Q,R},$$

qui constitue le corollaire 8.3.2 de la preuve. Ce corollaire signifie que, s'il y a un faible nombre de visites de P vers Q et si T est visité régulièrement par rapport à R, alors il y a également un faible nombre de visites de Q vers R.

L'argument de la preuve du lemme 8.2.4 est très simple : il repose sur l'idée que s'il y a peu de visites de P vers Q et peu de visites de R vers Q, alors il y a forcément peu de visites en Q, c'est à dire que Q est visité un nombre fini de fois. Plus précisément

$$\begin{aligned} \Upsilon_{P,Q} \cap \Upsilon_{R,Q} \quad &= \{\sum_{k \in \mathbb{N}^*} \frac{1_{\{X_{k-1}=P, X_k=Q\}}}{Q_k} < \infty\} \cap \{\sum_{k \in \mathbb{N}^*} \frac{1_{\{X_{k-1}=R, X_k=Q\}}}{Q_k} < \infty\} \\ &= \{\sum_{k \in \mathbb{N}^*} \frac{1_{\{X_k=Q\}}}{Q_k} < \infty\} = \{Q_\infty < \infty\}. \end{aligned}$$

Les trois lemmes précédents impliquent bien que $\alpha > 0$ p.s sur Ω_A : en effet

$$\Omega_A \cap \{\alpha = 0\} \subset \Omega_A \cap \Upsilon_{C,B} \subset \Upsilon_{A,B} \cap \Upsilon_{C,B} = \emptyset \text{ p.s.}$$

Etudions maintenant le cas $\beta = 0$. Alors $\beta' = \lim H_n/(H_n + J_n)$ existe car le lemme 8.6.1 implique que

$$\Upsilon_{F,G} \cap \{\lim \frac{G_n}{G_n + I_n} = 0\} \subset \Upsilon_{G,H} \subset \{\frac{H_n}{H_n + J_n} \to \beta' \in [0, 1[\}.$$

Et $\beta' > 0$, car les lemmes 8.2.4 et 8.6.1 impliquent que

$$\Upsilon_{G,H} \cap \{\lim \frac{H_n}{H_n + J_n} = 0\} \subset \Upsilon_{H,J} = \Upsilon_{H,G},$$

et

$$\begin{aligned} &\Omega_A \cap \{\lim \frac{G_n}{G_n + I_n} = 0\} \cap \{\lim \frac{H_n}{H_n + J_n} = 0\} \subset \Omega_A \cap \Upsilon_{F,G} \cap \Upsilon_{H,G} \\ &\subset \Omega_A \cap \{G_\infty < \infty\} \cap \{\lim \frac{H_n}{H_n + J_n} = 0\} = \emptyset \text{ p.s.} \end{aligned}$$

En résumé

$$\Omega_A \subset \Omega_{A,F} \cup \Omega_{A,1} \cup \Omega_{A,2} \text{ p.s}$$

où

$$\Omega_{A,F} = \Omega_A \cap \{G_\infty < \infty\},$$

$$\Omega_{A,1} = \Upsilon_{A,B} \cap \Upsilon_{B,C} \cap \Upsilon_{F,G} \cap \{\exists \alpha \in]0, 1[\ \exists \beta \in]0, 1[/ \ \frac{C_n}{C_n + E_n} \to \alpha \text{ et } \frac{G_n}{G_n + I_n} \to \beta\} \cap \{G_\infty = \infty\}$$

$$\begin{aligned} \Omega_{A,2} = &\{\exists \alpha \in]0, 1[,\ \exists \beta' \in]0, 1[/ \ \frac{C_n}{C_n + E_n} \to \alpha \text{ et } \frac{H_n}{H_n + J_n} \to \beta'\} \\ &\cap \Upsilon_{A,B} \cap \Upsilon_{B,C} \cap \Upsilon_{F,G} \cap \Upsilon_{G,H} \cap \{G_\infty = \infty\}. \end{aligned}$$

Il reste à prouver que, pour tout $A \in \mathbb{Z}$, $\Omega_{A,1}$ et $\Omega_{A,2}$ ont pour probabilité 0.

Ces deux preuves reposent sur des idées similaires. Commençons par exemple par $\Omega_{A,1}$.

L'argument se divise en deux étapes : sur $\Omega_{A,1}$, d'une part $H_n/D_n \not\to 1$ est p.s impossible, d'autre part $H_n/D_n \to 1$ est p.s impossible. Ce qui conduit à une contradiction.

La première étape se décompose de la manière suivante. En premier lieu nous montrons, en utilisant des idées reliées aux résultats de convergence avec probabilité non nulle vers un attracteur (voir par exemple [3], [24]), que $H_n/D_n \not\to 1$ implique $H_n/D_n \to 0$ ou $H_n/D_n \to \infty$. En deuxième lieu nous estimons la vitesse de convergence vers 0 ou ∞, ce qui nous permet de prouver que $\ln H_n/\ln D_n$ tend vers 0 ou ∞ lorsque n tend vers l'infini. Enfin en troisième lieu nous utilisons ce dernier résultat pour prouver que la marche aléatoire reste à partir d'un certain temps bloquée entre B et F ou à droite de F, ce qui conduit à une contradiction.

La preuve de la deuxième étape ($H_n/D_n \to 1$ p.s impossible) repose sur un principe de non-convergence vers les points instables similaire à celui de la preuve du lemme 8.6.2.

La preuve de $P(\Omega_{A,2}) = 0$ est du même type, en remplaçant I_n par D_n. Cependant nous avons besoin auparavant de montrer que les points B à L sont tous régulièrement visités sur $\Omega_{A,2}$. Nous utilisons donc un résultat (le lemme 8.8.1) nous permettant de distinguer entre le cas où la marche aléatoire reste finalement bloquée dans un sous-ensemble strict de $[B, L]$ et celui où les fréquences des visites aux points B à L sont suffisamment importantes pour pouvoir estimer les variations de I_n/D_n.

Les détails de ces arguments sont donnés dans le résumé de la preuve, chapitre 7, et la démonstration complète dans les sections 8.7 et 8.8.

Bibliographie

[1] K. Athreya. Doctoral dissertation. *Stanford University*, 1967.

[2] K. Athreya and P. Ney. *Branching processes.* Springer Verlag, Berlin, 1972.

[3] M. Benaïm. A dynamical systems approach to stochastic approximations. *SIAM Journal on Control and Optimization*, 34, 2 :437–472, 1996.

[4] M. Benaïm. Vertex-reinforced random walks and a conjecture of Pemantle. *Ann. Probab.*, 25 :361–392, 1997.

[5] M. Benaïm. Dynamics of stochastic approximation algorithms. *Séminaire de Probabilités XXXIII, Lecture Notes in Math., Springer*, 1709 :1–68, 1999.

[6] M. Benaïm and M. W Hirsch. Chain recurrence in surface flows. *Discrete and Continuous Dynamic Systems*, 1, 1 :1–17, 1995.

[7] M. Benaïm and M. W Hirsch. Dynamics of morse-smale urn processes. *Ergod. Th. Dynam. Sys.*, 15 :105–1030, 1995.

[8] M. Benaïm and M. W Hirsch. Asymptotic pseudotrajectories and chain recurrent flows, with applications. *J. Dynam. Differential Equations*, 8 :141–176, 1996.

[9] M. Benaïm and M. W Hirsch. Differential and stochastic epidemics models. *The Field Institute Communications*, 21 :31–43, 1999.

[10] M. Benaïm and M. W Hirsch. Mixed equilibria and dynamical systems arising from fictitious play in perturbed games. *Game and Economic Behavior*, 29 :36–72, 1999.

[11] M. Benaïm, M. Ledoux, and O. Raimond. Self-interacting diffusions. *A paraître dans Probab. Theor. Relat. Fields*, 2000.

[12] A. Benveniste, M. Métivier, and P. Priouret. *Stochastic Approximation and Adaptive Algorithms.* Springer-Verlag, Berlin and New York, 1990.

[13] A. Bienvenüe. Contribution à l'étude des marches aléatoires avec mémoire. *Doctoral dissertation, Université Claude Bernard-Lyon 1*, 1999.

[14] R. Bowen. ω-limit sets of axiom A diffeomorphisms. *J. Diff.Eq.*, 18 :333–339, 1975.

[15] O. Brandière. The dynamic system method and the traps. *Adv. in Appl. Probab.*, t. 30, no.1 :137–151, 1998.

[16] O. Brandière. Some pathological traps for stochastic approximation. *SIAM J. Control Optim.*, t. 36, no.4 :1293–1314, 1998.

[17] O. Brandière and M. Duflo. Les algorithmes stochastiques contournent-ils les pièges ? *Ann. Inst. Henri. Poincaré Probab. Statist.*, t. 32, no 3 :395–427, 1996.

[18] C. C. Conley. Isolated invariant sets and the morse index. *CBMS Regional conference series in mathematics, American Mathematical Society, Providence*, 38, 1978.

[19] D. Coppersmith and P. Diaconis. Random walks with reinforcement. *Unpublished manuscript*, 1986.

[20] M. Cranston and Y. Le Jan. Self-attracting diffusions : two cases studies. *Math. Ann*, 303 :87–93, 1995.

[21] D. Dacunha-Castelle and M. Duflo. *Probabilités et statistiques*, volume 2. Masson, 1993.

[22] B. Davis. Reinforced random walk. *Probab. Th. Rel. Fields*, 84 :203–229, 1990.

[23] M. Duflo. *Algorithmes stochastiques.* Mathématiques et Applications, Springer-Verlag, 1996.

[24] M. Duflo. Cibles atteignables avec probabilité positive d'après M. Benaïm. *Unpublished manuscript*, 1997.

[25] J. C. Fort et G. Pagès. Réseaux de neurones : des méthodes connexionnistes d'apprentissage. *Matapli*, 37 :31–48, 1994.

[26] W. Feller. *An introduction to probability theory and its applications 1, 2.* Wiley, New York, 1957, 1968.

[27] D. Fudenberg and D. Levine. *Theory of Learning in Games.* MIT Press, Cambridge, MA, 1998.

[28] S. Herrmann and B. Roynette. Boundedness and convergence of some self-attractive diffusions. *Prépublications de l'Institut Elie Cartan, 10*, 2001.

[29] J. Kiefer and J. Wolfowitz. Stochastic estimation of the maximum of a regression function. *Ann. Math. Statis.*, 23 :462–466, 1952.

[30] H. J. Kushner and C. C Clark. *Stochastic Approximation for Constrained and Unconstrained Systems.* Springer-Verlag, Berlin and New York, 1978.

[31] H. J. Kushner and G. Yin. *Stochastic Approximation Algorithms and Applications.* Springer-Verlag, New York, 1997.

[32] L. Ljung. Analysis of recursive stochastic algorithms. *IEEE Trans. Automat. Control*, AC-22 :551–575, 1977.

[33] L. Ljung. *System Identification Theory for the User.* Prentice Hall, Englewood Cliffs, NJ, 1986.

[34] L. Ljung and T. Sőderstrőm. *Theory and Practice of Recursive Identification.* MIT Press, Cambridge, MA, 1983.

[35] M. B. Nevelson and R. Z. Khasminskii. *Stochastic Approximation and Recursive Estimation*, volume 47. Translation of Math. Monographs American Mathematical Society, Providence, 1976.

[36] R. Pemantle. Phase transition in reinforced random walk and rwre on trees. *Annals of probability*, 16 :1229–1241, 1988.

[37] R. Pemantle. Random processes with reinforcement. *Massachussets Institute of Technology doctoral dissertation*, 1988.

[38] R. Pemantle. Nonconvergence to unstable points in urns models and stochastic approximation. *Annals of Probability*, 18 :698–712, 1990.

[39] R. Pemantle. Vertex reinforced random walk. *Probability Theory and Related Fields*, 92 :117–136, 1992.

[40] R. Pemantle and S. Volkov. Vertex reinforced random walk on $\mathbb{Z}$ has finite range. *Annals of probability*, 27 :No.3,1368–1388, 1999.

[41] O. Raimond. Self-attracting diffusions : case of the constant interaction. *Prob. Theor. Relat. Fields*, 107 :177–196, 1996.

[42] H. Robbins and S. Monro. A stochastic approximation method. *Ann. Math. Statis.*, 22 :400–407, 1951.

[43] C. Robinson. *Introduction to the Theory of Dynamical Systems.* Studies in Advances Mathematics. CRC Press. Boca Raton, 1995.

[44] S. J. Schreiber. Urn models, replicator processes, and random genetic drift. *SIAM J. Appl. Math.*, 61 :2148–2167, 2001.

[45] T. Sellke. Reinforced random walk on the d-dimensionnal integer lattice. *Technical report, no.94-26, Department of statistics, Purdue university*, 1994.

[46] A. Shiryaev. *Probability (2nd. ed.).* Springer, New York, 1989.

[47] M. Shub. *Global Stability of Dynamical Systems.* Springer, NewYork, 1987.

[48] B. Skyrms and R. Pemantle. A dynamic model of social network formation. *Preprint*, 2000.

[49] S. Volkov. Vertex reinforced random walk on arbitrary graphs. *To appear in Annals of probability*, 2000.

[50] H. White. *Artificial Neural Networks : Approximation and Learning Theory.* Blackwell, Cambridge, Massachussets, 1992.

Chapitre 2

Pièges répulsifs

C. R. Acad. Sci. Paris, t. 330, Série I, p. 125–130, 2000
Probabilités/*Probability Theory*
(Statistique/*Statistics*)

Pièges répulsifs

Pierre TARRÈS

Centre de mathématiques et leurs applications, École normale supérieure de Cachan, 61, avenue du Président-Wilson, 94230 Cachan, France
Courriel : tarres@cmla.ens-cachan.fr

(Reçu le 6 juin 1999, accepté après révision le 9 novembre 1999)

Résumé. Nous donnons des hypothèses faibles assurant qu'un algorithme stochastique contourne les pièges en dimension 1. Ce qui permet, dans le cas plus général de la dimension finie, d'élargir le cadre de théorèmes montrant qu'on évite les points instables, les orbites périodiques et ensembles normalement hyperboliques. © 2000 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

Repulsive traps

Abstract. *We give weak assumptions ensuring that a stochastic algorithm does not fall into a repulsive trap in the one-dimensionnal case. This enables us, in the more general finite-dimensionnal case, to extend theorems relevant to unstable points, periodic orbits and normally hyperbolic sets.* © 2000 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

Abridged English Version

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ a nondecreasing sequence of sub-σ-algebras of $\mathcal{F}$. Consider a discrete time process (X_n) living in $\mathbb{R}$ and satisfying equation (1), where $V \in \mathbb{R}$ is a neighbourhood of 0, F a function: $V \to \mathbb{R}$, (ε_n), (r_n) are adapted random variables on $\mathbb{R}$ (i.e. ε_n, r_n are mesurable with respect to $\mathcal{F}_n$ for each $n \geq 0$), (γ_n), (c_n) deterministic sequences of nonnegative numbers, (γ_n) having infinitely positive terms.

Let $\alpha_n = \sum_{i=n}^{+\infty} c_i^2$ for $n \in \mathbb{N}$. We make the following assumptions:

H-1 on perturbations (equations in the french text);
H-2 (γ_n) bounded and $[\gamma_n = \mathrm{O}(\sqrt{\alpha_n})$ or $|F(x)| = \mathrm{O}(|x|)]$;
H-3 F is a bounded Borel function, nonnegative on $V \cap \mathbb{R}_*^+$ and nonpositive on $V \cap \mathbb{R}_*^-$.

THEOREM 1. – *Let* (X_n) *be an algorithm satisfying equation* (1). *Under hypotheses* H-1, H-2 *and* H-3, $P\big(\lim_{n\to\infty} X_n = 0\big) = 0$.

Brandière and Duflo have proved that $P(\lim_{n\to\infty} X_n = 0) = 0$ under the assumptions (a1) or (a2) of the French text ([2–4], and [5], paragraph 3.IV).

Note présentée par Paul DEHEUVELS.

This one-dimensionnal result enables us to extend a result of Benaïm giving conditions ensuring that a Robbins–Monro algorithm does not fall into normally hyperbolic sets. The case of normally hyperbolic sets includes those of linearly unstable equilibria and periodic orbits.

Consider a sequence (x_n) of adapted random variables living in $\mathbb{R}^m$ and satisfying equation (2), where $F:\mathbb{R}^m\to\mathbb{R}^m$ is a smooth vector field generating the flow $\Phi=\{\Phi_t\}$, and define $V_{n+1}=F(x_n)+U_{n+1}+r_{n+1}$. Let Γ denote a nonempty compact set invariant under Φ. We define assumptions H-4 on steps and perturbations, H-5 : Γ is a normally hyperbolic set, and H-6 : there exist a neighbourhood $\mathcal{N}(\Gamma)$ of Γ and $b>0$ such that, for all unit vector $v\in\mathbb{R}^m$, $\mathrm{E}\big(|\langle V_{n+1},v\rangle|\mid\mathcal{F}_n\big)\geq b\,1_{\{x_n\in\mathcal{N}(\Gamma)\}}$.

THEOREM 2. – *Let* $\{x_n\}$ *be an algorithm given by* (2). *Suppose assumptions* H-4, H-5 *and* H-6 *are satisfied, and* F *and* S *are* $\mathrm{C}^{1+\alpha}$ *with* $1/2<\alpha\leq 1$*; then* $P\big(\lim_{n\to\infty}\mathrm{d}(x_n,\Gamma)=0\big)=0$.

---

1. Introduction

Nous nous plaçons dans un espace de probabilité $(\Omega,\mathcal{F},P)$ muni d'une filtration $\mathbb{F}=(\mathcal{F}_n)_{n\in\mathbb{N}}$.

Le principal résultat présenté ici concerne les algorithmes stochastiques à valeurs dans $\mathbb{R}$; nous explicitons des conditions sous lesquelles un tel algorithme contourne les pièges.

Plus précisément, nous considérons une suite de variables aléatoires (X_n) adaptées à la filtration $\mathbb{F}$ vérifiant :

$$X_{n+1}=X_n+\gamma_{n+1}\,F(X_n)+c_{n+1}\,(\varepsilon_{n+1}+r_{n+1})\ \text{ si } X_n\in V, \tag{1}$$

où V est un voisinage de 0 dans $\mathbb{R}$, F une fonction de V dans $\mathbb{R}$, (ε_n), (r_n) sont des suites réelles adaptées à la filtration $\mathbb{F}$, (γ_n), (c_n) des suites déterministes à valeurs dans $\mathbb{R}^+$ et (c_n) a une infinité de termes non nuls.

Posons $\alpha_n=\sum_{i=n}^{+\infty}c_i^2$ pour $n\in\mathbb{N}$. Appelons H-1 l'hypothèse suivante sur les perturbations :

- $\mathrm{E}(\varepsilon_{n+1}\mid\mathcal{F}_n)=0$, $\liminf\limits_{n\to\infty}\mathrm{E}(\varepsilon_{n+1}^2\mid\mathcal{F}_n)>0$ et $\exists\, a>2$, $\limsup\limits_{n\to\infty}\mathrm{E}(|\varepsilon_{n+1}|^a\mid\mathcal{F}_n)<+\infty$,
- $\sum_n r_n^2<+\infty$,

et H-2 l'hypothèse : (γ_n) bornée et $[\gamma_n=\mathrm{O}\big(\sqrt{\alpha_n}\big)$ ou $|F(x)|=\mathrm{O}(|x|)]$.

Il s'agit de savoir quelles conditions sur les perturbations, les pas et la fonction F au voisinage de 0 entraînent que l'algorithme contourne presque sûrement ce point 0.

Les travaux de Brandière et Duflo permettent d'affirmer que l'on a bien $P\big(\lim_{n\to\infty}X_n=0\big)=0$ dans les deux cas suivants (*voir* [2], [3], [4], et [5], paragraphe 3.IV) :

(a1) F est différentiable, F' est lipschitzienne et $F'(0)>0$, $\sum r_n^2<+\infty$, $\mathrm{E}(\varepsilon_{n+1}\mid\mathcal{F}_n)=0$, $\liminf\limits_{n\to\infty}\mathrm{E}(|\varepsilon_{n+1}|\mid\mathcal{F}_n)>0$, $\limsup\limits_{n\to\infty}\mathrm{E}(\varepsilon_{n+1}^2\mid\mathcal{F}_n)<+\infty$, $\gamma_n=\mathrm{O}(c_n)$, $\sum\gamma_n=+\infty$, $\sum c_n^2<+\infty$

(a2) F est C^1 et $F(x)\sim\alpha x^{2k+1}$ avec $\alpha>0$ et $k\in\mathbb{N}$, $\gamma_n=c_n=\frac{g}{n}$, $r_n=0$ et H-1.

Le théorème montré en deuxième partie vient en prolongement de ces résultats. Il se base sur l'hypothèse suivante sur la fonction F :

H-3 $\quad F:\mathbb{R}\longrightarrow\mathbb{R}$ borélienne bornée positive ou nulle sur $V\cap\mathbb{R}_*^+$ et négative ou nulle sur $V\cap\mathbb{R}_*^-$,

et sur H-1 et H-2.

Il permet d'affaiblir les conditions nécessaires à des résultats déjà obtenus sur les points instables, les orbites périodiques et ensembles normalement hyperboliques (*voir* [1]) : c'est l'objet de la troisième partie.

2. Les pièges en dimension 1

THÉORÈME 1. – *Soit* (X_n) *un algorithme vérifiant* (1) ; *alors, sous les hypothèses* H-1, H-2 *et* H-3, $P\big(\lim_{n\to+\infty}X_n=0\big)=0$.

La démonstration repose sur deux lemmes. Le premier s'inspire fortement d'un résultat de Benaïm (*voir* [1], lemme 9.6, également utilisé dans le corps de la démonstration), basé sur des travaux de Pemantle ([6] et [7]), et le généralise. Le deuxième reproduit l'idée qui sous-tend ce travail ; il utilise des estimations de $\sum c_i\, \varepsilon_i$ pour le contrôle du processus.

Suivant une astuce due à Lai et Wei (*voir* [5], 2.III.2), on peut se limiter à établir le théorème dans le cadre simplifié suivant :

(α) pour des constantes A, $B \in \mathbb{R}_*$, $\forall n \in \mathbb{N}$, $\mathrm{E}(\varepsilon_{n+1}^2 \mid \mathcal{F}_n) \geq 2A$, $\mathrm{E}(|\varepsilon_{n+1}|^a \mid \mathcal{F}_n) \leq B^{\frac{a}{2}}$.

De plus, quitte à remplacer (X_n) par (X_{n+n_1}) pour n_1 assez grand et à prolonger F sur $\mathbb{R} \setminus V$ par une fonction bornée positive sur $\mathbb{R}^+ \setminus V$ et négative sur $\mathbb{R}^- \setminus V$, on peut supposer sans perte de généralité, $\eta > 0$ étant fixé :

(β) $\sum_{n \geq 0} r_{n+1}^2 \leq \eta$ et $V = \mathbb{R}$.

Nous commençons par envisager le cas $\sum c_n^2 < +\infty$. Dans ce qui suit, jusqu'à la démonstration du théorème, nous considérons une suite adaptée (X_n) qui ne vérifie pas nécessairement (1).

Notons $s_n = \sum_{i=n}^{+\infty} r_i^2$. Nous désignerons, pour tout $n \in \mathbb{N}$, par S_n^L, P_n^L, N_n^L et T_n^L les temps d'arrêt définis par $S_n^L = \inf\{m \geq n \mid X_m^2 \geq L\, \alpha_{m+1}\}$, $P_n^L = \inf\{m \geq n \mid X_m \geq \sqrt{L\, \alpha_{m+1}}\}$, $N_n^L = \inf\{m \geq n \mid X_m \leq -\sqrt{L\, \alpha_{m+1}}\}$ et $T_n^L = \inf\{m \geq n \mid X_m^2 \geq L\}$. Nous noterons $H = \{\liminf |X_n| > 0\}$.

Nous notons H'-1 l'hypothèse : $\mathrm{E}(|\varepsilon_{n+1}|^a \mid \mathcal{F}_n) \leq B^{\frac{a}{2}}$, $\sum_{n \geq 0} r_{n+1}^2 \leq \eta$ et $V = \mathbb{R}$.

LEMME 1. – *Supposons $\sum c_n^2 < +\infty$, H'-1 avec $\eta \leq \frac{A^2}{16L}$, et il existe $c \in \mathbb{R}_*^+$ tel que :*
(i) $\mathrm{E}(X_{i+1}^2 - X_i^2 \mid \mathcal{F}_i) \geq A\, c_{i+1}^2 - 2c_{i+1}\mathrm{E}(|r_{i+1}| \mid \mathcal{F}_i)\, |X_i|$
(ii) $(X_{i+1} - X_i)^2 \leq c\, (\alpha_{n+1} + c_{i+1}^2 \varepsilon_{i+1}^2)$ *pour tout $n \in \mathbb{N}$ tel que $n \leq i < S_n^L$.*

Alors il existe une constante $C \in]0,1[$ telle que pour tout $n \in \mathbb{N}$, $P\big(S_n^L < \infty \mid \mathcal{F}_n\big) \geq C$.

Démonstration. – Remplaçons, pour simplifier, S_n^L par S. Sur $\{n \leq k < S\}$, $|X_k| \leq \sqrt{L\alpha_{k+1}} \leq \sqrt{L\alpha_{n+1}}$. On en déduit, en posant $Z_k = X_k^2 - A \sum_{i=n}^k c_i^2 + 2\sqrt{L\alpha_{n+1}} \sum_{i=n}^k c_i\, |r_i|$ et en utilisant (i), que $\{Z_{k \wedge S}\}_{k \geq n}$ est une sous-martingale. Donc, si $m \geq n$, $\mathrm{E}(Z_{m\wedge S} - Z_n \mid \mathcal{F}_n) \geq 0$.

Ainsi,

$$\mathrm{E}(X_{m\wedge S}^2 - X_n^2 \mid \mathcal{F}_n) \geq \left(A \sum_{i=n+1}^{m} c_i^2 - 2\sqrt{L\alpha_{n+1}}\, \frac{\sqrt{\alpha_{n+1}}A}{4\sqrt{L}}\right) P(S > m \mid \mathcal{F}_n)$$
$$- 2\sqrt{L\alpha_{n+1}}\, \frac{\sqrt{\alpha_{n+1}}A}{4\sqrt{L}} \big(1 - P(S > m \mid \mathcal{F}_n)\big)$$

puisque, d'après l'inégalité de Cauchy–Schwarz, $\sum_{i=n+1}^m c_i\, |r_i| \leq \sqrt{\alpha_{n+1} s_{n+1}} \leq \frac{\sqrt{\alpha_{n+1}}A}{4\sqrt{L}}$. Donc

$$\mathrm{E}\big(X_{m\wedge S}^2 - X_n^2 \mid \mathcal{F}_n\big) \geq A\big(\tfrac{\alpha_{n+1}}{2} - \alpha_{m+1}\big) P(S > m \mid \mathcal{F}_n) - \tfrac{A}{2}\alpha_{n+1}\left(1 - P(S > m \mid \mathcal{F}_n)\right).$$

D'autre part, $X_{m\wedge S}^2 = X_m^2 \leq L\, \alpha_{m+1}$ si $S > m$. Donc $X_{m\wedge S}^2 - X_n^2 \leq X_S^2\, 1_{n<S\leq m} + L\alpha_{m+1} 1_{S>m} \leq 2\big[X_{S-1}^2\, 1_{n<S\leq m} + (X_S - X_{S-1})^2\, 1_{n<S\leq m}\big] + L\, \alpha_{m+1}\, 1_{S>m}$. Puisque $X_{S-1}^2 \leq L\, \alpha_S \leq L\, \alpha_{n+1}$ si $n < S \leq m$, on peut écrire, en utilisant (ii)

$$\begin{aligned}
&\mathrm{E}\big(X_{m\wedge S}^2 - X_n^2 \mid \mathcal{F}_n\big) \\
&\leq 2L\big(P(S > m \mid \mathcal{F}_n)\alpha_{m+1} + \big(1 - P(S > m \mid \mathcal{F}_n)\big)\alpha_{n+1}\big) + 2\mathrm{E}\big((X_S - X_{S-1})^2\, 1_{n<S\leq m} \mid \mathcal{F}_n\big) \\
&\leq 2(L + c)\big(P(S > m \mid \mathcal{F}_n)\alpha_{m+1} + \big(1 - P(S > m \mid \mathcal{F}_n)\big)\alpha_{n+1}\big) + 2c \sum_{i=n+1}^{m} c_i^2\, \mathrm{E}\big(\varepsilon_i^2\, 1_{S=i} \mid \mathcal{F}_n\big)
\end{aligned}$$

D'après l'inégalité de Hölder, $\mathrm{E}(\varepsilon_i^2\, 1_{S=i} \mid \mathcal{F}_n) \leq \mathrm{E}(|\varepsilon_i|^a \mid \mathcal{F}_n)^{\frac{2}{a}} P(S = i \mid \mathcal{F}_n)^{1-\frac{2}{a}} \leq B\, P(S = i \mid \mathcal{F}_n)^{1-\frac{2}{a}}$ si $i > n$. Donc $\sum_{i=n+1}^m c_i^2\, \mathrm{E}(\varepsilon_i^2\, 1_{S=i} \mid \mathcal{F}_n) \leq B\, \alpha_{n+1}\big(1 - P(S > m \mid \mathcal{F}_n)\big)^{1-\frac{2}{a}}$. Finalement,

$$P(S > m \mid \mathcal{F}_n)\left[\frac{A}{2}\, \alpha_{n+1} - \big[A + 2(L + c)\big]\alpha_{m+1}\right] \leq \big(1 - P(S > m \mid \mathcal{F}_n)\big)^{1-\frac{2}{a}} \left[\frac{A}{2} + 2(cB + L + c)\right]\alpha_{n+1}$$

soit, en posant $b = \frac{1}{1-\frac{2}{a}} \geq 1$, $p = P(S = +\infty \mid \mathcal{F}_n)$ et $t = \left(\frac{A+4(cB+L+c)}{A}\right)^b$, $p^b \leq (1-p)t$.

Donc il existe $C \in \mathbb{R}_*^+$, ne dépendant que de A, B, L et c tel que $P(S = +\infty \mid \mathcal{F}_n) \leq 1 - C$. □

LEMME 2. – *Supposons* $\sum c_i^2 < +\infty$, H'-1 *avec* $\eta \leq 1$, $L \geq \left(\sqrt{8B}+2\right)^2$ *et, pour tout* $i \in \mathbb{N}$, $X_{i+1} \geq X_i + \gamma_{i+1}F(X_i) + c_{i+1}(\varepsilon_{i+1} + r_{i+1})$ *et* $\mathrm{E}(\varepsilon_{i+1} \mid \mathcal{F}_i) \geq 0$.

Alors, pour tout $n \in \mathbb{N}$ *et* $m \geq n$, $P(H \mid \mathcal{F}_m)\, 1_{P_n^L = m} \geq \frac{1}{2}\, 1_{P_n^L = m}$.

Démonstration. – a) Nous avons supposé $\mathrm{E}(\varepsilon_{i+1}^2 \mid \mathcal{F}_i) \leq B$. Posons $R_{m+1} = \inf_{p>m} \sum_{i=m+1}^p c_i \varepsilon_i$.

Soit $V_i = \varepsilon_i - \mathrm{E}(\varepsilon_i \mid \mathcal{F}_{i-1})$: alors, pour tout i, $V_i \leq \varepsilon_i$, $\mathrm{E}(V_{i+1} \mid \mathcal{F}_i) = 0$ et $\mathrm{E}(V_{i+1}^2 \mid \mathcal{F}_i) \leq \mathrm{E}(\varepsilon_{i+1}^2 \mid \mathcal{F}_i)$. D'après l'inégalité de Doob :

$$\mathrm{E}\left(\sup_{p>m} \Big| \sum_{i=m+1}^{p} c_i\, V_i\Big|^2 \mid \mathcal{F}_m\right) \leq 4\mathrm{E}\Big(\sum_{i=m+1}^{\infty} c_i^2\, V_i^2 \mid \mathcal{F}_m\Big) \leq 4 \sup_{i \geq m} \mathrm{E}(V_{i+1}^2 \mid \mathcal{F}_m)\alpha_{m+1} \leq 4B\, \alpha_{m+1},$$

donc, si $k \in \mathbb{R}_*$, $P\big(R_{m+1} \leq -k\sqrt{\alpha_{m+1}} \mid \mathcal{F}_m\big) \leq \frac{4B}{k^2}$. On en déduit, en posant $k_0 = \sqrt{8B}$,

$$P\left(R_{m+1} \geq -k_0\sqrt{\alpha_{m+1}} \mid \mathcal{F}_m\right) \geq \frac{1}{2}.$$

b) L'inégalité de Cauchy–Schwarz permet d'écrire $\sum_{i=m+1}^p |c_i\, r_i| \leq \sqrt{\alpha_{m+1}\, s_{m+1}} \leq \sqrt{\alpha_{m+1}}$. Si $P_n^L = m$ (donc $X_m \geq (k_0+2)\sqrt{\alpha_{m+1}}$) et $R_{m+1} \geq -k_0\sqrt{\alpha_{m+1}}$, alors, pour tout $p > m$,

$$X_p - X_m \geq \sum_{i=m}^{p-1} \gamma_{i+1}\, F(X_i) + \sum_{i=m+1}^{p} c_i\, \varepsilon_i - \sum_{i=m+1}^{p} |c_i\, r_i| \geq -(k_0+1)\sqrt{\alpha_{m+1}}\,;$$

par conséquent, $\forall\, p > m$, $X_p \geq \sqrt{\alpha_{m+1}}$. En résumé, $H \supset \big\{R_{m+1} \geq -k_0\sqrt{\alpha_{m+1}}$ et $P_n^L = m\big\}$. D'où il résulte que $P(H \mid \mathcal{F}_m)\, 1_{P_n^L=m} \geq \frac{1}{2}\, 1_{P_n^L=m}$. □

Démonstration du théorème 1. – Pour des raisons qui seront clarifiées par la démonstration, nous choisissons $L = \left(\sqrt{8B}+2\right)^2$ et $\eta = \min\left(\frac{A^2}{4B}, \frac{A^2}{16L}, 1\right)$. X_n vérifie (1), donc en développant l'espérance du carré de X_{n+1} :

$$\mathrm{E}\big(X_{n+1}^2 \mid \mathcal{F}_n\big) - 2\gamma_{n+1}X_n\, F(X_n) = X_n^2 + c_{n+1}^2\, \mathrm{E}\big(\varepsilon_{n+1}^2 \mid \mathcal{F}_n\big) + 2c_{n+1}\, \mathrm{E}(r_{n+1} \mid \mathcal{F}_n)X_n$$
$$+2c_{n+1}^2\, \mathrm{E}(\varepsilon_{n+1}r_{n+1} \mid \mathcal{F}_n) + \gamma_{n+1}F(X_n)(\gamma_{n+1}F(X_n) + 2c_{n+1}\, \mathrm{E}(r_{n+1} \mid \mathcal{F}_n)) + c_{n+1}^2\, \mathrm{E}\big(r_{n+1}^2 \mid \mathcal{F}_n\big)$$

L'hypothèse H-3 nous permet d'affirmer que $\mathrm{E}\big(X_{n+1}^2 \mid \mathcal{F}_n\big)$ est supérieur ou égal au deuxième membre de l'égalité. D'autre part, $\gamma_{n+1}\, F(X_n)\big(\gamma_{n+1}F(X_n) + 2c_{n+1}\, \mathrm{E}(r_{n+1} \mid \mathcal{F}_n)\big) \geq -c_{n+1}^2\, \mathrm{E}\big(r_{n+1}^2 \mid \mathcal{F}_n\big)$ et $\mathrm{E}(\varepsilon_{n+1}r_{n+1} \mid \mathcal{F}_n) \geq -\sqrt{B\, \mathrm{E}\big(r_{n+1}^2 \mid \mathcal{F}_n\big)}$ (inégalité de Cauchy–Schwarz). Ainsi,

$$\mathrm{E}\big(X_{n+1}^2 - X_n^2 \mid \mathcal{F}_n\big) \geq c_{n+1}^2\Big(2A - 2\sqrt{B\eta}\Big) - 2c_{n+1}\, \mathrm{E}(|r_{n+1}| \mid \mathcal{F}_n)\, |X_n|\,;$$

donc X_n vérifie la condition (i) du lemme 1 à cause du choix de η.

La condition (ii) de ce lemme est vérifiée : $(X_{i+1} - X_i)^2 \leq 3\big(\gamma_{i+1}^2(F(X_i))^2 + c_{i+1}^2\, r_{i+1}^2 + c_{i+1}^2\, \varepsilon_{i+1}^2\big)$ et, si $S_n^L > i \geq n$, l'hypothèse H-2 implique $\gamma_{i+1}^2(F(X_i))^2 = \mathrm{O}(\alpha_{n+1})$: c'est immédiat si $\gamma_p = \mathrm{O}\big(\sqrt{\alpha_p}\big)$ (F est supposée bornée). Si (γ_p) est seulement supposée bornée et $|F(x)| = \mathrm{O}(|x|)$, alors $i < S_n^L \Rightarrow |X_i| \leq \sqrt{L\alpha_{n+1}}$, ce qui permet de conclure.

Par conséquent, $P\big(S_n^L < +\infty \mid \mathcal{F}_n\big) \geq C$.

Remarquons que $1_{S_n^L=m} = 1_{P_n^L=m} + 1_{N_n^L=m}$. D'après le lemme 2, $P(H \mid \mathcal{F}_m)\, 1_{P_n^L=m} \geq \frac{1}{2}\, 1_{P_n^L=m}$; par un raisonnement analogue, $P(H \mid \mathcal{F}_m)\, 1_{N_n^L=m} \geq \frac{1}{2}\, 1_{N_n^L=m}$. Finalement, $P(H \mid \mathcal{F}_m)\, 1_{S_n^L=m} \geq \frac{1}{2}\, 1_{S_n^L=m}$. Donc

$$\mathrm{E}(1_H \mid \mathcal{F}_n) \geq \sum_{i \geq n} \mathrm{E}\left(1_H 1_{S_n^L = i} \mid \mathcal{F}_n\right) \geq \sum_{i \geq n} \mathrm{E}\left(\mathrm{E}(1_H \mid \mathcal{F}_i) 1_{S_n^L = i} \mid \mathcal{F}_n\right) \geq \frac{1}{2} \sum_{i \geq n} \mathrm{E}\left(1_{S_n^L = i} \mid \mathcal{F}_n\right) \geq \frac{C}{2}.$$

Puisque H est $\mathcal{F}_\infty$-mesurable, $\lim_{n \to \infty} \mathrm{E}(1_H \mid \mathcal{F}_n) = 1_H$ presque partout, on en déduit $P(H) = 1$. □

Dans le cas $\sum c_n^2 = +\infty$, la condition (i) du lemme 1 reste satisfaite et l'on peut également vérifier (ii)$'$: il existe $c \in \mathbb{R}_*^+$ tel que $(X_{i+1} - X_i)^2 \leq c\big(1 + c_{i+1}^2 + c_{i+1}^2\, \varepsilon_{i+1}^2\big)$. Cela permet de montrer, sur le même principe que le lemme 1, qu'il existe une constante $C \in]0,1[$ telle que, pour tout $n \in \mathbb{N}$, $P\big(T_n^L < \infty \mid \mathcal{F}_n\big) \geq C$. En effet, si l'on pose $Z_k = X_k^2 - A\sum_{i=n}^k c_i^2 + 2\sqrt{L}\sum_{i=n}^k c_i\,|r_i|$ et $S_{n,m} = \sum_{i=n+1}^m c_i^2$, on déduit que

$$\mathrm{E}\left(X_{m \wedge T_n^L}^2 - X_n^2 \mid \mathcal{F}_n\right) \geq \frac{A}{2} S_{n,m} P\big(T_n^L > m \mid \mathcal{F}_n\big) - \frac{A}{2}\sqrt{S_{n,m}}\,\big(1 - P\big(T_n^L > m \mid \mathcal{F}_n\big)\big) \text{ si } S_{n,m} \geq 1$$

puis, en majorant $X_{m \wedge T_n^L}^2$:

$$\left(\frac{A}{2} S_{n,m} - L) P(T_n^L > m \mid \mathcal{F}_n\right) \leq \left[(2c(1+B) + \frac{A}{2}) S_{n,m} + 2(L+c)\right] (1 - P(T_n^L > m \mid \mathcal{F}_n)^{1-\frac{2}{a}},$$

ce qui permet de conclure.

Remarque 1 . – La preuve, dans le cas $\sum c_n^2 < +\infty$, montre un peu plus que $P(X_n \to 0) = 0$: elle permet d'affirmer que l'algorithme reste p.s. sur une demi-droite à partir d'un certain rang. Ceci n'est plus vrai lorsque $\sum c_n^2 = +\infty$: le lemme 2 n'a pas d'équivalent vérifié de façon générale dans ce cas.

3. Les ensembles normalement hyperboliques en dimension finie

Le principe développé en dimension 1 permet d'élargir le cadre d'application d'un résultat dû à Benaïm ([1], théorème 9.1) donnant des conditions suffisantes pour qu'un algorithme de Robbins–Monro contourne presque sûrement des ensembles normalement hyperboliques en dimension finie.

La notion d'ensembles normalement hyperboliques inclut celles d'équilibres et d'orbites périodiques linéairement instables ([1], paragraphe 9).

Il s'agit d'étudier (x_n), suite de variables aléatoires à valeurs dans $\mathbb{R}^m$ adaptées à la filtration $\mathbb{F}$, définie par la forme itérative :

$$x_{n+1} - x_n = \gamma_{n+1}\,(F(x_n) + U_{n+1} + r_{n+1}), \tag{2}$$

F étant un champ de vecteurs continu : $\mathbb{R}^m \to \mathbb{R}^m$ engendrant le flot $\Phi = \{\Phi_t\}$. Nous appelons H-4 l'hypothèse : $\{\gamma_n\}_{n \in \mathbb{N}}$ est une suite déterministe à valeurs dans $\mathbb{R}^+$ ayant une infinité de termes non nuls, $\{U_n\}$, $\{r_n\}$ sont des perturbations aléatoires adaptées à la filtration $\mathbb{F}$, $\mathrm{E}(U_{n+1} \mid \mathcal{F}_n) = 0$ et (U_n) bornée ($\exists\, K > 0$, $\forall\, n \in \mathbb{N}$, $\|U_n\| \leq K$), $\sum \|r_i\|^2 < +\infty$.

Le théorème s'intéresse à des ensembles Γ invariants par Φ et vérifiant les conditions suivantes H-5 : $\Gamma \subset S$, où S est une sous-variété de $\mathbb{R}^m$ de dimension $(m-d)$, $d \in \{1, \ldots, m\}$, localement invariante (il existe un voisinage U de Γ dans $\mathbb{R}^m$ et $t_0 \in \mathbb{R}_*$ tels que, $\forall\, t \in \mathbb{R}$, $|t| \leq t_0$, $\Phi_t(U \cap S) \subset S$), enfin $\mathbb{R}^m = \mathrm{T}_p S \oplus E_p^u$ pour tout point $p \in \Gamma$ avec : (i) $p \mapsto E_p^u$ continue de Γ dans la variété grassmannienne $\mathbb{G}(d,m)$ des plans de dimension d dans $\mathbb{R}^m$; (ii) $\mathrm{D}\Phi_t(p)\, E_p^u = E_{\Phi_t(p)}^u$ pour tout $t \in \mathbb{R}$, $p \in E_p^u$; (iii) il existe $\lambda > 0$ et $C > 0$ tels que, pour tout $p \in \Gamma$, $w \in E_p^u$ et $t \geq 0$, $\|\mathrm{D}\Phi_t(p) w\| \geq C\,\mathrm{e}^{\lambda t}\,\|w\|$.

Définissons $V_{n+1} = F(x_n) + U_{n+1} + r_{n+1}$. Appelons H-6 l'hypothèse : il existe un voisinage $\mathcal{N}(\Gamma)$ de Γ et $b > 0$ tels que, pour tout vecteur unité $v \in \mathbb{R}^m$, $\mathrm{E}\big(|\langle V_{n+1}, v\rangle| \mid \mathcal{F}_n\big) \geq b\, 1_{\{x_n \in \mathcal{N}(\Gamma)\}}$.

THÉORÈME 2. – *Soit* $\{x_n\}$ *un algorithme défini par* (2). *Nous faisons les hypothèses* H-4, H-5 *et* H-6, *et supposons que* F *et* S *sont* $\mathrm{C}^{1+\alpha}$ *avec* $1/2 < \alpha \leq 1$ *; alors* $P\left(\lim_{n\to\infty} \mathrm{d}(x_n, \Gamma) = 0\right) = 0$.

Le changement par rapport au théorème déjà existant ([1], théorème 9.1) est de ne pas faire l'hypothèse sur les pas $\lim\limits_{n\to\infty} \gamma_{n+1}^{\alpha} / \sqrt{\lim\limits_{m\to\infty} \sum_{i=n+1}^{m} \gamma_i^2} = 0$ et d'ajouter un terme de perturbation $\{r_n\}$ avec $\sum \|r_i\|^2 < +\infty$ (le résultat est énoncé dans [1] avec $0 < \alpha \leq 1$ mais la preuve nécessite l'hypothèse $\alpha > 1/2$).

Démonstration. – Benaïm définit une fonction η lipschitzienne (*voir* [1], proposition 9.5) permettant, si l'on considère le processus $X_n = \eta(x_n)$ à valeurs dans $\mathbb{R}^+$, d'écrire

$$X_{n+1} - X_n = \eta(x_{n+1}) - \eta(x_n) \geq \gamma_{n+1} \beta X_n + \gamma_{n+1}(U'_{n+1} + r'_{n+1}) - k\gamma_{n+1}^{1+\alpha},$$

avec $U'_n = \mathrm{D}\eta(x_{n-1})\, U_n$ et $r'_n = \mathrm{D}\eta(x_{n-1})\, r_n$: U'_n est bornée sur un voisinage compact de Γ et l'inégalité de Jensen permet d'affirmer $\mathrm{E}(U'_{n+1} \mid \mathcal{F}_n) \geq \mathrm{D}\eta(x_n)\, \mathrm{E}(U_{n+1} \mid \mathcal{F}_n) = 0$. $\|\mathrm{D}\eta\|$ étant bornée sur $\mathcal{N}(\Gamma)$, nous pouvons, $\eta > 0$ étant fixé, supposer $\sum {r'_{n+1}}^2 \leq \eta$ par le principe expliqué précédemment.

Le bruit U'_n ainsi obtenu en dimension 1 est assez excitant : on montre que $\mathrm{E}({U'_{n+1}}^+ \mid \mathcal{F}_n) \geq \frac{c_1 b}{3}$ si $x \in \mathcal{N}(\Gamma)$, où c_1 est une constante strictement positive définie dans [1]. En effet, puisque $\mathrm{E}(U'_{n+1} \mid \mathcal{F}_n) \geq 0$, il suffit de montrer que $\mathrm{E}(|U'_{n+1}| \mid \mathcal{F}_n) \geq \frac{2c_1 b}{3}$, et, puisque pour β, $\gamma > 0$, $\gamma\, \eta(x) \geq \mathrm{D}\eta(x) \cdot F(x) \geq \beta\, \eta(x)$ ([1], proposition 9.5), que $\mathrm{E}(|\mathrm{D}\eta(x_n) \cdot V_{n+1}| \mid \mathcal{F}_n) \geq c_1\, b$ (en réduisant $\mathcal{N}(\Gamma)$ si nécessaire, et en prenant η assez petit). Cette dernière inégalité se déduit de H-6 par le même principe que celui développé dans [1], théorème 9.1.

Cela permet d'affirmer, par application de l'inégalité de Jensen en envisageant séparément les cas $X_n \geq \frac{k\,\gamma_{n+1}^{\alpha}}{\beta}$ et $X_n \leq \frac{k\,\gamma_{n+1}^{\alpha}}{\beta}$, que $\mathrm{E}\left(X_{n+1}^2 - X_n^2 \mid \mathcal{F}_n\right) \geq \frac{(c_1\, b)^2}{18}\, \gamma_{n+1}^2 - 2\gamma_{n+1}\, \mathrm{E}(|r'_{n+1}| \mid \mathcal{F}_n)\, |X_n|$ pour η assez petit.

La condition (i) du lemme 1 est donc bien vérifiée avec $c_i = \gamma_i$. La condition (ii) est, elle aussi, vérifiée puisque η lipschitzienne implique $|X_{n+1} - X_n| = |\eta(x_{n+1}) - \eta(x_n)| = \mathrm{O}(\gamma_{n+1})$. Les conditions du lemme 2 sont assurées lorsque $\sum \gamma_n^2 < +\infty$, à ceci près qu'il y a en plus de r'_i un terme $r''_i = \gamma_i^{\alpha}$, qui ne vérifie pas nécessairement $\sum {r''_i}^2 < +\infty$. Ce lemme reste vrai à cause de la forme particulière de F : $F(x) = \beta\, x$. En effet, d'après l'inégalité de Hölder, pour tout $\varepsilon > 0$, $\sum_{i=m+1}^{p} \gamma_i^{1+\alpha} \leq \sqrt{\sum_{i=m+1}^{p} \gamma_i^2} \sqrt{\sum_{i=m+1}^{p} \gamma_i^{2\alpha}} \leq \varepsilon \sqrt{\alpha_{m+1}} \sqrt{\sum_{i=m+1}^{p} \gamma_i} \leq \varepsilon \sqrt{\alpha_{m+1}} \left(1 + \sum_{i=m+1}^{p} \gamma_i\right)$ pour n assez grand car $\alpha > 1/2$.

Nous avons donc, dans la partie (ii) de la démonstration de ce lemme, en choisissant $\varepsilon = \min\left(\beta, \frac{1}{2}\right)$: si $p > m$, $X_p - X_m \geq \sum_{i=m+1}^{p} c_i \varepsilon_i - \frac{3}{2}\sqrt{\alpha_{m+1}}$, et la conclusion reste vraie. □

Remerciements. Je tiens à exprimer ma reconnaissance à mon directeur de thèse Michel Benaïm, qui m'a beaucoup appris.

References

[1] Benaïm M., Dynamics of Stochastic Approximation Algorithms, in: Séminaire de Probabilités, Lect. Notes in Math., Azema, M. Yor (Eds.), Springer-Verlag, 1999 (à paraître).

[2] Brandière O., Some pathological traps for stochastic approximation, SIAM J. Control Optim. 36 (4) (1998) 1293–1314.

[3] Brandière O., The dynamic system method and the traps, Adv. in Appl. Probab. 30 (1) (1998) 137–151.

[4] Brandière O., Duflo M., Les algorithmes stochastiques contournent-ils les pièges ?, Ann. Inst. H.-Poincaré Probab. Statis. 32 (3) (1996) 395–427.

[5] Duflo M., Algorithmes stochastiques, Mathématiques et applications, Springer-Verlag, 1996.

[6] Pemantle R., Nonconvergence to unstable points in urns models and stochastic approximation, Ann. Probab. 18 (1990) 698–712.

[7] Pemantle R., Vertex reinforced random walk, Probab. Th. and Rel. Fields 92 (1992) 117–136.

Chapitre 3

Pièges répulsifs : additif

3.1 Introduction

Nous nous plaçons dans un espace de probabilité $(\Omega, \mathcal{F}, P)$ muni d'une filtration $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$.

Le but de ce travail est de généraliser certains résultats obtenus dans [8], au sujet des conditions sous lesquelles un algorithme stochastique contourne les pièges. Nous expliquons pourquoi une des hypothèses du théorème 1 de [8] sur les pièges dans le cas unidimensionnel est inutile.

Ce travail se réfère à de nombreuses reprises à l'article [8] : il est souhaitable d'avoir lu précédemment ce dernier article, en particulier pour connaître d'autres résultats sur les pièges.

3.2 Les pièges en dimension 1

Nous considérons des suites de variables aléatoires (X_n), (Y_n), adaptées à la filtration $\mathbb{F}$ vérifiant

$$X_{n+1} = Y_n + c_{n+1}(\epsilon_{n+1} + r_{n+1}) \quad \text{si } X_n \in V \tag{3.1}$$

où V est un voisinage de 0 dans $\mathbb{R}$, (ϵ_n), (r_n) sont des suites réelles adaptées à la filtration $\mathbb{F}$, (c_n) une suite déterministe à valeurs dans $\mathbb{R}^+$ et (c_n) a une infinité de termes non nuls.

Posons $\alpha_n = \sum_{i=n}^{+\infty} c_i^2$ pour $n \in \mathbb{N}$. Appelons **A-1** l'hypothèse suivante sur les perturbations

- $E(\epsilon_{n+1}|\mathcal{F}_n) = 0$, $\liminf_{n\to\infty} E(\epsilon_{n+1}^2|\mathcal{F}_n) > 0$ et
  $\exists a > 2 \ \limsup_{n\to\infty} E(|\epsilon_{n+1}|^a|\mathcal{F}_n) < +\infty$
- $\sum_n r_n^2 < +\infty$,

Nous pouvons énoncer le théorème suivant :

Théorème 3.2.1 *Soient (X_n) et (Y_n) des suites de variables aléatoires vérifiant (3.1). Supposons* **A-1** *et, pour tout $n \in \mathbb{N}$, $|Y_n| \geq |X_n|$; alors*

$$P(\lim_{n\to+\infty} X_n = 0) = 0.$$

Une application directe du théorème 3.2.1 est le théorème suivant : conservons les notations précédentes et considérons une suite de variables aléatoires (X_n) adaptées à la filtration $\mathcal{F}$ vérifiant

$$X_{n+1} = X_n + \gamma_{n+1}F(X_n) + c_{n+1}(\epsilon_{n+1} + r_{n+1}) \quad \text{si } X_n \in V, \tag{3.2}$$

où (γ_n) est une suite déterministe à valeurs dans $\mathbb{R}^+$.

Notons **A-2** l'hypothèse suivante sur la fonction F :

$F : \mathbb{R} \to \mathbb{R}$ est borélienne positive ou nulle sur $V \cap \mathbb{R}_*^+$ et négative ou nulle sur $V \cap \mathbb{R}_*^-$.

Théorème 3.2.2 *Soit (X_n) un algorithme vérifiant (3.2); sous les hypothèses* **A-1** *et* **A-2**,

$$P(\lim_{n\to+\infty} X_n = 0) = 0.$$

En comparaison aux hypothèses de [8], **A-1** correspond à l'ancienne hypothèse **H-1**, et **A-2** à **H-3** sans la supposition F bornée.

Nous commençons maintenant la preuve du théorème 3.2.1.

Nous pouvons, sans perte de généralité et pour tout $\eta \in \mathbb{R}_+^*$, nous placer dans le cadre simplifié suivant (voir [8]) :

(α) pour des constantes $A, B \in \mathbb{R}_*^+$, $\forall n \in \mathbb{N}$, $E(\epsilon_{n+1}^2|\mathcal{F}_n) \geq 2A$, $E(|\epsilon_{n+1}|^a|\mathcal{F}_n) \leq B^{\frac{a}{2}}$.

(β) $\sum_{n\geq 0} r_{n+1}^2 \leq \eta$ et $V = \mathbb{R}$.

Nous montrons ce théorème 3.2.2 sous l'hypothèse $\sum c_n^2 < +\infty$, les changements à effectuer étant du même type lorsque $\sum c_n^2 = +\infty$. Les lemmes qui suivent se placent dans l'hypothèse où les processus $(X_n)_{n\in\mathbb{N}}$ et $(Y_n)_{n\in\mathbb{N}}$ sont $\mathbb{F}$-adaptés mais ne vérifient pas forcément les équations (3.2) ou (3.1).

Nous notons $s_n = \sum_{i=n}^{+\infty} r_i^2$, et définissons, pour tout $n \in \mathbb{N}$, les temps d'arrêt S_n^L, $S_n'^L$, P_n^L, N_n^L et T_n^L par $S_n^L = \inf\{m \geq n / X_m^2 \geq L\alpha_{m+1}\}$, $S_n'^L = \inf\{m \geq n / Y_m^2 \geq L\alpha_{m+1}\}$, $P_n^L = \inf\{m \geq n / X_m \geq \sqrt{L\alpha_{m+1}}\}$, $N_n^L = \inf\{m \geq n / X_m \leq -\sqrt{L\alpha_{m+1}}\}$ et $T_n^L = \inf\{m \geq n / X_m^2 \geq L\}$. Nous notons également $H = \{\liminf |X_n| > 0\}$.

Nous notons enfin **A'-1** l'hypothèse : $E(|\epsilon_{n+1}|^a|\mathcal{F}_n) \leq B^{\frac{a}{2}}$, $\sum_{n\geq 0} r_{n+1}^2 \leq \eta$ et $V = \mathbb{R}$.

Alors le lemme 1 de [8] s'adapte de la manière suivante :

Lemme 3.2.1 *Supposons $\sum c_n^2 < +\infty$,* **A'-1** *avec $\eta \leq A^2/16L$, et il existe $c \in \mathbb{R}_*^+$ tel que*

(i) $E(X_{i+1}^2 - Y_i^2|\mathcal{F}_i) \geq Ac_{i+1}^2 - 2c_{i+1}E(|r_{i+1}||\mathcal{F}_i)|Y_i|$

(ii) $(X_{i+1} - Y_i)^2 \leq c_{i+1}^2 c(\epsilon_{i+1}^2 + 1)$

Alors il existe une constante $C \in]0,1[$ telle que pour tout $n \in \mathbb{N}$,

$$P(S_n'^L < \infty|\mathcal{F}_n) \geq C.$$

PREUVE: Cette preuve repose sur le même principe que celle du lemme 1 de [8]. Nous remplaçons pour simplifier $S_n'^L$ par S'.

Sur $\{n \leq k < S'\}$, $|Y_k| \leq \sqrt{L\alpha_{k+1}} \leq \sqrt{L\alpha_{n+1}}$.

Par conséquent, en posant $Z_k = X_k^2 - A\sum_{i=n}^k c_i^2 + 2\sqrt{L\alpha_{n+1}}\sum_{i=n}^k c_i|r_i|$, en utilisant (i) et le fait que $Y_k^2 \geq X_k^2$ pour tout $k \in \mathbb{N}$, nous obtenons que $\{Z_{k\wedge S'}\}_{k\geq n}$

est une sous-martingale. Donc, si $m \geq n$, $E(Z_{m\wedge S'} - Z_n|\mathcal{F}_n) \geq 0$. Ainsi

$$\begin{aligned} E(X^2_{m\wedge S'} - X^2_n|\mathcal{F}_n) \geq \quad & (A\sum_{i=n+1}^{m} c_i^2 - 2\sqrt{L\alpha_{n+1}}\frac{\sqrt{\alpha_{n+1}}A}{4\sqrt{L}})P(S' > m|\mathcal{F}_n) \\ & -2(1 - P(S' > m|\mathcal{F}_n))\sqrt{L\alpha_{n+1}}\frac{\sqrt{\alpha_{n+1}}A}{4\sqrt{L}} \end{aligned}$$

puisque, d'après l'inégalité de Cauchy-Schwarz, $\sum_{i=n+1}^{m} c_i|r_i| \leq \sqrt{\alpha_{n+1}s_{n+1}} \leq \frac{\sqrt{\alpha_{n+1}}A}{4\sqrt{L}}$. Donc

$$E(X^2_{m\wedge S'} - X^2_n|\mathcal{F}_n) \geq A(\frac{\alpha_{n+1}}{2} - \alpha_{m+1})P(S' > m|\mathcal{F}_n) - (1 - P(S' > m|\mathcal{F}_n))\frac{A}{2}\alpha_{n+1}.$$

D'autre part, $X_m^2 \leq Y_m^2 \leq L\alpha_{m+1}$ si $S' > m$. Par conséquent

$$X^2_{m\wedge S'} - X^2_n \leq L\alpha_{m+1}1_{S'>m} + 2[Y^2_{S'-1}1_{n<S'\leq m} + (X_{S'} - Y_{S'-1})^2 1_{n<S'\leq m}],$$

et

$$\begin{aligned} & E(X^2_{m\wedge S'} - X^2_n|\mathcal{F}_n) \\ & \leq 2L(P(S' > m|\mathcal{F}_n)\alpha_{m+1} + (1 - P(S' > m|\mathcal{F}_n))\alpha_{n+1}) + 2E((X_{S'} - Y_{S'-1})^2 1_{n<S'\leq m}|\mathcal{F}_n) \\ & \leq 2(L + c)(P(S' > m|\mathcal{F}_n)\alpha_{m+1} + (1 - P(S' > m|\mathcal{F}_n))\alpha_{n+1}) + 2c\sum_{i=n+1}^{m} c_i^2 E(\epsilon_i^2 1_{S'=i}|\mathcal{F}_n). \end{aligned}$$

La fin de la preuve est identique à celle du lemme 1 de [8].

□

Nous adaptons le lemme 2 de [8] de la façon suivante :

Lemme 3.2.2 *Supposons* $\sum c_i^2 < +\infty$, **A'-1** *avec* $\eta \leq 1$, $L \geq (\sqrt{8B} + 2)^2$ *et, pour tout* $i \in \mathbb{N}$

- $X_{i+1} = Y_i + c_{i+1}(\epsilon_{i+1} + r_{i+1})$
- $|Y_i| \geq |X_i|$
- $E(\epsilon_{i+1}|\mathcal{F}_i) = 0$

Alors, pour tous $n \in \mathbb{N}$ *et* $m \geq n$, $P(H|\mathcal{F}_m)1_{S'^L_n = m} \geq \frac{1}{2}1_{S'^L_n = m}$.

La preuve de ce lemme est identique à celle du lemme 2 de [8].

Démonstration du théorème 1.- Nous choisissons $L = (\sqrt{8B} + 2)^2$ et $\eta = \min(\frac{A^2}{4B}, \frac{A^2}{16L}, 1)$. Il nous reste à prouver que les variables aléatoires $(X_n)_{n\in\mathbb{N}}$ et $(Y_n)_{n\in\mathbb{N}}$ vérifient bien les hypothèses du lemme 3.2.1. Remarquons que

$$\begin{aligned} E(X^2_{n+1}|\mathcal{F}_n) = \quad & Y_n^2 + c^2_{n+1}E(\epsilon^2_{n+1}|\mathcal{F}_n) + 2c_{n+1}E(r_{n+1}|\mathcal{F}_n)Y_n \\ & +2c^2_{n+1}E(\epsilon_{n+1}r_{n+1}|\mathcal{F}_n) + c^2_{n+1}E(r^2_{n+1}|\mathcal{F}_n). \end{aligned}$$

En appliquant l'inégalité de Cauchy-Schwarz, $E(\epsilon_{n+1} r_{n+1}|\mathcal{F}_n) \geq -\sqrt{BE(r_{n+1}^2|\mathcal{F}_n))}$. Ainsi

$$E(X_{n+1}^2 - Y_n^2|\mathcal{F}_n) \geq c_{n+1}^2(2A - 2\sqrt{B\eta}) - 2c_{n+1}E(|r_{n+1}||\mathcal{F}_n)|X_n|,$$

donc l'hypothèse (i) du lemme 3.2.1 est vérifiée à cause du choix de η. La condition (ii) de ce lemme est également vérifiée car pour tout $i \in \mathbb{N}$,

$$(X_{i+1} - Y_i)^2 \leq 2c_{i+1}^2(\epsilon_{i+1}^2 + r_{i+1}^2).$$

Ainsi les hypothèses du lemme 3.2.1 sont bien réunies et $P(S_n'^L < +\infty|\mathcal{F}_n) \geq C$. D'autre part le lemme 3.2.2 nous permet d'affirmer que $P(H|\mathcal{F}_n)1_{S_n'^L=m} \geq \frac{1}{2}1_{S_n'^L=m}$. Donc

$$E(1_H|\mathcal{F}_n) \geq \sum_{i\geq n} E(1_H 1_{S_n^L=i}|\mathcal{F}_n) \geq \sum_{i\geq n} E(E(1_H|\mathcal{F}_i)1_{S_n^L=i}|\mathcal{F}_n) \geq \frac{1}{2}\sum_{i\geq n} E(1_{S_n^L=i}|\mathcal{F}_n) \geq \frac{C}{2}.$$

Puisque H est $\mathcal{F}_\infty$-mesurable, $\lim_{n\to\infty} E(1_H|\mathcal{F}_n) = 1_H$ presque partout, on en déduit $P(H) = 1$. $\square$

Bibliographie

[1] M. Benaïm. Dynamics of stochastic approximation algorithms. *A paraitre dans Le Séminaire de Probabilités, (Azema and Yor, Ed) Springer-Verlag.*, 2000.

[2] O. Brandière. The dynamic system method and the traps. *Adv. in Appl. Probab.*, t. 30, no.1 :137–151, 1998.

[3] O. Brandière. Some pathological traps for stochastic approximation. *SIAM J. Control Optim.*, t. 36, no.4 :1293–1314, 1998.

[4] M. Duflo. *Algorithmes stochastiques.* Mathématiques et Applications, Springer Verlag, 1996.

[5] O. Brandière et M. Duflo. Les algorithmes stochastiques contournent-ils les pièges ? *Ann. Inst. Henri. Poincaré Probab. Statist.*, t. 32, no 3 :395–427, 1996.

[6] R. Pemantle. Nonconvergence to unstable points in urns models and stochastic approximation. *Annals of Probability*, 18 :698–712, 1990.

[7] R. Pemantle. Vertex-reinforced random walk. *Probability Theory and Related Fields*, 92 :117–136, 1992.

[8] P. Tarrès. Pièges répulsifs. *C. R. Acad. Sci.*, t. 330, Série I :125–130, 2000.

Chapitre 4

Bandit à deux bras

4.1 Introduction

Nous nous plaçons dans un espace de probabilité $(\Omega, \mathcal{F}, P)$ muni d'une filtration $\mathbb{F} = (\mathcal{F}_n)_{n \in \mathbb{N}}$.

Nous disposons d'un bandit à deux bras A et B : en appuyant sur A (resp. B), on gagne un franc avec probabilité θ_A (resp. θ_B), et rien sinon.

L'algorithme de Narendra (voir par exemple [2]) a pour but de choisir le bras pour lequel la probabilité de gagner est la plus grande, et est défini comme suit : à chaque instant $n \in \mathbb{N}^*$, je joue le bras A avec probabilité X_n, et le bras B avec probabilité $1 - X_n$. Initialement $X_1 = x_1 \in]0,1[$; à l'instant n, je détermine X_{n+1} de la manière suivante :

$$X_{n+1} = \begin{cases} (1-\gamma_n)X_n + \gamma_n \text{ si j'ai joué } A \text{ et } A \text{ est gagnant} \\ (1-\gamma_n)X_n \text{ si j'ai joué } B \text{ et } B \text{ est gagnant} \\ X_n \text{ sinon,} \end{cases}$$

étant donnée une suite déterministe $(\gamma_n)_{n \in \mathbb{N}^*}$. On obtient en résumé

$$X_{n+1} - X_n = \begin{cases} \gamma_n(1-X_n) \text{ avec probabilité } X_n.\theta_A \\ -\gamma_n X_n \text{ avec probabilité } (1-X_n).\theta_B \\ 0 \text{ avec probabilité } 1 - (X_n\theta_A + (1-X_n)\theta_B), \end{cases}$$

ces probabilités étant entendues conditionnellement à $\mathcal{F}_n$.

Pour l'étude de cet algorithme nous définissons, pour tout $n \in \mathbb{N}^*$, la variable aléatoire $\mathcal{F}_{n+1}$-mesurable U_{n+1} à valeurs dans $\{A, B, N\}$ comme suit : dans les équations précédentes, $U_{n+1} = A$ sous la première hypothèse, $U_{n+1} = B$ sous la deuxième hypothèse, enfin $U_{n+1} = N$ sous la dernière hypothèse.

Nous supposons dorénavant $\theta_A > \theta_B$. Le théorème suivant affirme que, si $\gamma_n = c/n^\alpha$ (avec $c \in]0,1]$, $\alpha \in]0,1]$), alors la probabilité que l'algorithme converge vers le choix du mauvais bras B, c'est à dire que X_n converge vers 0, est strictement positive lorsque $\alpha < 1$.

Théorème 4.1.1 *Si $x_1 < 1$ et $\alpha < 1$, alors $P(\lim_{n\to\infty} X_n = 0) > 0$.*

Par contre cette probabilité que l'algorithme converge vers le choix du mauvais bras B, c'est à dire que X_n converge vers 0, est nulle lorsque $\alpha = 1$.

Théorème 4.1.2 *Si $\alpha = 1$ et $x_1 > 0$, $P(\lim_{n\to\infty} X_n = 0) = 0$.*

4.2 Preuve des résultats

Nous notons $u_n = v_{1,n} + \ldots + v_{k,n} + o(\triangleleft)$ lorsque $u_n = v_{1,n} + \ldots + v_{k,n} + o(v_{k,n})$ quand n tend vers $+\infty$. Nous notons par Cst une constante universelle strictement positive quelconque et, pour $a_1, \ldots, a_k \in \mathbb{R}$, par $Cst(a_1, \ldots, a_k)$ une constante strictement positive ne dépendant que de $a_1, \ldots, a_k$.

Nous utilisons pour la preuve du théorème 4.1.1 le résultat suivant, qui donne des inégalités exponentielles pour les lois de Bernoulli.

Lemme 4.2.1 *Soit Z_i une suite de variables de Bernoulli 0 et 1 i.i.d avec probabilité de succès p. Posons*

$$Z = \frac{Z_1 + Z_2 + \ldots + Z_n}{n},$$

alors pour tout $a \in]0, p[$,

$$P(Z \leq a) \leq \exp\{-nH(a,p)\}$$

avec

$$H(a,p) = a \ln \frac{a}{p} + (1-a) \ln \frac{1-a}{1-p} > 0.$$

Preuve du théorème 4.1.1.- Posons, pour tout $n \in \mathbb{N}$, $t_n = 2^n$. Soient $k_0 \in \mathbb{N}$ que nous fixerons dans la suite de la démonstration, et $\epsilon > 0$ quelconque.

La démonstration repose sur le principe suivant : si à un instant t_n, X_{t_n} est petit, alors avec une grande probabilité on ne gagne jamais avec le bras A sur l'intervalle $]t_n, t_{n+1}]$, ce qui implique, puisque d'autre part avec une grande probabilité on gagne un assez grand nombre de fois avec le bras B sur ce même intervalle, qu'il est très probable que $X_{t_{n+1}}$ soit inférieur à X_{t_n} d'un facteur d'au moins $2^{-(1+\epsilon)}$. En appliquant ce raisonnement pour tout $n \in \mathbb{N}$, on montre qu'avec une probabilité non nulle, on ne gagne jamais avec le bras A à partir d'un instant donné.

Plus précisément, définissons le temps d'arrêt

$$T = \inf\{n > t_{k_0} /\ U_n = A\}$$

et, pour tout $k \geq k_0$, les événements probabilistes suivants

$$\begin{aligned} E_1(k) &: X_{t_k} \leq x_1 2^{-(k-k_0)(1+\epsilon)} \\ E_2(k) &: T > t_k, \end{aligned}$$

enfin $E(k) = E_1(k) \cap E_2(k)$.

Remarquons d'une part, en posant

$$p_k = (1 - 2^{-(k-k_0)(1+\epsilon)} x_1 \theta_A)^{2^k} = 1 - \theta_A x_1 2^{k_0(1+\epsilon)} 2^{-k\epsilon} + o(\triangleleft),$$

que

$$P(E_2(k+1) \mid \mathcal{F}_{t_k}) \geq 1_{E(k)} p_k.$$

D'autre part, notons

$$q_k = P(E(k+1) \mid \mathcal{F}_{t_k})$$

et montrons que

$$q_k \geq 1_{E(k)}(p_k + o(2^{-k\epsilon})).$$

En effet, notons N le nombre d'entiers $n \in]t_k, t_{k+1}]$ tels que $U_n = B$. Observons que sur $E_2(k+1)$, $X_{t_{k+1}} \leq x_1$ et, pour tout entier $n \in [t_k, t_{k+1}[$,

$$P(U_{n+1} = B \mid \mathcal{F}_n) \geq 1_{\{x_n \leq x_1\}}\theta_B(1-x_1).$$

Ce qui implique, en appliquant le lemme 4.2.1, que

$$P(\{N \geq 2^{k-1}\theta_B(1-x_1)\} \cap E_2(k+1) \mid \mathcal{F}_{t_k}) \geq p_k - \exp(-2^k Cst(\theta_B, x_1)).$$

Or $X_{n+1} = X_n(1-\gamma_n)$ lorsque $U_{n+1} = B$. Par conséquent, sur

$$E_2(k+1) \cap \{N \geq 2^{k-1}\theta_B(1-x_1)\},$$

$$\begin{aligned} X_{t_{k+1}} &\leq X_{t_k}(1 - c2^{-(k+1)\alpha})^{2^{k-1}\theta_B(1-x_1)} \leq X_{t_k}\exp(-2^{-(k+1)\alpha}2^{k-1}c\theta_B(1-x_1)) \\ &\leq 2^{-(k-k_0)(1+\epsilon)}\exp(-2^{k(1-\alpha)}c2^{-(\alpha+1)}\theta_B(1-x_1)) \leq 2^{-(k+1-k_0)(1+\epsilon)} \end{aligned}$$

si k_0 est supérieur à une constante ne dépendant que de α, ϵ, c, θ_B et x_1, de sorte que $\exp(-2^{k(1-\alpha)}c2^{-(\alpha+1)}\theta_B(1-x_1)) \leq 2^{-(1+\epsilon)}$.

En résumé

$$P[\cap_{k \geq k_0} E(k) \mid \mathcal{F}_{t_{k_0}}] \geq 1_{X_{t_{k_0}} \leq x_1} \Pi_{k=k_0}^{+\infty} q_k > 0,$$

car q_k est strictement positif pour tout $k \geq k_0$ et

$$q_k = 1 - \theta_A x_1 2^{k_0(1+\epsilon)} 2^{-k\epsilon} + o(\triangleleft).$$

Or

$$\cap_{k > k_0} E(k) \subset \{\lim_{n \to \infty} X_n = 0\}.$$

Ce qui permet de conclure.

□

La preuve du théorème 4.1.2 utilise le lemme suivant.

Lemme 4.2.2 *Soit* $F : [0,1] \to \mathbb{R}_+$ *une fonction borélienne positive sur* $[0,1]$. *Soit* $(\gamma_n)_{n \in \mathbb{N}^*}$ *une suite définie par* $\gamma_n = c/n$, *où* $c \in \mathbb{R}_+^*$.

Notons

$$B = \{\forall b \in \mathbb{R}_+^*,\ \forall n \in \mathbb{N}^*,\ \exists p \geq n /\ X_p \geq b/p\}.$$

Soit $(X_n)_{n \in \mathbb{N}^*}$ *un algorithme à valeurs dans* $[0,1]$ *; notons* $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$. *Supposons que*

$$X_{n+1} - X_n = \gamma_n(F(X_n) + \epsilon_{n+1})$$

et qu'il existe une constante $a \in \mathbb{R}_+^*$ telle que, pour tout $n \in \mathbb{N}^*$,

$$E(\epsilon_{n+1} \mid \mathcal{F}_n) = 0, \quad E(\epsilon_{n+1}^2 \mid \mathcal{F}_n) \leq aX_n.$$

Alors

$$P(\{\liminf_{n\to\infty} X_n = 0\} \cap B) = 0.$$

PREUVE: Nous allons montrer que, pour tout $p \in \mathbb{N}$ tel que $p \geq 2$, si $X_p \geq b/p$, alors

$$P(\liminf_{n\to\infty} X_n > 0 \mid \mathcal{F}_p) \geq 1/2$$

si b a été supposé plus grand qu'une constante ne dépendant que de a et de c. Ce qui permettra de conclure.

Nous fixons donc $p \in \mathbb{N}$ tel que $p \geq 2$ et $X_p \geq b/p$.

Le principe de l'argument utilisé est le suivant : toutes les fois que X_n est supérieur à X_p pour $n \geq p$, la probabilité que X_{n+1} soit inférieur à $X_p/2$ est très faible. D'autre part l'hypothèse sur les moments conditionnels d'ordre 2 du bruit implique que la somme des termes de bruit $\gamma_n \epsilon_{n+1}$ pour $X_n \leq X_p$ est, avec une grande probabilité, inférieure à un terme de l'ordre de $\sqrt{X_p/p}$ et est donc inférieure à $X_p/4$ si b a été supposé assez grand. En résumé, avec une grande probabilité, on reste alors supérieur à $X_p/4$.

Plus précisément, posons, pour tout $n \in \mathbb{N}^*$, $\tilde{\epsilon}_{n+1} = \epsilon_{n+1} 1_{X_n \leq X_p}$.

Appliquons l'inégalité de Doob :

$$E(\sup_{k>p} (\sum_{i=p}^{k-1} \gamma_i \tilde{\epsilon}_{i+1})^2 \mid \mathcal{F}_p) \leq 4 \sum_{i=p}^{+\infty} \gamma_i^2 E(\tilde{\epsilon}_{i+1}^2 \mid \mathcal{F}_p) \leq \frac{4ac^2 X_p}{p-1}.$$

Posons

$$\Delta_1 = \{\forall j, k / \ k > j \geq p, \ \sum_{i=j}^{k-1} \gamma_i \tilde{\epsilon}_{i+1} \geq -X_p/4\},$$

alors

$$P(\Delta_1^c \mid \mathcal{F}_p) \leq \frac{256ac^2}{(p-1)X_p} \leq \frac{512ac^2}{b} \leq \frac{1}{4}$$

en supposant b assez grand.

D'autre part, pour tout $k \geq p$, si $X_k > X_p \geq b/p$, alors

$$P(X_{k+1} < X_k/2 \mid \mathcal{F}_k) \leq P(|\epsilon_{k+1}| \geq X_k/2\gamma_k | \mathcal{F}_k) \leq \frac{4a\gamma_k^2 X_k}{X_k^2} \leq \frac{4ac^2}{k^2 X_k} \leq \frac{4ac^2}{k^2 X_p}.$$

Donc, en posant

$$\Delta_2 = \{\forall k \geq p / \ X_k > X_p \Longrightarrow X_{k+1} \geq X_k/2\},$$

nous déduisons

$$P(\Delta_2^c \mid \mathcal{F}_p) \leq \frac{4ac^2}{(p-1)X_p} \leq \frac{8ac^2}{b} \leq \frac{1}{4}$$

en supposant b assez grand.

En résumé, notons $\Delta = \Delta_1 \cap \Delta_2$. Alors $P(\Delta \mid \mathcal{F}_p) \geq 1/2$ et, sur Δ, pour tout $k \geq p$, $X_k \geq X_p/4$.

□

PREUVE DU THÉORÈME 4.1.2.- Nous appliquons le lemme 4.2.2. Il nous reste à montrer que $P(B) = 1$.

Remarquons que

$$X_n \geq x_1 \Pi_{k=1}^{n-1}(1-c/k)\Pi_{k\in[1,n-1]/\ U_{k+1}=N}\frac{1}{1-c/k} \geq \frac{x_1 Cst(c)}{k^c}\Pi_{k\in[1,n-1]/\ U_{k+1}=N}\frac{1}{1-c/k}.$$

Or, pour tout $k \in [1, n-1]$, $P(U_{k+1} = N \mid \mathcal{F}_k) \geq \min(1-\theta_A, 1-\theta_B)/2$. Par conséquent, en appliquant le lemme de Borel-Cantelli conditionnel (voir par exemple [3]),

$$\begin{aligned}\{\Pi_{k\in\mathbb{N}^*/\ X_{k+1}=N}\frac{1}{1-c/k} = +\infty\} \quad &= \{\sum_{k\in\mathbb{N}^*/\ X_{k+1}=N}\frac{c}{k} = +\infty\} \\ &= \{\sum_{k\in\mathbb{N}^*}\frac{c}{k}P(U_{k+1} = N|\mathcal{F}_k) = +\infty\} = \Omega,\end{aligned}$$

ce qui permet de conclure, car $c \leq 1$.

□

Bibliographie

[1] M. Benaïm. Dynamics of stochastic approximation algorithms. *Séminaire de probabilités XXXIII, Lect. Notes in Math.*, 1709 :1–68, 1999.

[2] M. Duflo. *Algorithmes stochastiques.* Mathématiques et Applications, Springer Verlag, 1996.

[3] D. Dacunha-Castelle et M. Duflo. *Probabilités et statistiques*, volume 2. Masson, 1993.

Chapitre 5

Article en collaboration avec M. Benaïm et S. Schreiber : Generalized urn models of evolutionary processes

Abstract Generalized Polya urn models can describe the dynamics of finite populations of interacting genotypes. Three basic questions these models can address are : Under what conditions does a population exhibit growth ? On the event of growth, at what rate does the population increase ? What is the long term behavior of the distribution of genotypes ? To address these questions, we associate a mean limit ordinary differential equation (ODE) with the urn model. Previously, it has been shown that on the event of population growth, the limiting distribution of genotypes is a connected internally chain recurrent set for the mean limit ODE. To determine when growth and convergence occurs with positive probability, we prove two results. First, if the mean limit ODE has an "attainable" attractor at which growth is expected, then growth and convergence toward this attractor occurs with positive probability. Second, the population distribution almost surely does not converge to sets where growth is not expected, and almost surely does not converge to "non-degenerate" unstable equilibria or periodic orbits of the mean limit ODE. Applications to stochastic analogs of the replicator equations and fertility-selection equations of population genetics are given.

5.1 Introduction

The founder-effect in population genetics refers to the establishment of a new population consisting of a few founders that is isolated from the original population. A founder-effect can occur when a small number of individuals colonize a place previously uninhabited by their species. In this case, the founding population is geographically isolated from the original population. A founder-effect due to temporal isolation can occur when a population passes through a bottle neck after which only a few individuals survive. Several fundamental questions surrounding the founder-effect include : What is the probability that a founding population successfully establishes itself ? If a founding population establishes itself, what is the population's growth rate and what is the long-term genotypic or phenotypic composition the population ? How does the initial genotypic composition and initial population size influence the likelihoods of the various outcomes ?

To address these questions, we consider a general class of Polya urn models. Traditionally, Polya urn models are described as involving an urn which contains a finite number of balls of different colors. At discrete moments in time, balls are added or removed from the urn according to probabilities that only depend on their distribution and number at that point in time. The pertinence of these models to evolutionary questions is self-evident if we view the balls as individuals whose color represents their genotype or phenotype, adding or removing balls as replication or death of individuals, and updates of the urn as interactions between individuals. When balls are added at a constant rate (i.e. a fixed number of individuals are added at every update of the process), Polya urn models have been studied extensively by [1, 5, 6, 8, 10, 11] amongst others. What makes the models we discuss here more relevant to population processes is that they permit the removal of balls, as well as the addition of balls at non-constant rates. Consequently, extinction of the entire population or one or more subpopulations is possible in finite time.

The article is structured as follows. In section 2, we introduce the generalized urn models of interest. As examples, replicator processes and fertility-selection processes with and without mutations are introduced. These urn models typically predict that either the populations go extinct or grow, and that demographic stochasticity is most pronounced when the population is small. Once the population starts to get large, it tends to grow in an essentially deterministic fashion. For this reason, on the event of non-extinction, the dynamics of the distribution of types in the population (i.e. the distribution of the color of balls in the urn) are strongly correlated to the asymptotic behavior of an appropriately chosen ordinary differential equation (ODE), commonly called the *mean limit ODE.* In section 3, we describe the mean limit ODE, and recall a theorem [12] that on the event of growth relates the asymptotic behaviors of the stochastic process to its mean limit ODE. Using this result, we derive a time averaging principle and a competitive exclusion principle for replicator processes. We also show that additive fertility-selection processes almost surely converge on the event of non-extinction to a fixed point of the mean limit ODE. These results, however, provide no insight into when population growth occurs with positive probability and which limiting behaviors occur with positive probability. These issues are addressed in sections 4 and 5. In section 4, we prove that the existence of an "attainable" attractor (for the mean limit ODE) at which growth is expected is a sufficient condition for growth and convergence to the attractor with positive probability. In addition, we provide an estimate for the rate of growth. In section 5, we prove non-convergence to invariant subsets of the mean limit ODE where growth is not expected, and non-convergence to "non-degenerate" unstable equilibria and periodic orbits of the mean limit ODE. In section 6, we combine our results to give necessary and sufficient conditions for growth with positive probability for process with gradient-like mean limit ODEs, and apply these conditions to additive fertility processes with mutation.

5.2 Generalized Urn Models

In this section, we introduce a class of generalized urn models. Several examples of evolutionary processes that fall into this class of generalized urn models are given in the subsections below. Due to the fact that we are dealing finite populations consisting of individuals that are one of k types, we consider Markov chains on the positive cone

$$\mathbb{Z}^k_+ = \{z = (z^1, \dots, z^k) \in \mathbb{Z}^k : z^i \geq 0 \text{ for all } i\}$$

of the set $\mathbb{Z}^k$ of k-tuples of integers. Given a vector $w = (w^1, \dots, w^k) \in \mathbb{Z}^k$ define

$$|w| = |w^1| + \cdots + |w^k| \text{ and } \alpha(w) = w^1 + \cdots + w^k.$$

We shall always write $\|\cdot\|$ for the Euclidean norm on $\mathbb{R}^k$.

Let $z_n = (z^1_n, \dots, z^k_n)$ be a homogeneous Markov chain with state space $\mathbb{Z}^k_+$. In our context, z^i_n corresponds to the number of balls of color i at the n-th update. Associated with z_n is the random process x_n defined by

$$x_n = \begin{cases} \frac{z_n}{|z_n|} & \text{if } z_n \neq 0 \\ 0 & \text{if } z_n = 0 \end{cases}$$

which is the distribution of balls at the n-th update. Note that when there are no balls at the n-th update, we set x_n to zero which we view as the "null" distribution. Let $S_k \subset \mathbb{R}^k$ denote the unit $k-1$ simplex, i.e.

$$S_k = \{x = (x^1, \dots, x^k) \in \mathbb{R}^k : x^i \geq 0, \sum_{i=1}^k x^i = 1\}.$$

Let $\Pi : \mathbb{Z}_+^k \times \mathbb{Z}_+^k \mapsto [0,1]$ denote the transition kernel of the Markov chain z_n. In other words, $\Pi(z, z') = P[z_{n+1} = z' | z_n = z]$. We place the following assumptions on the Markov chains z_n of interest :

(A1) At each update, there is a maximal number of balls that can be added or removed. In other words, there exists a positive integer m such that $|z_{n+1} - z_n| \leq m$ for all n.

(A2) There exist Lipschitz maps

$$\{p_w : S_k \to [0,1] \ : w \in \mathbb{Z}^k, \ |w| \leq m\}$$

and a real number $a > 0$ such that

$$|p_w(z/|z|) - \Pi(z, z+w)| \leq a/|z|$$

for all non-zero $z \in \mathbb{Z}_+^k$ and $w \in \mathbb{Z}^k$ with $|w| \leq m$.

Since we view updates of the Markov chain to correspond to the effect of interactions between individuals, assumption (A1) implies that each interaction results in the addition or removal of no more than a maximum number of individuals. Assumption (A2) assures that there is a well defined mean limit ODE for the urn models.

5.2.1 Replicator Processes

Consider a system consisting of a finite population of individuals playing k different strategies. At each update of the population, pairs of individuals are chosen randomly with replacement from the population. The chosen individuals replicate and die according to probabilities that only depend on their strategies. More precisely, let m be a non-negative integer that represents the maximum number of progeny that any individual can produce in one update. Let $\{R_n\}_{n\geq 0}$ and $\{\tilde{R}_n\}_{n\geq 0}$ be sequences of independent identically distributed random $k \times k$ matrices whose entries take values in the set $\{-1, 0, 1, \dots, m\}$. Let R_n^{ij} and $\tilde{R}_n^{ij}$ denote the ij-th entry of R_n and $\tilde{R}_n$, respectively. Let $\{r_n\}_{n\geq 0}$ be a sequence of independent identically distributed random $k \times 1$ matrices whose entries take values in the set $\{-1, 0, 1, \dots, m\}$. We define a replicator process according to the following rules :

1. Two individuals are chosen at random with replacement from the the population. Make note of the individuals chosen and return them to the population.
2. If the same individual is chosen twice and it plays strategy i, then r_n^i individuals of strategy i are added to population.
3. If two distinct individuals are chosen, say strategy i and strategy j, then add R_n^{ij} individuals of strategy i and add $\tilde{R}_n^{ji}$ individuals of strategy j.

To show that this process satisfies assumptions (A1) and (A2), let $z \in \mathbb{Z}_+^k\backslash\{0\}$ and $x = z/|z|$. We break up the transition probabilities into four cases. First, if $w = (0,\dots,0,w^i,0,\dots,0)$ with $w^i \neq 0$, then

$$\Pi(z,z+w) = x^i\frac{x^i|z|-1}{|z|}P[R_1^{ii}+\tilde{R}_1^{ii} = w^i]+\sum_{j\neq i}2x^ix^jP[R_1^{ij} = w^i]P[\tilde{R}_1^{ji} = 0]+\frac{x^i}{|z|}P[r_1^i = w^i]$$

and

$$p_w(x) = x^ix^iP[R_1^{ii} + \tilde{R}_1^{ii} = w^i] + \sum_{j\neq i}2x^ix^jP[R_1^{ij} = w^i]P[\tilde{R}_1^{ji} = 0].$$

Second, if $w = (0,\dots,0,w^i,0,\dots,0,w^j,0,\dots,0)$ with $w^i \neq 0$, $w^j \neq 0$, and $i \neq j$, then

$$\Pi(z,z+w) = p_w(x) = 2x^ix^jP[R_1^{ij} = w^i]P[\tilde{R}_1^{ji} = w^j].$$

Thirdly, if w has three or more non-zero coordinates, then $\Pi(z,z+w) = p_w(x) = 0$. Finally, $\Pi(z,z)$ and $p_0(x)$ are given by the complementary probabilities, $1-\sum_{w\neq 0}\Pi(z,z+w)$ and $1-\sum_{w\neq 0}p_w(x)$, respectively. From these expressions, it follows that (A2) holds.

5.2.2 Fertility Selection Processes

Consider a population of diploid individuals that has k distinct alleles $A_1,\dots,A_k$ that occupy a single locus. Assume that the population is monoecious (i.e. there is only one "sex") and that each individual chooses its mate randomly from the population. We assume that individuals die immediately after mating. Although there is no distinction between individuals of genotype A_iA_j and A_jA_i, we develop an urn model of the form $z_n = (z_n^{ij}) \in \mathbb{Z}_+^{k\times k}$ as it is notationally more convenient. For $i \neq j$, let $z_n^{ij} = z_n^{ji}$ denote the number of individuals of genotype A_iA_j at the n-th update of the population. Alternatively, let z_n^{ii} denote twice the number of individuals of genotype A_iA_i at the n-th update of the population. Hence $|z_n| = \sum_{i,j} z_n^{ij}$ equals twice the total number of individuals in the population.

For every pair of genotypes, say A_iA_j and A_rA_s, we associate a sequence of i.i.d random variables $G_n(ij,rs) = G_n(ij,rs)$ that take values in $\{0,\dots,m\}$ where m represents the maximal number progeny produced by a mating and where $G_n(ij,rs)$ represents the number of progeny produced by a mating between genotypes A_iA_j and A_rA_s at update n. Let $z_n \in \mathbb{Z}_+^{k\times k}$ be a Markov chain satisfying $z_n^{ij} = z_n^{ji}$ and updated according to the following rules :

1. If the population size is less than two, then the population goes extinct. In other words, if $|z_n| < 4$ then $z_{n+1} = 0$.
2. Pick two individuals at random without replacement from the population, say genotypes A_rA_s and A_uA_v.
3. Remove the chosen individuals from the population (i.e. they die).
4. Add $G_n(ij,rs)$ individuals to the population. The genotype, say A_uA_v, of each added individual is independently determined by random mating probabilities (i.e. u equals i or j with equal probability and v equals r or s with equal probability).

Define

$$x_n^{ij} = \begin{cases} \frac{z_n^{ij}}{|z_n|} & \text{if } z_n \neq 0 \\ \\ 0 & \text{if } z_n = 0. \end{cases}$$

Hence, if $i \neq j$, then $2x^{ij} = 2x^{ji}$ equals the proportion of the population with genotype A_iA_j. Alternatively, x^{ii} is the proportion of the population with genotype A_iA_i.

To see that this process satisfies assumption (A2), let w and z be in $\mathbb{Z}_+^{k\times k}$ such that $w^{ij} = w^{ji}$ and $z^{ij} = z^{ji}$ for all $1 \leq i, j \leq k$. Assume that $|z| > 0$. $p_w(z/|z|)$ is given by a linear combination of the terms $x^{ij}x^{rs}$. On the other hand, $\Pi(z, z + w)$ is given by the corresponding linear combination of the terms $x^{ij}x^{rs}$ $x^{ij}x^{rs}|z|/(|z| - 2)$ when $\{i, j\} \neq \{r, s\}$ and $x^{ij}(x^{rs}|z| - 2)/(|z| - 2)$. From these observations it follows that (A2) is satisfied.

5.2.3 Fertility Selection Process with Mutations

To account for mutations in the fertility selection process, let $\mu(rs, ij) \geq 0$ for $1 \leq i, j, r, s \leq k$ be such that $\sum_{i\leq j} \mu(rs, ij) = 1$ for all $1 \leq r \leq s \leq k$. The quantity $\mu(ij, rs)$ represents the probability that the genotype A_iA_j mutates to the genotype A_rA_s. The fertility process with mutation is given by the first three rules of the fertility process without mutation and replacing the fourth rule with : Add $G_n(ij, rs)$ individuals to the population. For each added individual, its genotype is determined by two steps. First, determine a genotype A_uA_v according to random mating probabilities. The probability that the added individual has genotype $A_{\tilde{u}}A_{\tilde{v}}$ is given by $\mu(uv, \tilde{u}\tilde{v})$. For reasons similar to the fertility selection process without mutation, this process also satisfies assumptions (A1) and (A2).

5.3 Mean Limit ODEs

To understand the limiting behavior of the x_n, we express x_n as a stochastic algorithm.

Lemma 5.3.1 *Let z_n be a Markov chain on $\mathbb{Z}_+^k$ satisfying assumptions (A1) and (A2) with mean limit transition probabilities $p_w : S_k \to [0, 1]$. Let $\mathcal{F}_n$ denote the σ field generated by $\{z_0, z_1, \ldots, z_n\}$. There exists sequences of random variables $\{U_n\}$ and $\{b_n\}$ adapted to $\mathcal{F}_n$, and a real number $K > 0$ such that*

(i) if $z_n \neq 0$, then

$$x_{n+1} - x_n = \frac{1}{|z_n|}\left(\sum_{w\in\mathbb{Z}^k} p_w(x_n)(w - x_n\alpha(w)) + U_{n+1} + b_{n+1}\right). \tag{5.1}$$

(ii) $E[U_{n+1}|z_n] = 0$.

(iii) $\|U_n\| \leq 4m$ and $E[\|U_{n+1}\|^2|\mathcal{F}_n] \leq 4m^2$.

(iv) $\|b_{n+1}\| \leq \frac{K}{\max\{1,|z_n|\}}$.

PROOF: Define

$$\begin{aligned} g(x) &= \sum_{w \in \mathbb{Z}^k} p_w(x)(w - x\alpha(w)) \\ U_{n+1} &= (x_{n+1} - x_n - E[x_{n+1} - x_n | z_n]) |z_n| \\ b_{n+1} &= |z_n| E[x_{n+1} - x_n | z_n] - g(x_n). \end{aligned}$$

From these definitions, it follows that (i) and (ii) hold.

For the remainder of the proof, we write $z = z_n$ and $x = x_n$. To prove (iii), notice that if $z_{n+1} \neq 0$, then

$$\begin{aligned} \|(x_{n+1} - x)|z|\| &= \|\frac{z_{n+1}|z| - z|z_{n+1}|}{|z_{n+1}|}\| \\ &\leq \|\frac{z_{n+1}|z| - z_{n+1}|z_{n+1}|}{|z_{n+1}|}\| + \|\frac{z_{n+1}|z_{n+1}| - z|z_{n+1}|}{|z_{n+1}|}\| \\ &\leq 2m \end{aligned} \tag{5.2}$$

where the last line follows from the fact that no more than m balls are being added or removed at any update. Alternatively if $z_{n+1} = 0$, then it must be that $|z_n| \leq m$ since no more than m balls can be removed at a single update. In which case, $x_{n+1} = 0$ and $\|(x_{n+1} - x_n)|z_n|\| = \|z_n\| \leq m$. Thus, we get that $\|U_{n+1}\| \leq 4m$. Furthermore, $E[\|U_{n+1}\|^2 | \mathcal{F}_n] \leq E[\|(x_{n+1} - x_n)|z_n|\|^2 | \mathcal{F}_n] \leq 4m^2$.

To prove (iv), notice that if $z \neq 0$ and $z_{n+1} - z = w$, then $|z_{n+1}| = |z| + \alpha(w)$. Consequently,

$$(x_{n+1} - x)|z| = \begin{cases} \frac{|z|(w - \alpha(w)x)}{|z| + \alpha(w)} & \text{if } w \neq -z \\ -z & \text{if } w = -z. \end{cases}$$

Therefore, for $z \neq 0$

$$\begin{aligned} |z| E[x_{n+1} - x | z_n = z] &= \sum_{w \neq -z} \frac{|z|(w - \alpha(w)x)}{|z| + \alpha(w)} \Pi(z, z + w) \\ &\quad - z\Pi(z, 0). \end{aligned} \tag{5.3}$$

Notice that $\Pi(z, 0) > 0$ only when $|z| \leq m$. Assume that $|z| > m$. Equation (5.3) implies that

$$\begin{aligned} \|b_{n+1}\| &= \||z| E[x_{n+1} - x_n | z_n] - g(x)\| \\ &= \| \sum_{|w| \leq m} \frac{|z|}{|z| + \alpha(w)} \Pi(z, z + w)(w - \alpha(w)x) - g(x)\| \\ &\leq \| \sum_{|w| \leq m} \left(\frac{|z|}{|z| + \alpha(w)} \Pi(z, z + w) - p_w(x) \right) (w - \alpha(w)x)\|. \end{aligned}$$

Applying assumptions (A1) and (A2) implies that there is $K_1 > 0$ such that $\|b_{n+1}\| \leq K_1/|z_n|$ whenever $|z_n| > m$. On the other hand, if $|z| \leq m$, then the definition of b_{n+1} implies that $\|b_{n+1}\| \leq 2m + \sup_{x \in S_n} \|g(x)\|$. Choosing K sufficiently larger than K_1 completes the proof of (iv).

□

The recurrence relationship (5.1) can be viewed as a "noisy" Cauchy-Euler approximation scheme with step size $1/|z_n|$ for solving the ordinary differential equation

$$\frac{dx}{dt} = \sum_{w \in \mathbb{Z}^k} p_w(x)(w - x\alpha(w)) \tag{5.4}$$

which we call the *mean limit ODE.* When the number of individuals in the population grow without bound, the step size decreases to zero and it seems reasonable that there is a strong relationship between the limiting behavior of the mean limit ODE and the distribution of balls x_n. To make the relationship between the stochastic process x_n and the mean limit ODE more transparent, it is useful to define a continuous time version of x_n where time is scaled in an appropriate manner. Since the number of events (updates) that occur in a given time interval is likely to be proportional to the size of the population, we define *the time* τ_n *that has elapsed by update* n as

$$\begin{aligned} \tau_0 &= 0 \\ \tau_{n+1} &= \begin{cases} \tau_n + \frac{1}{|z_{n+1}|} & \text{if } z_{n+1} \neq 0 \\ \tau_n + 1 & \text{if } z_{n+1} = 0. \end{cases} \end{aligned}$$

The continuous time version of x_n is given by

$$X_t = x_n \qquad \text{for } \tau_n \leq t < \tau_{n+1}. \tag{5.5}$$

To relate the limiting behavior of the flow $\phi_t(x)$ of (5.4) to the limiting behavior of X_t, Schreiber (2001) proved the next theorem using techniques of Benaïm (1996). Recall, a set C is called *invariant* for the flow ϕ_t provided that $\phi_t(C) = C$ for all $t \in \mathbb{R}$. A compact invariant set C is *internally chain recurrent* provided that for every $x \in C$, $T > 0$ and $\epsilon > 0$, there exist points $x_1, x_2 \ldots, x_s$ in C and times $t_1, \ldots, t_s$ greater than T such that $x_1 = x_s = x$ and $\|\phi_{t_i}(x_i) - x_{i+1}\| < \epsilon$ for $1 \leq i \leq s-1$. Given a function $X_t : \mathbb{R}_+ \to \mathbb{R}^k$ or a sequence $\{x_n\}_{n \geq 0}$ in $\mathbb{R}^k$, we define the *limit sets*, $L(X_t)$ and $L(x_n)$, of X_t and x_n. $L(X_t)$ is the set of $p \in \mathbb{R}^k$ such that $\lim_{k\to\infty} X_{t_k} = p$ for some subsequence $\{t_k\}_{k\geq 0}$ with $\lim_{k\to\infty} t_k = \infty$. $L(x_n)$ is the set of $p \in \mathbb{R}^k$ such that $\lim_{k\to\infty} x_{n_k} = p$ for some subsequence $\{n_k\}_{k\geq 0}$ with $\lim_{k\to\infty} n_k = \infty$.

Theorem 5.3.1 (Schreiber 2001) *Let z_n be a Markov process satisfying (A1) and (A2) with mean limit ODE (5.4). Then on the event* $\{\liminf_{n\to\infty} \frac{|z_n|}{n} > 0\}$

1. *the interpolated process X_t is almost surely an* asymptotic pseudotrajectory *for the flow ϕ_t of the mean limit ODE. In other words, X_t almost surely satisfies*

$$\lim_{t\to\infty} \sup_{0\leq h\leq T} \|\phi_h X_t - X_{t+h}\| = 0$$

   for any $T > 0$.

2. *the limit set $L(X_t)$ of X_t is almost surely an internally chain recurrent set for the mean limit ODE.*

The first assertion of the theorem roughly states that X_t tracks the flow of the mean limit ODE with increasing accuracy far into the future. The second assertion of the theorem states that the only candidates for limit sets of the process x_n corresponding to the distribution of balls are connected compact internally chain recurrent sets for the mean limit flow.

To give a sense of the utility of this result, we derive some corollaries for the replicator processes and the fertility selection process in the next two subsections.

5.3.1 Implications for Replicator Processes

Let $\{R_n\}_{n\geq 0}$ and $\{\tilde{R}_n\}_{n\geq 0}$ be sequences of independent identically distributed random $k\times k$ matrices whose entries take values in the set $\{-1,0,1,\ldots,m\}$. Let $\{r_n\}_{n\geq 0}$ be a sequence of independent identically distributed random $k\times 1$ matrices whose entries take values in the set $\{-1,0,1,\ldots,m\}$. Let $z_n \in \mathbb{Z}_+^k$ be the replicator process associated with these random matrices. Define the mean payoff matrix by $A = E[R_0 + \tilde{R}_0]$. The limiting mean ODE associated with this process is given by a replicator equation [9]

$$\frac{dx}{dt} = \text{diag}(x)Ax - (x.Ax)x \qquad i = 1,\ldots,k \tag{5.6}$$

where $x.A$ denotes multiplying the left hand side of A by the transpose of x and $\text{diag}(x)$ is a diagonal $k\times k$ matrix with diagonal entries x^i. The dynamics of (5.6) are well-studied and have two remarkable properties whose proofs can be found in [9].

Theorem 5.3.2 (Exclusion principle) *If the replicator equation (5.6) has no equilibrium in* $\text{int}S_k$*, then every orbit of (5.6) converges to* ∂S_k*.*

Theorem 5.3.3 (Time averaging principle) *If the replicator equation (5.6) has a unique equilibrium* p *in* $\text{int}S_k$ *and if* $x(t)$ *is a solution of (5.6) such that* $L(x(t)) \subset \text{int}S_k$*, then*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T x(t)dt = p.$$

It turns out that Theorem 5.3.1 provides us the tool in which to transfer these theorems to replicator processes.

Theorem 5.3.4 *Let* z_n *be a replicator process on* $\mathbb{Z}_+^k$ *with mean payoff matrix* A*. If mean limit replicator equation has no equilibria in* $\text{int}S_k$*, then* $L(x_n)\cap\partial S_k \neq \emptyset$ *almost surely on the event* $\{\liminf_{n\to\infty}\frac{|z_n|}{n} > 0\}$*.*

PROOF: The proof of Theorem 5.3.2 implies there exists a vector $c \in \mathbb{R}^k$ such that the function $V(x) = \sum c^i \log x^i$ is strictly increasing along the forward orbits of the mean limit replicator equation that lie in $\text{int}S_k$. Consequently every compact connected internally chain recurrent set intersects ∂S_k. Applying Theorem 5.3.1 completes the proof.

□

Theorem 5.3.5 *Let z_n be a replicator process on $\mathbb{Z}_+^k$ with mean payoff matrix A. Let X_t be continuous-time process associated with z_n that is defined by (5.5). Assume that (5.6) has a unique rest point p in $\mathrm{int}S_k$. Then*

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T X_t\,dt = p$$

almost surely on the event

$$\{\liminf_{n\to\infty}\frac{|z_n|}{n} > 0\}\cap\{L(X_t)\subset \mathrm{int}S_k\}.$$

PROOF: Consider a trajectory X_t from the event $\{\liminf_{n\to\infty}\frac{|z_n|}{n} > 0\}\cap\{L(X_t)\subset \mathrm{int}S_k\}$. Theorem 5.3.1 implies that X_t is almost surely an asymptotic pseudotrajectory for the flow ϕ_t of the mean limit replicator equation. Theorem 5.3.1 implies that $L(X_t)$ is a compact internally chain recurrent set for the flow ϕ_t of (5.6). Theorem 5.3.3 implies that

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t \phi_s x\,ds = p$$

for all $x\in L(X_t)$, and this convergence is uniform. Therefore given $\epsilon > 0$, we can choose $T > 0$ and a compact neighborhood U of $L(X_t)$ such that

$$\|\frac{1}{T}\int_0^T \phi_s x\,ds - p\| < \frac{\epsilon}{3}$$

whenever $x\in U$. Since X_t is an asymptotic pseudotrajectory, there exists an $l\geq 1$ such that

$$\sup_{0\leq h\leq T}\|X_{t+h}-\phi_h X_t\| < \frac{\epsilon}{3}$$

for all $t\geq lT$. For any $i\in\mathbb{Z}_+$, define

$$\psi(i) = \|\int_0^T (X_{iT+s}-\phi_s X_{iT})\,ds\| + \|\int_0^T (\phi_s X_{iT}-p)\,ds\|.$$

Since $L(X_t)\subset U$, there is an $N\geq l$ such that $X_t\in U$ for all $t\geq NT$. Given any $t\in\mathbb{R}$, let $[t]$ denote the integer part of t. For any $t > (N+1)T$, we get

$$\begin{aligned}\|\frac{1}{t}\int_0^t X_s\,ds - p\| &\leq \frac{1}{t}\Big(\|\int_0^{NT}(X_s-p)\,ds\| + \psi(N)+\psi(N+1)+\dots \\ &\quad +\psi([\frac{t}{T}]-1) + \|\int_{[t/T]T}^t (X_s-p)\,ds\|\Big) \\ &\leq \frac{1}{t}(2NT+\epsilon T([t/T]-N)+2(t-[\frac{t}{T}]T)).\end{aligned}$$

Taking the limit as $t\to\infty$ we get that

$$\limsup_{t\to\infty}\|\frac{1}{t}\int_0^t X_s\,ds - p\| \leq \epsilon.$$

Taking the limit as $\epsilon\to 0$ completes the proof of the theorem.

□

5.3.2 Implications for Additive Fertility-Selection Processes

Let $z_n \in \mathbb{Z}_+^{k\times k}$ be a fertility-selection process defined by the sequence of random variables $G_n(ij, rs)$ with $1 \le i, j, r, s \le k$. Define $g(ij, rs) = E[G_0(ij, rs)]$. Define $x_n = z_n/|z_n|$ whenever $z_n \neq 0$ and $x_n = 0$ otherwise.

The mean limit ODE for this selection-fertility process is given by the fertility selection equations (see, e.g., Hofbauer and Sigmund (1998))

$$\frac{dx^{ij}}{dt} = \sum_{r,s=1}^{k} g(ir, js)x^{ir}x^{js} - x^{ij}\bar{g} \tag{5.7}$$

where

$$\bar{g} = \sum_{1\le i,j,r,s\le k} g(ir, js)x^{ir}x^{js}.$$

Now consider the special case, when the each allele contributes additively to number of progeny produced by a mating. In this case, if γ_{ij} is genotype A_iA_j's contribution to fertility, then the mating between genotypes A_iA_j and A_rA_s produces on average

$$g(ij, rs) = \gamma_{ij} + \gamma_{rs}$$

progeny. Under this additional assumption, equation (5.7) simplifies to

$$\frac{dx^{ij}}{dt} = x^j\gamma_i + x^i\gamma_j - 2x^{ij}\bar{\gamma} \tag{5.8}$$

where $x^i = \sum_{j=1}^{k} x^{ij}$ is the frequency of the allele A_i in the population,

$$\gamma_i = \sum_{r=1}^{k} \gamma_{ir}x^{ir}$$

is the average fertility of allele i in the population, and

$$\bar{\gamma} = \sum_{i=1}^{k} \gamma_i = \sum_{i,j=1}^{k} \gamma_{ij}x^{ij}$$

is the average fertility of the population.

Theorem 5.3.6 *If z_n is an additive fertility-selection process and the mean-limit ODE (5.8) has only a finite number of equilibria, then on the event $\{\liminf_{n\to\infty} \frac{|z_n|}{n} > 0\}$, x_n almost surely converges to an equilibrium of (5.8).*

The proof of this theorem follows from the work of Hofbauer and Sigmund (1998) that we include here for the reader's convenience.

PROOF: Define the Hardy-Weinberg manifold by

$$H = \{x : x^{ij} = x^ix^j \text{ for all } 1 \le i, j \le k\}.$$

Since for any solution $x(t)$ to (5.8)

$$\frac{d}{dt}\left(x^{ij}(t) - x^i(t)x^j(t)\right) = -(x^{ij}(t) - x^i(t)x^j(t))2\bar{\gamma},$$

$x^{ij}(t) - x^i(t)x^j(t)$ converges exponentially to zero. Hence, all compact connected internally chain recurrent sets lie in the Hardy-Weinberg manifold. On the Hardy-Weinberg manifold, the dynamics of (5.8) are determined by the Hardy-Weinberg relations $x^{ij} = x^i x^j$ and the differential equation

$$\frac{dx^i}{dt} = \gamma_i - x^i\bar{\gamma}. \tag{5.9}$$

This differential equation is the continuous-time selection equations with selection parameters γ_{ij} and, consequently, the mean fertility $\bar{\gamma}$ is a strict Lyapunov function for (5.9) (see, e.g., [9]). Hence, all compact connected internally chain recurrent sets correspond to compact connected sets of equilibria. Since we have assumed that there are only a finite number of equilibria, the only compact connected internally chain recurrent sets are individual equilibria. Applying Theorem 5.3.1 completes the proof of this theorem.

□

5.4 Growth and Convergence with Positive Probability

Theorem 5.3.1 helps to determine the limiting behavior of the genotypic composition of a population on the event of growth. However, it does not indicate which limiting behaviors occur with positive probability and sheds no insight into conditions that ensure that the population grows with positive probability. The goal of this section is to show that when the mean limit ODE admits an attractor at which growth is expected, the population grows with positive probability and its genotypic composition converges to the attractor with positive probability. Prior to stating and proving this result, we prove the following proposition that estimates the rate of growth on the event of convergence to a set where growth is expected.

Proposition 5.4.1 *Let z_n be a generalized urn process satisfying (A1) and (A2). Let $K \subset S_k$ be a compact set. If*

$$a = \inf_{x\in K}\sum_w p_w(x)\alpha(w) > 0,$$

then

$$\liminf_{n\to\infty}\frac{|z_n|}{n} \geq a$$

on the event $\{L(\{x_n\})_{n\geq 0} \subset K\} \cap \{\lim_{n\to\infty}|z_n| = \infty\}$.

Remark. If K is an equilibrium, $\liminf_{n\to\infty} \frac{|z_n|}{n} = a$ on the event $\{L(\{x_n\})_{n\geq 0} \subset K\} \cap \{\lim_{n\to\infty} |z_n| = \infty\}$.

PROOF: Let $\mathcal{E} = \{L(\{x_n\})_{n\geq 0} \subset K\} \cap \{\lim_{n\to\infty} |z_n| = \infty\}$. We will show that $\liminf_{n\to\infty} |z_n|/n \geq a - \epsilon$ on the event $\mathcal{E}$ for every $\epsilon > 0$. Let $\epsilon > 0$ be given. The definition of a, compactness of K, continuity of p_w, and assumption (A2) imply that there exist an integer I and compact neighborhood U of K such that $E[|z_{n+1}|-|z_n||z_n = z] \geq a-\epsilon$ whenever $|z| \geq I$ and $z/|z| \in U$. For each natural number m, define the event $\mathcal{E}_m = \{|z_n| \geq I, z_n/|z_n| \in U \text{ for } n \geq m\}$. Notice that $\mathcal{E} \subset \cup_{m=1}^{\infty} \mathcal{E}_m$. Define a sequence of random variables by $N_0 = 0$ and

$$N_{n+1} = \begin{cases} |z_{n+1}| - |z_n| & \text{if } |z_n| \geq I \text{ and } z_n/|z_n| \in U \\ a & \text{else} \end{cases}$$

for $n \geq 0$. Let

$$M_n = \sum_{i=1}^{n} \frac{1}{i}(N_i - E[N_i|\mathcal{F}_{i-1}]).$$

M_n is a martingale that satisfies

$$\sup_n E[M_n^2] \leq 4m^2 \sum_{i\geq 1} \frac{1}{i^2}$$

as $|N_i| \leq 2m$. Therefore by Doob's convergence theorem $\{M_n\}_{n\geq 1}$ converges almost surely. By Kronecker's lemma

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} N_i - E[N_i|\mathcal{F}_{i-1}] = 0 \tag{5.10}$$

almost surely. Since $\sum_{i=m+1}^{n} N_i = |z_n| - |z_m|$ for all $n \geq m$ on the event $\mathcal{E}_m$ and $E[N_i|\mathcal{F}_{i-1}] \geq a - \epsilon$ for all $i \geq 1$, equation (5.10) implies that $\liminf_{n\to\infty} \frac{|z(n)|}{n} \geq a - \epsilon$ almost surely on the event $\mathcal{E}_m$. It follows that $\liminf_{n\to\infty} \frac{|z(n)|}{n} \geq a - \epsilon$ almost surely on the event $\mathcal{E}$.

□

Let $\phi_t(x)$ denote the flow of the mean limit ODE in (5.4). A compact invariant set $\mathcal{A} \subset S_k$ is called an *attractor* provided that there is an open neighborhood $U \subset S_k$ of $\mathcal{A}$ such that

$$\cap_{t>0}\overline{\cup_{s\geq t}\phi_t U} = \mathcal{A}.$$

The basin of attraction $\mathcal{B}(\mathcal{A})$ of $\mathcal{A}$ is the set of points $x \in S_k$ satisfying $\inf_{y\in\mathcal{A}} \|\phi_t x - y\| \to 0$ as $t \to \infty$.

Define $Att_\infty(X)$ as the set of points $x \in S_k$ such that, for all $M \in \mathbb{N}$ and every open neighborhood U of x

$$P[|z_n| \geq M \text{ and } x_n \in U \text{ for some } n] > 0.$$

Theorem 5.4.1 *Let z_n be a generalized urn process satisfying (A1) and (A2). Let $\mathcal{A}$ be an attractor for mean limit ODE with basin of attraction $B(\mathcal{A})$. Assume that*

$$a = \inf_{x\in\mathcal{A}} \sum_w p_w(x)\alpha(w) > 0$$

and define

$$\mathcal{C} = \{\liminf_{n\to\infty} \frac{|z_n|}{n} \geq a\} \cap \{L(\{x_n\}_{n\geq 0}) \subseteq \mathcal{A}\}.$$

If U is an open set whose closure is contained in $B(\mathcal{A})$, then there exists a constant $K > 0$ such that for all $M \in \mathbb{N}$,

$$P[\mathcal{C}] \geq \left(1 - \frac{K}{M}\right) P[|z_n| \geq M \text{ and } x_n \in U \text{ for some } n\ ].$$

In particular, if

$$B(\mathcal{A}) \cap Att_\infty(X) \neq \emptyset,$$

then $P[\mathcal{C}] > 0$.

PROOF: Assumption $\inf_{x\in\mathcal{A}} \sum_w p_w(x)\alpha(w) > 0$ means that the population grows in a neighborhood of the attractor $\mathcal{A}$ when the population size is sufficiently large. It implies that there exist a_1, $a_2 > 0$ and a neighborhood $\mathcal{N}$ of $\mathcal{A}$ such that $E[|z_{n+1}| - |z_n||z_n] \geq a_1 1_{\{x_n\in\mathcal{N}, |z_n|\geq a_2\}}$. The proof relies on the following principle : remaining in a neighborhood of the attractor increases the population size and this increase in population size increases the likelihood of remaining near the attractor. Let U be an open set such that $\overline{U}$ is a compact subset of $B(\mathcal{A})$. Assume $M \in \mathbb{N}$ and $r \in \mathbb{N}$ are such that $P[|z_r| \geq M\ ,\ x_r \in U] > 0$. Choose a neighborhood V of $\mathcal{A}$ such that $\overline{V}$ is a compact subset of $\mathcal{N} \cap B(\mathcal{A})$. Since $\mathcal{A}$ is an attractor there exists a time $T_0 \geq 1$ such that the trajectories coming from $U \cup V$ rejoin the neighborhood V after time T_0. More precisely, there exists $\delta > 0$ such that, if $X_t \in U\cup V$, $T \geq T_0$ and $\|\Phi_T(X_t) - X_{t+T}\| < \delta$, then $X_{t+T} \in V$.

To avoid double subscripts, we let $z(r)$ denote z_r, $\tau(r)$ denote τ_r, etc. Given any $r \in \mathbb{N}$, define $r_0 = r$ and

$$r_k = \inf\{r > r_{k-1}\ ,\ \tau(r) - \tau(r_{k-1}) \geq T_0\}$$

for all $k \geq 1$. Since $\tau_{n+1} - \tau_n \geq \frac{1}{|z_0|+mn}$ for all n where m is the integer in assumption A1, $\lim_{n\to\infty} \tau_n = \infty$ and $r_n < +\infty$ for all n. Define $A = e^{T_0 m}$, $B = 3A$, and the following events for all $k \geq 1$

$E_1(k)$ $|z(r_k)| \geq \zeta^{k-1}B^{-1}M$

$E_2(k)$ for all $r \in [r_k, r_{k+1}]$, $x_r \in V$

where $\zeta = 1 + a_1T_0/2B$. Let $E(0)$ be the event $\{|z_r| \geq M\ ,\ x_r \in U\}$. For $k \geq 1$, define $E(k) = E(k-1) \cap E_1(k) \cap E_2(k)$. We will show that there exists a constant $F > 0$ such that $P[E(k+1)|E(0)] \geq P[E(k)|E(0)] - F/M\zeta^k$ for all $k \geq 0$. The proof of this estimate relies on three lemmas. The first consists in observing that the population size $|z_r|$ remains bounded on a time intervals of order T_0, namely between $B^{-1}|z(r_k)|$ and $B|z(r_k)|$ on $[r_k, r_{k+1}]$. The second one makes use of this claim to underestimate probability of being inside V on $[r_{k+1}, r_{k+2}]$ if $x(r_k) \in U\cup V$. The third lemma estimates the probability that the population grows sufficiently.

Lemma 5.4.1 *For enough large $|z(r_k)|$:*

1. *$B|z(r_k)| \geq |z_r| \geq B^{-1}|z(r_k)|$ for all $r \in [r_k, r_{k+1}]$*

2. $T_0 \leq \tau(r_{k+1}) - \tau(r_k) \leq 2T_0$
3. $T_0 B^{-1}|z(r_k)| \leq r_{k+1} - r_k \leq T_0 B|z(r_k)|$

PROOF: Suppose $|z(r_k)| > Am$. Let $u = \inf\{n \in \mathbb{N} : |z(r_k + n)| \leq A^{-1}|z(r_k)|\}$. Then

$$\begin{aligned}
\tau(r_k + u) - \tau(r_k) &\geq \sum_{0 \leq j < |z(r_k)|(1-A^{-1})/m} \frac{1}{A^{-1}|z(r_k)| + mj} \\
&\geq \int_0^{|z(r_k)|(1-A^{-1})/m} \frac{dx}{A^{-1}|z(r_k)| + mx} \\
&= \frac{1}{m} \ln(A) = T_0.
\end{aligned}$$

This proves that $A^{-1}|z(r_k)| \leq |z(r)|$ for all $r \in [r_k, r_{k+1})$, which implies for sufficiently large $|z(r_k)|$ that $B^{-1}|z(r_k)| \leq |z(r)|$ for all $r \in [r_k, r_{k+1}]$. One can show similarly that $2A|z(r_k)| \geq |z(r)|$ for all $r \in [r_k, r_{k+1}]$. The definition of r_k immediately implies that $\tau(r_{k+1}) - \tau(r_k) \geq T_0$. Since $T_0 \geq 1$ and $\tau(n+1) - \tau(n) \leq 1$ for all n, the definition of r_k also implies that $\tau(r_{k+1}) - \tau(r_k) \leq T_0 + 1 \leq 2T_0$. The proof of claim 1 and the fact that $T_0 \geq 1$ imply that $\frac{r_{k+1} - r_k}{|z(r_k)|B^{-1}} \geq \tau(r_{k+1}) - \tau(r_k) \geq T_0$ and $\frac{r - r_k}{2|z(r_k)|A} \leq \tau(r) - \tau(r_k) \leq T_0$ for all $r \in [r_k, r_{k+1})$. Claim 3 follows for sufficiently large $|z(r_k)|$.

□

Lemma 5.4.2 *There exists a $C > 0$ depending only on p_w, T_0 , δ, a_1 and m such that*

$$P[E_2(k+1)|E_1(k), x(r_k) \in U \cup V] \geq 1 - \frac{C}{M\zeta^k}$$

for all $k \geq 0$ and $M > 0$ sufficiently large.

PROOF: Suppose $E_1(k)$ is satisfied and $x(r_k) \in U \cup V$. Lemma 5.4.1 implies that for large enough M, $|z(r)| \geq B^{-2}|z(r_k)|$ for $r \in [r_k, r_{k+2}]$, $4T_0 \geq \tau(r) - \tau(r_k) \geq T_0$ for $r \in [r_{k+1}, r_{k+2}]$, and $r_{k+2} - r_k \leq 2B^2 T_0 |z(r_k)|$. Define $g(x) = \sum_w p_w(x)(w - x\alpha(w))$. Let L be the Lipschitz constant for g. Using Gronwall's inequality, we prove the following estimate :

$$\sup_{r \in [r_{k+1}, r_{k+2}]} \|\Phi_{\tau(r)-\tau(r_k)} x(r_k) - x(r)\| \leq e^{4LT_0} \left(\Gamma_1(r_k, r_{k+2}) + \Gamma_2(r_k, r_{k+2})\right) \tag{5.11}$$

where

$$\Gamma_1(r_k, r_{k+2}) = \sup_{r_k \leq l \leq r_{k+2}-1} \| \sum_{i=r_k}^{l} \frac{U(i+1)}{|z(i)|} \|$$

and

$$\Gamma_2(r_k, r_{k+2}) = \frac{2 \sup \|g(x)\|}{\inf_{r_k \leq r \leq r_{k+2}} |z(r)|} + \sup_{r_k \leq l \leq r_{k+2}-1} \| \sum_{i=r_k}^{l} \frac{b(i+1)}{|z(i)|} \|.$$

To prove (5.11), let $X(t) = X_t$ denote the continuous time version of x_n defined in (5.5) and $c(t) = \sup\{n \in \mathbb{Z}_+ : t \geq \tau(n)\}$. Notice that for any $h \geq 0$ and $t \geq 0$

$$\begin{aligned} X(t+h) - X(t) = \quad & x(c(t+h)) - x(c(t)) = \textstyle\sum_{i=c(t)}^{c(t+h)-1} x(i+1) - x(i) \\ = \quad & \textstyle\sum_{i=c(t)}^{c(t+h)-1} \frac{g(x(i))+U(i+1)+b(i+1)}{|z(i)|} \\ = \quad & \textstyle\int_{\tau(c(t))}^{\tau(c(t+h))} g(X(s))\,ds + \sum_{i=c(t)}^{c(t+h)-1} \frac{U(i+1)+b(i+1)}{|z(i)|} \\ = \quad & \textstyle\int_{\tau(c(t))-t}^{\tau(c(t+h))-t} g(X(t+s))\,ds + \sum_{i=c(t)}^{c(t+h)-1} \frac{U(i+1)+b(i+1)}{|z(i)|} \end{aligned}$$

Since $\phi_h X(t) = X(t) + \int_0^h g(\phi_s X(t))\,ds$, the previous equalities imply that

$$\begin{aligned} \|\phi_h(X(t)) - X(t+h)\| \leq \quad & \textstyle\int_0^h \|g(\phi_s(X(t)) - g(X(t+s))\|ds + \int_{\tau(c(t))-t}^0 \|g(X(t+s))\|ds \\ & \textstyle + \int_{\tau(c(t+h))-t}^h \|g(X(t+s))\|ds + \| \sum_{i=c(t)}^{c(t+h)-1} \frac{U(i+1)+b(i+1)}{|z(i)|} \| \\ \leq \quad & \textstyle L\int_0^h \|\phi_s(X(t)) - X(t+s)\|ds + \frac{\|g\|_0}{|z(c(t))|} \\ & \textstyle \frac{\|g\|_0}{|z(c(t+h))|} + \| \sum_{i=c(t)}^{c(t+h)-1} \frac{U(i+1)+b(i+1)}{|z(i)|} \|. \end{aligned}$$

Choosing $t = \tau(r_k)$ and applying Gronwall's inequality to the previous inequality over the interval $0 \leq h \leq \tau(r_{k+2}) - \tau(r_k)$ gives the desired estimate.

Since $E_1(k)$ holds, $|z(r)| \geq |z(r_k)|B^{-2} \geq \zeta^{k-1}B^{-3}M$ for all $r \in [r_k, r_{k+2}]$. This observation plus the fact that there exists $K > 0$ such that $\|b(n+1)\| \leq \frac{K}{|z(n)|}$ imply that $e^{4LT_0}\Gamma_2(r_k, r_{k+2}) < \frac{\delta}{2}$ for M sufficiently large. On the other hand, Doob's inequality and Lemma 5.3.1 imply that

$$\begin{aligned} E[\sup_{r_{k-1} \leq l \leq r_{k+2}-1} \|\sum_{i=r_k}^{l} \frac{U(i+1)}{|z(i)|}\|^2 \mid z(r_k)] \quad & \leq \quad 16m^2 E[\sum_{i=r_k}^{r_{k+2}-1} \frac{1}{|z(i)|^2}|z(r_k)] \\ & \leq \quad \frac{16m^2B^4}{|z(r_k)|^2} E[r_{k+2} - r_k|z(r_k)] \\ & \leq \quad \frac{32m^2T_0B^6}{|z(r_k)|} \leq \frac{32m^2B^7T_0}{M\zeta^{k-1}} \end{aligned}$$

Therefore

$$P[\sup_{r_k \leq l \leq r_{k+2}-1} \|\sum_{i=r_k}^{l} \frac{U(i+1)}{|z(i)|}\| \geq e^{-4LT_0}\frac{\delta}{2} \mid z(r_k)] \leq \frac{128m^2T_0B^7e^{8LT_0}}{\delta^2 M\zeta^{k-1}}. \qquad (5.12)$$

Define $\mathcal{E} = \{\sup_{r \in [r_{k+1}, r_{k+2}]} d(\Phi_{\tau(r)-\tau(r_k)}x(r_k), x(r)) \leq \delta\}$. Since $\tau(r_{k+1}) - \tau(r_k) \geq T_0$ and $x(r_k) \in U \cup V$, our choice of T_0 implies that $x(r) \in V$ for all $r \in [r_{k+1}, r_{k+2}]$ on the event $\mathcal{E}$. Inequalities (5.11) and (5.12) imply that $P[\mathcal{E}] \geq 1 - \frac{128m^2T_0B^7e^{8LT_0}}{\delta^2 M\zeta^{k-1}}$ for M sufficiently large.

□

Lemma 5.4.3 *There exists $D > 0$ depending only on p_w, T_0, δ, a_1 and m such that*

$$P[E_1(k+1) \cup E_2(k)^c | E_1(k)] \geq 1 - \frac{D}{M\zeta^k}$$

for all $k \geq 1$ and $M > 0$ sufficiently large.

PROOF: Define $N(i) = |z(i)| - |z(i-1)|$, $D(i) = N(i) - E[N(i)|z(i-1)]$ and $G(k+1) = \frac{1}{r_{k+1}-r_k}\sum_{i=r_k+1}^{r_{k+1}} D(i)$. Observe that $|D(n)| \leq 2m$. Therefore

$$\begin{aligned} E\left[G(k+1)^2|z(r_k)\right] &\leq \frac{B^2}{T_0^2|z(r_k)|^2}E[(\sum_{i=r_k+1}^{r_k+T_0B|z(r_k)|} D(i)1_{\{i-1<r_{k+1}\}})^2 \mid z(r_k)] \\ &\leq \frac{B^2}{T_0^2|z(r_k)|^2}\sum_{i=r_k+1}^{r_k+T_0B|z(r_k)|} E[D(i)^2 \mid z(r_k)] \leq \frac{4m^2B^3}{T_0|z(r_k)|} \end{aligned}$$

where we have used the fact that $T_0B^{-1}|z(r_k)| \leq r_{k+1} - r_k \leq T_0B|z(r_k)|$. It follows that $P[G(k+1) \leq -\frac{a_1}{2}] \leq P[G(k+1)^2 \geq \frac{a_1^2}{4}] \leq \frac{D}{M\zeta^k}$, with $D = \frac{16m^2B^4\zeta}{a_1^2T_0}$. Since $\zeta = 1 + a_1T_0/2B$, it follows that

$$E_1(k+1)^c \cap E_1(k) \cap E_2(k) \subset \{G(k+1) \leq -a_1/2\}.$$

□

These three lemmas imply that there exists $F > 0$ depending only on p_w, T_0 , δ, a_1 and m such that

$$P[E(k+1)|E(0)] \geq P[E(k)|E(0)] - \frac{F}{M\zeta^k}$$

for all $k \geq 0$ and $M > 0$ sufficiently large. Indeed, due to the fact that for all $k \geq 1$ $E(k)$ equals the disjoint union of $E(k+1)$, $E_1(k+1)^c \cap E(k)$, and $E_2(k+1)^c \cap E_1(k+1) \cap E(k)$, we get

$$\begin{aligned} P[E(k+1)|E(0)] &= P[E(k)|E(0)] - P[E_2(k+1)^c \cap E_1(k+1) \cap E(k)|E(0)] \\ &\quad -P[E_1(k+1)^c \cap E(k)|E(0)] \\ &\geq P[E(k)|E(0)] - P[E_2(k+1)^c \cap E_1(k) \cap \{x(r_k) \in U \cup V\}|E(0)] \\ &\quad -P[E_1(k+1)^c \cap E_2(k)|E(0)] \\ &\geq P[E(k)|E(0)] - \frac{F}{M\zeta^k} \end{aligned}$$

where the first inequality follows from the inclusions $E(k) \cap E_1(k+1) \subset E_1(k) \cap \{x(r_k) \in U \cup V\}$ and $E(k) \subset E_2(k)$, and the second equality follows from Lemmas 5.4.2 and 5.4.3 with $F = C + D$. These inequalities remain true for $k = 0$ since $E_1(1)$ always holds.

It follows that

$$\begin{aligned} P[\lim_{k\to\infty} E(k)] &\geq P[E(0)]\left(1 - \sum_{k=0}^{\infty}\frac{F}{M\zeta^k}\right) \\ &\geq P[E(0)]\left(1 - \frac{\zeta F}{M(\zeta-1)}\right). \end{aligned}$$

The definition of $E(k)$ implies that $P[\mathcal{C}] \geq P[E(0)]\left(1 - \frac{\zeta F}{M(\zeta-1)}\right)$ where $\mathcal{C} = \{\liminf_{n\to\infty}\frac{|z_n|}{n} > 0\} \cap \{x_n \in U \cup V \text{ i.o.}\}$. On the event $\mathcal{C}$, Theorem 5.3.1 implies that $L(\{x_n\})$ is a compact internally chain recurrent set for the mean limit ODE. Since $L(\{x_n\}) \cap B(\mathcal{A}) \neq \emptyset$ on the event $\mathcal{C}$, a basic result about internally chain recurrent sets (see, e.g., [4, Cor. 5.4]) implies that $L(\{x_n\}) \subset \mathcal{A}$ on the event $\mathcal{C}$. Setting $K = \frac{\zeta F}{M(\zeta-1)}$ and applying Proposition 5.4.1 completes the proof of the first assertion of the theorem. To prove the second assertion, assume that $p \in Att_\infty(X) \cap \mathcal{B}(\mathcal{A})$, choose U an open neighborhood of p such that $\bar{U} \subset \mathcal{B}(\mathcal{A})$, and apply the first assertion of the theorem.

□

5.5 Non-Convergence

In this section, we show that there are two types of invariant sets of the mean limit ODE toward which the generalized urn process does not converge. The first type corresponds to a compact set where growth of the process is not expected, and the second type corresponds to a "non-degenerate" equilibrium or periodic orbit.

Proposition 5.5.1 *Let z_n be a Markov process on $\mathbb{Z}_+^k$ satisfying (A1) and (A2). If $K \subset S_k$ is a compact set satisfying*

$$\sup_{x \in K} \sum_w p_w(x)\alpha(w) < 0, \tag{5.13}$$

then $P[\{L(x_n) \subset K\} \cap \{\lim_{n\to\infty} |z_n| = +\infty\}] = 0$.

PROOF: Equation (5.13) implies that we can choose a neighborhood U of K, $N \in \mathbb{N}$, and $\epsilon > 0$ such that

$$\sup_{|z|\geq N, z/|z| \in U} E[|z_{n+1}| - |z_n||z_n = z] \leq -\epsilon. \tag{5.14}$$

Given any $l \in \mathbb{N}$ such that $x_l \in U$ and $|z_l| \geq N$, define the stopping time

$$T = \inf\{n \geq l : x_n \notin U \text{ or } |z_l| < N\}.$$

For any $n \geq l$, we get

$$\begin{aligned} 0 &\leq E[|z_{T\wedge n}|] = E[\sum_{i=l+1}^{T\wedge n} |z_i| - |z_{i-1}|] + E[|z_l|] \\ &= \sum_{i=l+1}^{n} E[|z_{i\wedge T}| - |z_{(i-1)\wedge T}|] + E[|z_l|] \\ &\leq -\epsilon \sum_{i=l+1}^{n} P[T \geq i] + E[|z_l|] \end{aligned}$$

Taking the limit as $n \to \infty$, we get that $\sum_{i=l+1}^{\infty} P[T \geq i] \leq E[|z_l|]/\epsilon$. The Borel-Cantelli lemma implies that $P[T = \infty] = 0$. It follows that $P[\{L(x_n) \subset K\} \cap \{\lim_{n\to\infty} |z_n| = +\infty\}] = 0$.

□

If A is a subset of $\mathbb{R}^k$, then we let $\mathrm{Span}(A) \subset \mathbb{R}^k$ denote the vector space spanned by the points in A. Given a compact subset $\mathcal{U} \subset \mathrm{int}(S_k)$ we say that the process $\{z_n\}$ is *nondegenerate at* $\mathcal{U}$ if for all $x \in \mathcal{U}$,

$$\mathrm{Span}\{w \in \mathbb{Z}^k : p_w(x) > 0\} = \mathbb{R}^k.$$

Recall, a periodic orbit or an equilibrium of an ODE is *linearly unstable* provided that one of its characteristic exponents is greater than zero.

Theorem 5.5.1 *Let $\{z_n\}$ be a generalized urn process satisfying (A1) and (A2). Let $\mathcal{U} \subset \mathrm{int}(S_k)$ be a linearly unstable equilibrium or a hyperbolic linearly unstable orbit for the mean limit ODE. Assume*

(a) There exists $\beta > 1/2$ such that the functions p_w are $C^{1+\beta}$ in a neighborhood of $\mathcal{U}$.

(b) $\{z_n\}$ is nondegenerate at $\mathcal{U}$.

Then $P[(L(x_n) \subset \mathcal{U}) \cap \{\liminf_{n\to\infty} \frac{|z_n|}{n} > 0\}] = 0$.

PROOF: Let $N(\mathcal{U})$ be a neighborhood of $\mathcal{U}$. The event

$$\{L(x_n) \subset \mathcal{U} \text{ and } \liminf_{n\to\infty} \frac{|z_n|}{n} > 0\}$$

is contained in the event

$$\bigcup_{N\in\mathbb{Z}_+, a\in\mathbb{Q}_+^*} E_{N,a}$$

where $\mathbb{Q}_+^*$ denotes the positive rationals and

$$E_{N,a} = \{L(x_n) \subset \mathcal{U}\} \cap \{\forall n \geq N\, |z_n| \geq na \text{ and } x_n \in N(\mathcal{U})\}$$

In order to prove Theorem 5.5.1 it then suffices to prove that for N large enough and $a \in \mathbb{Q}_+^*$

$$P[E_{N,a}] = 0.$$

Let $\mathcal{F}_n$ denote the sigma field generated by $z_0, \ldots, z_n$, and $V_{n+1} = |z_n|(x_{n+1} - x_n)$. Let F denote the vector field on S_k defined by $F(x) = \sum_w p_w(x)(w - x\alpha(w))$. Let $\{\epsilon_n\}$ denote a sequence of bounded, zero-mean i.i.d. random variables taking values in

$$TS_k = \{u \in \mathbb{R}^k : \sum u_i = 0\}$$

whose covariance matrix is nondegenerate (i.e has rank $k-1$).

Define the sequence $\{\tilde{x}_n\}_{n\geq N}$ as follows :

$$\tilde{x}_N = x_N,$$

$$\tilde{x}_{n+1} - \tilde{x}_n = \begin{cases} \frac{1}{|z_n|}(F(\tilde{x}_n) + V_{n+1} - F(x_n)) & \text{if } x_n \in N(\mathcal{U}) \text{ and } |z_n| \geq na \\ \frac{1}{an}(F(\tilde{x}_n) + \epsilon_{n+1}) & \text{otherwise} \end{cases} \tag{5.15}$$

The processes $\{x_n\}$ and $\{\tilde{x}_n\}$ coincide on the event $E_{N,a}$. On the other hand,

$$P[\lim_{n\to\infty} dist(\tilde{x}_n, \mathcal{U}) = 0] = 0$$

in view of the following theorem whose proof is an easy adaptation of [13] , Theorem 2.

Theorem 5.5.2 *Let $\{\mathcal{F}_n\}_{n\in\mathbb{Z}^+}$ denote a nondecreasing sequence of sub-σ-algebras of $\mathcal{F}$, and $(\tilde{x}_n)$ a sequence of adapted random variables given by*

$$\tilde{x}_{n+1} - \tilde{x}_n = \beta_n(F(\tilde{x}_n) + \tilde{U}_{n+1} + \tilde{b}_{n+1}) \tag{5.16}$$

where F is a $C^{1+\beta}$ vector field, with $1/2 < \beta \leq 1$, $\{\tilde{U}_n\}$, $\{\tilde{b}_n\}$ and $\{\beta_n\}$ are adapted random variables.

We define $\tilde{V}_{n+1} = F(\tilde{x}_n) + \tilde{U}_{n+1} + \tilde{b}_{n+1}$.

Assume

(i) $\exists K > 0,\ \forall n \in \mathbb{Z}_+\ \|\tilde{U}_n\| \leq K$ and $E(\tilde{U}_{n+1}|\mathcal{F}_n) = 0$.

(ii) There exist a, $b > 0$ and a deterministic sequence $\{\gamma_n\}$ of nonnegative numbers having infinitely positive terms, such that $\forall n \in \mathbb{Z}_+$, $a\gamma_n \leq \beta_n \leq b\gamma_n$

(iii) $\sum \tilde{r}_i^2 < +\infty$

(iv) $\mathcal{U} \subset Int(S_k)$ is a linearly unstable periodic orbit or equilibrium for F

(v) There exist a neighbourhood $N(\mathcal{U})$ of $\mathcal{U}$ and $c > 0$ such that, for all unit vector $v \in \mathbb{R}^m$, $E(| < \tilde{V}_{n+1}, v > ||\mathcal{F}_n) \geq c1_{\{\tilde{x}_n\in N(\mathcal{U})\}}$.

Then $P[L(\tilde{x}_n) \subset \mathcal{U}] = 0$.

Let

$$A_n = \{x_n \in N(\mathcal{U}) \text{ and } |z_n| \geq na\}.$$

Using notation of Lemma 5.3.1 set

$$\tilde{U}_{n+1} = U_{n+1}, \tilde{b}_{n+1} = b_{n+1}, \beta_n = 1/|z_n| \text{ on } A_n,$$

and

$$\tilde{U}_{n+1} = \epsilon_{n+1},\ \tilde{b}_{n+1} = 0,\ \beta_n = 1/an \text{ on } A_n^c.$$

Then, by Lemma 5.3.1, the process $\{\tilde{x}_n\}$ defined by (5.15) verifies resursion (5.16) and assertions (i),(ii),(iii) and (iv) of Theorem 5.5.2 are satisfied.

It remains to verify assertion (v). Let $B(1) = \{v \in S_k : ||v|| = 1\}$. Let $G_1 : S_k \times B(1) \to \mathbb{R}_+$ and $G_2 : \mathcal{U} \times S_k \times B(1) \to \mathbb{R}_+$ be the functions defined by

$$G_1(\tilde{x}, v) = E(| < F(\tilde{x}) + \epsilon_n, v > |)$$

and

$$G_2(x, \tilde{x}, v) = \sum_e | < F(\tilde{x}) - F(x) + Q_x(e), v > |p_e(x)$$

where Q_x denote the projection operator $Q_x : \text{Span}\{x\} \oplus TS_k \to TS_k$.

Then it is not hard to verify that

$$E(< \tilde{V}_{n+1}, v > ||\mathcal{F}_n) = G_1(\tilde{x}_n, v)\mathbf{1}_{A_n^c} + G_2(x, \tilde{x}_n, v)\mathbf{1}_{A_n} + O(1/n).$$

By continuity of G_1, G_2 compactness of the sets $\mathcal{U}, S_k, B(1)$ and assumption (A2), there exists $b > 0$ and a the neighborhood $N(\mathcal{U})$ such that $G_1(\tilde{x}, v) > b$, $G_2(x, \tilde{x}, v) > b$ for all $x \in N(\mathcal{U}), \tilde{x} \in S_k$ and $v \in B(1)$.

Assumption (v) is thus verified for $n \geq N$ and N large enough.

□

5.6 Non-degenerate Processes with Gradient-Like Mean Limit ODEs

Using the results from the previous two sections, we can prove the following result.

Theorem 5.6.1 *Let z_n be a generalized urn process satisfying (A1) and (A2), p_w are $C^{1+\beta}$ for some $\beta > 1/2$, z_n is non-degenerate on S_k, $Att_\infty(X) = S_k$, and the chain recurrent set for the mean limit ODE consists of hyperbolic equilibria q that satisfy $\sum_w p_w(q)\alpha(w) \neq 0$. Then $P[\liminf_{n\to\infty} \frac{|z_n|}{n} > 0]$ if and only if there exists a linearly stable equilibrium q such that $\sum_w p_w(q)\alpha(w) > 0$. Furthermore, $P[\{L(x_n) = q\}] > 0$ for any linearly stable equilibrium q satisfying $\sum_w p_w(q)\alpha(w) > 0$, and $P[\{L(x_n) = q\}] = 0$ for any equilibrium q which is linearly unstable.*

PROOF: Define $\mathcal{G} = \{\liminf_{n\to\infty} |z_n|/n > 0\}$. Since z_n is non-degenerate, Theorem 5.3.1 and Theorem 5.5.1 imply that $L(x_n)$ is contained almost surely in the set of linearly stable equilibria on the event $\mathcal{G}$. If $\sum_w p_w(q)\alpha(w) < 0$ for all linearly stable equilibria q, then Proposition 5.5.1 implies $P[\mathcal{G}] = 0$. Alternatively, if q is a linearly stable equilibrium and $\sum_w p_w(q)\alpha(w) > 0$, then Theorem 5.4.1 implies $P[\mathcal{G} \cap \{L(x_n) = q\}] > 0$. Finally, if q is an equilibrium which is linearly unstable, then Proposition 5.5.1 implies $P[L(x_n) = q] = 0$ if $\sum_w p_w(q)\alpha(w) < 0$ and Proposition 5.4.1 and Theorem 5.5.1 imply $P[L(x_n) = q] = 0$ if $\sum_w p_w(q)\alpha(w) > 0$.

□

As an application of this result, we consider an additive fertility selection processes z_n with mutation where $g(ij, rs) = E[G_n(ij, rs)]$ is the the expected number of progeny produced by a mating between genotypes A_iA_j and A_rA_s, and $\mu(ij, rs)$ is the probability genotype A_iA_j mutates to genotype A_rA_s.

Corollary 5.6.1 *Let z_n be an additive fertility selection process with mutation such that $\mu(rs, ij)$ are positive for all $1 \leq i, j, r, s \leq k$ and sufficiently small whenever $\{r, s\} \neq \{i, j\}$, $P[G_1(ij, rs) \geq 3] > 0$ for all $1 \leq i, j, r, s \leq k$, and such that the fertility selection equation (5.8) without mutation has hyperbolic equilibria q satisfying $\sum_w p_w(q)\alpha(w) \neq 0$. Then $P[\liminf_{n\to\infty} \frac{|z_n|}{n} > 0]$ if and only if there exists a linearly stable equilibrium q such that $\sum_w p_w(q)\alpha(w) > 0$. Furthermore, $P[\{L(\{x_n\}) = q\}] > 0$ for any linearly stable equilibrium q satisfying $\sum_w p_w(q)\alpha(w) > 0$, and $P[\{L\{x_n\} = q\}] = 0$ for any equilibrium q which is linearly unstable.*

PROOF: Since the mean limit ODE corresponding to the fertility selection process without mutation is gradient-like and has only hyperbolic equilibria, the chain-recurrent set for this mean limit ODE equals the set of equilibria. Consequently, the mean limit ODE corresponding to the fertility selection process with sufficiently small mutation rates also has a chain-recurrent set consisting only hyperbolic equilibria. Due to the fact that all mutation rates are positive, this process is non-degenerate on the entire simplex. Since $P[G_1(ij,rs) \geq 3] > 0$ for all $1 \leq i,j,r,s \leq k$, $Att_\infty(X_t)$ is the entire simplex. Applying Theorem 5.6.1 completes the proof.

□

Acknowledgements. This research was supported in part by National Science Foundation Grant DMS-0077986 to SJS.

Bibliographie

[1] B. Arthur, Y. Ermol'ev, and Y. Kaniovskii, *A generalized urn problem and its applications*, Cybernetics **19** (1983), 61–71.

[2] M. Benaïm, *A dynamical systems approach to stochastic approximations*, SIAM J. Control Optim. **34** (1996), 437–472.

[3] ———, *Vertex-reinforced random walks and a conjecture of Pemantle*, Ann. Probab. **25** (1997), 361–392.

[4] ———, *Dynamics of stochastic approximation algorithms*, Le Séminaire de Probabilites, Lectures Notes in Mathematics, **1709** (1999), 1–68.

[5] M. Benaïm and M. W. Hirsch, *Dynamics of Morse-Smale urn processes*, Ergodic Theory Dynam. Systems **15** (1995), 1005–1030.

[6] ———, *Asymptotic pseudotrajectories and chain recurrent flows, with applications*, J. Dynam. Differential Equations **8** (1996), 141–176.

[7] R. Durrett, *Probability : Theory and examples (2nd edition)*, Duxbury Press, Boston, 1996.

[8] B. M. Hill, D. Lane, and W. Sudderth, *A strong law for some generalized urn processes*, Ann. Prob. **8** (1980), 214–226.

[9] J. Hofbauer and K. Sigmund, *Evolutionary games and population dynamics*, Cambridge University Press, 1998.

[10] R. Pemantle, *Nonconvergence to unstable points in urn models and stochastic approximations*, Ann. Prob. **18** (1990), 698–712.

[11] M. Posch, *Cycling in a stochastic learning algorithm for normal form games*, Journal of Evolutionary Economics **7** (1997), 193–207.

[12] S. J. Schreiber, *Urn models, replicator processes, and random genetic drift*, SIAM J. Appl. Math., **61** (2001), 2148–2167

[13] P. Tarrès, *Pièges répulsifs*, Comptes Rendus de l'Académie des Sciences Sér. I Math **330** (2000), 125–130.

Chapitre 6

Un théorème de Pemantle et Volkov

Résumé : Les marches aléatoires renforcées par sommets, définies par Pemantle (1988,[3]), sont des processus évoluant dans un environnement qu'ils contribuent à faire changer par le fait qu'ils ont une probabilité plus grande de revenir aux endroits déjà visités. Pemantle et Volkov ([5]) ont montré que, lorsque le graphe sous-jacent est $\mathbb{Z}$, la marche reste presque-sûrement bloquée dans un ensemble fini d'au moins cinq points et, avec une probabilité strictement positive, dans cinq points. Nous donnons ici une preuve rapide de ces résultats.

6.1 Introduction

Soient G un graphe localement fini, $\sim$ la relation d'adjacence de G et $S(G)$ l'ensemble de ses sommets.

Soient $(\Omega, \mathcal{F}, P)$ un espace de probabilité, et $(X_n)_{n\in\mathbb{N}}$ un processus à valeurs dans $S(G)$; notons $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$.

Pour tous $v \in S(G)$ et $n \in \mathbb{N} \cup \{\infty\}$, soit $Z_n(v)$ le nombre de fois à l'instant n où ce sommet a été visité plus un, c'est à dire

$$Z_n(v) = 1 + \sum_{i=0}^{n} 1_{X_i = v}.$$

Considérons une application $r : S(G) \times S(G) \longrightarrow \mathbb{R}^+$ telle que, pour tous v, $w \in S(G)$, $r(v,w) > 0 \Longleftrightarrow v \sim w$.

Alors $(X_n)_{n\in\mathbb{N}}$ est une *marche aléatoire renforcée par sommets* de renforcement r et issue de $v_0 \in S(G)$ si $X_0 = v_0$ et, pour tout $n \in \mathbb{N}$,

$$P(X_{n+1} = v \mid \mathcal{F}_n) = 1_{v\sim X_n} \frac{r(X_n, v) Z_n(v)}{\sum_{w\sim X_n} r(X_n, w) Z_n(w)}.$$

Ce type de marche aléatoire avec mémoire a été introduit par Pemantle en 1988 ([3]), et analysé en détail sur des graphes finis par Pemantle (1992,[4]) et Benaïm (1997,[1]). Une description détaillée des applications (auto-organisation, apprentissage, économie) et des résultats connus sur le sujet est donnée dans [5] (voir aussi [6]).

Nous nous plaçons dans la suite de l'article dans l'hypothèse où $G = \mathbb{Z}$ et $r(v,w) = 1_{v\sim w}$. Dans ce cadre, Pemantle et Volkov ont obtenu des résultats très précis sur le comportement asymptotique de X_n (partiellement généralisés par Volkov ([8]) à une classe assez large de graphes localement finis.), que nous rappelons ci-dessous.

Définissons

$$R = \{v \in \mathbb{Z} / \ \exists n \in \mathbb{N},\ X_n = v\} \ , \ R' = \{v \in \mathbb{Z} / \ Z_\infty(v) = \infty\}.$$

Etants donnés $v \in \mathbb{Z}$ et $\alpha \in]0,1[$, considérons les six événements :

(i) $R' = \{v-2, v-1, v, v+1, v+2\}$;

(ii) $\ln Z_n(v-2)/\ln n \to \alpha$;

(iii) $\ln Z_n(v+2)/\ln n \to 1-\alpha$;

(iv) $Z_n(v-1)/n \to \alpha/2$;

(v) $Z_n(v+1)/n \to (1-\alpha)/2$;

(vi) $Z_n(v)/n \to 1/2$.

Théorème 6.1.1 *([5])* $P(|R| < \infty) = 1$.

Théorème 6.1.2 *([5])* $P(|R'| \leq 4) = 0$.

Théorème 6.1.3 *([5]) Pour tout ouvert* $I \subset]0,1[$ *et tout* $v \in \mathbb{Z}$*, il existe avec une probabilité strictement positive un réel* $\alpha \in I$ *tel que les événements* **(i)** *à* **(vi)** *se réalisent.*

Conjecture 6.1.1 *([5]) Il existe presque-sûrement* $v \in \mathbb{Z}$ *et* $\alpha \in]0,1[$ *tels que les événements* **(i)** *à* **(vi)** *se réalisent.*

Nous prouvons cette conjecture dans [7]. L'objet de cet article est de présenter une preuve rapide des théorèmes 6.1.1, 6.1.2 et 6.1.3. Nous utilisons en partie l'heuristique de l'article initial de Pemantle et Volkov. Cependant l'intérêt de cette nouvelle preuve est de faire apparaître les théorèmes 1 et 3 comme conséquences d'un même résultat (le lemme 1) donnant à chaque instant des conditions suffisantes pour que la marche reste bloquée à droite (ou à gauche) d'un point fixé. Et inversement de déduire le théorème 2 d'un résultat simple (le lemme 2, énoncé initialement dans [7]) donnant, lorsque la marche reste à partir d'un certain temps bloquée à droite d'un point fixé, des indications sur les fréquences des visites à droite de ce point. Ces deux résultats sont en grande partie basés sur la construction de deux sur-martingales, nommées ici V_n et W_n.

6.2 Démonstration des théorèmes 1 à 3

Soient A, B, C, D, E, F et G sept points successifs, dans l'ordre croissant ou décroissant. Définissons, pour tout $n \in \mathbb{N} \cup \{\infty\}$, $A_n = Z_n(A)$, $B_n = Z_n(B)$, etc..., et si $n \geq 2$, l'instant de $(n-1)$-ième visite en C

$$t_n = \inf\{m \in \mathbb{N} / \ C_m = n\}.$$

Soient $a \in \mathbb{N}^*$, $\mu > 1$, ζ, $\xi \in]0,1[$ tels que $\xi > \zeta$. Pour tout $n \in \mathbb{N}$, définissons les événements suivants :

$\mathcal{E}_1(n)_a$: $A_n = a$,

$\mathcal{E}_2(n)_{\mu,\xi}$: $B_n \leq \mu C_n^{\xi}$,

$\mathcal{E}_3(n)_\zeta$: $C_n/D_n \leq \zeta$,

$\mathcal{E}(n)_{a,\mu,\zeta,\xi} = \mathcal{E}_1(n)_a \cap \mathcal{E}_2(n)_{\mu,\xi} \cap \mathcal{E}_3(n)_\zeta$.
Enfin, pour tout $k \in \mathbb{N}$, définissons l'événement

$$\Lambda^k_{a,\mu,\zeta,\xi} = (\cap_{n\geq t_k}\mathcal{E}(n)_{a,\mu,\zeta,\xi}) \cap \{\exists \text{ une v.a. } \alpha \text{ t.q. } \mathrm{supp}(\alpha) \subset]0,\xi[\text{ et } \lim_{n\to\infty} C_n/D_n = \alpha = \lim_{n\to\infty} \ln B_n/\ln C_n < \xi\}.$$

Heuristiquement, ce dernier événement correspond à ce que la marche aléatoire reste bloquée à droite de A (ou à gauche, lorsque les entiers A à G sont en ordre décroissant), le point B étant rarement visité en comparaison du point C, et la proportion de visites en C par rapport aux visites en D restant inférieure à la constante ζ.

L'événement $\mathcal{E}(n)_{a,\mu,\zeta,\xi}$ (resp. $\Lambda^k_{a,\mu,\zeta,\xi}$) est parfois noté $\mathcal{E}(n)^{A,\pm}_{a,\mu,\zeta,\xi}$ (resp. $\Lambda^{k,A,\pm}_{a,\mu,\zeta,\xi}$) dans le but de désigner, lorsqu'il y a ambiguïté, les choix de A et de l'ordre sur les entiers A à G (+ si croissant, − si décroissant). De façon similaire, nous désignons parfois l'instant t_k par t_k^C.

Lemme 6.2.1 *Supposons $a \in \mathbb{N}^*$, μ_0, $\mu > 1$, ζ, $\xi \in]0,1[$ fixés tels que $\mu > \mu_0$ et $\xi > \zeta$. Alors, pour tous $\nu < \min(1/2, 1-\xi)$ et $c \in]0,1[$, il existe $k_0 \in \mathbb{N}$ ne dépendant que de a, μ_0, μ, ζ, ξ, ν et c tel que pour tout $k \geq k_0$,*

$$P(\Lambda^k_{a,\mu,\zeta+k^{-\nu},\xi}|\mathcal{F}_{t_k}) \geq c1_{\mathcal{E}(t_k)_{a,\mu_0,\zeta,\xi}}.$$

Les preuves des théorèmes 1 et 3 reposent pour l'essentiel sur le lemme 1, dont nous donnons la démonstration en annexe.

PREUVE DU THÉORÈME 6.1.1 : Montrons que R a p.s une borne inférieure. La situation étant équivalente en inversant l'ordre des entiers, on en déduira de la même manière que R a p.s une borne supérieure.

Il nous suffit de prouver qu'il existe une constante $\delta > 0$ telle que, pour tout $v \in \mathbb{Z}$ tel que $v \leq v_0$ (point de départ de la marche),

$$P(v-2 \notin R \mid v \in R) \geq \delta,$$

car cela implique, pour tout $k \in \mathbb{N}$,

$$P(v_0 - 2k \in R) \leq (1-\delta)^k.$$

Fixons $C = v$ (et les points A à G correspondants dans l'ordre croissant), $a = 1$, $\mu_0 = 1$, $\mu = 2$, $\zeta = 1/2$, $\xi = 3/4$, $\nu = 1/8$ et $c = 1/2$, et choisissons $k_0 \in \mathbb{N}$ (indépendant de v) tel que les conditions du lemme 6.2.1 soient satisfaites pour ce choix de variables a, μ_0, μ, ζ, ξ, ν et c. Observons que

$$P(\mathcal{E}(t_{k_0})_{1,1,1/2,3/4} \mid v \in R) \geq 1/6^{k_0-1}.$$

En effet, à la première visite en C, les points A et B à gauche de C n'ont jamais été visités par la marche ; donc la probabilité qu'il y ait au moins $k_0 - 1$ visites en D avant de retourner en C est supérieure à $1/2 \times 1/3^{k_0-1}$, et la probabilité qu'il n'y ait ensuite pas de visite en B jusqu'à la (k_0-1)-ième visite en C est supérieure à $1/2^{k_0-2}$. D'autre part, le lemme 6.2.1 implique que

$$P(v-2 \notin R \mid \mathcal{F}_{t_{k_0}}) \geq P((\cap_{n \geq t_{k_0}} \mathcal{E}_1(n)_a | \mathcal{F}_{t_{k_0}}) \geq P(\Lambda^{k_0}_{1,2,1/2+k_0^{-\nu},3/4} \mid \mathcal{F}_{t_{k_0}}) \geq 1_{\mathcal{E}(t_{k_0})_{1,1,1/2,3/4}}/2.$$

En résumé, $P(v-2 \notin R \mid v \in R) \geq 1/6^{k_0}$.

□

PREUVE DU THÉORÈME 6.1.2 : Notons, pour tout $A \in \mathbb{Z}$,

$$\Omega_A = \{\inf R' = A+1\}.$$

Nous savons, d'après le théorème 6.1.1, que

$$P(\cup_{A \in \mathbb{Z}} \Omega_A) = P(\inf R' > -\infty) = 1.$$

Choisissons $A \in \mathbb{Z}$ et appelons B, C, D, E et F les entiers qui suivent A par ordre croissant. Nous utilisons le lemme suivant, énoncé initialement dans [7]. Nous voulons montrer que F est visité un nombre infini de fois sur Ω_A.

Lemme 6.2.2

$$\Omega_A \subset \left\{\exists \alpha \in]0,1]/ \ \frac{E_n}{C_n+E_n} \to \alpha\right\} \subset \left\{\liminf \frac{E_n}{D_n} > 0\right\} \ p.s.$$

PREUVE: En appliquant deux fois le lemme de Borel-Cantelli conditionnel (voir par exemple [2]), nous pouvons montrer que

$$\Omega_A \subset \left\{\sum \frac{1_{X_i=B,X_{i+1}=A}}{A_i} < \infty\right\} = \left\{\sum \frac{1_{X_i=B}}{A_i+C_i} < \infty\right\} = \left\{\sum \frac{1_{X_i=B,X_{i+1}=C}}{C_i} < \infty\right\} \tag{6.1}$$

Ce qui signifie qu'il y a "peu" de visites de B vers C sur Ω_A. D'autre part, il est facile de vérifier que le processus $(W_n)_{n\in\mathbb{N}}$ défini par

$$W_n = \ln(C_n+E_n) - \ln(E_n-1) - \sum_{i \leq n-1} \frac{1}{C_i+E_i} 1_{\{X_i=B,X_{i+1}=C\}}$$

est une sur-martingale (la preuve est laissée au lecteur). Par conséquent, en définissant le temps d'arrêt

$$T_M = \inf\left\{n \in \mathbb{N}^* / \sum_{i \leq n-1} \frac{1}{C_i+E_i} 1_{\{X_i=B,X_{i+1}=C\}} \geq M\right\},$$

le processus $W_{n \wedge T_M}$ est une sur-martingale minorée par $-M-1$, et donc converge p.s. Ce qui permet de conclure que $C_n/(C_n+E_n)$ a une limite finie sur Ω_A puisque, d'après 6.1, il existe alors p.s $M > 0$ tel que $T_M = \infty$.

□

Le lemme 2 implique, en appliquant à nouveau le lemme de Borel-Cantelli conditionnel et en utilisant le fait que $F_n/(D_n+F_n) \geq 1/(D_n+1)$,

$$\Omega_A \subset \Big\{\sum \frac{1_{X_n=E}F_n}{D_n+F_n} \geq \sum \frac{1_{X_n=E}}{D_n+1} = \infty\Big\} \subset \Big\{\sum 1_{X_n=E,X_{n+1}=F} = \infty\Big\} \subset \Big\{F_\infty = \infty\Big\}.$$

□

PREUVE DU THÉORÈME 6.1.3 : Nous pouvons supposer sans perte de généralité $I =]\alpha_*, \alpha^*[$ avec α_*, $\alpha^* \in]0,1[$, $\alpha^* > \alpha_*$, et $v \geq v_0$.

Notons $A = v-3$, et B à G les entiers qui suivent A par ordre croissant.

Le principe de cette preuve est d'appliquer le lemme 6.2.1 d'une part sur les points A à G avec $\xi = \alpha^*$, et d'autre part sur les points G à A avec $\xi = 1-\alpha_*$.

Soient ζ_1 et ζ_2 tels que $\alpha^* > \zeta_1 > 1-\zeta_2 > \alpha_*$. Notons, pour tous $k_0 \in \mathbb{N}^*$ et $m \in \mathbb{N}$,

$$\Gamma_m = \{X_m = D,\ \min(C_m, D_m, E_m) \geq k_0,\ A_m = 2,\ B_m \leq 3,\ F_m = G_m = 1,$$
$$C_m/D_m \leq \zeta_1 - k_0^{-1},\ E_m/D_m \leq \zeta_2 - k_0^{-1}[\}.$$

Alors, d'une part, pour tout $m \in \mathbb{N}$,

$$\Gamma_m \subset \mathcal{E}(t^C_{C_m+1})^{A,+}_{2,3,\zeta_1,\alpha^*} \cap \mathcal{E}(t^E_{E_m+1})^{G,-}_{1,1,\zeta_2,1-\alpha_*}$$

et d'autre part, si l'on suppose k_0 assez grand pour que $1-(\zeta_2 - k_0^{-1}) < \zeta_1 - k_0^{-1}$, il existe $m \in \mathbb{N}$ tel que $P(\Gamma_m) > 0$.

Nous obtenons, en appliquant le lemme 6.2.1, que

$$P(\Lambda^{C_m+1,A,+}_{2,4,\zeta_1+k_0^{-\nu},\alpha^*} \mid \mathcal{F}_m) \geq 3/4.1_{\Gamma_m}\ ,\ P(\Lambda^{E_m+1,G,-}_{1,2,\zeta_2+k_0^{-\nu},1-\alpha_*} \mid \mathcal{F}_m) \geq 3/4.1_{\Gamma_m},$$

si $\nu < \min(1/2, 1-\alpha^*, \alpha_*)$ et k_0 est supposé supérieur à une constante ne dépendant que de ν, ζ_1, ζ_2, α_* et α^*.

Par conséquent, en notant

$$\Delta_m = \Lambda^{C_m+1,A,+}_{2,4,\zeta_1+k_0^{-\nu},\alpha^*} \cap \Lambda^{E_m+1,G,-}_{1,2,\zeta_2+k_0^{-\nu},1-\alpha_*},$$

nous obtenons que

$$P(\Delta_m) \geq P(\Gamma_m)/2 > 0.$$

Plaçons nous sur Δ_m. Alors, pour $n \geq m$, $B_n \leq 4D_n^{\alpha^*}$, $F_n \leq 2D_n^{1-\alpha_*}$, $D_n \leq C_n + E_n \leq D_n + B_n + F_n \leq D_n(1+6D_n^{-\min(1-\alpha^*,\alpha_*)})$, et les événements **(i)** à **(vi)** sont bien réalisés.

□

6.3 Annexe : preuve du lemme 1

Nous nous plaçons dorénavant sur $\mathcal{E}(t_k)_{a,\mu_0,\zeta,\xi}$. En particulier, pour tout $n \geq t_k$, les probabilités conditionnelles sachant $\mathcal{F}_n$ considérées sont implicitement restreintes à $\mathcal{E}(t_k)_{a,\mu_0,\zeta,\xi}$. Nous supposons dans toute la suite que la constante k_0 a été choisie suffisamment grande, en fonction des autres constantes a, μ_0, μ, ζ, ξ, ν et c.

Nous notons $\zeta' = \zeta + k^{-\nu}$ (nous supposons k_0 assez grand pour que $\zeta' < \xi$), et définissons les temps d'arrêt

$$\begin{aligned} T_1 &= \inf\{n \geq t_k / \ A_n > a\}, \\ T_2 &= \inf\{n \geq t_k / \ B_n > \mu C_n^{\xi}\}, \\ T_3 &= \inf\{n \geq t_k / \ C_n/D_n > \zeta'\}, \end{aligned}$$

ainsi que $T = T_1 \wedge T_2 \wedge T_3$.

Considérons les ensembles de probabilité

$$\Omega_1 = \{T_1 \leq T_2 \wedge T_3\}, \ \Omega_2 = \{T_2 \leq T_1 \wedge T_3\}, \ \Omega_3 = \{T_3 \leq T_1 \wedge T_2\}.$$

Nous allons montrer que, pour tout $c \in]0,1[$, il existe k_0 tel que, pour tout $k \geq k_0$, $P(\Omega_i \cap \{T < \infty\} \mid \mathcal{F}_{t_k}) \leq (1-c)/3$ pour $i \in \{1,2,3\}$. Ce qui permettra de conclure car $\Omega_1 \cup \Omega_2 \cup \Omega_3 = \Omega$.

Définissons, pour tout $n \geq 2$,

$$\theta_n = \inf\{m \in \mathbb{N} / \ C_m + D_m = n\} \wedge T, \ \kappa_n = t_n \wedge T.$$

La preuve est divisée en trois étapes, notées **1)**, **2)** et **3)**.

1) Démonstration de $P(\Omega_1 \cap \{T < \infty\} \mid \mathcal{F}_{t_k}) \leq (1-c)/3$.

Il nous suffit de prouver que, avec une grande probabilité, il n'y a pas de visite en A venant de C jusqu'à l'instant T.

Supposons $n \geq k \geq k_0$ et $t_n < T$. Alors

$$P(X_{t_n+2} = A \mid \mathcal{F}_{t_n}) = \frac{B_{t_n}}{B_{t_n} + D_{t_n}} \frac{A_{t_n}}{A_{t_n} + C_{t_n}} \leq \frac{B_{t_n}}{C_{t_n}} \frac{A_{t_n}}{C_{t_n}} \leq \frac{\mu n^{\xi}}{n} \frac{a}{n} \leq \frac{a\mu}{n^{2-\xi}}.$$

Par conséquent

$$P(\Omega_1 \cap \{T < \infty\} \mid \mathcal{F}_{t_k}) \leq \sum_{n \geq k} \frac{a\mu}{n^{2-\xi}} \leq \frac{a\mu}{(1-\xi)} \frac{1}{(k-1)^{1-\xi}} \leq \frac{c}{3}.$$

2) Démonstration de $P(\Omega_2 \cap \{T < \infty\} \mid \mathcal{F}_{t_k}) \leq (1-c)/3$.

Soit $n \geq k \geq k_0$ tel que $t_n < T$.

Nous voulons contrôler l'évolution de B_{t_n}/n^{ξ} : d'une part, si A n'est pas visité en $t_n + 2$, B est visité au plus une fois entre les instants t_n et t_{n+1}, et d'autre part, si $t_n < T$,

$$P(X_{t_n+1} = B \mid \mathcal{F}_{t_n}) = \frac{B_{t_n}}{B_{t_n} + D_{t_n}} \leq \frac{B_{t_n}}{C_{t_n}} \frac{C_{t_n}}{D_{t_n}} \leq \frac{\zeta' B_{t_n}}{n},$$

ce qui implique
$$E(B_{\kappa_{n+1}} \mid \mathcal{F}_{t_n}) \leq B_{t_n}(1+\zeta'/n).$$

Donc le processus $(R_n)_{n\in\mathbb{N}}$, défini pour tout $n \in \mathbb{N}$ par $R_n = B_{\kappa_n}/n^{\xi}$, est une sur-martingale :

$$E(R_{n+1} - R_n \mid \mathcal{F}_{t_n}) \leq \frac{B_{t_n}}{n^{\xi}} \frac{1+\zeta'/n}{(1+1/n)^{\xi}} - \frac{B_{t_n}}{n^{\xi}} \leq 0$$

car $\zeta < \xi$ implique $1+\zeta'/n \leq (1+1/n)^{\xi}$ si k_0 a été choisi assez grand.

Notons enfin, pour tout $n \in \mathbb{N}$, $\tilde{R}_{n+1} = R_{n+1} - E(R_{n+1} \mid \mathcal{F}_{t_n})$. Alors

$$E(\tilde{R}_{n+1}^2 \mid \mathcal{F}_{t_n}) \leq \frac{1}{(n+1)^{2\xi}} P(X_{t_n+1} = B \mid \mathcal{F}_{t_n}) \leq \frac{1}{(n+1)^{2\xi}} \frac{\zeta' B_{t_n}}{n} \leq \frac{\mu\xi}{n^{1+\xi}}$$

en utilisant le fait que $B_{t_n} \leq \mu n^{\xi}$.

Donc, en utilisant l'inégalité de Doob,

$$E(\sup_{n\geq k}(\sum_{i=k}^{n} \tilde{R}_{i+1})^2) \leq 4\sum_{n=k}^{\infty} \frac{\mu\xi}{n^{1+\xi}} \leq \frac{4\mu}{(k-1)^{\xi}}.$$

Notons $\Omega_2' = \{\sup_{n\geq k} \sum_{i=k}^{n} \tilde{R}_{i+1} > \mu - \mu_0\}$, alors $P(\Omega_2' \mid \mathcal{F}_{t_k}) \leq (1-c)/3$. Sur $\Omega_2'^{c}$, pour tout $n \geq t_k$ tel que $n \leq T$,

$$R_n \leq R_k + \mu - \mu_0 \leq \mu,$$

ce qui permet de conclure.

Remarquons enfin que le raisonnement utilisé au début de cette étape permet de prouver également que, sur $\{T = \infty\} \cap \{C_n/D_n \to \alpha < \xi\}$, $\ln B_n/\ln C_n$ converge vers α. En effet, pour tout $\beta > \alpha$, $S_n = B_{\kappa_n}/n^{\beta}$ converge puisque S_n est alors une sur-martingale à partir d'un certain rang. Et, de façon similaire, si $\alpha \neq 0$, $T_n = n/B_{\kappa_n}^{\gamma}$ converge pour tout $\gamma > 1/\alpha$.

3) Démonstration de $P(\Omega_3 \cap \{T < \infty\} \mid \mathcal{F}_{t_k}) \leq (1-c)/3$.

Le principe de cette partie est d'observer que, en moyenne, C_n/D_n n'augmente que lors des visites de B à C, puisqu'à chaque visite en D la probabilité d'aller en C est $C_n/(C_n + E_n)$ et est donc strictement inférieure à C_n/D_n.

Plus précisément, le processus $(V_n)_{n\in\mathbb{N}}$ défini pour tout $n \in \mathbb{N}$ par

$$V_n = \frac{C_{\theta_n}}{D_{\theta_n - 1}} - \sum_{i\leq n} \frac{B_{\theta_{i-1}} 1_{X_{\theta_{i-1}}=C}}{(B_{\theta_{i-1}} + D_{\theta_{i-1}}) D_{\theta_{i-1}-1}}$$

est une sur-martingale.

D'autre part, notons $\tilde{k} = C_{t_k} + D_{t_k} \geq k$ et, pour tout $n \in \mathbb{N}$, $\tilde{V}_{n+1} = V_{n+1} - E(V_{n+1} \mid \mathcal{F}_{\theta_n})$. Alors, si $n \geq \tilde{k}$ et $\theta_n < T$,

$$E(\tilde{V}_{n+1}^2 \mid \mathcal{F}_{\theta_n}) \leq \frac{C_{\theta_n}}{(C_{\theta_n} + E_{\theta_n}) D_{\theta_n}^2} 1_{X_{\theta_n}=D} + \frac{B_{\theta_n}}{(B_{\theta_n} + D_{\theta_n}) D_{\theta_n}^2} 1_{X_{\theta_n}=C} \leq \frac{2}{D_{\theta_n}^2} \leq \frac{2(1+\xi)^2}{n^2}.$$

Donc en appliquant l'inégalité de Doob,

$$E(\sup_{n\geq\tilde{k}}(\sum_{i=\tilde{k}}^{n}\tilde{V}_{i+1})^2 \mid \mathcal{F}_{t_k}) \leq \frac{8(1+\xi)^2}{k-1}.$$

Notons $\Omega'_3 = \{\sup_{n\geq\tilde{k}}\sum_{i=\tilde{k}}^{n}\tilde{V}_{i+1} > k^{-\nu}/3\}$, alors $P(\Omega'_3 \mid \mathcal{F}_{t_k}) \leq (1-c)/3$ en supposant k_0 assez grand, puisque l'on a fait l'hypothèse $\nu < 1/2$. Sur $\Omega_3'^{c}$, pour tout $n \geq \tilde{k}$ tel que $\theta_n \leq T$,

$$\begin{aligned}\frac{C_{\theta_n}}{D_{\theta_n}} &\leq \frac{C_{\theta_n}}{D_{\theta_n - 1}} \leq \frac{C_{t_k}}{D_{t_k - 1}} + \frac{k^{-\nu}}{3} + \sum_{\tilde{k}\leq i\leq n-1}\frac{B_{\theta_i}}{(B_{\theta_i}+D_{\theta_i})D_{\theta_i - 1}} \\ &\leq \zeta + \frac{\zeta}{k-1} + \frac{k^{-\nu}}{3} + \sum_{i\geq k}\frac{\mu(1+\xi)^2}{(i-1)i}i^{\xi} \leq \zeta + k^{-\nu}\end{aligned}$$

puisque l'on a supposé $\nu < 1-\xi$.

Remarquons enfin que, sous les hypothèses précédentes, V_n est une sur-martingale minorée par $-k^{-\nu}/3$, et converge donc p.s. Par conséquent C_n/D_n converge p.s sur $\{T=\infty\}$.

Remerciements. Je tiens à remercier mon directeur de thèse Michel Benaïm pour ses précieux conseils au cours de la rédaction de ce travail.

Bibliographie

[1] M. Benaïm. Vertex-reinforced random walks and a conjecture of Pemantle. *Annals of probability*, 25 :No.1,361–392, 1997.

[2] D. Dacunha-Castelle et M. Duflo. *Probabilités et statistiques*, volume 2. Masson, 1993.

[3] R. Pemantle. Random processes with reinforcement. *Massachussets Institute of Technology doctoral dissertation*, 1988.

[4] R. Pemantle. Vertex-reinforced random walk. *Probability Theory and Related Fields*, 92 :117–136, 1992.

[5] R. Pemantle and S. Volkov. Vertex-reinforced random walk on $\mathbb{Z}$ has finite range. *Annals of probability*, 27 :No.3,1368–1388, 1999.

[6] B. Skyrms and R. Pemantle. A dynamic model of social network formation. *Proc. NAS*, 97 :9340–9346, 2000.

[7] P. Tarrès. Vertex-reinforced random walk on $\mathbb{Z}$ eventually gets stuck at a set of five points. *Preprint*, 2001.

[8] S. Volkov. Vertex-reinforced random walk on arbitrary graphs. *Annals of probability*, 29 :No.1,66–91, 2001.

Chapitre 7

Detailed summary of the article "VRRW on $\mathbb{Z}$ eventually gets stuck at a set of five points"

7.1 Proposition 8.3.1

The beginning of the proof is devoted to estimates on the number of visits to the smallest points infinitely visited by the random walk. More precisely, we let B denote the smallest point visited infinitely often, $A = B - 1$, $C = B + 1$, and D, E, F, G... be the points following C.

We define, for all $n \in \mathbb{N} \cup \{+\infty\}$ and for all A, B ...$\in \mathbb{Z}$, A_n, B_n ... the number of times A (resp. B ...) has been visited at time n.

The two following results are claimed in order to show that there exists an $\alpha \in [0,1[$ such that $C_n/(C_n + E_n) \to \alpha$, and that if $\alpha > 0$, then there exists $\beta \in [0,1[$ such that $G_n/(G_n + I_n) \to \beta$.

Let us define, for all successive numbers L and M ($M = L \pm 1$),

$$\Upsilon_{L,M} = \{\sum_{k \geq 1} \frac{1_{\{X_{k-1}=L, X_k=M\}}}{M_{k-1}} < +\infty\}.$$

Lemma 8.2.3 *Let K, L and M be three successive numbers, increasing or decreasing. Then $\Upsilon_{L,K} = \Upsilon_{L,M}$.*

PROOF: Using conditional Borel-Cantelli lemma,

$$\Upsilon_{L,M} = \{\sum_{k \geq 1} \frac{1_{\{X_{k-1}=L, X_k=M\}}}{M_{k-1}} < +\infty\} = \{\sum_{k \geq 1} \frac{1_{X_{k-1}=L}}{K_{k-1} + M_{k-1}} < +\infty\} = \Upsilon_{L,K}.$$

$\square$

Proposition 8.3.1 *Let Q, R, S, T, U be five successive numbers, increasing or decreasing. Then*

a)

$$Y_n = \ln(R_n + T_n) - \ln(T_n - 1) - \sum_{k=1}^{n} \frac{1}{R_{k-1} + T_{k-1}} 1_{\{X_{k-1}=Q, X_k=R\}}$$

and

$$\begin{aligned} Z_n = \ln(R_n + T_n) - \ln(T_n - 1) - \quad & \sum_{k=1}^{n} \frac{1}{R_{k-1} + T_{k-1}} 1_{\{X_{k-1}=Q, X_k=R\}} \\ & + \sum_{k=1}^{n} \frac{R_{k-1}}{(R_{k-1} + T_{k-1}) T_{k-1}} 1_{\{X_{k-1}=U, X_k=T\}} \end{aligned}$$

are supermartingales.

b)

$$\begin{aligned} \Upsilon_{Q,R} \quad & \subset \{\frac{R_n}{R_n + T_n} \to \alpha \in [0,1[\} \\ \Upsilon_{Q,R} \cap \{\alpha > 0\} \quad & \subset \Upsilon_{U,T}. \end{aligned}$$

PROOF: The first statement of **b)** follows from the first statement of **a)**, since

$$\begin{aligned}
\Upsilon_{Q,R} \quad & \subset \{\sum_{k=1}^{n} \frac{1}{R_{k-1}+T_{k-1}} 1_{\{X_{k-1}=Q, X_k=R\}} < +\infty\} \subset \{Y_n \text{ bounded below }\} \\
& \subset \{Y_n \text{ converges }\} \subset \{\frac{R_n+T_n}{T_n} \text{ converges }\} \subset \{\frac{R_n}{R_n+T_n} \to \alpha \in [0,1[\},
\end{aligned}$$

The second statement of **b)** follows from the second statement of **a)**, since

$$\begin{aligned}
\Upsilon_{Q,R} \cap \{\alpha > 0\} \quad & \subset \{\sum_{k=1}^{n} \frac{1_{\{X_{k-1}=Q, X_k=R\}}}{R_{k-1}+T_{k-1}} < +\infty\} \cap \{\alpha > 0\} \\
& \subset \{Z_n \text{ converges }\} \cap \{\alpha > 0\} \cap \{\sum_{k=1}^{n} \frac{1_{\{X_{k-1}=Q, X_k=R\}}}{R_{k-1}+T_{k-1}} < +\infty\} \\
& \subset \{\sum_{k=1}^{n} \frac{R_{k-1}}{(R_{k-1}+T_{k-1})T_{k-1}} 1_{\{X_{k-1}=U, X_k=T\}} < +\infty\} \cap \{\alpha > 0\} \\
& \subset \{\sum_{k=1}^{n} \frac{1_{\{X_{k-1}=U, X_k=T\}}}{T_{k-1}} < +\infty\} = \Upsilon_{U,T}.
\end{aligned}$$

$\square$

Let us define the set R' of points visited infinitely often and, for all $A \in \mathbb{Z}$,

$$\Omega_A = \{A = \inf R' - 1\}.$$

Let B, C... be the integers following A. Lemma 8.2.3 and proposition 8.3.1 imply that

$$\Omega_A \subset \Upsilon_{B,A} = \Upsilon_{B,C} \subset \{\exists \alpha \in [0,1[/ \ \frac{C_n}{C_n+E_n} \to \alpha \in [0,1[\}$$

and

$$\Omega_A \cap \{\alpha > 0\} \subset \Upsilon_{B,C} \cap \{\alpha > 0\} \subset \Upsilon_{F,E} = \Upsilon_{F,G} \subset \{\frac{G_n}{G_n+I_n} \to \beta \in [0,1[\}.$$

7.2 Proof of lemma 8.6.2

We now need to study the case $\alpha = 0$, and similarly the case $\beta = 0$. In fact, we will prove that $\alpha > 0$ a.s on Ω_A, and that nevertheless the case $\beta = 0$ is a priori possible but implies $H_n/(H_n + J_n) \to \beta' > 0$.

We begin with the proof of the following lemma.

Lemma 8.6.2 *Let B, C, D, E and F be five consecutive points. Then*

$$\Upsilon_{B,C} \cap \{\lim \frac{C_n}{C_n + E_n} > 0\} \subset \{\liminf \frac{F_n}{D_n + F_n} > 0\}.$$

The heuristic underlying this lemma is that the case

$$\Gamma_0 = \Upsilon_{B,C} \cap \{\lim \frac{C_n}{C_n + E_n} = 0\} \cap \{\liminf \frac{F_n}{D_n + F_n} = 0\}$$

can be compared to an unstable equilibria. More precisely, it is natural to think that there is a "competition" between the numbers of visits to points B and F. We will prove on one hand that if we start from a configuration where F is more visited than B, then the number of visits to F with respect to the number of visits to E will tend to increase, and eventually $\liminf E_n/D_n > 1$ with a reasonnable probability, which leads to a contradiction. On the other hand, if for $\mu > 1$, F_n remains smaller than μC_n for large n, then the noise inherent to the behavior of the random walk implies that eventually one of the terms D_n and E_n becomes more important, i.e $\liminf E_n/D_n \neq 1$.

Henceforth assume we belong to Γ_0. Let us define, for all $n \in \mathbb{N}$,

$$p_n = \frac{C_n}{C_n + E_n} \ , \ \tilde{p}_n = \max(\frac{C_n}{C_n + E_n}, \frac{F_n}{D_n + F_n}).$$

The first result we need to prove is that $\tilde{p}_n$ remains enough large to generate a stable noise.

Lemma 8.3.8 *Let L, M, N and O be four consecutive integers. Then, for all $b < 1/2$, there exists almost surely $n \in \mathbb{N}$ such that, for all $k \in \mathbb{N}$ such that $M_p + N_p > n$,*

$$L_k + O_k \geq b \frac{M_k + N_k}{\ln(L_k + O_k)}.$$

This lemma implies that, for enough large $n \in \mathbb{N}$,

$$\tilde{p}_n \geq \frac{1}{4 \ln D_n}.$$

Indeed, for enough large $n \in \mathbb{N}$,

$$C_n + F_n \geq \frac{3}{8} \frac{D_n + E_n}{\ln(C_n + F_n)}.$$

Now, either $\tilde{p}_n \geq 1/5$ or ($4C_n \leq E_n$ and $4F_n \leq D_n$). Similarly, either $\tilde{p}_n \geq 1/3$ or $D_n \in [2E_n/3, 3E_n/2]$ (observe for instance that $D_n \leq C_n + E_n \leq 3E_n/2$ when $\tilde{p}_n \leq 1/3$).

This implies in particular that, either $\tilde{p}_n \geq 1/5$ or $2\max(C_n, F_n) \leq \max(D_n, E_n)/2 \leq \min(D_n, E_n)$.

Assume for instance $C_n \geq F_n$. Then

$$\frac{C_n}{C_n+E_n} \geq \frac{3}{16}\frac{1}{\ln(2C_n)}\frac{D_n+E_n}{C_n+E_n} \geq \frac{3}{16}\frac{1}{\ln D_n}\frac{5E_n/3}{5E_n/4} \geq \frac{1}{4\ln D_n}.$$

The other case is similar.

First, we apply this underestimate of $\tilde{p}_n$ to bound the term $C_n/(C_n+E_n)$ for enough large n. Afterwards we estimate the variations of $\ln E_n/D_n$, and use these estimates to show for all $\mu > 1$ that $F_n \leq \mu C_n$ for enough large n. Then we show that this last case is a.s impossible.

First, let us show that, given $\mu > 1$, there exists a.s $i_0 \in \mathbb{N}$ such that for all n, $j_0 \in \mathbb{N}$ with $n \geq j_0 \geq i_0$, p_n remains smaller than $\mu\tilde{p}_{j_0}$.

The result is a consequence of the following lemma.

Lemma 8.3.3 *Let, for all $n \in \mathbb{N}$,*

$$\Pi_n = \prod_{1\leq k\leq n}(1+\frac{1_{\{X_{k-1}=B, X_k=C\}}}{C_k}),\ \ M_n = \frac{C_n}{\Pi_n(C_n+E_n)},$$

with the convention that $\Pi_0 = 1$.

Then, for all $\nu < 1/2$,

1) *M_n is a nonnegative supermartingale.*

2) $\sup_{n\geq j_0}(M_n - M_{j_0}) \leq o(1/(C_{j_0}+E_{j_0})^\nu)$.

In order to simplify the summary, the stopping times T_0, T_1, T_2 and T_3 defined here are not the same as those defined in the article.

Given $\epsilon > 0$, define the stopping time

$$T_0 = \inf\{n \geq i_0/\ \sum_{k=i_0}^{n}\frac{1_{\{X_{k-1}=B, X_k=C\}}}{C_k} > \epsilon\} \wedge \inf\{n \geq i_0/\ \tilde{p}_n < \frac{1}{4\ln D_n}\}.$$

For all $\nu \in]0, 1/2[$, we obtain that for $n < T_0$ and $j_0 \in [i_0, n]$, if we assume D_{j_0} enough large,

$$\frac{p_n}{\Pi_n} \leq \frac{p_{j_0}}{\Pi_{j_0}} + \frac{1}{(C_{j_0}+E_{j_0})^\nu},$$

which implies

$$p_n \leq \tilde{p}_{j_0}\frac{\Pi_n}{\Pi_{j_0}} + \frac{\Pi_n}{D_{j_0}^\nu} < \mu\tilde{p}_{j_0}$$

for enough small ϵ and enough large D_{j_0}, since $\tilde{p}_{j_0} \geq 1/4\ln D_{j_0}$ and $\Pi_n < \exp(\sum_{k\leq n}\frac{1_{\{X_{k-1}=B, X_k=C\}}}{C_k}) \leq e^\epsilon$.

We now define stopping time

$$T_1 = T_0 \wedge \inf\{n \geq i_0/\ \exists j_0 \in [i_0, n]/\ p_n > \mu\tilde{p}_{j_0}\}.$$

In order to estimate the variations of $\ln E_n/D_n$ we define, for all $j_0 \in \mathbb{N}$ such that $X_{j_0} = D$, a sequence α_n as follows :

$$\begin{cases} \alpha_0 = j_0 \\ \forall n \in \mathbb{N},\ \ \alpha_{n+1} = \inf\{k > \alpha_n /\ X_k = D \text{ and } X_{k-1} = E\} \end{cases}$$

and our goal is to give an estimate of

$$\ln \frac{E_{\alpha_{n+1}}}{D_{\alpha_{n+1}}} - \ln \frac{E_{\alpha_n}}{D_{\alpha_n}}.$$

At each visit to point D, the probability to go to E is $E_k/(C_k + E_k)$ at time k. Similarly at each visit to point E, the probability to go to D is $D_k/(D_k + F_k)$ at time k.

First assume $D_k/(D_k+F_k)$ is near $D_{\alpha_n}/(D_{\alpha_n}+F_{\alpha_n})$ for all $k \in [\alpha_n, \alpha_{n+1}]$, the law of the number of visits to E between times α_n and α_{n+1} is near the law of first success with probability $D_{\alpha_n}/(D_{\alpha_n}+F_{\alpha_n})$, and therefore has an expectation of nearly $(D_{\alpha_n} + F_{\alpha_n})/D_{\alpha_n}$. The similar remark is true for the number of visits to D between time $[\alpha_n, \alpha_{n+1}]$. Under this assumption, we obtain

$$\begin{aligned} E(\ln \frac{E_{\alpha_{n+1}}}{D_{\alpha_{n+1}}} - \ln \frac{E_{\alpha_n}}{D_{\alpha_n}} \mid \mathcal{F}_{\alpha_n}) \quad &\approx \frac{1}{E_{\alpha_n}} \frac{D_{\alpha_n} + F_{\alpha_n}}{D_{\alpha_n}} - \frac{1}{D_{\alpha_n}} \frac{C_{\alpha_n} + E_{\alpha_n}}{E_{\alpha_n}} \\ &\approx \frac{D_{\alpha_n} + F_{\alpha_n} - (C_{\alpha_n} + E_{\alpha_n})}{D_{\alpha_n} E_{\alpha_n}}. \end{aligned}$$

Unfortunately, we have to deal with the possibility that $F_k/(D_k + F_k)$ or $C_k/(C_k + E_k)$ change significantly between time α_n and α_{n+1}.

To this end we define, given $q \in]0,1[$ and $j_0 \geq i_0$, a stopping time

$$T_3 = T_1 \wedge \inf\{n \geq j_0 /\ \tilde{p}_n > q\},$$

and construct for instance a random variable V (given by lemma 8.5.4) whose expectation and variance may be estimated by

$$E(V \mid \mathcal{F}_{\alpha_n}) = E(\frac{C_{\alpha_{n+1} \wedge T_3} + E_{\alpha_n}}{E_{\alpha_n}} \mid \mathcal{F}_{\alpha_n})\ ,\ \ V(V \mid \mathcal{F}_{\alpha_n}) \leq 5q$$

and such that $V = D_{\alpha_{n+1}} - D_{\alpha_n}$ when $\alpha_{n+1} < T_3$.

Using this tool, we prove the following lemma.

Let us define, for all $k \in \mathbb{N}$,

$$R_k = D_k + F_k - (C_k + E_k).$$

Lemma 8.6.3 *Let $j_0 \in \mathbb{N}$ be such that $X_{j_0} = D$; let $q \in]0, Cst[$. Then there* *exist random variables* $(\tilde{D}_n)_{n\in\mathbb{N}}$, $(\tilde{E}_n)_{n\in\mathbb{N}}$, $(\epsilon_n)_{n\in\mathbb{N}}$, $(r_n)_{n\in\mathbb{N}}$ *adapted to the filtration* $(\mathcal{F}_{\alpha_n})_{n\in\mathbb{N}}$ *such that, if* $\alpha_n < T_3$*, then*

$$\mathbf{1)}\tilde{D}_n = D_{\alpha_n} \ , \ \tilde{E}_n = E_{\alpha_n}$$

$$\mathbf{2)}\ln\frac{\tilde{E}_{n+1}}{\tilde{D}_{n+1}} - \ln\frac{\tilde{E}_n}{\tilde{D}_n} = \frac{R_{\alpha_{n+1}\wedge T_3}}{D_{\alpha_n}E_{\alpha_n}} + \epsilon_{n+1} + r_n,$$

$$\mathbf{3)}E(\epsilon_{n+1} \mid \mathcal{F}_{\alpha_n}) = 0 \ , \ E(\epsilon_{n+1}^2 \mid \mathcal{F}_{\alpha_n}) \leq \frac{Cst.q}{(n+D_{j_0})^2} \ , \ |r_n| \leq \frac{Cst}{(n+D_{j_0})^2}.$$

A first consequence of this result is the following lemma.
Given $i_0 \in \mathbb{N}$ and $\epsilon > 0$, let

$$\Gamma_{1,i_0,\epsilon} = \{\tilde{p}_{i_0} \leq \epsilon/2 \text{ and } \min(D_{i_0}, E_{i_0}) \geq 1/\epsilon^4\}.$$

Let us define the stopping time

$$T_2 = T_1 \wedge \inf\{n \geq i_0 /\ F_n \geq \mu C_n\}.$$

Lemma 8.6.5 $\Gamma_0 \subset \cup_{i_0\in\mathbb{N}}(\{T_2 = +\infty\} \cap \Gamma_{1,i_0,\epsilon})$.

As seen above, the proof of this lemma relies on the fact that, each time $F_n > \mu C_n$ for a given constant μ, the probability that $\liminf_{k\to\infty} E_k/D_k > 0$ is greater than a constant. This fact needs technical arguments, since it has to deal with cases where F is too large in comparison with C_n, which complicates the estimates of

$$\ln\frac{E_{\alpha_{n+1}}}{D_{\alpha_{n+1}}} - \ln\frac{E_{\alpha_n}}{D_{\alpha_n}}.$$

This proof also needs an estimate of a sum involving the number of visits from B to C between times j_0 and α_n, given $j_0 \in \mathbb{N}$ (recall that $(\alpha_n)_{n\in\mathbb{N}}$ depends on j_0 since $\alpha_0 = j_0$). This corresponds to lemma 8.6.4, claimed hereafter.

Let us define, for all successive integers B and C ($C = B \pm 1$) and i, $j \in \mathbb{N}$ such that $j \geq i$,

$$(BC)_{i,j} = \sum_{k=i+1}^{j} 1_{\{X_{k-1}=B, X_k=C\}}.$$

Lemma 8.6.4 *Let us define, for all* $r \in \mathbb{N}$*, the stopping time*

$$W_r = \inf\{n \geq r /\ s_n > \lambda\}.$$

Suppose $\mu \in]1, Cst[$, $\lambda \in]0, Cst[$ *and* $\epsilon \in]0, Cst(\mu)[$*. Then, on* $\Gamma_{1,i_0,\epsilon}$, $\alpha_{n+1} < T_1 \wedge W_{\alpha_k}$ *implies*

$$\sum_{j=k}^{n} \frac{(BC)_{\alpha_k,\alpha_{j+1}}}{D_{\alpha_j}E_{\alpha_j}} \leq 2\tilde{p}_{\alpha_k}\epsilon.$$

Next we prove that, when $F_n \leq \mu C_n$ for sufficiently large n, the conclusion $\lim E_n/D_n = 1$ (since $C_n/(C_n + E_n) \to 0$ on Γ_0) holds with probability 0.

The idea behind this last part of the proof of lemma 8.6.2 is to make use of the arguments developed in the proofs of non-convergence towards repulsive traps.

If we consider a stochastic algorithm $(x_n)_{n\in\mathbb{N}}$ taking values in $\mathbb{R}$, whose behavior can be written as

$$x_{n+1} - x_n = y_n + c_{n+1}(\epsilon_{n+1} + r_{n+1})$$

and satisfying the following conditions, given a filtration $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$,

1) $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$, $(\epsilon_n)_{n\in\mathbb{N}}$ and $(r_n)_{n\in\mathbb{N}}$ are $\mathbb{F}$-adapted random variables taking values in $\mathbb{R}$, $(\gamma_n)_{n\in\mathbb{N}}$ and $(c_n)_{n\in\mathbb{N}}$ are deterministic sequences taking values in $\mathbb{R}_+$, $(c_n)_{n\in\mathbb{N}}$ having an infinite number of positive terms,

2) for all $n \in \mathbb{N}$, x_n and y_n are of the same sign,

3) $E(\epsilon_{n+1} \mid \mathcal{F}_n) = 0$, $\liminf_{n\to\infty} E(\epsilon_{n+1}^2 \mid \mathcal{F}_n) > 0$ and there exists $a > 2$ such that $\limsup_{n\to\infty} E(|\epsilon_{n+1}|^a \mid \mathcal{F}_n) < +\infty$, $\sum r_n^2 < +\infty$,

then the trap $x = 0$ is avoided with probability 1 (see theorem 3.2.2, chapter 3), i.e $P(\lim_{n\to\infty} x_n = 0) = 0$.

In the case of VRRW, we can write that

$$x_{n+1} - x_n = y_{n+1} + \epsilon_{n+1} + r_n,$$

where

$$x_n = \ln \frac{\tilde{E}_n}{\tilde{D}_n} \ , \ y_{n+1} = \frac{R_{\alpha_{n+1}\wedge T_3}}{D_{\alpha_n} E_{\alpha_n}},$$

and ϵ_{n+1}, r_{n+1} are given by lemma 8.6.3.

The problem here is that we can't say that x_n and y_{n+1} are of the same sign.

However, we can observe that, roughly, the term $R_{\alpha_{n+1}}$ only increases with the visits from G to F and decreases with the visits from B to C.

Heuristically, we therefore can say that, when x_n "tends" to increase (resp. to decrease) the random walk tends to go "more to the right" (resp. "to the left"), which implies that y_n also "tends" to increase.

The precise tool behind these remarks is the definition of a partial order on the random walks. Lemma 8.5.1 claims the following result.

Assume we deal with two random walks $\mathcal{M}$ and $\mathcal{M}'$ such that at each point $j \in \mathbb{Z}$, for a same number of visits to j, if $\mathcal{M}'$ has more visited $j+1$ than $\mathcal{M}$ and less visited $j-1$, then $\mathcal{M}'$ has a greater probability than $\mathcal{M}$ to go right. Then we can couple $\mathcal{M}$ and $\mathcal{M}'$ such that for all $j \in \mathbb{Z}$, for a same number of visits to j, $\mathcal{M}'$ has more visited $j+1$ than $\mathcal{M}$ and less visited $j-1$. In this case, we write that $\mathcal{M}' \gg \mathcal{M}$.

It is easy to prove that, given two random walks $\mathcal{M}$ and $\mathcal{M}'$ such that $\mathcal{M}' \gg \mathcal{M}$, if we keep notations A_n, B_n, ..., and R_n, α_n, p_n, $\tilde{p}_n$ concerning $\mathcal{M}$, and make use of notations A'_n, B'_n, ... and R'_n, α'_n, p_n, $\tilde{p}'_n$ for the corresponding definitions with $\mathcal{M}'$, then, for all $n \in \mathbb{N}$,

$$R'_{\alpha'_n} \geq R_{\alpha_n}.$$

Having defined a partial order on random walks on $\mathbb{Z}$, we observe that the noise inherent to the behavior of $\ln E_{\alpha_n}/D_{\alpha_n}$ is related to the uncertainty on the visits from D to C during the time D_n remains of the order of D_{j_0}. Indeed, for all $n \geq j_0 \geq i_0$ such that $n < T_2$, $p_n \leq \mu \tilde{p}_{j_0}$ and $F_n \leq \mu C_n$, which implies $\tilde{p}_n \leq 2p_n \leq 2\mu\tilde{p}_{j_0} \leq 4\mu p_{j_0}$.

This latter remark gives the idea to define, concurrently to the VRRW called $\mathcal{M}$, a random walk $\mathcal{M}'$ with the same conditional probabilities as $\mathcal{M}$ starting from all points different from D, but such that, given g_2 and $g_3 > 0$,

$$\begin{cases} P(X'_{j+1} = C \mid X'_j = D, \mathcal{F}_j) = \overline{p'_j} & \text{if } j \in [j_0, \alpha_{\lfloor g_3 r \rfloor}] \\ P(X'_{j+1} = C \mid X'_j = D, \mathcal{F}_j) = p'_j & \text{otherwise,} \end{cases}$$

where $r = D_{j_0}$ and, for all $n \in \mathbb{N}$,

$$\overline{p_n} = p_n(1 - \frac{g_2}{\sqrt{2p_{j_0}D_{j_0}}}) \ , \ \overline{p'_n} = p'_n(1 - \frac{g_2}{\sqrt{2p_{j_0}D_{j_0}}}).$$

The advantage of such a definition is to generate a trend towards the right for $\mathcal{M}'$ in comparison with $\mathcal{M}$, enough significant to cover the noise and nevertheless not too large, so that the probabilities of a same group of paths for $\mathcal{M}$ and for $\mathcal{M}'$ are of the same order.

More precisely, this new random walk $\mathcal{M}'$ satisfies two properties. First, the behavior of $\ln(E'_{\alpha'_n}/D'_{\alpha'_n})$ can be written similarly as the corresponding behavior of $\ln(E_{\alpha_n}/D_{\alpha_n})$ but with an additionnal term for $n \leq \lfloor g_3 r \rfloor$ corresponding to the lower probability to go to C starting from D. Second, given any group of paths, we can precisely bound the probability for it to hold for $\mathcal{M}'$ given the corresponding probability for $\mathcal{M}$.

Statement There exists $(\mathcal{F}'_{\alpha'_n})_{n\in\mathbb{N}}$-adapted random variables $(\epsilon'_n)_{n\in\mathbb{N}}$, $(r'_n)_{n\in\mathbb{N}}$ such that for all $n \in \mathbb{N}$,

$$E(\epsilon'_{n+1} \mid \mathcal{F}_{\alpha_n}) = 0 \ , \ E(\epsilon'^2_{n+1} \mid \mathcal{F}'_{\alpha'_n}) \leq \frac{Cst.p_{j_0}}{(n + D_{j_0})^2} \ , \ |r'_n| \leq \frac{Cst}{(n + D_{j_0})^2},$$

and, when $\alpha'_{n+1} < T'_2$,

$$-\ln\frac{E'_{\alpha'_{n+1}}}{D'_{\alpha'_{n+1}}} - \ln\frac{E'_{\alpha'_n}}{D'_{\alpha'_n}} \geq \frac{R'_{\alpha'_{n+1}}}{D'_{\alpha'_n}E'_{\alpha'_n}} + \frac{Cst.g_2}{D_{j_0}}\sqrt{\frac{p_{j_0}}{D_{j_0}}} + \epsilon'_{n+1} + r'_n \text{ if } n \in [0, \lfloor g_3 D_{j_0} \rfloor - 1],$$

$$-\ln\frac{E'_{\alpha'_{n+1}}}{D'_{\alpha'_{n+1}}} - \ln\frac{E'_{\alpha'_n}}{D'_{\alpha'_n}} \geq \frac{R'_{\alpha'_{n+1}}}{D'_{\alpha'_n}E'_{\alpha'_n}} + \epsilon'_{n+1} + r'_n \text{ otherwise.}$$

Lemma 8.5.2 *For all $\tilde{\epsilon} > 0$ and for all mesurable group of paths $\mathcal{C}$,*

$$P_{\mathcal{M}'}(\mathcal{C} \mid \mathcal{F}_{j_0}) \leq M(g_2, \tilde{\epsilon}) P_{\mathcal{M}}(\mathcal{C} \mid \mathcal{F}_{j_0}) + \tilde{\epsilon},$$

$M(g_2, \tilde{\epsilon})$ being a constant depending only on g_2 and $\tilde{\epsilon}$.

(the lemma is stated with k_0 instead of j_0 in the proof).

These two lemmas will imply that, for all $j_0 \geq i_0$, the probability that $\liminf E_{\alpha_n}/D_{\alpha_n} \neq 1$ or $T_2 < +\infty$ conditionally to $\mathcal{F}_{j_0}$ is greater than a constant.

Indeed, taking g_2 enough large implies that the sum of the $\lfloor g_3 D_{j_0} \rfloor$ terms

$$\frac{Cst.g_2}{D_{j_0}} \sqrt{\frac{p_{j_0}}{D_{j_0}}}$$

is with large probability greater than the contributions of the noises (ϵ_n), (ϵ'_n) and of the terms (r_n), (r'_n), which ensures that $\liminf E'_{\alpha'_n}/D'_{\alpha'_n} > \liminf E_{\alpha_n}/D_{\alpha_n}$ with large probability. On the other hand, the paths followed by $\mathcal{M}'$ are followed by $\mathcal{M}$ with a significant probability(decreasing as g_2 increases).

More precisely, for enough large $c \in \mathbb{R}_*^+$, letting

$$\Delta = \{ \quad \sup_{k_1 \in \mathbb{N},\ n \geq k_1} \left| \sum_{k=k_1}^{n-1} \epsilon_{k+1} \right| \leq c \sqrt{\frac{p_{j_0}}{D_{j_0}}} \ , \ \sup_{k_1 \in \mathbb{N},\ n \geq k_1} \left| \sum_{k=k_1}^{n-1} \epsilon'_{k+1} \right| \leq c \sqrt{\frac{p_{j_0}}{D_{j_0}}} \ ,$$

we obtain by Doob's inequality that $P(\Delta^c \mid \mathcal{F}_{j_0}) \leq \tilde{\epsilon}$ if we choose $c = Cst/\sqrt{\tilde{\epsilon}}$ and, if we assume D_{j_0} large enough,

$$\sum_{k \in \mathbb{N}^*} |r_k| \leq c \sqrt{\frac{p_{j_0}}{D_{j_0}}} \ , \ \sum_{k \in \mathbb{N}^*} |r'_k| \leq c \sqrt{\frac{p_{j_0}}{D_{j_0}}}.$$

Define, for all $n \in \mathbb{N}$,

$$m_n = \ln \frac{E'_{\alpha'_n}/D'_{\alpha'_n}}{E_{\alpha_n}/D_{\alpha_n}}$$

and $\hat{m}_n = \sup_{k \in [0,n]} m_k$. Then, if we belong to Δ,

$$m_n \geq \sum_{k=0}^{n-1} \left(\frac{R'_{\alpha'_{k+1}}}{D'_{\alpha'_k} E'_{\alpha'_k}} - \frac{R_{\alpha_{k+1}}}{D_{\alpha_k} E_{\alpha_k}} \right) + Cst.g_2 g_3 \sqrt{\frac{p_{j_0}}{D_{j_0}}} - 4c \sqrt{\frac{p_{j_0}}{D_{j_0}}}.$$

Now, an easy calculation (detailed in the article) shows that

$$\sum_{k=0}^{n-1} \left(\frac{R'_{\alpha'_{k+1}}}{D'_{\alpha'_k} E'_{\alpha'_k}} - \frac{R_{\alpha_{k+1}}}{D_{\alpha_k} E_{\alpha_k}} \right) \geq -2\hat{m}_{n-1} \sum_{k=0}^{n-1} \frac{|R_{\alpha_{k+1}}|}{D_{\alpha_k} E_{\alpha_k}}$$

and that

$$\sum_{k \leq n-1} \frac{|R_{\alpha_{k+1}}|}{D_{\alpha_k} E_{\alpha_k}} \leq 8 p_{j_0}$$

for all $n \in \mathbb{N}$ such that $\alpha_n < T_2$ on Δ.

In summary, on Δ,

$$m_n \geq (Cst.g_2 g_3 - 4c - 96cp_{j_0})\sqrt{\frac{p_{j_0}}{D_{j_0}}} \geq c\sqrt{\frac{p_{j_0}}{D_{j_0}}}$$

if $\hat{m}_n \leq 6c\sqrt{p_{j_0}/D_{j_0}}$, taking $g_2 = Cst.c/g_3$. We similarly prove this inequality in the other case for $\hat{m}_n$.

In summary, we obtain that, on $\Gamma_0 \cap \Delta$,

$$\liminf E'_{\alpha'_n}/D'_{\alpha'_n} > 1$$

when $T_2 = \infty$ and $T'_2 = \infty$. We let $\mathcal{C}_1$ denote the group of paths followed by $\mathcal{M}'$ on $\Gamma_0 \cap \Delta$. Then $P_{\mathcal{M}'}(\mathcal{C}_1 \mid \mathcal{F}_{j_0}) \geq 1/2$ if we take $\tilde{\epsilon}$ enough small. This implies by lemma 8.5.2 that

$$P_{\mathcal{M}}(\mathcal{C} \mid \mathcal{F}_{j_0}) \geq M(g_2, \tilde{\epsilon})(1/2 - \tilde{\epsilon}).$$

Therefore the paths followed by $\mathcal{M}'$ on $\Gamma_0 \cap \Delta$ are also followed by $\mathcal{M}$ with a reasonable probability, and

$$\liminf E_{\alpha_n}/D_{\alpha_n} > 1$$

on these paths. This finishes the proof of lemma 8.6.2.

7.3 Proof of $\Omega_A \subset \Omega_{A,F} \cup \Omega_{A,1} \cup \Omega_{A,2}$

We now wish to deduce from lemma 8.6.2 that $\alpha = 0$ on Ω_A, and that $\beta' = \lim H_n/(H_n + J_n) \neq 0$ when $\beta = 0$. The first result we use to this end is that if we belong to the second event involved in lemma 8.6.2, i.e if $\liminf F_n/(D_n + F_n) > 0$ and if we also belong to $\Upsilon_{B,C}$, then there is a few number of visits from C to D, i.e we belong to $\Upsilon_{C,D}$. The second result we use in the proof of this fact is that if we belong, given three consecutive integers P, Q and R, to $\Upsilon_{P,Q} \cap \Upsilon_{R,Q}$, the fact that there is a few number of visits from P to Q and from R to Q is equivalent to the fact that there is a few visits to Q, i.e that Q is visited finitely often.

Corollary 8.3.2 *Let P, Q, R, S and T be five successive integers, increasing or decreasing. Then*

$$\Upsilon_{P,Q} \cap \{\limsup \frac{R_n}{R_n + T_n} < 1\} \subset \Upsilon_{Q,R}.$$

Lemma 8.2.4 *Let P, Q and R be three consecutive integers. Then*

$$\Upsilon_{P,Q} \cap \Upsilon_{R,Q} = \{Q_\infty < +\infty\}.$$

Let us now prove that these results imply that $\alpha > 0$ on Ω_A. Indeed, using lemma 8.6.2, corollary 8.3.2 and lemma 8.2.4,

$$\begin{aligned}\Omega_A \cap \{\alpha = 0\} \quad & \subset \Omega_A \cap \Upsilon_{B,C} \cap \{\liminf \frac{F_n}{D_n + F_n} > 0\} \\ & \subset \Omega_A \cap \Upsilon_{C,B} \subset \Upsilon_{A,B} \cap \Upsilon_{C,B} \cap \Omega_A = \emptyset \text{ a.s.}\end{aligned}$$

We now prove that $\beta' = 0$ when $\beta = 0$. First, β' exists in this case, using lemma 8.6.2 and corollary 8.3.2 :

$$\Upsilon_{F,G} \cap \{\lim \frac{G_n}{G_n + I_n} = 0\} \subset \Upsilon_{G,H} \subset \{\frac{H_n}{H_n + J_n} \to \beta' \in [0, 1[\}.$$

Second, applying again these two results,

$$\Upsilon_{G,H} \cap \{\lim \frac{H_n}{H_n + J_n} = 0\} \subset \Upsilon_{H,J} = \Upsilon_{H,G}.$$

Therefore

$$\begin{aligned}&\Omega_A \cap \Upsilon_{F,G} \cap \{\lim \frac{G_n}{G_n + I_n} = 0\} \cap \{\lim \frac{H_n}{H_n + J_n} = 0\} \subset \Omega_A \cap \Upsilon_{F,G} \cap \Upsilon_{H,G} \\ &\subset \Omega_A \cap \{G_\infty < +\infty\} \cap \{\lim \frac{H_n}{H_n + J_n} = 0\} = \emptyset \text{ a.s.}\end{aligned}$$

In summary, we have proved that

$$\Omega_A = \Omega_{A,F} \cup \Omega_{A,1} \cup \Omega_{A,2}$$

a.s where

$$\Omega_{A,F} = \Omega_A \cap \{G_\infty < \infty\},$$

$$\Omega_{A,1} = \Upsilon_{A,B} \cap \Upsilon_{B,C} \cap \Upsilon_{F,G} \cap \{\exists \alpha \in]0, 1[\ \exists \beta \in]0, 1[/ \frac{C_n}{C_n + E_n} \to \alpha \text{ and } \frac{G_n}{G_n + I_n} \to \beta\} \cap \{G_\infty = \infty\}$$

$$\begin{aligned}\Omega_{A,2} = & \{\exists \alpha \in]0, 1[,\ \exists \beta' \in]0, 1[/ \frac{C_n}{C_n + E_n} \to \alpha \text{ and } \frac{H_n}{H_n + J_n} \to \beta'\} \\ & \cap \Upsilon_{A,B} \cap \Upsilon_{B,C} \cap \Upsilon_{F,G} \cap \Upsilon_{G,H} \cap \{G_\infty = \infty\}.\end{aligned}$$

Let us now give some indications on the proofs of corollary 8.3.2 and lemma 8.2.4. Corollary 8.3.2 is a direct consequence of the lemma 8.3.6.

Lemma 8.3.6 *Let P, Q, R, S and T be five successive integers, increasing or decreasing. Then, for all $\gamma_- > 0$ and $\gamma_+ < 1$,*

(1)
$$\begin{aligned}&\Upsilon_{P,Q} \cap \{Q_\infty = \infty\} \cap \{\limsup \frac{R_n}{R_n + T_n} \le \gamma_+\} \\ &\subset \{\limsup \frac{\ln Q_n}{\ln R_n} \le \limsup \frac{\ln Q_n}{\ln S_n} \le \gamma_+\} \cap \{Q_\infty = \infty\} \\ &\subset \Upsilon_{Q,R} \cap \{R_\infty = S_\infty = \infty\} \ \textit{a.s}\end{aligned}$$

(2)
$$\Upsilon_{P,Q} \cap \{Q_\infty = \infty\} \cap \{\liminf \frac{R_n}{R_n + T_n} \ge \gamma_-\} \subset \{\liminf \frac{\ln Q_n}{\ln S_n} \ge \gamma_-\} \ \textit{a.s.}$$

Let us give a short proof of lemma 8.2.4 :

$$\begin{aligned}\Upsilon_{P,Q} \cap \Upsilon_{R,Q} &= \{\sum_{k\in\mathbb{N}^*} \frac{1_{\{X_{k-1}=P, X_k=Q\}}}{Q_k} < \infty\} \cap \{\sum_{k\in\mathbb{N}^*} \frac{1_{\{X_{k-1}=R, X_k=Q\}}}{Q_k} < \infty\} \\ &= \{\sum_{k\in\mathbb{N}^*} \frac{1_{\{X_k=Q\}}}{Q_k} < \infty\} = \{Q_\infty < \infty\}.\end{aligned}$$

7.4 Proofs of $P(\Omega_{A,1}) = 0$ and $P(\Omega_{A,2}) = 0$

7.4.1 Introduction

It remains to prove that $P(\Omega_{A,1}) = 0$ and $P(\Omega_{A,2}) = 0$.

These two proofs rely on the same principle. Let us first explain the case $\Omega_{A,1}$. Firstly we show, using ideas related to convergence with positive probability towards an attractor, that $H_n/D_n \not\to 1$ implies $H_n/D_n \to 0$ or $H_n/D_n \to \infty$. Secondly we also can estimate the rate of convergence towards 0 or ∞, which leads us to prove that in this case $\ln H_n/\ln D_n$ tends to 0 or to $+\infty$ as n tends to infinity. Thirdly we make use of this latter result to prove that the VRRW eventually gets stuck either between B and F or right hand from F, which leads to a contradiction.

The proof is similar on $\Omega_{A,2}$, if we take I_n in place of D_n. However, the case of $\Omega_{A,2}$ needs beforehand an argument showing that points B to L are all regularly visited. To this end, we need to discriminate between the case the random walk eventually gets stuck at a strict subset of $[B, L]$ and the case the frequency of visits at points B to L is enough large to ensure that we can estimate the variations of I_n/D_n. These two cases correspond respectively to $\{\beta' \leq \alpha\}$ and $\beta' > \alpha\}$.

Let us now claim first a lemma and its corollary giving conditions ensuring that the random walk eventually gets stuck left hand from a fixed point, and second a result of convergence towards an attractor suited to this part of the proof.

Let us endow the following notation. Given two nonnegative sequences $(u_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$, we write $u_n \asymp v_n$ iff either $\limsup u_n < +\infty$ and $\limsup v_n < +\infty$ or $u_n/v_n \to 1$. We write $u_n \succeq v_n$ iff $\liminf u_n/v_n \geq 1$.

Lemma 8.4.1 *Let L, M, N and O be four consecutive states. Let a_1, $a_2 > 0$ be two constants. Let $p_0 \in \mathbb{N}$ and $(r(n))_{n\geq p_0}$ be a process taking values in $]0,1[$ adapted to $(\mathcal{F}_n)_{n\in\mathbb{N}}$.*

Given m_0, p_0, $q_0 \in \mathbb{N}$ and $\nu > 0$, define the probability sets

$$\Pi^{N,O}_{0,p_0,a_1,d,r} = \{r(p_0) \leq 1-d \text{ and } \forall n \geq p_0,\ N_n \leq a_1 L_n^{r(n)} \text{ and, } \forall n \geq p,\ r(n) \leq r(p)+L_p^{-\nu}\},$$

$$\Pi^{N,O}_{1,q_0} = \{\forall n \geq q_0,\ X_n < O\}$$

and

$$\Pi^{N,O}_{2,m_0,a_2,r} = \{\exists i \geq m_0/ \ X_i = M \ \textit{and} \ O_i \leq a_2 M_i^{1-r(i)}\}.$$

Then

$$\Pi^{N,O}_{0,p_0,a_1,d,r} \cap (\cap_{m_0 \geq p_0} \Pi^{N,O}_{2,m_0,a_2,r}) \subset \cup_{q_0 \in \mathbb{N}} \Pi^{N,O}_{1,q_0}.$$

Corollary 8.4.1 *Let L, M, N and O be four consecutive states.*

$$\{\liminf_{n\to\infty} \frac{\ln O_n}{\ln M_n} + \limsup_{n\to\infty} \frac{\ln N_n}{\ln L_n} < 1\} \subset \{O_\infty < +\infty\}.$$

Lemma 8.4.2 *Let $(\mathcal{G}_n)_{n\in\mathbb{N}}$ be a filtration. Let $(\gamma_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$, $(\epsilon_n)_{n\in\mathbb{N}}$ and $(s_n)_{n\in\mathbb{N}}$ be sequences of $\mathcal{G}_n$-mesurable random variables, taking values in $\mathbb{R}^+$ (resp. $\mathbb{R}^+_*$, $\mathbb{R}$, $\mathbb{R}$), and such that*

$$\sum_{k\in\mathbb{N}} |s_k| < \infty \ \textit{and} \quad \sum_{k\in\mathbb{N}} E(\epsilon^2_{k+1} \mid \mathcal{F}_k) < \infty \ \textit{a.s.}$$

Let $\lambda > 1$ possibly infinite, and let $\mu \in [1, \sqrt{\lambda}]$.
Assume

$$\begin{cases} \ln y_{n+1} - \ln y_n = \gamma_n(y_n - 1) + \epsilon_{n+1} + s_n & \textit{if } y_n \in [\lambda^{-1}, \lambda] \\ y_{n+1} \geq \mu^{-1}\lambda & \textit{if } y_n \geq \lambda \\ y_{n+1} \leq \mu\lambda^{-1} & \textit{if } y_n \leq \lambda^{-1} \end{cases}$$

Then y_n enjoys the following properties.

1)

$$\{y_n \not\to 1\} \subset \{\liminf y_n > 1\} \cap \{\limsup y_n < 1\}.$$

2) *If $\lambda = \infty$, then*

$$\{y_n \not\to 1\} \subset \{\ln y_n \asymp \sum_{k=1}^n \gamma_k(y_k - 1)\}.$$

If $\lambda < \infty$ and

$$\sum_{k\in\mathbb{N}} \gamma_k = +\infty \ \textit{on} \ \bigcup_{n\in\mathbb{N}} \bigcap_{k\geq n} \{y_k \in [\lambda^{-1}, \lambda]\},$$

then

$$\{y_n \not\to 1\} \subset \{\liminf y_n \geq \mu^{-2}\lambda > 1\} \cup \{\limsup y_n \leq \mu^2\lambda^{-1} < 1\}.$$

3)

$$\{y_n \to 1\} \subset \{\sum_{k\in\mathbb{N}} \gamma_k |y_k - 1| < +\infty\}.$$

7.4.2 Proof $P(\Omega_{A,1}) = 0$

Let us define, for all $\mu > 1$, a, $b \in]0,1[$ and $i_0 \in \mathbb{N}$ the stopping time

$T_1 = \inf\{n \geq i_0/\ E_n/(C_n + E_n) \notin [\mu^{-1}a, \mu a]$ or $G_n/(G_n + I_n) \notin [\mu^{-1}b, \mu b]\}$.

Then

$$\Omega_{A,1} \subset \cup_{i_0\in\mathbb{N},a,b\in]0,1[}\{T_1 = +\infty\}.$$

Lemma 8.7.1

$$\Omega_{A,1} \subset \{\frac{B_n}{E_n}, \frac{F_n}{E_n}, \frac{F_n}{G_n}, \frac{J_n}{I_n} \to 0\} \cap \{\sum \frac{1_{X_n=F}}{F_n} \frac{B_n}{E_n + G_n} < \infty\} \cap \{\sum \frac{1_{X_n=F}}{F_n} \frac{J_n}{E_n + G_n} < \infty\}.$$

Let us define, for all $n \in \mathbb{N}$, the time κ_n of n-th return to state F. Let us define the stopping times

$$U_1 = \inf\{n \geq i_0/\ F_n/G_n > \epsilon \text{ or } F_n/E_n > \epsilon\},$$

and $T_2 = T_1 \wedge U_1$.

Lemma 8.7.2 *Let $a_1 < Cst(\mu, b)$. Assume $\mu \geq 1 + \epsilon$ and $\kappa_n \geq i_0$. Then*

$$P(\{G_{\kappa_{(a_1+1)n} \wedge T_2} > \mu G_{\kappa_n}\} \mid \mathcal{F}_{\kappa_n}) \leq \exp(-Cst.(\mu - 1)\mu^{-5}bn).$$

Let us define, given $i_0 \in \mathbb{N}$ and $a_1 < Cst(\mu, b)$, such that $X_{i_0} = F$, the stopping times

$$U_2 = \inf\{n \geq i_0/\ \exists V \in \{D, E, G, H\}/\ \exists k \in \mathbb{N}/\ \kappa_k \geq i_0 \text{ and } \kappa_{(1+a_1)k} \leq n \text{ and } V_{\kappa_{(1+a_1)k}} \geq \mu V_{\kappa_k}\}$$

and $T_3 = T_2 \wedge U_2$.

Then, for all $\mu > 1$ and $\epsilon < \mu - 1$,

$$\Omega_{A,1} \subset \cup_{i_0\in\mathbb{N},a,a_1,b\in]0,1[}\{T_3 = +\infty\}.$$

Lemma 8.7.3 *Let $i_0 \in \mathbb{N}$, a, $b \in]0,1[$, $\mu \in]1,2[$ and $\epsilon \in]0, \mu - 1[$. There exist sequences of $\mathcal{F}_{\kappa_n}$-mesurable random variables $(\tilde{D}_n)_{n\geq k_0}$, $(\tilde{H}_n)_{n\geq k_0}$, $(\epsilon_n)_{n\geq m}$, $(r_n)_{n\geq k_0}$ such that, if $\kappa_n \leq T_3$, then*

1) $\tilde{D}_n = D_{\kappa_n}, \tilde{H}_n = H_{\kappa_n}$,

2) $\ln \frac{\tilde{H}_{n+1}}{\tilde{D}_{n+1}} - \ln \frac{\tilde{H}_n}{\tilde{D}_n} = \frac{G_{\kappa_n} + I_{\kappa_n} - (C_{\kappa_n} + E_{\kappa_n})}{n(G_{\kappa_n} + E_{\kappa_n})} + \epsilon_{n+1} + r_n$,

3) $E(\epsilon_{n+1} \mid \mathcal{F}_{\kappa_n}) = 0$, $E(\epsilon^2_{n+1} \mid \mathcal{F}_{\kappa_n}) \leq Cst.\frac{a^{-2} + b^{-2}}{n^2}$, $|r_n| \leq Cst.\frac{a^{-2} + b^{-2}}{n^2}$.

Let, for all $k \in \mathbb{N}$,

$$R_k = G_k + I_k - (C_k + E_k).$$

Lemma 8.7.4

$$\Omega_{A,1} \subset \{\lim_{n\to\infty} H_n/D_n = 1\} \cap \{\sum_{n\in\mathbb{N}} \frac{|R_{\kappa_n}|}{n(E_{\kappa_n} + G_{\kappa_n})} < +\infty\}.$$

Sketch of the proof. We observe that, for all $k \in \mathbb{N}$, $H_k \leq G_k + I_k \leq F_k + H_k + J_k$ and $D_k \leq C_k + E_k \leq B_k + D_k + F_k$.

Now lemma 8.7.3 implies that, when $\kappa_n < T_3$,

$$\ln \frac{\tilde{H}_{n+1}}{\tilde{D}_{n+1}} - \ln \frac{\tilde{H}_n}{\tilde{D}_n} = \frac{\tilde{H}_n - \tilde{D}_n}{n(E_{\kappa_n} + G_{\kappa_n})} + \epsilon_{n+1} + r_n + \frac{S_{\kappa_n}}{n(E_{\kappa_n} + G_{\kappa_n})},$$

where

$$\sum \frac{|S_{\kappa_n}|}{n(E_{\kappa_n} + G_{\kappa_n})} < +\infty$$

by lemma 8.7.1.

The end of the proof relies on lemma 8.4.2 and corollary 8.4.1.

□

The proof of the non-convergence of H_n/D_n towards 1 is similar to the proof of lemma 8.6.2. For further details, see section 8.7.2.

7.4.3 Proof of $P(\Omega_{A,2}) = 0$

We first need martingales estimates for the VRRW, proved in subsection 8.3.2 of chapter 8.

Lemma 8.3.4 *On the event* $\{\limsup \ln G_n / \ln I_n < 1\}$*, there exists almost surely* $\nu > 0$ *such that*

$$\sup_{n\geq k}\left(\frac{H_n}{I_n} - \frac{H_k}{I_k}\right) \leq o\left(\frac{1}{I_k^\nu}\right).$$

Lemma 8.3.7

$$\Upsilon_{F,G} \cap \Upsilon_{L,K} \cap \{\frac{H_n}{H_n + J_n} \to \beta' \in]0, 1[\} \subset \{\ln G_n - \frac{H_n}{H_n + J_n} \ln I_n \textit{ converges}\}.$$

Let L, M and N be the integers following K. We assume, without loss of generality, that we belong to $\Upsilon_{L,K}$. Indeed, on $\Omega_{A,2}$, we belong to $\Upsilon_{F,G}$, $\Upsilon_{G,H}$ and

$\Upsilon_{K,L}$, which implies by lemma 8.6.2 and corollary 8.3.2 that either $\lim L_n/(L_n+N_n)>0$ or we belong to $\Upsilon_{L,K}$. The first case is impossible, since $P(\Omega_{F,1})=0$.

Let us define, for all $n\in\mathbb{N}$,

$$l(n)=\frac{E_n}{C_n+E_n}\ ,\ m(n)=\frac{H_n}{H_n+J_n}\ ,\ \tilde{l}(n)=\frac{E_n}{D_n}\ ,\ \tilde{m}(n)=\frac{H_n}{I_n}.$$

Let us define, given $\mu>1$, a, $a'\in\mathbb{R}_*^+$, l, $m\in]0,1[$, the stopping time

$$\begin{aligned}T_1=\inf\{n\geq i_0/\qquad & F_n\notin[\mu^{-1}aD_n^{l(n)}\mu aD_n^{l(n)}]\ \text{ or }\ G_n\notin[\mu^{-1}a'I_n^{m(n)},\mu aI_n^{m(n)}]\\ & \text{or } l(n),\tilde{l}(n)\notin[\mu^{-1}l,\mu l] \text{ or } m(n),\tilde{m}(n)\notin[\mu^{-1}m,\mu m]\\ & \text{or } \exists k\in[i_0,n]/\ |l(n)-l(k)|>D_k^{-\nu} \text{ or } |m(n)-m(k)|>I_k^{-\nu}\}.\end{aligned}$$

Then lemmas 8.3.4 and 8.3.7 imply that, for all $\mu>1$,

$$\Omega_{A,2}\subset\cup_{i_0\in\mathbb{N},a,a',\nu>0,l,m\in]0,1[}\{T_1=\infty\}.$$

The following lemma 8.8.1 implies that, on $\Omega_{A,2}$, points B to K are regularly visited by the random walk.

Lemma 8.8.1 $P(\beta'\leq\alpha\}\cap\Omega_{A,2})=0$.

Let $b\in]0,1/2[$. Define the stopping time

$$U_1=\inf\{n\geq i_0/\ l(n)+m(n)-1<b \text{ or } \tilde{l}(n)+\tilde{m}(n)-1<b\},$$

and $T_2=T_1\wedge U_1$. Then, for all $\mu>1$,

$$\Omega_{A,2}\subset\cup_{i_0\in\mathbb{N},a,a',\nu>0,b\in]0,1/2[,l,m\in]0,1[}\{T_2=+\infty\}.$$

We assume, without loss of generality, B_{i_0}, $\ldots K_{i_0}$ as large as necessary. We suppose i_0 is large enough, and take b smaller if necessary, so that, for $n\in[i_0,T_2[$, $l(n)\leq 1-b$, $F_n\leq D_n^{1-b}\leq\epsilon bD_n$, and similarly $m(n)\leq 1-b$, $G_n\leq I_n^{1-b}\leq\epsilon bI_n$.

Let us define, for all $n\in\mathbb{N}$, $v(n)=\mu^{-1}(l(n)+m(n)-1)$. Observe that, for all $n\in[i_0,T_2[$, $v(n)\geq\mu^{-1}b$.

Given $\lambda>\mu>1$, let us define two random sequences $(\xi_n)_{n\in\mathbb{N}}$, $(\tilde{\xi_n})_{n\in\mathbb{N}}$, as follows : $\xi_0=j_0$ and, for all $n\in\mathbb{N}$,

$$\begin{cases}\tilde{\xi_n}=\xi_n & \text{if } I_{\xi_n}\in[\lambda^{-1}D_{\xi_n},\lambda D_{\xi_n}]\\ \tilde{\xi_n}=\inf\{k>\xi_n/\ I_k\in[\lambda^{-1}D_k,\lambda D_k]\} & \text{otherwise,}\end{cases}$$

and

$$\xi_{n+1}=\inf\{k>\tilde{\xi_n}/\ X_k=G \text{ and } X_{k-1}=F\}$$

in all cases.

For all $k\in\mathbb{N}$, let

$$T_{+,k}=\inf\{n\geq k\ X_n=F\}\ ,\ T_{-,k}=\inf\{n\geq k/\ X_n=G\}.$$

Lemma 8.8.2 *If $I_k \le \lambda D_k$, then*

$$P(H_{T_{+,k}} > H_k + H_k^{1-v(k)} \mid \mathcal{F}_k) \le \exp[-\lambda^{-1}\mu^{-7} maa' I_k^{(\mu-1)b}].$$

We obtain similar inequalities if we replace H by G, I or J for T_+, and by C, D, E or F for T_-.

Therefore, let us define, for all $n \in \mathbb{N}$, the stopping time

$$\begin{aligned} U_2 = \inf\{n \ge i_0/\ \exists V \in \{G,H,I,J\}/\ \exists k \ge i_0/\ I_k \le \lambda D_k\ ,\ T_{+,k} \le i_0 \text{ and } V_{T_{+,k}} - V_k \ge V_k^{1-v(k)} \\ \exists V \in \{C,D,E,F\}/\ \exists k \ge i_0/\ D_k \le \lambda I_k\ ,\ T_{-,k} \le i_0 \text{ and } V_{T_{-,k}} - V_k \ge V_k^{1-v(k)}\}. \end{aligned}$$

Then, for all $\mu > 1$ and $\lambda > 1$,

$$\Omega_{A,2} \subset \cup_{i_0 \in \mathbb{N}, a, a', \nu > 0, b \in]0,1/2[, l, m \in]0,1[} \{T_3 = +\infty\},$$

where $T_3 = T_2 \wedge U_2$.

Remark that in particular, for all $V \in \{C,D,E,F,G,H,I,J\}$, for all $\xi_n < T_3$ such that $\xi_{n-1} \ge i_0$, then $V_{\xi_n} \le V_{\tilde{\xi}_{n-1}} + V_{\tilde{\xi}_{n-1}}^{1-v(\tilde{\xi}_{n-1})}$. A consequence is that $I_{\xi_n} \in [\lambda^{-1}\mu^{-1} D_{\xi_n}, \lambda\mu D_{\xi_n}]$ as long as $\xi_n < T_3$.

Lemma 8.8.3 *There exists a constant $e > 0$, depending only on b, λ, μ, l and m such that*

$$\frac{F_{\xi_n} G_{\xi_n}}{D_{\xi_n}} \succeq en1_{\xi_n < T_3}.$$

Moreover,

$$\sum_{\xi_n < T_3} F_{\xi_n}^{-1} < \infty\ ,\ \sum_{\xi_n < T_3} G_{\xi_n}^{-1} < \infty\ ,\ \sum_{\xi_n < T_3} \frac{K_{\xi_n}}{F_{\xi_n} G_{\xi_n}} < \infty\ ,\ \sum_{\xi_n < T_3} \frac{B_{\xi_n}}{F_{\xi_n} G_{\xi_n}} < \infty.$$

Let us define the stopping times

$$U_3 = \inf\{n \ge i_0/\ F_{\xi_n} G_{\xi_n} / D_{\xi_n} < en\}\ ,\ T_4 = U_3 \wedge T_3.$$

Let, for all $k \in \mathbb{N}$,

$$R_k = H_k + J_k - (C_k + E_k).$$

Lemma 8.8.4 *Let $i_0 \in \mathbb{N}$, a, a', $\nu > 0$, $b \in]0,1/2[$, l, $m \in]0,1[$. There exist sequences of $\mathcal{F}_{\xi_n}$-mesurable random variables $(\tilde{D}_n)_{n \ge k_0}$, $(\tilde{I}_n)_{n \ge k_0}$, $(\epsilon_n)_{n \ge m}$, $(r_n)_{n \ge k_0}$ such that, if $\xi_n \le T_4$, then*

1) $\tilde{D}_n = D_{\xi_n}, \tilde{I}_n = I_{\xi_n}$,

2) $\ln \dfrac{\tilde{I}_{n+1}}{\tilde{D}_{n+1}} - \ln \dfrac{\tilde{I}_n}{\tilde{D}_n} = \dfrac{R_{\xi_n}}{F_{\xi_n} G_{\xi_n}} + \epsilon_{n+1} + r_n$ *if* $I_{\xi_n} \in [\lambda^{-1} D_{\xi_n}, \lambda D_{\xi_n}]$,

3) $E(\epsilon_{n+1} \mid \mathcal{F}_{\kappa_n}) = 0\ ,\ E(\epsilon_{n+1}^2 \mid \mathcal{F}_{\kappa_n}) \le Cst(\lambda).\left(\dfrac{D_{\xi_n}}{F_{\xi_n} G_{\xi_n}}\right)^2\ ,\ |r_n| \le Cst(\lambda).\left(\dfrac{D_{\xi_n}}{F_{\xi_n} G_{\xi_n}}\right)^2.$

Lemma 8.8.5

$$\Omega_{A,2} \subset \{I_{\xi_n}/D_{\xi_n} \to 1\} \cap \{\sum_{n\in\mathbb{N}} \frac{|R_{\xi_n}|}{F_{\xi_n} G_{\xi_n}} < +\infty\}.$$

The proof of the non-convergence of I_n/D_n toward 1 is similar to the proof of lemma 8.6.2. For further details, see section 8.8.2.

Chapitre 8

VRRW on $\mathbb{Z}$ eventually gets stuck at a set of five points

Abstract : Vertex-Reinforced Random Walk (VRRW), defined by Pemantle (1988,[7]), is a random process taking values in the vertex set of a graph G, which is more likely to visit vertices it has visited before. Pemantle and Volkov (1997,[9]) considered the case where the underlying graph is the one-dimensional integer lattice $\mathbb{Z}$. They proved that the range is almost surely finite, and that with positive probability the range contains exactly five points. They conjectured that this second event holds with probability one ; the proof of this conjecture is the main purpose of this paper.

8.1 General introduction

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let G be a locally finite graph, $\sim$ be its neighbor relation, and $V(G)$ be its vertex set.

Let $(X_n)_{n\in\mathbb{N}}$ be a process taking values in $V(G)$. Let $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$ denote the filtration generated by the process, i.e $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$ for all $n \in \mathbb{N}$.

For any $v \in V(G)$, let $Z(n, v)$ be the number of times plus one that the process visits site v up through time $n \in \mathbb{N} \cup \{\infty\}$, i.e $Z(n, v) = 1 + \sum_{i=0}^{n} 1_{X_i = v}$.

Then $(X_n)_{n\in\mathbb{N}}$ is called Vertex-Reinforced Random Walk (VRRW) with starting point $v \in V(G)$ if $X_0 = v$ and, for all $n \in \mathbb{N}$,

$$P(X_{n+1} = x \mid \mathcal{F}_n) = 1_{x\sim X_n} \frac{Z(n, x)}{\sum_{w\sim X_n} Z(n, w)}.$$

In other words, moves are restricted to the edges of G, with the probability of a move to a neighbor y being proportionnal to the augmented occupation $Z(n, y)$ of y at that time.

The reinforced random walks have been introduced in 1988 by Pemantle ([7]), in the spirit of the seminal work of Coppersmith and Diaconis ([4]) who defined the notion of edge-reinforced random walks, having at each step a probability to move along an edge proportional to the number of times plus one that the process has visited this edge.

Reinforced processes are useful in models involving self-organization and learning behavior ; they can also describe spatial monopolistic competition in economics. For more details on applications and known results in connection with these models, we refer the reader to Pemantle and Volkov ([8],[9]).

Reinforced vertex random walks on finite graphs, with reinforcements weighted by factors associated to each edge of the graph, have been studied in Pemantle (1992,[8]) and Benaïm (1997,[1]).

Pemantle and Volkov achieved in 1997 results on reinforced random walks on $\mathbb{Z}$ ([9]), which are described in the following. Recently, Volkov ([11]) generalized

[1]AMS 1991 subject classification. Primary 60G17 ; secondary 34F05, 60G50.

Keywords : Vertex-reinforced random walks, urn model, random perturbations of dynamical systems, repulsive traps.

some of these results to a fairly broad class of locally finite graphs (containing the graphs of bounded degree) by proving that, when the weights associated to vertices are 1 (which is the case in this paper), the VRRW has finite range with positive probability.

The remainder of this paper is devoted to vertex-reinforced random walks on $\mathbb{Z}$.

Define $R = \{v \in \mathbb{Z} / \ \exists n \in \mathbb{N},\ X_n = v\}$, $R' = \{v \in \mathbb{Z} / \ X_n = v$ infinitely often$\}$ and, for $k \in \mathbb{Z}$ and $\alpha \in]0,1[$, the six events :

(i) $R' = \{k-2, k-1, k, k+1, k+2\}$;

(ii) $\ln Z(n, k-2)/\ln n \to \alpha$;

(iii) $\ln Z(n, k+2)/\ln n \to 1-\alpha$;

(iv) $Z(n, k-1)/n \to \alpha/2$;

(v) $Z(n, k+1)/n \to (1-\alpha)/2$;

(vi) $Z(n, k)/n \to 1/2$.

Pemantle and Volkov ([9]), proved the following results.

Theorem 8.1.1 $P(|R| = 5) > 0$ *and* $P(|R| < +\infty) = 1$.

Theorem 8.1.2 $P(|R'| \leq 4) = 0$.

Theorem 8.1.3 *For any open set* $I \subset [0,1]$ *and any integer* $k \in \mathbb{Z}$*, there exists with positive probability* $\alpha \in I$ *such that the events (i) to (vi) occur.*

They also proposed the following conjecture.

Conjecture 8.1.1 *There exist almost surely* $k \in \mathbb{Z}$ *and* $\alpha \in]0,1[$ *such that the events (i) to (vi) occur.*

The main purpose of the present article is the proof of this conjecture. In fact, we prove the following result which is a little more accurate. Define the two events

(ii') $\exists C_1 \in \mathbb{R}_+^* / \ Z(n, k-2)/n^\alpha \to C_1$;

(iii') $\exists C_2 \in \mathbb{R}_+^* / \ Z(n, k+2)/n^{1-\alpha} \to C_2$.

Theorem 8.1.4 *There exist almost surely* $k \in \mathbb{Z}$ *and* $\alpha \in]0,1[$ *such that the events (i), (ii'), (iii'), and (iv) to (vi) occur.*

We have used some results of Pemantle and Volkov ([9], [11]) : we point it out in the concerned parts. We have also used a result of Benaïm about convergence with positive probability towards an attractor.

8.2 Introduction to the proof

8.2.1 Notations

For a sequence (γ_n) of $\mathbb{F}$-adapted random variables (γ_n is $\mathcal{F}_n$-mesurable for all $n \in \mathbb{N}$), let $\mathbb{F}_{(\gamma_n)_{n\in\mathbb{N}}} = (\mathcal{F}_{\gamma_n})_{n\in\mathbb{N}}$ denote the filtration defined as follows : for all $n \in \mathbb{N}$, $A \in \mathcal{F}_{\gamma_n} \iff \forall q \in \mathbb{N},\ A \cap \{\gamma_n \leq q\} \in \mathcal{F}_q$.

Let us denote, for all $x \in \mathbb{R}$, $x^+ = \max(x, 0)$ and $x^- = -\max(-x, 0)$.

The notation Cst is used to denote an arbitrary universal constant ; similarly, we write $Cst(a_1, a_2, \ldots, a_p)$ for a constant depending only on a_1, a_2, $\ldots$ a_p.

We say for simplicity that a property holds for $x < Cst(a_1, \ldots, a_p)$ (resp. for $x > Cst(a_1, \ldots, a_p)$) when there exists a constant c, which depends only on a_1, $\ldots$ a_p so that this property holds for $x < c$ (resp. for $x > c$). The same remark holds when we first say we assume $x < Cst(a_1, \ldots, a_p)$ (resp. $x > Cst(a_1, \ldots, a_p)$) and further consider properties involving x.

Given sequences $(u_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$, we write $u_n = O(v_n)$ when, for all $n \in \mathbb{N}$, $|u_n| \leq Cst.|v_n|$; and $u_n = o(v_n)$ when $\lim_{n\to\infty} u_n/v_n = 0$, with the convention that $u_n/v_n = 0$ when $u_n = v_n = 0$. If $(u_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$ are taking values in $\mathbb{R}_*^+$, we write $u_n \asymp v_n$ iff either $\limsup u_n < +\infty$ and $\limsup v_n < +\infty$ or $u_n/v_n \to 1$, and $u_n \succeq v_n$ iff $\liminf u_n/v_n \geq 1$. We write $u_n \equiv v_n$ iff $\limsup |u_n - v_n| < +\infty$.

Given u, $v \in \mathbb{R}_*^+ \cup \{\infty\}$, we write $u \approx v$ iff $u = v = \infty$, or $u < \infty$ and $v < \infty$.

We also write, to simplify notation, $f = g_1 + \ldots + g_k + o(\triangleleft)$ whenever $f = g_1 + \ldots + g_k + o(g_k)$.

We write for simplicity, for all integer $I \in \mathbb{Z}$, for all m, $n \in \mathbb{N} \cup \{\infty\}$, $I_n = Z(n, I)$, and $I_{m,n} = Z(n, I) - Z(m, I)$.

We also write, given integers I, J and K such that $J = I \pm 1$, $K = J \pm 1$, for all m, $n \in \mathbb{N}$,

$$(IJ)_n = \sum_{k=1}^{n} 1_{X_{k-1}=I, X_k=J},\ (IJ)_{m,n} = (IJ)_n - (IJ)_m,$$

$$(IJK)_n = \sum_{k=1}^{n} 1_{X_{k-2}=I, X_{k-1}=J, X_k=K},\ (IJK)_{m,n} = (IJK)_n - (IJK)_m,$$

and given two integers P and Q not necessarily successive,

$$(P/Q)(n) = P(\sup\{k \in \mathbb{N} / \ Q_k = n\}).$$

We will make use of the following conditional Borel-Cantelli lemma at various steps of the proof.

Lemma 8.2.1 *Let $\mathbb{F} = (\mathcal{F}_n)_{n\geq 0}$ be a filtration, and consider a sequence $(\xi_n)_{n\geq 1}$ of $\mathbb{F}$-adapted random variables.*

Assume that the r.v $(\xi_n)_{n\in\mathbb{N}}$ are nonnegative and bounded by a constant C (for instance $\xi_n = 1_{\Gamma_n}$, with $\Gamma_n \in \mathcal{F}_n$). Then

$$\{\sum \xi_n < +\infty\} = \{\sum E(\xi_n \mid \mathcal{F}_{n-1}) < +\infty\} \ a.s.$$

8.2.2 Introduction to the proof

Let us define, for all $A \in \mathbb{Z}$,

$$\Omega_A = \{A = \inf R' - 1\}.$$

Theorem 8.1.1(Pemantle and Volkov) implies that

$$\Omega \subset \cup_{A\in\mathbb{Z}}\Omega_A.$$

Therefore proving theorem 8.1.4 is equivalent to showing that, for all $A \in \mathbb{Z}$, conditions (i), (ii'), (iii'), and (iv) to (vi) are satisfied on Ω_A, for $k = A+3$.

From now on we fix $A \in \mathbb{Z}$, and let B, C, D, E, F, G and so on until letter K in alphabetical order denote the integers following A.

We first observe that there is a few visits from B to C on Ω_A.

Let us define, for all successive numbers L and M ($M = L \pm 1$),

$$\Upsilon_{L,M} = \{\sum_{k\geq 1} \frac{1_{\{X_{k-1}=L, X_k=M\}}}{M_{k-1}} < +\infty\}.$$

The event $\Upsilon_{L,M}$ corresponds to the fact that there is a few number of visits from L to M. We mention in the two following lemmas a few properties satisfied on $\Upsilon_{L,M}$.

Lemma 8.2.2 *Let L, M, and N be three successive integers, increasing or decreasing. Suppose we belong to $\Upsilon_{L,M} \cap \{M_\infty = \infty\}$. Then*

$$\sum_{k=1}^{n} 1_{X_{k-1}=L, X_k=M} = o(M_n) \ , \ \sum_{k=1}^{n} \frac{1_{X_{k-1}=N, X_k=M}}{M_{k-1}} \equiv \ln M_n.$$

PROOF: The first estimate follows from Kronecker lemma. The second estimate follows from

$$\ln M_n \equiv \sum_{k=1}^{n} \frac{1_{X_k=M}}{M_n} = \sum_{k=1}^{n} \frac{1_{X_{k-1}=L, X_k=M}}{M_n} + \sum_{k=1}^{n} \frac{1_{X_{k-1}=N, X_k=M}}{M_n}.$$

□

Lemma 8.2.3 *Let K, L and M be three successive integers, increasing or decreasing. Then $\Upsilon_{L,K} = \Upsilon_{L,M}$.*

PROOF: Using conditional Borel-Cantelli lemma,

$$\Upsilon_{L,M} = \{\sum_{k\geq 1} \frac{1_{\{X_{k-1}=L, X_k=M\}}}{M_{k-1}} < +\infty\} = \{\sum_{k\geq 1} \frac{1_{X_{k-1}=L}}{K_{k-1}+M_{k-1}} < +\infty\} = \Upsilon_{L,K}.$$

□

The first important result of the proof is Proposition 8.3.1, which implies that

$$\Omega_A \subset \Upsilon_{B,A} = \Upsilon_{B,C} \subset \{\exists \alpha \in [0,1[/ \ \frac{C_n}{C_n+E_n} \to \alpha\},$$

and

$$\Upsilon_{B,C} \cap \{0 < \alpha < 1\} \subset \Upsilon_{F,E} = \Upsilon_{F,G} \subset \{\exists \beta \in [0,1[/ \ \frac{G_n}{G_n+I_n} \to \beta\}.$$

This means that, when there is a few visits coming from B to C, then the visits from B have a negligible influence on the behavior of $C_n/(C_n+E_n)$, which therefore tends to decrease. Conversely, if there is a few visits from B to C and $C_n/(C_n+E_n)$ converges towards a positive number, then there is a few visits from F to E.

The next goal is the study of the case $\alpha = 0$, and similarly the case $\beta = 0$. This study relies on lemma 8.6.1, which claims that, for all consecutive integers P, Q, R and S,

$$\Upsilon_{P,Q} \cap \{\lim \frac{Q_n}{Q_n+S_n} = 0\} \subset \Upsilon_{Q,R} = \Upsilon_{Q,P}.$$

This means that when $Q_n/(Q_n+S_n)$ tends to 0 on $\Upsilon_{P,Q}$, then Q becomes negligible with respect to R.

This gives us a tool to analyse the case $\alpha = 0$, since it implies that there is a few number of visits from C to B (we belong to $\Upsilon_{C,B}$). Since there is also a few number of visits from A to B (we belong to $\Omega_A \subset \Upsilon_{A,B}$), we can deduce from it that there is a few number of visits to B, i.e that B is visited finitely often.

More precisely, let us state the following lemma.

Lemma 8.2.4 *Let P, Q and R be three consecutive integers. Then*

$$\Upsilon_{P,Q} \cap \Upsilon_{R,Q} = \{Q_\infty < +\infty\}.$$

PROOF:

$$\begin{aligned} \Upsilon_{P,Q} \cap \Upsilon_{R,Q} \quad &= \{\sum_{k\in\mathbb{N}^*} \frac{1_{\{X_{k-1}=P, X_k=Q\}}}{Q_{k-1}} < +\infty\} \cap \{\sum_{k\in\mathbb{N}^*} \frac{1_{\{X_{k-1}=R, X_k=Q\}}}{Q_{k-1}} < +\infty\} \\ &= \{\sum_{k\in\mathbb{N}^*} \frac{1_{\{X_{k-1}=Q\}}}{Q_{k-1}} < +\infty\} = \{Q_\infty < +\infty\}. \end{aligned}$$

□

In summary, it follows from lemmas 8.6.1 and 8.2.4 that

$$\Omega_A \cap \{\alpha = 0\} \subset \Omega_A \cap \Upsilon_{C,B} \subset \Upsilon_{A,B} \cap \Upsilon_{C,B} = \emptyset \text{ a.s.}$$

Let us now consider the case $\beta = 0$. In this case, $\beta' = \lim H_n/(H_n+J_n)$ exists. Indeed, applying lemmas 8.6.1 :

$$\Upsilon_{F,G} \cap \{\lim \frac{G_n}{G_n + I_n} = 0\} \subset \Upsilon_{G,H} \subset \{\frac{H_n}{H_n + J_n} \to \beta' \in [0,1[\}.$$

And $\beta' \neq 0$, since lemma 8.2.4, and lemma 8.6.1 on $\Upsilon_{G,H}$, imply

$$\Upsilon_{F,G} \cap \Upsilon_{G,H} \cap \{\lim \frac{H_n}{H_n + J_n} = 0\} \subset \Upsilon_{F,G} \cap \Upsilon_{H,G} = \emptyset \text{ a.s.}$$

In summary,

$$\Omega_A \subset \Omega_{A,F} \cup \Omega_{A,1} \cup \Omega_{A,2} \text{ a.s}$$

where

$$\Omega_{A,F} = \Omega_A \cap \{G_\infty < \infty\},$$

$$\Omega_{A,1} = \Upsilon_{A,B} \cap \Upsilon_{B,C} \cap \Upsilon_{F,G} \cap \{\exists \alpha \in]0,1[\ \exists \beta \in]0,1[/\ \frac{C_n}{C_n + E_n} \to \alpha \text{ and } \frac{G_n}{G_n + I_n} \to \beta\} \cap \{G_\infty = \infty\}$$

$$\begin{aligned}\Omega_{A,2} =& \{\exists \alpha \in]0,1[,\ \exists \beta' \in]0,1[/\ \frac{C_n}{C_n + E_n} \to \alpha \text{ and } \frac{H_n}{H_n + J_n} \to \beta'\} \\ & \cap \Upsilon_{A,B} \cap \Upsilon_{B,C} \cap \Upsilon_{F,G} \cap \Upsilon_{G,H} \cap \{G_\infty = \infty\}.\end{aligned}$$

It remains to prove that, for all $A \in \mathbb{Z}$, $\Omega_{A,1}$ and $\Omega_{A,2}$ are of probability 0. In sections 8.7.1 and 8.8.1, we prove that H_n/D_n (resp. I_n/D_n) does converge towards 1 on $\Omega_{A,1}$ (resp. on $\Omega_{A,2}$). In sections 8.7.2 and 8.8.2, we prove that these two cases are almost surely impossible.

In summary, we obtain that events (i), (iv), (v) and (vi) take place on Ω_A a.s, with $k = A + 3$. We conclude for (ii') and (iii') by applying lemma 8.3.7, in the case $P := A$, $Q := B$.

8.3 Martingales results

8.3.1 General martingales results

Let us recall the Chow theorem.

Theorem 8.3.1 *(Chow). Let $(M_n)_{n\in\mathbb{N}}$ be a square integrable martingale. Let*

$$\alpha_n = E((M_{n+1} - M_n)^2 \mid \mathcal{F}_n) < +\infty \ , \ < M >_n = \sum_{k=0}^{n} \alpha_i$$

Then, for all $r > 1/2$,

$$\begin{array}{ll} (\mathbf{1}) & \{< M >_\infty < \infty\} \subset \{\exists M_\infty \in \mathbb{R} / \ M_n \to M_\infty\} \ \ a.s. \\ (\mathbf{2}) & \{< M >_\infty = \infty\} \subset \{M_n = o(< M >_{n-1}^{1/2} (\ln < M >_{n-1})^r)\} \ a.s. \end{array}$$

Now observe that, given a supermartingale $(M_n)_{n\in\mathbb{N}}$ such that the variance of $M_{n+1} - M_n$ decreases as $1/n^\xi$ with $\xi > 1$, M_n converges a.s towards a r.v M_∞ taking values in $\mathbb{R} \cup \{-\infty\}$ (by Chow theorem). The following lemma aims to give an overestimate of the difference $M_n - M_\infty$ as $n \to \infty$.

Corollary 8.3.1 *Let $(M_n)_{n\in\mathbb{N}}$ be a supermartingale such that, given $\xi > 1$,*

$$V(M_{n+1} - M_n \mid \mathcal{F}_n) = O(1/n^\xi).$$

Then, for all $\nu < (\xi - 1)/2$,

(1) $M_n \to M_\infty \in \mathbb{R} \cup \{-\infty\}$ *a.s*

(2) $M_n - M_\infty \leq o(1/n^\nu)$ *a.s*

(3) $\sup_{n\geq p}(M_n - M_p) \leq o(1/p^\nu)$ *a.s*

PROOF: Let us define, for all $k \in \mathbb{N}$,

$$U_k = M_k - M_{k-1} - E(M_k - M_{k-1} \mid \mathcal{F}_{k-1}).$$

Then, for all n, $p \in \mathbb{N}$ such that $n > p$,

$$\begin{aligned} M_n - M_p \quad &= \sum_{k=p+1}^{n} (M_k - M_{k-1}) = \sum_{k=p+1}^{n} U_k + \sum_{k=p+1}^{n} E(M_k - M_{k-1} \mid \mathcal{F}_k) \\ &\leq \sum_{k=p+1}^{n} U_k. \end{aligned}$$

Let $\lambda > 0$ be such that $\lambda < (\xi - 1)2$, and define a random process $(S_n)_{n\in\mathbb{N}}$ by

$$S_n = \sum_{i=1}^{n} i^\lambda U_i \ , \ S_0 = 0.$$

Then $(S_n)_{n\in\mathbb{N}}$ is a martingale such that, for all $n \in \mathbb{N}$,

$$\alpha_n = E((S_{n+1} - S_n)^2 \mid \mathcal{F}_n) = n^{2\lambda} V(M_{n+1} - M_n \mid \mathcal{F}_n) \leq n^{2\lambda - \xi}.$$

It implies, by using Chow theorem, that a.s S_n converges a.s towards a r.v $S_\infty \in \mathbb{R}$ taking values in $\mathbb{R}$.

Now, for all p, $n \in \mathbb{N}$ such that $n \geq p$,

$$\sum_{k=p+1}^{n} U_k = \sum_{k=p+1}^{n} \frac{1}{k^\lambda}(S_k - S_{k-1}) = \sum_{k=p+1}^{n} S_k(\frac{1}{k^\lambda} - \frac{1}{(k+1)^\lambda}) + \frac{S_n}{n^\lambda} - \frac{S_p}{(p+1)^\lambda}.$$

Hence, if we assume $p > Cst(\lambda)$,

$$|\sum_{k=p+1}^{n} U_k| \leq \sup_{k \geq p+1} |S_k|(\frac{1}{(p+1)^\lambda} - \frac{1}{(n+1)^\lambda} + \frac{2}{(p+1)^\lambda}) \leq o(\frac{1}{p^\nu})$$

for all $\nu < \lambda$.

□

We now state a lemma giving an exponential inequality for Bernoulli laws.

Lemma 8.3.1 *Let $(Z_i)_{i \in \mathbb{N}}$ be a sequence of i.i.d Bernoulli 0 and 1 variables with probability of success p. Let, for all $n \in \mathbb{N}^*$,*

$$Z = \frac{Z_1 + Z_2 + \ldots + Z_N}{N},$$

then

$$P(|Z - p| \geq a) \leq \exp\{-NH(a, p)\}$$

with

- *if $a \ll p$, $H(a,p) = \dfrac{a^2}{2p(1-p)} + O\left(\dfrac{a^3}{p^2(1-p)^2}\right)$*
- *if $a = rp$, $r > 0$, a and p are small, $H(a,p) = p((r+1)\ln(r+1) - r) + O(p^2) > 0$.*

8.3.2 Martingales results for the VRRW

General martingales results for the VRRW

Let us first recall a short proof given by Volkov ([11],part 2.1) on the Pólya urn model, which led us to improve and shorten significantly the proof of the proposition stated in this subsection.

Consider an urn containing balls of n different colors. At each time t a ball is drawn at random, and is then replaced together with another ball of the same color. Let $Z(t,i)$ denote the number of balls of color i at time t, and let $Z(t) := \sum_i Z(t,i)$ be the total number of balls. Consider

$$\overline{\alpha}(t) := (\frac{Z(t,1)}{Z(t)}, \ldots, \frac{Z(t,n)}{Z(t)})$$

to be the relative distribution of colors at time t.

Lemma 8.3.2 *(Pólya urns). The vector $\overline{\alpha}(t)$ converges a.s towards some random element in the interior of $(n-1)$-simplex $\Delta \subset \mathbb{R}^n$.*

In fact, it can be proved that $\lim_{t\to\infty} \overline{\alpha}(t)$ has a Dirichlet distribution with parameters depending on the initial distribution of the colors (see Pemantle, [6], lemma 1).

PROOF: (Volkov) For all $i \in [1, n]$,

$$\xi_i(t) := \log Z(t) - \log(Z(t,i) - 1)$$

is a nonnegative supermartingale. This implies that $\xi_i(t)$ converges a.s to a nonnegative random value $\xi_i(\infty)$, and therefore

$$\lim_{t\to\infty} \overline{\alpha}(t) = (\exp(-\xi_1(\infty)), \dots, \exp(-\xi_n(\infty))).$$

□

Given five successive numbers Q, R, S, T and U, if we ignore the visits from Q to R, the term $\ln(R_n + T_n)/(T_n - 1)$ is nonnegative and tends to decrease. Therefore, if there is a few number of visits from Q to R, $T_n/(R_n + T_n)$ converges towards a positive number. Conversely, we can prove that if $R_n/(R_n + T_n)$ converges towards a positive number and if there is a few number of visits from Q to R, then there is few number of visits from U to T.

Proposition 8.3.1 *Let Q, R, S, T, U be five successive numbers, increasing or decreasing. Then*

a)

$$Y_n = \ln(R_n + T_n) - \ln(T_n - 1) - \sum_{k=1}^{n} \frac{1}{R_{k-1} + T_{k-1}} 1_{\{X_{k-1}=Q, X_k=R\}}$$

and

$$\begin{aligned} Z_n = \ln(R_n + T_n) - \ln(T_n - 1) - &\sum_{k=1}^{n} \frac{1}{R_{k-1} + T_{k-1}} 1_{\{X_{k-1}=Q, X_k=R\}} \\ &+ \sum_{k=1}^{n} \frac{R_{k-1}}{(R_{k-1} + T_{k-1})T_{k-1}} 1_{\{X_{k-1}=U, X_k=T\}} \end{aligned}$$

are supermartingales.

b)

$$\begin{aligned} \Upsilon_{Q,R} \quad &\subset \{\frac{R_n}{R_n + T_n} \to \alpha \in [0, 1[\} \\ \Upsilon_{Q,R} \cap \{\alpha > 0\} \quad &\subset \Upsilon_{U,T}. \end{aligned}$$

PROOF: The first statement of **b)** follows from the first statement of **a)**, since

$$\begin{aligned}\Upsilon_{Q,R} \quad &\subset \{\sum_{k=1}^{n} \frac{1}{R_{k-1}+T_{k-1}} 1_{\{X_{k-1}=Q,X_k=R\}} < +\infty\} \subset \{Y_n \text{ bounded below }\} \\ &\subset \{Y_n \text{ converges }\} \subset \{\frac{R_n+T_n}{T_n} \text{ converges }\} \subset \{\frac{R_n}{R_n+T_n} \to \alpha \in [0,1[\},\end{aligned}$$

The second statement of **b)** follows from the second statement of **a)**, since

$$\begin{aligned}\Upsilon_{Q,R} \cap \{\alpha > 0\} \quad &\subset \{\sum_{k=1}^{n} \frac{1_{\{X_{k-1}=Q,X_k=R\}}}{R_{k-1}+T_{k-1}} < +\infty\} \cap \{\alpha > 0\} \\ &\subset \{Z_n \text{ converges }\} \cap \{\alpha > 0\} \cap \{\sum_{k=1}^{n} \frac{1_{\{X_{k-1}=Q,X_k=R\}}}{R_{k-1}+T_{k-1}} < +\infty\} \\ &\subset \{\sum_{k=1}^{n} \frac{R_{k-1}}{(R_{k-1}+T_{k-1})T_{k-1}} 1_{\{X_{k-1}=U,X_k=T\}} < +\infty\} \cap \{\alpha > 0\} \\ &\subset \{\sum_{k=1}^{n} \frac{1_{\{X_{k-1}=U,X_k=T\}}}{T_{k-1}} < +\infty\} = \Upsilon_{U,T}.\end{aligned}$$

□

Remark 8.3.1 A. Bienvenüe also worked on this conjecture of Pemantle and Volkov in his doctorate thesis (1999,[3]) and proved that on the event $R' = \{k-2, k-1, k, k+1, k+2\}$, the asymptotic behaviour of the random walk was really given by the events (ii) to (vi) (using ideas related to the theory of continous vertex-reinforced random walks, e.g Sellke,[10]). Part of proposition 8.3.1, and lemma 8.2.2, give us a short proof of this result : indeed, on $\{R' = \{k-2, k-1, k, k+1, k+2\}\}$, we belong to $\Upsilon_{k-2,k-1}$ and $\Upsilon_{k+2,k+1}$ by lemma 8.2.2. Therefore, applying twice the proposition ($Q = k-2$, $R = k-1$, and $Q = k+2$, $R = k+1$) implies $Z(n, k-1)/(Z(n, k-1) + Z(n, k+1)) \to \alpha \in]0, 1[$. Now we prove that events (ii) and (iii) hold by applying lemma 3.5,[9] (with $X_n = Z(\gamma_n, k-2)$ and $Y_n = Z(\gamma_n, k)$ for event (ii), γ_n being the n-th return to state $k-1$).

□

Martingales estimates for the VRRW

Similarly as in Proposition 8.3.1, given four consecutive integers Q, R, S and T, $R_n/(R_n + T_n)$ tends to decrease if we ignore the visits from Q to R. We use this principle together with corollary 8.3.1 of Chow lemma to overestimate, for all p, n large enough such that $n \geq p$, $R_n/(R_n + T_n)$ given $R_p/(R_p + T_p)$.

Lemma 8.3.3 *Let Q, R, S and T be four successive integers. Let $(M_n)_{n\in\mathbb{N}}$ be the process defined for all $n \in \mathbb{N}$ by*

$$\Pi_n = \prod_{1\le k\le n}(1 + \frac{1_{\{X_{k-1}=Q, X_k=R\}}}{R_k}) \ , \ M_n = \frac{R_n}{\Pi_n(R_n+T_n)}.$$

Then, for all $\nu < 1/2$,

(1) $(M_n)_{n\in\mathbb{N}}$ *is a positive supermartingale,*

(2) $\sup_{n\ge p}|M_n - M_p| \le o(1/(R_p+T_p)^\nu)$.

PROOF: Let us prove **(2)**. Let t_n be the time of $n-2$-th visit to R or T, i.e

$$t_n = \inf\{k \in \mathbb{N}/\ R_k + T_k = n\}.$$

Then (M_{t_n}) is a supermartingale and, for all $n \ge 2$,

$$|M_{t_{n+1}} - M_{t_n}| \le \frac{1}{R_{t_n}+T_{t_n}} = \frac{1}{n}.$$

We conclude by using corollary 8.3.1 of Chow theorem.

□

Following the same heuristics, when $\limsup \ln Q_n/\ln S_n < 1$, there is a very few number of visits from Q to R, which implies that the ratio R_n/S_n tends to decrease. The aim of the following lemma is to give a sharp overestimate of $R_n/S_n - R_p/S_p$ for enough large n and p such that $n \ge p$.

Lemma 8.3.4 *Let Q, R and S be three successive integers, increasing or decreasing. On the event $\{\limsup \ln Q_n/\ln S_n < 1\}$, there exists almost surely $\nu > 0$ such that*

$$\sup_{n\ge p}(\frac{R_n}{S_n} - \frac{R_p}{S_p}) \le o(\frac{1}{S_p^\nu}).$$

PROOF: Let u_n be the time of $n-2$-th visit to R or S, i.e

$$u_n = \inf\{k \in \mathbb{N}/\ R_k + S_k = n\}.$$

Given $\gamma \in \mathbb{R}_+^*$ and $k \in \mathbb{N}^* \setminus \{1\}$, define the stopping time (for filtration $(\mathcal{F}_{u_n})_{n\in\mathbb{N}}$)

$$T = \inf\{n \ge k/\ Q_{u_n} > S_{u_n}^\gamma\}.$$

Observe that

$$\{\limsup \ln Q_n/\ln S_n < 1\} \subset \cup_{\gamma\in\mathbb{Q}_*^+, k\in\mathbb{N}}\{T = +\infty\}.$$

Let, for all $n \in \mathbb{N}^* \setminus \{1\}$, $\theta_n = u_{n\wedge T}$.

Let $(M_n)_{n\in\mathbb{N}}$ be the process defined for all $n \in \mathbb{N}\setminus\{1\}$ by

$$M_n = \frac{R_{\theta_n}}{S_{\theta_n - 1}} - \sum_{i\le n} \frac{Q_{\theta_{n-1}} 1_{X_{\theta_{n-1}}=R}}{(Q_{\theta_{n-1}} + S_{\theta_{n-1}}) S_{\theta_{n-1}-1}}.$$

Then $(M_n)_{n\in\mathbb{N}}$ is a supermartingale. Moreover, for all $n \ge k$,

$$V(M_{n+1} - M_n \mid \mathcal{F}_{\theta_n}) \le \frac{R_{\theta_n} 1_{X_{\theta_n}=S}}{(R_{\theta_n} + E_{\theta_n}) S^2_{\theta_n}} + \frac{Q_{\theta_n} 1_{X_{\theta_n}=R}}{(Q_{\theta_n} + S_{\theta_n}) S^2_{\theta_n}} \le \frac{2}{S^2_{\theta_n}} \le \frac{18}{n^2},$$

the last inequality following from the fact that, when $n < T$, then $n = R_{u_n} + S_{u_n} \le Q_{u_n} + 2S_{u_n} \le 3S_{u_n}$.

Therefore (M_n) satisfies the assumptions of corollary 8.3.1 of Chow theorem with $\xi = 2$. This implies that, for all $\nu < 1/2$,

$$\sup_{n\ge p}(M_n - M_p) \le o(1/p^\nu).$$

On the other hand, for all enough large $p \in \mathbb{N}$,

$$\sum_{n\ge p} \frac{Q_{\theta_n}}{(Q_{\theta_n} + S_{\theta_n}) S_{\theta_n - 1}} \le 2 \sum_{n\ge p} S^{\gamma-2}_{\theta_n} \le 18 \sum_{n\ge p} n^{\gamma-2} = O(p^{\gamma-1}).$$

□

Given three successive integers P, Q and R, the following lemma provides estimates on sums taken over visits from Q to R.

Lemma 8.3.5 *Let P, Q and R be three successive integers, increasing or decreasing. Then, for all $r > 1/2$,*

$$\textbf{(1)} \quad \sum_{k=1}^{n} 1_{X_{k-1}=Q, X_k=R} \asymp \sum_{k=1}^{n} \frac{R_{k-1}}{P_{k-1} + R_{k-1}} 1_{X_{k-1}=Q}$$

$$\textbf{(2)} \quad \sum_{k=1}^{n} 1_{X_{k-1}=Q, X_k=R} - \sum_{k=1}^{n} \frac{R_{k-1}}{P_{k-1} + R_{k-1}} 1_{X_{k-1}=Q} = o(Q_n^{1/2} (\ln Q_n)^r)$$

$$\textbf{(3)} \quad \sum_{k=1}^{n} \frac{1_{X_{k-1}=Q, X_k=R}}{R_{k-1}} \equiv \sum_{k=1}^{n} \frac{1_{X_{k-1}=Q}}{P_{k-1} + R_{k-1}} \equiv \sum_{k=1}^{n} \frac{1_{X_{k-1}=Q, X_k=P}}{P_{k-1}}$$

Observe that **(3)** is a generalization of lemma 8.2.3.

PROOF: We endow notations of Chow theorem. Let us prove first **(1)** and **(2)**.

$$H_n = \sum_{k=1}^{n} 1_{X_{k-1}=Q, X_k=R} - \sum_{k=1}^{n} \frac{R_{k-1}}{P_{k-1} + R_{k-1}} 1_{X_{k-1}=Q}$$

is a martingale, and

$$< H >_n = \sum_{k=1}^{n} \frac{R_{k-1}P_{k-1}1_{X_{k-1}=Q}}{(P_{k-1}+R_{k-1})^2} \leq \sum_{k=1}^{n} \frac{R_{k-1}1_{X_{k-1}=Q}}{P_{k-1}+R_{k-1}} \leq Q_{n-1}.$$

We can conclude for **(1)** and **(2)** by applying the second part of Chow theorem.

Let us now prove **(3)** :

$$G_n = \sum_{k=1}^{n} \frac{1_{X_{k-1}=Q,X_k=R}}{R_{k-1}} - \sum_{k=1}^{n} \frac{1_{X_{k-1}=Q}}{P_{k-1}+R_{k-1}}$$

is a martingale, and

$$< G >_n = \sum_{k=1}^{n} \frac{R_{k-1}P_{k-1}1_{X_{k-1}=Q}}{R_{k-1}^2(P_{k-1}+R_{k-1})^2} \leq \sum_{k=1}^{n} \frac{1_{X_{k-1}=Q}}{R_{k-1}(P_{k-1}+R_{k-1})} < +\infty.$$

Indeed, the last sum is finite by Borel-Cantelli lemma :

$$\sum_{k=1}^{n} \frac{1_{X_{k-1}=Q}}{R_{k-1}(P_{k-1}+R_{k-1})} \asymp \sum_{k=1}^{n} \frac{1}{R_{k-1}^2} 1_{X_{k-1}=Q,X_k=R} < +\infty.$$

□

Given successive integers P, Q, R and S, the following lemma gives estimates on the number of visits to Q with respect to the number of visits to R (resp. to S), when we belong to $\Upsilon_{P,Q}$ and $\limsup R_n/(R_n+T_n) < 1$.

Lemma 8.3.6 *Let P, Q, R, S and T be five successive integers, increasing or decreasing. Then, for all $\gamma_- > 0$ and $\gamma_+ < 1$,*

(1) $$\Upsilon_{P,Q} \cap \{Q_\infty = \infty\} \cap \{\limsup \frac{R_n}{R_n+T_n} \leq \gamma_+\}$$
$$\subset \{\limsup \frac{\ln Q_n}{\ln R_n} \leq \limsup \frac{\ln Q_n}{\ln S_n} \leq \gamma_+\} \cap \{Q_\infty = \infty\}$$
$$\subset \Upsilon_{Q,R} \cap \{R_\infty = S_\infty = \infty\} \ a.s$$

(2) $$\Upsilon_{P,Q} \cap \{Q_\infty = \infty\} \cap \{\liminf \frac{R_n}{R_n+T_n} \geq \gamma_-\} \subset \{\liminf \frac{\ln Q_n}{\ln S_n} \geq \gamma_-\} \ a.s.$$

PROOF: Assume we belong to $\Upsilon_{P,Q} \cap \{Q_\infty = \infty\}$. Then

$$\ln Q_n \equiv \sum_{k=1}^{n} \frac{1_{X_k=Q}}{Q_{k-1}} \equiv \sum_{k=1}^{n} \frac{1_{X_{k-1}=R,X_k=Q}}{Q_{k-1}} \equiv \sum_{k=1}^{n} \frac{1_{X_{k-1}=R,X_k=S}}{S_{k-1}} \equiv \sum_{k=1}^{n} \frac{1_{X_{k-1}=S,X_k=R}}{S_{k-1}}.$$

The second equivalence follows from lemma 8.2.2, the third from part **(3)** of lemma 8.3.5, and the fourth is straightforward. On the other hand, following the notations of Chow theorem,

$$M_n = \sum_{k=1}^{n} \frac{1_{X_{k-1}=S,X_k=R}}{S_{k-1}} - \sum_{k=1}^{n} \frac{1_{X_{k-1}=S}R_{k-1}}{S_{k-1}(R_{k-1}+T_{k-1})}$$

is a martingale such that

$$< M >_n = \sum_{k=1}^{n} \frac{1_{X_{k-1}=S}}{S_{k-1}^2} \frac{R_{k-1}T_{k-1}}{(R_{k-1}+T_{k-1})^2} \leq \sum_{k=1}^{n} \frac{1_{X_{k-1}=S}}{S_{k-1}^2} < \infty.$$

Therefore $\limsup \ln Q_n / \ln S_n \leq \gamma_+$ on $\{\limsup R_n/(R_n+T_n) \leq \gamma_+\} \cap \{Q_\infty = \infty\}$, and $\limsup \ln Q_n / \ln S_n \geq \gamma_-$ on $\{\limsup R_n/(R_n+T_n) \geq \gamma_+\} \cap \{Q_\infty = \infty\}$.

We now assume we belong to $\Upsilon_{P,Q} \cap \{Q_\infty = \infty\} \cap \{\limsup R_n/(R_n+T_n) \leq \gamma_+\}$, with $\gamma_+ < 1$. Let $\tilde{\gamma}_+ = \limsup \ln Q_n / \ln S_n$. We need to prove that, for all $\hat{\gamma}_+ > \tilde{\gamma}_+$, $\limsup \ln Q_n / \ln R_n \leq \hat{\gamma}_+$. Observe that

$$\begin{aligned} Q_n \quad & \asymp \sum_{k=1}^{n} 1_{X_{k-1}=R, X_k=Q} \asymp \sum_{k=1}^{n} \frac{Q_{k-1}}{Q_{k-1}+S_{k-1}} 1_{X_{k-1}=R} \\ & \preceq \sum_{k=1}^{n} \frac{1_{X_{k-1}=R}}{S_{k-1}^{1-\hat{\gamma}_+}} \leq 2 \sum_{k=1}^{n} \frac{1_{X_{k-1}=R}}{R_{k-1}^{1-\hat{\gamma}_+}} \preceq 2 R_n^{\hat{\gamma}_+} / \hat{\gamma}_+. \end{aligned}$$

Indeed, the first equivalence follows from lemma 8.2.2, the second from part **(2)** of lemma 8.3.5, and the inequality follows from $R_k \leq Q_k + S_k \leq 2S_k$ for enough large k.

It remains to prove the last inclusion of **(1)** :

$$\{\limsup \frac{\ln Q_n}{\ln R_n} < 1\} \subset \{\sum \frac{1_{X_{k-1}=Q}}{P_{k-1}+R_{k-1}} < \infty\} = \Upsilon_{Q,R}.$$

□

Corollary 8.3.2 *Let P, Q, R, S and T be five successive integers, increasing or decreasing. Then*

$$\Upsilon_{P,Q} \cap \{\limsup \frac{R_n}{R_n+T_n} < 1\} \subset \Upsilon_{Q,R}.$$

Lemma 8.3.7 *Let P, Q, R, S, T, U and V be seven successive integers, increasing or decreasing. Then*

$$\Upsilon_{P,Q} \cap \Upsilon_{V,U} \cap \{\lim \frac{R_n}{R_n+T_n} \text{ exists and } \in]0,1[\} \subset \{\ln Q_n \equiv \frac{R_n}{R_n+T_n} \ln S_n\}.$$

PROOF: We assume in all the proof we belong to $\Upsilon_{P,Q} \cap \Upsilon_{V,U} \cap \{\lim R_n/(R_n+T_n)$ exists and $\in]0,1[\}$.

Applying lemma 8.3.6, we obtain that $\limsup \ln Q_n / \ln S_n < 1$ and $\limsup \ln U_n / \ln S_n < 1$, i.e that there exists a.s $\nu > 0$ such that $Q_n = o(S_n^{1-\nu})$ and $U_n = o(S_n^{1-\nu})$. Therefore $S_n \leq R_n + T_n \leq S_n + Q_n + U_n = S_n + o(S_n^{1-\nu})$, and

$$\frac{R_n}{R_n+T_n} = \frac{R_n}{S_n} + o(S_n^{-\nu}).$$

Applying lemma 8.3.4 twice (for the behaviors of R_n/S_n and T_n/S_n), we deduce that

$$\frac{R_n}{R_n+T_n} - \lim \frac{R_n}{R_n+T_n} = o(S_n^{-\nu}).$$

On the other hand, it follows from the proof of lemma 8.3.6 that

$$\ln Q_n \equiv \sum_{k=1}^{n} \frac{1_{X_{k-1}=S}}{S_{k-1}} \frac{R_{k-1}}{R_{k-1}+T_{k-1}}.$$

Therefore

$$\ln Q_n \equiv \sum_{k=1}^{n} \frac{1_{X_{k-1}=S}}{S_{k-1}} \frac{R_{k-1}}{R_{k-1}+T_{k-1}} \equiv \lim \frac{R_n}{R_n+T_n} \ln S_n \equiv \frac{R_n}{R_n+T_n} \ln S_n.$$

□

The following lemma 8.3.8 claims that, given four consecutive points L to O, the boundary states have an occupation measure greater than the inverse of logarithm of the number of visits to these boundary states. This result is stated in [9] (Pemantle and Volkov, for the proof of theorem 8.1.2) with the assumption that, after a finite time, the random walk visits these four points only, which implies that $L_n - M_n + N_n - O_n$ can only take two values. We prove it without this assumption.

Lemma 8.3.8 *Let L, M, N and O be four consecutive points. Then, for all $b < 1/2$, there exists almost surely $n \in \mathbb{N}$ such that, for all $k \in \mathbb{N}$ such that $M_k + N_k > n$,*

$$L_k + O_k \geq b \frac{M_k + N_k}{\ln(L_k + O_k)}.$$

We make use of the following lemma 8.3.9 ([9], theorem 2.3, Pemantle and Volkov) for the proof of lemma 8.3.8.

Lemma 8.3.9 *Let $c > 0$. Let us define the random process (P_n, Q_n) as follows. For all n, with probability $P_n/(P_n+Q_n)$, $P_{n+1} = P_n + 1$ and $Q_{n+1} = Q_n$; and, with probability $Q_n/(P_n+Q_n)$, $P_{n+1} = P_n + c$ and $Q_{n+1} = Q_n + 1$.*

Then $P_n/(cQ_n) - \ln Q_n$ converges to a random limit in $(-\infty, +\infty)$.

PROOF OF LEMMA 8.3.8 : Let τ_m be the time of the m-th visit to M or N skipping at least one step,

$$\tau_m = \inf\{n > \tau_{m-1} + 1 : X_n \in \{M, N\}\}$$

and $\tau_0 = 0$.

Define

$$U_m = \frac{M_{\tau_m} + N_{\tau_m}}{2}, \; V_m = L_{\tau_m} + O_{\tau_m}.$$

At time τ_m, suppose for instance that we are at state M : VRRW will either go to the left, or go twice to the right, or go to the right and make one step back.

On the two first events, $U_{m+1} \leq U_m + 1$ and $V_{m+1} \geq V_m + 1$, and on the third $U_{m+1} = U_m + 1$ and $V_{m+1} = V_m$.

Let us call F the union of the two first events, and show that $P(F) \geq V_m/(U_m + V_m)$. This will imply that the process (U, V) can be coupled with the urn model (P, Q) described in lemma 8.3.9, with $U_m \leq P_m$, $V_m \geq Q_m$ and $c = 1$, and will therefore complete the proof of this lemma.

It is equivalent to prove that

$$P(F) = \frac{L}{L+N} + \frac{N}{L+N} . \frac{O}{M+O} \geq \frac{L+O}{L+O+\frac{M+N}{2}} = \frac{V}{U+V}.$$

which can be written as $[L(M+O)+NO][L+O+\frac{M+N}{2}] \geq (L+N)(L+O)(M+O)$.

Expanding the left-hand side of this latter inequality :

$$\begin{aligned}
&[L(M+O)+NO][L+O+\frac{M+N}{2}] \\
&= L(M+O)(L+O) + N(L+O)O + L(M+O)\frac{M+N}{2} + NO\frac{M+N}{2}.
\end{aligned}$$

Now, since $M+O \geq N$, $L(M+O)\frac{(M+N)}{2} \geq L(M+O)\frac{N}{2} + LN\frac{M}{2} \geq LMN + \frac{LNO}{2}$. Therefore

$$\begin{aligned}
&L(M+O)\frac{M+N}{2} + NO\frac{M+N}{2} \geq LMN + \frac{LNO}{2} + \frac{NOM}{2} + \frac{NON}{2} \\
&\geq LMN + NOM \text{ (indeed } \frac{LNO}{2} + \frac{NON}{2} \geq \frac{NOM}{2} \text{ since } L+N \geq M)
\end{aligned}$$

Finally

$$\begin{aligned}
&[L(M+O)+NO][L+O+\frac{M+N}{2}] \\
&\geq L(M+O)(L+O) + N(L+O)O + N(L+O)M \\
&\geq (L+N)(L+O)(M+O)
\end{aligned}$$

□

8.4 Other general results for the VRRW

The following lemma provides, given four consecutive states L, M, N and O, conditions ensuring that the process will eventually remain left hand from O.

Lemma 8.4.1 *Let L, M, N and O be four successive states, increasing or decreasing. Let a_1, $a_2 > 0$ be two constants. Let $(r(n))_{n\in\mathbb{N}}$ be a process taking values in $]0,1[$ adapted to $(\mathcal{F}_n)_{n\in\mathbb{N}}$.*

Given m_0, p_0, $q_0 \in \mathbb{N}$ and $\nu > 0$, define the probability sets

$$\Pi^{N,O}_{0,p_0,a_1,d,r} = \{r(p_0) \le 1-d,\ \forall n \ge p_0,\ N_n \le a_1 L_n^{r(n)} \text{ and, for all } n \ge p \ge p_0,\ r(n) \le r(p)+L_p^{-\nu}\},$$

$$\Pi^{N,O}_{1,m_0,a_2,r} = \{\exists i \ge m_0/\ X_i = M \text{ and } O_i \le a_2 M_i^{1-r(i)}\}$$

and

$$\Pi^{N,O}_{2,q_0} = \{\forall n \ge q_0,\ X_n < O\}.$$

Then

$$\Pi^{N,O}_{0,p_0,a_1,d,r} \cap (\cap_{m_0\ge p_0}\Pi^{N,O}_{1,m_0,a_2,r}) \subset \cup_{q_0\in\mathbb{N}}\Pi^{N,O}_{2,q_0}.$$

PROOF: For simplicity, we may write Π_{0,p_0}, Π_{1,m_0}, and Π_{2,q_0} respectively, the probability sets $\Pi^{N,O}_{0,p_0,a_1,d,r}$, resp. $\Pi^{N,O}_{1,m_0,a_2,r}$ and $\Pi^{N,O}_{2,q_0}$. Let us define the stopping time

$$R = \inf\{n \ge m_0/\ X_i = M \text{ and } O_i \le a_2 M_i^{1-r(i)}\}$$

associated with Π_{2,m_0}.

For $i \ge p_0$, let

$$S_i = \inf\{n \ge i/\ X_n = O\}$$

and

$$T_i = \inf\{n \ge i/\ r(n) > 1-d \text{ or } N_n > a_1 L_n^{r(n)} \text{ or } \exists p \in [i,n]/\ r(n) > r(p) + L_p^{-\nu}\}$$

be the stopping times associated with Π_{2,q_0} and Π_{1,m_0}. It suffices to show that there exists a constant $c < 1$ such that, for all $m_0 \ge p_0$,

$$P(\Pi_{0,p_0} \cap (\cap_{m_0'\ge p_0}\Pi_{1,m_0'}) \cap (\cap_{q_0\in\mathbb{N}}\Pi^c_{2,q_0}) \mid \mathcal{F}_{m_0}) \le c.$$

Now

$$\begin{aligned} P(\Pi_{0,p_0} \cap \Pi_{1,m_0} \cap (\cap_{q_0\in\mathbb{N}}\Pi^c_{2,q_0}) \mid \mathcal{F}_{m_0}) &\le \sum_{i=m_0}^{+\infty} E(1_{R=i}E(1_{\{T_i=+\infty \text{ and } S_i<+\infty\}} \mid \mathcal{F}_i) \mid \mathcal{F}_{m_0}) \\ &\le \sum_{i=m_0}^{+\infty} E(1_{R=i}E(1_{\{S_i<T_i\}} \mid \mathcal{F}_i) \mid \mathcal{F}_{m_0}). \end{aligned}$$

Therefore it suffices to show that there exists a constant $c > 0$ such that, if $X_i = M$ and $O_i \le a_2 M_i^{1-r(i)}$, then $P(S_i \ge T_i \mid \mathcal{F}_i) \ge c$. We write $S = S_i$ and $T = T_i$.

Let ζ_n be the time of n-th visit to state M after time i, i.e $\zeta_0 = i$ and, for all $n \in \mathbb{N}$,

$$\zeta_{n+1} = \inf\{k > \zeta_n/\ X_k = M\}.$$

We want an estimate of $P(\{X_{\zeta_n+2} = O\} \mid \mathcal{F}_{\zeta_n})$ for $\zeta_n < S \wedge T$.

If $j < S \wedge T$, $X_j = M$ and $j \geq i$, then

$$P(X_{j+2} = O \mid \mathcal{F}_j) \leq N_j/L_j.O_j/M_j \leq a_1 L_j^{r(j)}/L_j.a_2 M_i^{1-r(i)}/M_j \leq 2a_1a_2 M_j^{r(i)+L_i^{-\nu}-2} M_i^{1-r(i)},$$

observing that $M_n \leq L_n + N_n \leq a_1 L_n^{1-d} + L_n \leq 2L_n$ if L_{p_0} was supposed large enough, which can be assumed without loss of generality. Therefore

$$\begin{aligned} P(S \geq T \mid \mathcal{F}_i) &\geq \prod_{n=0}^{+\infty}(1 - 2a_1a_2(n+M_i)^{r(i)+L_i^{-\nu}-2} M_i^{1-r(i)}) \\ &\geq \exp(-4a_1a_2 M_i^{1-r(i)} \sum_{n=0}^{+\infty}(n+M_i)^{r(i)+L_i^{-\nu}-2}) \geq \exp(-8a_1a_2 d^{-1}) \end{aligned}$$

(if M_{p_0} is greater than a constant which depends only on a_1, a_2, d and ν).

$\square$

Corollary 8.4.1 *Let L, M, N and O be four successive states, increasing or decreasing.*

$$\{\liminf_{n\to\infty} \frac{\ln O_n}{\ln M_n} + \limsup_{n\to\infty} \frac{\ln N_n}{\ln L_n} < 1\} \subset \{O_\infty < \infty\}.$$

PROOF: Let

$$\Gamma = \{\liminf_{n\to\infty} \frac{\ln O_n}{\ln M_n} + \limsup_{n\to\infty} \frac{\ln N_n}{\ln L_n} < 1\} \cap \{O_\infty = \infty\}.$$

Let us endow notations of lemma 8.4.1 in the particular case where $r(n)$ equals the constant function $r \in]0,1[$ over $\mathbb{N}$. Then

$$\Gamma \subset \cup_{r\in]0,1[,a_1,a_2\in\mathbb{R}_*^+}(\Pi_{0,p_0,a_1,d,r}^{N,O} \cap (\cap_{m_0\geq p_0} \Pi_{1,m_0,a_2,r}^{N,O})) \subset \cup_{q_0\in\mathbb{N}} \Pi_{2,q_0}^{N,O} \cap \{L_\infty = \infty\} = \emptyset \text{ a.s.}$$

$\square$

The following lemma gives conditions ensuring that a certain class of processes a.s remains far from 1 after a finite time. It makes use of the heuristic of a result of Benaïm about convergence with positive probability towards an attractor ([2], chapter 7).

Lemma 8.4.2 *Let $(\mathcal{G}_n)_{n\in\mathbb{N}}$ be a filtration. Let $(\gamma_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$, $(\epsilon_n)_{n\in\mathbb{N}}$ and $(s_n)_{n\in\mathbb{N}}$ be sequences of $\mathcal{G}_n$-mesurable random variables, taking values in $\mathbb{R}^+$ (resp. $\mathbb{R}_*^+$, $\mathbb{R}$, $\mathbb{R}$), and such that*

$$\sum_{k\in\mathbb{N}} |s_k| < \infty \text{ and } \sum_{k\in\mathbb{N}} E(\epsilon_{k+1}^2 \mid \mathcal{F}_k) < \infty \text{ a.s.}$$

Let $\lambda > 1$ possibly infinite, and let $\mu \in [1, \sqrt{\lambda}]$.

Assume

$$\begin{cases} \ln y_{n+1} - \ln y_n = \gamma_n(y_n - 1) + \epsilon_{n+1} + s_n & \text{if } y_n \in [\lambda^{-1}, \lambda] \\ y_{n+1} \geq \mu^{-1}\lambda & \text{if } y_n \geq \lambda \\ y_{n+1} \leq \mu\lambda^{-1} & \text{if } y_n \leq \lambda^{-1} \end{cases}$$

Then y_n enjoys the following properties.

1)

$$\{y_n \not\to 1\} \subset \{\liminf y_n > 1\} \cap \{\limsup y_n < 1\}.$$

2) *If $\lambda = \infty$, then*

$$\{y_n \not\to 1\} \subset \{\ln y_n \asymp \sum_{k=1}^{n} \gamma_k(y_k - 1)\}.$$

If $\lambda < \infty$ and

$$\sum_{k\in\mathbb{N}} \gamma_k = +\infty \text{ on } \bigcup_{n\in\mathbb{N}} \bigcap_{k\geq n} \{y_k \in [\lambda^{-1}, \lambda]\},$$

then

$$\{y_n \not\to 1\} \subset \{\liminf y_n \geq \mu^{-2}\lambda > 1\} \cup \{\limsup y_n \leq \mu^2\lambda^{-1} < 1\}.$$

3)

$$\{y_n \to 1\} \subset \{\sum_{k\in\mathbb{N}} \gamma_k |y_k - 1| < +\infty\}.$$

PROOF: We begin with the proofs of **1)** and **2)**. By definition,

$$\{\limsup y_n > 1\} = \cup_{\delta>1} \cap_{k_0\in\mathbb{N}} \{\exists n_0 \geq k_0/\ y_{n_0} \geq \delta\}.$$

Fix $\delta > 1$, and suppose we belong to $\cap_{k_0\in\mathbb{N}}\{\exists n_0 \geq k_0/\ y_{n_0} \geq \delta\}$.

Now we can choose without loss of generality a nonnegative integer $p_0 \in \mathbb{N}$ such that $\sum_{n\geq p_0} |r_n| < \min(\ln\delta/3, \ln\mu/2, 1)$ and, applying Chow lemma, take p_0 such that

$$\sup_{p\geq p_0, n\geq p} |\sum_{k=p}^{n} \epsilon_{k+1} 1_{y_k\in[\lambda^{-1},\lambda]}| \leq \min(\ln\delta/3, \ln\mu/2).$$

We know that there exists $n_0 \geq p_0$ such that $y_{n_0} \geq \delta$; this implies, for all $n \geq n_0$, as long as $y_n \in [1, \lambda]$,

$$\ln y_n = \ln y_{n_0} + \sum_{k=n_0}^{n-1} \gamma_k(y_k - 1) + \sum_{k=n_0}^{n-1} \epsilon_{k+1} + \sum_{k=n_0}^{n-1} s_n \geq \ln\delta - 2\ln\delta/3 \geq \ln\delta/3,$$

This proves **1)** and **2)** when $\lambda = \infty$ (the argument being the same in the case $\liminf y_n < 1$).

Let us now prove **2)** when $\lambda < \infty$. We can derive from the latter inequality that

$$\ln y_n \geq \ln \delta/3 + (\sqrt[3]{\delta} - 1) \sum_{k=n_0}^{n-1} \gamma_k,$$

and therefore that there exists $q_0 \geq n_0$ such that $y_{q_0} \geq \lambda$. For all $n \geq q_0$, let $\tilde{n}$ ne the greatest integer $k \in [q_0, n]$ such that $y_k \geq \lambda$. If $\tilde{n} \neq n$, then $y_{\tilde{n}+1} \geq \mu^{-1}\lambda$ and, for all $k \in [\tilde{n}+1, n[$,

$$\ln y_k \geq \ln \mu^{-1}\lambda + \sum_{k=\tilde{n}+1}^{n-1} \epsilon_{k+1} + \sum_{k=\tilde{n}+1}^{n-1} s_k \geq \ln \mu^{-1}\lambda - \ln \mu \geq \ln \mu^{-2}\lambda > 0.$$

Therefore $y_n \geq \mu^{-2}\lambda$.

The proof of **1)** when $\lambda < \infty$ is similar.

Let us now prove statement **3)**. Let us define, for all $x \in \mathbb{R}_*^+$, $\delta(x) = 1_{x \geq 1} - 1_{x<1}$.

Suppose we belong to $\{y_n \to 1\}$. Then there exists $q_0 \in \mathbb{N}$ such that $y_n \in [\lambda^{-1}, \lambda]$ for $n \geq q_0$. We can take q_0 large enough so that $\sum_{k \geq q_0} |r_k| \leq 1$ and

$$\sup_{p \geq q_0, n \geq p} |\sum_{k=p}^{n} \epsilon_{k+1} 1_{y_k \in [\lambda^{-1}, \lambda]} \delta(y_k)| \leq 1.$$

For all $n \geq q_0$,

$$|\ln y_{n+1}| \geq |\ln y_n| + \gamma_n |y_n - 1| + \delta(y_n)(\epsilon_{n+1} + s_n).$$

Therefore, for all m, $n \geq q_0$,

$$\sum_{k=m}^{n} \gamma_k |y_k - 1| \leq |\ln y_{n+1}| - |\ln y_m| - \sum_{k=m}^{n} \epsilon_{k+1} \delta(y_k) - \sum_{k=m}^{n} \delta(y_k) s_k \leq 2|\ln \lambda| + 2.$$

This completes the proof of **3)**.

□

8.5 Coupling results for the VRRW

We first define a partial order relation on a certain class of random walks on $\mathbb{Z}$, and give sufficient conditions for two random walks to be coupled in order to be comparable by the order relation.

Definition 8.5.1 *Let $k \in \mathbb{N}$; we say $(x_0, \ldots, x_k) \in \mathbb{Z}^{k+1}$ is a path on $\mathbb{Z}$ if and only if (iff), for all $0 \leq n \leq k-1$, there exists $\epsilon_n \in \{-1, 1\}$ such that $x_{n+1} - x_n = \epsilon_n$.*

Let $i \in \mathbb{N}^*$ and $j \in \mathbb{Z}$; let $n_{i,j}(x_0, \ldots, x_k)$ be the i-th visit time to state j of the sequence $(x_n)_{0 \le n \le k}$, with the convention that $n_{i,j}(x_0, \ldots, x_k) = +\infty$ if j has been visited less than i times.

We also define $Z(n,j)(x_0, \ldots, x_k)$, $n \le k$, the number of times plus one site j has been visited by the sequence x_i until time n.

These definitions extend to the case of a sequence $(x_p)_{p \in \mathbb{N}}$ indexed by $\mathbb{N}$.

When this is unambiguous, we write $n_{i,j}$ for $n_{i,j}(x_0, \ldots, x_k)$, $n'_{i,j}$ for $n_{i,j}(x'_0, \ldots, x'_{k'})$; likewise we write $Z(n,j)$ and $Z'(n,j)$ without indicating the sequences $(x_0, \ldots, x_k)$ and $(x'_0, \ldots, x'_{k'})$.

Definition 8.5.2 *Let $\mathcal{P}$ be the set of paths on $\mathbb{Z}$ indexed by $\mathbb{N}$, and let $\mathcal{T}$ be the smallest σ-field containing the cylinders $C_{z_0,\ldots,z_p} = \{(x_k)_{k \in \mathbb{N}} / x_0 = z_0, \ldots, x_p = z_p\}$.*

Definition 8.5.3 *Let $(x_0, \ldots, x_k) \in \mathbb{Z}^{k+1}$ and $(x'_0, \ldots, x'_{k'}) \in \mathbb{Z}^{k'+1}$ be two paths on $\mathbb{Z}$ (k or k' may be infinite). Let $i \in \mathbb{N}^*$ and $j \in \mathbb{Z}$; we define property $E_{i,j}((x_i)_{0 \le i \le k}, (x'_i)_{0 \le i \le k'})$ (written $E_{i,j}$ when this is unambiguous) as follows :*

$$Z'(n'_{i,j}, j+1) \ge Z(n_{i,j}, j+1) \text{ and } Z'(n'_{i,j}, j-1) \le Z(n_{i,j}, j-1)$$

with the convention that $E_{i,j}$ holds whenever $n_{i,j} = +\infty$ or $n'_{i,j} = +\infty$.

Definition 8.5.4 *Two random walks $\mathcal{M} = (X_k)_{k \in \mathbb{N}}$ and $\mathcal{M}' = (X'_{k'})_{k' \in \mathbb{N}}$ taking values on $\mathbb{Z}$ and defined on the same probability space $(\Omega, \mathcal{F}, P)$ are said to be coupled iff, for almost every $\omega \in \Omega$, $E_{i,j}((X_k(\omega))_{k \in \mathbb{N}}, (X'_{k'}(\omega))_{k' \in \mathbb{N}})$ holds for all $i \in \mathbb{N}^*$ and $j \in \mathbb{Z}$. We let $\mathcal{M}' \gg \mathcal{M}$ denote this property.*

Thus $\mathcal{M}' \gg \mathcal{M}$ means that, for a same number i of visits to j, $\mathcal{M}'$ has more visited point $j+1$ (right hand from j) than $\mathcal{M}$, and less visited $j-1$.

Lemma 8.5.1 *Let $\mathcal{M}$ and $\mathcal{M}'$ denote two random walks on $\mathbb{Z}$, starting from the same point $X_0 = X'_0 = v$.*

Suppose, for all $i \in \mathbb{N}^*$ and $j \in \mathbb{Z}$, for all paths $(x_0, \ldots, x_k) \in \mathbb{Z}^{k+1}$, $(x_0, \ldots, x'_{k'}) \in \mathbb{Z}^{k'+1}$ (with k, $k' \in \mathbb{N}$) starting from v, ending at the same point $x_k = x'_{k'} = j$, and having visited j a same number i of times,

$$P(X'_{k'+1} = j+1 | X'_k = x'_{k'}, \ldots, X'_0 = x_0) \ge P(X_{k+1} = j+1 | X_k = x_k, \ldots, X_0 = x_0)$$

whenever $E_{i,j}$ holds.

Then the processes $\mathcal{M}$ and $\mathcal{M}'$ can be coupled so that $\mathcal{M}' \gg \mathcal{M}$.

PROOF: Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $(\omega_{i,j})_{i \in \mathbb{N}^*, j \in \mathbb{Z}}$ be i.i.d random variables of density 1 on $[0,1]$.

Let $\mathcal{Z} = (Z_k)_{k\in\mathbb{N}}$ be an arbitrary random walk on $\mathbb{Z}$ such that $Z_0 = v$. Let $q(\mathcal{Z}) = (q(\mathcal{Z})(z_0, \dots, z_k))_{(z_i)_{i\in\mathbb{N}}\in\mathcal{P}, k\in\mathbb{N}}$ be the assembly of conditional probabilities for $\mathcal{Z}$, defined for all $(z_i)_{i\in\mathbb{N}} \in \mathcal{P}$ and $k \in \mathbb{N}$ by

$$q(\mathcal{Z})(z_0, \dots, z_k) = P(Z_{k+1} = z_k + 1 | Z_k = z_k, \dots, Z_0 = z_0).$$

We construct a random walk on $(\Omega, \mathcal{F}, P)$ having the expected conditional probabilities $q(\mathcal{Z})$, also written $\mathcal{Z}$ by a slight abuse of language, as follows : $Z_0 = v$ and, for all $n \in \mathbb{N}$, given $(Z_0, \dots, Z_n)$:

$$\begin{cases} Z_{n+1} = Z_n + 1 & \text{if } \omega_{Z(n,Z_n)-1, Z_n} \leq q(\mathcal{Z})(Z_0, \dots, Z_n) \\ Z_{n+1} = Z_n - 1 & \text{otherwise .} \end{cases}$$

We make use of this construction to define $\mathcal{M} = (X_k)_{k\in\mathbb{N}}$ and $\mathcal{M}' = (X'_{k'})_{k'\in\mathbb{N}}$ on $(\Omega, \mathcal{F}, P)$ given the conditional probabilities $q(\mathcal{M})$ and $q(\mathcal{M}')$.

Now consider an arbitrary element $(\omega_{i,j})_{\mathbb{N}^*\times\mathbb{Z}}$ of $[0,1]^{\mathbb{N}^*\times\mathbb{Z}}$, and consider the sequence $(x_i)_{i\in\mathbb{N}}$ (resp. $(x'_i)_{i\in\mathbb{N}}$) corresponding to the values of $(X_i)_{i\in\mathbb{N}}$ (resp. $(X'_i)_{i\in\mathbb{N}}$) given $(\omega_{i,j})_{\mathbb{N}^*\times\mathbb{Z}}$.

Let

$$p_{i,j} = q(\mathcal{M})(x_0, \dots, x_{n_{i,j}}) \ , \ p'_{i,j} = q(\mathcal{M}')(x'_0, \dots, x'_{n'_{i,j}}).$$

We want to prove that $E_{i,j}$ holds for (x_i) and (x'_i), for all $(i,j) \in \mathbb{N}^* \times \mathbb{Z}$. Observe that :

• $x_0 = v$

• There are two instances for x_{p+1}, knowing $(x_0, \dots, x_p)$ for $p \geq 1$: there exists $(i,j) \in \mathbb{N}^* \times \mathbb{Z}$ such that $p = n_{i,j}$,

– if $\omega_{i,j} \leq p_{i,j}$ then $x_{p+1} = x_p + 1$

– if $\omega_{i,j} > p_{i,j}$ then $x_{p+1} = x_p - 1$

The same remark holds for $\mathcal{M}'$, replacing $p_{i,j}$ by $p'_{i,j}$ and $n_{i,j}$ by $n'_{i,j}$.

Let us introduce the property

$$P_k = \{\forall (i,j) \in \mathbb{N}^* \times \mathbb{Z} / n_{i,j} \leq k \text{ and } n'_{i,j} \leq k, E_{i,j} \text{ holds } \}.$$

Let us prove by induction on k that P_k holds for all $k \in \mathbb{N}$. P_0 follows from $X_0 = X'_0 = v$. Suppose P_{k-1} ; we want to deduce P_k, which is different from P_{k-1} if there exists $(i,j) \in \mathbb{N}^* \times \mathbb{Z}$ such that $n_{i,j} = k$ and $n'_{i,j} \leq k$, or $n'_{i,j} = k$ and $n_{i,j} \leq k$. Select such a couple (i,j).

If $i = 1$, then suppose for instance that $j > x_0 = v$ (the case $j < v$ is analogous, and $j = v$ is obvious). Then $Z'(n'_{i,j}, j+1) = Z(n_{i,j}, j+1) = 1$, and we aim to prove that $Z'(n'_{i,j}, j-1) \leq Z(n_{i,j}, j-1)$.

Suppose the contrary $Z'(n'_{i,j}, j-1) > Z(n_{i,j}, j-1)$ and let $a = Z(n_{i,j}, j-1) - 1$. Since P_{k-1} holds, $E_{a,j-1}$ holds and therefore $p'_{a,j-1} \geq p_{a,j-1}$. Now $x_{n_{a,j-1}+1} = x_{n_{a,j-1}} + 1$ since $a = Z(n_{i,j}, j-1) - 1$; this implies $x'_{n_{a,j-1}+1} = x'_{n_{a,j-1}} + 1$ and leads to a contradiction.

If $i > 1$, take $\mathcal{M}$ at time $n_{i-1,j}$ and $\mathcal{M}'$ at time $n'_{i-1,j}$. We make use of the notation $\mathcal{M}$ or $\mathcal{M}' \to l$ or r in order to indicate that $\mathcal{M}$ or $\mathcal{M}'$ goes to the left or to the right at this time $n_{i-1,j}$ (resp. $n'_{i-1,j}$). Since $E_{i-1,j}$ is satisfied, $p'_{i-1,j} \geq p_{i-1,j}$ and it is impossible that $\mathcal{M} \to r$ and $\mathcal{M}' \to l$. Hence, there are three cases :

–$\mathcal{M} \to l$ and $\mathcal{M}' \to r$, and conclusion follows.

–$\mathcal{M} \to r$ and $\mathcal{M}' \to r$, and conclusion follows from an analogous proof to case $i = 1$.

–$\mathcal{M} \to l$ and $\mathcal{M}' \to l$, and conclusion follows from an analogous proof.

$\square$

We define A'_n (resp. B'_n, etc...) the number of times plus one site A (resp. B, etc...) has been visited by $\mathcal{M}'$ at time n, similarly as A_n, B_n, etc...for $\mathcal{M}$.

In the sequel, we replace the previous definition of the filtration $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$ by, for all $n \in \mathbb{N}$, $\mathcal{F}_n = \sigma(\omega_{i,j})_{(i,j)\in\mathbb{N}^*\times\mathbb{Z}/\ n_{i,j}\leq n}$. Similarly, we define the filtration $\mathbb{F}' = (\mathcal{F}'_n)_{n\in\mathbb{N}}$ by, for all $n \in \mathbb{N}$, $\mathcal{F}'_n = \sigma(\omega_{i,j})_{(i,j)\in\mathbb{N}^*\times\mathbb{Z}/\ n'_{i,j}\leq n}$.

We also define a notion of stopping time on paths.

If m, $m' \in \mathbb{N}$ and $m' > m$, we write, for all subset A of $\mathbb{Z}^{m'+1}$, $A|_{\mathbb{Z}^{m+1}} = \{(x_0, \ldots, x_m) \in \mathbb{Z}^{m+1},\ \exists(x_{m+1}, \ldots, x_{m'}) \in \mathbb{Z}^{m'-m}/\ (x_0, \ldots, x_{m'}) \in A\}$.

Fix a family of subsets $\mathcal{T}_m$ of $\mathbb{Z}^{m+1}$ such that, for all $m\ m' \in \mathbb{N}$, $m' > m$, $\mathcal{T}_{m'}|_{\mathbb{Z}^{m+1}} \cap \mathcal{T}_m = \emptyset$, define a stopping time t on the paths defined by $\{\mathrm{t} = m\} = \mathcal{T}_m$. We can define stopping times T and T' corresponding to t for the random walks $\mathcal{M}$ and $\mathcal{M}'$ by $T = m \iff (X_1, \ldots, X_m) \in \mathcal{T}_m = \{\mathrm{t} = m\}$ and $T' = m \iff (X'_1, \ldots, X'_m) \in \mathcal{T}_m = \{\mathrm{t} = m\}$. We will say in the remainder of the proof that two stopping times, respectively for the random walks $\mathcal{M}$ and $\mathcal{M}'$, are associated if they are corresponding to a same stopping time on the paths.

We also say that two sequences of $(\mathcal{F}_n)_{n\in\mathbb{N}}$-stopping times $(T_k)_{k\in\mathbb{N}}$ (for $(\mathcal{F}_n)_{n\in\mathbb{N}}$) and $(\mathcal{F}'_n)_{n\in\mathbb{N}}$-stopping times $(T'_k)_{k\in\mathbb{N}}$ (for $(\mathcal{F}'_n)_{n\in\mathbb{N}}$) are associated iff, for all $k \in \mathbb{N}$, T_k and T'_k are associated.

Similarly we say that an $(\mathcal{F}_n)_{n\in\mathbb{N}}$-adapted process $(\Delta_n)_{n\in\mathbb{N}}$ corresponds to a process defined on the paths iff there exists a sequence of functions $F_n :\ \mathcal{P}|_{\mathbb{Z}^{n+1}} \to \mathbb{R}$, $n \in \mathbb{N}$, such that $\Delta_n = F_n(X_0, \ldots, X_n)$. Observe that if such a sequence of functions exists, it is uniquely defined. In this case we say that the $(\mathcal{F}_n)_{n\in\mathbb{N}}$-adapted process $(\mathcal{D}_n)_{n\in\mathbb{N}}$ and a $(\mathcal{F}'_n)_{n\in\mathbb{N}}$-adapted process $(\mathcal{D}'_n)_{n\in\mathbb{N}}$ are associated iff, for all $n \in \mathbb{N}$, $\Delta'_n = F_n(X_0, \ldots, X'_n)$.

We will also make use of notations $\mathcal{I} : \Omega \to \mathcal{P}$, $\omega \in \Omega \mapsto (X_i(\omega))_{i\in\mathbb{N}}$ and $\mathcal{I}' : \Omega \to \mathcal{P}$, $\omega \in \Omega \mapsto (X'_i(\omega))_{i\in\mathbb{N}}$.

Lemma 8.5.2 *Let L, M and N be three consecutive integers. Let $k_0 \in \mathbb{N}$, $g_1 \in]0, 1]$ and $g_2 > 0$.*

We will define a random walk $\mathcal{M}'$ coupled to the original VRRW $\mathcal{M}$.

Set $r = M_{k_0}$ and define

$$p_n = \frac{L_n}{L_n + N_n},\ \ \delta_n = g_2 \frac{p_n}{\sqrt{2p_{k_0} r}},\ \ \overline{p}_n = p_n - \delta_n$$
$$p'_n = \frac{L'_n}{L'_n + N'_n},\ \ \delta'_n = g_2 \frac{p'_n}{\sqrt{2p_{k_0} r}},\ \ \overline{p'}_n = p'_n - \delta'_n.$$

Let us consider a stopping time on paths u, and let s, v, w be stopping times on paths defined by

$$s_1 = \inf\{n \geq k_0 / M_n > (g_1 + 1) M_{k_0}\}\ ,\ \ s_2 = \inf\{n \geq k_0 / (LM)_{k_0,n} > g_1 p_{k_0} M_{k_0}\}$$
$$v = \inf\{j \geq k_0 /\ p_j \notin [p_{k_0}/2, 2p_{k_0}]\}\ ,\ \ w_i = s_i \wedge u \wedge v,\ \ i \in \{1,2\}.$$

Let S_1 and S'_1 (resp. S_2 and S'_2, U and U', V and V', W_1 and W'_1, W_2 and W'_2) be the stopping times for $\mathcal{M}$ and $\mathcal{M}'$ corresponding to s_1 (resp. s_2, u, v, w_1, w_2).

Let $i \in \{1,2\}$. The random walk $\mathcal{M}'$ is defined with the same conditional probabilities as $\mathcal{M}$ from all points different from M, the conditional laws starting from M being

$$P(X'_{j+1} = L \mid X'_j = M, \mathcal{F}_j) = \overline{p'_n} \quad \text{if } j \in [k_0, W'_i[$$
$$P(X'_{j+1} = L \mid X'_j = M, \mathcal{F}_j) = p'_n \quad \text{if } j \geq W'_i \text{ or } j < k_0$$

By lemma 8.5.1, these two processes can be coupled so that $\mathcal{M}' \gg \mathcal{M}$. Assume $p_{k_0} \in]0, Cst[$, and $p_{k_0} M_{k_0} > Cst(\tilde{\epsilon})$ in the case $i = 2$. Then, for all $\tilde{\epsilon} > 0$ and $\mathcal{C} \in \mathcal{T}$,

$$P_{\mathcal{M}'}(\mathcal{C} \mid \mathcal{F}_{k_0}) \leq M(g_2, \tilde{\epsilon}) P_{\mathcal{M}}(\mathcal{C} \mid \mathcal{F}_{k_0}) + \tilde{\epsilon},$$

where $M(g_2, \tilde{\epsilon})$ only depends on g_2 and $\tilde{\epsilon}$.

PROOF: Let us prove the lemma for $i = 2$, the case $i = 1$ being similar. Let

$$Z = (z_{k_0}, \ldots, z_k)$$

be a path such that $k \leq \mathrm{s} \wedge \mathrm{u} \wedge \mathrm{v}$. Set, when $X_{k_0} = z_{k_0}$,

$$P_{\mathcal{M}}(Z) = P_{\mathcal{M}}((X_{k_0}, \ldots, X_k) = (z_{k_0}, \ldots, z_k) \mid \mathcal{F}_{k_0})$$

for simplicity, and similarly for $P_{\mathcal{M}'}(Z)$. The probabilities $P_{\mathcal{M}}(Z)$ and $P_{\mathcal{M}'}(Z)$ are different because of the different probabilities for $\mathcal{M}$ and $\mathcal{M}'$ starting from M.

The proof consists to bound the rate $P_{\mathcal{M}'}(Z)/P_{\mathcal{M}}(Z)$ on a set of paths of large probability for $\mathcal{M}'$.

More precisely, let us define $I = P_{\mathcal{M}'}(Z)/P_{\mathcal{M}}(Z)$. Then

$$I = \prod_{z_j=M,z_{j+1}=L} (1-\frac{\delta_j}{p_j}) \prod_{z_j=M,z_{j+1}=N} (1+\frac{\delta_j}{1-p_j}),$$

and

$$\ln I \leq \sum_{z_j=M,z_{j+1}=N} \frac{\delta_j}{1-p_j} - \sum_{z_j=M,z_{j+1}=L} \frac{\delta_j}{p_j} = \sum_{z_j=M} \frac{\delta_j}{1-p_j} - \sum_{z_j=M,z_{j+1}=L} \frac{\delta_j}{p_j(1-p_j)}$$
$$\leq \sum_{j=k_0}^{k} \frac{\delta_j}{p_j(1-p_j)}[p_j - 1_{z_{j+1}=L}]1_{z_j=M} = \frac{g_2}{\sqrt{2p_{k_0}r}} \sum_{j=k_0}^{k} \frac{1}{1-p_j}[p_j - 1_{z_{j+1}=L}]1_{z_j=M}$$

Let us define the stopping times $P = \inf\{n \geq k_0/ \ M'_n > 8M_{k_0}\}$, and $T' = P' \wedge W'$.

Let, for all $j \geq k_0$,

$$S'_n = \sum_{j=k_0+1}^{k} \frac{1}{1-p'_{j-1}}[p'_{j-1} - 1_{X'_j=L}]1_{X'_{j-1}=M}.$$

Then $(S'_n)_{n\geq k_0+1}$ is a martingale, and

$$E(S'^2_{T'} \mid \mathcal{F}_{k_0}) \leq 8p_{k_0}E(\sum_{j=k_0+1}^{T'} 1_{X'_{j-1}=M} \mid \mathcal{F}_{k_0}) \leq 128 p_{k_0}M_{k_0} = 128p_{k_0}r,$$

since $p_j \leq 2p_{k_0} \leq 1/2$.

Let $\{\Delta = \{|S'_{T'}| \leq B\sqrt{2p_{k_0}r}\}$, then $P(\Delta^c \mid \mathcal{F}_{k_0}) \leq 64/B^2 < \tilde{\epsilon}$ for $B = 8/\sqrt{\tilde{\epsilon}}$.

It remains to prove that $W < P$ on Δ. Indeed, assume the contrary, i.e $M'_{T'} > 8M_{k_0}$. Then

$$\sum_{j=k_0+1}^{T'} \frac{1_{X'_{j-1}=M,X'_j=L}}{1-p'_{j-1}} \geq \sum_{j=k_0+1}^{T} \frac{p'_{j-1}}{1-p'_{j-1}}1_{X_j=M} - B\sqrt{2p_{k_0}r} \geq 2p_{k_0}M_{k_0} - B\sqrt{2p_{k_0}M_{k_0}},$$

which implies $(LM)'_{k_0,T} > p_{k_0}M_{k_0}$ and leads to a contradiction.

Therefore conclusion holds with $M(g_2, \tilde{\epsilon}) = \exp(8g_2/\sqrt{\tilde{\epsilon}})$.

$\square$

Lemma 8.5.3 *Let $(R_n)_{n\in\mathbb{N}}$, $(S_n)_{n\in\mathbb{N}}$ and $(U_n)_{n\in\mathbb{N}}$ be three $(\mathcal{F}_n)_{n\in\mathbb{N}}$-adapted processes associated to a process defined on paths.*

Let $(\delta_n)_{n\in\mathbb{N}}$ be an increasing sequence of stopping times associated to paths. Let T be a stopping time, associated to a stopping time on paths t. Let L, M and N be three consecutive integers, and let $g_1 \in]0,1]$, $g_3 \in]0,1[$, $i \in \{1,2\}$.

Let $u(\delta_n)_{n\in\mathbb{N}}$, $v(\delta_n)_{n\in\mathbb{N}}$ be $(\mathcal{F}_{\delta_n})_{n\in\mathbb{N}}$-adapted processes.

Given g_1, L, M and N, let us consider, for all $g_2 > 0$ and all $m \in \mathbb{N}$, the random walk $\mathcal{M}'$ defined in lemma 8.5.2 starting at $k_0 := \delta_m$, with the stopping time on paths $u = t \wedge \delta_{m+\lfloor g_3 v(\delta_m)\rfloor}$. Let us define the $(\mathcal{F}'_n)_{n\in\mathbb{N}}$-adapted processes $(R'_n)_{n\in\mathbb{N}}$, $(S'_n)_{n\in\mathbb{N}}$ and $(U'_n)_{n\in\mathbb{N}}$ associated to $(R_n)_{n\in\mathbb{N}}$, $(S_n)_{n\in\mathbb{N}}$ and $(U_n)_{n\in\mathbb{N}}$, the stopping time T' associated to T and the increasing sequence $(\delta'_n)_{n\geq m}$ associated to $(\delta_n)_{n\geq m}$. Note that $\mathcal{M}'$ depends on m.

Assume that, for all $m \in \mathbb{N}$,

a) *There exist $(\mathcal{F}_{\delta_n})_{n\geq m}$-adapted processes $(\tilde{U}_n)_{n\geq m}$, $(\epsilon_n)_{n>m}$, $(r_n)_{n>m}$ with $\tilde{U}_n = U_{\delta_n}$ if $\delta_n < T$, $(\mathcal{F}'_{\delta'_n})_{n\geq m}$-adapted processes $(\tilde{U}'_n)_{n\geq m}$, $(\epsilon'_n)_{n>m}$, $(r'_n)_{n>m}$ with $\tilde{U}'_n = U'_{\delta'_n}$ if $\delta'_n < T'$, sequences of $\mathcal{F}_{\delta_m}$-mesurable positive reals $(m(\delta_m, n))_{n\in\mathbb{N}}$, $(n(\delta_m, n))_{n\in\mathbb{N}}$, such that the following holds for $n \geq m$.*

$$E(\epsilon_{n+1} \mid \mathcal{F}_{\delta_n}) = 0\ ,\ E(\epsilon_{n+1}^2 \mid \mathcal{F}_{\delta_n}) \leq m(\delta_m, n)^2\ ,\ |r_n| \leq n(\delta_m, n)^2$$
$$E(\epsilon'_{n+1} \mid \mathcal{F}_{\delta_n}) = 0\ ,\ E(\epsilon'^2_{n+1} \mid \mathcal{F}_{\delta_n}) \leq m(\delta_m, n)^2\ ,\ |r'_n| \leq n(\delta_m, n)^2$$

$$\ln \frac{\tilde{U}_{n+1}}{\tilde{U}_n} = \frac{R_{\delta_{n+1}\wedge T}}{S_{\delta_n}} + \epsilon_{n+1} + r_n \text{ if } \delta_n < T,$$

$$\begin{cases} \ln \frac{\tilde{U}'_{n+1}}{\tilde{U}'_n} \geq \frac{R'_{\delta'_{n+1}\wedge T'}}{S'_{\delta'_n}} + g_2 u(\delta_m) + \epsilon'_{n+1} + r'_n & \text{when } n \in [m, m + \lfloor g_3 v(\delta_m)\rfloor[\text{ and } \delta'_n < T \\ \ln \frac{\tilde{U}'_{n+1}}{\tilde{U}'_n} = \frac{R'_{\delta'_{n+1}\wedge T'}}{S'_{\delta'_n}} + \epsilon'_{n+1} + r'_n & \text{when } n \geq \lfloor g_3 v(\delta_m)\rfloor \text{ and } \delta'_n < T. \end{cases}$$

b) There exists a constant $C_1 \in \mathbb{R}_+^*$ such that

$$\max\left(\sqrt{\sum_{n=m}^{\infty} m(\delta_m, n)^2}, \sqrt{\sum_{n=m}^{\infty} n(\delta_m, n)^2}\right) \leq C_1 u(\delta_m) v(\delta_m),$$

and $u(\delta_m)v(\delta_m) \to_{m\to\infty} 0$, $v(\epsilon_m) \to_{m\to\infty} \infty$.

c) For all $n \in \mathbb{N}$,

$$R'_{\delta'_n} \geq R_{\delta_n}\ ,\ U'_{\delta'_n} \geq U_{\delta_n}.$$

d)

$$\text{(i)} \sum_{\delta_{n+1}<T} \frac{R^+_{\delta_{n+1}}}{S_{\delta_n}} < +\infty \text{ or (ii)} \sum_{\delta_{n+1}<T} \frac{R^-_{\delta_{n+1}}}{S_{\delta_n}} > -\infty.$$

e) Let, for all $n \geq m$ such that $\delta_n < T$ and $\delta'_n < T'$, $g_n = U'_{\delta'_n}/U_{\delta_n}$ and $h_n = S'_{\delta'_n}/S_{\delta_n}$. Then

$$\text{(i) } \exists C_2 \in \mathbb{R}_*^+ /\ h_n \in [C_2 g_n^{-1}, C_2 g_n]$$

or

(ii) There exists $C_2 \geq 72C_1$ and $C_3 \in \mathbb{R}_*^+$ such that

$$\ln g_n \leq C_2 u(\delta_m) v(\delta_m) \implies |\ln h_n| \leq C_3 u(\delta_m) v(\delta_m),$$

and, for all $n \geq m$,

$$g_{n+1} \geq g_n - C_2 u(\delta_m) v(\delta_m)/2.$$

Then

$$P(\{\lim_{n\to\infty} U_{\delta_n} = 1\} \cap \{T = \infty\}) = 0,$$

with the convention that $\lim_{n\to\infty} U_{\delta_n} \neq 1$ when there exists $n \in \mathbb{N}$ such that $\delta_n = \infty$.

The lemma remains true if we replace, in hypotheses **a)** and **d)**, $R_{\delta_{n+1}\wedge T}$ (resp. $R'_{\delta'_{n+1}\wedge T'}$) by $R_{\delta_n \wedge T}$ (resp. $R'_{\delta'_n \wedge T'}$).

Remark 8.5.1 Assumption **d)** is independent of m and $\mathcal{M}'$.

Condition $C_2 \geq 72 C_1$ in assumption **e)**(ii) is only technical, and the assumption is in the case of application satisfied for all $C_2 \in \mathbb{R}_*^+$. This condition **e)**(ii) corresponds to the fact that, when $\ln g_n$ is in the order of the noise generated by (ϵ_n), (ϵ'_n), (r_n), (r'_n), $\ln h_n$ is of the same order. It is related, in section 8.8, to a kind of Lipschitz condition of $\ln S_n$ with respect to $\ln U_n$, i.e $|\ln S_n - \ln S_p|$ is in the order of $|\ln U_n - \ln U_p|$ on small time intervals. However, such a definition would not be suitable, since we have to deal with the noise inherent to the random walk : a small change of $\ln U_n$ does not generate with probability 1 a similar change of $\ln S_n$.

□

PROOF: We give the proof under assumption **d)**(i), the case **d)**(ii) being similar.

It suffices to prove that there exists a constant $\eta > 0$ (which will be chosen in the following) such that, for all $m_0 \in \mathbb{N}$ and $m \geq m_0$,

$$P[\{\lim_{n\to\infty} U_{\delta_n} \neq 1\} \cup \{Q \wedge T < \infty\}) \mid \mathcal{F}_{\delta_m}) \geq \eta,$$

where $Q = \inf\{n \geq \delta_{m_0} / \ \exists k \geq m_0, \delta_k \leq n / \ |\ln U_{\delta_k}| > \iota \text{ or } u(\delta_k)v(\delta_k) > \iota \text{ or } \sum_{i=m_0}^{k-1} R^+_{\delta_{i+1}}/S_{\delta_i} > \iota\}$, where ι will be fixed in the following.

Let us assume the contrary : we will prove that it leads to a contradiction for a good choice of η.

Firstly, this implies $P(T < \infty \mid \mathcal{F}_{\delta_m}) \leq \eta$. Applying lemma 8.5.2, this implies $P(T' < \infty) \mid \mathcal{F}_{\delta_m}) \leq M(g_2, \tilde{\epsilon}) P(T < \infty \mid \mathcal{F}_{\delta_m}) + \tilde{\epsilon}$. We choose $\eta = \tilde{\epsilon}/M(g_2, \tilde{\epsilon})$ in order to have $P(T' < \infty) \leq 2\tilde{\epsilon}$. We assume without loss of generality that $\eta \leq \tilde{\epsilon}$.

Now let $c > 0$ and define

$$\{\Delta_1 = \sup_{k_1 \geq m, n \geq k_1} |\sum_{k=k_1}^{n-1} \epsilon_{k+1}| \leq c u(\delta_m) v(\delta_m) \ , \ \sup_{k_1 \geq m, n \geq k_1} |\sum_{k=k_1}^{n-1} \epsilon'_{k+1}| \leq c u(\delta_m) v(\delta_m)\}.$$

Applying Doob's inequality and assumption **c)**,

$$E(\sup_{n\geq m}|\sum_{k=m}^{n-1}\epsilon_{k+1}|^2 \mid \mathcal{F}_{\delta_m}) \leq 4\sum_{n=m}^{\infty} m(\delta_m,n)^2 \leq 4C_1^2(u(\delta_m)v(\delta_m))^2.$$

Hence, using the same argument for $(\epsilon'_{k+1})_{k\geq m}$, $P(\Delta_1^c \mid \mathcal{F}_{\delta_m}) \leq 1/2$ if we choose $c = 8C_1$. Moreover, recall that $\sum_{k=m}^{\infty}|r_k| \leq C_1 u(\delta_m)v(\delta_m)$, $\sum_{k=m}^{\infty}|r'_k| \leq C_1 u(\delta_m)v(\delta_m)$.

Let us define

$$\Delta_2 = \{\lim_{n\to\infty} R_{\delta_n} = 1\} \cap \{Q \wedge T = \infty\} \cap \{T' = \infty\}\ ,\ \Delta = \Delta_1 \cap \Delta_2.$$

We obtain that $P(\Delta \mid \mathcal{F}_{\delta_m}) \geq 1/2 - 3\tilde{\epsilon}$. In the sequel, assume we belong to Δ.

Let, for all $k \in \mathbb{N}$, $m_k = \ln g_k = \ln U'_{\delta'_k}/U_{\delta_k} \geq 0$ (assumption **c)**), $n_k = \ln h_k = \ln S'_{\delta'_k}/S_{\delta_k}$, $\hat{m}_k = \sup_{i\in[m,k]} m_i$.

We now estimate m_k for $k \geq m + \lfloor g_3 v(\delta_m)\rfloor$:

$$\begin{aligned} m_k \geq& (\frac{g_2 g_3}{2} - 2c - 2C_1)u(\delta_m)v(\delta_m) + \sum_{i=m}^{k-1}(\frac{R'_{\delta'_{i+1}}}{S'_{\delta'_i}} - \frac{R_{\delta_{i+1}}}{S_{\delta_i}}) \\ \geq& (\frac{g_2 g_3}{2} - 2c - 2C_1)u(\delta_m)v(\delta_m) + \sum_{i=m}^{k-1} R_{\delta_{i+1}}(\frac{1}{S'_{\delta'_i}} - \frac{1}{S_{\delta_i}}), \end{aligned}$$

the second inequality following from assumption **c)**.

We now need to estimate the sum of the terms $S_{\delta_{n+1}}(1/T'_{\delta'_n} - 1/T_{\delta_n})$. Let us assume that **e)**(ii) is satisfied, the proof in the case **e)**(i) being similar.

We first deal with the case $\hat{m}_k \leq C_2 u(\delta_m)v(\delta_m)$, which implies, for $i \in [m,k]$, $|n_i| \leq C_3 u(\delta_m)v(\delta_m)$ and therefore

$$|R_{\delta_{i+1}}(\frac{1}{S'_{\delta'_i}} - \frac{1}{S_{\delta_i}})| \leq \frac{|R_{\delta_{i+1}}|}{S_{\delta_i}}\max(|e^{n_i}-1|, |1-e^{-n_i}|) \leq 2C_3 u(\delta_m)v(\delta_m)\frac{|R_{\delta_{i+1}}|}{S_{\delta_i}}.$$

On the other hand, for all $k > m$,

$$\ln\frac{U_{\delta_k}}{U_{\delta_m}} = \sum_{i=m}^{k-1}\frac{R_{\delta_{i+1}}}{S_{\delta_i}} + \sum_{i=m}^{k-1}\epsilon_{i+1} + \sum_{i=m}^{k-1} r_i,$$

which implies

$$|\sum_{i=m}^{k-1}\frac{R^+_{\delta_{i+1}}}{S_{\delta_i}} + \sum_{i=m}^{k-1}\frac{R^-_{\delta_{i+1}}}{S_{\delta_i}}| = |\sum_{i=m}^{k-1}\frac{R_{\delta_{i+1}}}{S_{\delta_i}}| \leq 2cu(\delta_m)v(\delta_m) + 2\iota \leq 2\iota(c+2).$$

Therefore
$$\sum_{i=m}^{k-1} \frac{|R_{\delta_{i+1}}|}{S_{\delta_i}} \leq 2\iota(c+1).$$

We obtain in summary
$$m_k \geq [g_2 g_3/2 - 2(c+C_1) - 4\iota C_3(c+2)]u(\delta_m)v(\delta_m).$$

In the other case $\hat{m}_k \geq C_2 u(\delta_m)v(\delta_m)$, we consider k_0 to be the greatest j such that $m_j \geq C_2 u(\delta_m)v(\delta_m)$. This implies $m_{j+1} \geq C_2 u(\delta_m)v(\delta_m)/2$ and, if $k \neq j+1$,
$$m_k \geq [C_2/2 - 2(c+C_1) - 4\iota C_3(c+2)]u(\delta_m)v(\delta_m).$$

In summary, if we choose $c = 8C_1$, $\iota = C_2/(32C_3(c+2))$, $g_2 = C_2/g_3$, we obtain $m_k \geq C_2 u(\delta_m)v(\delta_m)/4$ in the two cases (recall that $C_2 \geq 72C_1$). Therefore $\liminf_{n\to\infty} U'_{\delta'_n}/U_{\delta_n} > 1$ on Δ.

Let us now choose $\tilde{\epsilon} = 1/10$; recall that $\eta = \tilde{\epsilon}/M(g_2, \tilde{\epsilon})$.

Let $\mathcal{C}_1 = \mathcal{I}'(\Gamma)$. It follows that $P_{\mathcal{M}'}(\mathcal{C}_1 \mid \mathcal{F}_{\delta_m}) \geq P(\Delta \mid \mathcal{F}_{\delta_m}) \geq 1/2 - 3\tilde{\epsilon}$, which implies $P_{\mathcal{M}}(\mathcal{C}_1 \mid \mathcal{F}_{\delta_m}) \geq (1/2 - 3\tilde{\epsilon})M(g_2,\tilde{\epsilon})^{-1} \geq 2\tilde{\epsilon}M(g_2,\tilde{\epsilon})^{-1} > \eta$.

Accordingly, assumption $\eta = \tilde{\epsilon}/M(g_2, \tilde{\epsilon})$ is contradictory, since $\liminf U_{\delta_n} > 1$ on $\mathcal{C}_1$. This completes the proof.

□

The following lemma provides a tool to compute the expectation and variance of the number of visits to a state before the first return to the previous state for VRRW. It is useful at various steps of the proof.

Lemma 8.5.4 *Let L, M and N be three consecutive integers and suppose $X_{n_0} = M$. Let $p \in]0,1[$. Define the stopping times*
$$R = \inf\{n \geq n_0 \ , \ \frac{N_n + 1}{L_n + N_n + 1} \geq p\},$$
and
$$S = \inf\{n \geq n_0 \ , \ X_n = L\}.$$
Let T be a stopping time satisfying $T \leq R$.

Let $W_{n_0,L,M,N}$ denote the number of visits to M before the first visit to state L after time n_0 ($W_{n_0,L,M,N}$ is $\mathcal{F}_S$-mesurable).

Then $W_{n_0,L,M,N}$ can be coupled with a $\mathcal{F}_S$-mesurable r.v $V_{n_0,L,M,N}$ so that, setting $V = V_{n_0,L,M,N}$ and $W = W_{n_0,L,M,N}$ for simplicity, when $n_0 < R$,
$$V = W \ \text{on} \ \{S \leq T\} \cap \{S < +\infty\},$$
$$E(V \mid \mathcal{F}_{n_0}) = E(\frac{L_{n_0} + N_{S\wedge T}}{L_{n_0}} \mid \mathcal{F}_{n_0}),$$
and
$$Var(V \mid \mathcal{F}_{n_0}) = E((V - E(V))^2 \mid \mathcal{F}_{n_0}) \leq E(V(V-1) \mid \mathcal{F}_{n_0}) \leq \frac{2p}{(1-p)^2}.$$

PROOF: We endow the notations of this section (in particular $n_{i,j}$ and the r.v $\omega_{i,j}$), and set $L = L_{n_0}$, $m_0 = M_{n_0}$ for simplicity.

Define the random variable

$$V_{n_0,L,M,N} = \inf\{n \geq m_0 - 1/\ \omega_{n,M} > \frac{N_{n_{n,M}\wedge T}}{L_{n_0} + N_{n_{n,M}\wedge T}}\} - (m_0 - 2)$$

(recall that at time n_0, M has been visited $m_0 - 1$ times).

Set $V = V_{n_0,L,M,N}$ and $W = W_{n_0,L,M,N}$ for simplicity. Notice that

$$W = \inf\{n \geq m_0 - 1/\ \omega_{n,M} > \frac{N_{n_{n,M}}}{L_{n_0} + N_{n_{n,M}}} \text{ or } N_{n_{n+1,T}} = \infty\} - (m_0 - 2).$$

Therefore $V \leq W$ on $\{S < +\infty\}$, V is $\mathcal{F}_S$-mesurable, and

$$V = W \text{ on } \{S \leq T\} \cap \{S < +\infty\}.$$

Moreover,

$$\begin{aligned} E(V \mid \mathcal{F}_n) &= \sum_{k=0}^{+\infty} P(V > k) = \sum_{k,r \geq 0} P(V > k, N_{n_{k+m_0-1,M}\wedge T} = r) \\ &= \sum_{k,r \geq 0} P(V = k+1, N_{n_{k+m_0-1,M}\wedge T} = r)\frac{L+r}{L} = E(\frac{L_{n_0} + N_{S\wedge T}}{L_{n_0}} \mid \mathcal{F}_{n_0}) \end{aligned}$$

Moreover, when $n_0 < R$, note that $N_n/(L_n + N_n) \leq p$ for all $n \in [n_0, R]$; this implies that $P(V > k) \leq p^k$; then

$$\begin{aligned} &E(V(V-1) \mid \mathcal{F}_{n_0}) = \sum_{j=1}^{+\infty} P(V = j) \sum_{k \leq j-1} 2k \leq \sum_{k=1}^{+\infty} (\sum_{j>k} P(V = j))2k \\ &\leq \sum_{k=1}^{+\infty} P(V > k)2k \leq 2\sum_{k=1}^{+\infty} kp^k = \frac{2p}{(1-p)^2} \end{aligned}$$

□

Remark 8.5.2 Let L, M, N and O be four consecutive integers, and T be a stopping time; suppose $X_{n_0} = N$. Let W denote the number of visits to N before the first visit to state L.

Define the stopping times

$$R = \inf\{n \geq n_0/\ \frac{L_n}{L_n + N_n + 1}\frac{M_n}{M_n + O_n + 1} \leq 1 - p\}$$

and $S = \inf\{n \geq n_0/\ X_n = L\}$.

We can similarly couple W with a random variable V (called $V_{n_0,L,M,N,O}$) so that $V = W$ on $\{S \leq T\} \cap \{S < +\infty\}$,

$$E(V \mid \mathcal{F}_{n_0}) = E(\frac{L_{n_0} + N_{S\wedge T}}{L_{n_0}} \frac{M_{S\wedge T} + O_{S\wedge T}}{M_{S\wedge T}}),$$

and

$$Var(V \mid \mathcal{F}_{n_0}) \leq E(V(V-1) \mid \mathcal{F}_{n_0}) \leq 2p/(1-p)^2.$$

$\square$

8.6 Proof of lemma 8.6.1

This section is devoted to the proof of the following lemma.

Lemma 8.6.1 *Let P, Q, R and S be four consecutive integers. Then*

$$\Upsilon_{P,Q} \cap \{\lim \frac{Q_n}{Q_n + S_n} = 0\} \subset \Upsilon_{Q,R} = \Upsilon_{Q,P}.$$

Lemma 8.6.1 is a consequence of the following lemma.

Lemma 8.6.2 *Let B, C, D, E and F be five consecutive points. Then*

$$\Upsilon_{B,C} \cap \{\lim \frac{C_n}{C_n + E_n} = 0\} \subset \{\limsup \frac{D_n}{D_n + F_n} < 1\}.$$

Indeed, applying lemma 8.6.2 with $B := P$, $C := Q$, $D := R$, $E := S$ and $F := T$ implies

$$\Upsilon_{P,Q} \cap \{\lim \frac{Q_n}{Q_n + S_n} = 0\} \subset \{\limsup \frac{R_n}{R_n + T_n} < 1\},$$

and on the other hand corollary 8.3.2 implies that

$$\Upsilon_{P,Q} \cap \{\limsup \frac{R_n}{R_n + T_n} < 1\} \subset \Upsilon_{Q,R}.$$

Let us now begin with the proof of lemma 8.6.2. Note that this proof is absolutely independant of the proof of the conjecture, and that the choice of integers B, etc... is not connected to a belonging to Ω_A.

Set

$$\Gamma_0 = \Upsilon_{B,C} \cap \{\lim \frac{C_n}{C_n + E_n} = 0\} \cap \{\liminf \frac{F_n}{D_n + F_n} = 0\}.$$

We have to prove that $P(\Gamma_0) = 0$.

Observe that on Γ_0, $C_n/(C_n + E_n) \to 0$ and $\liminf F_n/(D_n + F_n) = 0$ imply D_n, $E_n \to \infty$.

Let, for all $n \in \mathbb{N}$,

$$p_n = \frac{C_n}{C_n + E_n} \text{ and } \tilde{p}_n = \max(\frac{C_n}{C_n + E_n}, \frac{F_n}{D_n + F_n}).$$

Let $i_0 \in \mathbb{N}$ and $\epsilon > 0$; let

$$\Gamma_{1,i_0,\epsilon} = \{\tilde{p}_{i_0} \leq \epsilon/2 \text{ and } \min(D_{i_0}, E_{i_0}) \geq 1/\epsilon^4\}$$

and define the stopping time

$$T_0 = \inf\{n \geq i_0/ \sum \frac{1_{X_{n-1}=B, X_n=C}}{C_n} > \epsilon\}.$$

The definition of Γ_0 implies that

$$\Gamma_0 \subset \cup_{i_0 \in \mathbb{N}}(\{T_0 = +\infty\} \cap \Gamma_{1,i_0,\epsilon})$$

for all $\epsilon > 0$.

Let us define, for all $n \in \mathbb{N}$,

$$\Pi_n = \prod_{1 \leq k \leq n} (1 + \frac{1_{\{X_{k-1}=B, X_k=C\}}}{C_k}),\ M_n = \frac{C_n}{\Pi_n(C_n + E_n)}.$$

Applying lemmas 8.3.3 (with $Q := B$, $R := C$, $S := D$, $T := E$) and 8.3.8 (with $L := C$, $M := D$, $N := E$, $O := F$), we obtain that $(M_n)_{n \in \mathbb{N}}$ is a supermartingale and that, given the stopping times

$$U_1 = \inf\{n \geq i_0/ \exists k \in [i_0, n]/ |M_n - M_k| > 1/(C_k + E_k)^{1/4}\},$$

$$U_2 = \inf\{n \geq i_0/ C_n + F_n < \frac{3}{8} \frac{D_n + E_n}{\ln(C_n + F_n)}\},$$

$$T_1 = T_0 \wedge U_1 \wedge U_2,$$

$$\Gamma_0 \subset \cup_{i_0 \in \mathbb{N}}(\{T_1 = +\infty\} \cap \Gamma_{1,i_0,\epsilon}\}).$$

Now define, given $\mu \in]1, 2[$, the stopping time

$$U_3 = \inf\{n \geq i_0/ \exists j_0 \in [i_0, n],\ p_n > \mu \tilde{p}_{j_0}\}.$$

Let us prove that $U_3 \geq T_1$ for $\epsilon < Cst(\mu)$. First note that, for all $j_0 < U_2$,

$$\tilde{p}_{j_0} \geq \frac{1}{4 \ln D_{j_0}}.$$

Indeed, either $\tilde{p}_{j_0} \geq 1/5$ or ($4C_{j_0} \leq E_{j_0}$ and $4F_{j_0} \leq D_{j_0}$). Similarly, either $\tilde{p}_{j_0} \geq 1/3$ or $D_{j_0} \in [2E_{j_0}/3, 3E_{j_0}/2]$ (observe for instance that $D_{j_0} \leq C_{j_0} + E_{j_0} \leq 3E_{j_0}/2$ when $\tilde{p}_{j_0} \leq 1/3$).

This implies in particular that, either $\tilde{p}_{j_0} \geq 1/5$ or $2\max(C_{j_0}, F_{j_0}) \leq \max(D_{j_0}, E_{j_0})/2 \leq \min(D_{j_0}, E_{j_0})$.

Assume for instance $C_{j_0} \geq F_{j_0}$. Then

$$\frac{C_{j_0}}{C_{j_0}+E_{j_0}} \geq \frac{3}{16}\frac{1}{\ln(2C_{j_0})}\frac{D_{j_0}+E_{j_0}}{C_{j_0}+E_{j_0}} \geq \frac{3}{16}\frac{1}{\ln D_{j_0}}\frac{5E_{j_0}/3}{5E_{j_0}/4} \geq \frac{1}{4\ln D_{j_0}}.$$

The other case is similar. Therefore, $i_0 \leq j_0 \leq n < T_1$ implies

$$\frac{C_n}{C_n+E_n} \leq e^{\epsilon}M_n \leq e^{\epsilon}(M_{j_0} + 1/(C_{j_0}+E_{j_0})^{1/4}) \leq e^{\epsilon}(\tilde{p}_{j_0} + 1/D_{j_0}^{1/4}) \leq \mu\tilde{p}_{j_0}$$

if we assume $\epsilon < Cst(\mu)$ (which implies $D_{j_0} \geq Cst(\mu)$). We make this choice of $\epsilon < Cst(\mu)$, so that $T_1 = T_1 \wedge U_3$.

Let us define, for $j_0 \geq i_0$ such that $X_{j_0} = D$, a sequence α_n as follows :

$$\begin{cases} \alpha_0 = j_0 \\ \forall n \in \mathbb{N},\ \alpha_{n+1} = \inf\{k > \alpha_n /\ X_k = D \text{ and } X_{k-1} = E\}. \end{cases}$$

We estimate the variations of $\ln E_{\alpha_n}/D_{\alpha_n}$, and use these estimates to show for all $\mu > 1$ that $F_n \leq \mu C_n$ for enough large n. Then we show that this last case is a.s impossible.

Pick $q \in]0,1[$ and define the stopping times

$$T_2 = T_1 \wedge \inf\{n \geq i_0 /\ F_n > \mu C_n\}\ ,\ T_3 = \inf\{n \geq j_0 /\ \tilde{p}_n > q\}.$$

Set, for all $k \in \mathbb{N}$,

$$R_k = D_k + F_k - (C_k + E_k).$$

Observe that $-C_k \leq R_k \leq F_k$.

Lemma 8.6.3 *Let i_0, $j_0 \in \mathbb{N}$ be such that $j_0 \geq i_0$ and $X_{j_0} = D$; let $q \in]0,1[$, $\epsilon > 0$ and $\mu \in]1,2[$. Suppose $q < Cst$. Then there exist random variables $(\tilde{D}_n)_{n\in\mathbb{N}}$, $(\tilde{E}_n)_{n\in\mathbb{N}}$, $(\epsilon_n)_{n\in\mathbb{N}}$, $(r_n)_{n\in\mathbb{N}}$ adapted to the filtration $(\mathcal{F}_{\alpha_n})_{n\in\mathbb{N}}$ such that, if $\alpha_n < T_3$, then $\tilde{D}_n = D_{\alpha_n}$, $\tilde{E}_n = E_{\alpha_n}$, and*

$$\ln\frac{\tilde{E}_{n+1}}{\tilde{D}_{n+1}} - \ln\frac{\tilde{E}_n}{\tilde{D}_n} = \frac{R_{\alpha_{n+1}\wedge T_3}}{D_{\alpha_n}E_{\alpha_n}} + \epsilon_{n+1} + r_n, \tag{8.1}$$

with $E(\epsilon_{n+1} \mid \mathcal{F}_{\alpha_n}) = 0$, $E(\epsilon_{n+1}^2 \mid \mathcal{F}_{\alpha_n}) \leq \frac{Cst.q}{(n+D_{j_0})^2}$ and $|r_n| \leq \frac{Cst}{(n+D_{j_0})^2}$.

PROOF: We assume $q < Cst$, where Cst is as small as necessary.

Let us construct $\tilde{D}_n$ iteratively as follows : $\tilde{D}_0 = D_{j_0}$ and, for all $n \in \mathbb{N}$:

$$\begin{cases} \tilde{D}_{n+1} - \tilde{D}_n = V_{\alpha_n,E,D,C} & \text{if } T_3 > \alpha_n \\ \tilde{D}_{n+1} = \tilde{D}_n & \text{otherwise} \end{cases}$$

where $V_{\alpha_n,E,D,C}$ is the random variable defined in lemma 8.5.4 with $T := T_3$, which satisfies $V_{\alpha_n,E,D,C} = D_{\alpha_{n+1}} - D_{\alpha_n}$ if $\alpha_{n+1} \leq T_3$. Hence $\tilde{D}_n = D_{\alpha_n}$ if $\alpha_n \leq T_3$.

Define $\tilde{E}_n$ identically, iteratively out of $V_{\beta_n,D,E,F}$, where β_n is the first visit to state E after time α_n.

Suppose $T_3 > \alpha_n$ in the following computations.

We wish to compute

$$E(\ln \frac{\tilde{E}_{n+1}}{\tilde{D}_{n+1}} - \ln \frac{\tilde{E}_n}{\tilde{D}_n} \mid \mathcal{F}_{\alpha_n}) = E(\ln \frac{\tilde{E}_{n+1}}{\tilde{E}_n} \mid \mathcal{F}_{\alpha_n}) - E(\ln \frac{\tilde{D}_{n+1}}{\tilde{D}_n} \mid \mathcal{F}_{\alpha_n}),$$

and to estimate the variance of $\ln \tilde{E}_{n+1}/\tilde{D}_{n+1}$. Write $V_{\alpha_n,D,E,F} = V$ for simplicity. Now

$$E(\ln(\frac{\tilde{D}_{n+1}}{\tilde{D}_n}) \mid \mathcal{F}_{\alpha_n}) = \frac{E(V \mid \mathcal{F}_{\alpha_n})}{D_{\alpha_n}} + \frac{O(E(V^2 \mid \mathcal{F}_{\alpha_n}))}{D^2_{\alpha_n}}.$$

Thus our first goal is to estimate the conditional expectations of V and V^2.

First, the factor p associated to the stopping time T (through the stopping time R) in lemma 8.5.4 is of order q. Indeed, if $\alpha_n < k < T$, then

$$\frac{C_k + 1}{C_k + E_k + 1} \leq 2q.$$

Therefore, using the estimate of lemma 8.5.4 (we assume $q < Cst$),

$$E(V(V-1) \mid \mathcal{F}_{\alpha_n}) \leq \frac{4q}{(1-2q)^2} \leq 5q,$$

moreover $E(V \mid \mathcal{F}_{\alpha_n}) \leq 3/2$ and $E(V^2 \mid \mathcal{F}_{\alpha_n}) \leq 3/2$. Finally

$$E(\ln \frac{\tilde{D}_{n+1}}{\tilde{D}_n} \mid \mathcal{F}_{\alpha_n}) = \frac{E(E_{\alpha_n} + C_{\alpha_{n+1}\wedge T} \mid \mathcal{F}_{\alpha_n})}{D_{\alpha_n} E_{\alpha_n}} + r^0_n$$

where

$$|r^0_n| \leq \frac{1}{D^2_{\alpha_n}} \leq \frac{1}{(D_{j_0} + n)^2}.$$

Similarly, apply lemma 8.5.4 to estimate $V_{\beta_n,E,D,C}$: if $\alpha_n < k \leq T$, then

$$E(\ln \frac{\tilde{E}_{n+1}}{\tilde{E}_n} \mid \mathcal{F}_{\beta_n \wedge T}) = \frac{E(D_{\beta_n \wedge T} + F_{\alpha_{n+1}\wedge T} \mid \mathcal{F}_{\beta_n\wedge T})}{E_{\alpha_n} D_{\beta_n \wedge T}} + r^1_n$$

where

$$|r^1_n| \leq \frac{3/2}{(D_{j_0} + n)^2}.$$

Note that

$$|E(\frac{1}{D_{\beta_n \wedge T}} - \frac{1}{D_{\alpha_n}} \mid \mathcal{F}_{\alpha_n})| \leq |\frac{E(V_{\alpha_n,E,D,C} \mid \mathcal{F}_{\alpha_n})}{D_{\alpha_n} D_{\beta_n \wedge T}}| \leq \frac{3/2}{(D_{j_0} + n)^2},$$

and that

$$\frac{E(E_{\alpha_{n+1}\wedge T} - E_{\alpha_n} \mid \mathcal{F}_{\beta_n\wedge T})}{D_{\alpha_n}E_{\alpha_n}} \le \frac{E(F_{\alpha_{n+1}\wedge T} + D_{\beta_n\wedge T} \mid \mathcal{F}_{\beta_n\wedge T})}{D_{\alpha_n}E_{\alpha_n}D_{\beta_n\wedge T}} \le \frac{2}{(D_{j_0}+n)^2}.$$

To recap, we obtain

$$E(\ln\frac{\tilde{E}_{n+1}}{\tilde{D}_{n+1}} - \ln\frac{\tilde{E}_n}{\tilde{D}_n} \mid \mathcal{F}_{\alpha_n}) = \frac{E(R_{\alpha_{n+1}\wedge T} \mid \mathcal{F}_{\alpha_n})}{D_{\alpha_n}E_{\alpha_n}} + r_n,$$

where $|r_n| \le \frac{8}{(D_{j_0}+n)^2}$.

Let

$$U_{n+1} = \ln\frac{\tilde{E}_{n+1}}{\tilde{D}_{n+1}} - \ln\frac{\tilde{E}_n}{\tilde{D}_n} - \frac{R_{\alpha_{n+1}\wedge T}}{D_{\alpha_n}E_{\alpha_n}}\ ,\ \epsilon_{n+1} = U_{n+1} - E(U_{n+1} \mid \mathcal{F}_{\alpha_n}).$$

We wish to bound $Var(U_{n+1} \mid \mathcal{F}_{\alpha_n})$. Observe that

$$\begin{aligned} Var(\ln\frac{\tilde{D}_{n+1}}{\tilde{D}_n} \mid \mathcal{F}_{\alpha_n}) &\le E((\ln(1+\frac{V_{\alpha_n,E,D,C}}{\tilde{D}_n}) - \ln(1+\frac{1}{\tilde{D}_n}))^2 \mid \mathcal{F}_{\alpha_n}) \\ &\le \frac{E((V_{\alpha_n,E,D,C}-1)^2 \mid \mathcal{F}_{\alpha_n})}{\tilde{D}_n^2} \le \frac{5q}{(D_{j_0}+n)^2}. \end{aligned}$$

We can similarly bound the conditional variance of $\ln(\tilde{E}_{n+1}/\tilde{E}_n)$.

Now observe that

$$Var(\frac{R_{\alpha_{n+1}\wedge T}}{D_{\alpha_n}E_{\alpha_n}} \mid \mathcal{F}_{\alpha_n}) \le \frac{5q^2}{(D_{j_0}+n)^2},$$

since

$$-2qE_{\alpha_{n+1}\wedge T} \le -C_{\alpha_{n+1}\wedge T} \le R_{\alpha_{n+1}\wedge T} - R_{\alpha_n} \le F_{\alpha_{n+1}\wedge T} \le 2qE_{\alpha_{n+1}\wedge T}$$

and

$$E(E^2_{\alpha_{n+1}\wedge T} \mid \mathcal{F}_{\alpha_n}) \le E^2_{\alpha_n} + 4E_{\alpha_n} + 2.$$

In summary, $E(\epsilon_{n+1} \mid \mathcal{F}_{\alpha_n}) = 0$, $E(\epsilon^2_{n+1} \mid \mathcal{F}_{\alpha_n}) \le \frac{Cst.q}{(D_{j_0}+n)^2}$, $|r_n| \le \frac{Cst}{(D_{j_0}+n)^2}$, (observe that $r_n = E(U_{n+1} \mid \mathcal{F}_{\alpha_n})$). This completes the proof.

□

Let us define, for all $n \in \mathbb{N}$,

$$s_n = \frac{E_n - D_n}{D_n}.$$

The following lemma makes use an Abel transform :

Lemma 8.6.4 *Let us define, for all $r \in \mathbb{N}$, the stopping time*

$$W_r = \inf\{n \geq r/\ s_n > \lambda\}.$$

Suppose $\mu \in]1, Cst[$, $\lambda \in]0, Cst[$ and $\epsilon \in]0, Cst(\mu)[$. Then, on $\Gamma_{1,i_0,\epsilon}$, $\alpha_{n+1} < T_1 \wedge W_{\alpha_k}$ implies

$$\sum_{j=k}^{n} \frac{(BC)_{\alpha_k,\alpha_{j+1}}}{D_{\alpha_j} E_{\alpha_j}} \leq 2\tilde{p}_{\alpha_k}\epsilon.$$

PROOF: Let us write, for all $i \in \mathbb{N}$, $N_i = (BC)_{\alpha_k,i}$, and $r = \alpha_k$. Using an Abel transform, we obtain for all $s > r$,

$$\sum_{i=r+1}^{s} \frac{1_{\{X_{i-1}=B, X_i=C\}}}{D_i} = \sum_{i=r+1}^{s} \frac{N_i - N_{i-1}}{D_i} = \sum_{i=r}^{s-1} \frac{N_i}{D_i(D_i+1)} 1_{X_{i+1}=D} + \frac{N_s}{D_{s+1}}.$$

Observe on the other hand that, for all $j \geq k$, $N_{\alpha_j} = N_{\alpha_j - 1}$.

It follows that, for all $a > 1$, assuming λ, $\mu \in]1, Cst(a)[$ and $\epsilon \in]0, Cst(\mu, a)[$, $\alpha_{n+1} < T_1 \wedge W_{\alpha_k}$ implies

$$\begin{aligned}
\sum_{j=k}^{n} \frac{N_{\alpha_{j+1}}}{D_{\alpha_j} E_{\alpha_j}} &\leq a \sum_{j=k}^{n} \frac{N_{\alpha_{j+1}}}{D_{\alpha_j}(D_{\alpha_j}+1)} \leq a^2 \sum_{j=k}^{n} \frac{N_{\alpha_{j+1}-1}}{D_{\alpha_{j+1}-1}(D_{\alpha_{j+1}-1}+1)} \\
&\leq a^2 \sum_{j=\alpha_k+1}^{\alpha_{n+1}} \frac{1_{\{X_{j-1}=B, X_j=C\}}}{D_j} \leq a^2 \sum_{j=\alpha_k+1}^{\alpha_{n+1}} \frac{1_{\{X_{j-1}=B, X_j=C\}}}{C_j} \frac{C_j}{D_j} \leq a^3 \tilde{p}_{\alpha_k}\epsilon.
\end{aligned}$$

□

Lemma 8.6.5 $\Gamma_0 \subset \cup_{i_0 \in \mathbb{N}}(\{T_2 = +\infty\} \cap \Gamma_{1,i_0,\epsilon})$.

PROOF: It suffices to show that on $\Gamma_{1,i_0,\epsilon}$, if $\mu \in]1, Cst[$ and $\epsilon \in]0, Cst(\mu)[$, $F_{j_0} > \mu C_{j_0}$ implies

$$P(\Gamma_0^c \cup \{T_1 < +\infty\} \mid \mathcal{F}_{j_0}) \geq 1/2.$$

Assume $\mu \in]1, Cst[$ and $\epsilon \in]0, Cst(\mu)[$, Cst and $Cst(\mu)$ being as small as expected.

Suppose $F_{j_0} \geq \mu C_{j_0}$. If $X_{j_0} \leq C$, then $C_{k_0} \leq e^{\epsilon} C_{j_0}$ (if $k_0 < T_1$) where k_0 denotes the first visit time to state D after time j_0. Therefore we can suppose $X_{j_0} = D$ without loss of generality.

Let us define the sequence $(\alpha_n)_{n \in \mathbb{N}}$ starting at $\alpha_0 = j_0$.

Let us define, for all $n \in \mathbb{N}$,

$$F_n^0 = F_{j_0} + (EF)_{j_0,n} \leq F_n, \ R_n^0 = D_n + F_n^0 - (C_n + E_n).$$

Observe that

$$C_n + E_n \geq D_n + (BC)_n + (FE)_n - 1,$$

which implies

$$F_n^0 \le F_{j_0} + C_n + E_n - D_n + 3.$$

Therefore

$$\frac{F_n^0}{D_n + F_n^0} \le \frac{F_{j_0} + C_n + E_n - D_n + 3}{F_{j_0} + C_n + E_n} \le \frac{F_{j_0} + 3}{D_{j_0} + F_{j_0}} + \frac{C_n}{E_n} + s_n \le 3\tilde{p}_{j_0} + s_n.$$

We need an estimate of

$$\ln \frac{E_{\alpha_{n+1}}}{D_{\alpha_{n+1}}} - \ln \frac{E_{\alpha_n}}{D_{\alpha_n}}.$$

This variation is greater than the one that would exist by taking for all $n \in \mathbb{N}$, between time α_n and α_{n+1}, $F_{\alpha_n}^0/(D_{\alpha_n}+F_{\alpha_n}^0)$ for the probability to go to F starting from E.

Using lemma 8.6.3, this implies that

$$E(\ln \frac{E_{\alpha_{n+1}}}{D_{\alpha_{n+1}}} - \ln \frac{E_{\alpha_n}}{D_{\alpha_n}} \mid \mathcal{F}_{\alpha_n}) \ge \frac{E(D_{\alpha_{n+1}\wedge T_1} + F_{\alpha_n}^0 - (E_{\alpha_n} + C_{\alpha_{n+1}\wedge T_1}) \mid \mathcal{F}_{\alpha_n})}{D_{\alpha_n} E_{\alpha_n}} + O(\frac{1}{(D_{j_0} + n)^2}).$$

Now observe that

$$D_{\alpha_{n+1}\wedge T_1} + F_{\alpha_n}^0 - (E_{\alpha_n} + C_{\alpha_{n+1}\wedge T_1}) \ge R_{\alpha_n}^0 + (C_{\alpha_n} - C_{\alpha_{n+1}\wedge T_1}),$$

and, for all $k \le n$,

$$C_{\alpha_k,\alpha_n} + E_{\alpha_k,\alpha_n} - (D_{\alpha_k,\alpha_n} + (BC)_{\alpha_k,\alpha_n} + (FE)_{\alpha_k,\alpha_n}) \le 1,$$

which implies

$$R_{\alpha_n}^0 \ge R_{\alpha_k} - (BC)_{\alpha_k,\alpha_n} - 2.$$

In summary, we can fit lemma 8.6.3 and prove that there exists sequences of random variables $(\epsilon_n)_{n\in\mathbb{N}}$ and $(r_n)_{n\in\mathbb{N}}$ such that, if $s_{\alpha_n} \le 4\tilde{p}_{j_0}$ and $\alpha_{n+1} < T_1$,

$$\ln \frac{E_{\alpha_{n+1}}}{D_{\alpha_{n+1}}} - \ln \frac{E_{\alpha_n}}{D_{\alpha_n}} \ge \frac{R_{\alpha_k} - (BC)_{\alpha_k,\alpha_n}}{D_{\alpha_n} E_{\alpha_n}} + O(\frac{1}{(D_{j_0} + n)^2}) + \epsilon_{n+1} + r_n,$$

for all $k \le n$, with $E(\epsilon_{n+1} \mid \mathcal{F}_{\alpha_n}) = 0$, $E(\epsilon_{n+1}^2 \mid \mathcal{F}_{\alpha_n}) \le \frac{Cst.\tilde{p}_{j_0}}{(n+D_{j_0})^2}$ and $|r_n| \le \frac{Cst}{(n+D_{j_0})^2}$.

Now, let us recall that, according to lemma 8.6.4, for all n, $k \in \mathbb{N}$ such that $n \ge k$, $\alpha_{n+1} < T_1 \wedge W_{\alpha_k}$,

$$\sum_{j=k}^{n} \frac{(BC)_{\alpha_k,\alpha_{j+1}}}{D_{\alpha_j} E_{\alpha_j}} \le 2\tilde{p}_{\alpha_k}\epsilon.$$

Also observe that, if $k \in \mathbb{N}$ is such that for all $i \leq k$, $s_{\alpha_i} \leq 4\tilde{p}_{j_0}$, then

$$\sum_{i=0}^{k} \left| \frac{1}{D_{\alpha_i}^2} - \frac{1}{D_{\alpha_i} E_{\alpha_i}} \right| \leq \frac{5\tilde{p}_{j_0}}{D_{j_0} - 1}.$$

and

$$\sum_{i=0}^{k} \left| \frac{1}{D_{\alpha_i}^2} - \frac{1}{(D_{j_0}+i)^2} \right| \leq 2 \sum_{i=j}^{k} \frac{D_{\alpha_i} - (D_{j_0}+i)}{D_{\alpha_i}(D_{j_0}+i)^2} \leq 2 \sum_{i=0}^{k} \frac{C_{\alpha_i}}{D_{\alpha_i}} \frac{1}{(D_{j_0}+i)^2} \leq \frac{3\tilde{p}_{j_0}}{D_{j_0} - 1}.$$

Given $\nu \in]0, 1/2[$, let $\Delta = \{\sup_{n \geq 0} | \sum_{k=0}^{n} \epsilon_{k+1} 1_{s_{\alpha_k} \leq 4\tilde{p}_{j_0}} | \leq D_{j_0}^{-\nu}\}$, then Doob's inequality implies that for enough large D_{j_0}, $P(\Delta \mid \mathcal{F}_{j_0}) \geq 1/2$.

We now assume that we belong to Δ, and prove that $\ln E_{\alpha_n}/D_{\alpha_n} \geq \tilde{p}_{j_0}/2$ for enough large n, which will complete the proof of this lemma.

Given, $n \in \mathbb{N}$, let $k_0 = \inf\{k \in [0, n]/\ s_{\alpha_{n-k}} > 4\tilde{p}_{j_0}\}$, and let $n_0 = n - k_0$. We deal with three cases :

1) n_0 does not exist. Then

$$\begin{aligned} \ln \frac{E_{\alpha_n}}{D_{\alpha_n}} & \geq \ln \frac{E_{j_0}}{D_{j_0}} + \sum_{k=0}^{n} \frac{R_{j_0}}{D_{\alpha_n} E_{\alpha_n}} - 2\tilde{p}_{j_0}\epsilon - 2D_{j_0}^{-\nu} \\ & \geq \frac{E_{j_0} - D_{j_0}}{D_{j_0}} + \frac{F_{j_0} + D_{j_0} - (C_{j_0} + E_{j_0})}{D_{j_0}} - 3\tilde{p}_{j_0}\epsilon - 2D_{j_0}^{-\nu} \geq \frac{\tilde{p}_{j_0}}{2}, \end{aligned}$$

using the fact that $\tilde{p}_{j_0} \geq 1/4 \ln D_{j_0}$.

2) $n_0 = n$. Then $s_{\alpha_n} \geq 4\tilde{p}_{j_0}$.

2) n_0 exists and $n_0 < n$. Then $s_{\alpha_{n_0}} > 4\tilde{p}_{j_0}$, which implies $s_{\alpha_{n_0+1}} \geq 2\tilde{p}_{j_0}$ since $D_{\alpha_{n_0+1}} \leq D_{\alpha_{n_0}} + C_{\alpha_{n_0+1}} - C_{\alpha_{n_0}} + 1 \leq D_{\alpha_{n_0}} + \mu\tilde{p}_{j_0} E_{\alpha_{n_0}}$. Now

$$\ln \frac{E_{\alpha_n}}{D_{\alpha_n}} \geq \ln \frac{E_{\alpha_{n_0+1}}}{D_{\alpha_{n_0+1}}} - (1+\epsilon)\frac{C_{\alpha_{n_0+1}}}{D_{\alpha_{n_0+1}}} - 2\tilde{p}_{j_0}\epsilon - 2D_{j_0}^{-\nu} \geq \frac{3\tilde{p}_{j_0}}{2}.$$

□

Let us now prove that, for all $\mu > 1$, $\epsilon < Cst(\mu)$,

$$P(\Gamma_0 \cap \Gamma_{1,i_0,\epsilon} \cap \{\lim_{n \to \infty} E_{\alpha_n}/D_{\alpha_n} = 1\} \cap \{T_2 = \infty\}) = 0.$$

This will imply that $P(\Gamma_0) = 0$ since, for all $\mu > 1$, $\Gamma_0 \subset \cup_{i_0 \in \mathbb{N}}(\{T_2 = \infty\} \cap \Gamma_{1,i_0,\epsilon})$, and $E_{\alpha_n}/D_{\alpha_n} \to 1$ on $\Gamma_0 \cap \{T_2 = \infty\}$.

We now consider the sequence $(\alpha_n)_{n \in \mathbb{N}}$ starting at i_0.

Let us first observe that, for all j_0, $n \in \mathbb{N}$ such that $i_0 \leq j_0 \leq n < T_2$, $F_n \leq \mu C_n$ and $p_n \leq \mu\tilde{p}_{j_0}$ imply $F_n/(D_n + F_n) \leq \mu C_n/(D_n + \mu C_n) \leq \mu C_n/(E_n + (\mu - 1)C_n) \leq \mu^2 \tilde{p}_{j_0}/(1 - \mu\tilde{p}_{j_0}) \leq \mu^3 \tilde{p}_{j_0}$ for $\epsilon < Cst(\mu)$. The same proof implies that $\tilde{p}_n \leq \mu^2 p_n/(1 - p_n) \leq \mu^3 p_n$, which also implies $p_n \leq \mu^4 p_{j_0}$.

Let us now apply lemma 8.5.3 with $T := T_2$, $\delta_n := \alpha_n$ (such that $\alpha_0 = i_0$), $U_n := E_n/D_n$, $R_n := R_n = D_n + F_n - (C_n + E_n)$, $S_n = D_n E_n$, $v(\alpha_m) := D_{\alpha_m}$ for all $m \in \mathbb{N}$, $L := C$, $M := D$, $N := E$, $g_1 := 1$, $g_3 \in]0, 1/2[$ and $i := 1$.

First observe that $U = T_2 \wedge \alpha_{m+\lfloor g_3 D_{\alpha_m} \rfloor} \leq S_1 \wedge V$, where S_1 and V denote the stopping times defined in lemma 8.5.2. Indeed, $U \leq V$ since $p_n \leq \mu^4 p_{\alpha_m} < 2p_{\alpha_m}$ for $n < T_2$ (if $\mu < Cst$) and, for all $n \in [\alpha_m, U[$, $E_n \leq \sum_{k=0}^n 1_{X_{k-1}=D, X_k=E} + \sum_{k=0}^n 1_{X_{k-1}=F, X_k=E} \leq F_n + m + D_{i_0} + \lfloor g_3 D_{\alpha_m} \rfloor \leq F_n + 3D_{\alpha_m}/2$, which implies $E_n \leq 5D_{\alpha_m}/3$ (if $\epsilon < Cst$) and $p_n = C_n/(C_n + E_n) \geq C_{\alpha_m}/(C_{\alpha_m} + 5D_{\alpha_m}/3) \geq p_{\alpha_m}/2$. And $U \leq S_1$, since when $n < U$ a similar argument implies $D_n - C_n \leq 3D_{\alpha_m}/2$, hence $D_n < 2D_{\alpha_m}$.

Applying lemma 8.6.3 for $\mathcal{M}$, and for $\mathcal{M}'$ in the case $n \geq m + \lfloor g_3 D_{\alpha_m} \rfloor$, we obtain that condition **a)** of lemma 8.5.3 is satisfied in these cases, with $m(\alpha_m, n) = Cst p_{\alpha_m}/(D_{\alpha_m} + n - m)^2$, $n(\alpha_m, n) = Cst/(D_{\alpha_m} + n - m)^2$

On the other hand $U \leq S_1 \wedge V$ implies that assumption **a)** of lemma 8.5.3 is satisfied in the case $n \in [m, m + \lfloor g_3 D_{\alpha_m} \rfloor]$ and $\alpha'_n < T'_2$ with $u(\alpha_m) = Cst\sqrt{p_{\alpha_m}}/D_{\alpha_m}^{3/2}$. Indeed,

$$\ln \frac{\tilde{E}'_{n+1}}{\tilde{D}'_{n+1}} - \ln \frac{\tilde{E}'_n}{\tilde{D}'_n} \geq \frac{D'_{\alpha'_{n+1} \wedge T'_2} + F'_{\alpha'_{n+1} \wedge T'_2}}{D'_{\alpha'_n} E'_{\alpha'_n}} - \frac{1}{D'_{\alpha'_n}(1 - p'_{\alpha'_{n+1} \wedge T'_2}(1 - \frac{g_2}{\sqrt{2p_{\alpha_m} D_{\alpha_m}}}))} + \epsilon'_{n+1} + r'_n$$
$$\geq \frac{R'_{\alpha'_{n+1} \wedge T'_2}}{D'_{\alpha'_n} E'_{\alpha'_n}} + \frac{g_2}{8\sqrt{2}} \sqrt{\frac{p_{\alpha_m}}{D_{\alpha_m}}} \frac{1}{D_{\alpha_m}} + \epsilon'_{n+1} + r'_n$$

Easy calculation shows that assumption **b)** holds with $C_1 = Cst > 0$ (for the sum of $n(\alpha_m, n)$, note that $p_{\alpha_m} \geq 1/4 \ln D_{\alpha_m}$).

Assumption **c)** holds since $\mathcal{M}' \gg \mathcal{M}$ implies on one hand $E'_{\alpha'_{n+1}} \geq E_{\alpha_{n+1}}$, $D'_{\alpha'_{n+1}} \leq D_{\alpha_{n+1}}$, $C'_{\alpha'_{n+1}} \leq C_{\alpha_{n+1}}$, $F'_{\alpha'_{n+1}} \geq F_{\alpha_{n+1}}$, and on the other hand $R_t \leq R'_{t'}$ for all t, $t' \in \mathbb{N}$ satisfying $X_t = X'_{t'}$, $F'_{t'} \geq F_t$ and $C'_{t'} \leq C_t$.

Assumption **d)** holds by lemma 8.6.4 :

$$\sum_{\alpha_{n+1} < T} \frac{R^-_{\alpha_{n+1}}}{D_{\alpha_n} E_{\alpha_n}} \geq - \sum_{\alpha_{n+1} < T} \frac{(BC)_{i_0, \alpha_{n+1}}}{D_{\alpha_n} E_{\alpha_n}} - (1 + \epsilon) \frac{B_{i_0}}{D_{i_0} - 1} > -\infty.$$

Let us prove that **e)**(i) holds :

$$D_{\alpha_n} E_{\alpha_n} g_n = \frac{U'_{\alpha'_n}}{U_{\alpha_n}} D^2_{\alpha_n} \frac{E_{\alpha_n}}{D_{\alpha_n}} \geq D'^2_{\alpha'_n} \frac{E'_{\alpha'_n}}{D'_{\alpha'_n}} = D'_{\alpha'_n} E'_{\alpha'_n} = E'^2_{\alpha'_n} \frac{D'_{\alpha'_n}}{E'_{\alpha'_n}} \geq \frac{E^2_{\alpha_n}}{U'_{\alpha'_n}/U_{\alpha_n}} \frac{D_{\alpha_n}}{E_{\alpha_n}} = \frac{D_{\alpha_n} E_{\alpha_n}}{g_n}$$

implies $h_n \in [g_n^{-1}, g_n]$.

Conclusion of lemma 8.5.3 completes the proof.

8.7 Proof of $P(\Omega_{A,1}) = 0$

8.7.1 Proof of $H_n/D_n \to 1$ on $\Omega_{A,1}$

Let us define, for all $\mu > 1$, a, $b \in]0,1[$ and $i_0 \in \mathbb{N}$ the stopping time

$$T_1 = \inf\{n \geq i_0 / \ E_n/(C_n + E_n) \notin [\mu^{-1}a, \mu a] \text{ or } G_n/(G_n + I_n) \notin [\mu^{-1}b, \mu b]\}.$$

Then

$$\Omega_{A,1} \subset \cup_{i_0 \in \mathbb{N}, a,b \in]0,1[} \{T_1 = +\infty\}.$$

Lemma 8.7.1

$$\Omega_{A,1} \subset \{\frac{B_n}{E_n}, \frac{F_n}{E_n}, \frac{F_n}{G_n}, \frac{J_n}{I_n} \to 0\} \cap \{\sum \frac{1_{X_n=F}}{F_n}\frac{B_n}{E_n+G_n} < \infty\} \cap \{\sum \frac{1_{X_n=F}}{F_n}\frac{J_n}{E_n+G_n} < \infty\}.$$

PROOF: Let us first prove for instance that $F_n/G_n \to 0$ on $\Omega_{A,1}$: remark that

$$\Omega_{A,1} \subset \Upsilon_{F,G} = \{\sum_{n \in \mathbb{N}} \frac{1_{X_n=G, X_{n+1}=F}}{G_n} < \infty\} = \{\sum_{n \in \mathbb{N}} \frac{F_n}{G_n(F_n + H_n)} 1_{X_n=G} < \infty\}.$$

On the other hand, fix $\mu > 1$, a, b, $\in]0,1[$ and $i_0 \in \mathbb{N}$. On $\Omega_{A,1}$, for all $p \geq i_0$,

$$\sum_{n \geq p} \frac{F_n 1_{X_n=G}}{G_n(F_n + H_n)} \geq \sum_{n \geq p} \frac{F_p 1_{X_n=G}}{G_n(F_p + H_n)} \geq \mu^{-1} b \sum_{i \in \mathbb{N}} \frac{F_p}{(F_p + G_p + i)^2} \geq \mu^{-1} b \frac{F_p}{F_p + G_p}.$$

We now make use of F_n/E_n, ...$\to$ 0. Let us prove the second part of the inclusion, the proof for the third being simlar, taking calculations symetrically. Applying Borel-Cantelli lemma, on $\Omega_{A,1} \subset \Upsilon_{B,A}$,

$$\begin{aligned}
\infty > \sum_{n \in \mathbb{N}} \frac{1_{X_n=B}}{A_n + C_n} &\approx \sum_{n \in \mathbb{N}} \frac{1_{X_n=B, X_{n+2}=D}}{D_n} \frac{B_n + D_n}{C_n} \approx \sum_{n \in \mathbb{N}} \frac{1_{X_n=B, X_{n+2}=D}}{D_n} \\
&\approx \sum_{n \in \mathbb{N}} \frac{1_{X_n=D, X_{n+2}=B}}{D_n} \approx \sum_{n \in \mathbb{N}} \frac{1_{X_n=D, X_{n+2}=B}}{D_n} \frac{B_n + D_n}{C_n} \approx \sum_{n \in \mathbb{N}} \frac{1_{X_n=D}}{D_n} \frac{B_n}{C_n + E_n} \\
&\approx \sum_{n \in \mathbb{N}} \frac{1_{X_n=D, X_{n+2}=F}}{F_n} \frac{B_n}{E_n} \frac{D_n + F_n}{D_n} \approx \sum_{n \in \mathbb{N}} 1_{X_n=D, X_{n+2}=F} \frac{B_n}{E_n F_n}.
\end{aligned}$$

Given $p_0 \in \mathbb{N}$ such that $X_{p_0} = D$, let q_0 be the first $k \geq p_0$ such that $X_k = F$. Then

$$\sum_{n \geq q_0} 1_{X_n=F, X_{n+2}=D} \frac{B_n}{E_n F_n} \leq \sum_{n \geq p_0} 1_{X_n=D, X_{n+2}=F} \frac{B_n}{E_n F_n}.$$

Therefore

$$\infty > \sum_{n \in \mathbb{N}} \frac{1_{X_n=F, X_{n+2}=D}}{F_n} \frac{B_n}{E_n} \frac{D_n + F_n}{D_n} \approx \sum_{n \in \mathbb{N}} \frac{1_{X_n=F}}{F_n} \frac{B_n}{E_n + G_n}.$$

□

Let us define, for all $n \in \mathbb{N}$, the time κ_n of n-th return to state F. Let us define the stopping times

$$U_1 = \inf\{n \geq i_0 /\ F_n/G_n > \epsilon \text{ or } F_n/E_n > \epsilon\},$$

and $T_2 = T_1 \wedge U_1$.

Lemma 8.7.2 *Let $a_1 < Cst(\mu, b)$. Assume $\mu \geq 1 + \epsilon$ and $\kappa_n \geq i_0$. Then*

$$P(\{G_{\kappa_{(a_1+1)n} \wedge T_2} > \mu G_{\kappa_n}\} \mid \mathcal{F}_{\kappa_n}) \leq \exp(-Cst.(\mu - 1)\mu^{-5} bn).$$

PROOF: Let $k \in \mathbb{N}$ be the first $i \in \mathbb{N}$ such that $G_i > \mu G_{\kappa_n}$ or $i = T_2$. For all $j \in [\kappa_n, k[$,

$$H_j \leq \mu b^{-1} G_j \leq \mu^2 b^{-1} G_{\kappa_n} \leq \mu^3 H_{\kappa_n},$$

which implies

$$\frac{F_j}{F_j + H_j} \geq \frac{n}{n + \mu^3 H_{\kappa_n}} \geq \mu^{-4} \frac{n}{H_{\kappa_n}}$$

since $\epsilon < \mu - 1$.

We need to prove that with large probability, there is more than $a_1 F_{\kappa_n}$ visits to state F between time κ_n and time k. Let us define the event $\Delta = \{F_{k-1} \geq (1 + a_1) F_{\kappa_n}\}$, and apply lemma 8.3.1 with $N = (\mu - 1) G_{\kappa_n} - 1$, $p = \mu^{-4} n / H_{\kappa_n}$, $a = p/2$: observe that $pN \geq (\mu - 1)\mu^{-5} bn - 1$. Accordingly, choosing $a_1 < (\mu - 1)\mu^{-5} b/2$ implies $P(\Delta) \leq \exp(-Cst.(\mu - 1)\mu^{-5} nb)$.

□

Let us define, given $i_0 \in \mathbb{N}$ and $a_1 < Cst(\mu, b)$, such that $X_{i_0} = F$, the stopping times

$$U_2 = \inf\{n \geq i_0 /\ \exists V \in \{D, E, G, H\} /\ \exists k \in \mathbb{N} /\ \kappa_k \geq i_0 \text{ and } \kappa_{(1+a_1)k} \leq n \text{ and } V_{\kappa_{(1+a_1)k}} \geq \mu V_{\kappa_k}\}$$

and $T_3 = T_2 \wedge U_2$.

Then, for all $\mu > 1$ and $\epsilon < \mu - 1$,

$$\Omega_{A,1} \subset \cup_{i_0 \in \mathbb{N}, a, a_1, b \in]0,1[} \{T_3 = +\infty\}.$$

Lemma 8.7.3 *Let $i_0 \in \mathbb{N}$, a, $b \in]0, 1[$, $\mu \in]1, 2[$ and $\epsilon \in]0, \mu - 1[$. There exist sequences of $\mathcal{F}_{\kappa_n}$-mesurable random variables $(\tilde{D}_n)_{n \geq k_0}$, $(\tilde{H}_n)_{n \geq k_0}$, $(\epsilon_n)_{n \geq m}$, $(r_n)_{n \geq k_0}$ such that, if $\kappa_n \leq T_3$, then*

1) $\tilde{D}_n = D_{\kappa_n}, \tilde{H}_n = H_{\kappa_n}$,

2) $\ln \dfrac{\tilde{H}_{n+1}}{\tilde{D}_{n+1}} - \ln \dfrac{\tilde{H}_n}{\tilde{D}_n} = \dfrac{G_{\kappa_n} + I_{\kappa_n} - (C_{\kappa_n} + E_{\kappa_n})}{n(G_{\kappa_n} + E_{\kappa_n})} + \epsilon_{n+1} + r_n$,

3) $E(\epsilon_{n+1} \mid \mathcal{F}_{\kappa_n}) = 0$, $E(\epsilon_{n+1}^2 \mid \mathcal{F}_{\kappa_n}) \leq Cst. \dfrac{a^{-2} + b^{-2}}{n^2}$, $|r_n| \leq Cst. \dfrac{a^{-2} + b^{-2}}{n^2}$.

PROOF: The proof is the same as for lemma 8.6.3 (using remark 8.5.2 after lemma 8.5.4). The only difference is that we take G_{κ_n}, H_{κ_n}... instead of $G_{\kappa_{n+1}\wedge T_3}$, $H_{\kappa_{n+1}\wedge T_3}$... at the right hand side of the inequality since, for instance for G,

$$E(G_{\kappa_{n+1}\wedge T_3} - G_{\kappa_n} \mid \mathcal{F}_{\kappa_n}) \le Cst(\mu)\frac{H_{\kappa_n}}{n}.$$

□

Let, for all $k \in \mathbb{N}$,

$$R_k = G_k + I_k - (C_k + E_k).$$

Lemma 8.7.4

$$\Omega_{A,1} \subset \{\lim_{n\to\infty} H_n/D_n = 1\} \cap \{\sum_{n\in\mathbb{N}} \frac{|R_{\kappa_n}|}{n(E_{\kappa_n} + G_{\kappa_n})} < +\infty\}.$$

PROOF: Observe that, for all $k \in \mathbb{N}$, $H_k \le G_k + I_k \le F_k + H_k + J_k$ and $D_k \le C_k + E_k \le B_k + D_k + F_k$.

Therefore, applying lemma 8.7.3, when $\kappa_n < T_3$,

$$\ln\frac{\tilde{H}_{n+1}}{\tilde{D}_{n+1}} - \ln\frac{\tilde{H}_n}{\tilde{D}_n} = \frac{\tilde{H}_n - \tilde{D}_n}{n(E_{\kappa_n} + G_{\kappa_n})} + \epsilon_{n+1} + r_n + \frac{S_{\kappa_n}}{n(E_{\kappa_n} + G_{\kappa_n})},$$

where $|S_{\kappa_n}| \le 2F_{\kappa_n} + B_{\kappa_n} + J_{\kappa_n}$. Hence, it follows from lemma 8.7.1 that

$$\sum \frac{|S_{\kappa_n}|}{n(E_{\kappa_n} + G_{\kappa_n})} < +\infty.$$

Now conditions of lemma 8.4.2 are fulfilled in the case $\lambda = \infty$, $y_n = \tilde{H}_n/\tilde{D}_n$, $s_n = r_n + S_{\kappa_n}/n(E_{\kappa_n} + G_{\kappa_n})$, $\gamma_n = D_{\kappa_n}/n(E_{\kappa_n} + G_{\kappa_n})$.

Hence

$$\begin{aligned}\Omega_{A,1} \cap \{T_3 = \infty\} \cap \{H_n/D_n \not\to 1\} \subset \quad & \{\liminf H_n/D_n > 1 \text{ or } \limsup H_n/D_n < 1\} \\ & \cap\{\ln\frac{H_{\kappa_n}}{D_{\kappa_n}} \asymp \sum_{k\le n} \frac{H_{\kappa_k} - D_{\kappa_k}}{k(E_{\kappa_k} + G_{\kappa_k})}\}.\end{aligned}$$

Assume for instance we belong, given $\delta > 1$, to $\{\liminf H_n/D_n > \delta\} \cap \{T_3 = \infty\}$. This implies that

$$\sum_{k\le n} \frac{H_{\kappa_k} - D_{\kappa_k}}{k(E_{\kappa_k} + G_{\kappa_k})} \succeq \frac{1-\delta^{-1}}{\beta + \delta^{-1}(1-\alpha)} \sum_{k\le n} \frac{1}{k} \succeq \frac{1-\delta^{-1}}{\beta + \delta^{-1}(1-\alpha)} \ln n,$$

since $\limsup (E_{\kappa_k} + G_{\kappa_k})/H_{\kappa_k} \le \delta^{-1}(1-\alpha) + \beta$.

On the other hand,

$$\ln n = \ln F_{\kappa_n} \asymp \sum_{k\le\kappa_n} \frac{1_{X_k=F}}{F_k} \succeq \sum_{k\le\kappa_n} \frac{1_{X_{k-1}=G, X_k=F}}{F_k} \asymp \beta \ln H_{\kappa_n},$$

the last estimate being a consequence of statement **(3)** of lemma 8.3.5 :

$$\sum_{k\leq\kappa_n}\frac{1_{X_{k-1}=G,X_k=F}}{F_k}\asymp\sum_{k\leq\kappa_n}\frac{1_{X_{k-1}=G}}{F_k+H_k}\asymp\beta\sum_{k\leq\kappa_n}\frac{1_{X_{k-1}=G}}{G_k}\asymp\beta\ln G_{\kappa_n}.$$

Hence

$$\ln H_{\kappa_n}-\ln D_{\kappa_n}\succeq(1-\delta^{-1})\beta\ln H_{\kappa_n}/(\beta+\delta^{-1}(1-\alpha)).$$

Thus $\limsup\ln D_n/\ln H_n<1$, which implies in particular that $\lim D_n/H_n=0$. We apply the same inequality with $\delta=\infty$, and obtain

$$\ln H_{\kappa_n}-\ln D_{\kappa_n}\succeq\ln H_{\kappa_n}\iff\limsup\ln D_n/\ln H_n=0.$$

This implies

$$\limsup\ln D_n/\ln F_n=0\ ,\ \ \limsup\ln E_n/\ln G_n=0.$$

Applying corollary 8.4.1 with $L:=G$, $M:=F$, $N:=E$, $O:=D$, we deduce that D is visited finitely often, which leads to a contradiction and completes the proof of $\Omega_{A,1}\subset\{\lim H_n/D_n=1\}$.

The fact that

$$\Omega_{A,1}\cap\{\lim H_n/D_n=1\}\cap\{T_3=\infty\}\subset\{\sum_{\kappa_n<T_3}\frac{|R_{\kappa_n}|}{n(E_{\kappa_n}+G_{\kappa_n})}<\infty\}$$

follows directly from statement **3)** of lemma 8.4.2.

□

8.7.2 End of the proof of $P(\Omega_{A,1})=0$

We now prove that, for all $i_0\in\mathbb{N}$, a, a_1, $b\in]0,1[$,

$$P(\{\lim H_n/D_n=1\}\cap\{T_3=\infty\})=0.$$

This implies $P(\Omega_{A,1})=0$ by lemma 8.7.4. Let us apply lemma 8.5.3 with $T:=T_3$, $\delta_n:=\kappa_n$, $U_n:=H_n/D_n$, $R_n:=R_n=G_n+I_n-(C_n+E_n)$, $S_n=F_n(C_n+E_n)$, $v(\kappa_m)=F_{\kappa_m}=m$ for all $m\in\mathbb{N}$, $L:=F$, $M:=G$, $N:=H$, $g_1:=1$, $g_3\in]0,1/2[$, and $i:=1$.

First observe that $U=T_3\wedge\kappa_{\lfloor(1+g_3)m\rfloor}\leq S_1\wedge V$ for $g_3<a_1$ and $\mu<2$. Indeed, $U\leq V$ since, on one hand, if $n<\kappa_{\lfloor(1+g_3)m\rfloor}$, $p_n=F_n/(F_n+H_n)\leq(1+g_3)F_{\kappa_m}/(F_{\kappa_m}+H_{\kappa_m})<2p_{\kappa_m}$ and, on the other hand, $n<U$ implies $p_n=F_n/(F_n+H_n)\geq F_{\kappa_m}/(F_{\kappa_m}+\mu H_{\kappa_m})>p_{\kappa_m}/2$, since $g_3<a_1$ and $\mu<2$. The same assumptions imply $U\leq S_1:G_{\kappa_{\lfloor(1+g_3)m\rfloor}}<\mu G_{\kappa_m}<2G_{\kappa_m}$.

Applying lemma 8.7.3 for $\mathcal{M}$, and for $\mathcal{M}'$ in the case $n \geq \lfloor (1+g_3)m \rfloor$, we obtain that condition **a)** of lemma 8.5.3 is satisfied in these cases, with $m(\kappa_m, n) = Cst(a^{-2}+b^{-2})/n^2$, $n(\kappa_m, n) = Cst(a^{-2}+b^{-2})/n^2$.

On the other hand, $U \leq S_1 \wedge V$ implies that **a)** is satisfied in the case $n \in [m, \lfloor (1+g_3)m \rfloor]$ and $\kappa'_n < T'_3$, with $u(\kappa_m) = Cst/m^{3/2}$. Indeed, in this case,

$$\ln \frac{\tilde{H}'_{n+1}}{\tilde{E}'_{n+1}} - \ln \frac{\tilde{H}'_n}{\tilde{E}'_n} \geq \frac{G'_{\kappa'_n} + I'_{\kappa'_n}}{n(E'_{\kappa'_n} + G'_{\kappa'_n})(1 - \frac{g_2}{\sqrt{2p_{\kappa_m} G_{\kappa_m}}})} - \frac{C'_{\kappa'_n} + E'_{\kappa'_n}}{n(E'_{\kappa'_n} + G'_{\kappa'_n})} + \epsilon'_{n+1} + r'_n$$

$$\geq \frac{R'_{\kappa'_n}}{n(E'_{\kappa'_n} + G'_{\kappa'_n})} + \frac{g_2}{4\sqrt{2}m^{3/2}} + \epsilon'_{n+1} + r'_n.$$

Easy calculation shows that condition **b)** holds with $C_1 = Cst(a,b)$. Assumption **c)** follows directly from the definition of the coupling, and assumption **d)** follows from lemma 8.7.4.

Let us now prove that **e)**(i) holds. Let, for all $n \in \mathbb{N}$, $q_n = \frac{G'_{\kappa'_n}/E'_{\kappa'_n}}{G_{\kappa_n}/E_{\kappa_n}} \geq 1$. Then

$$nE_{\kappa_n}(1 + G'_{\kappa'_n}/E'_{\kappa'_n}) \geq S'_{\kappa'_n} = n(E'_{\kappa'_n} + G'_{\kappa'_n}) \geq nG_{\kappa_n}(1 + E'_{\kappa'_n}/G'_{\kappa'_n})$$

implies $h_n \in [q_n^{-1}, q_n]$.

Let us now prove that $\ln q_n \leq 2\mu \sup(a^{-1}, b^{-1}) \ln g_n$, when g_n is close enough to 1.

Indeed,

$$\ln q_n = \ln G'_{\kappa'_n}/G_{\kappa_n} + \ln E_{\kappa_n}/E'_{\kappa'_n}.$$

On the other hand, $G'_{\kappa'_n} - G_{\kappa_n} \leq H'_{\kappa'_n} - H_{\kappa_n}$ implies

$$\ln \frac{G'_{\kappa'_n}}{G_{\kappa_n}} = \ln(1 + \frac{G'_{\kappa'_n} - G_{\kappa_n}}{G_{\kappa_n}}) \leq \ln(1 + \mu b^{-1}\frac{H'_{\kappa'_n} - H_{\kappa_n}}{H_{\kappa_n}}) \leq 2\mu b^{-1} \ln(1 + \frac{H'_{\kappa'_n} - H_{\kappa_n}}{H_{\kappa_n}}) = 2\mu b^{-1} \ln \frac{H'_{\kappa'_n}}{H_{\kappa_n}},$$

when g_n is close enough to 1.

A similar inequality holds for $\ln E_{\kappa_n}/E'_{\kappa'_n}$. This proves **e)**(i). Conclusion of lemma 8.5.3 completes the proof of $P(\Omega_{A,1}) = 0$.

8.8 Proof of $P(\Omega_{A,2}) = 0$

8.8.1 Proof of $I_n/D_n \to 1$ on $\Omega_{A,2}$

Let L, M and N be the integers following K. We assume, without loss of generality, that we belong to $\Upsilon_{L,K}$. Indeed, on $\Omega_{A,2}$, we belong to $\Upsilon_{F,G}$, $\Upsilon_{G,H}$ and $\Upsilon_{K,L}$, which implies by lemma 8.6.1 that either $\lim L_n/(L_n+N_n) > 0$ or we belong to $\Upsilon_{L,K}$. We proved in section 8.7 that the first case is impossible, since $P(\Omega_{F,1}) = 0$.

Let us define, for all $n \in \mathbb{N}$,

$$l(n) = \frac{E_n}{C_n + E_n}\ ,\ m(n) = \frac{H_n}{H_n + J_n}\ ,\ \tilde{l}(n) = \frac{E_n}{D_n}\ ,\ \tilde{m}(n) = \frac{H_n}{I_n}.$$

Let us define, given $\mu > 1$, a, $a' \in \mathbb{R}_*^+$, l, $m \in]0,1[$, the stopping time

$$\begin{aligned} T_1 = \inf\{n \geq i_0/ \quad & F_n \notin [\mu^{-1} a D_n^{l(n)} \mu a D_n^{l(n)}] \text{ or } G_n \notin [\mu^{-1} a' I_n^{m(n)}, \mu a I_n^{m(n)}] \\ & \text{or } l(n), \tilde{l}(n) \notin [\mu^{-1} l, \mu l] \text{ or } m(n), \tilde{m}(n) \notin [\mu^{-1} m, \mu m] \\ & \text{or } \exists k \in [i_0, n]/\ |l(n) - l(k)| > D_k^{-\nu} \text{ or } |m(n) - m(k)| > I_k^{-\nu}\}. \end{aligned}$$

Then lemmas 8.3.4 and 8.3.7 (applied with $P = F$, $Q = G$...) imply that, for all $\mu > 1$,

$$\Omega_{A,2} \subset \cup_{i_0 \in \mathbb{N}, a, a', \nu > 0, l, m \in]0,1[} \{T_1 = \infty\}.$$

Lemma 8.8.1 $P(\beta' \leq \alpha\} \cap \Omega_{A,2}) = 0$.

PROOF: It suffices to prove that, for all $i_0 \in \mathbb{N}$, l, $m \in]0,1[$ such that $\mu l < 1$ and $\mu m < 1$, the event $\Omega_{A,2} \cap \{\beta' \leq \alpha\} \cap \{T_1 = \infty\}$ is of probability 0. We now assume we belong to this event.

Let $m_0 \geq i_0$, then there exists $n \geq m_0$ such that $X_n = E$ and $D_n \geq I_n$, or $X_n = H$ and $D_n \leq I_n$. Indeed, assume for instance that $D_{m_0} \geq I_{m_0}$. At time n_0 of first return to state E after m_0 (which exists since we belong to $\Omega_{A,2}$), $D_{n_0} \geq I_{n_0}$ or $D_{n_0} < I_{n_0}$; in this second case, let $n_0' = \inf\{n \geq m_0/\ I_n = I_{n_0}\}$, and then $n_0'' = \inf\{n \geq n_0'/\ X_n = H\}$. Then $D_{n_0''} < I_{n_0''}$ and $X_{n_0''} = H$.

Suppose for instance that $X_n = E$ and $D_n \geq I_n$. Observe that $\beta' < \alpha$ implies $m(n) \leq 1 - l(n) + 2I_n^{-\nu}$. Therefore, for sufficiently large I_{i_0} (which can be assumed without loss of generality),

$$G_n \leq \mu a' I_n^{m(n)} \leq \mu a' I_n^{1-l(n)+2I_n^{-\nu}} \leq \mu^2 a' I_n^{1-l(n)} \leq \mu^2 a' D_n^{1-l(n)} \leq \mu^2 a' b^{-1} E_n^{1-l(n)}.$$

Finally, using notations of lemma 8.4.1,

$$\begin{aligned} & \Omega_{A,2} \cap \{\beta' \leq \alpha\} \cap \{T_1 = \infty\} \\ & \subset \Pi^{F,G}_{0,i_0,a\mu,1-\mu l,l} \cap \Pi^{G,F}_{0,i_0,a'\mu,1-\mu m,m} \cap ((\cap_{m_0 \geq i_0} \Pi^{F,G}_{1,m_0,\mu^2 a' b^{-1},l}) \cup (\cap_{m_0 \geq i_0} \Pi^{G,F}_{1,m_0,\mu^2 a b^{-1},m})) \text{ a.s.} \end{aligned}$$

Therefore it follows from lemma 8.4.1 that

$$\Omega_{A,2} \cap \{\beta' \leq \alpha\} \cap \{T_1 = \infty\} \subset \Omega_{A,2} \cap [(\cup_{q_0 \in \mathbb{N}} \Pi^{F,G}_{2,q_0}) \cup (\cup_{q_0 \in \mathbb{N}} \Pi^{G,F}_{2,q_0})] \subset \emptyset \text{ a.s.}$$

□

Let $b \in]0, 1/2[$. Define the stopping time

$$U_1 = \inf\{n \geq i_0/\ l(n) + m(n) - 1 < b \text{ or } \tilde{l}(n) + \tilde{m}(n) - 1 < b\},$$

and $T_2 = T_1 \wedge U_1$. Then, for all $\mu > 1$,

$$\Omega_{A,2} \subset \cup_{i_0 \in \mathbb{N}, a, a', \nu > 0, b \in]0,1/2[, l, m \in]0,1[} \{T_2 = +\infty\}.$$

We assume, without loss of generality, $B_{i_0}, \ldots K_{i_0}$ as large as necessary. We suppose i_0 is large enough, and take b smaller if necessary, so that, for $n \in [i_0, T_2[$, $l(n) \leq 1-b$, $F_n \leq D_n^{1-b} \leq \epsilon b D_n$, and similarly $m(n) \leq 1-b$, $G_n \leq I_n^{1-b} \leq \epsilon b I_n$.

Let us define, for all $n \in \mathbb{N}$, $v(n) = \mu^{-1}(l(n) + m(n) - 1)$. Observe that, for all $n \in [i_0, T_2[$, $v(n) \geq \mu^{-1} b$.

Given $\lambda > \mu > 1$, let us define two random sequences $(\xi_n)_{n \in \mathbb{N}}$, $(\tilde{\xi_n})_{n \in \mathbb{N}}$, as follows : $\xi_0 = j_0$ and , for all $n \in \mathbb{N}$,

$$\begin{cases} \tilde{\xi_n} = \xi_n & \text{if } I_{\xi_n} \in [\lambda^{-1} D_{\xi_n}, \lambda D_{\xi_n}] \\ \tilde{\xi_n} = \inf\{k > \xi_n / \ I_k \in [\lambda^{-1} D_k, \lambda D_k]\} & \text{otherwise,} \end{cases}$$

and

$$\xi_{n+1} = \inf\{k > \tilde{\xi_n} / \ X_k = G \text{ and } X_{k-1} = F\}$$

in all cases.

For all $k \in \mathbb{N}$, let

$$T_{+,k} = \inf\{n \geq k \ X_n = F\} \ , \ T_{-,k} = \inf\{n \geq k / \ X_n = G\}.$$

Lemma 8.8.2 *If $I_k \leq \lambda D_k$, then*

$$P(H_{T_{+,k}} > H_k + H_k^{1-v(k)} \mid \mathcal{F}_k) \leq \exp[-\lambda^{-1} \mu^{-7} m a a' I_k^{(\mu-1)b}].$$

PROOF: Let us define the stopping time

$$S = \inf\{n \geq k / \ H_n > H_k + H_k^{1-v(k)}\}.$$

Assume $n \in [k, T_{+,k}]$, then $H_n \leq \mu H_k$, which implies $I_n \leq \mu m^{-1} H_n \leq \mu^2 m^{-1} H_k \leq \mu^3 I_k$. Hence $G_n/(G_n + I_n) \geq \mu^{-3} G_k / I_k$, $F_n/(F_n + H_n) \geq \mu^{-1} F_k / H_k$. Therefore

$$\begin{aligned} P(H_{T_{+,k}} > H_k + H_k^{1-v(k)} \mid \mathcal{F}_k) &\leq (1 - \mu^{-4} \frac{F_k G_k}{H_k I_k})^{H_k^{1-v(k)}} \\ &\leq \exp[\mu^{-7} \lambda^{-1} m^{-1} a a' I_k^{l(k)+m(k)-1}] \\ &\leq \exp[-\lambda^{-1} \mu^{-7} m a a' I_k^{(\mu-1)b}]. \end{aligned}$$

□

We obtain similar inequalities if we replace H by G, I or J for T_+, and by C, D, E or F for T_-.

Therefore, let us define, for all $n \in \mathbb{N}$, the stopping time

$$\begin{aligned} U_2 = \inf\{n \geq i_0 / \ &\exists V \in \{G, H, I, J\} / \ \exists k \geq i_0 / \ I_k \leq \lambda D_k \ , \ T_{+,k} \leq i_0 \text{ and } V_{T_{+,k}} - V_k \geq V_k^{1-v(k)} \\ &\exists V \in \{C, D, E, F\} / \ \exists k \geq i_0 / \ D_k \leq \lambda I_k \ , \ T_{-,k} \leq i_0 \text{ and } V_{T_{-,k}} - V_k \geq V_k^{1-v(k)}\}. \end{aligned}$$

Then, for all $\mu > 1$ and $\lambda > 1$,

$$\Omega_{A,2} \subset \cup_{i_0 \in \mathbb{N}, a, a', \nu > 0, b \in]0,1/2[, l, m \in]0,1[} \{T_3 = +\infty\},$$

where $T_3 = T_2 \wedge U_2$.

Remark that in particular, for all $V \in \{C, D, E, F, G, H, I, J\}$, for all $\xi_n < T_3$ such that $\xi_{n-1} \geq i_0$, then $V_{\xi_n} \leq V_{\tilde{\xi}_{n-1}} + V_{\tilde{\xi}_{n-1}}^{1-v(\tilde{\xi}_{n-1})}$. A consequence is that $I_{\xi_n} \in [\lambda^{-1}\mu^{-1}D_{\xi_n}, \lambda\mu D_{\xi_n}]$ as long as $\xi_n < T_3$.

Lemma 8.8.3 *There exists a constant $e > 0$, depending only on b, λ, μ, l and m such that*

$$\frac{F_{\xi_n} G_{\xi_n}}{D_{\xi_n}} \succeq en 1_{\xi_n < T_3}.$$

Moreover,

$$\sum_{\xi_n < T_3} F_{\xi_n}^{-1} < +\infty \ , \ \sum_{\xi_n < T_3} G_{\xi_n}^{-1} < +\infty \ , \ \sum_{\xi_n < T_3} \frac{K_{\xi_n}}{F_{\xi_n} G_{\xi_n}} < +\infty \ , \ \sum_{\xi_n < T_3} \frac{B_{\xi_n}}{F_{\xi_n} G_{\xi_n}} < +\infty.$$

PROOF: Let us first prove the first statement. Let us define the set $\mathcal{I}$ of integers $n \in \mathbb{N}$ such that $I_n \in [\lambda^{-1}\mu^{-1}D_n, \lambda\mu D_n]$. It suffices to compare

$$\sum_{n \leq p, n \in \mathcal{I}} 1_{X_n = F, X_{n+1} = G} \text{ with } \frac{F_p G_p}{D_p}.$$

Now, let us apply statement **(1)** of lemma 8.3.5 :

$$\sum_{n \leq p, n \in \mathcal{I}} 1_{X_n = F, X_{n+1} = G} \asymp \sum_{n \leq p, n \in \mathcal{I}} 1_{X_n = F} \frac{G_n}{E_n + G_n} \preceq \sum_{n \leq p, n \in \mathcal{I}} \frac{1_{X_n = F}}{F_n^{1-b/2}} \frac{F_n^{1-b/2} G_n}{E_n}.$$

The term $F_n^{1-b/2} G_n / E_n$ roughly can only increase when n increases. More precisely, for all $n \leq p$ such that $n \in \mathcal{I}$,

$$\begin{aligned} \frac{F_n^{1-b/2} G_n}{E_n} &\leq \lambda\mu^3 a^{1-b/2} a' m^{-1} D_n^{(1-b/2)l(n) + m(n) - 1} \leq \lambda\mu^4 a^{1-b/2} a' m^{-1} D_n^{(1-b/2)l(p) + m(p) - 1} \\ &\leq \lambda\mu^4 a^{1-b/2} a' m^{-1} D_p^{(1-b/2)l(p) + m(p) - 1} \leq \lambda^2 \mu^7 m^{-1} \frac{F_p^{1-b/2} G_p}{D_p}. \end{aligned}$$

Therefore

$$\sum_{n \leq p, n \in \mathcal{I}} 1_{X_n = F, X_{n+1} = G} \preceq \lambda^2 \mu^7 m^{-1} \frac{F_p^{1-b/2} G_p}{D_p} \sum_{n \leq p} \frac{1_{X_n = F}}{F_n^{1-b/2}} \preceq 2b^{-1} \lambda^2 \mu^7 m^{-1} \frac{F_p G_p}{D_p}.$$

It remains to prove the second statement. Remark that there exists a.s $\nu > 0$ such that, for all $n \in \mathbb{N}$, $\max(B_{\xi_n}, F_{\xi_n}, G_{\xi_n}, K_{\xi_n}) \leq D_{\xi_n}^{1-\nu}$ (recall that $I_{\xi_n} \in [\lambda^{-1}\mu^{-1}D_{\xi_n}, \lambda\mu D_{\xi_n}]$). Hence

$$\sum_{n\in\mathbb{N}/\ \xi_n<T_3} \frac{\max(B_{\xi_n}, F_{\xi_n}, G_{\xi_n}, K_{\xi_n})}{F_{\xi_n}G_{\xi_n}} \leq \sum_{n\in\mathbb{N}/\ \xi_n<T_3} \frac{D_{\xi_n}}{F_{\xi_n}G_{\xi_n}} \frac{1}{D_{\xi_n}^{\nu}} \preceq \sum_{n\in\mathbb{N}} \frac{1}{en^{1+\nu}} < \infty.$$

□

Let us define the stopping times

$$U_3 = \inf\{n \geq i_0/\ F_{\xi_n}G_{\xi_n}/D_{\xi_n} < en\}\ ,\ T_4 = U_3 \wedge T_3.$$

Let, for all $k \in \mathbb{N}$,

$$R_k = H_k + J_k - (C_k + E_k).$$

Lemma 8.8.4 *Let $i_0 \in \mathbb{N}$, a, a', $\nu > 0$, $b \in]0, 1/2[$, l, $m \in]0, 1[$. There exist sequences of $\mathcal{F}_{\xi_n}$-mesurable random variables $(\tilde{D}_n)_{n\geq k_0}$, $(\tilde{I}_n)_{n\geq k_0}$, $(\epsilon_n)_{n\geq m}$, $(r_n)_{n\geq k_0}$ such that, if $\xi_n \leq T_4$, then*

1) $\tilde{D}_n = D_{\xi_n}, \tilde{I}_n = I_{\xi_n}$,

2) $\ln \frac{\tilde{I}_{n+1}}{\tilde{D}_{n+1}} - \ln \frac{\tilde{I}_n}{\tilde{D}_n} = \frac{R_{\xi_n}}{F_{\xi_n}G_{\xi_n}} + \epsilon_{n+1} + r_n$ *if* $I_{\xi_n} \in [\lambda^{-1}D_{\xi_n}, \lambda D_{\xi_n}]$,

3) $E(\epsilon_{n+1} \mid \mathcal{F}_{\kappa_n}) = 0\ ,\ E(\epsilon_{n+1}^2 \mid \mathcal{F}_{\kappa_n}) \leq Cst(\lambda).\left(\frac{D_{\xi_n}}{F_{\xi_n}G_{\xi_n}}\right)^2\ ,\ |r_n| \leq Cst(\lambda).\left(\frac{D_{\xi_n}}{F_{\xi_n}G_{\xi_n}}\right)^2.$

PROOF: The proof is similar to the proofs of lemmas 8.6.3 and 8.7.2.

□

Lemma 8.8.5

$$\Omega_{A,2} \subset \{I_{\xi_n}/D_{\xi_n} \to 1\} \cap \{\sum_{n\in\mathbb{N}} \frac{|R_{\xi_n}|}{F_{\xi_n}G_{\xi_n}} < +\infty\}.$$

PROOF: The proof is divided in two steps. In the first step we prove, using lemma 8.4.2, that

$$\Omega_{A,2} \cap \{I_{\xi_n}/D_{\xi_n} \not\to 1\} \subset \Omega_{A,2} \cap (\{\liminf I_n/D_n \geq \mu^{-3}\lambda\} \cup \{\limsup I_n/D_n \leq \mu^3\lambda^{-1}\}),$$

and

$$\Omega_{A,2} \cap \{T_4 = \infty\} \cap \{\lim I_{\xi_n}/D_{\xi_n} = 1\} \subset \{\sum_{n\in\mathbb{N}} \frac{|R_{\xi_n}|}{F_{\xi_n}G_{\xi_n}} < \infty\}.$$

We prove in the second step that, if λ was chosen large enough, then

$$\Omega_{A,2} \cap (\{\liminf I_n/D_n \geq \mu^{-3}\lambda\} \cup \{\limsup I_n/D_n \leq \mu^3\lambda^{-1}\}) = \emptyset \text{ a.s.}$$

First step Observe that, for all $k \in \mathbb{N}$, $I_k \leq H_k + J_k \leq I_k + G_k + K_k$ and $D_k \leq C_k + E_k \leq D_k + B_k + F_k$.

Now lemma 8.8.4 implies that, when $\tilde{I}_n \in [\lambda^{-1}\tilde{D}_n, \lambda\tilde{D}_n]$, then

$$\ln\frac{\tilde{I}_{n+1}}{\tilde{D}_{n+1}} - \ln\frac{\tilde{I}_n}{\tilde{D}_n} = \frac{D_{\xi_n}}{F_{\xi_n}G_{\xi_n}}(\frac{\tilde{I}_n}{\tilde{D}_n} - 1) + \epsilon_{n+1} + r_n + \frac{U_{\xi_n}}{F_{\xi_n}G_{\xi_n}},$$

where $U_{\xi_n} \leq B_{\xi_n} + F_{\xi_n} + G_{\xi_n} + K_{\xi_n}$. Therefore it follows from the second statement of lemma 8.8.3 that

$$\sum_{n\in\mathbb{N}} \frac{|U_{\xi_n}|}{F_{\xi_n}G_{\xi_n}} < \infty.$$

Therefore conditions of lemma 8.4.2 are fulfilled in the case $y_n = \tilde{I}_n/\tilde{D}_n$, $\gamma_n = D_{\xi_n}/F_{\xi_n}G_{\xi_n}$, $s_n = r_n + U_{\xi_n}/F_{\xi_n}G_{\xi_n}$, $\lambda = \lambda$, $\mu = \mu$ (if we assume I_{i_0} large enough). Let us prove that the assumptions of statement **2)** of this lemma are satisfied. Indeed, an easy adaptation of lemma 8.8.3 implies there exists $f \in \mathbb{R}_*^+$ such that $F_{\xi_n}G_{\xi_n}/D_{\xi_n} \preceq fn$ on $\cup_{p\in\mathbb{N}}\{\forall p \geq p_0, y_k \in [\lambda^{-1}, \lambda]\}$ (since in this case G_n/E_n roughly can only decrease when n increases), which implies $\sum_{k\in\mathbb{N}} \gamma_k = \infty$. This completes the proof of this first step.

Second step Let us prove for instance that

$$\Omega_{A,2} \cap \{\liminf I_n/D_n \geq \mu^{-3}\lambda\}$$

is of probability 0. We now suppose we belong to this event.

We know that $\limsup \ln G_n/\ln I_n \leq 1 - b$ on $\{T_4 = \infty\}$. Let us compare $\ln F_n$ and $\ln H_n$.

Claim $\limsup \ln F_n/\ln G_n \leq \mu^5 l m^{-1}\lambda^{-1}$.

This claim will imply that $\limsup \ln F_n/\ln H_n \leq \mu^6 l\lambda^{-1}$ and therefore, applying corollary 8.4.1 with an appropriate choice of λ, that F is visited only finitely often, which leads to a contradiction.

Proof of the claim

$$M_n = \sum_{k=1}^{n} \frac{1_{X_{k-1}=F}}{F_{k-1}} - \sum_{k=1}^{n} \frac{1_{X_{k-1}=F, X_k=G}}{F_{k-1}} \frac{E_{k-1} + G_{k-1}}{G_{k-1}}$$

is a martingale and, endowing notations of Chow theorem,

$$< M >_n = \sum_{k=1}^{n} \frac{1_{X_{k-1}=F}}{F_k^2} \frac{E_{k-1}}{G_{k-1}} \preceq \sum_{k=1}^{n} \frac{1_{X_{k-1}=F}}{F_k}$$

since $E_n/F_nG_n \leq 1$ for sufficiently large E_{i_0} (recall that $l(n) + m(n) - 1 \geq b$ and $D_n/I_n \leq \mu^3\lambda^{-1}$).

Therefore

$$\sum_{k=1}^{n\wedge T_3} \frac{1_{X_{k-1}=F}}{F_{k-1}} \asymp \sum_{k=1}^{n\wedge T_3} \frac{1_{X_{k-1}=F,X_k=G}}{F_{k-1}} \frac{E_{k-1} + G_{k-1}}{G_{k-1}} \asymp \sum_{k=1}^{n\wedge T_3} \frac{1_{X_{k-1}=G,X_k=F}}{F_{k-1}} \frac{E_{k-1} + G_{k-1}}{G_{k-1}}.$$

Now

$$N_n = \sum_{k=1}^{n} \frac{1_{X_{k-1}=G,X_k=F}}{F_{k-1}} \frac{E_{k-1}+G_{k-1}}{G_{k-1}} - \sum_{k=1}^{n} \frac{1_{X_{k-1}=G}}{G_{k-1}} \frac{E_{k-1}+G_{k-1}}{F_{k-1}+H_{k-1}}$$

is a martingale, and

$$\begin{aligned} < N >_n &= \sum_{k=1}^{n} \frac{1_{X_{k-1}=G}}{G_{k-1}} \frac{E_{k-1}+G_{k-1}}{F_{k-1}+H_{k-1}} \frac{1}{F_{k-1}} \frac{E_{k-1}+G_{k-1}}{G_{k-1}} \frac{H_{k-1}}{F_{k-1}+H_{k-1}} \\ &\leq \sum_{k=1}^{n} \frac{1_{X_{k-1}=G}}{G_{k-1}} \frac{E_{k-1}+G_{k-1}}{F_{k-1}+H_{k-1}} \left(\frac{E_{k-1}}{F_{k-1}G_{k-1}} + \frac{1}{F_{k-1}}\right) \\ &\preceq \sum_{k=1}^{n} \frac{1_{X_{k-1}=G}}{G_{k-1}} \frac{E_{k-1}+G_{k-1}}{F_{k-1}+H_{k-1}} \end{aligned}$$

since $E_n/F_nG_n \leq 1$ for sufficiently large E_{i_0}.

Therefore

$$\sum_{k=1}^{n\wedge T_3} \frac{1_{X_{k-1}=F}}{F_{k-1}} \asymp \sum_{k=1}^{n\wedge T_3} \frac{1_{X_{k-1}=G}}{G_{k-1}} \frac{E_{k-1}+G_{k-1}}{F_{k-1}+H_{k-1}}.$$

It remains to observe that $\limsup(E_n + G_n)/(F_n + H_n) \leq \limsup(E_n/H_n + G_n/H_n) \leq \mu^5 lm^{-1}\lambda^{-1}$ on $\liminf I_n/D_n \geq \mu^{-3}\lambda$.

□

8.8.2 End of the proof of $P(\Omega_{A,2}) = 0$

Using the fact that $I_n/D_n \to 1$, we can adapt the proof of lemma 8.8.3 to estimate more precisely the behaviour of $F_{\xi_n}G_{\xi_n}/D_{\xi_n}$ and prove that, for all $\mu > 1$, there exists a.s $e > 0$ such that $\mu^{-1}en \leq F_{\xi_n}G_{\xi_n}/D_{\xi_n} \leq \mu en$ for enough large n.

Let us define the stopping times

$$U_4 = \inf\{k \geq \xi_{i_0}/\ \exists n \leq k/F_{\xi_n}G_{\xi_n}/D_{\xi_n} \notin [\mu^{-1}en, \mu en] \text{ or } I_{\xi_n} \notin [\mu^{-1}D_{\xi_n}, \mu D_{\xi_n}]\},$$

and

$$T_5 = T_3 \wedge U_4.$$

Then, for all $\mu > 1$,

$$\Omega_{A,2} \subset \cup_{i_0\in\mathbb{N},a,a',\nu,e>0,b\in]0,1/2[,l,m\in]0,1[}\{T_5 = \infty\} \text{ a.s.}$$

We now need to estimate the small variations of G with respect to H. We endow notation $(G/I)(n)$.

Lemma 8.8.6 *For all $j,\ k \in \mathbb{N}^*$ such that $k \geq j$ and $\xi_k < T_5$,*

$$P[(G/I)(I_{\xi_k}(1 + 1/\sqrt{j})) \leq G_{\xi_k}(1 + 2/\sqrt{j}) \mid \mathcal{F}_{\xi_k}) \leq \exp(-Cst(b)k/\sqrt{j}),$$

if I_{i_0} is sufficiently large in the stopping time T_5.

PROOF:

Let $\eta > 1$. Assumption $k < T_5$ implies $I_{\xi_k} \geq H_{\xi_k} \geq G_{\xi_k} \geq k$. Let us define the stopping times $R = \inf\{n \geq \xi_k / \ I_n \geq I_{\xi_k}(1 + 1/\sqrt{j}) + 1\}$, $S = \inf\{n \geq \xi_k / \ G_n \geq G_{\xi_k}(1 + 1/\sqrt{j})\}$, and $T = R \wedge S \wedge T_5$.

We want to prove that $S > T$ with large probability.

Firstly, for all $n < T$, $G_n \leq \eta G_{\xi_k}$ (for $j \geq Cst(\eta)$), which implies that the probability to go to G starting from I is less than $\eta G_{\xi_k}/(G_{\xi_k} + I_{\xi_k}) \leq \eta G_{\xi_k}/I_{\xi_k}$.

Therefore, applying lemma 8.3.1 with $N := I_{\xi_k}/\sqrt{j} + 1$, $p := \eta G_{\xi_k}/I_{\xi_k}$, $a := \eta(\eta - 1)p$ implies

$$P((IHG)_{\xi_k,T} \geq \eta^2 G_{\xi_k}/\sqrt{j} \mid \mathcal{F}_{\xi_k}) \leq \exp(-Cst(\eta)G_{\xi_k}/\sqrt{j}) \leq \exp(-Cst(\eta)k/\sqrt{j}).$$

On the other hand, at each visit to G at time n, the probability to go to F is $F_n/(F_n + H_n) \leq (\eta - 1)/2$ (for I_{i_0} sufficiently large), and the probability to go to H and return to G is also less than $(\eta - 1)/2$ for I_{i_0} large. Applying lemma 8.3.1 with $N := 2G_{\xi_k}/\sqrt{j}$, $p := \eta - 1$, $a := p/2$:

$$P((FG)_{\xi_k,T} + (GHG)_{\xi_k,T} \geq 3(\eta - 1)G_{\xi_k}/\sqrt{j} \mid \mathcal{F}_{\xi_k}) \leq \exp(-Cst(\eta)G_{\xi_k}/\sqrt{j}) \leq \exp(-Cstk/\sqrt{j}).$$

This completes the proof, since

$$G_{\xi_k,T} \leq (IHG)_{\xi_k,T} + (FG)_{\xi_k,T} + (GHG)_{\xi_k,T} + 1 \leq \frac{G_{\xi_k}}{\sqrt{j}}\eta[3(\eta - 1) + \eta^2] \leq 2\frac{G_{\xi_k}}{\sqrt{j}},$$

for enough small η.

□

Let us define the stopping times

$$U_5 = \inf\{n \geq \xi_{i_0}/\exists j, k \in \mathbb{N}/\xi_k \leq n, j \leq k[I_n \leq I_{\xi_k}(1 + 1/\sqrt{j}) \text{ and } G_n > G_{\xi_k}(1 + 2/\sqrt{j})] \\ \text{or } [D_n \leq D_{\xi_k}(1 + 1/\sqrt{j}) \text{ and } F_n > F_{\xi_k}(1 + 2/\sqrt{j})\},$$

and $T_6 = T_5 \wedge U_5$.

Then

$$P(\{U_5 < \infty\} \cap \{T_5 = \infty\}) \leq \sum_{k \geq j \geq i_0} \exp(-Cstk/\sqrt{j}) \leq \sum_{j \geq i_0} Cst.\sqrt{j}\exp(-Cst\sqrt{j}) < 1/2$$

for enough large i_0. Therefore, for all $\mu > 1$,

$$\Omega_{A,2} \subset \cup_{i_0 \in \mathbb{N}, a, a', \nu, e > 0, b \in]0,1/2[, l, m \in]0,1[} \{T_6 = \infty\} \text{ a.s.}$$

Let us now apply lemma 8.5.3 with $T := T_6$, $\delta_n := \xi_n$ (such that $\xi_0 = i_0$), $U_n := I_n/D_n$, $R_n := R_n = H_n + J_n - (C_n + E_n)$, $S_n = F_n G_n$, $v(\xi_m) = m$ for all $m \geq i_0$, $L := F$, $M := G$, $N := H$, $g_1 = 1$, $g_3 \in]0, 1/2[$, and $i := 2$.

First observe that $U = T_6 \wedge \xi_{\lfloor(1+g_3)m\rfloor} \leq S_2 \wedge V$ for a good choice of g_3. Indeed, $U \leq S_2$ for $g_3 < \mu^{-1}g_1 e = \mu^{-1}e$, since $p_{\xi_m}G_{\xi_m} \geq \mu^{-1}F_{\xi_m}G_{\xi_m}/H_{\xi_m} \geq \mu^{-1}em$.

And $U \leq V$ since, for all j, $k \in \mathbb{N}$ such that $j \geq k$ and $\xi_j < T_6$, $k = F_{\xi_k} G_{\xi_k} / H_{\xi_k} \in [\mu^{-5} a m^{-1} D_{\xi_k}^{l(\xi_k)+m(\xi_k)-1}, \mu^5 a m^{-1} D_{\xi_k}^{l(\xi_k)+m(\xi_k)-1}]$, $p_{\xi_k} \in [\mu^{-4} a m^{-1} D_{\xi_k}^{l(\xi_k)-1}, \mu^4 a m^{-1} D_{\xi_k}^{l(\xi_k)-1}]$, and $|l(\xi_k) - l(\xi_j)| \leq D_{\xi_k}^{-\nu}$, $|m(\xi_k) - m(\xi_j)| \leq I_{\xi_k}^{-\nu}$ imply $p_{\xi_{\lfloor (1+g_3)m \rfloor}} \in [p_{\xi_m}/2, 2p_{\xi_m}]$ for $\mu < Cst(b)$ and $g_3 < Cst(b)$, which can be assumed without loss of generality.

Applying lemma 8.8.4 for $\mathcal{M}$, and for $\mathcal{M}'$ in the case $n \geq \lfloor (1+g_3)m \rfloor$, we obtain that assumption **a)** of lemma 8.5.3 is satisfied in these cases, with $m(\xi_m, n) = Cst(e)/n^2$, $n(\xi_m, n) = Cst(e)/n^2$.

On the other hand, $U \leq S_2 \wedge V$ implies that **a)** is satisfied for $n \in [m, \lfloor (1+g_3)m \rfloor]$ such that $\xi'_n < T'_6$, with $u(\xi_m) = Cst(e)/m^{3/2}$. Indeed,

$$\begin{aligned}
\ln \frac{\tilde{I}'_{n+1}}{\tilde{D}'_{n+1}} - \ln \frac{\tilde{I}'_n}{\tilde{D}'_n} &\geq \frac{H'_{\xi'_n} + J'_{\xi'_n}}{F'_{\xi'_n} G'_{\xi'_n}} \frac{1}{1 - \frac{g_2}{\sqrt{2 p_{\xi_m} G_{\xi_m}}}} - \frac{C'_{\xi'_n} + E'_{\xi'_n}}{F'_{\xi'_n} G'_{\xi'_n}} + \epsilon'_{n+1} + r'_n \\
&\geq \frac{R'_{\xi'_n}}{F'_{\xi'_n} G'_{\xi'_n}} + \frac{g_2}{2 e^{3/2} m^{3/2}} + \epsilon'_{n+1} + r'_n
\end{aligned}$$

Easy calculation shows that assumption **b)** holds with $C_1 = Cst(e)$. Assumption **c)** follows directly from the definition of the coupling, and assumption **d)** follows from lemma 8.8.5.

Let us prove that **e)** (ii) holds. Let $n_k = \ln \frac{G'_{\xi'_k} / F'_{\xi'_k}}{G_{\xi_k} / F_{\xi_k}}$. Observe that

$$e^{n_k} F_{\xi_k} G_{\xi_k} \geq F'^2_{\xi'_k} G'_{\xi'_k} / F'_{\xi'_k} = F'_{\xi'_k} G'_{\xi'_k} = G'^2_{\xi'_k} F'_{\xi'_k} / G'_{\xi'_k} \geq e^{-n_k} F_{\xi_k} G_{\xi_k},$$

which implies $|\ln h_k| \leq \ln n_k$.

Now if $\xi_k < T_6$, $\ln g_k \leq C_2 u(\xi_m) v(\xi_m) = C_2 Cst(e)/\sqrt{m}$ implies $\ln I'_{\xi'_k} / I_{\xi_k} \leq C_3/\sqrt{m}$ and $\ln D_{\xi_k} / D'_{\xi'_k} \leq C'_2/\sqrt{m}$, with $C'_2 = C_2 Cst(e)$. We can assume C'_2 large enough in **e)**(ii).

Let us use for instance the first inequality. Let $\tilde{\xi}_k$ be the first $t \in \mathbb{N}$ such that $G'_{\xi'_k} = G_t$. Then $\tilde{\xi}_k \geq \xi_k$ and $I'_{\xi'_k} \geq I_{\tilde{\xi}_k}$. Therefore $I_{\tilde{\xi}_k} \leq I_{\xi_k}(1 + 2C_3/\sqrt{m})$ which implies, since $\xi_k < T_6$, that $G_{\tilde{\xi}_k} = G'_{\xi'_k} \leq G_{\xi_k}(1 + 4C'_2/\sqrt{m})$. We can apply the same argument for estimating $\ln F'_{\xi'_k} / F_{\xi_k}$. In summary, we obtain that, for sufficiently large m,

$$|\ln h_k| \leq \ln n_k \leq 10 C'_2/\sqrt{p}.$$

Now, $\xi_n < T_6$ and $\xi'_n < T_6$ imply, assuming C'_2 large enough,

$$g_{n+1} \geq g_n - C_2 / D_{\xi_m}^{v(\xi_m)} \geq g_n - C_2 u(\delta_m) v(\delta_m)/2$$

for all $\mu < 1/2$, since $v(\xi_m) > (l(\xi_m) + m(\xi_m) - 1)/2$ implies $D_{\xi_m}^{v(\xi_m)} \geq Cst(\mu, a, a', e)\sqrt{m}$.

Conclusion of lemma 8.5.3 completes the proof of $P(\Omega_{A,2}) = 0$.

Acknowledgments. I am very grateful to my thesis director Michel Benaïm for helpful discussions and for inspiring the choice of this subject. I would also like to thank Marie Duflo for her remarks on estimates, which have been very useful for improving some parts of this article.

Bibliographie

[1] M. Benaïm. Vertex-reinforced random walks and a conjecture of Pemantle. *Annals of probability*, 25 :No.1,361–392, 1997.

[2] M. Benaïm. Dynamics of stochastic approximation algorithms. *Séminaire de probabilités XXXIII, Lect. Notes in Math.*, 1709 :1–68, 1999.

[3] A. Bienvenüe. Contribution à l'étude des marches aléatoires avec mémoire. *Doctoral dissertation, Université Claude Bernard-Lyon 1*, 1999.

[4] D. Coppersmith and P. Diaconis. Random walks with reinforcement. *Unpublished manuscript*, 1986.

[5] B. Davis. Reinforced random walk. *Probab. Th. Rel. Fields*, 84 :203–229, 1990.

[6] R. Pemantle. Phase transition in reinforced random walk and RWRE on trees. *Annals of probability*, 16 :1229–1241, 1988.

[7] R. Pemantle. Random processes with reinforcement. *Massachussets Institute of Technology doctoral dissertation*, 1988.

[8] R. Pemantle. Vertex-reinforced random walk. *Probability Theory and Related Fields*, 92 :117–136, 1992.

[9] R. Pemantle and S. Volkov. Vertex-reinforced random walk on $\mathbb{Z}$ has finite range. *Annals of probability*, 27 :No.3,1368–1388, 1999.

[10] T. Sellke. Reinforced random walk on the d-dimensionnal integer lattice. *Technical report, no.94-26, Department of statistics, Purdue university*, 1994.

[11] S. Volkov. Vertex-reinforced random walk on arbitrary graphs. *To appear in Annals of probability*, 2000.

Résumé

La première partie de cette thèse est consacrée aux algorithmes stochastiques. Nous présentons de nouveaux résultats de convergence et de non-convergence vers certains ensembles "instables" du système dynamique associé. Nous étudions, en collaboration avec M. Benaïm et S. Schreiber, la modélisation de dynamiques génétiques par des processus d'urnes de Pólya généralisées.

La deuxième partie est consacrée à une conjecture de Pemantle et Volkov énonçant que les marches aléatoires renforcées par sommets sur les entiers restent p.s. bloquées dans cinq points au bout d'un certain temps. Nous prouvons cette conjecture, et donnons une nouvelle preuve des résultats obtenus précédemment sur le sujet par Pemantle et Volkov, en utilisant en partie certaines des techniques mises en oeuvre pour la conjecture.

Mots clés : Algorithmes stochastiques. Pièges répulsifs. Modèles d'urnes. Dynamiques génétiques. Marches aléatoires renforcées.

Traps of stochastic algorithms and vertex-reinforced random walks

The first part of the thesis focuses on stochastic algorithms. We present new results concerning convergence and non-convergence towards "unstable" sets of the associated dynamical system. In a joint work with M. Benaïm and S. Schreiber we study certain evolutionary processes described by generalized Pólya urn models.

The second part of thesis is devoted to a conjecture of Pemantle and Volkov, stating that vertex-reinforced random walks on integers eventually get stuck at a set of five points a.s. We prove this conjecture, and give a new proof of the results previously obtained by Pemantle and Volkov on this subject, by using in part some tools constructed for the conjecture.

Keywords : Stochastic algorithms. Repulsive traps. Urn models. Evolutionary processes. Reinforced random walks.