



Convergence abrupte et métastabilité

Olivier Bertoncini

► To cite this version:

Olivier Bertoncini. Convergence abrupte et métastabilité. Mathématiques [math]. Université de Rouen, 2007. Français. tel-00218132

HAL Id: tel-00218132

<https://tel.archives-ouvertes.fr/tel-00218132>

Submitted on 25 Jan 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

en vue de l'obtention du titre de

Docteur de l'Université de Rouen

présentée par

Olivier Bertoncini

Discipline : Mathématiques Appliquées
Spécialité : Probabilités

Convergence abrupte et Métastabilité

Date de soutenance : 29 novembre 2007

Composition du Jury

Président	:	P. Collet,	Professeur, École Polytechnique
Rapporteurs	:	A. Galves,	Professeur, Universidade de São Paulo
	:	B. Ycart,	Professeur, Université de Grenoble
Examineurs	:	E. Janvresse,	Chargée de recherche CNRS, Université de Rouen
	:	T. de la Rue,	Chargé de recherche CNRS, Université de Rouen
Directeur de Thèse	:	R. Fernández,	Professeur, Université de Rouen

Thèse préparée à l'Université de Rouen
Laboratoire de Mathématiques Raphaël Salem, UMR-CNRS 6085

Remerciements

Je veux tout d'abord remercier mon directeur de thèse Roberto Fernández pour sa disponibilité et pour l'écoute et la confiance qu'il m'a accordées. Cette thèse est le fruit d'un long travail en collaboration, j'y ai découvert le métier de chercheur, et il m'a transmis son approche et sa vision des mathématiques.

Merci à Antonio Galves et Bernard Ycart d'avoir accepté de relire mon travail, et de l'intérêt qu'ils y ont porté. Merci à Pierre Collet d'avoir accepté d'être le président de mon jury.

Je veux aussi remercier les membres du laboratoire Raphael Salem qui m'ont accompagné durant ces (quelques) années de thèse. Marguerite, Edwidge et Catherine, toujours disponibles et qui m'ont orienté dans les dédales de l'administration universitaire. Isabelle pour son aide précieuse dans la recherche documentaire (oups ! je crois que j'ai des livres en retard !). Marc qui a imprimé et relié cette thèse. Paul qui m'a fait passer du deuxième au troisième cycle en étant successivement tuteur de mon mémoire de maîtrise puis de mon monitorat. Claude dont j'ai bu et je bois encore les paroles, depuis les cours de DEA jusqu'à ses (nombreux) exposés à l'atelier. Et enfin Thierry et Elise d'avoir accepté de faire partie de mon jury, et pour leur précieux conseils lors de la préparation de mes exposés et posters. Merci aussi à François, Mustapha, Margherita, Vlad, Jean-Baptiste et Olivier d'avoir répondu présents lorsque j'avais une question mathématique ou informatique.

Je remercie encore une fois Claude, Thierry et Elise pour l'organisation de l'atelier des doctorants. Cela m'a permis de porter un regard critique sur la façon de concevoir et de présenter un exposé.

Je remercie également tous mes collègues et ex-collègues doctorants du laboratoire pour les bons moments, pour les nombreuses discussions (mathématiques et autres !), pour les repas et les pauses cafés, ou pour avoir partagé mon bureau. Dans le désordre : Sara, Greg, Mohammed, Etienne, Patrice, Lahcen, Islam, Nico, Lamia, Nico, Vincent, Olivier, Nico...

Merci à mes parents de m'avoir toujours soutenu et suivi dans mes choix.

Enfin, une pensée spéciale à mes deux amours Sylvie et Titouan.

Résumé

Le but de cette thèse est de relier deux phénomènes relatifs au comportement asymptotique des processus stochastiques, qui jusqu'à présent étaient restés dissociés. La convergence abrupte ou phénomène de cutoff d'une part, et la métastabilité d'autre part. Dans le cas du cutoff, une convergence abrupte vers la mesure d'équilibre du processus a lieu à un instant que l'on peut déterminer, alors que la métastabilité est liée à une grande incertitude sur l'instant où l'on va sortir d'un certain équilibre. On propose un cadre commun pour étudier et comparer les deux phénomènes : celui des chaînes de naissance et de mort à temps discret sur \mathbb{N} , qui possèdent une dérive vers zéro.

On va montrer que sous l'hypothèse de dérive, il y a convergence abrupte vers zéro, métastabilité dans l'autre sens, et que de plus, la dernière excursion dans la métastabilité est la renversée temporelle d'une trajectoire typique de cutoff.

On étend ensuite notre approche au modèle d'Ehrenfest, ce qui nous permet de montrer la convergence abrupte et la métastabilité sous une hypothèse de dérive plus faible.

Mots clefs

Convergence abrupte, cutoff, métastabilité, chaînes de Markov, marches aléatoires, chaînes de naissance et de mort, temps d'atteinte, couplages.

Abstract

The aim of this thesis is to link two phenomena concerning the asymptotical behavior of stochastic processes, which were disjoined up to now. The abrupt convergence or cutoff phenomenon on one hand, and metastability on the other hand. In the cutoff case an abrupt convergence towards the equilibrium measure occurs, at a time which can be determined, whereas metastability is linked to a great uncertainty of the time at which we leave some equilibrium. We propose a common framework to compare and study both phenomena : that of discrete time birth and death chains on \mathbb{N} , with drift towards zero.

Under the drift hypothesis, we prove that there is an abrupt convergence towards zero, metastability in the other direction, and that the last exit in the metastability is the time reverse of a typical cutoff path.

Then we extend our approach to the Ehrenfest model, which allows us to prove abrupt convergence and metastability with a weaker drift hypothesis.

Keywords

Abrupt convergence, cutoff, metastability, Markov chains, random walks, birth and death chains, hitting times, coupling

Table des matières

1	Introduction	3
2	Cadre	7
2.1	Chaînes de Markov	7
2.2	Chaînes de Naissance et de Mort	10
2.3	Distance à l'équilibre pour les modèles d'absorption	12
2.4	Couplages	14
3	Convergence abrupte et phénomène de Cutoff	15
3.1	Historique et définitions	16
3.2	Convergence abrupte pour un temps d'atteinte	21
3.3	Exemples extrêmes de cutoff	24
4	Métastabilité	31
4.1	Un peu de physique	31
4.2	Théorie de Curie-Weiss	34
4.3	Modèle de Curie-Weiss	35
4.4	Approche "pathwise"	36
4.5	Métastabilité pour les chaînes de Markov	37
4.6	Autres approches de la métastabilité	44
5	Convergence abrupte et métastabilité pour les Chaînes de Naissance et de Mort	47
5.1	Modèle	47
5.2	Dérive, énergie et décroissance exponentielle	48
5.3	Temps d'atteinte et mesure invariante	50
5.4	Convergence abrupte	56
5.5	Métastabilité de la trajectoire de sortie	57
5.6	Dernière sortie	64
6	Modèle d'Ehrenfest	71
6.1	Introduction	71
6.2	Modèle	73

6.3	Quantités remarquables	74
6.4	Résultats	79
7	Cadre général	83
7.1	L'hypothèse (\mathcal{H})	83
7.2	Modèles à transitions “indépendantes” de a	84
7.3	Cas Ehrenfest	86
7.4	Modèle hybride	88
7.5	Généralisation	90
7.6	Temps d'atteinte et de sortie d'une région	91
8	Conclusion et perspectives	95
9	Appendice	97
9.1	Temps d'atteinte pour les marches aléatoires	97
9.2	Marche aléatoire sur l'hypercube	102
9.2.1	Description réduite	102
9.2.2	Métastabilité de la trajectoire de sortie	104

Chapitre 1

Introduction

Le but de cette thèse est de relier deux phénomènes relatifs au comportement asymptotique des processus stochastiques, qui jusqu'à présent étaient restés dissociés : la convergence abrupte ou phénomène de cutoff d'une part, et la métastabilité d'autre part.

Le phénomène de cutoff est un phénomène asymptotique de convergence abrupte à l'équilibre découvert au début des années 80 par Aldous et Diaconis en étudiant des modèles de battage de cartes [2]. Cela a permis à Diaconis de répondre à la question "Combien de fois doit-on battre un jeu de 52 cartes afin qu'il soit bien mélangé ?" : sept fois suffisent ("seven shuffles are enough").

En effet, il existe un temps autour duquel, asymptotiquement en un certain paramètre (ici le nombre de cartes dans le paquet), la loi du processus converge de façon abrupte vers la mesure d'équilibre. Si on mesure l'écart entre la loi du processus au temps t et sa mesure d'équilibre à l'aide d'une distance (séparation, variation totale), il est proche de 1 avant cet instant, et proche de 0 juste après, la transition se faisant asymptotiquement de plus en plus abrupte.

Ce temps, qu'ils ont appelé temps de cutoff, est déterministe, et tend vers l'infini avec le paramètre asymptotique. Dans les modèles qu'ils ont étudiés, Aldous et Diaconis le déterminent à l'aide des propriétés spectrales de la matrice de transition du modèle.

Cela n'est pas satisfaisant si l'on veut généraliser l'approche, car cela nécessite la connaissance des valeurs propres de la matrice de transition. Notre travail a pour but de déterminer l'existence du phénomène à partir de données plus générales sur le processus, et notamment la connaissance de sa mesure invariante et le profil énergétique qui en résulte.

On utilise l'approche de Martínez et Ycart [38], qui ont montré l'équivalence entre le cutoff pour la convergence vers la mesure d'équilibre avec une notion de cutoff pour le temps d'atteinte d'un état absorbant, et ont donné une condition suffisante pour l'existence d'un tel cutoff. Sous une condition de décroissance exponentielle de la mesure invariante d'une chaîne de Markov sur un arbre, ils montrent l'existence d'un cutoff pour le temps d'atteinte de la racine de l'arbre.

Cela nous amène à définir une notion de convergence abrupte pour les temps d'atteinte : un temps d'atteinte converge abruptement au temps t_a ($t_a \rightarrow \infty$), s'il est asymptotiquement équivalent à t_a en probabilité (définition 3.1).

On se place dans le cadre des marches aléatoires à temps discret sur l'axe des entiers qui ont une dérive vers 0. On considère des familles $X_a(t)$ de chaînes de naissance et de mort sur $\llbracket 0, a \rrbracket = \{0, \dots, a\}$, et le paramètre asymptotique a est donc ici la taille de l'espace d'états. La dérive vers 0 est caractérisée par le fait que la mesure invariante de la chaîne possède une queue à décroissance exponentielle (condition (\mathcal{H}) , (5.3)), et qu'elle est donc concentrée autour de 0. Zéro est “caractéristique” de la mesure d'équilibre, dans le sens où elle y atteint son maximum. Nous considérerons dans notre approche que 0 est “l'état d'équilibre” du système dans l'acceptation physicienne du mot “état”. En effet, en physique les états d'un système sont les configurations du système, alors que pour les probabilistes ce sont des mesures de probabilité.

Le profil énergétique associé à une telle mesure possède une pente vers zéro (section 5.2), et on comprend intuitivement qu'une chaîne initialement en a va vite atteindre 0 en suivant la pente. On montre en effet (section 5.4, théorème 5.3) que $T_{a \rightarrow 0}$, le temps d'atteinte de 0 pour la chaîne initialement en a , admet une convergence abrupte au temps $\mathbb{E}[T_{a \rightarrow 0}]$.

Cette approche n'est pas directement compatible avec celle de Diaconis, car dans les modèles de marches aléatoires sur les groupes qu'il étudie (dont les modèles de battage de cartes sont des cas particuliers), la mesure invariante est la mesure uniforme sur le groupe et son profil énergétique est plat. Mais on voit avec le modèle d'Ehrenfest (chapitre 6), qu'on peut étudier ces phénomènes en le ramenant dans le cadre des marches aléatoires unidimensionnelles.

Cette approche nous permet aussi de faire le lien avec la métastabilité. A la différence de la convergence abrupte qui est quasi déterministe, la métastabilité est un phénomène très aléatoire. Elle est caractérisée par le fait que le système va rester un temps exponentiellement long dans un état dit métastable, avant de basculer vers l'équilibre.

L'approche “pathwise” de la métastabilité consiste à regarder la loi des trajectoires typiques entre l'état métastable et l'état stable du système (section 4.4).

Elle a été utilisée pour la première fois par Cassandro, Galves, Olivieri et Vares [12], pour décrire le comportement métastable du modèle de Curie-Weiss (voir sections 4.2 et 4.3). C'est un modèle qui décrit l'évolution de la magnétisation d'un système ferromagnétique soumis à un faible champ externe positif. Son profil énergétique possède deux puits d'attraction, avec un minimum local m^- de magnétisation négative et un minimum global m^+ de magnétisation positive qui correspond à l'état d'équilibre du système. Le minimum local m^- correspond lui à un état de pseudo-équilibre : si le système est dans le puits d'attraction de m^- , il est attiré par ce minimum et devra en sortir pour atteindre le vrai état d'équilibre m^+ . On dit que m^- est un état métastable. Si T est le temps d'atteinte de m^+ partant de m^- , le comportement métastable de T est caractérisé par le fait que $T/\mathbb{E}[T]$ converge en loi vers une exponentielle de paramètre 1, asymptotiquement en un certain paramètre (voir (4.14) et la définition 4.1).

Et ce qui fait la métastabilité est en fait l'imprédictibilité du temps de sortie du puits d'attraction de m^- . En effet, il faudra au système un grand nombre d'essais pour sortir de ce puits, mais une fois qu'il en est sorti, il est dans le puits d'attraction de m^+ , et il va y tomber de façon quasi-déterministe. On retrouve le phénomène de convergence abrupte ! Que le puits d'attraction corresponde à un minimum local ou global ne change pas la problématique, qui est donc la suivante : tomber dans un puits d'attraction correspond à la convergence abrupte, alors qu'en sortir correspond à la métastabilité.

De plus, pour des raisons de symétrie, le côté par lequel on rentre ou sort du puits importe peu. On peut donc transposer cette problématique aux chaînes de naissance et de mort entre zéro et a avec dérive vers zéro.

C'est ce que l'on fait au chapitre 5. Avec notre hypothèse de dérive (condition (\mathcal{H}) , (5.3)), on est en fait en présence d'un demi-puits d'attraction, avec minimum en zéro qui est donc l'état d'équilibre du système. Il n'y a pas ici d'état métastable, la métastabilité vient du fait qu'il faut "remonter la pente" pour sortir du puits d'attraction de 0.

On montre la convergence abrupte dans le "sens de la pente", c'est-à-dire de a vers zéro (théorème 5.3), et la métastabilité de 0 vers a (théorème 5.6).

Pour montrer la convergence abrupte de $T_{a \rightarrow 0}$ au temps $\mathbb{E}[T_{a \rightarrow 0}]$ (au sens de la définition 3.1), on montre que sous la condition (\mathcal{H}) , la chaîne satisfait les hypothèses de la proposition 3.2, qui est une condition suffisante due à Martínez et Ycart.

L'existence de deux échelles de temps est une condition suffisante pour la métastabilité (théorème 4.2). En collaboration avec J. Barrera, nous avons utilisé un couplage entre une chaîne initialisée en 0 et une initialisée en a pour décomposer le temps de sortie $T_{0 \rightarrow a}$ de 0 à a en une somme géométrique de N_a "tombées à l'équilibre" $T_{a \rightarrow 0}$ (section 5.5). On montre ainsi que le temps moyen $\mathbb{E}[T_{0 \rightarrow a}]$ pour aller de 0 à a est de l'ordre de $\mathbb{E}[N_a] \mathbb{E}[T_{a \rightarrow 0}]$, qui est donc beaucoup plus grand (au sens de (4.19)) que le temps de cutoff $\mathbb{E}[T_{a \rightarrow 0}]$.

On voit donc que ces deux phénomènes sont complémentaires, et on explore plus encore cette complémentarité en montrant que la dernière sortie dans la trajectoire métastable est la renversée temporelle d'une trajectoire typique de cutoff. Pour atteindre a , le système initialement en zéro va réaliser de nombreuses tentatives infructueuses (c'est ce qui "fait" la métastabilité), mais la dernière sortie est "sans hésitation" : elle emmène le système en a de façon quasi-déterministe, de la même façon que pour "tomber" de a à 0 dans la convergence abrupte.

La thèse est organisée de la manière suivante :

Nous rappelons au chapitre 2 les notions de bases sur les chaînes de Markov qui vont nous être utiles par la suite.

Dans le chapitre 3, nous faisons un survol historique du phénomène de cutoff (section 3.1). Puis nous définissons la notion de convergence abrupte pour des temps d'atteinte (définition 3.1), et donnons une condition suffisante pour son existence (proposition 3.2), dans l'esprit des travaux de Martínez et Ycart (section 3.2). Enfin, dans la dernière section de ce chapitre (section 3.3), on illustre notre approche du phénomène en étudiant des exemples très simples.

Le chapitre 4 est consacré à la métastabilité. Après avoir présenté le phénomène physique dans la section 4.1, on introduit le modèle de Curie-Weiss (sections 4.2 et 4.3) et l’approche “pathwise” (section 4.4). Dans la section 4.5 on formalise la métastabilité dans l’esprit de l’approche “pathwise” (définition 4.1), et on montre que l’existence de deux échelles de temps est une condition suffisante pour son existence (théorème 4.2). Enfin, dans la dernière section de ce chapitre (section 4.6), on présente brièvement les autres approches utilisées pour formaliser la métastabilité.

On montre au chapitre 5 la convergence abrupte (section 5.4) et la métastabilité (section 5.5) pour les chaînes de naissance et de mort avec dérive vers zéro. On discute dans la section 5.2 les interprétations en termes de mesure invariante et de profil énergétique de la condition (\mathcal{H}) . Et dans la section 5.6, on montre que la dernière sortie dans la trajectoire métastable est la renversée temporelle d’une trajectoire typique de cutoff (propositions 5.7, 5.8 et 5.9).

Dans le chapitre 6, on montre comment notre approche s’étend à un modèle qui ne satisfait pas la condition (\mathcal{H}) : le modèle d’Ehrenfest. On rappelle le modèle original (section 6.1), puis on en présente notre version qui rentre dans le cadre des chaînes de naissance et de mort du chapitre précédent (section 6.2). Les résultats sont dans la section 6.4, et les calculs dans la section 6.3.

On discute au chapitre 7 les différences entre les chaînes de naissance et de mort du chapitre 5 et le modèle d’Ehrenfest (section 7.1 à 7.3), et on introduit un modèle hybride (section 7.4) qui conserve certaines des propriétés de chacun de ces modèles de telle façon qu’on puisse lui appliquer l’approche développée au chapitre 5. On présente ensuite le cadre le plus général possible dans lequel on peut utiliser cette approche (section 7.5), et dans la section 7.6 on discute des conditions suffisantes pour montrer la convergence abrupte et la métastabilité si on regarde les temps d’atteinte et de sortie d’une région “caractéristique” de l’équilibre (et non plus seulement d’un état), pour le modèle d’Ehrenfest et le modèle hybride.

Enfin, on conclut notre propos dans le chapitre 8, et on discute des orientations possibles à donner à notre travail.

On trouvera en appendice (chapitre 9), l’expression des temps moyens d’atteinte pour une marche aléatoire en fonction de ses transitions, et également le lien entre le modèle d’Ehrenfest et la marche aléatoire sur l’hypercube.

Chapitre 2

Cadre

L'objet de notre travail est de comparer le phénomène de convergence abrupte et la métastabilité pour des familles $X_a(t)$ de chaînes de naissance et de mort sur l'intervalle $\{0, \dots, a\}$, où le paramètre a va tendre vers l'infini.

Les chaînes de naissance et de mort sont des chaînes de Markov irréductibles récurrentes positives. Dans ce chapitre, nous allons donc rappeler ce qu'est une chaîne de Markov, une mesure invariante, une mesure réversible. Nous allons également définir les notions d'irréductibilité, de récurrence positive, et les principales notations utilisées dans cette thèse (section 2.1).

Dans la section 2.2, après avoir rappelé les conditions de récurrence positive pour une chaîne de naissance et de mort, on définit les modèles que l'on utilisera au chapitre 5.

Ensuite (section 2.3), on définit la distance en variation totale, et on montre que pour un modèle d'absorption on peut relier la distance à l'équilibre et le temps d'absorption (lemme 2.3).

Enfin, dans la section 2.4 on rappelle la notion de couplage pour des chaînes de Markov.

2.1 Chaînes de Markov

Définitions

Les chaînes de Markov sont des processus à temps discret qui satisfont la propriété de Markov, c'est-à-dire telles que l'évolution future de la chaîne ne dépend du passé qu'au travers de l'état présent du processus.

Généralement on va les noter $X(t)$, où le temps $t \in \mathbb{N}$. C'est donc une suite de variables aléatoires sur un certain espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans un espace d'états \mathcal{S} , telles que

$$\mathbb{P}(X(t+1) = x_{t+1} \mid X(t) = x_t, \dots, X(0) = x_0) = \mathbb{P}(X(t+1) = x_{t+1} \mid X(t) = x_t), \quad (2.1)$$

pour tout $t \in \mathbb{N}$ et tous $x_0, \dots, x_{t+1} \in \mathcal{S}$.

Cette propriété s'appelle la propriété de Markov, et les probabilités

$$p_t(x, y) := \mathbb{P}(X(t+1) = y \mid X(t) = x) \quad (2.2)$$

s'appellent les probabilités de transition de la chaîne.

Nous allons nous intéresser uniquement aux chaînes homogènes, c'est-à-dire lorsque les transitions sont indépendantes de t ($p_t(x, y) = p(x, y)$, $\forall t$).

Dans ce cas, la propriété de Markov (2.1) devient

$$\mathbb{P}(X(t+1) = x_{t+1} \mid X(t) = x_t, \dots, X(0) = x_0) = p(x_t, x_{t+1}) . \quad (2.3)$$

La probabilité d'une trajectoire est alors le produit des transitions entre les états successifs de la trajectoire

$$\mathbb{P}(X(0) = x_0, \dots, X(t) = x_t) = \mathbb{P}(X(0) = x_0) p(x_0, x_1) \dots p(x_{t-1}, x_t) . \quad (2.4)$$

On suppose de plus que la chaîne est irréductible, c'est-à-dire que tous les états communiquent :

$$\forall x, y \in \mathcal{S}, \quad \exists t > 0 \text{ tel que } p^t(x, y) > 0 , \quad (2.5)$$

où $p^t(x, y)$ est le coefficient (x, y) de la matrice $P^t = P \times \dots \times P$, où P est la matrice de transition de la chaîne (dont les coefficients sont les transitions de la chaîne $p(x, y)$).

Nous considérerons également que l'espace d'états \mathcal{S} est discret (on parle alors de marche aléatoire).

Une variable $T : \Omega \rightarrow \mathbb{N}$ est un temps d'arrêt pour la chaîne $X(t)$ s'il ne dépend que du passé de la chaîne, autrement dit si l'évènement $\{T \leq t\}$ est mesurable par rapport au passé (i.e. si $\{T \leq t\} \in \mathcal{F}_t$, où $\mathcal{F}_t = \sigma(X(0), \dots, X(t))$).

Remarque : la propriété de Markov se généralise dans le cas où l'instant t est déterminé par un temps d'arrêt T , c'est la propriété de Markov forte :

soit T un temps d'arrêt pour la chaîne $X(t)$, alors la loi de $X(T+1)$ sachant que T est fini dépend uniquement de la valeur de $X(T)$

$$\mathbb{P}(X(T+1) = y \mid X(T) = x, T < \infty) = p(x, y) . \quad (2.6)$$

Probabilité invariante

Une mesure μ sur \mathcal{S} est invariante pour la chaîne si

$$\mu(x) = \sum_{y \in \mathcal{S}} \mu(y) p(y, x) , \quad \forall x \in \mathcal{S} . \quad (2.7)$$

Lorsqu'une mesure de probabilité invariante μ existe pour la chaîne $X(t)$, alors la loi $\mathcal{L}(X(t))$ de la chaîne au temps t converge vers μ lorsque t tend vers l'infini.

Les cas qui nous intéressent sont ceux où la mesure invariante existe et est unique. C'est notamment le cas lorsque la chaîne est récurrente positive.

Soit $T_{x \rightarrow y}$ le temps d'atteinte de y partant de x

$$T_{x \rightarrow y} = \inf\{t > 0 : X^x(t) = y\} , \quad (2.8)$$

où l'on note par $X^x(t)$ la chaîne initialisée en x (i.e. telle que $X^x(0) = x$).

On appelle $T_{x \rightarrow x}$ le temps de retour en x , et $\mathbb{E}[T_{x \rightarrow x}]$ est donc le temps moyen de retour en x .

Remarque : $T_{x \rightarrow y}$ et $T_{x \rightarrow x}$ sont des temps d'arrêt pour la chaîne $X(t)$.

On va maintenant définir les notions de récurrence et de transience. Dans le cas général, on parle d'états récurrents ou transients, et une chaîne est récurrente ou transiente lorsque tous ses états le sont. Mais, ces propriétés étant des propriétés de classe (ensemble des états qui communiquent entre eux), pour une chaîne irréductible (qui ne possède donc qu'une seule classe d'états) il suffit de les vérifier pour un seul état de l'espace \mathcal{S} .

Définition 2.1. Soit $x \in \mathcal{S}$, la chaîne est récurrente si les temps de retour en x sont presque sûrement finis : $\mathbb{P}(T_{x \rightarrow x} < \infty) = 1$.

Si de plus $\mathbb{E}[T_{x \rightarrow x}] < \infty$, la chaîne est récurrente positive.

Et si $\mathbb{E}[T_{x \rightarrow x}] = \infty$, la chaîne est récurrente nulle.

Sinon ($\mathbb{P}(T_{x \rightarrow x} < \infty) < 1$) la chaîne est transiente.

Remarque :

- D'après la définition, on voit que si un état est récurrent alors, comme on y revient presque sûrement en un temps fini, on va y revenir une infinité de fois avec probabilité un.
- Lorsque la chaîne est récurrente positive, les temps moyens d'atteinte $\mathbb{E}[T_{x \rightarrow y}]$ sont également finis $\forall x, y \in \mathcal{S}$.

Le théorème suivant nous donne une condition d'existence et d'unicité de la loi invariante de la chaîne.

Théorème 2.2. *Une chaîne de Markov irréductible, récurrente positive admet une unique probabilité invariante μ , qui est donnée par la formule suivante (lemme de Kac)*

$$\mu(i) = \frac{1}{\mathbb{E}[T_{i \rightarrow i}]} , \quad \forall i \in \mathcal{S} \quad (2.9)$$

Remarque : pour une chaîne irréductible, lorsque l'espace d'états \mathcal{S} est fini, les temps moyens de retour sont finis (et donc les temps de retour sont presque sûrement finis). Ainsi

une chaîne irréductible à espace d'états fini est récurrente positive et le théorème précédent s'applique. C'est le cas que nous allons rencontrer le plus souvent dans notre travail.

Un cas particulier de mesures invariantes sont les mesures réversibles. On dit qu'une mesure de probabilité μ est réversible pour la chaîne si pour tous x et y dans \mathcal{S} on a

$$\mu(x)p(x, y) = \mu(y)p(y, x) . \quad (2.10)$$

c'est-à-dire que sous la mesure μ la probabilité pour la chaîne d'être en x et d'aller en y est égale à la probabilité d'être en y et d'aller en x . Autrement dit, sous la mesure μ la loi de la chaîne reste la même si on renverse le temps.

Familles de chaînes

Le phénomène de convergence abrupte et la métastabilité sont des phénomènes asymptotiques. Nous allons donc considérer des familles de chaînes de Markov indexées par un paramètre a qui va ensuite tendre vers l'infini. Dans ce cas on met l'indice de la famille en bas : $X_a(t)$.

Lorsqu'on veut spécifier l'état initial de la chaîne, on le met en indice en haut : $X^x(t)$ ou $X_a^x(t)$. Et lorsqu'on ne le précise pas, c'est que la propriété que l'on regarde est indépendante de l'état initial.

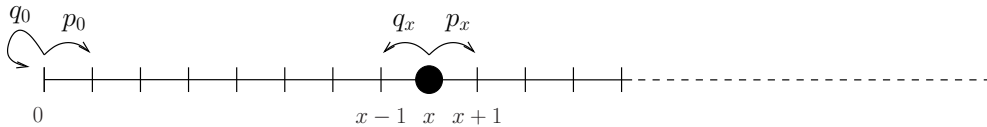
2.2 Chaînes de Naissance et de Mort

Cas général

Une chaîne de naissance et de mort est une marche aléatoire irréductible $X(t)$ à temps discret sur \mathbb{N} . C'est donc une chaîne de Markov dont les transitions sont données pour tout $x \in \mathbb{N}$, par :

$$\begin{cases} p_x = p(x, x+1) = \mathbb{P}(X(t+1) = x+1 \mid X(t) = x) , \\ q_x = p(x, x-1) = \mathbb{P}(X(t+1) = x-1 \mid X(t) = x) , \forall x > 0, \text{ et } q_0 = p(0, 0) , \end{cases} \quad (2.11)$$

avec $p_x, q_x > 0$ et $p_x + q_x = 1$, $\forall x \in \mathbb{N}$. (q_0 est ici la probabilité de rester en zéro).



On peut considérer que $X(t)$ représente l'effectif d'une population à l'instant t , et qu'à l'instant $t+1$, on a naissance d'un individu avec probabilité p_x , ou mort d'un individu avec probabilité q_x . C'est de là que vient l'appellation chaîne de naissance et de mort.

Remarque : on peut aussi penser qu'une particule se déplace sur l'axe des entiers. A chaque instant, la particule située à l'état x saute vers la droite avec probabilité p_x , ou vers la gauche avec probabilité $q_x = 1 - p_x$.

Pour les chaînes de naissance et de mort, on peut exprimer des conditions de récurrence et de transience en fonction des transitions p_x et q_x . On trouve une démonstration de ces résultats dans le livre de Hoel Port et Stone "Introduction to Stochastic Processes" [30] où les chaînes de naissance et de mort sont traitées en détail.

La chaîne est transiente si et seulement si

$$\sum_{x=1}^{\infty} \prod_{i=1}^x \frac{q_i}{p_i} < \infty . \quad (2.12)$$

Une condition nécessaire et suffisante pour que la chaîne soit récurrente est donc

$$\sum_{x=1}^{\infty} \prod_{i=1}^x \frac{q_i}{p_i} = \infty . \quad (2.13)$$

De plus

$$\sum_{x=1}^{\infty} \prod_{i=1}^x \frac{p_{i-1}}{q_i} < \infty , \quad (2.14)$$

est une condition nécessaire et suffisante pour que la chaîne soit récurrente positive. Dans ce cas, elle admet une unique probabilité invariante π déterminée pour tout x par

$$\pi(x) = \prod_{i=1}^x \frac{p_{i-1}}{q_i} \pi(0) , \quad (2.15)$$

avec $\pi(0) = 1 / \sum_{x=1}^{\infty} \prod_{i=1}^x \frac{p_{i-1}}{q_i}$.

Remarque : on a donc une condition nécessaire et suffisante pour que la chaîne soit récurrente nulle :

$$\sum_{x=1}^{\infty} \prod_{i=1}^x \frac{q_i}{p_i} = \infty , \quad \text{et} \quad \sum_{x=1}^{\infty} \prod_{i=1}^x \frac{p_{i-1}}{q_i} = \infty . \quad (2.16)$$

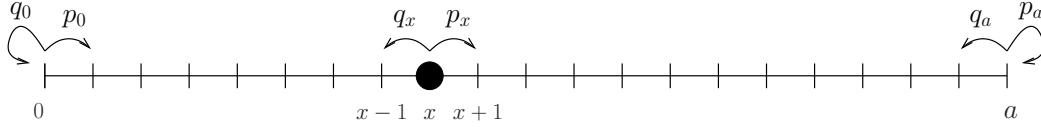
Chaînes de naissance et de mort sur un intervalle

Pour ce qui nous intéresse, nous allons considérer une famille $X_a(t)$ de chaînes de naissance et de mort sur l'intervalle $\llbracket 0, a \rrbracket = \{0, \dots, a\}$ de \mathbb{N} .

Lorsque a est fixé, on a donc une chaîne dont les transitions sont telles que

$$\begin{cases} p_x = p(x, x+1) = \mathbb{P}(X_a(t+1) = x+1 \mid X_a(t) = x), x = 0, \dots, a-1, \\ q_x = p(x, x-1) = \mathbb{P}(X_a(t+1) = x-1 \mid X_a(t) = x), x = 1, \dots, a, \end{cases} \quad (2.17)$$

avec $p_x + q_x = 1, \forall x$, et $p_a = p(a, a)$, $q_0 = p(0, 0)$ aux bords.



Comme l'espace d'états de la chaîne est fini, elle admet une unique mesure invariante, que l'on note π_a . De plus cette mesure est réversible, et on peut l'exprimer à l'aide des transitions. En effet par définition de la réversibilité, pour tout x on a

$$\pi_a(x-1) p_{x-1} = \pi_a(x) q_x \quad (2.18)$$

Soit

$$\pi_a(x) = \frac{p_{x-1}}{q_x} \pi_a(x-1), \quad (2.19)$$

et en itérant, on trouve, pour tout x :

$$\pi_a(x) = \prod_{i=1}^x \frac{p_{i-1}}{q_i} \pi_a(0). \quad (2.20)$$

2.3 Distance à l'équilibre pour les modèles d'absorption

Distance variationnelle

Une façon de mesurer l'écart entre deux mesures de probabilités μ et ν sur un espace \mathcal{X} est d'utiliser la distance variationnelle

$$\|\mu - \nu\|_{VT} = \max_{A \subset \mathcal{X}} \|\mu(A) - \nu(A)\| = \frac{1}{2} \sum_{x \in \mathcal{X}} |\mu(x) - \nu(x)|. \quad (2.21)$$

Cette distance prend donc en compte le "pire" évènement c'est-à-dire celui qui maximise la différence entre les deux mesures. Il existe d'autres distances pour mesurer l'écart entre deux mesures de probabilité, notamment la séparation (3.8), avec laquelle Diaconis et Aldous ont travaillé sur le phénomène de cutoff. On se reportera à l'article de Barrera, Lachaud et Ycart "Cutoff for n -tuples of processes" [4] pour une discussion sur les relations entre les différentes distances entre mesures (Variation totale, Hellinger, Chi-deux, Kullback) et la notion de cutoff pour ces distances.

Modèles d'absorption

On parle de modèle d'absorption pour une chaîne de Markov sur un espace \mathcal{X} , lorsqu'il existe un état $x_0 \in \mathcal{X}$ tel que $p(x_0, x_0) = 1$. On appelle un tel état un état absorbant pour la chaîne puisque lorsqu'il est atteint on ne peut plus en repartir. Lorsque c'est le cas, la masse de Dirac δ_{x_0} au point x_0 est une mesure de probabilité invariante pour la chaîne. De plus on peut relier la distance variationnelle de la loi de la chaîne à l'équilibre avec le temps d'atteinte de x_0 .

Lemme 2.3. *Soit X une chaîne de Markov sur un espace \mathcal{X} absorbée en $x_0 \in \mathcal{X}$.*

On note P^t sa loi au temps t et $\pi = \delta_{x_0}$. Soit T le temps d'atteinte de x_0 .

On a :

$$\|P^t - \pi\|_{VT} = \mathbb{P}(T > t) \quad (2.22)$$

Démonstration :

Comme T est le temps d'atteinte de x_0 et que cet état est absorbant (on parle de temps d'absorption), conditionné à $\{T \leq t\}$, la chaîne est à l'équilibre : pour tout $x \in \mathcal{X}$ on a

$$\mathbb{P}(X(t) = x \mid T \leq t) = \pi(x) = \delta_{x_0}(x) = \begin{cases} 1 & \text{si } x = x_0, \\ 0 & \text{sinon.} \end{cases} \quad (2.23)$$

Et par définition de la distance variationnelle

$$\begin{aligned} \|P^t - \pi\|_{VT} &= \frac{1}{2} \sum_{x \in \mathcal{X}} |P^t(x) - \pi(x)| \\ &= \frac{1}{2} \sum_{x \in \mathcal{X}} \left| \mathbb{P}(X(t) = x \mid T > t) \mathbb{P}(T > t) + \mathbb{P}(X(t) = x \mid T \leq t) \mathbb{P}(T \leq t) - \pi(x) \right| \\ &= \frac{1}{2} \sum_{x \in \mathcal{X}} \left| \mathbb{P}(X(t) = x \mid T > t) \mathbb{P}(T > t) + \pi(x) (1 - \mathbb{P}(T > t)) - \pi(x) \right| \\ &= \frac{1}{2} \mathbb{P}(T > t) \sum_{x \in \mathcal{X}} \left| \mathbb{P}(X(t) = x \mid T > t) - \pi(x) \right|. \end{aligned} \quad (2.24)$$

Or x_0 n'appartient pas au support de la loi du processus conditionné à ne pas être absorbé

$$\mathbb{P}(X(t) = x_0 \mid T > t) = 0 \quad \text{et} \quad \sum_{x \neq x_0} \mathbb{P}(X(t) = x \mid T > t) = 1. \quad (2.25)$$

Et π ne charge que x_0 , d'où

$$\sum_{x \in \mathcal{X}} \left| \mathbb{P}(X(t) = x \mid T > t) - \pi(x) \right| = \sum_{x \neq x_0} \mathbb{P}(X(t) = x \mid T > t) + \pi(x_0) = 2, \quad (2.26)$$

ce qui termine la démonstration du lemme.

□

2.4 Couplages

De façon générale, un couplage de deux chaînes de Markov $X(t)$ et $Y(t)$ est un processus couplé $Z(t) = (U(t), V(t))$ tel que $Z(t)$ soit une chaîne de Markov sur l'espace produit et tel que les marginales $U(t)$ et $V(t)$ ont même loi que $X(t)$ et $Y(t)$.

Les couplages sont souvent utilisés pour trouver des bornes supérieures à la distance entre les lois de deux processus, ou entre la loi du processus au temps t et sa mesure invariante (voir (3.7) et (3.9)). Dans ce cas on associe au couplage $Z(t) = (X(t), Y(t))$ un temps d'arrêt T dit temps de couplage, tel que $X(t) = Y(t)$ lorsque $t \geq T$.

Dans [3], Aldous et Diaconis prennent $X(t)$ une chaîne de transition P , $Y(t)$ une chaîne de transition P , démarrée sous la mesure π invariante pour P (à tout instant t , $Y(t)$ est donc à l'équilibre, on dit que c'est une chaîne stationnaire), et T tel que $X(t) = Y(t)$ lorsque $t \geq T$.

Pour le couplage indépendant coalescent que nous allons utiliser pour prouver la métastabilité (5.54), on couple deux versions de la même chaîne avec des états initiaux différents, et le temps de couplage est le premier instant où les deux copies se rencontrent.

Chapitre 3

Convergence abrupte et phénomène de Cutoff

Le phénomène de cutoff est un phénomène asymptotique de convergence abrupte à l'équilibre. Il existe un temps autour duquel, asymptotiquement en un certain paramètre, la loi du processus converge de façon abrupte vers la mesure d'équilibre. C'est donc un phénomène quasi-déterministe : déterministe parce qu'on sait à partir de quel instant on va atteindre l'équilibre, quasi car cette convergence abrupte a lieu asymptotiquement (l'instant où la convergence a lieu tend vers l'infini avec le paramètre asymptotique, et elle n'est abrupte qu'à la limite).

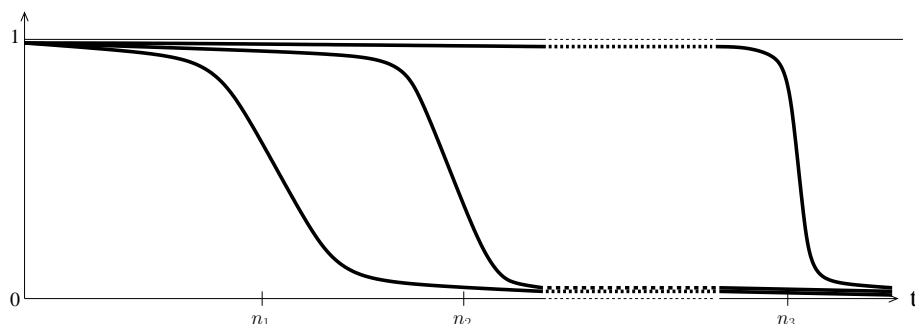


FIG. 3.1 – La convergence est asymptotiquement de plus en plus abrupte ($n_1 < n_2 < n_3$)

Aldous et Diaconis ont découvert ce phénomène au début des années 80, et ils ont montré son existence pour différents modèles (notamment les modèles de battage de cartes), en reliant le temps de cutoff avec les propriétés spectrales de la matrice de transition du modèle.

Ensuite Martínez et Ycart [38] ont relié le cutoff pour la convergence vers la mesure d'équilibre avec une notion de cutoff pour des temps d'absorption.

Après un survol historique des travaux concernant le phénomène (section 3.1), nous définissons la notion de convergence abrupte pour des temps d'atteinte (définition 3.1), et

donnons une condition suffisante pour son existence (proposition 3.2), dans l'esprit des travaux de Martínez et Ycart (section 3.2). Cette approche du phénomène est plus physique que mathématicienne dans le sens où en physique, les états d'un système sont les configurations, alors que pour les probabilistes les états sont les mesures de probabilité. Cela va nous permettre de faire le lien avec la métastabilité définie au chapitre 4, et on verra au chapitre 5 la complémentarité des deux phénomènes.

Dans la dernière section de ce chapitre (section 3.3), on illustre en étudiant des exemples très simples, notre approche du phénomène.

3.1 Historique et définitions

Combien de fois doit-on battre un jeu de cartes afin qu'il soit bien mélangé ?

C'est pour répondre à cette question que Persi Diaconis s'est intéressé au phénomène de convergence abrupte vers l'équilibre qu'avec David Aldous ils ont dénommé “cutoff phenomenon” [2], ou “threshold phenomenon” [3], ou encore “abrupt switch” [1].

Les modèles de battage de cartes, sont des cas particuliers de marches aléatoires sur le groupe symétrique, et le théorème de Perron-Frobenius nous dit que leurs lois convergent exponentiellement vers la mesure uniforme sur le groupe. Mais cela ne nous dit pas précisément la façon dont cette convergence a lieu, et ne permet pas de répondre à la question initialement posée.

Le phénomène de Cutoff permet de répondre que pour le battage “riffle shuffle” par exemple, sept battages sont suffisants pour mélanger un jeu de 52 cartes .

Si on mesure l'écart entre la distribution P^t des cartes au temps t et la mesure uniforme U à l'aide d'une distance sur les mesures (variation totale, séparation...) :

$$d(t) = \|P^t - U\|, \quad (3.1)$$

le théorème de Perron-Frobenius nous dit que $d(t)$ tend vers zéro exponentiellement en la seconde valeur propre lorsque t tend vers l'infini : $d(t) \sim C\lambda^t$, où C est une constante positive et λ la plus grande (en valeur absolue) valeur propre différente de 1 de la matrice de transition de la chaîne.

Cela ne donne aucune indication sur l'instant à partir duquel on est proche de l'équilibre, mais seulement qu'il faut attendre “suffisamment longtemps” pour l'atteindre.

Le phénomène de cutoff, nous dit qu'asymptotiquement en n , où n est ici le nombre de cartes dans le jeu, il existe un temps t_n qui tend vers l'infini avec n , auquel la convergence a lieu de manière abrupte : la fonction $d(t)$ est proche de 1 avant t_n , et tombe à zéro juste après. Dans “Shuffling cards and stopping times” [2] Aldous et Diaconis le formalisent de la manière suivante :

$$d(t_n + o(t_n)) \approx 0 \text{ et } d(t_n - o(t_n)) \approx 1. \quad (3.2)$$

Dans cet article ils définissent la notion de temps uniforme fort “Strong Uniform Time” (qu'ils généraliseront plus tard en temps stationnaire fort “Strong Stationary Time”) :

un temps d'arrêt T est un SUT pour le processus $X(t)$ si conditionné à $\{T = t\}$, le processus est distribué suivant la mesure uniforme (et c'est un SST pour une mesure stationnaire quelconque).

$$\mathbb{P}(X(t) = x \mid T = t) = U(x) \quad (3.3)$$

Cela leur permet d'obtenir des bornes supérieures pour $d(t)$, où la norme utilisée est la variation totale :

$$\|\mu - \nu\|_{VT} = \sup_{A \subset \mathcal{S}} |\mu(A) - \nu(A)| \quad (3.4)$$

$$= \frac{1}{2} \sum_{x \in \mathcal{S}} |\mu(x) - \nu(x)|, \quad (3.5)$$

pour μ et ν mesures sur l'espace \mathcal{S} . En effet, on a

$$\begin{aligned} P^t(A) - U(A) &= \mathbb{P}(X(t) \in A, T > t) + \mathbb{P}(X(t) \in A, T \leq t) - U(A) \\ &= \mathbb{P}(X(t) \in A \mid T > t) \mathbb{P}(T > t) + \mathbb{P}(X(t) \in A \mid T \leq t) \mathbb{P}(T \leq t) - U(A) \\ &= \mathbb{P}(X(t) \in A \mid T > t) \mathbb{P}(T > t) + U(A) (1 - \mathbb{P}(T > t)) - U(A) \\ &= (\mathbb{P}(X(t) \in A \mid T > t) - U(A)) \mathbb{P}(T > t). \end{aligned} \quad (3.6)$$

Et donc

$$\|P^t - U\|_{VT} \leq \mathbb{P}(T > t). \quad (3.7)$$

En construisant des SUT, ils obtiennent ainsi des bornes supérieures pour plusieurs modèles de marches aléatoires sur les groupes ("Top in at random shuffle", "riffle shuffle"...)

et montrent que le modèle de battage de cartes "Top in at random shuffle" (on prend la carte du dessus du paquet et on l'insère au hasard) admet un cutoff au temps $n \log n$.

Dans ce cas, le SUT est le premier instant où la carte initialement en dessous arrive sur le dessus du paquet, et pour obtenir une borne inférieure pour $d(t)$, ils exhibent un événement A tel que $P^{t_n}(A) \xrightarrow{n \rightarrow \infty} 1$, où $t_n = n \log n - c_n n$, avec $c_n \rightarrow \infty$.

Ils introduisent aussi la "distance" (qui n'est pas une métrique) de séparation

$$s(\mu, \nu) = \max_{x \in \mathcal{S}} \left(1 - \frac{\mu(x)}{\nu(x)} \right), \quad (3.8)$$

qui est une borne supérieure pour la variation totale ($d(t) \leq s(t)$) et montrent que l'on a également

$$s(t) \leq \mathbb{P}(T > t), \quad (3.9)$$

si T est un temps uniforme fort pour le processus.

Dans "Strong Uniform Times and Finite Random Walks" [3], ils montrent qu'il existe toujours un SUT pour lequel on a égalité dans (3.9), et qu'il existe un couplage tel qu'on ait égalité dans (3.7) pour le temps de couplage T .

Remarques :

1. les temps uniformes forts sont un cas particulier de couplage : étant donné un SUT T , il existe un couplage de temps de couplage T .
2. l'inégalité (3.7) reste vraie pour un temps de couplage T .
3. on peut donc construire des temps uniformes forts (resp. des couplages) pour obtenir des bornes supérieures pour la séparation (resp. la variation totale). Dans leur article Aldous et Diaconis en donnent pour la marche aléatoire sur le cube (ou modèle d'Ehrenfest), mais la construction d'un temps minimal (i.e. qui donne égalité dans (3.9)) ou d'un couplage qui donne égalité dans (3.7) ("maximal" au sens de Griffeath [27, 28, 29], voir aussi Pitman [46], Goldstein [26], Thorisson [53]) est uniquement théorique, car elle requiert la connaissance de la loi du processus à tout instant.
4. Dans "Random walks on finite groups and rapidly mixing Markov chains" [1], Aldous donne des couplages pour la marche sur le cube et différents modèles de battage de cartes et obtient des estimations de leurs temps de cutoff (ici il considère le temps "de mélange" (mixing time) $\tau(\varepsilon) = \min\{t : d(t) \leq \varepsilon\}$).

Pour un survol de ces résultats sur le phénomène de cutoff, on pourra lire "The cutoff phenomenon in finite Markov chains" [17], dans lequel Diaconis rappelle les différents exemples qu'il a étudiés et les résultats qu'il a obtenus ([2, 3, 5, 13, 19, 20, 22]). Il discute aussi des questions de symétrie de la matrice de transition et de multiplicité de la seconde valeur propre de cette matrice qui selon lui sous-tendent l'existence du phénomène ("*cutoff phenomena occur because of high multiplicity of the second eigenvalue*", section 4 de [17]). Pour une discussion plus détaillée sur les marches aléatoires sur les groupes, on lira "Group Representations in Probability and Statistics" [16]. Pour un panorama récent sur les battages de cartes, on lira "Mathematical developments from the analysis of riffle shuffling" [18].

"Random walks on finite groups" [48] est un autre panorama sur les marches aléatoires sur les groupes et le phénomène de cutoff. Dans cet article, Saloff-Coste présente les techniques de couplage, de temps de mélange et de temps uniforme fort pour différents modèles de battage de cartes, ainsi qu'une grande variété d'autres techniques venant de différents champs tels que l'algèbre, la combinatoire, mais surtout la géométrie, l'analyse fonctionnelle et la théorie de la représentation. Il définit aussi une notion de précutoff, plus faible, qui impose seulement que la distance en variation totale à l'équilibre soit strictement positive avant le temps t_n , et qui donne l'ordre de grandeur d'un possible temps de cutoff. De la même façon, dans "Lectures on finite Markov chains" [47], il avait défini une notion de cutoff faible ℓ^p en considérant la chaîne continue associée à la matrice de transition K_n de la chaîne en temps discret. Une famille de chaînes de Markov admet un cutoff faible ℓ^p au temps critique t_n , si $t_n \rightarrow \infty$ et

$$\lim_{n \rightarrow \infty} \|h_{n,t_n} - 1\| > 0 \text{ et } \lim_{n \rightarrow \infty} \|h_{n,(1+\varepsilon)t_n} - 1\| = 0, \quad (3.10)$$

où $h_{n,t}$ est la densité de la chaîne continue associée au temps t par rapport à sa mesure invariante π_n , et la norme considérée est la norme $\ell^p(\pi_n)$.

Le résultat principal de ce papier est de relier le trou spectral λ_n (i.e. la plus petite valeur propre non nulle de $I - K_n$) avec le temps de mélange ℓ^p : $\tau_n = \inf\{t > 0 : \|h_t - 1\|_p < \varepsilon\}$. Si $\lim_n \lambda_n t_n = \infty$, alors la famille de chaînes admet un cutoff faible ℓ^p au temps τ_n , et si $p = 2$, la condition est nécessaire et suffisante. De plus, il peut estimer que la taille de la fenêtre de convergence (ordre de grandeur de la transition abrupte) est de l'ordre de $\frac{1}{\lambda_n}$. Ce qui a amené Peres a conjecturer que si $\lambda_n \tau_n \rightarrow \infty$, alors la convergence abrupte a lieu. Ce qui n'est pas vrai en général pour la variation totale (contre exemple d'Aldous), mais que Diaconis et Saloff-Coste ont démontré pour une famille de chaînes de naissance et de mort en considérant la distance de séparation dans "Separation cutoff for birth and death chains [21].

Un autre auteur à s'être intéressé au phénomène est Ycart. Dans "Cutoff for samples of Markov chains" [55] il montre le cutoff en variation totale pour des échantillons de taille n de chaînes discrètes irréductibles et réversibles sur un espace d'états fini E , ainsi que pour les chaînes en temps continu associées, et donne leurs temps de cutoff en fonction du spectre de leurs matrices de transition. Ce sont donc des chaînes sur l'espace produit E^n , et leur mesure invariante est la mesure produit $\pi_n = \pi^{\otimes n}$ si π est la mesure invariante de la chaîne initiale sur E .

Une chaîne d'échantillonnage parallèle (n copies indépendantes qui évoluent simultanément) d'une chaîne discrète irréductible et réversible sur un espace d'états fini admet un cutoff au temps

$$t_n = \frac{\log n}{2 \log(1/\alpha)}, \quad (3.11)$$

où α est la valeur propre différente de 1 la plus grande en valeur absolue de la matrice de transition P de la chaîne. Si maintenant, on couple n copies indépendantes d'une chaîne continue de générateur Λ , alors il y a cutoff au temps

$$t_n = \frac{\log n}{2\beta}, \quad (3.12)$$

où β est la plus petite valeur propre non nulle de Λ en valeur absolue. Dans ce cas continu, on change une seule coordonnée à la fois, et l'équivalent en temps discret est la chaîne d'échantillonnage séquentielle de matrice de transition $Q_n = I_{E^n} + \frac{1}{n\lambda}\Lambda_n$, où $\lambda = \max_{i \in E} \sum_{j \neq i} \lambda_{ij}$, pour laquelle on a cutoff au temps

$$t_n = \frac{n\lambda \log n}{2\beta}. \quad (3.13)$$

"Cutoff for Markov chains : some examples and applications" [56] regroupe ces résultats avec d'autres exemples de cutoff, des tables numériques et des algorithmes pour produire des échantillons, ainsi que les résultats de "Decay rates and cutoff for convergence and hitting times of Markov chains with countably infinite state space" [38], en collaboration avec Martínez sur lequel on reviendra un peu plus loin.

On peut trouver un autre survol rapide sur la notion de cutoff et ces résultats dans le livre

de Ycart “Modèles et algorithmes markoviens” [57] au chapitre “Convergence abrupte d’un échantillon”.

Plus récemment, dans “Cutoff for n -tuples of processes” [4] Ycart avec J. Barrera et B. Lachaud se sont intéressés au cutoff en termes de taux exponentiels de convergence à l’équilibre et non plus en termes de valeurs propres de la matrice de transition.

Ils considèrent un échantillon (X_1, \dots, X_n) de processus exponentiellement convergents au sens où chaque processus X_i vérifie

$$\lim_{t \rightarrow \infty} \frac{\log d_i(t)}{t} = -\rho_i, \quad (3.14)$$

où $d_i(t)$ est la distance à l’équilibre de la loi du processus X_i au temps t pour la distance considérée (Variation totale, Hellinger, Chi-deux, Kullback).

Sous certaines hypothèses de régularité (notamment que la convergence vers zéro des $d_i(t)$ aux taux exponentiels ρ_i soit uniforme en i), ils montrent que l’échantillon admet un cutoff au temps t_n défini par

$$t_n = \max_{i=1, \dots, n} \left\{ \frac{\log i}{2\rho_{i,n}} \right\} \quad (3.15)$$

où les $\rho_{i,n}$ sont les taux de convergence ρ_i rangés par ordre croissant (Th 4.1).

Si de plus l’échantillon est i.i.d. de taux de convergence exponentiel ρ , alors la convergence a lieu dans un intervalle de temps de longueur $O(1)$ autour du temps de cutoff, qui dans ce cas vaut (Th 4.4 et 4.5) :

$$t_n = \frac{\log n}{2\rho} \quad (3.16)$$

Ce résultat généralise la proposition 7.7 de “Strong Uniform Times and Finite Random Walks” [3], dans laquelle Aldous et Diaconis montrent que le produit indépendant d’une chaîne exponentiellement convergente ($d(t) \sim \exp(-n/\tau_\varepsilon)$) admet un cutoff au temps $\tau_d = \tau_\varepsilon \log n$ pour la variation totale et au temps $\tau_s = \frac{1}{2}\tau_\varepsilon \log n$ pour la séparation.

Mais la nouveauté ici est que contrairement à Aldous et Diaconis, le temps de cutoff n’est pas exprimé en fonction des valeurs propres de la matrice de transition ($\tau_\varepsilon = -1/\log |\lambda_2|$, où λ_2 est la seconde valeur propre), mais est relié à la concentration de la mesure produit.

Une autre approche du phénomène développée par Ycart [55, 56], est de relier le temps de cutoff avec un temps d’atteinte de moyenne empirique. Il considère la moyenne empirique de l’échantillon

$$M^{(n)}(f)(t) = \frac{1}{n} \sum_{i=1}^n f(X_i(t)), \quad (3.17)$$

pour une fonction test f . Cette moyenne converge vers $\pi_n(f)$ et il montre que le temps d’atteinte

$$T^{(n)}(f) = \inf\{t \geq 0 : M^{(n)}(f)(t) = \pi_n(f)\} \quad (3.18)$$

est asymptotiquement équivalent au temps de cutoff pour les chaînes à espace d’états fini. Ceci fournit une règle d’arrêt pour la simulation de la mesure π par méthode de Monte-Carlo. Pour simuler cette mesure, pour un certain n fixé, on génère un échantillon i.i.d.

de taille n d'une chaîne de Markov admettant π pour mesure invariante, et on s'arrête lorsque ce temps d'atteinte empirique est atteint. Comme l'échantillon converge de manière abrupte et que le temps de cutoff est atteint, on obtient donc une réalisation de la mesure π .

Ce résultat est généralisé dans “Convergence times for parallel Markov chains” [35] pour des processus à espace d'états infini, dans le cadre des processus exponentiellement convergents de [4]. La méthode est illustrée avec les processus d'Ornstein-Uhlenbeck, pour lesquels B. Lachaud a démontré le phénomène de cutoff (“Cutoff and hitting times for a sample of Ornstein-Uhlenbeck processes and its average” [34]).

3.2 Convergence abrupte pour un temps d'atteinte

Nous revenons dans cette partie sur le travail de Martínez et Ycart dans “Decay rates and cutoff for convergence and hitting times of Markov chains with countably infinite state space” [38], afin de motiver notre définition de convergence abrupte pour un temps d'atteinte.

Dans cet article Martínez et Ycart se placent sur un espace d'états infini dénombrable et considèrent une chaîne $X_a(t)$ irréductible récurrente positive d'état initial a . C'est ici a qui est le paramètre asymptotique qui tend vers l'infini dans le phénomène de cutoff. Ils se placent également en temps continu, mais les résultats ne diffèrent pas du cas discret. Cette chaîne admet une mesure invariante π (car elle est récurrente positive), et le résultat principal de cet article (Th 4.1) est que la chaîne admet un cutoff pour la convergence à l'équilibre au temps t_a si et seulement si la chaîne absorbée associée admet un cutoff au même instant pour le temps d'atteinte d'un état absorbant.

Supposons que l'on choisisse que 0 devienne un état absorbant. On obtient une nouvelle chaîne $X_a^{(0)}(t)$ dite chaîne absorbée associée, qui a δ_0 pour mesure invariante et dont la loi $P_{(0),a}^t$ est égale à celle de $X_a(t)$ tant que zéro n'est pas atteint. C'est-à-dire que $P_{(0),a}^t = P_a^t$ tant que t est strictement inférieur au temps d'atteinte de zéro T_a défini par

$$T_a = \inf\{t > 0 : X_a(t) = 0\}. \quad (3.19)$$

Ce théorème nous dit que $\|P_a^t - \pi\|_{VT}$ admet un cutoff au temps t_a si et seulement si $\|P_{(0),a}^t - \delta_0\|_{VT}$ admet un cutoff au même instant.

Or on a vu (lemme 2.3) que pour un modèle d'absorption on a

$$\|P_{(0),a}^t - \delta_0\|_{VT} = \mathbb{P}(T_a > t). \quad (3.20)$$

Donc la chaîne $X_a^{(0)}(t)$ admet un cutoff au temps t_a si et seulement si le temps d'absorption T_a est asymptotiquement équivalent en probabilité au temps de cutoff t_a :

$$\lim_{a \rightarrow \infty} \mathbb{P}\left(\left|\frac{T_a}{t_a} - 1\right| > \varepsilon\right) = 0, \quad \forall \varepsilon > 0. \quad (3.21)$$

Car (3.21) est équivalent à

$$\lim_{a \rightarrow \infty} \mathbb{P}(T_a > ct_a) = \begin{cases} 1 & \text{si } c < 1, \\ 0 & \text{si } c > 1. \end{cases} \quad (3.22)$$

En effet, supposons que l'on ait (3.22), et soit $\varepsilon > 0$. On a

$$\begin{aligned} \mathbb{P}\left(\left|\frac{T_a}{t_a} - 1\right| > \varepsilon\right) &= \mathbb{P}(T_a > (1 + \varepsilon)t_a, T_a > t_a) + \mathbb{P}(T_a < (1 - \varepsilon)t_a, T_a < t_a) \\ &\leq \mathbb{P}(T_a > (1 + \varepsilon)t_a) + \mathbb{P}(T_a \leq (1 - \varepsilon)t_a) \\ &= \mathbb{P}(T_a > (1 + \varepsilon)t_a) + 1 - \mathbb{P}(T_a > (1 - \varepsilon)t_a) \end{aligned} \quad (3.23)$$

qui tend vers zéro lorsque a tend vers l'infini, par hypothèse en prenant $c = 1 + \varepsilon$ pour le premier terme et $c = 1 - \varepsilon$ pour le dernier.

Supposons maintenant que l'on ait (3.21), soit $c > 1$, i.e. $c = 1 + \varepsilon$ avec $\varepsilon > 0$, on a

$$\begin{aligned} \mathbb{P}(T_a > ct_a) &= \mathbb{P}(T_a > (1 + \varepsilon)t_a) \\ &= \mathbb{P}\left(\frac{T_a}{t_a} - 1 > \varepsilon\right) \\ &\leq \mathbb{P}\left(\left|\frac{T_a}{t_a} - 1\right| > \varepsilon\right) \end{aligned} \quad (3.24)$$

qui tend vers zéro lorsque a tend vers l'infini par hypothèse.

Et soit $c < 1$, i.e. $c = 1 - \varepsilon$ avec $\varepsilon > 0$, et soit $\varepsilon' < \varepsilon$, on a

$$\begin{aligned} \mathbb{P}(T_a > ct_a) &= \mathbb{P}(T_a > (1 - \varepsilon)t_a) \\ &\geq \mathbb{P}(T_a \geq (1 - \varepsilon')t_a) \\ &= 1 - \mathbb{P}(T_a < (1 - \varepsilon')t_a) \\ &= 1 - \mathbb{P}\left(1 - \frac{T_a}{t_a} > \varepsilon'\right) \\ &\geq 1 - \mathbb{P}\left(\left|1 - \frac{T_a}{t_a}\right| > \varepsilon'\right) \end{aligned} \quad (3.25)$$

qui tend vers 1 lorsque a tend vers l'infini, puisque par hypothèse, le dernier terme tend vers zéro.

On a donc équivalence entre cutoff pour la convergence à l'équilibre et cutoff pour le temps d'atteinte de l'état absorbant, c'est-à-dire que la chaîne initiale $X_a(t)$ admet un cutoff au temps t_a si et seulement si le temps d'absorption T_a satisfait (3.21).

Ceci motive notre définition de convergence abrupte pour un temps d'atteinte :

Définition 3.1. Le temps d'arrêt T_a admet une convergence abrupte au temps t_a si

$$\lim_{a \rightarrow \infty} \mathbb{P} \left(\left| \frac{T_a}{t_a} - 1 \right| > \varepsilon \right) = 0, \quad \forall \varepsilon > 0. \quad (3.26)$$

Remarque : cette condition implique aussi que

$$\lim_{a \rightarrow \infty} \mathbb{E} \left[\frac{T_a}{t_a}, \frac{T_a}{t_a} \leq K \right] = 1, \quad (3.27)$$

pour une constante $K > 1$.

En effet la variable $\tau_a = (T_a/t_a) \mathbf{1}_{\{T_a/t_a \leq K\}}$ est bornée par K uniformément en a , et donc la famille $(\tau_a)_a$ est uniformément intégrable (voir le chapitre 13 de “Probability with martingales” de Williams [54]).

Comme par hypothèse T_a/t_a converge vers 1 en probabilité, τ_a aussi ($K > 1$), et d'après le théorème 13.7 de [54], $\mathbb{E}[\tau_a]$ converge vers 1.

Donc un éventuel temps de cutoff sera équivalent à $\mathbb{E}[T_a]$ dans ce cas.

Ainsi, on obtient une condition suffisante pour qu'un temps d'atteinte T_a admette une convergence abrupte au temps $\mathbb{E}[T_a]$ (voir [38]) :

Proposition 3.2. Si

$$\lim_{a \rightarrow \infty} \mathbb{E}[T_a] = \infty \quad \text{et} \quad \lim_{a \rightarrow \infty} \text{Var} \left(\frac{T_a}{\mathbb{E}[T_a]} \right) = 0, \quad (3.28)$$

alors le temps d'atteinte T_a converge abruptement au temps $\mathbb{E}[T_a]$.

Démonstration : La première condition nous assure que le temps de cutoff tend bien vers l'infini.

Soit $\varepsilon > 0$, supposons que l'on ait

$$\lim_{a \rightarrow \infty} \text{Var} \left(\frac{T_a}{\mathbb{E}[T_a]} \right) = 0,$$

comme

$$\mathbb{E} \left[\frac{T_a}{\mathbb{E}[T_a]} \right] = 1, \quad (3.29)$$

on applique l'inégalité de Tchebychev

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq t) \leq \frac{\text{Var } X}{t^2} \quad (3.30)$$

à la variable $X = \frac{T_a}{\mathbb{E}[T_a]}$. Cela donne

$$\mathbb{P} \left(\left| \frac{T_a}{t_a} - 1 \right| > \varepsilon \right) \leq \frac{1}{\varepsilon^2} \text{Var} \left(\frac{T_a}{\mathbb{E}[T_a]} \right), \quad (3.31)$$

et donc

$$\lim_{a \rightarrow \infty} \mathbb{P} \left(\left| \frac{T_a}{t_a} - 1 \right| > \varepsilon \right) = 0. \quad (3.32)$$

□

3.3 Exemples extrêmes de cutoff

Nous allons illustrer cette approche avec deux modèles d'absorption très simples. Ici la loi invariante du processus est une mesure de Dirac, et on a donc cutoff au temps moyen d'absorption.

Évolution déterministe

Pour illustrer la notion de convergence abrupte, le plus simple est de considérer un modèle d'évolution déterministe. On regarde une particule qui se déplace sur l'axe des entiers entre zéro et a . À chaque instant, la particule initialement en a saute d'un pas vers la gauche, et elle s'arrête lorsqu'elle atteint zéro. Il n'y a donc pas d'aléatoire, et la particule va atteindre zéro au temps $t = a$. C'est un cas particulier de chaîne de naissance et de mort, d'état initial a , dont les transitions sont pour tout i : $q_i = 1$ vers la gauche, et $p_i = 0$ vers la droite.

La loi de la chaîne au temps t est donc $P_a^t = \delta_{a-t \vee 0}$, et sa mesure invariante est la masse de Dirac en zéro $\pi_a = \delta_0$ (on dit que zéro est un état absorbant).

Soit T_a le temps d'atteinte de zéro :

$$T_a = \inf\{t > 0 : X_a(t) = 0\}. \quad (3.33)$$

Il est ici déterministe et vaut a ($\mathbb{P}(T_a = a) = 1$). En particulier on a

$$\mathbb{P}(T_a > t) = \begin{cases} 1 & \text{si } t < a, \\ 0 & \text{si } t \geq a. \end{cases} \quad (3.34)$$

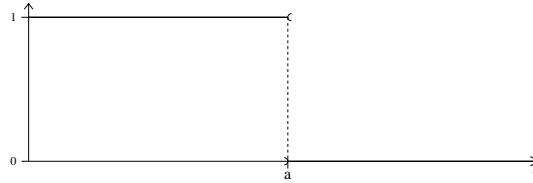
Or, comme pour tout modèle d'absorption on a égalité entre la distance variationnelle à l'équilibre et la queue de la distribution du temps d'atteinte de l'état absorbant (lemme 2.3)

$$\|P_a^t - \pi_a\|_{VT} = \mathbb{P}(T_a > t) \quad (3.35)$$

Donc

$$\|P_a^t - \pi_a\|_{VT} = \begin{cases} 1 & \text{si } t < a, \\ 0 & \text{si } t \geq a, \end{cases} \quad (3.36)$$

et on a un exemple extrême de transition abrupte au temps $t_a = a$:



De plus on a bien $t_a = \mathbb{E}[T_a] (= a)$ et la proposition 3.2 s'applique de manière brutale puisque

$$\text{Var}\left(\frac{T_a}{\mathbb{E}[T_a]}\right) = 0 \quad (3.37)$$

Modèle basique d'absorption

On considère un modèle simple d'absorption avec lequel on peut mettre en évidence le phénomène de cutoff dans le cadre des chaînes parallèles définies par Ycart.

Soit $p > 0$, on prend n copies indépendantes (X^1, \dots, X^n) d'une chaîne X^0 sur $\{0, 1\}$, initialisée en 1 et de transitions :

$$p(1, 0) = p, \quad p(1, 1) = 1 - p, \quad p(0, 0) = 1. \quad (3.38)$$

A chaque instant, on a une probabilité $1 - p$ de rester en 1, et une probabilité p d'être absorbée en 0, et la mesure invariante de chaque copie est la masse de Dirac en zéro δ_0 .

Soit $X = (X^1, \dots, X^n)$ la chaîne parallèle, c'est une chaîne sur $\{0, 1\}^n$, initialisée en $\underline{1} = (1, \dots, 1)$ et absorbée en $\underline{0} = (0, \dots, 0)$, de mesure invariante $\pi_n = \delta_{\underline{0}}$ et de matrice de transition donnée par le produit de la matrice P de la chaîne X^0 .

Soit T^i le temps d'absorption en 0 de la i^{eme} copie X^i , et T_n le temps d'absorption en $\underline{0}$ de la chaîne parallèle X :

$$T^i = \inf\{t > 0 : X^i(t) = 0\} \quad \text{et} \quad T_n = \sup_{i=1, \dots, n} T^i, \quad (3.39)$$

et soit P_n^t sa loi au temps t . On a

$$\|P_n^t - \pi_n\|_{VT} = \mathbb{P}(T_n > t) = 1 - (1 - (1 - p)^t)^n. \quad (3.40)$$

En effet $\mathbb{P}(T^0 > t) = (1 - p)^t$, et on a :

$$\begin{aligned} \mathbb{P}(T_n > t) &= 1 - \mathbb{P}(\sup T^i \leq t) \\ &= 1 - (\mathbb{P}(T^0 \leq t))^n \\ &= 1 - (1 - \mathbb{P}(T^0 > t))^n \\ &= 1 - (1 - (1 - p)^t)^n. \end{aligned} \quad (3.41)$$

Le graphe de $\|P_n^t - \pi_n\|_{VT}$ a donc l'allure suivante :

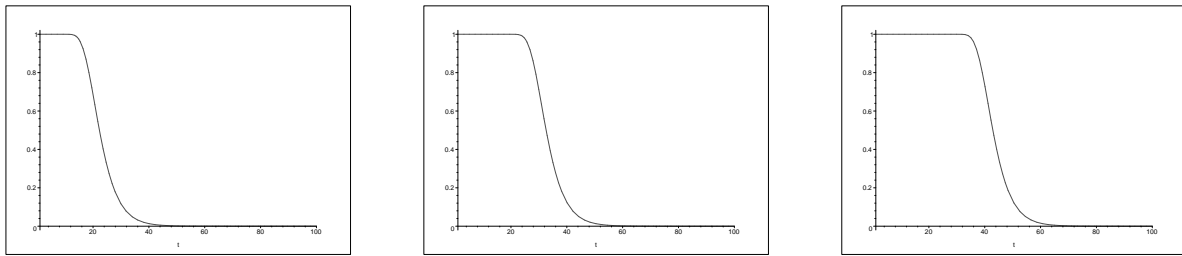


FIG. 3.2 - $\|P_n^t - \pi_n\|_{VT}$ avec $p = 0.2$ pour $n = 100, 1000$ et 10000

et la chaîne admet un cutoff au temps $t_n = \frac{-\log n}{\log(1 - p)}$.

Ici aussi notre définition de convergence abrupte pour le temps d'absorption T_n s'applique, puisque

$$t_n \sim \mathbb{E}[T_n], \quad (3.42)$$

où le symbole \sim signifie que les deux quantités comparées sont asymptotiquement équivalentes à des constantes près : $u_a \sim v_a$ équivaut à $\exists C_1, C_2 > 0$ telles que

$$C_1 \leq \lim_{a \rightarrow \infty} \frac{u_a}{v_a} \leq C_2 . \quad (3.43)$$

De plus ici $C_1 = C_2 = 1$.

En effet,

$$\mathbb{E}[T_n] = \sum_{t \geq 0} \mathbb{P}(T_n > t) = \sum_{t \geq 0} [1 - (1 - (1 - p)^t)^n] , \quad (3.44)$$

et la fonction $t \rightarrow [1 - (1 - (1 - p)^t)^n]$ est décroissante positive, et pour toute fonction décroissante positive f , on peut encadrer $\sum_{t \geq 0} f(t)$ par comparaison somme intégrale :

$$\int_{t \geq 0} f(t) dt \leq \sum_{t \geq 0} f(t) \leq 1 + \int_{t \geq 0} f(t) dt . \quad (3.45)$$

Or, en utilisant la formule du binôme pour $(1 - (1 - p)^t)^n$, on a

$$\begin{aligned} (1 - (1 - p)^t)^n &= \sum_{k=0}^n \binom{n}{k} (-1)^k (1 - p)^{kt} \\ &= 1 + \sum_{k=1}^n \binom{n}{k} (-1)^k (1 - p)^{kt} . \end{aligned} \quad (3.46)$$

Donc

$$\begin{aligned} \int_{t \geq 0} [1 - (1 - (1 - p)^t)^n] dt &= \int_{t \geq 0} - \sum_{k=1}^n \binom{n}{k} (-1)^k (1 - p)^{kt} dt \\ &= - \sum_{k=1}^n \binom{n}{k} (-1)^k \int_{t \geq 0} (1 - p)^{kt} dt \\ &= \sum_{k=1}^n \binom{n}{k} (-1)^k \frac{1}{k \log(1 - p)} , \end{aligned} \quad (3.47)$$

car

$$\int_{t \geq 0} (1 - p)^{kt} dt = \frac{-1}{k \log(1 - p)} . \quad (3.48)$$

D'où

$$\frac{1}{\log(1 - p)} \sum_{k=1}^n \binom{n}{k} \frac{(-1)^k}{k} \leq \mathbb{E}[T_n] \leq 1 + \frac{1}{\log(1 - p)} \sum_{k=1}^n \binom{n}{k} \frac{(-1)^k}{k} . \quad (3.49)$$

Or on a le résultat suivant :

Lemme 3.3.

$$\sum_{k=1}^n \binom{n}{k} \frac{(-1)^k}{k} = - \sum_{k=1}^n \frac{1}{k} \quad (3.50)$$

Démonstration :

Soit

$$S_n = \sum_{k=1}^n \binom{n}{k} \frac{(-1)^k}{k} . \quad (3.51)$$

On a

$$\begin{aligned} S_{l+1} - S_l &= \sum_{k=1}^{l+1} \binom{l+1}{k} \frac{(-1)^k}{k} - \sum_{k=1}^l \binom{l}{k} \frac{(-1)^k}{k} \\ &= \frac{(-1)^{l+1}}{l+1} + \sum_{k=1}^l \left(\binom{l+1}{k} - \binom{l}{k} \right) \frac{(-1)^k}{k} \\ &= \frac{(-1)^{l+1}}{l+1} + \sum_{k=1}^l \frac{k}{l+1} \binom{l+1}{k} \frac{(-1)^k}{k} \\ &= \sum_{k=1}^{l+1} \binom{l+1}{k} \frac{(-1)^k}{l+1} \\ &= \frac{1}{l+1} \left(\sum_{k=0}^{l+1} \binom{l+1}{k} (-1)^k - 1 \right) \\ &= \frac{-1}{l+1} , \end{aligned} \quad (3.52)$$

en utilisant la formule du binôme.

On conclut à l'aide de l'identité télescopique

$$S_n = \sum_{l=1}^{n-1} (S_{l+1} - S_l) + S_1 = - \sum_{l=1}^{n-1} \frac{1}{l+1} - 1 \quad (3.53)$$

□

Donc

$$\frac{-1}{\log(1-p)} \sum_{k=1}^n \frac{1}{k} \leq \mathbb{E}[T_n] \leq 1 + \frac{-1}{\log(1-p)} \sum_{k=1}^n \frac{1}{k} . \quad (3.54)$$

Et finalement, en utilisant l'approximation somme intégrale $\log n \sim \sum_{k=1}^n \frac{1}{k}$:

$$\mathbb{E}[T_n] \sim - \frac{\log n}{\log(1-p)} = t_n . \quad (3.55)$$

De même, on peut estimer que

$$\mathbb{E} [T_n^2] \sim (t_n)^2. \quad (3.56)$$

En effet T_n est à valeurs positives donc

$$\mathbb{E} [T_n^2] = \sum_{t \geq 0} \mathbb{P} (T_n^2 > t) = \sum_{t \geq 0} \mathbb{P} (T_n > \sqrt{t}) = \sum_{t \geq 0} \left[1 - \left(1 - (1-p)^{\sqrt{t}} \right)^n \right], \quad (3.57)$$

et le même calcul que (3.47) donne

$$\begin{aligned} \int_{t \geq 0} \left[1 - \left(1 - (1-p)^{\sqrt{t}} \right)^n \right] dt &= - \sum_{k=1}^n \binom{n}{k} (-1)^k \int_{t \geq 0} (1-p)^{k\sqrt{t}} dt \\ &= \frac{-2}{(\log(1-p))^2} \sum_{k=1}^n \binom{n}{k} \frac{(-1)^k}{k^2}. \end{aligned} \quad (3.58)$$

Or

Lemme 3.4. *On a*

$$\sum_{k=1}^n \binom{n}{k} \frac{(-1)^k}{k^2} = - \sum_{l=1}^n \frac{1}{l} \sum_{k=1}^l \frac{1}{k} \sim -\frac{1}{2} \log^2 n \quad (3.59)$$

Démonstration :

Soit

$$R_n = \sum_{k=1}^n \binom{n}{k} \frac{(-1)^k}{k^2} \quad (3.60)$$

On a

$$\begin{aligned} R_{l+1} - R_l &= \frac{(-1)^{l+1}}{(l+1)^2} + \sum_{k=1}^l \left(\binom{l+1}{k} - \binom{l}{k} \right) \frac{(-1)^k}{k^2} \\ &= \frac{(-1)^{l+1}}{(l+1)^2} + \sum_{k=1}^l \frac{k}{l+1} \binom{l+1}{k} \frac{(-1)^k}{k^2} \\ &= \frac{1}{l+1} + \sum_{k=1}^{l+1} \binom{l+1}{k} \frac{(-1)^k}{k}, \end{aligned} \quad (3.61)$$

d'où

$$R_{l+1} - R_l = -\frac{1}{l+1} \sum_{k=1}^{l+1} \frac{1}{k}, \quad (3.62)$$

d'après le lemme précédent.

Et on obtient la première égalité en sommant comme dans le lemme précédent à l'aide de l'identité télescopique

$$\begin{aligned}
R_n &= \sum_{l=1}^{n-1} (R_{l+1} - R_l) + R_1 \\
&= - \sum_{l=1}^{n-1} \frac{1}{l+1} \sum_{k=1}^{l+1} \frac{1}{k} - 1 \\
&= - \sum_{l=1}^n \frac{1}{l} \sum_{k=1}^l \frac{1}{k} .
\end{aligned} \tag{3.63}$$

Et pour obtenir la deuxième partie du lemme, on procède par comparaison somme intégrale :

pour toute fonction décroissante positive f on a

$$\int_{t=1}^m f(t) dt \leq \sum_{k=1}^m f(k) \leq 1 + \int_{t=1}^m f(t) dt . \tag{3.64}$$

Donc

$$\log l = \int_{t=1}^l \frac{1}{t} dt \leq \sum_{k=1}^l \frac{1}{k} \leq 1 + \log l , \tag{3.65}$$

d'où

$$\sum_{l=1}^n \frac{1}{l} \sum_{k=1}^l \frac{1}{k} \geq \sum_{l=1}^n \frac{\log l}{l} \geq \int_{t=1}^n \frac{\log t}{t} dt = \frac{1}{2} \log^2 n , \tag{3.66}$$

et

$$\begin{aligned}
\sum_{l=1}^n \frac{1}{l} \sum_{k=1}^l \frac{1}{k} &\leq \sum_{l=1}^n \frac{1}{l} (1 + \log l) = \sum_{l=1}^n \frac{1}{l} + \sum_{l=1}^n \frac{\log l}{l} \\
&\leq 1 + \int_{t=1}^n \frac{1}{t} dt + 1 + \int_{t=1}^n \frac{\log t}{t} dt \\
&= 2 + \log n + \frac{1}{2} \log^2 n .
\end{aligned} \tag{3.67}$$

Et finalement

$$\sum_{l=1}^n \frac{1}{l} \sum_{k=1}^l \frac{1}{k} \sim \frac{1}{2} \log^2 n . \tag{3.68}$$

□

On a donc

$$\mathbb{E} [T_n^2] \sim \frac{\log^2 n}{(\log(1-p))^2} = t_n^2 , \tag{3.69}$$

et la proposition 3.2 s'applique :

$$\lim_{n \rightarrow \infty} \text{Var}\left(\frac{T_n}{\mathbb{E}[T_n]}\right) = \lim_{n \rightarrow \infty} \frac{\mathbb{E}[T_n^2]}{\mathbb{E}[T_n]^2} - 1 = \lim_{n \rightarrow \infty} \frac{t_n^2}{t_n^2} - 1 = 0. \quad (3.70)$$

Donc T_n admet une convergence abrupte au temps $\mathbb{E}[T_n]$ qui est équivalent au temps de cutoff $t_n = -\frac{\log n}{\log(1-p)}$.

Chapitre 4

Métastabilité

A la différence de la convergence abrupte, la métastabilité est un phénomène très aléatoire. Elle est caractérisée par le fait que le système va rester un temps exponentiellement long dans un état métastable, avant de basculer vers l'équilibre. Elle est liée à l'existence de minima locaux de l'énergie potentielle du système, et on va voir que le comportement métastable résulte de la difficulté à sortir du puits d'attraction de ces minima.

Après avoir présenté le phénomène physique dans la section 4.1, on introduit un modèle de mécanique statistique pour décrire les systèmes ferromagnétiques dans un régime métastable (sections 4.2 et 4.3). C'est dans ce cadre qu'a été formalisée l'approche "pathwise" [12], qui consiste à regarder la loi des trajectoires typiques entre l'état métastable et l'état stable du système (section 4.4).

Ensuite, dans la section 4.5 on formalise le comportement métastable (définition 4.1), dans l'esprit de l'approche "pathwise". Et on montre que l'existence de deux échelles de temps (le temps moyen de tombée est très inférieur à celui de sortie), est une condition suffisante pour qu'une chaîne de Markov observe un tel comportement (théorème 4.2).

Enfin, dans la dernière section de ce chapitre (section 4.6), on présente brièvement les autres approches utilisées pour formaliser la métastabilité.

4.1 Un peu de physique

Certains systèmes physico-chimiques ont la propriété de rester longtemps dans un état différent de l'équilibre thermodynamique vers lequel ils sont censés évoluer. C'est le cas par exemple du diamant dont l'état stable est le carbone graphite, mais qui à l'oeil humain semble stable. On parle alors de pseudo-équilibre ou encore d'état métastable.

La métastabilité est la capacité de cet état de non-équilibre à persister un certain temps. Un état métastable est stable cinétiquement, mais pas thermodynamiquement, et si on caractérise le système par son énergie potentielle, un état métastable correspond à un pseudo-minimum (minimum local) de l'énergie alors que l'état stable du système correspond au vrai minimum (minimum global) de cette énergie potentielle.

Typiquement, un système dans un état métastable va y rester un temps exponentiellement long, avant d'atteindre le vrai équilibre.

En thermodynamique, on observe la métastabilité pour des systèmes proches d'une transition de phase de premier ordre, comme les transitions liquide/solide ou liquide/gaz. Ces transitions dépendent de paramètres thermodynamiques comme la température ou la pression, et on observe un état métastable lorsqu'au lieu de changer de phase quand le paramètre varie, le système reste dans son état précédent. Les exemples d'états métastables sont les solutions et vapeurs sursaturées, et les liquides et vapeurs "supercooled" tel l'eau en surfusion (brouillards givrants) : des gouttelettes d'eau liquide en suspension dans l'air à température négative.

Afin de bien comprendre le phénomène, regardons un exemple de métastabilité que l'on peut observer expérimentalement au niveau de la transition liquide/gaz.

Cet exemple est discuté dans le livre de Enzo Olivieri et Maria Eulalia Vares "Large deviations and metastability" [45], dans une très belle introduction sur le phénomène physique de la métastabilité (chap. 4).

On considère la compression isotherme d'un gaz à température T inférieure à la température critique T_c . On regarde le graphe de la pression P par rapport au volume v (cf fig).



Il y a deux volumes critiques : v_l en dessous duquel l'état stable est la phase purement liquide, et v_g dit point de condensation, en dessus duquel l'état stable est la phase purement gazeuse.

Entre les deux ($v_l < v < v_g$), les deux phases coexistent. Maintenant, si on commence l'expérience avec un volume de gaz $v > v_g$, et que l'on compresse suffisamment lentement le gaz, on peut obtenir une phase purement gazeuse avec $v < v_g$, que l'on appelle vapeur sursaturée. On peut continuer jusqu'à un volume v^* , à partir duquel le gaz devient instable, une phase liquide se forme et coexiste avec la phase gazeuse. En apparence les propriétés de l'état métastable ($v^* < v < v_g$) sont similaires à celles du gaz dans son état normal ($v > v_g$), mais une petite perturbation (impureté, fluctuation thermique spontanée ou perturbation extérieure) peut entraîner la nucléation de la phase liquide : un petit noyau de condensation se crée localement, et il grandit rapidement en absorbant les molécules de vapeur supersaturée autour de lui jusqu'à atteindre l'équilibre où les deux phases coexistent dans certaines proportions.

C'est le même phénomène qui se produit avec les brouillards givrants : les gouttelettes d'eau liquide en suspension dans l'air restent à l'état liquide tant que la température est supérieure à -39°C (qui correspond au volume v^* dans l'exemple précédent), mais elles gèlent brusquement dès qu'elles entrent en contact avec un corps étranger (le sol, un cristal de glace...).

Les systèmes ferromagnétiques fournissent un autre exemple de métastabilité. On regarde la magnétisation m du système en fonction du champ magnétique extérieur h , pour un système en dessous de la température critique T_c dite point de Curie, au-dessus de laquelle la magnétisation n'a plus lieu. L'état d'équilibre du système correspond à une magnétisation de même signe que le champ externe. Pourtant, autour de $h = 0$, on observe un "hysteresis loop" (cf fig) : si on part de l'état d'équilibre avec un champ externe négatif mais proche de zéro, donc avec $m(h) < 0$, et qu'on fait varier le champ externe jusqu'à une petite valeur positive, la magnétisation du système reste négative. C'est donc un état métastable, puisque le vrai équilibre serait une magnétisation positive. On peut continuer jusqu'à atteindre une valeur h^* dit champ coercitif, à laquelle la magnétisation redevient positive. On a le même phénomène en parcourant la boucle dans l'autre sens.

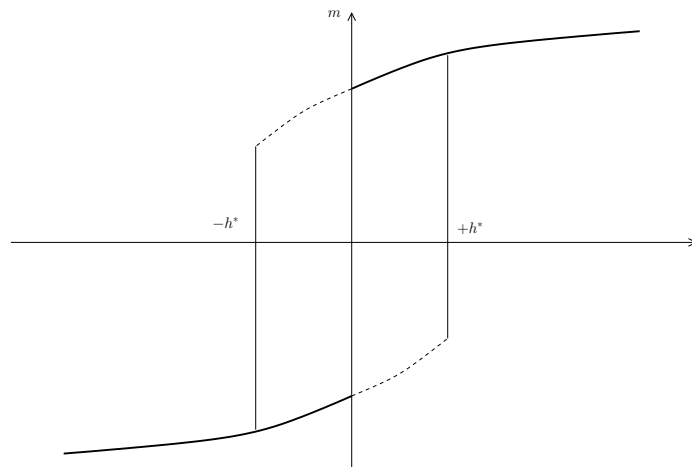


FIG. 4.1 – Hysteresis loop

Toujours en suivant Olivieri et Vares ([45], chap 4), on va maintenant introduire le modèle de Curie-Weiss, pour décrire le comportement métastable des systèmes ferromagnétiques entre $-h^*$ et h^* . Cela va nous amener à une description mathématique de la métastabilité suivant une approche dite "pathwise approach", qui va nous permettre ensuite de faire le lien avec le phénomène de cutoff.

4.2 Théorie de Curie-Weiss

On se place dans le cadre de la mécanique statistique, et on considère un système de N spins qui peuvent prendre la valeur $+1$ ou -1 . Une configuration du système est donc la donnée de $\sigma = (\sigma_1, \dots, \sigma_N) \in \{-1, +1\}^N$. On considère que les interactions sont de type “champ moyen”, c’est-à-dire qu’il n’y a pas de structure spatiale, et que les spins interagissent entre eux indépendamment de la distance qui les sépare.

Ainsi l’énergie associée à une configuration dépend de ces interactions et du champ magnétique externe h :

$$H_N(\sigma) = -\frac{J}{2N} \sum_{i,j=1}^N \sigma_i \sigma_j - h \sum_{i=1}^N \sigma_i , \quad (4.1)$$

où J est un paramètre fixé du système.

Cette énergie ne dépend de la configuration σ qu’au travers de la magnétisation $m_N(\sigma) = N^{-1} \sum_{i=1}^N \sigma_i$ qui lui est associée :

$$H_N(\sigma) = -N \left(\frac{J}{2} (m_N(\sigma))^2 + h m_N(\sigma) \right) . \quad (4.2)$$

Cette magnétisation évolue dans

$$\mathcal{Y}_N = \{-1, -1 + 2/N, \dots, 1 - 2/N, 1\} , \quad (4.3)$$

et on lui associe l’énergie libre grand canonique

$$F_N(\beta, h) = -\frac{1}{\beta N} \log Z_N(\beta, h) , \quad (4.4)$$

où $\beta = 1/kT$, avec T la température et k la constante de Boltzmann, et

$$Z_N(\beta, h) = \sum_{\sigma} e^{-\beta H_N(\sigma)} . \quad (4.5)$$

On est intéressé par la limite de cette énergie lorsque N tend vers l’infini. En appliquant la théorie des grandes déviations, on trouve que

$$\lim_N F_N(\beta, h) = \min_{m \in [-1, +1]} f_{\beta, h}(m) \quad (4.6)$$

où $f_{\beta, h}$ est l’énergie libre canonique du système ou encore énergie libre de Gibbs.

En fonction des paramètres β et J , on peut déterminer le profil de cette fonction.

Si $\beta J \leq 1$, elle possède un seul minimum $m(\beta, h)$ de même signe que h , cela correspond au cas de température surcritique $T > T_c$.

Si $\beta J > 1$, il existe une valeur critique h_0 pour le champ magnétique (dite champ coercitif) au delà de laquelle ($|h| > h_0$) il n’y a à nouveau qu’un seul minimum $m(\beta, h)$ de même signe que h .

Et lorsque le champ est inférieur au champ coercitif ($|h| < h_0$), l'énergie libre canonique possède deux minima qui constituent deux puits d'attraction dont la profondeur dépend de h :

- si $h = 0$, alors le profil est symétrique et le système a donc deux vrais états d'équilibre.
- et si $0 < |h| < h_0$, on est toujours en présence de deux puits, mais de profondeurs différentes. Il y a un minimum local qui correspond à un état métastable du système, et un minimum global de même signe que h , qui correspond au vrai équilibre du système (voir Fig. 4.2).

C'est ce dernier cas, qui correspond à la boucle d'hysteresis, qui va nous intéresser dans le cadre de notre approche “pathwise” de la métastabilité.

4.3 Modèle de Curie-Weiss

On considère la distribution $\nu_{N,\beta,h}$ de $m_N(\sigma)$ sous la mesure de Gibbs

$$\mu_{N,\beta,h}(\sigma) = \frac{1}{Z_N(\beta, h)} e^{-\beta H_N(\sigma)} . \quad (4.7)$$

On a

$$\nu_{N,\beta,h}(m) = \frac{1}{Z_N(\beta, h)} e^{-N\beta f_{N,\beta,h}(m)} , \quad (4.8)$$

où $f_{N,\beta,h}(m)$ tend vers l'énergie libre de Gibbs $f_{\beta,h}$ définie plus haut, lorsque N tend vers l'infini.

On se place dans le cas où $J = 1$, $\beta > 1$ et $0 < h < h_0$, c'est-à-dire dans le cas où le profil de $f_{\beta,h}$ possède deux puits avec un minimum local en $m^-(\beta, h) < 0$, et un minimum global en $m^+(\beta, h) > 0$. Entre les deux, il y a un maximum local $m^0(\beta, h)$ qui délimite les deux puits (voir Fig. 4.2).

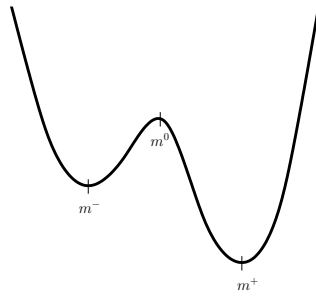


FIG. 4.2 – Modèle de Curie-Weiss, profil énergétique à deux puits.

Dans la suite, on omet les indices β et h pour alléger les notations.

Le modèle de Curie-Weiss consiste à prendre une chaîne de Markov ξ_N qui admet ν_N comme mesure invariante, et dont les transitions $Q_N(m, \tilde{m})$ sont nulles si $|m - \tilde{m}| > 2/N$ (c'est-à-dire que l'on autorise un seul changement de spin à chaque instant). Une telle chaîne est une chaîne de naissance et de mort et elle réversible par rapport à sa mesure invariante ν_N

$$\nu_N(m) Q_N(m, \tilde{m}) = \nu_N(\tilde{m}) Q_N(\tilde{m}, m), \quad \text{pour } m \text{ et } \tilde{m} \text{ tels que } |m - \tilde{m}| > 2/N. \quad (4.9)$$

Différents exemples de telles chaînes sont discutés dans “Large deviation and Metastability” [45].

4.4 Approche “pathwise”

Cette approche a été utilisée pour la première fois en 1984 dans l'article de M. Cassandro, A. Galves, E. Olivieri et M. E. Vares “Metastable behavior of stochastic dynamics : a pathwise approach” [12], pour étudier le modèle de Curie-Weiss.

Elle consiste à regarder les trajectoires typiques de la métastabilité et leurs statistiques. On s'intéresse tout particulièrement au temps moyen pour passer de l'état métastable m^- à l'état d'équilibre du système m^+ .

De manière heuristique, on peut interpréter ce processus comme modélisant une particule qui se déplace dans le profil d'énergie défini par f . Si la particule se trouve dans un puits, elle sera attirée par le minimum local qui s'y trouve. Cela se comprend intuitivement, et on verra dans le chapitre suivant que cela correspond au phénomène de convergence abrupte. Ainsi, ce qui fait la métastabilité dans la trajectoire qui va de m^- à m^+ , c'est la trajectoire de m^- à m^0 : c'est la sortie du puits de m^- , qui se fait dans le sens inverse de la dérive qui pousse la particule vers le fond.

Remarque : comme m^- , m^0 et m^+ n'appartiennent pas forcément à l'espace \mathcal{Y}_N , on définit les entiers i^- , i^0 et i^+ tels que

$$m_j := -1 + \frac{2i_j}{N} \leq m^j < -1 + \frac{2(i_j + 1)}{N}, \quad (4.10)$$

pour $j = -, +, 0$ (bien que ces entiers dépendent de N , on omet l'indice N).

On définit également l'entier i_* tel que

$$m_* := -1 + \frac{2i_*}{N} \leq m^0 + \frac{1}{N^{1/4}} < -1 + \frac{2(i_* + 1)}{N}, \quad (4.11)$$

qui correspond au point de \mathcal{Y}_N situé immédiatement à droite du sommet m^0 qui sépare les deux puits (voir Fig. 4.2).

On s'intéresse donc maintenant à la trajectoire de sortie de m_- à m_* .

On définit $T_N^{i,j}$, le temps d'atteinte de $-1 + 2j/N$ partant de $-1 + 2i/N$

$$T_N^{i,j} = \inf\{t \geq 0 : \xi_N(t) = -1 + 2j/N \mid \xi_N(0) = -1 + 2i/N\}, \quad (4.12)$$

et T_N le temps de sortie (encore appelé “tunnelling time”)

$$T_N = T_N^{i_-, i_*} . \quad (4.13)$$

Dans “Large deviation and metastability” ([45], Th 4.6), Olivieri et Vares, en reprenant les résultats de [12], montrent qu’asymptotiquement en N , le temps de sortie T_N est exponentiellement distribué autour de sa moyenne

$$\frac{T_N}{\mathbb{E}[T_N]} \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \exp(1) . \quad (4.14)$$

Ils donnent également un ordre de grandeur du temps moyen de sortie ($\mathbb{E}[T_N] \sim e^{N\Delta}$), où $\Delta = \beta(f(m^0) - f(m^-))$.

C’est ce qui pour nous caractérise le comportement métastable du processus. Le système reste un temps exponentiellement long (et donc imprévisible) dans l’état métastable.

Pour être plus précis, avant le temps T_N , le système se concentre autour de m_- ([45], Th 4.10), c’est la thermalisation. Une fois en m_* , avec probabilité 1, il atteint m_+ dans un temps beaucoup plus court que $e^{N\Delta}$ et avant de revenir en m_0 ([45], Corollaire 4.15).

Ces résultats sont synthétisés dans le théorème 2.1 de [12] de la manière suivante : si on regarde le processus empirique

$$A_{N, R_N}(s, x) = \frac{1}{R_N} \sum_{k=s\mathbb{E}[T_N]}^{s\mathbb{E}[T_N] + R_N} \delta_{\xi_N(k)}(x) , \quad (4.15)$$

qui compte le nombre moyen de passages en x autour du temps moyen $\mathbb{E}[T_N]$ du processus initialement en m_- . Alors, en prenant R_N tel que $R_N/\mathbb{E}[T_N] \rightarrow 0$ (par exemple $R_N = e^{\alpha N}$, avec $0 < \alpha < \Delta$), A_{N, R_N} converge en loi vers le processus A

$$A(s, \cdot) = \begin{cases} \delta_{m_-} & \text{si } s < T \\ \delta_{m_+} & \text{si } s \geq T , \end{cases} \quad (4.16)$$

où T est une variable de loi exponentielle de paramètre 1.

4.5 Métastabilité pour les chaînes de Markov

On a vu ci-dessus comment la métastabilité était liée au profil énergétique associé à la mesure invariante du processus : un état métastable correspond à un minimum local de cette énergie, et ce qui fait le comportement métastable du processus est l’imprédictibilité du temps de sortie du puits où se trouve ce minimum (4.14).

Cela nous amène à formaliser dans l’esprit de l’approche “pathwise”, et de manière générale le comportement métastable pour une suite T_N de variables aléatoires sur un espace dénombrable \mathcal{A}_N :

Définition 4.1. Une variable aléatoire T_N sur \mathcal{A}_N possède un comportement métastable si

$$\lim_{N \rightarrow \infty} \frac{T_N}{\mathbb{E}[T_N]} \stackrel{\mathcal{L}}{=} \exp(1) . \quad (4.17)$$

Cette définition est en fait plus large que le phénomène physique de la métastabilité. En effet elle décrit bien le comportement d'un système dans sa transition d'un état métastable vers le vrai équilibre. Mais ce problème est en fait celui de la sortie d'un puits du profil énergétique. Que celui-ci soit un minimum local ou global ne change en rien le comportement du processus : c'est le même phénomène que l'on observe !

On a vu également que le comportement métastable était liée à l'existence de deux échelles de temps différentes : le temps pour tomber dans le puits est très inférieur au temps nécessaire pour en sortir. C'est en fait une condition suffisante pour observer un comportement métastable. C'est ce qu'on va montrer dans le théorème 4.2, qui est une généralisation aux chaînes de Markov de la proposition 2.1 de [12], qui établit la limite exponentielle pour le modèle de Curie-Weiss.

On formalise l'existence de ces deux échelles de temps de la façon suivante : soit $X_N(t)$ une chaîne de Markov irréductible sur l'espace \mathcal{A}_N , et soient l_N et h_N deux états de \mathcal{A}_N , (appelés respectivement basse et haute énergie), le temps moyen pour aller de h_N à l_N est très inférieur à celui mis pour aller de l_N à h_N , et on note

$$\mathbb{E}[T_{h_N \rightarrow l_N}] \ll \mathbb{E}[T_{l_N \rightarrow h_N}] , \quad (4.18)$$

si les conditions suivantes sont satisfaites

1. il existe α_N tel que :

$$\lim_{N \rightarrow \infty} \frac{\mathbb{E}[T_{h_N \rightarrow l_N}]}{\alpha_N} = 0 , \quad \text{et} \quad \lim_{N \rightarrow \infty} \frac{\alpha_N}{\mathbb{E}[T_{l_N \rightarrow h_N}]} = 0 , \quad (4.19)$$

où $T_{x \rightarrow y}$ est le temps d'atteinte de y pour la chaîne partant de x à l'instant zéro

$$T_{x \rightarrow y} = \inf\{t > 0 : X_N^x(t) = y\} . \quad (4.20)$$

2. h_N maximise l'énergie, dans le sens où il maximise le temps moyen pour atteindre l_N :

$$\sup_{x_N \in \mathcal{A}_N} \mathbb{E}[T_{x_N \rightarrow l_N}] = \mathbb{E}[T_{h_N \rightarrow l_N}] . \quad (4.21)$$

3. la suite

$$S_N = \frac{T_{l_N \rightarrow h_N}}{\mathbb{E}[T_{l_N \rightarrow h_N}]} , \quad (4.22)$$

est uniformément intégrable (voir [54] chap. 13).

C'est à dire que pour tout $\varepsilon > 0$, il existe $K > 0$ tel que

$$\mathbb{E}[|S_N|; |S_N| > K] < \varepsilon , \quad \forall N . \quad (4.23)$$

Remarques :

- la première condition formalise l’existence des deux échelles de temps.
- la seconde hypothèse est nécessaire pour des raisons techniques, mais ne réduit pas la portée du résultat. Cela suppose seulement que l’état h_N est en haut du puits d’attraction de l_N . Si ça n’est pas le cas (par exemple si h_N se situe à mi-pente d’un puits), il suffit juste de considérer la restriction \tilde{A}_N de l’espace \mathcal{A}_N , de sorte que h_N maximise l’énergie sur \tilde{A}_N , ce qui ne change pas le comportement général et permet de montrer le comportement métastable de $T_{l_N \rightarrow h_N}$.
- la troisième condition va nous permettre d’appliquer le théorème 5.4 de “Convergence of Probability Measures” [7] :
pour toute suite X_N uniformément intégrable qui converge en loi vers X ,
on a $\mathbb{E}[X_N] \xrightarrow[N \rightarrow \infty]{} \mathbb{E}[X]$.

On montre ainsi à la fin de la démonstration du théorème que la limite S est unique.

Théorème 4.2. *Soit $N \in \mathbb{N}$, soit $X_N(t)$ une chaîne de Markov irréductible et récurrente positive sur un espace dénombrable \mathcal{A}_N .*

S’il existe deux états l_N et h_N tels que $\mathbb{E}[T_{h_N \rightarrow l_N}] \ll \mathbb{E}[T_{l_N \rightarrow h_N}]$, alors le temps $T_{l_N \rightarrow h_N}$ a un comportement métastable.

Démonstration du théorème 4.2 :

On va donc montrer que S_N converge en loi vers variable exponentielle de paramètre un ($\exp(1)$).

On va le faire en deux temps. Tout d’abord, on remarque que la suite S_N est équitendue. En effet, par définition de S_N

$$\mathbb{E}[S_N] = 1, \quad (4.24)$$

et donc par l’inégalité de Markov

$$\mathbb{P}(S_N > t) \leq \frac{\mathbb{E}[S_N]}{t} = \frac{1}{t} \xrightarrow[t \rightarrow \infty]{} 0. \quad (4.25)$$

En utilisant un résultat d’analyse (voir le th. 4.4.3. de [14]), on sait donc que toute sous-suite de S_N admet une sous suite qui converge en loi.

Il nous reste donc à montrer que cette limite est unique et que c’est une loi exponentielle $\exp(1)$.

Pour cela, on a besoin de définir les variables auxiliaires $T_{l_N}^s$ et $S_{l_N}^s$ et de quelques résultats techniques.

Soit

$$T_{l_N}^s = \inf\{t \geq s \mathbb{E}[T_{l_N \rightarrow h_N}] : X_N^{l_N}(t) = l_N\}, \quad (4.26)$$

le premier temps de retour en l_N après l’instant $s \mathbb{E}[T_{l_N \rightarrow h_N}]$, pour la chaîne $X_N^{l_N}$,
et

$$S_{l_N}^s = \frac{T_{l_N}^s}{\mathbb{E}[T_{l_N \rightarrow h_N}]} . \quad (4.27)$$

Ces temps sont bien définis : puisqu'on a supposé la chaîne récurrente positive, à N fixé, tous les temps moyens d'atteinte sont finis.

Par hypothèse, on suppose donc qu'il existe un α_N tel que

$$\beta_N = \frac{\alpha_N}{\mathbb{E}[T_{l_N \rightarrow h_N}]} \xrightarrow{N \rightarrow \infty} 0. \quad (4.28)$$

On a les lemmes suivants :

Lemme 4.3. *On a*

$$\lim_{N \rightarrow \infty} \mathbb{P}(S_{l_N}^s > s + \beta_N) = 0. \quad (4.29)$$

Démonstration du lemme 4.3 :

D'après la définition de $S_{l_N}^s$, on a

$$\mathbb{P}(S_{l_N}^s > s + \beta_N) = \mathbb{P}\left(X_N^{l_N}(k) \neq l_N, k = s\mathbb{E}[T_{l_N \rightarrow h_N}], \dots, s\mathbb{E}[T_{l_N \rightarrow h_N}] + \alpha_N\right). \quad (4.30)$$

Soit $s_N = s\mathbb{E}[T_{l_N \rightarrow h_N}]$, et $I_N = \{s_N, \dots, s_N + \alpha_N\}$.

En partitionnant sur toutes les valeurs possibles de $X_N^{l_N}(s_N)$, on a :

$$\begin{aligned} \mathbb{P}(S_{l_N}^s > s + \beta_N) &= \mathbb{P}\left(X_N^{l_N}(k) \neq l_N, \forall k \in I_N\right) \\ &= \sum_{x_N \in \mathcal{A}_N} \mathbb{P}\left(X_N^{l_N}(k) \neq l_N, \forall k \in I_N \mid X_N^{l_N}(s_N) = x_N\right) \mathbb{P}\left(X_N^{l_N}(s_N) = x_N\right) \\ &\leq \sup_{x_N \in \mathcal{A}_N} \mathbb{P}\left(X_N^{l_N}(k) \neq l_N, \forall k \in I_N \mid X_N^{l_N}(s_N) = x_N\right) \sum_{x_N \in \mathcal{A}_N} \mathbb{P}\left(X_N^{l_N}(s_N) = x_N\right) \\ &= \sup_{x_N \in \mathcal{A}_N} \mathbb{P}(T_{x_N \rightarrow l_N} > \alpha_N), \end{aligned} \quad (4.31)$$

On utilise maintenant l'inégalité de Markov :

$$\mathbb{P}(T_{x_N \rightarrow l_N} > \alpha_N) \leq \frac{\mathbb{E}[T_{x_N \rightarrow l_N}]}{\alpha_N}. \quad (4.32)$$

D'où :

$$\mathbb{P}(S_{l_N}^s > s + \beta_N) \leq \frac{\mathbb{E}[T_{h_N \rightarrow l_N}]}{\alpha_N} \xrightarrow{N \rightarrow \infty} 0, \quad (4.33)$$

puisque par hypothèse h_N maximise le temps d'atteinte de l_N , et que

$$\lim_{N \rightarrow \infty} \frac{\mathbb{E}[T_{h_N \rightarrow l_N}]}{\alpha_N} = 0. \quad (4.34)$$

□

Lemme 4.4. *On a*

$$\mathbb{P}(S_N > s + t, S_{l_N}^s \leq s + \beta_N) \geq \mathbb{P}(S_N > s + \beta_N, S_{l_N}^s \leq s + \beta_N) \mathbb{P}(S_N > t) . \quad (4.35)$$

Démonstration du lemme 4.4 :

En gardant les notations du lemme précédent et en partitionnant sur les valeurs possibles de $T_{l_N}^s$, on a

$$\begin{aligned} \mathbb{P}(S_N > s + t, S_{l_N}^s \leq s + \beta_N) &= \sum_{k \in I_N} \mathbb{P}(S_N > s + t, T_{l_N}^s = k) \\ &= \sum_{k \in I_N} \mathbb{P}(S_N > s + t, S_N > s, T_{l_N}^s = k) \\ &= \sum_{k \in I_N} \mathbb{P}(S_N > s + t \mid S_N > s, T_{l_N}^s = k) \mathbb{P}(S_N > s, T_{l_N}^s = k) . \end{aligned} \quad (4.36)$$

Or, par la propriété de Markov, pour tout $k \in I_N$

$$\mathbb{P}(S_N > s + t \mid S_N > s, T_{l_N}^s = k) = \mathbb{P}\left(S_N > s + t - \frac{k}{\mathbb{E}[T_{l_N \rightarrow h_N}]}\right) , \quad (4.37)$$

qu'on peut minorer par $\mathbb{P}(S_N > t)$ en prenant $k = s_N = s \mathbb{E}[T_{l_N \rightarrow h_N}]$.

D'où

$$\begin{aligned} \mathbb{P}(S_N > s + t, S_{l_N}^s \leq s + \beta_N) &\geq \mathbb{P}(S_N > t) \sum_{k \in I_N} \mathbb{P}(S_N > s, T_{l_N}^s = k) \\ &= \mathbb{P}(S_N > t) \mathbb{P}(S_N > s, S_{l_N}^s \leq s + \beta_N) \\ &\geq \mathbb{P}(S_N > t) \mathbb{P}(S_N > s + \beta_N, S_{l_N}^s \leq s + \beta_N) . \end{aligned} \quad (4.38)$$

□

Lemme 4.5. *On a*

$$\mathbb{P}(S_N > s + t, S_{l_N}^s \leq s + \beta_N) \leq \mathbb{P}(S_N > s) \mathbb{P}(S_N > t - \beta_N) . \quad (4.39)$$

Démonstration du lemme 4.5 :

On fait le même découpage que dans la démonstration du lemme précédent (4.36).

On majore (4.37) par $\mathbb{P}(S_N > t - \beta_N)$ en prenant $k = s_N + \alpha_N$.

Et finalement, on majore $\sum_{k \in I_N} \mathbb{P}(S_N > s, T_{l_N}^s = k)$ par $\mathbb{P}(S_N > s)$.

□

On sait que toute sous suite $S_{N'}$ de S_N admet une sous suite $S_{N''}$ qui converge. Soit S une limite faible de $S_{N''}$:

$$S_{N''} \xrightarrow[N \rightarrow \infty]{\mathcal{L}} S . \quad (4.40)$$

Remarque : les lemmes précédents restent vrais pour $S_{N''}$, et pour alléger les notations, on va omettre les "", et noter S_N pour $S_{N''}$.

Par définition de la convergence en loi, on a

$$\mathbb{P}(S_N > t) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathbb{P}(S > t) \quad (4.41)$$

pour tout point t de continuité de la fonction $t \rightarrow \mathbb{P}(S > t)$.

Soient donc s et t des points de continuité de cette fonction tels que $s + t$ le soit aussi.

On va montrer que

$$\mathbb{P}(S > s + t) = \mathbb{P}(S > s) \mathbb{P}(S > t) . \quad (4.42)$$

Majoration. On écrit :

$$\begin{aligned} \mathbb{P}(S_N > s + t) - \mathbb{P}(S_N > s) \mathbb{P}(S_N > t) = \\ \mathbb{P}(S_N > s + t) - \mathbb{P}(S_N > s) \mathbb{P}(S_N > t - \beta_N) \\ + \mathbb{P}(S_N > s) \left[\mathbb{P}(S_N > t - \beta_N) - \mathbb{P}(S_N > t) \right] . \end{aligned} \quad (4.43)$$

D'après le lemme 4.5, on peut majorer la première différence par

$$\mathbb{P}(S_N > s + t) - \mathbb{P}(S_N > s + t, S_{l_N}^s \leq s + \beta_N) , \quad (4.44)$$

qui est égal à

$$\mathbb{P}(S_N > s + t, S_{l_N}^s > s + \beta_N) , \quad (4.45)$$

qui tend vers zéro d'après le lemme 4.3. De plus,

$$\lim_{N \rightarrow \infty} \mathbb{P}(S_N > t - \beta_N) - \mathbb{P}(S_N > t) = 0 , \quad (4.46)$$

car $\beta_N = \frac{\alpha_N}{\mathbb{E}[T_{l_N \rightarrow h_N}]}$ tend vers zéro, et donc

$$\lim_{N \rightarrow \infty} \mathbb{P}(S_N > t - \beta_N) = \mathbb{P}(S > t) = \lim_{N \rightarrow \infty} \mathbb{P}(S_N > t) , \quad (4.47)$$

car la suite $S_N + \beta_N$ converge en loi vers S .

D'où

$$\begin{aligned} & \mathbb{P}(S > s + t) - \mathbb{P}(S > s) \mathbb{P}(S > t) \\ &= \lim_{N \rightarrow \infty} \mathbb{P}(S_N > s + t) - \mathbb{P}(S_N > s) \mathbb{P}(S_N > t) \\ &\leq \lim_{N \rightarrow \infty} \mathbb{P}(S_N > s + t, S_{l_N}^s > s + \beta_N) \\ &\quad + \lim_{N \rightarrow \infty} \mathbb{P}(S_N > s) \left[\mathbb{P}(S_N > t - \beta_N) - \mathbb{P}(S_N > t) \right] \\ &= 0 . \end{aligned} \quad (4.48)$$

Minoration. On écrit

$$\begin{aligned} & \mathbb{P}(S_N > s + t) - \mathbb{P}(S_N > s) \mathbb{P}(S_N > t) = \\ & \mathbb{P}(S_N > s + t) - \mathbb{P}(S_N > s + \beta_N, S_{l_N}^s \leq s + \beta_N) \mathbb{P}(S_N > t) \\ & + \mathbb{P}(S_N > t) \left[\mathbb{P}(S_N > s + \beta_N, S_{l_N}^s \leq s + \beta_N) - \mathbb{P}(S_N > s) \right]. \end{aligned} \quad (4.49)$$

On minore la première différence à l'aide du lemme 4.4, par

$$\mathbb{P}(S_N > s + t, S_{l_N}^s > s + \beta_N), \quad (4.50)$$

qui tend vers zéro d'après le lemme 4.3.

Pour le second terme on écrit :

$$\begin{aligned} & \mathbb{P}(S_N > s + \beta_N, S_{l_N}^s \leq s + \beta_N) - \mathbb{P}(S_N > s) \\ & = \mathbb{P}(S_N > s + \beta_N, S_{l_N}^s \leq s + \beta_N) - \mathbb{P}(S_N > s + \beta_N) \\ & + \mathbb{P}(S_N > s + \beta_N) - \mathbb{P}(S_N > s) \\ & = -\mathbb{P}(S_N > s + \beta_N, S_{l_N}^s > s + \beta_N) + \mathbb{P}(S_N > s + \beta_N) - \mathbb{P}(S_N > s). \end{aligned} \quad (4.51)$$

Et, toujours d'après le lemme 4.3 :

$$\lim_{N \rightarrow \infty} \mathbb{P}(S_N > s + \beta_N, S_{l_N}^s > s + \beta_N) = 0. \quad (4.52)$$

Et enfin

$$\lim_{N \rightarrow \infty} \mathbb{P}(S_N > s + \beta_N) - \mathbb{P}(S_N > s) = \mathbb{P}(S > s) - \mathbb{P}(S > s) = 0. \quad (4.53)$$

D'où

$$\begin{aligned} & \mathbb{P}(S > s + t) - \mathbb{P}(S > s) \mathbb{P}(S > t) \\ & \geq \lim_{N \rightarrow \infty} \mathbb{P}(S_N > s + t, S_{l_N}^s > s + \beta_N) \\ & + \lim_{N \rightarrow \infty} \mathbb{P}(S_N > t) \left[\mathbb{P}(S_N > s + \beta_N) - \mathbb{P}(S_N > s) - \mathbb{P}(S_N > s + \beta_N, S_{l_N}^s > s + \beta_N) \right] \\ & = 0. \end{aligned} \quad (4.54)$$

Conclusion :

On a donc

$$\mathbb{P}(S > s + t) = \mathbb{P}(S > s) \mathbb{P}(S > t), \quad (4.55)$$

en tout point de continuité de la fonction de répartition F de S ($F(t) = 1 - \mathbb{P}(S > t)$). De plus, comme F est continue à droite et admet au plus un nombre dénombrable de discontinuités, ces points de continuité sont denses et (4.55) est donc vraie pour tous s et t . Ce qui est caractéristique des lois exponentielles.

Pour montrer que la limite S est unique, il reste à montrer que $\mathbb{E}[S] = 1$.

Or d'après le le théorème 5.4 de [7], comme $S_{N''}$ converge en loi vers S , qu'elle est uniformément intégrable et que $\mathbb{E}[S_{N''}] = 1$, on a

$$\mathbb{E}[S] = \lim_{N \rightarrow \infty} \mathbb{E}[S_{N''}] = 1, \quad (4.56)$$

et donc S suit une loi exponentielle $\exp(1)$.

On a vu que la suite S_N est tendue, et donc que toute sous suite admet une sous suite qui converge. De plus la limite S est unique et exponentielle $\exp(1)$, d'après le théorème 2.3 de [7], la suite S_N converge en loi vers S :

$$S_N = \frac{T_{l_N \rightarrow h_N}}{\mathbb{E}[T_{l_N \rightarrow h_N}]} \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \exp(1). \quad (4.57)$$

C'est à dire que la temps de sortie de l_N vers h_N a un comportement métastable.

□

4.6 Autres approches de la métastabilité

Parmi les nombreux travaux mathématiques sur la métastabilité, une référence incontournable sont ceux de Freidlin et Wentzell ("Random perturbations of dynamical systems" [25]). Ils considèrent un système dynamique associé à un champ de vecteurs b sur \mathbb{R}^d , perturbé par un bruit blanc W_t . Typiquement $b(x) = -\nabla U(x)$, avec U un potentiel possédant deux puits d'attraction (w_1 qui contient un minimum local x_1 , et w_2 qui contient le minimum global x_2). C'est la partie déterministe du système, l'aléa étant représenté par un mouvement brownien standard d-dimensionnel W_t . Le processus de Markov $X_t^{x,\varepsilon}$ considéré est l'unique solution de l'équation différentielle stochastique

$$X_t^{x,\varepsilon} = x + \int_0^t b(X_s^{x,\varepsilon}) ds + \varepsilon W_t, \quad (4.58)$$

et il admet une mesure invariante μ_ε qui se concentre autour du puits le plus profond w_2 , et tend vers la masse de Dirac au point x_2 , lorsque ε tend vers zéro.

Par des méthodes de grandes déviations, Freidlin et Wentzell obtiennent des estimations du temps de sortie du domaine w_1 : ils montrent qu'il existe deux échelles de temps γ_ε et δ_ε telles que $\gamma_\varepsilon \rightarrow \infty$, $\delta_\varepsilon \rightarrow \infty$, et $\delta_\varepsilon/\gamma_\varepsilon \rightarrow 0$, lorsque ε tend vers zéro.

Le processus reste localisé autour de x_1 dans une échelle de temps d'ordre γ_ε , et sort du puits dans dans une échelle de temps d'ordre δ_ε . De plus on a la limite exponentielle du temps de sortie τ_ε , exprimée ici de la manière suivante :

$$\lim_{\varepsilon \rightarrow 0} \mathbb{P}(\tau_\varepsilon > t\gamma_\varepsilon) = e^{-t}, \quad (4.59)$$

avec $\gamma_\varepsilon \sim \mathbb{E}[\tau_\varepsilon]$.

Dans “Markov Chains with Exponentially Small Transitions Probabilities : First Exit Problem from a General Domain” I et II [42, 43], E. Olivieri et E. Scoppola donnent une description probabiliste complète du temps de sortie et du tube de sortie pour les modèles de Freidlin et Wentzell, en utilisant les techniques de renormalisation de “Metastability for Markov chains : a general procedure based on renormalization group ideas” [51]. Dans le cadre de l’approche “pathwise”, ils considèrent une famille de chaînes de Markov dont les probabilités de transition décroissent exponentiellement en un paramètre β qui tend vers l’infini

$$p(x, y) \sim \exp(-\beta \Delta(x, y)) . \quad (4.60)$$

Typiquement, β est l’inverse de la température, et Δ représente une différence d’énergie $\Delta(x, y) = [H(y) - H(x)]_+$, pour un profil d’énergie $H \geq 0$ quelconque, où $[a]_+ = a \vee 0$ est la partie positive du réel a .

Pour un domaine \mathcal{Q} contenant plusieurs minima locaux, ils regardent le temps de sortie du domaine (“exit time”). En se basant sur les propriétés du profil énergétique et notamment la notion de cycle (un cycle est un ensemble d’états tel que le maximum de l’énergie à l’intérieur de l’ensemble est inférieur au minimum de l’énergie sur la frontière de l’ensemble), ils montrent la limite exponentielle du temps de sortie.

De plus ils caractérisent le tube de sortie (“exit tube”). C’est à dire la suite d’états par lequel le processus passe pour sortir du domaine. Ils montrent que ce tube est la renversée temporelle d’un chemin typique pour atteindre le fond du domaine (“standart cascade”).

C’est la même idée que nous allons développer au chapitre suivant pour caractériser la trajectoire typique de sortie. La différence majeure entre ces travaux et les nôtres réside dans le fait qu’asymptotiquement en β , les transitions tendent exponentiellement vers zéro où un (4.60). Intuitivement, c’est comme si les pentes du profil d’énergie tendaient à être verticales, alors que dans les cas que nous allons considérer la pente est constante, mais c’est l’état initial (ou d’arrivée selon la trajectoire que l’on regarde) qui tend vers l’infini.

Pour un survol de ces résultats, on pourra lire “Metastability and typical exit paths in stochastic dynamics” [44], et pour une discussion plus générale des principaux aspects de la métastabilité “On the essential features of metastability : tunnelling time and critical configurations” [37].

Déterminer le tube de sortie, c’est s’intéresser aux états par lesquels le système passe dans la transition d’un état métastable vers l’état stable du système. C’est l’objet des travaux de Schonmann, Neves et Schonmann, Kotecký et Olivieri, et den Hollander, Olivieri et Scoppola [49, 50, 40, 41, 33, 15], de déterminer la forme et la taille de la “goutte critique” de spins $+1$ qu’il faut atteindre pour basculer vers l’équilibre (tous les spins $+1$), pour un modèle d’Ising à très basse température soumis à un faible champ externe positif, initialement dans l’état métastable avec tous les spins -1 . C’est le phénomène de nucléation de la phase stable.

Une autre question est de caractériser les états métastables. Pour nous ce sont les minima locaux du profil énergétique.

Dans ses travaux (“Metastability and ageing in stochastic dynamics” [9], “Metastability : A potential theoretic approach” [10], ou les lectures notes “Markov Processes and metastability” TUB 2003, et “Metastability” Prague 2006 pour un survol complet de ses résultats), A. Bovier propose une définition plus générale en termes de potentiels d’équilibre. Pour deux ensembles A et B , le potentiel $h_{A,B}$ est la probabilité d’atteindre A avant B

$$h_{A,B}(x) = \mathbb{P}_x(\tau_A < \tau_B) , \quad (4.61)$$

où τ_A et τ_B sont les temps d’atteinte de A et B respectivement.

Un processus de Markov est métastable par rapport à l’ensemble \mathcal{M} des états métastables du système si

$$\frac{\sup_{x \in \mathcal{M}} \mathbb{P}_x(\tau_{\mathcal{M} \setminus x} < \tau_x)}{\inf_{y \notin \mathcal{M}} \mathbb{P}_y(\tau_{\mathcal{M}} < \tau_y)} \leq \rho \ll 1 . \quad (4.62)$$

Ou encore en termes de temps moyens d’atteinte

$$\frac{\inf_{x \in \mathcal{M}} \mathbb{E}_x[\tau_{\mathcal{M} \setminus x}]}{\sup_{y \notin \mathcal{M}} \mathbb{E}_y[\tau_{\mathcal{M}}]} \geq 1/\rho \gg 1 . \quad (4.63)$$

Il utilise ensuite la notion de capacité de la théorie du potentiel pour exprimer ces temps moyens

$$\mathbb{E}_x[\tau_A] = \sum_{y \notin A} \frac{\mu(y) h_{x,A}(y)}{\text{cap}(A, x)} , \quad (4.64)$$

où μ est la mesure invariante du processus, et la capacité est définie de la manière suivante : pour deux ensembles A et B , c’est la probabilité partant de B d’atteindre A avant B

$$\text{cap}(A, B) = \sum_{x \in B} \mu(x) h_{A,B}(x) = \sum_{x \in B} \mu(x) \mathbb{P}_x(\tau_A < \tau_B) . \quad (4.65)$$

Et donc

$$\text{cap}(A, x) = \mu(x) \mathbb{P}_x(\tau_A < \tau_x) , \quad (4.66)$$

est la capacité de A à être atteint de x .

Remarque : la formule (4.64) est équivalente à notre formule (5.30) (voir chapitre suivant), exprimée ici en terme de capacité. Ce sont deux versions de la formule plus générale (9.1) (en appendice), qui ne nécessite pas forcément l’hypothèse de réversibilité.

A. Bovier relie ensuite la capacité avec les valeurs propres du générateur du processus, ce qui lui permet d’avoir des bornes inférieures et supérieures pour les temps de sorties des domaines métastables, et de relier ainsi de manière générale la métastabilité aux propriétés spectrales du processus.

Chapitre 5

Convergence abrupte et métastabilité pour les chaînes de naissance et de mort

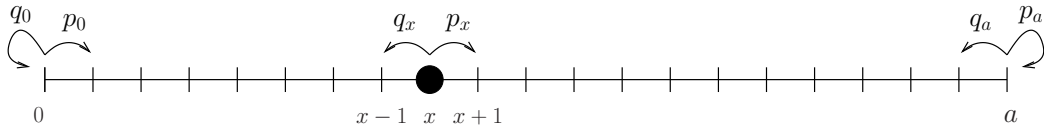
Dans ce chapitre on propose un modèle de marche aléatoire entre 0 et a , pour lequel on montre les phénomènes de convergence abrupte et de métastabilité sous une hypothèse de dérive. On suppose en effet qu'il y a une dérive vers zéro (hypothèse (\mathcal{H}) , (5.3)), dont on discute les interprétations en termes de mesure invariante et de profil énergétique dans la section 5.2.

En utilisant une approche due à Martínez et Ycart [38], on montre dans la section 5.4 la convergence abrupte vers zéro (théorème 5.3). Cette approche consiste à exprimer des quantités clés liées aux temps d'atteinte en fonction de la mesure invariante (section 5.3). Dans la section 5.5, on montre la métastabilité de la trajectoire de sortie (théorème 5.6). Pour ce faire, on utilise bien évidemment l'hypothèse (\mathcal{H}) , mais on s'appuie fortement sur la convergence abrupte au travers d'un couplage (5.54). Cela nous permet de décomposer le temps de sortie en un nombre géométrique de “tombées à l'équilibre”, typiques de la convergence abrupte (5.62).

On explore plus encore la complémentarité des deux phénomènes dans la section suivante (section 5.6), en montrant que la dernière sortie dans la trajectoire métastable est la renversée temporelle d'une trajectoire typique de cutoff (propositions 5.7, 5.8 et 5.9).

5.1 Modèle

On considère une famille $X_a(t)$ de chaînes de naissance et de mort sur $\llbracket 0, a \rrbracket$.



Lorsque a est fixé, c'est une chaîne à espace d'états fini, et donc elle est récurrente positive et admet une unique mesure invariante π_a . De plus cette mesure invariante est réversible, et on a vu au chapitre 2 qu'on pouvait donc l'exprimer en fonction des transitions de la chaîne :

$$\pi_a(x) = \prod_{i=1}^x \frac{p_{i-1}}{q_i} \pi_a(0). \quad (5.1)$$

On va de plus supposer que cette famille est uniformément irréductible, c'est-à-dire que pour tout a :

$$\inf_{x=0,\dots,a-1} p_x > 0, \text{ et } \inf_{x=1,\dots,a} q_x > 0. \quad (5.2)$$

Cette hypothèse nous permet de considérer les trajectoires de zéro vers a et de a vers zéro (définies par les temps d'atteinte $T_{0 \rightarrow a}$ et $T_{a \rightarrow 0}$) et leur comportement asymptotique en a .

Rappel : $T_{x \rightarrow y}$ est le temps d'atteinte de y partant de x : $T_{x \rightarrow y} = \inf\{t > 0 : X_a^x(t) = y\}$, où l'on note par $X_a^x(t)$ la chaîne initialisée en x .

5.2 Dérive, profil énergétique et décroissance exponentielle de la mesure invariante

On suppose qu'il existe K tel que pour tout x et pour tout a ,

$$(\mathcal{H}) \quad \frac{\pi_a(B_x)}{\pi_a(x)} < K, \quad (5.3)$$

où $B_x = \llbracket x, a \rrbracket$.

Remarque : d'après (5.1), cette quantité s'écrit

$$\frac{\pi_a(B_x)}{\pi_a(x)} = \sum_{k=x}^a \prod_{i=x+1}^k \frac{p_{i-1}}{q_i}. \quad (5.4)$$

C'est la condition (\mathcal{H}) qui donne la dérive vers zéro et on peut le voir de deux façons :

En termes de mesure invariante, cela impose que π_a possède une queue à décroissance exponentielle, ou en d'autres termes, qu'elle est concentrée vers zéro.

On parle de décroissance exponentielle puisque

$$\frac{\pi_a(B_x)}{\pi_a(x)} \leq K$$

équivalent à

$$\pi_a(B_x) \leq K \pi_a(x) = K(\pi_a(B_x) - \pi_a(B_{x+1})), \quad (5.5)$$

car $\{x\} = B_x \setminus B_{x+1}$.

On a donc pour tout x

$$\pi_a(B_{x+1}) \leq \pi_a(B_x) \left(1 - \frac{1}{K}\right). \quad (5.6)$$

Or $\pi_a(x) \leq \pi_a(B_x)$ et $\pi_a(B_0) = 1$, donc en itérant (5.6), on obtient

$$\pi_a(x) \leq \left(1 - \frac{1}{K}\right)^x = e^{-\alpha x}, \quad (5.7)$$

avec $\alpha = -\log\left(1 - \frac{1}{K}\right) > 0$.

En termes de profil énergétique, si on définit une énergie H associée à la mesure invariante de la manière suivante :

$$\pi_a(x) = \frac{1}{Z} e^{-H(x)}. \quad (5.8)$$

En comparant avec l'expression de π_a (5.1), on prend $1/Z = \pi_a(0)$ comme normalisation, ce qui nous donne

$$H(x) = -\log \left(\prod_{i=1}^x \frac{p_{i-1}}{q_i} \right). \quad (5.9)$$

Exemple : (p, q) random walk

C'est l'exemple le plus simple d'une telle chaîne de naissance et de mort. Les transitions sont indépendantes de l'état de la chaîne

$$p_x = p, \text{ et } q_x = q \quad \forall x. \quad (5.10)$$

La mesure invariante s'écrit

$$\pi_a(x) = \left(\frac{p}{q}\right)^x \pi_a(0), \quad (5.11)$$

où, puisque $\sum_{x=0}^a \pi_a(x) = 1$, le facteur de normalisation $\pi_a(0)$ est

$$\pi_a(0) = \frac{1 - (p/q)}{1 - (p/q)^{a+1}}. \quad (5.12)$$

Pour satisfaire notre hypothèse (\mathcal{H}) de dérive vers zéro, il faut évidemment supposer que $q > p$.

Comme on le voit sur la figure 5.1, on a donc bien la décroissance exponentielle de π_a . Elle est concentrée autour de zéro, et c'est pour cela qu'on parle de dérive. La loi de la chaîne tend vers π_a lorsque t tend vers l'infini, la marche aléatoire va donc se diriger vers les états les plus typiques, c'est-à-dire vers les états les plus probables sous π_a .

L'énergie associée est

$$H(x) = x \log(q/p) \quad (5.13)$$

Si on interprète la marche comme le déplacement d'une particule : le profil énergétique étant linéaire (voir Fig. 5.1), on comprend la dérive comme la pente sur laquelle la particule se déplace.

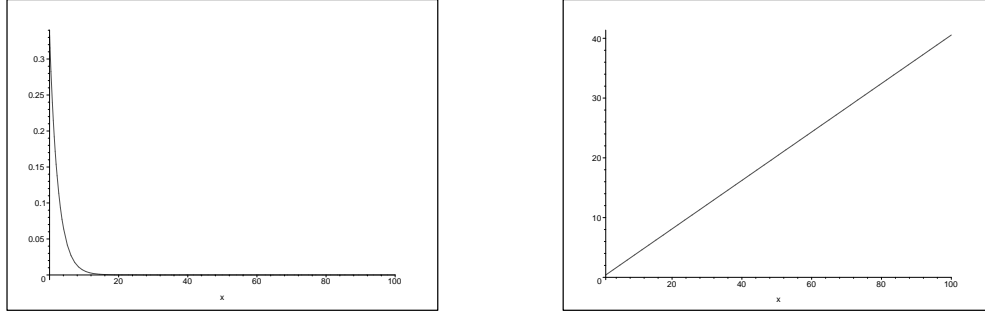


FIG. 5.1 – (p, q) random walk : mesure invariante et énergie (avec $p = 0.4$ et $a = 100$).

On a donc ici un modèle basique pour comprendre notre hypothèse (\mathcal{H}) , et la dérive vers zéro qui en découle. Elle implique la décroissance exponentielle de la mesure invariante (5.7), ou en d'autres termes sa concentration autour de zéro. Mais remarquons aussi que (5.7) nous donne une minoration du profil énergétique. En effet, pour tout modèle qui satisfait (\mathcal{H}) , on a

$$\pi_a(0) e^{-H(x)} = \pi_a(x) \leq \left(1 - \frac{1}{K}\right)^x = e^{-\alpha x}, \quad (5.14)$$

et donc

$$H(x) \geq \log \pi_a(0) + \alpha x. \quad (5.15)$$

Avant de montrer que cette dérive implique la convergence abrupte vers zéro et la métastabilité dans l'autre sens, nous avons besoin d'exprimer certaines quantités en fonction des transitions et de la mesure invariante.

5.3 Temps d'atteinte et mesure invariante

On peut calculer la probabilité partant de x d'atteindre a avant zéro $\mathbb{P}(T_{x \rightarrow a} < T_{x \rightarrow 0})$, en fonction des transitions de la chaîne.

Lemme 5.1. *On a, pour tout $x = 1, \dots, a-1$:*

$$\mathbb{P}(T_{x \rightarrow a} < T_{x \rightarrow 0}) = \frac{\sum_{k=0}^{x-1} \prod_{i=1}^k \frac{q_i}{p_i}}{\sum_{k=0}^{a-1} \prod_{i=1}^k \frac{q_i}{p_i}} \quad (5.16)$$

Démonstration

Soit $P_x = \mathbb{P}(T_{x \rightarrow a} < T_{x \rightarrow 0})$.

En partitionnant sur les positions possibles de la chaîne au premier pas, on montre que P_x satisfait à l'équation de différence

$$P_x = p_x P_{x+1} + q_x P_{x-1}, \quad (5.17)$$

pour $x = 2, \dots, a-2$.

En effet, on a :

$$\begin{aligned} P_x &= \mathbb{P}(T_{x \rightarrow a} < T_{x \rightarrow 0}, X_a^x(1) = x+1) + \mathbb{P}(T_{x \rightarrow a} < T_{x \rightarrow 0}, X_a^x(1) = x-1) \\ &= p_x \mathbb{P}(T_{x \rightarrow a} < T_{x \rightarrow 0} \mid X_a^x(1) = x+1) + q_x \mathbb{P}(T_{x \rightarrow a} < T_{x \rightarrow 0} \mid X_a^x(1) = x-1) \\ &= p_x \mathbb{P}(T_{x+1 \rightarrow a} < T_{x+1 \rightarrow 0}) + q_x \mathbb{P}(T_{x-1 \rightarrow a} < T_{x-1 \rightarrow 0}) \end{aligned} \quad (5.18)$$

Le même calcul pour $x = 1$ et $x = a-1$ donne les conditions aux bords

$$\begin{cases} P_1 = p_1 P_2 \\ P_{a-1} = p_{a-1} + q_{a-1} P_{a-2}. \end{cases} \quad (5.19)$$

En effet, comme il faut au moins $a-1$ pas pour aller de 1 à a , $T_{1 \rightarrow a} \geq a-1$.

Et si $X_a^1(1) = 0$ alors $T_{1 \rightarrow 0} = 1$, donc

$$\mathbb{P}(T_{1 \rightarrow a} < T_{1 \rightarrow 0} \mid X_a^1(1) = 0) = 0. \quad (5.20)$$

De même $T_{a-1 \rightarrow 0} \geq a-1$ et $X_a^{a-1}(1) = a$ implique $T_{a-1 \rightarrow a} = 1$, donc $\mathbb{P}(T_{a-1 \rightarrow a} < T_{a-1 \rightarrow 0} \mid X_a^{a-1}(1) = a) = 1$.

De même, en zéro et en a , on a

$$\begin{cases} P_0 = p_0 P_1 \\ P_a = p_a + q_a P_{a-1}. \end{cases} \quad (5.21)$$

Comme $p_k + q_k = 1$, on peut réécrire (5.17)

$$(p_k + q_k)P_k = p_k P_{k+1} + q_k P_{k-1}, \quad (5.22)$$

pour $k \in \{2, \dots, a-2\}$.

On obtient la relation de récurrence

$$P_{k+1} - P_k = \frac{q_k}{p_k} (P_k - P_{k-1}), \quad (5.23)$$

que l'on itère pour trouver

$$P_{k+1} - P_k = \prod_{i=2}^k \frac{q_i}{p_i} (P_2 - P_1). \quad (5.24)$$

L'identité télescopique

$$P_x = \sum_{k=1}^{x-1} (P_{k+1} - P_k) + P_1, \quad (5.25)$$

nous permet d'obtenir la solution générale de (5.17)

$$P_x = \sum_{k=1}^{x-1} \prod_{i=2}^k \frac{q_i}{p_i} (P_2 - P_1) + P_1. \quad (5.26)$$

Or $P_2 - P_1 = \frac{q_1}{p_1} P_1$, d'où

$$P_x = \sum_{k=0}^{x-1} \prod_{i=1}^k \frac{q_i}{p_i} P_1. \quad (5.27)$$

Et l'autre condition de bord $P_{a-1} = p_{a-1} + q_{a-1} P_{a-2}$, donne

$$P_1 = \frac{1}{\sum_{k=0}^{a-1} \prod_{i=1}^k \frac{q_i}{p_i}}. \quad (5.28)$$

□

A l'aide de la mesure invariante, on peut réécrire $\mathbb{P}(T_{x \rightarrow a} < T_{x \rightarrow 0})$, et on a également des expressions explicites pour le temps moyen d'atteindre $x - 1$ partant de x et son moment d'ordre 2.

Lemme 5.2. *Soit $B_x = \llbracket x, a \rrbracket$, on a*

1. $\forall x \in \llbracket 1, a-1 \rrbracket$

$$\mathbb{P}(T_{x \rightarrow a} < T_{x \rightarrow 0}) = \frac{\sum_{k=1}^x \frac{1}{q_k \pi_a(k)}}{\sum_{k=1}^a \frac{1}{q_k \pi_a(k)}} \quad (5.29)$$

2. $\forall x \in \llbracket 1, a \rrbracket$

$$\mathbb{E}[T_{x \rightarrow x+1}] = \frac{\pi_a(B_x)}{q_x \pi_a(x)} \quad (5.30)$$

3. $\forall x \in \llbracket 1, a \rrbracket$

$$\mathbb{E}[T_{x \rightarrow x+1}^2] = \frac{1}{q_x \pi_a(x)} \left(\sum_{c=x}^a \frac{2\pi_a(B_c)^2}{q_c \pi_a(c)} - \pi_a(c) \right) \quad (5.31)$$

Remarque : grâce à (5.30), on obtient une expression du temps moyen pour aller de a à zéro

$$\mathbb{E}[T_{a \rightarrow 0}] = \sum_{x=1}^a \mathbb{E}[T_{x \rightarrow x-1}] = \sum_{x=1}^a \frac{\pi_a(B_x)}{q_x \pi_a(x)}. \quad (5.32)$$

Démonstration

1. A l'aide du lemme 5.1, on peut réécrire $\mathbb{P}(T_{x \rightarrow a} < T_{x \rightarrow 0})$, en fonction de la mesure invariante π_a .

En effet

$$\begin{aligned} \pi_a(k) &= \prod_{i=1}^k \frac{p_{i-1}}{q_i} \pi_a(0) \\ &= \frac{p_0}{p_k} \prod_{i=1}^k \frac{p_i}{q_i} \pi_a(0) \end{aligned} \quad (5.33)$$

D'où

$$\prod_{i=1}^k \frac{q_i}{p_i} = \frac{p_0 \pi_a(0)}{p_k \pi_a(k)} = \frac{p_0 \pi_a(0)}{q_{k+1} \pi_a(k+1)}, \quad (5.34)$$

en utilisant la réversibilité. Et finalement, pour tout $x = 1, \dots, a-1$ on a

$$\mathbb{P}(T_{x \rightarrow a} < T_{x \rightarrow 0}) = \frac{\sum_{k=1}^x \frac{1}{q_k \pi_a(k)}}{\sum_{k=1}^a \frac{1}{q_k \pi_a(k)}}. \quad (5.35)$$

2. $t_x = \mathbb{E}[T_{x \rightarrow x-1}]$ vérifie l'équation de récurrence

$$q_x t_x = 1 + p_x t_{x+1} \quad (5.36)$$

pour $x = 1, \dots, a-1$.

En effet, en partitionnant sur la position au premier pas, on a

$$\begin{aligned} t_x &= p_x \mathbb{E}[T_{x \rightarrow x-1} \mid X(1) = x+1] + q_x \mathbb{E}[T_{x \rightarrow x-1} \mid X(1) = x-1] \\ &= p_x (\mathbb{E}[T_{x+1 \rightarrow x-1}] + 1) + q_x \\ &= p_x (t_{x+1} + t_x + 1) + q_x. \end{aligned} \quad (5.37)$$

D'où l'équation pour tout $x = 1, \dots, a-1$.

Le même calcul au point a , donne $q_a t_a = 1$, que l'on réécrit

$$t_a = \frac{\pi_a(B_a)}{\pi_a(a) q_a}, \quad (5.38)$$

car $B_a = \{a\}$.

Supposons que l'on ait

$$t_{x+1} = \frac{\pi_a(B_{x+1})}{\pi_a(x+1)q_{x+1}}, \quad (5.39)$$

on applique (5.36) et la définition de la réversibilité pour prouver la récurrence :

$$\begin{aligned} q_x t_x &= 1 + p_x t_{x+1} \\ &= 1 + \frac{p_x}{q_{x+1}} \frac{\pi_a(B_{x+1})}{\pi_a(x+1)} \\ &= 1 + \frac{\pi_a(B_{x+1})}{\pi_a(x)} \\ &= \frac{\pi_a(B_x)}{\pi_a(x)}. \end{aligned} \quad (5.40)$$

3. Soit $u_x = \mathbb{E}[T_{x \rightarrow x-1}^2]$, on obtient l'équation de récurrence en partitionnant de la même façon que pour le lemme précédent

$$\begin{aligned} u_x &= \mathbb{E}[(T_{x \rightarrow x-1} \mathbf{1}_{\{X(1)=x+1\}} + T_{x \rightarrow x-1} \mathbf{1}_{\{X(1)=x-1\}})^2] \\ &= \mathbb{E}[(T_{x \rightarrow x-1} \mathbf{1}_{\{X(1)=x+1\}})^2] + \mathbb{E}[(T_{x \rightarrow x-1} \mathbf{1}_{\{X(1)=x-1\}})^2] \\ &= p_x (\mathbb{E}[(1 + T_{x+1 \rightarrow x-1})^2]) + q_x \\ &= p_x (1 + \mathbb{E}[(T_{x+1 \rightarrow x-1})^2] + 2\mathbb{E}[T_{x+1 \rightarrow x-1}]) + q_x, \end{aligned} \quad (5.41)$$

et par la propriété de Markov forte, on a $T_{x+1 \rightarrow x-1} = T_{x+1 \rightarrow x} + T_{x \rightarrow x-1}$, donc $\mathbb{E}[T_{x+1 \rightarrow x-1}] = t_{x+1} + t_x$, ce qui donne

$$u_x = p_x (1 + (u_{x+1} + u_x + 2t_{x+1}t_x) + 2(t_{x+1} + t_x)) + q_x. \quad (5.42)$$

En utilisant l'équation de récurrence pour t_x (5.36), on a

$$u_x = 2q_x t_x^2 + p_x (u_{x+1} + u_x) - 1, \quad (5.43)$$

d'où

$$q_x u_x = 2q_x t_x^2 + p_x u_{x+1} - 1, \quad (5.44)$$

pour tout $x = 1, \dots, a-1$.

On va montrer (5.31), par récurrence.

En a , on a, $q_a u_a = 1 + 2p_a t_a$. Et $t_a = \frac{1}{q_a}$, donc

$$\begin{aligned} u_a &= \frac{1}{q_a} (1 + 2p_a t_a) \\ &= \frac{1}{q_a^2} (q_a + 2p_a) \\ &= \frac{1}{q_a^2} (2 - q_a) \end{aligned} \quad (5.45)$$

Qu'on peut réécrire

$$u_a = \frac{1}{\pi_a(a)q_a} \left(\frac{2\pi_a(a)^2}{\pi_a(a)q_a} - \pi_a(a) \right). \quad (5.46)$$

Supposons que l'on ait

$$u_{x+1} = \frac{1}{\pi_a(x+1)q_{x+1}} \left(\sum_{c=x+1}^a \frac{2\pi_a(B_c)^2}{\pi_a(c)q_c} - \pi_a(c) \right), \quad (5.47)$$

d'après (5.44), on a

$$\begin{aligned} q_x u_x &= 2q_x t_x^2 + p_x u_{x+1} - 1 \\ &= 2q_x \left(\frac{\pi_a(B_x)}{\pi_a(x)q_x} \right)^2 - 1 + \frac{p_x}{\pi_a(x+1)q_{x+1}} \left(\sum_{c=x+1}^a \frac{2\pi_a(B_c)^2}{\pi_a(c)q_c} - \pi_a(c) \right) \\ &= \frac{1}{\pi_a(x)} \left(\frac{2\pi_a(B_x)^2}{\pi_a(x)} - \pi_a(x) \right) + \frac{1}{\pi_a(x)} \left(\sum_{c=x+1}^a \frac{2\pi_a(B_c)^2}{\pi_a(c)q_c} - \pi_a(c) \right) \\ &= \frac{1}{\pi_a(x)} \left(\sum_{c=x}^a \frac{2\pi_a(B_c)^2}{\pi_a(c)q_c} - \pi_a(c) \right). \end{aligned} \quad (5.48)$$

D'où la formule (5.31).

□

Remarques :

- la démonstration des formules (5.30) et (5.31) est adaptée de l'article de Martínez et Ycart [38] où ils donnent des expressions analogues pour les processus de naissance et de mort sur les arbres (lemme 5.1).
- ces formules ((5.30) et (5.31)) permettent d'obtenir des expressions pour les moments d'ordre 1 et 2 des temps d'atteinte $T_{j \rightarrow n}$ pour $j < n$, en fonction de la mesure invariante π_a .
Elles sont équivalentes aux formules (9.12) et (9.21) de la première section de l'appendice obtenues par la méthode des équations de différence (voir Feller "An Introduction to Probability Theory and Its Applications, vol.1" [24]), qui expriment ces quantités en fonction des transitions.
- par la même méthode, on a l'analogue de (5.30) pour $\mathbb{E}[T_{x \rightarrow x+1}]$:

$$\mathbb{E}[T_{x \rightarrow x+1}] = \frac{\pi_a(A_x)}{p_x \pi_a(x)}, \quad (5.49)$$

où $A_x = [0, x]$.

5.4 Convergence abrupte

Sous l'hypothèse (\mathcal{H}) de dérive vers zéro, on a vu que la mesure d'équilibre π_a charge les points autour de zéro : les états typiques de π_a sont concentrés autour de zéro (ce sont ceux qui ont la plus forte probabilité sous cette mesure). Par analogie avec la terminologie physicienne, on appelle donc 0 l'état d'équilibre du système.

On s'intéresse ici à la trajectoire de la particule partant initialement de a . A cause de la dérive, elle va atteindre l'équilibre rapidement et de manière presque déterministe, c'est le phénomène de convergence abrupte. On appelle cette trajectoire typique la "tombée à l'équilibre". Pour montrer qu'elle admet une convergence abrupte, nous allons vérifier que le temps d'atteinte $T_{a \rightarrow 0}$ vérifie les conditions de la proposition 3.2.

Théorème 5.3. *Sous l'hypothèse (\mathcal{H}) , le temps de tombée à l'équilibre $T_{a \rightarrow 0}$ converge abruptement au temps $\mathbb{E}[T_{a \rightarrow 0}]$,*

Démonstration

Nous devons donc montrer que

$$\lim_{a \rightarrow \infty} \mathbb{E}[T_{a \rightarrow 0}] = \infty \text{ et } \lim_{a \rightarrow \infty} \text{Var}\left(\frac{T_{a \rightarrow 0}}{\mathbb{E}[T_{a \rightarrow 0}]}\right) = 0. \quad (5.50)$$

La première condition est satisfaite de manière évidente. Puisque il faut au moins a pas pour aller de a à zéro, on a $T_{a \rightarrow 0} \geq a$, et donc $\mathbb{E}[T_{a \rightarrow 0}] \geq a$ et tend vers l'infini lorsque a tend vers l'infini.

Montrons maintenant que $\lim_{a \rightarrow \infty} \text{Var}\left(\frac{T_{a \rightarrow 0}}{\mathbb{E}[T_{a \rightarrow 0}]}\right) = 0$.

Soit K_q la borne supérieure de $(1/q_x)$, d'après le lemme précédent (formules (5.30) et (5.31)), on a

$$\begin{aligned} \mathbb{E}[T_{a \rightarrow 0}^2] &= \sum_{x=1}^a u_x \\ &= \sum_{x=1}^a \frac{1}{\pi_a(x) q_x} \left(\sum_{c=x}^a \frac{2 \pi_a(B_c)^2}{\pi_a(c) q_c} - \pi_a(c) \right) \\ &\leq \sum_{x=1}^a \frac{2}{\pi_a(x) q_x} \left(\sum_{c=x}^a \frac{K^2 \pi_a(c)}{q_c} \right) \\ &\leq \sum_{x=1}^a \frac{2 K^2 K_q}{\pi_a(x) q_x} \sum_{c=x}^a \pi_a(c) \\ &= 2 K^2 K_q \sum_{x=1}^a \frac{1}{\pi_a(x) q_x} \pi_a(B_x) \\ &= 2 K^2 K_q \sum_{x=1}^a t_x = 2 K^2 K_q \mathbb{E}[T_{a \rightarrow 0}] \end{aligned} \quad (5.51)$$

D'où

$$\text{Var}\left(\frac{T_{a \rightarrow 0}}{\mathbb{E}[T_{a \rightarrow 0}]}\right) \leq \frac{\mathbb{E}[T_{a \rightarrow 0}^2]}{(\mathbb{E}[T_{a \rightarrow 0}])^2} \leq \frac{2 K^2 K_q}{\mathbb{E}[T_{a \rightarrow 0}]}, \quad (5.52)$$

$$\text{et } \lim_{a \rightarrow \infty} \text{Var}\left(\frac{T_{a \rightarrow 0}}{\mathbb{E}[T_{a \rightarrow 0}]}\right) = 0.$$

□

Remarque : ce théorème est l'analogie du théorème 5.1 de Martínez et Ycart [38] qui montre le cutoff pour les processus de naissance et de mort sur les arbres.

5.5 Métastabilité de la trajectoire de sortie

Maintenant on s'intéresse à la trajectoire de 0 à a . Toujours dans la terminologie physique on l'appelle "trajectoire de sortie" (de l'équilibre), ou encore trajectoire de montée si on pense à la particule qui doit remonter la pente du profil énergétique. On va montrer qu'avec la dérive vers zéro (hypothèse \mathcal{H}), elle a un comportement métastable.

Partant de l'état d'équilibre zéro, comme elle doit "remonter le courant", la particule va rester un temps exponentiel proche de zéro, jusqu'à la réalisation d'une excursion atypique qui l'amènera en a .

Comme nous l'avons vu au chapitre 4, ceci est dû à la différence d'échelles de temps, entre le temps moyen de tombée $\mathbb{E}[T_{a \rightarrow 0}]$ (temps de cutoff) et le temps moyen de sortie $\mathbb{E}[T_{0 \rightarrow a}]$. On va donc montrer que la condition (4.19) est satisfaite :

$$\lim_{a \rightarrow \infty} \frac{\mathbb{E}[T_{a \rightarrow 0}]}{\mathbb{E}[T_{0 \rightarrow a}]} = 0 \quad (5.53)$$

En collaboration avec J. Barrera, nous avons développé une technique qui consiste à décomposer la trajectoire de sortie (de zéro à a) pour la chaîne initialisée en zéro, à l'aide du couplage indépendant coalescent suivant.

On couple deux copies indépendantes de notre chaîne, une initialisée en zéro et une autre en a , qui évoluent ensemble après couplage jusqu'à ce qu'elles atteignent zéro ou a pour la première fois.

Soit $\hat{X} = (\hat{X}^0, \hat{X}^a)$ le processus couplé (pour alléger les notations, on omet l'indice a).

Remarque : comme il n'y a pas d'ambiguïtés, on notera également par \mathbb{P} la probabilité associée à ce couplage.

On a donc :

$$\begin{cases} \hat{X}^0 \stackrel{\text{loi}}{=} X^0 \\ \hat{X}^a \stackrel{\text{loi}}{=} X^a \end{cases} \quad \text{et } \hat{X}^0(t) = \hat{X}^a(t), \quad \forall t \geq \hat{T}_{\text{coupl}} = \inf\{t > 0 : \hat{X}^0(t) = \hat{X}^a(t)\}. \quad (5.54)$$

Soient $\hat{T}_{a \rightarrow 0} = \inf\{t > 0 : \hat{X}^a(t) = 0\}$ et $\hat{T}_{0 \rightarrow a} = \inf\{t > 0 : \hat{X}^0(t) = a\}$, on a aussi $\hat{T}_{a \rightarrow 0} \stackrel{\text{loi}}{=} T_{a \rightarrow 0}$ et $\hat{T}_{0 \rightarrow a} \stackrel{\text{loi}}{=} T_{0 \rightarrow a}$.

On dit qu'il y a renouvellement si le processus couplé atteint zéro avant d'atteindre a , et on définit l'instant de renouvellement

$$\widehat{T} = \inf\{t > \widehat{T}_{coupl} : \widehat{X}^0(t) = \widehat{X}^a(t) = 0\}. \quad (5.55)$$

Remarque : on a donc $\widehat{T} = \widehat{T}_{a \rightarrow 0}$, si $\widehat{T}_{a \rightarrow 0} < \widehat{T}_{0 \rightarrow a}$.

On recommence cette construction à chaque renouvellement, jusqu'à ce qu'on atteigne a pour la première fois (voir Fig 5.2).

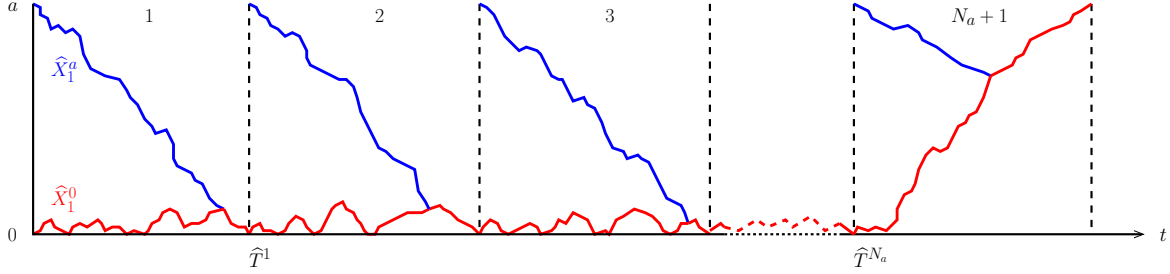


FIG. 5.2 – Couplage entre une chaîne initialisée en 0 et une en a

On construit donc itérativement une suite de processus couplés $\widehat{X}_i = (\widehat{X}_i^0, \widehat{X}_i^a)$ telle que $\widehat{X}_i \circ \theta_{\widehat{T}_{i-1}} \stackrel{loi}{=} \widehat{X}$, où θ_T est le shift, i.e. $\widehat{X}_i(t + \widehat{T}_{i-1}) \stackrel{loi}{=} \widehat{X}(t)$. Et on définit les temps d'arrêt :

$$\begin{cases} \widehat{T}_{coupl}^i = \inf\{t > 0 : \widehat{X}_i^a \circ \theta_{\widehat{T}_{i-1}}(t) = \widehat{X}_i^0 \circ \theta_{\widehat{T}_{i-1}}(t)\} \\ \widehat{T}_{a \rightarrow 0}^i = \inf\{t > 0 : \widehat{X}_i^a \circ \theta_{\widehat{T}_{i-1}}(t) = 0\} \\ \widehat{T}_{0 \rightarrow a}^i = \inf\{t > 0 : \widehat{X}_i^0 \circ \theta_{\widehat{T}_{i-1}}(t) = a\}, \end{cases} \quad (5.56)$$

où

$$\widehat{T}^i = \inf\{t > \widehat{T}_{coupl}^i : \widehat{X}_i^a \circ \theta_{\widehat{T}_{i-1}}(t) = \widehat{X}_i^0 \circ \theta_{\widehat{T}_{i-1}}(t) = 0\} + \widehat{T}^{i-1}. \quad (5.57)$$

Par la propriété de Markov forte les \widehat{X}_i sont i.i.d., et les temps d'arrêt correspondants le sont aussi.

Remarque : si $\forall j = 1, \dots, i : \widehat{T}_{a \rightarrow 0}^j < \widehat{T}_{0 \rightarrow a}^j$, alors $\widehat{T}^i = \sum_{j=1}^i \widehat{T}_{a \rightarrow 0}^j$.

On compte le nombre N_a d'instants de renouvellement avant d'atteindre a pour la première fois (ce n'est pas un temps d'arrêt).

C'est une variable géométrique de paramètre $\mathbb{P}(T_{a \rightarrow 0} < T_{0 \rightarrow a})$, où \mathbb{P} est ici la probabilité associée au couplage.

En effet

$$\{N_a = k\} = \{\widehat{T}_{a \rightarrow 0}^j < \widehat{T}_{0 \rightarrow a}^j, j = 1, \dots, k, \widehat{T}_{a \rightarrow 0}^{k+1} \geq \widehat{T}_{0 \rightarrow a}^{k+1}\}, \quad (5.58)$$

et comme les \widehat{X}_i^0 et les \widehat{X}_i^a sont indépendantes entres elles et iid, on a

$$\mathbb{P}(N_a = k) = \mathbb{P}(T_{a \rightarrow 0} < T_{0 \rightarrow a})^k \mathbb{P}(T_{a \rightarrow 0} \geq T_{0 \rightarrow a}), \quad (5.59)$$

et

$$\mathbb{P}(N_a \geq k) = \mathbb{P}(T_{a \rightarrow 0} < T_{0 \rightarrow a})^k. \quad (5.60)$$

D'où

$$\begin{aligned} \mathbb{E}[N_a] &= \sum_{k=0}^{\infty} \mathbb{P}(N_a \geq k) \\ &= \sum_{k=0}^{\infty} \mathbb{P}(T_{a \rightarrow 0} < T_{0 \rightarrow a})^k \\ &= \frac{1}{\mathbb{P}(T_{a \rightarrow 0} \geq T_{0 \rightarrow a})} \end{aligned} \quad (5.61)$$

A l'aide de ce couplage, on peut décomposer le temps $\widehat{T}_{0 \rightarrow a}$ de montée de zéro à a , comme la somme de N_a descentes de a à zéro plus une dernière montée :

$$\widehat{T}_{0 \rightarrow a} = \widehat{T}_{0 \rightarrow a}^1 = \widehat{T}^{N_a} + \widehat{T}_{0 \rightarrow a}^{N_a+1} = \sum_{i=1}^{N_a} \widehat{T}_{a \rightarrow 0}^i + \widehat{T}_{0 \rightarrow a}^{N_a+1}. \quad (5.62)$$

La décroissance exponentielle de la mesure invariante $\pi_a(\mathcal{H})$ implique deux échelles de temps très différentes. Avec probabilité un à la limite en a , le temps de tombée à l'équilibre est beaucoup plus court que le temps de sortie de l'équilibre.

Remarque : On a vu dans le théorème 5.3 que $\lim_{a \rightarrow \infty} \frac{T_{a \rightarrow 0}}{\mathbb{E}[T_{a \rightarrow 0}]} = 1$ en probabilité.

On montre aussi dans les lemmes suivants que le temps de sortie est plus grand que le temps moyen de tombée.

Sous l'hypothèse (\mathcal{H}) , on a

Lemme 5.4.

$$\lim_{a \rightarrow \infty} \mathbb{P}(T_{0 \rightarrow a} \leq (1 + \varepsilon) \mathbb{E}[T_{a \rightarrow 0}]) = 0 \quad (5.63)$$

Lemme 5.5.

$$\lim_{a \rightarrow \infty} \mathbb{P}(T_{a \rightarrow 0} \geq T_{0 \rightarrow a}) = 0 \quad (5.64)$$

Démonstration du lemme 5.4

Soit $ct_a = (1 + \varepsilon) \mathbb{E}[T_{a \rightarrow 0}]$.

On s'intéresse donc au processus $X_a^0(t)$ qui part de zéro, et à la probabilité qu'il atteigne a avant l'instant ct_a . On regarde les instants de retours en zéro pour ce processus.

Soient $T_{0 \rightarrow 0}^{(1)} = T_{0 \rightarrow 0}$, et pour tout $i > 1$:

$$T_{0 \rightarrow 0}^{(i)} = \inf\{t > 0 : X_a^0 \circ \theta_{T^{(i-1)}}(t) = 0\} \quad \text{où} \quad T^{(i)} = \sum_{j=1}^i T_{0 \rightarrow 0}^{(j)} \quad (5.65)$$

Par la propriété de Markov forte, chaque fois qu'il revient en zéro, il y a régénération du processus, dans le sens où pour tout i , la loi de $X_a^0(t + T^{(i)})$ est égale à celle de $X_a^0(t)$.

Soit aussi

$$T_{0 \rightarrow a}^{(i)} = \inf\{t > 0 : X_a^0 \circ \theta_{T^{(i-1)}}(t) = a\}, \quad (5.66)$$

le temps d'atteinte de a après le $(i-1)^{\text{ème}}$ retour en zéro.

Dire que le processus atteint a avant l'instant ct_a , revient à dire qu'il l'atteint entre deux retours à zéro avant cet instant.

Or, il y a au plus ct_a retours en zéro dans cet intervalle de temps.

On a donc

$$\begin{aligned} \mathbb{P}(T_{0 \rightarrow a} \leq ct_a) &\leq \mathbb{P}(T_{0 \rightarrow a} \leq \lceil ct_a \rceil) \\ &\leq \mathbb{P}(\exists i \text{ tq } T^{(i-1)} \leq \lceil ct_a \rceil : T_{0 \rightarrow a}^{(i)} < T_{0 \rightarrow 0}^{(i)}) \\ &\leq \sum_{i: T^{(i-1)} \leq \lceil ct_a \rceil} \mathbb{P}(T_{0 \rightarrow a}^{(i)} < T_{0 \rightarrow 0}^{(i)}) \\ &\leq \sum_{i \leq \lceil ct_a \rceil} \mathbb{P}(T_{0 \rightarrow a}^{(i)} < T_{0 \rightarrow 0}^{(i)}) \\ &= \lceil ct_a \rceil \mathbb{P}(T_{0 \rightarrow a} < T_{0 \rightarrow 0}). \end{aligned} \quad (5.67)$$

Et d'après le lemme 5.1, et la démonstration de (5.29)

$$\begin{aligned} \mathbb{P}(T_{0 \rightarrow a} < T_{0 \rightarrow 0}) &= p_0 \mathbb{P}(T_{1 \rightarrow a} < T_{1 \rightarrow 0}) \\ &= \frac{p_0}{q_1 \pi_a(1)} \frac{1}{\sum_{k=1}^a \frac{1}{q_k \pi_a(k)}} \\ &= \frac{1}{\sum_{k=1}^a \frac{\pi_a(0)}{q_k \pi_a(k)}}, \end{aligned} \quad (5.68)$$

puisque par réversibilité $q_1 \pi_a(1) = p_0 \pi_a(0)$.

Or on a vu (5.7) que pour tout x , $\pi_a(x) \leq e^{-\alpha x}$. On peut donc minorer

$$\sum_{k=1}^a \frac{1}{q_k \pi_a(k)} \geq \sum_{k=1}^a \frac{1}{\pi_a(k)} \geq \sum_{k=1}^a e^{\alpha k} > e^{\alpha a} \quad (5.69)$$

Et comme d'après (\mathcal{H})

$$t_a = \mathbb{E}[T_{a \rightarrow 0}] \leq K_q K a \quad \text{et} \quad (1/\pi_a(0)) < K, \quad (5.70)$$

où Kq est la borne uniforme des $(1/q_x)$. On a

$$\begin{aligned}
 ct_a \mathbb{P}(T_{0 \rightarrow a} < T_{0 \rightarrow 0}) &= \frac{ct_a}{\pi_a(0)} \frac{1}{\sum_{k=1}^a \frac{1}{q_k \pi_a(k)}} \\
 &< c K_q K^2 a \frac{1}{\sum_{k=1}^a \frac{1}{q_k \pi_a(k)}} \\
 &< c K_q K^2 a e^{-\alpha a},
 \end{aligned} \tag{5.71}$$

et

$$\lim_{a \rightarrow \infty} \mathbb{P}(T_{0 \rightarrow a} \leq (1 + \varepsilon) \mathbb{E}[T_{a \rightarrow 0}]) = 0.$$

□

Remarque : Dans la démonstration précédente, l'argument principal (5.71), est que

$$\lim_{a \rightarrow \infty} \mathbb{E}[T_{a \rightarrow 0}] \mathbb{P}(T_{0 \rightarrow a} < T_{0 \rightarrow 0}) = 0 \tag{5.72}$$

Démonstration du lemme 5.5

Pour montrer que $\mathbb{P}(T_{a \rightarrow 0} \geq T_{0 \rightarrow a})$ tend vers zéro, on va utiliser le lemme 5.4 plus la convergence abrupte. On écrit

$$\begin{aligned}
 \mathbb{P}(T_{0 \rightarrow a} \leq T_{a \rightarrow 0}) &= \mathbb{P}\left(\frac{T_{0 \rightarrow a}}{\mathbb{E}[T_{a \rightarrow 0}]} \leq \frac{T_{a \rightarrow 0}}{\mathbb{E}[T_{a \rightarrow 0}]}\right) \\
 &= \mathbb{P}\left(\frac{T_{0 \rightarrow a}}{\mathbb{E}[T_{a \rightarrow 0}]} \leq \frac{T_{a \rightarrow 0}}{\mathbb{E}[T_{a \rightarrow 0}]}, \frac{T_{a \rightarrow 0}}{\mathbb{E}[T_{a \rightarrow 0}]} \leq 1 + \varepsilon\right) \\
 &\quad + \mathbb{P}\left(\frac{T_{0 \rightarrow a}}{\mathbb{E}[T_{a \rightarrow 0}]} \leq \frac{T_{a \rightarrow 0}}{\mathbb{E}[T_{a \rightarrow 0}]}, \frac{T_{a \rightarrow 0}}{\mathbb{E}[T_{a \rightarrow 0}]} > 1 + \varepsilon\right) \\
 &\leq \mathbb{P}\left(\frac{T_{0 \rightarrow a}}{\mathbb{E}[T_{a \rightarrow 0}]} \leq 1 + \varepsilon\right) + \mathbb{P}\left(\frac{T_{a \rightarrow 0}}{\mathbb{E}[T_{a \rightarrow 0}]} > 1 + \varepsilon\right).
 \end{aligned} \tag{5.73}$$

D'après le lemme 5.4, le premier terme tend vers zéro, et d'après le théorème 5.3, le second aussi par définition de la convergence abrupte de $T_{a \rightarrow 0}$ (3.1).

□

Théorème 5.6. *Sous l'hypothèse (\mathcal{H}) , le temps de sortie $T_{0 \rightarrow a}$ a un comportement métastable.*

Démonstration

On doit montrer que les conditions du théorème 4.2 sont satisfaites. On a vu au début de ce chapitre que notre chaîne de naissance et de mort était irréductible et récurrente

positive. De plus la condition (4.21) est vérifiée de manière évidente ici. Et pour montrer que la suite $T_{0 \rightarrow a} / \mathbb{E}[T_{0 \rightarrow a}]$ est uniformément intégrable (condition (4.22)), on va montrer à l'aide des formules (9.1) et (9.13) qu'elle admet un moment d'ordre deux uniformément borné en a .

D'après (9.13) on a

$$\begin{aligned} \mathbb{E}[T_{0 \rightarrow a}^2] &= \sum_{k=0}^{a-1} \sum_{l=0}^k \frac{2 \mathbb{E}[T_{l \rightarrow a}] - 1}{p_l} \prod_{i=l+1}^k \frac{q_i}{p_i} \\ &< 2 \mathbb{E}[T_{0 \rightarrow a}] \sum_{k=0}^{a-1} \sum_{l=0}^k \frac{1}{p_l} \prod_{i=l+1}^k \frac{q_i}{p_i}, \end{aligned} \quad (5.74)$$

car $\mathbb{E}[T_{l \rightarrow a}] < \mathbb{E}[T_{0 \rightarrow a}] \forall l > 0$. Or le membre de droite est égal à $\mathbb{E}[T_{0 \rightarrow a}]$ (formule (9.1)), d'où

$$\mathbb{E}[T_{0 \rightarrow a}^2] < 2 (\mathbb{E}[T_{0 \rightarrow a}])^2, \quad (5.75)$$

c'est-à-dire que

$$\mathbb{E} \left[\left(\frac{T_{0 \rightarrow a}}{\mathbb{E}[T_{0 \rightarrow a}]} \right)^2 \right] < 2, \quad \forall a, \quad (5.76)$$

ce qui implique l'intégrabilité uniforme de la suite $T_{0 \rightarrow a} / \mathbb{E}[T_{0 \rightarrow a}]$ (voir [54], 13.3 (a)).

Il nous reste à montrer les deux échelles de temps (condition (4.19), formule (5.53)).

A l'aide de la décomposition (5.62) du temps de sortie $T_{0 \rightarrow a}$, on a

$$\begin{aligned} \mathbb{E}[T_{0 \rightarrow a}] &\geq \mathbb{E} \left[\sum_{i=1}^{N_a} T_{a \rightarrow 0}^{(i)} \right] \\ &= \sum_{k \geq 0} \sum_{i=1}^k \mathbb{E}[T_{a \rightarrow 0}^{(i)} \mid N_a = k] \mathbb{P}(N_a = k). \end{aligned} \quad (5.77)$$

Et d'après (5.59), pour tout n et pour tout i , on a

$$\begin{aligned} \mathbb{P}(T_{a \rightarrow 0}^{(i)} = n \mid N_a = k) &= \frac{\mathbb{P}(T_{a \rightarrow 0}^{(i)} = n, T_{a \rightarrow 0}^{(j)} < T_{0 \rightarrow a}^{(j)}, j = 1, \dots, k, T_{a \rightarrow 0}^{(k+1)} \geq T_{0 \rightarrow a}^{(k+1)})}{\mathbb{P}(T_{a \rightarrow 0}^{(j)} < T_{0 \rightarrow a}^{(j)}, j = 1, \dots, k, T_{a \rightarrow 0}^{(k+1)} \geq T_{0 \rightarrow a}^{(k+1)})} \\ &= \frac{\mathbb{P}(T_{a \rightarrow 0}^{(i)} = n, T_{a \rightarrow 0}^{(i)} < T_{0 \rightarrow a}^{(i)}) \mathbb{P}(T_{a \rightarrow 0} < T_{0 \rightarrow a})^{k-1} \mathbb{P}(T_{a \rightarrow 0} \geq T_{0 \rightarrow a})}{\mathbb{P}(T_{a \rightarrow 0} < T_{0 \rightarrow a})^k \mathbb{P}(T_{a \rightarrow 0} \geq T_{0 \rightarrow a})} \\ &= \frac{\mathbb{P}(T_{a \rightarrow 0} = n, T_{a \rightarrow 0} < T_{0 \rightarrow a})}{\mathbb{P}(T_{a \rightarrow 0} < T_{0 \rightarrow a})} \\ &= \mathbb{P}(T_{a \rightarrow 0} = n \mid T_{a \rightarrow 0} < T_{0 \rightarrow a}). \end{aligned} \quad (5.78)$$

Donc

$$\mathbb{E} [T_{a \rightarrow 0}^{(i)} \mid N_a = k] = \sum_{n \geq 0} n \mathbb{P} (T_{a \rightarrow 0} = n \mid T_{a \rightarrow 0} < T_{0 \rightarrow a}) , \quad \forall i . \quad (5.79)$$

D'où

$$\begin{aligned} \mathbb{E} [T_{0 \rightarrow a}] &\geq \sum_{k \geq 0} \sum_{i=1}^k \sum_{n \geq 0} n \mathbb{P} (T_{a \rightarrow 0} = n \mid T_{a \rightarrow 0} < T_{0 \rightarrow a}) \mathbb{P} (N_a = k) \\ &= \sum_{k \geq 0} k \mathbb{P} (N_a = k) \sum_{n \geq 0} n \mathbb{P} (T_{a \rightarrow 0} = n \mid T_{a \rightarrow 0} < T_{0 \rightarrow a}) \\ &= \mathbb{E} [N_a] \mathbb{E} [T_{a \rightarrow 0} \mid T_{a \rightarrow 0} < T_{0 \rightarrow a}] . \end{aligned} \quad (5.80)$$

Soit $t_a = \mathbb{E} [T_{a \rightarrow 0}]$ et soient $n_a^{\varepsilon-} = \lfloor (1 - \varepsilon)t_a \rfloor$, et $n_a^{\varepsilon+} = \lceil (1 + \varepsilon)t_a \rceil$, où $\lfloor x \rfloor$ désigne la partie entière inférieure de x , et $\lceil x \rceil$ sa partie entière supérieure.

Soit $C_a^\varepsilon = \llbracket n_a^{\varepsilon-}, n_a^{\varepsilon+} \rrbracket$.

On a

$$\begin{aligned} \mathbb{E} [T_{a \rightarrow 0} \mid T_{a \rightarrow 0} < T_{0 \rightarrow a}] &= \sum_{n \geq 0} n \mathbb{P} (T_{a \rightarrow 0} = n \mid T_{a \rightarrow 0} < T_{0 \rightarrow a}) \\ &\geq \sum_{n \in C_a^\varepsilon} n \mathbb{P} (T_{a \rightarrow 0} = n \mid T_{a \rightarrow 0} < T_{0 \rightarrow a}) \\ &\geq n_a^{\varepsilon-} \sum_{n \in C_a^\varepsilon} \frac{\mathbb{P} (T_{a \rightarrow 0} = n) \mathbb{P} (T_{0 \rightarrow a} > n)}{\mathbb{P} (T_{a \rightarrow 0} < T_{0 \rightarrow a})} \\ &\geq n_a^{\varepsilon-} \frac{\mathbb{P} (T_{0 \rightarrow a} > n_a^{\varepsilon+})}{\mathbb{P} (T_{a \rightarrow 0} < T_{0 \rightarrow a})} \mathbb{P} (T_{a \rightarrow 0} \in C_a^\varepsilon) . \end{aligned} \quad (5.81)$$

On peut maintenant montrer que (5.53) est satisfaite :

$$\lim_{a \rightarrow \infty} \frac{\mathbb{E} [T_{a \rightarrow 0}]}{\mathbb{E} [T_{0 \rightarrow a}]} = 0 .$$

En effet, on a

$$\begin{aligned} \frac{\mathbb{E} [T_{a \rightarrow 0}]}{\mathbb{E} [T_{0 \rightarrow a}]} &= \frac{\mathbb{E} [T_{a \rightarrow 0}]}{\mathbb{E} [T_{a \rightarrow 0} \mid T_{a \rightarrow 0} < T_{0 \rightarrow a}]} \frac{\mathbb{E} [T_{a \rightarrow 0} \mid T_{a \rightarrow 0} < T_{0 \rightarrow a}]}{\mathbb{E} [T_{0 \rightarrow a}]} \\ &\leq \frac{1}{1 - \varepsilon} \frac{\mathbb{P} (T_{a \rightarrow 0} < T_{0 \rightarrow a})}{\mathbb{P} (T_{0 \rightarrow a} > n_a^{\varepsilon+}) \mathbb{P} (T_{a \rightarrow 0} \in C_a^\varepsilon)} \frac{1}{\mathbb{E} [N_a]} . \end{aligned} \quad (5.82)$$

D'après le lemme 5.5, on a $\mathbb{P} (T_{a \rightarrow 0} < T_{0 \rightarrow a}) \xrightarrow{a \rightarrow \infty} 1$.

Par le même argument que celui de la démonstration du lemme 5.4 : $\mathbb{P} (T_{0 \rightarrow a} > n_a^{\varepsilon+}) \xrightarrow{a \rightarrow \infty} 1$.

Et de plus

$$\mathbb{P}(T_{a \rightarrow 0} \in C_a^\varepsilon) \geq \mathbb{P}\left(\left|\frac{T_{a \rightarrow 0}}{\mathbb{E}[T_{a \rightarrow 0}]} - 1\right| \leq \varepsilon\right) \xrightarrow{a \rightarrow \infty} 1, \quad (5.83)$$

d'après la définition de la convergence abrupte (définition 3.1) et le théorème 5.3.

D'où le résultat, puisque d'après le lemme 5.5 et la formule (5.61) : $\mathbb{E}[N_a] \xrightarrow{a \rightarrow \infty} +\infty$.

□

5.6 Dernière sortie

On appelle dernière sortie, l'excursion atypique qui emmène la particule de zéro à a . Elle correspond à la trajectoire de la particule entre la dernière visite à zéro et la première fois où elle atteint a .

Nous allons montrer que cette trajectoire est la renversée temporelle d'une trajectoire "typique" de cutoff.

Pour cela, nous allons considérer uniquement les trajectoires finies "strictement comprises entre zéro et a ", c'est-à-dire les trajectoires de a à zéro et de zéro à a qui ne repassent pas par les extrémités entre leurs points de départ et d'arrivée. Ce sont donc l'ensemble des réalisations $\{\tilde{T}_{a \rightarrow 0} < \infty\}$ et $\{\tilde{T}_{0 \rightarrow a} < \infty\}$ où $\tilde{T}_{a \rightarrow 0}$ et $\tilde{T}_{0 \rightarrow a}$ sont définis comme suit :

$$\tilde{T}_{a \rightarrow 0} = \inf\{t > 0 : X_a^a \circ \theta_{\tau_a}(t) = 0\}, \quad (5.84)$$

où τ_a est la dernière visite en a de la particule initialement en a avant qu'elle n'atteigne zéro pour la première fois

$$\tau_a = \sup\{0 < t < T_{a \rightarrow 0} : X_a^a(t) = a\}, \quad (5.85)$$

et

$$\tilde{T}_{0 \rightarrow a} = \inf\{t > 0 : X_a^0 \circ \theta_{\tau_0}(t) = a\}, \quad (5.86)$$

où τ_0 est la dernière visite en zéro de la particule initialement en zéro avant qu'elle n'atteigne a pour la première fois

$$\tau_0 = \sup\{0 < t < T_{0 \rightarrow a} : X_a^0(t) = 0\}. \quad (5.87)$$

Remarque : la dernière sortie caractérisée par $\tilde{T}_{0 \rightarrow a}$, n'est pas celle caractérisée par $\widehat{T}_{0 \rightarrow a}^{N_a+1}$ dans la décomposition (5.62) de $T_{0 \rightarrow a}$ à l'aide du couplage dans la section précédente. En effet dans $\widehat{T}_{0 \rightarrow a}^{N_a+1}$, on autorise d'éventuels retours en zéro, alors qu'ici on conditionne à ne plus repasser par zéro. On a donc $\tilde{T}_{0 \rightarrow a} \leq \widehat{T}_{0 \rightarrow a}^{N_a+1}$.

De même, $\tilde{T}_{a \rightarrow 0}$ n'est pas égal à $T_{a \rightarrow 0}$ puisque l'on conditionne à ne pas repasser par a .

Nous allons maintenant montrer le résultat en deux temps.

Premièrement, nous allons montrer que pour toute chaîne irréductible, les temps d'arrêt $\tilde{T}_{a \rightarrow 0}$ et $\tilde{T}_{0 \rightarrow a}$ définissent des trajectoires qui sont les renversées temporelles l'une de l'autre (proposition 5.7), et qu'elles ont la même loi (proposition 5.8).

Et dans un deuxième temps, nous montrerons que sous l'hypothèse de dérive (\mathcal{H}) , $\tilde{T}_{a \rightarrow 0}$ admet une convergence abrupte au temps $\mathbb{E} [\tilde{T}_{a \rightarrow 0}]$ (proposition 5.9).

L'approche suivante est due à R.H. Schonmann dans "The Pattern of Escape from Metastability of a Stochastic Ising Model" [50]. Elle permet de déterminer la trajectoire typique de sortie de l'état métastable vers une "goutte critique" pour un système ferromagnétique. Schonmann montre en utilisant la réversibilité du modèle, que cette trajectoire est la renversée temporelle d'une trajectoire typique de la goutte sous-critique vers l'état métastable.

La technique qui consiste à renverser le temps est indépendante du modèle et uniquement basée sur la réversibilité de la chaîne. Nous en donnons ici une version adaptée à nos modèles.

On définit \mathcal{U}_a l'ensemble des trajectoires finies de a à zéro "strictement entre zéro et a " et pour tout $k \in \mathbb{N}$ fixé, \mathcal{V}_a^k celles de longueur k . On note par $v^k = (v_1, \dots, v_k)$ un chemin de longueur k , et donc

$$\mathcal{V}_a^k = \{v^k : v_1 = a, v_k = 0, \text{ et } v_i \neq a, v_i \neq 0, \forall 0 < i < k\}, \quad \mathcal{U}_a = \bigcup_{k < \infty} \mathcal{V}_a^k. \quad (5.88)$$

Avec la terminologie de Schonmann, on dit que " X_a^a starts as v^k " si X_a^a suit le chemin v^k , i.e. si pour tout $i = 1, \dots, k$, $X_a^a(i-1) = v_i$, et pour un ensemble de trajectoires \mathcal{V} , on dit que " X_a^a starts in \mathcal{V} " si X_a^a suit un chemin $v \in \mathcal{V}$.

On définit maintenant la renversée temporelle d'une trajectoire, il s'agit tout simplement de la parcourir dans l'autre sens :

$$\mathcal{R}(v^k) = \{v_k, \dots, v_1\}, \quad (5.89)$$

et pour un ensemble de trajectoires \mathcal{V}

$$\mathcal{R}(\mathcal{V}) = \{\mathcal{R}(v), v \in \mathcal{V}\}. \quad (5.90)$$

$\mathcal{R}(\mathcal{U}_a)$ est donc l'ensemble des chemins finis de zéro à a "strictement entre zéro et a ", et $\mathcal{R}(\mathcal{V}_a^k)$ ceux de longueur k .

On peut relier ces définitions avec les temps $\tilde{T}_{a \rightarrow 0}$ et $\tilde{T}_{0 \rightarrow a}$: en effet $\tilde{T}_{a \rightarrow 0}$ est fini si et seulement si la chaîne X_a^a suit un chemin de \mathcal{U}_a après sa dernière visite en a

$$\{\tilde{T}_{a \rightarrow 0} < \infty\} = \{X_a^a \circ \theta_{\tau_a} \text{ starts in } \mathcal{U}_a\}, \quad (5.91)$$

et de même, pour tout k

$$\{\tilde{T}_{a \rightarrow 0} = k\} = \{X_a^a \circ \theta_{\tau_a} \text{ starts in } \mathcal{V}_a^k\}. \quad (5.92)$$

Et donc, pour toute chaîne irréductible, les temps $\tilde{T}_{a \rightarrow 0}$ et $\tilde{T}_{0 \rightarrow a}$ que l'on a définis ci-dessus sont tels que

Proposition 5.7.

$$\{\tilde{T}_{0 \rightarrow a} < \infty\} = \{X_a^0 \circ \theta_{\tau_0} \text{ starts in } \mathcal{R}(\mathcal{U}_a)\} = \mathcal{R}(\{\tilde{T}_{a \rightarrow 0} < \infty\}) , \quad (5.93)$$

et pour tout k

$$\{\tilde{T}_{0 \rightarrow a} = k\} = \{X_a^0 \circ \theta_{\tau_0} \text{ starts in } \mathcal{R}(\mathcal{V}_a^k)\} = \mathcal{R}(\{\tilde{T}_{a \rightarrow 0} = k\}) . \quad (5.94)$$

On dit alors que les trajectoires définies par $\tilde{T}_{0 \rightarrow a}$ sont les renversées temporelles de celles définies par $\tilde{T}_{a \rightarrow 0}$.

De plus, et toujours sans perte de généralité :

Proposition 5.8. *Les variables $\tilde{T}_{a \rightarrow 0}$ et $\tilde{T}_{0 \rightarrow a}$ suivent la même loi.*

Démonstration

En suivant la démonstration de la proposition 3 de Schonmann [50], on va d'abord montrer que

$$\mathbb{P}(X_a^a \text{ starts in } \mathcal{V}_a^k \mid X_a^a \text{ starts in } \mathcal{U}_a) = \mathbb{P}(X_a^0 \text{ starts in } \mathcal{R}(\mathcal{V}_a^k) \mid X_a^0 \text{ starts in } \mathcal{R}(\mathcal{U}_a)) . \quad (5.95)$$

Soit $v^k \in \mathcal{V}_a^k$, en utilisant la propriété de Markov et la réversibilité de la mesure invariante, on a

$$\begin{aligned} \mathbb{P}(X_a^a \text{ starts as } v^k) &= p(a, v_2) p(v_2, v_3) \cdots p(v_{k-2}, v_{k-1}) p(v_{k-1}, 0) \\ &= p(a, v_2) p(v_2, v_3) \cdots p(v_{k-2}, v_{k-1}) p(v_{k-1}, 0) (\pi_a(0))^{-1} \pi_a(0) \\ &= p(a, v_2) p(v_2, v_3) \cdots p(v_{k-2}, v_{k-1}) (\pi_a(v_{k-1}))^{-1} p(0, v_{k-1}) \pi_a(0) \\ &= p(a, v_2) p(v_2, v_3) \cdots (\pi_a(v_{k-2}))^{-1} p(v_{k-1}, v_{k-2}) p(0, v_{k-1}) \pi_a(0) \\ &= \dots\dots\dots \\ &= (\pi_a(a))^{-1} p(v_2, a) p(v_3, v_2) \cdots p(v_{k-1}, v_{k-2}) p(0, v_{k-1}) \pi_a(0) \\ &= (\pi_a(0) / \pi_a(a)) \mathbb{P}(X_a^0 \text{ starts as } \mathcal{R}(v^k)) . \end{aligned} \quad (5.96)$$

De même, en sommant sur les $v^k \in \mathcal{V}_a^k$,

$$\mathbb{P}(X_a^a \text{ starts in } \mathcal{V}_a^k) = (\pi_a(0) / \pi_a(a)) \mathbb{P}(X_a^0 \text{ starts in } \mathcal{R}(\mathcal{V}_a^k)) , \quad (5.97)$$

et en sommant sur k

$$\mathbb{P}(X_a^a \text{ starts in } \mathcal{U}_a) = (\pi_a(0) / \pi_a(a)) \mathbb{P}(X_a^0 \text{ starts in } \mathcal{R}(\mathcal{U}_a)) , \quad (5.98)$$

d'où (5.95) car

$$\mathbb{P}(X_a^a \text{ starts in } \mathcal{V}_a^k \mid X_a^a \text{ starts in } \mathcal{U}_a) = \frac{\mathbb{P}(X_a^a \text{ starts in } \mathcal{V}_a^k)}{\mathbb{P}(X_a^a \text{ starts in } \mathcal{U}_a)} . \quad (5.99)$$

Il nous reste à montrer que

$$\mathbb{P}(\tilde{T}_{a \rightarrow 0} = k) = \mathbb{P}(X_a^a \text{ starts in } \mathcal{V}_a^k \mid X_a^a \text{ starts in } \mathcal{U}_a) . \quad (5.100)$$

Or

$$\{\tau_a = n\} = \{X_a^a(t) \neq 0, t < n, X_a^a(n) = a, X_a^a(t) \neq a, t > n\} , \quad (5.101)$$

donc, pour tout $v^k \in \mathcal{V}_a^k$

$$\{\tau_a = n, X_a^a \circ \theta_n \text{ starts as } v^k\} = \{X_a^a(t) \neq 0, t < n, X_a^a \circ \theta_n \text{ starts as } v^k\} , \quad (5.102)$$

car $v_1 = a$ et $v_i \neq a$ pour tout $i = 2, \dots, k$.

En partitionnant sur les valeurs de τ_a et en utilisant la propriété de Markov, on en déduit

$$\begin{aligned} \mathbb{P}(\tilde{T}_{a \rightarrow 0} = k) &= \mathbb{P}(X_a^a \circ \theta_{\tau_a} \text{ starts in } \mathcal{V}_a^k) \\ &= \sum_n \mathbb{P}(\tau_a = n, X_a^a \circ \theta_n \text{ starts in } \mathcal{V}_a^k) \\ &= \sum_n \mathbb{P}(X_a^a(t) \neq 0, t < n, X_a^a \circ \theta_n \text{ starts in } \mathcal{V}_a^k) \\ &= \mathbb{P}(X_a^a \text{ starts in } \mathcal{V}_a^k) \sum_n \mathbb{P}(X_a^a(t) \neq 0, t < n) . \end{aligned} \quad (5.103)$$

En sommant sur k , on a

$$1 = \sum_n \mathbb{P}(X_a^a(t) \neq 0, t < n) \sum_k \mathbb{P}(X_a^a \text{ starts in } \mathcal{V}_a^k) . \quad (5.104)$$

Or

$$\bigcup_k \{X_a^a \text{ starts in } \mathcal{V}_a^k\} = \{X_a^a \text{ starts in } \mathcal{U}_a\} , \quad (5.105)$$

d'où

$$\sum_n \mathbb{P}(X_a^a(t) \neq 0, t < n) = (\mathbb{P}(X_a^a \text{ starts in } \mathcal{U}_a))^{-1} , \quad (5.106)$$

et on obtient (5.100).

De la même manière, on montre que

$$\mathbb{P}(\tilde{T}_{a \rightarrow a} = k) = \mathbb{P}(X_a^0 \text{ starts in } \mathcal{R}(\mathcal{V}_a^k) \mid X_a^0 \text{ starts in } \mathcal{R}(\mathcal{U}_a)) , \quad (5.107)$$

d'où l'égalité des lois en utilisant (5.95) : pour tout k , on a

$$\mathbb{P}(\tilde{T}_{a \rightarrow 0} = k) = \mathbb{P}(\tilde{T}_{0 \rightarrow a} = k) \quad (5.108)$$

□

Schonmann utilise cette propriété de la réversibilité dans une situation métastable. Il peut ainsi déterminer la trajectoire de sortie du puits d'attraction de l'état métastable comme la renversée d'une trajectoire typique de tombée dans ce puits.

Dans notre formalisme, on va montrer que cette trajectoire typique de tombée est en fait une trajectoire typique de cutoff.

Proposition 5.9. *Sous l'hypothèse (\mathcal{H}) de dérive vers zéro, le temps $\tilde{T}_{a \rightarrow 0}$ admet une convergence abrupte au temps $\mathbb{E} \left[\tilde{T}_{a \rightarrow 0} \right]$.*

Démonstration

Regarder $\tilde{T}_{a \rightarrow 0}$, c'est regarder le temps d'atteinte de zéro pour une particule initialement en a conditionnée à ne pas revenir en a ensuite. Le premier pas de la particule se fait donc vers $a - 1$, et une fois en $a - 1$, comme elle ne peut pas revenir en a , on peut définir son évolution à l'aide du processus \bar{X}_{a-1}^{a-1} sur $\llbracket 0, a - 1 \rrbracket$ dont les transitions \bar{p}_i et \bar{q}_i sont définies à l'aide de celles de X_a :

$$\begin{cases} \bar{p}_i = p_i & \text{et} \quad \bar{q}_i = q_i \quad \forall i < a - 1 \\ \bar{p}_{a-1} = 0 & \text{et} \quad \bar{q}_{a-1} = 1 \end{cases} \quad (5.109)$$

On peut donc écrire

$$\tilde{T}_{a \rightarrow 0} = 1 + \bar{T}_{a-1 \rightarrow 0} . \quad (5.110)$$

Et $\bar{T}_{a-1 \rightarrow 0}$ admet une convergence abrupte au temps $\mathbb{E} \left[\bar{T}_{a-1 \rightarrow 0} \right]$. En effet la mesure invariante $\bar{\pi}_{a-1}$ du processus \bar{X}_{a-1}^{a-1} s'écrit

$$\bar{\pi}_{a-1}(x) = \prod_{i=1}^x \frac{\bar{p}_{i-1}}{\bar{q}_i} \bar{\pi}_a(0) , \quad (5.111)$$

et donc le rapport $\bar{\pi}_{a-1}(\bar{B}_x) / \bar{\pi}_{a-1}(x)$ (où $\bar{B}_x = \llbracket x, a - 1 \rrbracket$) satisfait la condition (\mathcal{H})

$$\begin{aligned} \frac{\bar{\pi}_{a-1}(\bar{B}_x)}{\bar{\pi}_{a-1}(x)} &= \sum_{k=x}^{a-1} \prod_{i=x+1}^k \frac{\bar{p}_{i-1}}{\bar{q}_i} \\ &= \sum_{k=x}^{a-2} \prod_{i=x+1}^k \frac{p_{i-1}}{q_i} + \prod_{i=x+1}^{a-1} \frac{\bar{p}_{i-1}}{\bar{q}_i} \\ &= \sum_{k=x}^{a-2} \prod_{i=x+1}^k \frac{p_{i-1}}{q_i} + \prod_{i=x+1}^{a-1} \frac{p_{i-1}}{q_i} q_{a-1} \\ &< \sum_{k=x}^{a-1} \prod_{i=x+1}^k \frac{p_{i-1}}{q_i} , \end{aligned} \quad (5.112)$$

puisque dans cette somme la seule transition qui change entre X_{a-1}^{a-1} et \bar{X}_{a-1}^{a-1} est $\bar{q}_{a-1} = 1$, alors que $q_{a-1} < 1$ par hypothèse.

On a donc

$$\frac{\bar{\pi}_{a-1}(\bar{B}_x)}{\bar{\pi}_{a-1}(x)} < \frac{\pi_a(B_x)}{\pi_a(x)} = \sum_{k=x}^a \prod_{i=x+1}^k \frac{p_{i-1}}{q_i}, \quad (5.113)$$

et comme on a supposé que le rapport $\pi_a(B_x)/\pi_a(x)$ est borné uniformément en a (\mathcal{H}), il en est de même pour $\bar{\pi}_{a-1}(\bar{B}_x)/\bar{\pi}_{a-1}(x)$.

Par le théorème 5.3 on en déduit que $\bar{T}_{a-1 \rightarrow 0}$ admet une convergence abrupte au temps $\mathbb{E} [\bar{T}_{a-1 \rightarrow 0}]$. c'est-à-dire que $\bar{T}_{a-1 \rightarrow 0}/\mathbb{E} [\bar{T}_{a-1 \rightarrow 0}]$ converge vers 1 en probabilité.

Maintenant, comme $\tilde{T}_{a \rightarrow 0} = 1 + \bar{T}_{a-1 \rightarrow 0}$, on a aussi

$$\mathbb{E} [\tilde{T}_{a \rightarrow 0}] = 1 + \mathbb{E} [\bar{T}_{a-1 \rightarrow 0}], \quad (5.114)$$

et on a

$$\begin{aligned} \frac{\tilde{T}_{a \rightarrow 0}}{\mathbb{E} [\tilde{T}_{a \rightarrow 0}]} &= \frac{1 + \bar{T}_{a-1 \rightarrow 0}}{1 + \mathbb{E} [\bar{T}_{a-1 \rightarrow 0}]} \\ &= \left(\frac{1}{\mathbb{E} [\bar{T}_{a-1 \rightarrow 0}]} + \frac{\bar{T}_{a-1 \rightarrow 0}}{\mathbb{E} [\bar{T}_{a-1 \rightarrow 0}]} \right) \left(\frac{1}{1 + \frac{1}{\mathbb{E} [\bar{T}_{a-1 \rightarrow 0}]}} \right), \end{aligned} \quad (5.115)$$

or $1/\mathbb{E} [\bar{T}_{a-1 \rightarrow 0}]$ tend vers zéro, et donc $\tilde{T}_{a \rightarrow 0}/\mathbb{E} [\tilde{T}_{a \rightarrow 0}]$ converge vers 1 en probabilité, et $\tilde{T}_{a \rightarrow 0}$ admet une convergence abrupte au temps $\mathbb{E} [\tilde{T}_{a \rightarrow 0}]$.

□

Chapitre 6

Modèle d'Ehrenfest

On présente ici une version du modèle “historique” d'Ehrenfest. Après avoir rappelé le modèle original et discuté des nombreux travaux auquel il a donné lieu (section 6.1), on présente notre version du modèle qui rentre dans le cadre des chaînes de naissance et de mort du chapitre précédent (section 6.2). On va montrer que bien qu'il ne satisfasse pas les hypothèses du chapitre précédent, la convergence abrupte et la métastabilité ont lieu (section 6.4). Les calculs sont faits dans la section 6.3.

6.1 Introduction

Le modèle d'Ehrenfest, introduit par P. et T. Ehrenfest en 1907 (voir “The Conceptual Foundations of the Statistical Approach in Mechanics” [23]), modélise les échanges de particules de gaz entre deux réservoirs, ou encore les échanges de chaleur entre deux corps isolés de températures différentes. A l'origine, les Ehrenfest ont utilisé ce modèle pour défendre l'approche de Boltzmann qui expliquait la thermodynamique à partir de la théorie cinétique des gaz, et résoudre le paradoxe apparent entre la réversibilité des évolutions mécaniques (au niveau microscopique), et l'irréversibilité (macroscopique) des phénomènes thermodynamiques. M. Kac discute de ces aspects dans son article 1947 “Random Walk and the Theory of Brownian Motion” [31].

Essentiellement, le modèle d'Ehrenfest pour N particules est équivalent à la marche aléatoire sur le cube de dimension N $\{-1, 1\}^N$ (ou hypercube). On trouvera en appendice une définition précise de la marche aléatoire sur le cube en temps discret, ainsi que le lien avec le modèle d'Ehrenfest original.

Dans “Asymptotic Analysis of a Random Walk on a Hypercube with Many Dimensions” [20], Diaconis étudie la marche aléatoire sur le cube de dimension N . En utilisant le modèle à temps continu, et par des arguments d'analyse de Fourier, il montre qu'il y a un phénomène de cutoff au temps $(1/4) N \log N$ dans la convergence à l'équilibre du système.

La notion de cutoff utilisée (cf (3.2)) est celle de la convergence en variation totale de la loi du processus au temps t vers la mesure d'équilibre du système, qui dans ce cadre est la mesure uniforme sur le cube (9.24).

Notre approche de la convergence abrupte et de la métastabilité ne peut prendre en compte un tel modèle, puisqu'une énergie associée à la mesure uniforme (cf (5.8)) serait constante et ne présenterait donc pas de puits d'attraction. On ne pourrait donc pas caractériser l'équilibre par un état au sens où on l'entend en physique. En effet, dans l'approche physicienne les états sont des configurations du système, alors que pour les probabilistes les états sont des mesures de probabilité.

Par contre, notre approche physicienne cadre tout à fait avec le modèle original d'Ehrenfest. En effet, celui-ci consiste à compter le nombre de particules dans une des deux urnes (voir en appendice la description réduite du modèle). C'est une marche aléatoire entre 0 et N , avec une dérive vers $N/2$.

L'équilibre thermodynamique se situe donc autour de $N/2$, qui est l'état qui pour nous caractérise l'équilibre. C'est l'état qui minimise l'énergie (5.8) associée à la mesure invariante (9.27) ($N/2$ correspond à l_N au chapitre 4). Le profil énergétique est donc un puits centré autour de $N/2$. Si la chaîne commence en 0 (ou en N), elle est attirée vers le fond du puits, et on va montrer que cette trajectoire correspond au phénomène de convergence abrupte. Par contre, lorsqu'elle est en $N/2$, elle doit "remonter la pente" pour atteindre 0 ou $N/2$. On va montrer que cette trajectoire correspond à la trajectoire de sortie de l'équilibre (voir les chapitres 4 et 5), et qu'elle a un comportement métastable.

Pour des raisons de symétrie évidentes, on voit bien que le problème de la sortie du puits d'attraction de $N/2$, est en fait le problème de remonter la pente, et que le comportement ne diffère pas, suivant qu'on cherche à atteindre 0 ou N . Le modèle qu'on propose au paragraphe suivant est un demi modèle d'Ehrenfest, et rentre dans le cadre des chaînes de naissance et de mort entre 0 et a du chapitre précédent. C'est le profil entre $N/2$ et N du modèle original que l'on a translaté entre 0 et a (voir Fig. 6.1).

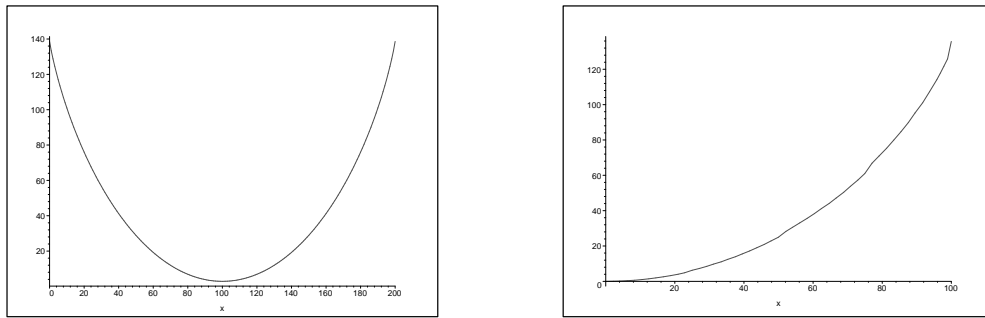


FIG. 6.1 – Ehrenfest : profils (modèle original avec $N = 200$, et demi modèle avec $a = 100$).

La seule différence avec les chaînes du chapitre 5, est que le rapport $\pi_a(B_x)/\pi_a(x)$ n'est pas uniformément borné en a par une constante K . La constante dépend ici de a , mais les

résultats de convergence abrupte et de métastabilité sont conservés, et nos démonstrations restent valables. Ceci est dû à la vitesse relative de la borne K_a par rapport à celle du temps moyen de tombée à l'équilibre. Nous développerons plus en détail la comparaison entre les deux modèles au chapitre suivant.

Le modèle d'Ehrenfest, la marche aléatoire sur l'hypercube et diverses généralisations de ces modèles ont été largement étudiés. Parmi ces nombreux travaux, citons M. Kac [31], qui calcule les probabilités de transition $p^k(x, y)$ pour tous x, y et tout k , et les temps de retours pour le modèle d'Ehrenfest original en temps discret. Ces calculs sont également démontrés et généralisés dans les articles de Takács [52, 36]. Karlin et McGregor s'intéressent aux généralisations à plusieurs urnes du modèle d'Ehrenfest en temps continu [32]. Plus en rapport avec notre problématique, Bellman et Harris regardent le modèle d'Ehrenfest à $2n$ particules en temps continu. Avec des outils d'analyse, ils montrent que si $k/n \leq \lambda < 1$, le rapport $T_{n \rightarrow k} / \mathbb{E}[T_{n \rightarrow k}]$ converge en loi vers une variable exponentielle de paramètre un [6]. Ils montrent en fait la métastabilité pour la trajectoire de sortie de l'équilibre (ici l'état n est l'état d'équilibre, et la condition $k/n \leq \lambda < 1$ impose que $(n - k) \xrightarrow[n \rightarrow \infty]{} \infty$).

Dans le même esprit, mais toujours sans parler de métastabilité, Bingham [8] montre la limite exponentielle pour le temps de passage au sommet opposé pour la marche aléatoire sur l'hypercube, c'est-à-dire le temps d'atteinte de $(1, \dots, 1)$ partant de $(0, \dots, 0)$. Il décompose la trajectoire en trois parties : de l'origine à la “bande centrale” des états typiques, puis de nombreuses excursions sans succès pour quitter cette “bande centrale”, et enfin une dernière excursion jusqu'en $(1, \dots, 1)$, et précise que le caractère exponentiel provient de la seconde partie. Au vu des résultats du chapitre 5, on peut maintenant préciser : la première partie correspond à la convergence abrupte, la seconde à la métastabilité, et la dernière excursion est la renversée temporelle d'une trajectoire de cutoff.

Citons enfin les travaux de Matthews [39], qui donne un couplage et un temps uniforme fort pour la variation totale et la séparation pour la marche aléatoire sur le cube. Ceux de Burdzy et Kendall [11], qui construisent un couplage efficace dont le temps de couplage est de l'ordre de $n \log n$. et enfin ceux de Bovier, qui discute certains aspects de la métastabilité pour la marche aléatoire sur le cube dans [9].

6.2 Modèle

On considère donc la chaîne de naissance et de mort sur $\llbracket 0, a \rrbracket$ à temps discret dont les transitions sont

$$p_i = \frac{a - i}{2a}, \quad \text{et} \quad q_i = \frac{a + i}{2a}. \quad (6.1)$$

Comme $q_i > 1/2$ pour tout $i > 0$, et que $p_i > 0$ pour tout $i < a$ elle est uniformément irréductible. De plus, lorsque a est fixé, la chaîne est à espace d'états fini, donc récurrente positive et admet une unique mesure de probabilité invariante π_a . On peut la déterminer car la chaîne est réversible par rapport à π_a .

Soit $x \in \llbracket 0, a \rrbracket$, par définition de la réversibilité on a $\pi_a(x-1) p_{x-1} = \pi_a(x) q_x$, d'où

$$\pi_a(x) = \prod_{i=1}^x \frac{p_{i-1}}{q_i} \pi_a(0) = \prod_{i=1}^x \frac{a-i+1}{a+i} \pi_a(0) . \quad (6.2)$$

avec, en utilisant la normalisation $\sum_{k=0}^a \pi_a(k) = 1$

$$\frac{1}{\pi_a(0)} = \sum_{k=0}^a \prod_{i=1}^k \frac{p_{i-1}}{q_i} = \sum_{k=0}^a \prod_{i=1}^k \frac{a-i+1}{a+i} . \quad (6.3)$$

On va montrer que ce modèle satisfait une condition (\mathcal{H}') un peu plus générale que la condition (\mathcal{H}) du chapitre 5 (Convergence abrupte et métastabilité pour les chaînes de naissance et de mort), et qu'on a aussi convergence abrupte de a vers 0, métastabilité de 0 vers a , et que la dernière sortie qui emmène la particule de 0 à a est la renversée temporelle d'une trajectoire de cutoff.

6.3 Quantités remarquables

On s'intéresse au comportement asymptotique en a du modèle, et on a besoin des estimations suivantes

Proposition 6.1. *Soit $B_x = \llbracket x, a \rrbracket$, on a*

1. *pour tout $x \in \llbracket 0, a \rrbracket$:*

$$\frac{\pi_a(B_x)}{\pi_a(x)} \leq \frac{1}{\pi_a(0)} , \quad (6.4)$$

2. *et asymptotiquement en a :*

$$\frac{1}{\pi_a(0)} \sim \sqrt{a} , \quad \mathbb{E}[T_{a \rightarrow 0}] \sim a \ln a , \quad (6.5)$$

Où le symbole \sim signifie que les deux quantités comparées sont asymptotiquement équivalentes à des constantes près : $u_a \sim v_a$ équivaut à $\exists C_1, C_2 > 0$ telles que

$$C_1 \leq \lim_{a \rightarrow \infty} \frac{u_a}{v_a} \leq C_2 . \quad (6.6)$$

Démonstration

On a

$$A_x := \frac{\pi_a(B_x)}{\pi_a(x)} = \sum_{k=x}^a \frac{\pi_a(k)}{\pi_a(x)} = \sum_{k=x}^a \prod_{i=x+1}^k \frac{p_{i-1}}{q_i} = \sum_{k=x}^a \prod_{i=x+1}^k \frac{a-i+1}{a+i} \quad (6.7)$$

A_x est décroissante en x , en effet

$$\begin{aligned}
 A_x &= 1 + \sum_{k=x+1}^a \prod_{i=x+1}^k \frac{a-i+1}{a+i} \\
 &= 1 + \frac{a-x}{a+x+1} \sum_{k=x+1}^a \prod_{i=x+2}^k \frac{a-i+1}{a+i} \\
 &= 1 + \frac{a-x}{a+x+1} A_{x+1}
 \end{aligned}$$

Donc

$$A_x - A_{x+1} = 1 - \frac{2x+1}{a+x+1} A_{x+1} \quad (6.8)$$

Or on peut majorer

$$\begin{aligned}
 A_{x+1} &= \sum_{k=x+1}^a \prod_{i=x+2}^k \frac{a-i+1}{a+i} \\
 &< \sum_{k=x+1}^a \left(\frac{a-x}{a+x+1} \right)^{(k-x-1)} \\
 &= \sum_{k=0}^{a-x-1} \left(\frac{a-x}{a+x+1} \right)^k \\
 &< \frac{1}{1 - \frac{a-x}{a+x+1}} = \frac{a+x+1}{2x+1}
 \end{aligned}$$

Et la différence $A_x - A_{x+1}$ est strictement positive, ce qui montre la première assertion de la proposition.

Montrons maintenant que $\frac{1}{\pi_a(0)} \sim \sqrt{a}$. En réordonnant les numérateurs de la façon suivante

$$\prod_{i=1}^k \frac{p_{i-1}}{q_i} = \prod_{i=1}^k \frac{p_{k-i}}{q_i} \quad (6.9)$$

on écrit

$$\frac{1}{\pi_a(0)} = \sum_{k=0}^a \prod_{i=1}^k \frac{a-i+1}{a+i} = \sum_{k=0}^a \prod_{i=1}^k \frac{a-k+i}{a+i} \quad (6.10)$$

Qu'on peut majorer

$$\begin{aligned}
\frac{1}{\pi_a(0)} &< \sum_{k=0}^a \left(\frac{a}{a+k} \right)^k \\
&= \sum_{k=0}^a \exp \left(-k \ln \left(1 + \frac{k}{a} \right) \right) \\
&\leq \sum_{k=0}^a \exp \left(-\frac{k^2}{a+k} \right) \\
&\leq \sum_{k=0}^a \exp \left(-\frac{k^2}{2a} \right)
\end{aligned}$$

en utilisant $\ln z \geq \frac{z-1}{z}$, puis $a+k \leq 2a$.

Finalement, on utilise la comparaison somme intégrale, pour une fonction f décroissante et positive, on a

$$\sum_{k=1}^{\infty} f(k) < \int_0^{\infty} f(y) dy \quad (6.11)$$

D'où

$$\frac{1}{\pi_a(0)} < 1 + \int_0^{\infty} \exp \left(-\frac{y^2}{2a} \right) dy = 1 + \sqrt{a} \int_0^{\infty} \exp \left(-\frac{y^2}{2} \right) dy = 1 + \sqrt{a} \sqrt{\frac{\pi}{2}} \quad (6.12)$$

Pour la minoration, on part de (6.10)

$$\begin{aligned}
\frac{1}{\pi_a(0)} &> \sum_{k=0}^a \left(\frac{a-k}{a} \right)^k \\
&= \sum_{k=0}^a \exp \left(k \ln \left(1 - \frac{k}{a} \right) \right) \\
&\geq \sum_{k=0}^{a-\frac{a}{\lambda}} \exp \left(-\lambda \frac{k^2}{a} \right)
\end{aligned}$$

En utilisant la minoration $\ln(1-z) \geq \frac{z}{z-1} \geq -\lambda z$, valable pour tout $z < 1 - \frac{1}{\lambda}$ avec $\lambda > 1$ fixé.

Donc, en utilisant la comparaison somme intégrale

$$\sum_{k=0}^b f(k) > \int_0^{b+1} f(y) dy \quad (6.13)$$

toujours pour une fonction f décroissante, on peut minorer

$$\frac{1}{\pi_a(0)} > \int_0^{a-\frac{a}{\lambda}} \exp \left(-\lambda \frac{y^2}{a} \right) dy = \sqrt{\frac{a}{\lambda}} \int_0^{f_{\lambda}(a)} \exp(-y^2) dy \quad (6.14)$$

avec $f_\lambda(a) = \sqrt{a\lambda} - \sqrt{\frac{a}{\lambda}} = \sqrt{a} \left(\frac{\lambda - 1}{\sqrt{\lambda}} \right)$, qui est croissante en a .

Donc pour tout c tel que $0 < c < \sqrt{\pi/2}$, et pour a assez grand

$$\int_0^{f_\lambda(a)} \exp(-y^2) dy > c \quad (6.15)$$

et finalement avec $C = \frac{c}{\sqrt{\lambda}}$

$$\frac{1}{\pi_a(0)} > C \sqrt{a} \quad (6.16)$$

L'approximation de $\frac{1}{\pi_a(0)}$ découle de (6.12) et (6.16).

Il nous reste maintenant à estimer $\mathbb{E}[T_{a \rightarrow 0}]$. D'après (5.30), on a

$$\mathbb{E}[T_{a \rightarrow 0}] = \sum_{x=1}^a \mathbb{E}[T_{x \rightarrow x-1}] = \sum_{x=1}^a \frac{\pi_a(B_x)}{q_x \pi_a(x)} = 2a \sum_{x=1}^a \frac{1}{a+x} \frac{\pi_a(B_x)}{\pi_a(x)}. \quad (6.17)$$

Donc on peut majorer

$$\begin{aligned} \frac{\pi_a(B_x)}{\pi_a(x)} &< \sum_{k=x}^a \left(\frac{a-x}{a+x} \right)^{(k-x)} \\ &= \sum_{k=0}^{a-x} \left(\frac{a-x}{a+x} \right)^k \\ &< \frac{1}{1 - \frac{a-x}{a+x}} = \frac{a+x}{2x}. \end{aligned} \quad (6.18)$$

D'où

$$\mathbb{E}[T_{a \rightarrow 0}] < a \sum_{x=1}^a \frac{1}{x} = a \left(1 + \sum_{x=2}^a \frac{1}{x} \right) < a (1 + \ln a) \quad (6.19)$$

Pour finir la démonstration de la proposition, on cherche une minoration asymptotique de $\mathbb{E}[T_{a \rightarrow 0}]$. Pour cela on va minorer $\frac{\pi_a(B_x)}{\pi_a(x)}$ par une fonction qui nous donne l'ordre de grandeur en $a \ln a$ de ce temps moyen.

On a

$$\frac{\pi_a(B_x)}{\pi_a(x)} = \sum_{k=x}^a \prod_{i=x+1}^k \frac{a-i+1}{a+i} = \sum_{k=x}^a \prod_{i=1}^{k-x} \frac{a-k+i}{a+x+i}, \quad (6.20)$$

en réordonnant les numérateurs comme dans (6.10).

On peut donc minorer

$$\frac{\pi_a(B_x)}{\pi_a(x)} > \sum_{k=x}^a \left(\frac{a+1-k}{a+x+1} \right)^{(k-x)} = \sum_{k=0}^{a-x} \left(\frac{a+1-x-k}{a+x+1} \right)^k \quad (6.21)$$

On va à nouveau utiliser une approximation de la somme par l'intégrale, et minorer la fonction à intégrer.

Pour tous a et x , on a

$$e^{y \ln\left(\frac{a+1-x-y}{a+x+1}\right)} \geq e^{-K y \frac{\sqrt{a}+x}{a+x}}, \quad (6.22)$$

pour tout $y \leq a+1-x-(a+x+1) e^{-K \frac{\sqrt{a}+x}{a+x}} := lys(a, x)$, où K est une constante fixée. En effet

$$\begin{aligned} & e^{y \ln\left(\frac{a+1-x-y}{a+x+1}\right)} \geq e^{-K y \frac{\sqrt{a}+x}{a+x}} \\ \iff & y \ln\left(\frac{a+1-x-y}{a+x+1}\right) \geq -K y \frac{\sqrt{a}+x}{a+x} \\ \iff & \ln\left(\frac{a+1-x-y}{a+x+1}\right) \geq -K \frac{\sqrt{a}+x}{a+x} \\ \iff & \left(\frac{a+1-x-y}{a+x+1}\right) \geq e^{-K \frac{\sqrt{a}+x}{a+x}} \\ \iff & a+1-x-y \geq (a+x+1) e^{-K \frac{\sqrt{a}+x}{a+x}} \\ \iff & y \leq lys(a, x). \end{aligned} \quad (6.23)$$

Et $lys(a, x) < a - x$, donc

$$\begin{aligned} \frac{\pi_a(B_x)}{\pi_a(x)} & > \int_0^{lys(a, x)} e^{y \ln\left(\frac{a+1-x-y}{a+x+1}\right)} dy \\ & \geq \int_0^{lys(a, x)} e^{-K y \frac{\sqrt{a}+x}{a+x}} dy \\ & = \frac{1}{K} \frac{a+x}{\sqrt{a}+x} \left(1 - e^{-K lys(a, x) \frac{\sqrt{a}+x}{a+x}}\right). \end{aligned} \quad (6.24)$$

Or $L(a, x) = lys(a, x) \frac{\sqrt{a}+x}{a+x}$ satisfait asymptotiquement en a

$$\sup_{x \leq \alpha a} L(a, x) \geq K, \quad (6.25)$$

pour tous α et K tels que $K > \frac{1+\alpha}{\alpha} \ln \frac{1+\alpha}{1-\alpha}$, avec $0 < \alpha < 1$.

En effet, cette fonction est concave et pour de tels α et K , on a

$$\lim_{a \rightarrow \infty} L(a, 0) = K, \text{ et } \lim_{a \rightarrow \infty} L(a, \alpha a) = +\infty. \quad (6.26)$$

Donc il existe $C = C_{\alpha, K}$, tel que pour tout $x \leq \alpha a$, on ait asymptotiquement en a

$$\frac{\pi_a(B_x)}{\pi_a(x)} > C \frac{a+x}{\sqrt{a}+x}, \quad (6.27)$$

d'où, pour a suffisamment grand

$$\begin{aligned}
\mathbb{E}[T_{a \rightarrow 0}] &= 2a \sum_{x=1}^a \frac{1}{a+x} \frac{\pi_a(B_x)}{\pi_a(x)} \\
&> 2C a \sum_{x=1}^{[\alpha a]} \frac{1}{\sqrt{a}+x} \\
&> 2C a \int_1^{\alpha a} \frac{1}{\sqrt{a}+x} dx \\
&= 2C a \int_{\sqrt{a}+1}^{\sqrt{a}+\alpha a} \frac{1}{x} dx \\
&= 2C a \ln \frac{\sqrt{a}+\alpha a}{\sqrt{a}+1} \\
&> 2C a \ln(\alpha \sqrt{a}) = C a (\ln a + \ln \alpha^2),
\end{aligned} \tag{6.28}$$

ce qui termine la démonstration de la proposition.

□

Remarque

Comme choix de α et K , on peut prendre $\alpha = \frac{1}{2}$ et $K = 4$, ce qui donne $C \sim \frac{1}{4}$. La fonction $lys(a, x)$ est elle aussi concave et $\lim_{a \rightarrow \infty} lys(a, 0) = \lim_{a \rightarrow \infty} lys(a, \alpha a) = +\infty$, donc elle est strictement positive pour tout $x \in \llbracket 0, \alpha a \rrbracket$ avec a assez grand, et la constante $C = C_{\alpha, K}$ dans (6.27) est positive.

6.4 Résultats

Comme on l'a précisé en introduction de ce chapitre, notre modèle ne satisfait pas la condition (\mathcal{H}) . Néanmoins, on a

$$(\mathcal{H}') \quad \frac{\pi_a(B_x)}{\pi_a(x)} < K_a, \tag{6.29}$$

où K_a tend vers l'infini lorsque a tend vers l'infini.

Bien qu'on ne puisse appliquer directement les résultats du chapitre 5, les démonstrations fonctionnent toujours.

Théorème 6.2. *Pour le modèle d'Ehrenfest entre 0 et a défini ci-dessus, on a*

- *convergence abrupte pour le temps de tombée $T_{a \rightarrow 0}$ au temps $\mathbb{E}[T_{a \rightarrow 0}] \sim a \ln a$,*
- *métastabilité pour le temps de sortie $T_{0 \rightarrow a}$,*
- *la dernière sortie est la renversée temporelle d'une trajectoire de cutoff.*

Démonstration**Convergence abrupte :**

Comme ici, on a pour tout $x \in \llbracket 0, a \rrbracket$

$$\frac{1}{q_x} < 2 \quad \text{et} \quad \frac{\pi_a(B_x)}{\pi_a(x)} \leq \frac{1}{\pi_a(0)} \sim \sqrt{a} := K_a, \quad (6.30)$$

en se reportant à la démonstration du théorème 5.3, on a

$$\lim_{a \rightarrow \infty} \text{Var} \left(\frac{T_{a \rightarrow 0}}{\mathbb{E}[T_{a \rightarrow 0}]} \right) < \lim_{a \rightarrow \infty} \frac{4K_a^2}{\mathbb{E}[T_{a \rightarrow 0}]} = \lim_{a \rightarrow \infty} \frac{a}{a \ln a} = 0 \quad (6.31)$$

Et d'après la proposition 3.2, il y a convergence abrupte pour la trajectoire de tombée à l'équilibre : $T_{a \rightarrow 0}$ converge abruptement au temps $\mathbb{E}[T_{a \rightarrow 0}]$, qui est de l'ordre de $a \ln a$ (proposition 6.1).

Métastabilité :

D'après le lemme 5.1, on a

$$\mathbb{P}(T_{0 \rightarrow a} < T_{0 \rightarrow 0}) = \frac{p_0}{\sum_{k=0}^{a-1} \prod_{i=1}^k \frac{q_i}{p_i}} = \frac{1}{2 \sum_{k=0}^{a-1} \prod_{i=1}^k \frac{a+i}{a-i}}. \quad (6.32)$$

Avec $\sum_{k=0}^{a-1} \prod_{i=1}^k \frac{a+i}{a-i} > \prod_{i=1}^{a-1} \frac{a+i}{a-i} > (a-1)(2a-1) \sim a^2$.

Et avec (6.19), on obtient pour a grand

$$\mathbb{E}[T_{a \rightarrow 0}] \mathbb{P}(T_{0 \rightarrow a} < T_{0 \rightarrow 0}) < \frac{1}{a} (1 + \ln a) \xrightarrow{a \rightarrow \infty} 0 \quad (6.33)$$

Donc le lemme 5.4 est vérifié. Comme on a la convergence abrupte, le lemme 5.5 l'est aussi, et on peut faire fonctionner la démonstration du théorème 5.6 : le temps de sortie $T_{0 \rightarrow a}$ a un comportement métastable.

Dernière sortie :

On a défini la dernière sortie $\tilde{T}_{0 \rightarrow a}$ dans la section 5.6. Le fait qu'elle soit la renversée d'une trajectoire typique de cutoff repose sur les propositions 5.7, 5.8 et 5.9.

Les propositions 5.7 et 5.8 ne dépendent pas de la condition (\mathcal{H}) , et sont vraies pour toute chaîne de naissance et de mort irréductible sur $\llbracket 0, a \rrbracket$. Donc les variables $\tilde{T}_{a \rightarrow 0}$ et $\tilde{T}_{0 \rightarrow a}$ définies au chapitre précédent (section 5.6) suivent la même loi.

Pour montrer que $\tilde{T}_{a \rightarrow 0}$ admet une convergence abrupte au temps $\mathbb{E}[\tilde{T}_{a \rightarrow 0}]$, on se reporte à la démonstration de la proposition 5.9 :

On a

$$\frac{\bar{\pi}_{a-1}(\bar{B}_x)}{\bar{\pi}_{a-1}(x)} < \frac{\pi_a(B_x)}{\pi_a(x)} \leq K_a, \quad (6.34)$$

et donc $\bar{T}_{a-1 \rightarrow 0}$ admet une convergence abrupte au temps $\mathbb{E} [\bar{T}_{a-1 \rightarrow 0}]$.

En effet, les transitions de \bar{X}_{a-1}^{a-1} et de X_a ne diffèrent que sur les dernières valeurs (a et $a-1$), et celles-ci n'interviennent pas dans les calculs de la proposition 6.1 pour minorer asymptotiquement $\mathbb{E} [T_{a \rightarrow 0}]$ par $C a (\ln a + \ln \alpha^2)$ (6.28).

On a donc $\mathbb{E} [\bar{T}_{a-1 \rightarrow 0}] \gtrsim a \ln a$,

et

$$\lim_{a \rightarrow \infty} \text{Var} \left(\frac{\bar{T}_{a-1 \rightarrow 0}}{\mathbb{E} [\bar{T}_{a-1 \rightarrow 0}]} \right) < \lim_{a \rightarrow \infty} \frac{a}{a \ln a} = 0. \quad (6.35)$$

Pour conclure, on utilise le fait que $\tilde{T}_{a \rightarrow 0} = 1 + \bar{T}_{a-1 \rightarrow 0}$, et on a la convergence abrupte de $\tilde{T}_{a \rightarrow 0}$ au temps $\mathbb{E} [\tilde{T}_{a \rightarrow 0}]$, comme dans la démonstration de la proposition 5.9.

□

Remarque : ce qui nous permet ici de montrer la convergence abrupte, la métastabilité et la dernière sortie, c'est que la borne K_a de la condition (\mathcal{H}') est telle que

$$\lim_{a \rightarrow \infty} \frac{K_a^2}{\mathbb{E} [T_{a \rightarrow 0}]} = 0. \quad (6.36)$$

Dans le chapitre suivant, nous discutons des différences entre les conditions (\mathcal{H}) et (\mathcal{H}') , et énonçons le résultat le plus général avec cette approche (théorème 7.2).

Chapitre 7

Cadre général

On vient de voir que le modèle d'Ehrenfest ne satisfaisait pas l'hypothèse (\mathcal{H}) du chapitre 5, mais qu'on pouvait adapter les démonstrations des théorèmes 5.3 et 5.6 dans ce cas et prouver la convergence abrupte et la métastabilité.

Dans ce chapitre on va discuter des différences entre les chaînes de naissance et de mort du chapitre 5 et le modèle d'Ehrenfest (section 7.1 à 7.3). On introduit ensuite un modèle dit “hybride”, qui conserve certaines des propriétés des chaînes de naissance et de mort du chapitre 5 et du modèle d'Ehrenfest, de telle façon qu'on puisse lui appliquer l'approche développée au chapitre 5 (section 7.4).

Cette discussion nous conduit à regrouper nos résultats, et on présente le cadre le plus général possible dans lequel on peut utiliser notre approche (section 7.5).

Dans la dernière section (section 7.6), on montre que pour le modèle d'Ehrenfest et le modèle “hybride”, on a des conditions plus faibles pour montrer la convergence abrupte et la métastabilité, si on regarde les temps d'atteinte et de sortie d'une région $\llbracket 0, b_a \rrbracket$ (et non plus seulement d'un état), “caractéristique” de l'équilibre.

7.1 L'hypothèse (\mathcal{H})

Les chaînes de naissance et de mort qu'on considère au chapitre 5 sont des marches aléatoires sur $\llbracket 0, a \rrbracket$, de transitions p_i (vers la droite) et q_i (vers la gauche), uniformément irréductibles, i.e. telles que uniformément en a les transitions ne s'annulent pas (sauf éventuellement p_a et q_0).

La mesure invariante d'une telle chaîne est

$$\pi_a(x) = \prod_{i=1}^x \frac{p_{i-1}}{q_i} \pi_a(0), \quad (7.1)$$

avec

$$\pi_a(0) = 1 / \sum_{x=0}^a \prod_{i=1}^x \frac{p_{i-1}}{q_i}. \quad (7.2)$$

On suppose que nos chaînes sont telles que la quantité

$$A_x := \frac{\pi_a(B_x)}{\pi_a(x)}, \quad (7.3)$$

où $B_x = \llbracket x, a \rrbracket$, est uniformément bornée en a .

C'est à dire qu'il existe K tel que pour tout x et pour tout a on ait

$$(\mathcal{H}) \quad A_x := \frac{\pi_a(B_x)}{\pi_a(x)} < K. \quad (7.4)$$

On a vu dans ce même chapitre que

$$A_x := \frac{\pi_a(B_x)}{\pi_a(x)} = \sum_{k=x}^a \prod_{i=x+1}^k \frac{p_{i-1}}{q_i}. \quad (7.5)$$

Supposer que cette quantité est bornée pour tout x implique que $A_0 = 1/\pi_a(0)$ est borné. Or on a vu au chapitre 2 (2.14), qu'une condition nécessaire et suffisante pour qu'une chaîne de naissance et de mort sur \mathbb{N} soit récurrente positive est

$$\sum_{x=1}^{\infty} \prod_{i=1}^x \frac{p_{i-1}}{q_i} < \infty. \quad (7.6)$$

La condition (\mathcal{H}) peut donc s'interpréter comme une sorte de "récence positive à la limite". Si de plus les transitions p_i et q_i ne dépendent pas de a , on peut parler de récence positive du modèle limite, car dans ce cas la quantité dans (7.6) est la limite de A_0 lorsque a tend vers l'infini.

Ces modèles, qu'on appellera modèles à transitions "indépendantes" de a par référence au modèle (p, q) présenté au chapitre 5 (section 5.2), nous fournissent l'exemple type de modèles qui satisfont l'hypothèse (\mathcal{H}) .

7.2 Modèles à transitions "indépendantes" de a

On rappelle le modèle (p, q) random walk de la section 5.2, où les transitions sont constantes pour tout x

$$p_x = p, \text{ et } q_x = q \quad \forall x = 0, \dots, a, \quad (7.7)$$

avec $p + q = 1$, et où pour avoir la dérive qui pousse vers zéro, on suppose que $q > p$.

Ce modèle satisfait l'hypothèse (\mathcal{H}) avec $K = q/(q - p)$, et on a vu dans la section 5.2, comment elle se traduisait en termes de mesure invariante et de profil énergétique.

Rappel :

– ce qui fait marcher la démonstration du théorème 5.3, c’est qu’on a

$$\text{Var}\left(\frac{T_{a \rightarrow 0}}{\mathbb{E}[T_{a \rightarrow 0}]}\right) \leq \frac{2K_q K^2}{\mathbb{E}[T_{a \rightarrow 0}]} \text{ et donc } \lim_{a \rightarrow \infty} \text{Var}\left(\frac{T_{a \rightarrow 0}}{\mathbb{E}[T_{a \rightarrow 0}]}\right) = 0,$$

car $\mathbb{E}[T_{a \rightarrow 0}] \geq a \rightarrow \infty$.

Donc les conditions de la proposition 3.2 sont vérifiées et $\frac{T_{a \rightarrow 0}}{\mathbb{E}[T_{a \rightarrow 0}]} \xrightarrow[a \rightarrow \infty]{\mathbb{P}} 1$.

– dans la démonstration du théorème 5.6, on montre d’abord que

$$\lim_{a \rightarrow \infty} \mathbb{P}(T_{0 \rightarrow a} < T_{0 \rightarrow 0}) \mathbb{E}[T_{a \rightarrow 0}] = 0.$$

Avec la convergence abrupte en plus, on montre que

$$\lim_{a \rightarrow \infty} \mathbb{P}(T_{0 \rightarrow a} \leq T_{a \rightarrow 0}) = 0,$$

et donc en utilisant le couplage et le nombre N_a de retours en zéro avant d’atteindre a pour la première fois, on a :

$$\frac{\mathbb{E}[T_{a \rightarrow 0}]}{\mathbb{E}[T_{0 \rightarrow a}]} \sim \frac{1}{\mathbb{E}[N_a]} = \mathbb{P}(T_{0 \rightarrow a} \leq T_{a \rightarrow 0}) \xrightarrow[a \rightarrow \infty]{} 0.$$

Il y a deux échelles de temps, et donc $\frac{T_{0 \rightarrow a}}{\mathbb{E}[T_{0 \rightarrow a}]} \xrightarrow[a \rightarrow \infty]{\mathcal{L}} \exp(1)$.

Ici, comme $q > 1/2$, on peut prendre $K_q = 2$ pour la borne supérieure des $(1/q_x)$ (voir la démonstration du théorème 5.3).

On a dans ce cas des estimations précises pour les temps moyen de tombée et de sortie :

$$\mathbb{E}[T_{a \rightarrow 0}] \sim \frac{a}{q-p}, \text{ et } \mathbb{E}[T_{0 \rightarrow a}] \sim \left(\frac{q}{p}\right)^a, \quad (7.8)$$

et pour la probabilité partant de zéro d’atteindre a avant de revenir en zéro (qui intervient dans la démonstration du théorème 5.6) :

$$\mathbb{P}(T_{0 \rightarrow a} < T_{0 \rightarrow 0}) \sim \left(\frac{p}{q}\right)^{a-1}. \quad (7.9)$$

On peut généraliser ce modèle en considérant un modèle à transitions uniformément bornées en a :

$$q = \min_i q_i > \max_i p_i = p. \quad (7.10)$$

Dans ce cas, on a les mêmes bornes K et K_q que précédemment et les estimations

$$\mathbb{E}[T_{a \rightarrow 0}] \lesssim K_q K a, \quad \mathbb{E}[T_{0 \rightarrow a}] \gtrsim \left(\frac{q}{p}\right)^a, \quad \text{et } \mathbb{P}(T_{0 \rightarrow a} < T_{0 \rightarrow 0}) \lesssim \left(\frac{p}{q}\right)^{a-1}. \quad (7.11)$$

On peut encore généraliser le modèle, en prenant $q = \min_i q_i > \max_i p_i = p$, pour i en dehors d'un ensemble fini (on autorise donc quelques irrégularités dans le profil énergétique).

Remarque : pour ces modèles, on a pas besoin d'avoir une approximation précise de $\mathbb{E}[T_{a \rightarrow 0}]$, pour montrer la convergence abrupte, il nous suffit d'utiliser $\mathbb{E}[T_{a \rightarrow 0}] \geq a \rightarrow \infty$.

7.3 Cas Ehrenfest

La différence du modèle d'Ehrenfest par rapport aux modèles précédents, c'est que ses transitions dépendent de a :

$$p_i = \frac{a-i}{2a}, \quad \text{et} \quad q_i = \frac{a+i}{2a}. \quad (7.12)$$

Dans ce cas, on a $\min_i q_i = \max_i p_i = \frac{1}{2}$, et

$$\lim_{a \rightarrow \infty} p_i = \lim_{a \rightarrow \infty} q_i = \frac{1}{2}, \quad (7.13)$$

pour tout $i \leq b_a$, avec $b_a/a \xrightarrow{a \rightarrow \infty} 0$.

C'est la raison pour laquelle A_x n'est pas uniformément borné en a et que la condition (\mathcal{H}) n'est pas satisfaite.

Néanmoins, on a

$$(\mathcal{H}') \quad \frac{\pi_a(B_x)}{\pi_a(x)} < K_a, \quad (7.14)$$

avec $K_a \xrightarrow{a \rightarrow \infty} \infty$.

Et le théorème 5.3 reste vrai, car $\frac{K_q K_a^2}{\mathbb{E}[T_{a \rightarrow 0}]} \xrightarrow{a \rightarrow \infty} 0$.

En effet, on a $1/q_x \leq K_q = 2$, et les approximations $K_a \sim \sqrt{a}$, et $\mathbb{E}[T_{a \rightarrow 0}] \sim a \log a$.

Avec la condition (\mathcal{H}') , on garde une certaine forme de décroissance exponentielle de la mesure invariante, puisqu'elle implique que

$$\pi_a(x) \leq \left(1 - \frac{1}{K_a}\right)^x = e^{-\alpha_a x}, \quad (7.15)$$

avec $\alpha_a = -\log\left(1 - \frac{1}{K_a}\right)$, qui est strictement positif lorsque a est fixé.

Et on a donc une minoration du profil énergétique

$$H(x) \geq \log \pi_a(0) + \alpha_a x. \quad (7.16)$$

La différence avec les modèles qui satisfont (\mathcal{H}) , est qu'ici on a $\alpha_a \xrightarrow{a \rightarrow \infty} 0$.

En analogie avec les chaînes de naissance et de mort sur \mathbb{N} vues au chapitre 2, on peut interpréter la condition (\mathcal{H}') comme une sorte de “réurrence nulle à la limite” (cf (2.16)).

Un exemple de chaîne récurrente nulle est la marche aléatoire symétrique sur \mathbb{N} , dont les transitions sont égales et constantes : $p = q = 1/2$. Mais si on prend la restriction de cette chaîne à $\llbracket 0, a \rrbracket$, nos résultats ne s’appliquent pas.

En effet, dans ce cas la mesure invariante est la mesure uniforme sur $\llbracket 0, a \rrbracket$

$$\pi_a(x) = \pi_a(0) = \frac{1}{a+1}, \quad \forall x \in \llbracket 0, a \rrbracket, \quad (7.17)$$

et on a donc $A_x = a - x + 1 \leq K_a = a + 1$ pour tout x .

Dans ce cas le temps moyen pour aller de a à zéro est

$$\mathbb{E}[T_{a \rightarrow 0}] = \sum_{x=0}^a \frac{A_x}{q_x} = 2 \sum_{x=0}^a x = a(a+1), \quad (7.18)$$

et on ne peut pas appliquer la démonstration du théorème 5.3, car $\frac{K_a^2}{\mathbb{E}[T_{a \rightarrow 0}]} \xrightarrow{a \rightarrow \infty} 1$.

On a pas non plus la métastabilité, puisque pour des raisons de symétrie évidentes on a $\mathbb{E}[T_{0 \rightarrow a}] = \mathbb{E}[T_{a \rightarrow 0}]$.

Ce qui est normal puisque comme on l’a déjà vu, notre approche du phénomène n’est pas compatible avec la mesure uniforme, qui donne un profil énergétique plat.

Mais on a vu (7.13) que les transitions en un site i du modèle d’Ehrenfest tendaient vers $1/2$, pour tout $i \leq b_a$ où $b_a/a \rightarrow 0$.

De plus, en regardant les calculs que l’on a fait au chapitre précédent pour estimer K_a et $\mathbb{E}[T_{a \rightarrow 0}]$ pour le modèle d’Ehrenfest, on s’aperçoit que les états proches de zéro contribuent peu dans les estimations, et que la contribution principale provient des états situés à partir d’un certain rang.

Lorsqu’on minore $\mathbb{E}[T_{a \rightarrow 0}]$, on peut supposer que $x \geq \alpha a^{1-\varepsilon}$, et on conserve le même ordre de grandeur.

En effet, d’après (6.27) on a

$$\frac{\pi_a(B_x)}{\pi_a(x)} > C \frac{a+x}{\sqrt{a}+x}, \quad (7.19)$$

pour tout $x \leq \alpha a$, où C est une constante qui dépend de $0 < \alpha < 1$ (rappel : on peut prendre $\alpha = 1/2$ et $C = 1/4$). Et on obtient la minoration asymptotique (6.28)

$$\mathbb{E}[T_{a \rightarrow 0}] > 2C a \sum_{x=1}^{\lfloor \alpha a \rfloor} \frac{1}{\sqrt{a}+x} \gtrsim 2C a \ln \sqrt{a} = C a \ln a, \quad (7.20)$$

en passant par l'approximation somme intégrale

$$\sum_{x=1}^{\lfloor \alpha a \rfloor} \frac{1}{\sqrt{a} + x} \gtrsim \ln \frac{\sqrt{a} + \alpha a}{\sqrt{a} + 1} \sim \ln(\alpha \sqrt{a}) . \quad (7.21)$$

Or on peut couper cette somme, en gardant les $x \geq \alpha a^{1-\varepsilon}$, ce qui donne

$$\sum_{x=\lceil \alpha a^{1-\varepsilon} \rceil}^{\lfloor \alpha a \rfloor} \frac{1}{\sqrt{a} + x} \gtrsim \ln \frac{\sqrt{a} + \alpha a}{\sqrt{a} + \alpha a^{1-\varepsilon}} \sim \ln a^\varepsilon , \quad (7.22)$$

et finalement, on a le même ordre de grandeur dans la minoration

$$\mathbb{E}[T_{a \rightarrow 0}] \gtrsim C \varepsilon a \ln a . \quad (7.23)$$

Remarque : on peut garder uniquement les $x \geq a^{1-\varepsilon}$, et on a $\mathbb{E}[T_{a \rightarrow 0}] \gtrsim C \varepsilon a \ln \alpha a$.

De même dans la majoration de K_a , on peut prendre les $x \geq a^{\frac{1}{2}-\varepsilon}$, et on garde l'ordre de grandeur en \sqrt{a} .

En effet, pour majorer $1/\pi_a(0)$ (voir 6.12), on utilise l'intégrale

$$\int_0^\infty \exp\left(-\frac{y^2}{2a}\right) dy = \sqrt{a} \int_0^\infty \exp\left(-\frac{y^2}{2}\right) dy . \quad (7.24)$$

Or si on coupe en $a^{\frac{1}{2}-\varepsilon}$, la première partie de l'intégrale est un $o(\sqrt{a})$:

$$\int_0^{a^{\frac{1}{2}-\varepsilon}} \exp\left(-\frac{y^2}{2a}\right) dy = \sqrt{a} \int_0^{a^{-\varepsilon}} \exp\left(-\frac{y^2}{2}\right) dy = o(\sqrt{a}) . \quad (7.25)$$

Et c'est la seconde partie de l'intégrale qui est d'ordre \sqrt{a} .

Les états qui contribuent le plus dans cette estimation sont donc les $x \geq a^{\frac{1}{2}-\varepsilon}$.

7.4 Modèle hybride

Les remarques que l'on a faites ci dessus nous amènent à considérer un modèle "hybride" entre le modèle d'Ehrenfest et les modèles à transitions indépendantes de a .

Puisque les états proches de zéros contribuent peu dans le comportement du modèle d'Ehrenfest, on veut garder de ce modèle le fait que les transitions en un site i tendent vers $1/2$, pour tout $i \leq b_a$ où $b_a/a \rightarrow 0$. Et on veut qu'à partir de ce rang b_a , notre modèle satisfasse la condition (\mathcal{H}) .

On suppose donc que $p_i = q_i = 1/2$ pour tout $i < b_a$, et qu'à partir de b_a , on a uniformément en a :

$$q = \min_{i \geq b_a} q_i > \max_{i \geq b_a} p_i = p \quad (7.26)$$

La question qu'on se pose est : jusqu'où peut on prendre $p_i = q_i = 1/2$ en conservant la convergence abrupte et la métastabilité ?

Quelques calculs nous permettent d'obtenir un encadrement de $A_x = \pi_a(B_x)/\pi_a(x)$.

Soit $K = q/(q - p)$, et $\bar{K} = \bar{q}/(\bar{q} - \bar{p})$, où $\bar{q} = \max_{i \geq b_a} q_i$ et $\bar{p} = \min_{i \geq b_a} p_i$, qui est strictement positif car on a supposé le modèle uniformément irréductible.

On a

$$\bar{K} < A_x < K, \quad \forall x \geq b_a, \quad (7.27)$$

et

$$b_a - x + \bar{K} < A_x < b_a - x + K, \quad \forall x < b_a. \quad (7.28)$$

La condition (\mathcal{H}) est satisfaite à partir de b_a , puisque pour tout $x \geq b_a$: $A_x < K$. Et pour $x < b_a$, (\mathcal{H}') est satisfaite avec $K_a = b_a + K$.

On obtient ainsi une condition suffisante sur b_a pour avoir la convergence abrupte et la métastabilité pour le modèle hybride.

Proposition 7.1. *Si b_a est tel que*

$$\frac{b_a^2}{a} \xrightarrow{a \rightarrow \infty} 0, \quad (7.29)$$

alors $T_{a \rightarrow 0}$ converge abruptement au temps $\mathbb{E}[T_{a \rightarrow 0}]$, le temps de sortie $T_{0 \rightarrow a}$ a un comportement métastable, et la dernière sortie $\tilde{T}_{a \rightarrow 0}$ est la renversée d'une trajectoire de cutoff.

Démonstration :

En effet, on a la minoration $\mathbb{E}[T_{a \rightarrow 0}] \gtrsim b_a^2 + a\bar{K}$, et donc la convergence abrupte car

$$\text{Var}\left(\frac{T_{a \rightarrow 0}}{\mathbb{E}[T_{a \rightarrow 0}]}\right) \leq \frac{K_a^2}{\mathbb{E}[T_{a \rightarrow 0}]} \leq \frac{(b_a + K)^2}{b_a^2 + a\bar{K}} \xrightarrow{a \rightarrow \infty} 0, \quad (7.30)$$

si $b_a^2/a \xrightarrow{a \rightarrow \infty} 0$.

Pour la métastabilité, on a $\mathbb{E}[T_{a \rightarrow 0}] \lesssim b_a^2 + 2aK$, et en se basant sur le lemme 5.1 et la démonstration de (5.29), on a $\mathbb{P}(T_{0 \rightarrow a} < T_{0 \rightarrow 0}) < (p/q)^{a-b_a}$, en minorant la somme $\sum_{x=0}^{a-1} \prod (q_i/p_i)$ par la somme des termes supérieurs à b_a .

Et donc $\mathbb{E}[T_{a \rightarrow 0}] \mathbb{P}(T_{0 \rightarrow a} < T_{0 \rightarrow 0}) \xrightarrow{a \rightarrow \infty} 0$.

Pour montrer que la dernière sortie est la renversée d'une trajectoire de cutoff, il suffit de montrer que $\bar{T}_{a-1 \rightarrow 0}$ admet une convergence abrupte (voir la démonstration de la proposition 5.9).

Or $\mathbb{E}[\bar{T}_{a-1 \rightarrow 0}] \geq a - 1$, et $\bar{\pi}_{a-1}(\bar{B}_x)/\bar{\pi}_{a-1}(x) < K_a = b_a + K$, d'où

$$\lim_{a \rightarrow \infty} \text{Var}\left(\frac{\bar{T}_{a-1 \rightarrow 0}}{\mathbb{E}[\bar{T}_{a-1 \rightarrow 0}]}\right) < \lim_{a \rightarrow \infty} \frac{(b_a + K)^2}{a - 1} = \lim_{a \rightarrow \infty} \frac{b_a^2}{a} = 0. \quad (7.31)$$

□

On prend donc $b_a \sim a^{1/2-\varepsilon}$, et on retrouve les valeurs qui contribuent le plus dans les estimations pour le modèle d'Ehrenfest.

7.5 Généralisation

On peut donc généraliser les résultats du chapitre 5.

En effet, pour le modèle hybride comme pour le modèle d'Ehrenfest, on voit que lorsque la condition (\mathcal{H}') est satisfaite avec une condition sur la vitesse relative de K_a et $\mathbb{E}[T_{a \rightarrow 0}]$ ($K_a^2/\mathbb{E}[T_{a \rightarrow 0}] \xrightarrow{a \rightarrow \infty} 0$), on peut montrer la convergence abrupte et la métastabilité.

On regroupe ces deux conditions :
s'il existe $K_a \xrightarrow{a \rightarrow \infty} \infty$ tel que $\forall x \in \llbracket 0, a \rrbracket$, on a

$$(\mathcal{H}'') \quad \frac{\pi_a(B_x)}{\pi_a(x)} < K_a, \quad \text{avec} \quad \frac{K_a^2}{\mathbb{E}[T_{a \rightarrow 0}]} \xrightarrow{a \rightarrow \infty} 0, \quad (7.32)$$

Et on a le théorème suivant qui synthétise nos résultats

Théorème 7.2. *Soit $X_a(t)$ une famille de chaînes de naissance et de mort uniformément irréductible sur $\llbracket 0, a \rrbracket$.*

Si la condition (\mathcal{H}'') est satisfaite, alors le temps de tombée à l'équilibre $T_{a \rightarrow 0}$ converge abruptement au temps $\mathbb{E}[T_{a \rightarrow 0}]$, et le temps de sortie $T_{0 \rightarrow a}$ a un comportement métastable.

Démonstration :

On donne ici seulement les arguments principaux de la démonstration. Pour les détails, on se reportera au rappel de la section 7.2 et aux démonstrations des théorèmes 5.3 et 5.6.

Pour montrer la convergence abrupte de $T_{a \rightarrow 0}$ au temps $\mathbb{E}[T_{a \rightarrow 0}]$ on a

$$\text{Var}\left(\frac{T_{a \rightarrow 0}}{\mathbb{E}[T_{a \rightarrow 0}]}\right) < \frac{K_q K_a^2}{\mathbb{E}[T_{a \rightarrow 0}]} \xrightarrow{a \rightarrow \infty} 0, \quad (7.33)$$

et la proposition 3.2 s'applique.

Pour montrer la métastabilité, on doit montrer que

$$\lim_{a \rightarrow \infty} \mathbb{P}(T_{0 \rightarrow a} < T_{0 \rightarrow 0}) \mathbb{E}[T_{a \rightarrow 0}] = 0 ,$$

et ensuite on conclut en utilisant la convergence abrupte de $T_{a \rightarrow 0}$.

On a

$$\mathbb{E}[T_{a \rightarrow 0}] \leq a K_a K_q . \quad (7.34)$$

D'après la formule (5.29)

$$\mathbb{P}(T_{0 \rightarrow a} < T_{0 \rightarrow 0}) = 1 / \sum_{k=1}^a \frac{\pi_a(0)}{q_k \pi_a(k)} . \quad (7.35)$$

Or par hypothèse $1/\pi_a(0) < K_a$, et d'après (7.15) $\pi_a(x) \leq \left(1 - \frac{1}{K_a}\right)^x$, $\forall x$.

D'où

$$\sum_{k=1}^a \frac{1}{q_k \pi_a(k)} \geq \sum_{k=1}^a \frac{1}{\pi_a(k)} > \frac{1}{\pi_a(a)} \geq \left(1 - \frac{1}{K_a}\right)^{-a} . \quad (7.36)$$

Or comme $K_a \rightarrow \infty$, on a

$$\left(1 - \frac{1}{K_a}\right)^a = e^{a \ln(1-1/K_a)} \sim e^{-a/K_a} . \quad (7.37)$$

Et $a/K_a \xrightarrow{a \rightarrow \infty} \infty$, car

$$\frac{K_a}{a} = \frac{K_q K_a^2}{a K_a K_q} \leq \frac{K_a^2}{\mathbb{E}[T_{a \rightarrow 0}]} \xrightarrow{a \rightarrow \infty} 0 . \quad (7.38)$$

Donc

$$\lim_{a \rightarrow \infty} \mathbb{P}(T_{0 \rightarrow a} < T_{0 \rightarrow 0}) \mathbb{E}[T_{a \rightarrow 0}] \leq \lim_{a \rightarrow \infty} a K_a^2 K_q e^{-a/K_a} = 0 . \quad (7.39)$$

□

7.6 Temps d'atteinte et de sortie d'une région

Comme on l'a remarqué avec le modèle d'Ehrenfest, la contribution des états proches de zéro est négligeable dans les calculs, et que $\lim_{a \rightarrow \infty} p_i = \lim_{a \rightarrow \infty} q_i = \frac{1}{2}$, pour tout $i \leq b_a$, avec $b_a/a \xrightarrow{a \rightarrow \infty} 0$. En fait, on a la convergence abrupte et la métastabilité entre b_a et a , pour un tel b_a .

Proposition 7.3. *Dans le modèle d'Ehrenfest, si b_a est tel que $b_a/a \xrightarrow{a \rightarrow \infty} 0$, alors il y a convergence abrupte pour $T_{a \rightarrow b_a}$ au temps $\mathbb{E}[T_{a \rightarrow b_a}]$, et $T_{b_a \rightarrow a}$ a un comportement métastable.*

Démonstration :

En se reportant aux calculs de la section 6.3, on a

$$\frac{\pi_a(B_x)}{\pi_a(x)} > C \frac{a+x}{\sqrt{a}+x}, \quad (7.40)$$

pour tout $x \leq \alpha a$ (formule (6.27)), et on peut minorer

$$\begin{aligned} \mathbb{E}[T_{a \rightarrow b_a}] &\geq 2 C a \sum_{x=b_a+1}^{\lfloor \alpha a \rfloor} \frac{1}{\sqrt{a}+x} \\ &\sim 2 C a \ln \frac{\sqrt{a} + \alpha a}{\sqrt{a} + b_a + 1} \\ &= 2 C a \ln \frac{(1/\sqrt{a}) + \alpha}{(1/\sqrt{a}) + (b_a/a) + (1/a)} \\ &\sim 2 C a \ln \alpha (a/b_a), \end{aligned} \quad (7.41)$$

où on peut prendre les constantes α et C telles que $\alpha < 1$ et $C < 1/2$.

Et on a la convergence abrupte de $T_{a \rightarrow b_a}$, car

$$\text{Var}\left(\frac{T_{a \rightarrow b_a}}{\mathbb{E}[T_{a \rightarrow b_a}]}\right) \leq \frac{K_a^2}{2 C a \ln \alpha (a/b_a)} \sim \frac{1}{2 C \ln \alpha (a/b_a)} \xrightarrow{a \rightarrow \infty} 0, \quad (7.42)$$

puisque $(a/b_a) \xrightarrow{a \rightarrow \infty} \infty$.

Pour la métastabilité, comme d'après (6.18)

$$\frac{\pi_a(B_x)}{\pi_a(x)} \leq \frac{a+x}{2x}, \quad (7.43)$$

pour tout x . On a

$$\mathbb{E}[T_{a \rightarrow b_a}] = \sum_{x=b_a+1}^a \frac{\pi_a(B_x)}{q_x \pi_a(x)} \leq a \sum_{x=b_a+1}^a \frac{1}{x} \sim a \ln(a/b_a) \quad (7.44)$$

Pour la majoration de $\mathbb{P}(T_{b_a \rightarrow a} < T_{b_a \rightarrow b_a})$, on a une formule analogue à celle du lemme 5.1 :

$$\mathbb{P}(T_{b_a \rightarrow a} < T_{b_a \rightarrow b_a}) = p_{b_a} \mathbb{P}(T_{b_{a+1} \rightarrow a} < T_{b_{a+1} \rightarrow b_a}) = p_{b_a} \frac{1}{\sum_{k=b_a}^{a-1} \prod_{b_{a+1}}^k \frac{q_i}{p_i}}. \quad (7.45)$$

Or on peut minorer $\sum_{k=b_a}^{a-1} \prod_{b_{a+1}}^k \frac{q_i}{p_i}$, en prenant les derniers termes de la somme comme dans la démonstration de la métastabilité pour le modèle d'Ehrenfest

$$\sum_{k=b_a}^{a-1} \prod_{b_{a+1}}^k \frac{q_i}{p_i} = \sum_{k=b_a}^{a-1} \prod_{b_{a+1}}^k \frac{a+i}{a-i}, > \prod_{b_{a+1}}^{a-1} \frac{a+i}{a-i} > (a-1)(2a-1) \sim a^2. \quad (7.46)$$

Donc

$$\mathbb{P}(T_{b_a \rightarrow a} < T_{b_a \rightarrow b_a}) < \frac{a - b_a}{2a} \frac{1}{a^2}, \quad (7.47)$$

et

$$\mathbb{E}[T_{a \rightarrow b_a}] \mathbb{P}(T_{b_a \rightarrow a} < T_{b_a \rightarrow b_a}) < \frac{a - b_a}{2a^2} \ln(a/b_a) \xrightarrow{a \rightarrow \infty} 0, \quad (7.48)$$

puisque $(a - b_a)/2a^2$ tend vers zéro.

□

On observe le même comportement avec le modèle hybride :

Proposition 7.4. *Dans le modèle hybride, si b_a est tel que $b_a/a \xrightarrow{a \rightarrow \infty} 0$, alors il y a convergence abrupte pour $T_{a \rightarrow b_a}$ au temps $\mathbb{E}[T_{a \rightarrow b_a}]$, et $T_{b_a \rightarrow a}$ a un comportement métastable.*

Démonstration :

Il y a convergence abrupte pour $T_{a \rightarrow b_a}$ car $\mathbb{E}[T_{a \rightarrow b_a}] \geq a - b_a$, et comme $A_x < K, \forall x \geq b_a$, on a

$$\text{Var}\left(\frac{T_{a \rightarrow b_a}}{\mathbb{E}[T_{a \rightarrow b_a}]}\right) \leq \frac{4K^2}{a - b_a} \xrightarrow{a \rightarrow \infty} 0. \quad (7.49)$$

Et pour la métastabilité de $T_{b_a \rightarrow a}$, on utilise une formule analogue à (5.29) pour montrer que $\mathbb{P}(T_{b_a \rightarrow a} < T_{b_a \rightarrow b_a}) < (1/b_a)e^{-\alpha a}$, et donc que $\mathbb{E}[T_{a \rightarrow b_a}] \mathbb{P}(T_{b_a \rightarrow a} < T_{b_a \rightarrow b_a}) \xrightarrow{a \rightarrow \infty} 0$.

□

On voit que notre problématique peut se ramener au temps d'atteinte et de sortie d'une région (ici $\llbracket 0, b_a \rrbracket$, avec $b_a/a \xrightarrow{a \rightarrow \infty} 0$), dès lors qu'il y a une dérive qui pousse la chaîne vers cette région, et qu'on est en présence de deux échelles de temps.

Notre approche de la convergence abrupte et de la métastabilité, qui consiste à regarder les temps d'atteinte d'états "caractéristiques" de l'équilibre (au sens où on l'entend en physique), peut s'étendre à une région qui concentre la mesure d'équilibre, et est donc plus caractéristique de cet "état" (au sens probabiliste) d'équilibre.

Chapitre 8

Conclusion et perspectives

Dans ce travail, nous avons utilisé une approche liée aux temps d’atteinte de certains états pour étudier les phénomènes de convergence abrupte et de métastabilité. Cette approche cadre tout à fait avec la métastabilité et l’approche “pathwise” développée par Cassandro, Galves, Olivieri et Vares pour la décrire [12]. Mais on a vu qu’elle était différente de celle utilisée pour l’étude du phénomène de cutoff par Diaconis, Aldous, Saloff-Coste et Ycart, basée sur les mesures de probabilité. Elle y est néanmoins équivalente lorsqu’on a affaire à des modèles d’absorption, et on a vu dans les travaux de Martínez et Ycart qu’on pouvait relier les deux (section 4. “Equivalence of cutoffs” de [38]). On a vu de plus, à la fin du chapitre précédent, qu’on pouvait étendre notre approche au temps d’atteinte d’une région. La taille de cette région étant de l’ordre de la variance de la mesure invariante pour le modèle d’Ehrenfest notamment, cela nous donne bon espoir de relier la convergence abrupte avec cette variance : on aurait convergence abrupte pour atteindre une zone de la taille de la variance dès lors qu’on part suffisamment loin de cette zone sous une condition du type $\text{Var}(\pi_a)/\mathbb{E}[T_a] \xrightarrow{a \rightarrow \infty} 0$.

Nos modèles sont des marches aléatoires unidimensionnelles avec dérive, c’est-à-dire dont le profil énergétique associé à la mesure invariante présente une pente ou un puits d’attraction. De nombreux modèles pour lesquels la convergence abrupte a été démontrée ne rentrent pas dans ce cadre, mais on a vu avec la marche aléatoire sur l’hypercube qu’on pouvait la ramener au modèle d’Ehrenfest. Pour les propriétés qui nous intéressent, on peut souvent (toujours ?) ramener l’étude du modèle dans un cadre unidimensionnel. Dès qu’on peut se ramener à un profil énergétique, sa connaissance nous suffit pour démontrer la convergence abrupte et la métastabilité. Notre approche doit donc pouvoir se généraliser à des profils plus complexes, du type de ceux étudiés par Olivieri et Scoppola [42, 43] (voir section 4.6). Leurs travaux prennent en compte des profils contenant plusieurs minimas locaux sur des espaces très généraux, et ils montrent la métastabilité pour sortir de domaines contenant plusieurs attracteurs appelés cycles. Mais la différence avec notre approche, est que dans leur cas, le paramètre asymptotique est l’inverse de la température β , et lorsqu’il tend vers l’infini, ce n’est plus l’état initial ou à atteindre qui tend vers l’infini, mais les pentes du profil qui tendent à être verticales (les transitions tendent vers

zéro ou un, suivant qu'on entre ou sort d'un puits d'attraction).

Nous devrions aussi pouvoir étendre notre approche aux chaînes de naissance et de mort sur un arbre pour lesquelles Martínez et Ycart ont démontré le cutoff. Nous aimerions également généraliser cette approche à des familles de chaînes de Markov quelconques.

Notre problématique de départ était de montrer la métastabilité à partir de la seule hypothèse de cutoff. Cette question reste ouverte, mais notre travail y répond en partie. En effet, bien que nous nous appuyons sur la dérive pour montrer la métastabilité, nous utilisons fortement la convergence abrupte au travers du couplage (5.54). Cela nous amène à conjecturer que la convergence abrupte vers l'équilibre implique la métastabilité de la sortie de l'équilibre par renversée temporelle. Il serait sûrement utile pour cela d'avoir des conditions nécessaires et suffisantes pour la convergence abrupte et la métastabilité.

L'autre question ouverte est de savoir s'il existe des systèmes pour lesquels on a la métastabilité, mais pas la convergence abrupte. Or la métastabilité semble requérir des hypothèses plus faibles, en effet elle repose sur l'existence d'échelles de temps différentes, alors que la convergence abrupte nécessite une connaissance plus fine de la façon dont la convergence a lieu. Par exemple dans les modèles de Olivieri et Scoppola, lorsqu'il y a plusieurs puits imbriqués les uns dans les autres : on doit avoir cutoff et métastabilité pour entrer et sortir de chacun des puits, mais si on considère le plus grand puits, il y a métastabilité pour en sortir, mais sûrement pas cutoff pour tomber au fond, puisqu'il faut entrer et sortir de plusieurs puits avant de l'atteindre.

Une réponse pourrait être fournie par D. Aldous, qui a proposé un contre-exemple à la conjecture de Peres en variation totale : il correspondrait à un cas avec métastabilité mais sans cutoff.

Une autre direction serait de relier notre travail avec la conjecture de Peres. Celle-ci postule que si le produit du temps de mélange par le trou spectral tend vers l'infini, alors il y a cutoff. Aldous propose un contre exemple quand la distance utilisée est la variation totale, mais elle a été prouvée pour les chaînes de naissance et de mort, si la notion de cutoff utilise la distance de séparation, par Diaconis et Saloff-Coste dans "Separation cut-offs for birth and death chains" [21].

Chapitre 9

Appendice

9.1 Temps d'atteinte pour les marches aléatoires

On considère une marche aléatoire $X_N^x(t)$ à temps discret sur $\{0, \dots, N\}$, $N \in \mathbb{N}$, d'état initial x , et dont les transitions

$$\begin{aligned} p_i &= \mathbb{P}(X(t+1) = i+1 \mid X(t) = i) , \\ q_i &= \mathbb{P}(X(t+1) = i-1 \mid X(t) = i) , \\ \text{et } r_i &= \mathbb{P}(X(t+1) = i \mid X(t) = i) , \end{aligned}$$

sont telles que : $p_i + r_i + q_i = 1$, $p_N = 0$, $q_0 = 0$, et $p_i, q_i > 0$, $i = 1, \dots, N-1$.

On s'intéresse au temps moyen pour aller d'un état x à y . On définit pour cela la variable aléatoire suivante :

$$T_{x \rightarrow y} = \inf\{t \geq 0 : X_N^x(t) = y\},$$

qui est le premier temps d'atteinte de y pour la chaîne partant de x à l'instant zéro, et on veut calculer son espérance en fonction des p_i et q_i . On le fait en utilisant la méthode des équations de différence (voir [24]).

Proposition 9.1. *On a*

$$\mathbb{E}[T_{j \rightarrow n}] = \sum_{k=j}^{n-1} \sum_{l=0}^k \frac{1}{p_l} \prod_{i=l+1}^k \frac{q_i}{p_i} \quad (9.1)$$

pour tous $n = 1, \dots, N$, et $j < n$.

Remarque : Pour alléger la notation, on prend par convention $\prod_{i=l+1}^l = 1$.

Démonstration :

Soit j compris entre 1 et $n-1$. On va montrer que $\mathbb{E}[T_{j \rightarrow n}]$, satisfait une équation de différence.

En partitionnant sur la valeur du processus au premier pas, on a :

$$\begin{aligned}\mathbb{E}[T_{j \rightarrow n}] &= p_j \mathbb{E}[T_{j \rightarrow n} \mid X(1) = j+1] + q_j \mathbb{E}[T_{j \rightarrow n} \mid X(1) = j-1] + r_j \mathbb{E}[T_{j \rightarrow n} \mid X(1) = j] \\ &= p_j (\mathbb{E}[T_{j+1 \rightarrow n}] + 1) + q_j (\mathbb{E}[T_{j-1 \rightarrow n}] + 1) + r_j (\mathbb{E}[T_{j \rightarrow n}] + 1)\end{aligned}$$

Car :

$$\mathbb{E}[T_{j \rightarrow n} \mid X(1) = i] = \mathbb{E}[T_{i \rightarrow n}] + 1, \quad (9.2)$$

pour $i = j-1, j, j+1$.

En effet, on a :

$$\begin{aligned}\mathbb{E}[T_{j \rightarrow n} \mid X(1) = i] &= \sum_{k \geq 0} k \mathbb{P}(T_{j \rightarrow n} = k \mid X(1) = i) \\ &= \sum_{k \geq 0} k \mathbb{P}(T_{j \rightarrow n} = k-1 \mid X(0) = i),\end{aligned}$$

d'après la propriété de Markov.

Et donc :

$$\begin{aligned}\mathbb{E}[T_{j \rightarrow n} \mid X(1) = i] &= \sum_{k \geq 0} k \mathbb{P}(T_{i \rightarrow n} = k-1) \\ &= \sum_{k \geq 0} (k+1) \mathbb{P}(T_{i \rightarrow n} = k) \\ &= \sum_{k \geq 0} k \mathbb{P}(T_{i \rightarrow n} = k) + \sum_{k \geq 0} \mathbb{P}(T_{i \rightarrow n} = k) \\ &= \mathbb{E}[T_{i \rightarrow n}] + 1.\end{aligned}$$

D'où :

$$\begin{aligned}\mathbb{E}[T_{j \rightarrow n}] &= p_j \mathbb{E}[T_{j+1 \rightarrow n}] + q_j \mathbb{E}[T_{j-1 \rightarrow n}] + r_j \mathbb{E}[T_{j \rightarrow n}] + p_j + q_j + r_j \\ &= p_j \mathbb{E}[T_{j+1 \rightarrow n}] + q_j \mathbb{E}[T_{j-1 \rightarrow n}] + r_j \mathbb{E}[T_{j \rightarrow n}] + 1.\end{aligned}$$

$\mathbb{E}[T_{j \rightarrow n}]$ satisfait donc à l'équation de différence suivante :

$$\mathbb{E}[T_{j \rightarrow n}] = \frac{p_j}{p_j + q_j} \mathbb{E}[T_{j+1 \rightarrow n}] + \frac{q_j}{p_j + q_j} \mathbb{E}[T_{j-1 \rightarrow n}] + \frac{1}{p_j + q_j}, \quad (9.3)$$

pour $j = 1, \dots, n-1$. Avec les conditions aux bords :

$$\mathbb{E}[T_{0 \rightarrow n}] = \mathbb{E}[T_{1 \rightarrow n}] + \frac{1}{p_0}, \text{ et } \mathbb{E}[T_{n \rightarrow n}] = 0 \quad (9.4)$$

On va maintenant résoudre cette équation :

Notons $D_j = \mathbb{E}[T_{j \rightarrow n}]$, et pour tout i , $B_i = \frac{q_i}{p_i}$, $\epsilon_i = \frac{1}{p_i}$.

Alors (9.3) se réécrit, pour tout $k < n-1$:

$$D_{k+1} - D_k = B_k (D_k - D_{k-1}) - \epsilon_k \quad (9.5)$$

En itérant, on trouve :

$$D_{k+1} - D_k = \prod_{i=1}^k B_i (D_1 - D_0) - \sum_{l=1}^k \epsilon_l \prod_{i=l+1}^k B_i \quad (9.6)$$

Et on a, pour tout $j = 1, \dots, n$,

$$D_j = \sum_{k=0}^{j-1} (D_{k+1} - D_k) + D_0$$

D'où la solution générale de (9.3) :

$$D_j = \sum_{k=0}^{j-1} \left(\prod_{i=1}^k B_i (D_1 - D_0) - \sum_{l=1}^k \epsilon_l \prod_{i=l+1}^k B_i \right) + D_0, \quad (9.7)$$

pour tout $j = 1, \dots, n$.

On fait maintenant intervenir les conditions aux bords (9.4) :

On a $D_1 - D_0 = -\frac{1}{p_0}$, ce qui donne

$$D_j = \sum_{k=0}^{j-1} \left(\prod_{i=1}^k B_i \frac{-1}{p_0} - \sum_{l=1}^k \epsilon_l \prod_{i=l+1}^k B_i \right) + D_0, \quad (9.8)$$

et comme $D_n = 0$, on a :

$$\sum_{k=0}^{n-1} \left(\prod_{i=1}^k B_i \frac{-1}{p_0} - \sum_{l=1}^k \epsilon_l \prod_{i=l+1}^k B_i \right) + D_0 = 0. \quad (9.9)$$

Donc

$$D_0 = \sum_{k=0}^{n-1} \left(\prod_{i=1}^k B_i \frac{1}{p_0} + \sum_{l=1}^k \epsilon_l \prod_{i=l+1}^k B_i \right). \quad (9.10)$$

Et en reportant dans (9.7), on obtient :

$$D_j = \sum_{k=j}^{n-1} \left(\prod_{i=1}^k B_i \frac{1}{p_0} + \sum_{l=1}^k \epsilon_l \prod_{i=l+1}^k B_i \right), \quad (9.11)$$

pour tout $j < n$.

□

On montre de la même façon que, pour $j > n$, on a :

Proposition 9.2.

$$\mathbb{E}[T_{j \rightarrow n}] = \sum_{k=n+1}^j \sum_{l=k}^N \frac{1}{q_l} \prod_{i=k}^{l-1} \frac{p_i}{q_i} \quad (9.12)$$

pour tous $n = 0, \dots, N-1$, et $j > n$.

On peut utiliser la même méthode pour calculer le moment d'ordre deux de $T_{x \rightarrow y}$.

Proposition 9.3. On a

$$\mathbb{E}[T_{j \rightarrow n}^2] = \sum_{k=j}^{n-1} \sum_{l=0}^k \frac{2 \mathbb{E}[T_{l \rightarrow n}] - 1}{p_l} \prod_{i=l+1}^k \frac{q_i}{p_i} \quad (9.13)$$

pour tous $n = 1, \dots, N$, et $j < n$.

Démonstration :

On a

$$\mathbb{E}[T_{j \rightarrow n}^2 \mid X(1) = i] = \mathbb{E}[T_{i \rightarrow n}^2] + 2 \mathbb{E}[T_{i \rightarrow n}] + 1, \quad (9.14)$$

pour $i = j-1, j, j+1$.

En effet :

$$\begin{aligned} \mathbb{E}[T_{j \rightarrow n}^2 \mid X(1) = i] &= \sum_{k>0} k^2 \mathbb{P}(T_{j \rightarrow n} = k \mid X(1) = i) \\ &= \sum_{k>0} k^2 \mathbb{P}(T_{j \rightarrow n} = k-1 \mid X(0) = i), \end{aligned}$$

d'après la propriété de Markov.

Et donc :

$$\begin{aligned} \mathbb{E}[T_{j \rightarrow n}^2 \mid X(1) = i] &= \sum_{k>0} k^2 \mathbb{P}(T_{i \rightarrow n} = k-1) \\ &= \sum_{k \geq 0} (k+1)^2 \mathbb{P}(T_{i \rightarrow n} = k) \\ &= \sum_{k \geq 0} k^2 \mathbb{P}(T_{i \rightarrow n} = k) + 2 \sum_{k \geq 0} k \mathbb{P}(T_{i \rightarrow n} = k) + \sum_{k \geq 0} \mathbb{P}(T_{i \rightarrow n} = k) \\ &= \mathbb{E}[T_{i \rightarrow n}^2] + 2 \mathbb{E}[T_{i \rightarrow n}] + 1. \end{aligned}$$

D'où :

$$\begin{aligned} \mathbb{E}[T_{j \rightarrow n}^2] &= p_j \left(\mathbb{E}[T_{j+1 \rightarrow n}^2] + 2 \mathbb{E}[T_{j+1 \rightarrow n}] + 1 \right) + q_j \left(\mathbb{E}[T_{j-1 \rightarrow n}^2] + 2 \mathbb{E}[T_{j-1 \rightarrow n}] + 1 \right) \\ &\quad + r_j \left(\mathbb{E}[T_{j \rightarrow n}^2] + 2 \mathbb{E}[T_{j \rightarrow n}] + 1 \right) \\ &= p_j \mathbb{E}[T_{j+1 \rightarrow n}^2] + q_j \mathbb{E}[T_{j-1 \rightarrow n}^2] + r_j \mathbb{E}[T_{j \rightarrow n}^2] \\ &\quad + 2 \left(p_j \mathbb{E}[T_{j+1 \rightarrow n}] + q_j \mathbb{E}[T_{j-1 \rightarrow n}] + r_j \mathbb{E}[T_{j \rightarrow n}] \right) + p_j + q_j + r_j \\ &= p_j \mathbb{E}[T_{j+1 \rightarrow n}^2] + q_j \mathbb{E}[T_{j-1 \rightarrow n}^2] + r_j \mathbb{E}[T_{j \rightarrow n}^2] + 2 \mathbb{E}[T_{j \rightarrow n}] - 1. \end{aligned}$$

$\mathbb{E} [T_{j \rightarrow n}^2]$ satisfait donc à l'équation de différence suivante :

$$\mathbb{E} [T_{j \rightarrow n}^2] = \frac{p_j}{p_j + q_j} \mathbb{E} [T_{j+1 \rightarrow n}^2] + \frac{q_j}{p_j + q_j} \mathbb{E} [T_{j-1 \rightarrow n}^2] + \frac{2 \mathbb{E} [T_{j \rightarrow n}] - 1}{p_j + q_j}, \quad (9.15)$$

pour $j = 1, \dots, n-1$. Avec les conditions aux bords :

$$\mathbb{E} [T_{0 \rightarrow n}^2] = \mathbb{E} [T_{1 \rightarrow n}^2] + \frac{2 \mathbb{E} [T_{0 \rightarrow n}] - 1}{p_0}, \text{ et } \mathbb{E} [T_{n \rightarrow n}^2] = 0 \quad (9.16)$$

On résout en posant $\Delta_j = \mathbb{E} [T_{j \rightarrow n}^2]$, $B_i = \frac{q_i}{p_i}$, et $\alpha_i = \frac{2 \mathbb{E} [T_{i \rightarrow n}] - 1}{p_i}$.
(9.15) équivaut à :

$$\Delta_{k+1} - \Delta_k = B_k (\Delta_k - \Delta_{k-1}) - \alpha_k \quad (9.17)$$

Et en itérant, on trouve :

$$\Delta_{k+1} - \Delta_k = \prod_{i=1}^k B_i (\Delta_1 - \Delta_0) - \sum_{l=1}^k \alpha_l \prod_{i=l+1}^k B_i \quad (9.18)$$

Or, pour tout $j = 1, \dots, n$,

$$\Delta_j = \sum_{k=0}^{j-1} (\Delta_{k+1} - \Delta_k) + \Delta_0$$

D'où la solution générale de (9.15) :

$$\Delta_j = \sum_{k=0}^{j-1} \left(\prod_{i=1}^k B_i (\Delta_1 - \Delta_0) - \sum_{l=1}^k \alpha_l \prod_{i=l+1}^k B_i \right) + \Delta_0, \quad (9.19)$$

pour tout $j = 1, \dots, n$.

On conclut, en faisant intervenir les conditions aux bords (9.16) :

$$\Delta_j = \sum_{k=j}^{n-1} \left(\prod_{i=1}^k B_i \alpha_0 + \sum_{l=1}^k \alpha_l \prod_{i=l+1}^k B_i \right), \quad (9.20)$$

pour tout $j < n$.

□

Et on a aussi, pour $j > n$:

Proposition 9.4.

$$\mathbb{E} [T_{j \rightarrow n}^2] = \sum_{k=n+1}^j \sum_{l=k}^N \frac{2 \mathbb{E} [T_{l \rightarrow n}] - 1}{q_l} \prod_{i=k}^{l-1} \frac{p_i}{q_i} \quad (9.21)$$

pour tous $n = 0, \dots, N-1$, et $j > n$.

9.2 Marche aléatoire sur l'hypercube

On considère deux urnes U_0 et U_1 , et N boules numérotées de 1 à N .

A chaque instant, on choisit une boule au hasard (avec probabilité $1/N$), et on la change d'urne.

Soit $\xi(t) = (\xi_1(t), \dots, \xi_N(t))$ la configuration des boules au temps t ,

où $\xi_i(t) = 0$ signifie que la $i^{\text{ème}}$ boule est dans l'urne U_0 au temps t .

C'est une chaîne de Markov à valeurs dans $\{0, 1\}^N$, et on l'appelle la marche aléatoire sur le cube de dimension N .

Cette marche saute sur l'un de ses N plus proches voisins avec probabilité $1/N$, c'est à dire que l'on a :

$$\mathbb{P}(\xi(t+1) = \xi \mid \xi(t) = \eta) = \frac{1}{N}, \quad (9.22)$$

pour η et ξ telles que $\sum_i |\xi_i - \eta_i| = 1$, et que toutes les autres transitions sont nulles.

Cette chaîne est réversible par rapport à sa mesure invariante π_N , qui est la mesure uniforme sur le cube.

En effet :

Soient η et ξ deux configurations voisines (i.e. telles que $\sum_i |\xi_i - \eta_i| = 1$). Elles ne diffèrent donc que sur une seule coordonnée,

c'est à dire qu'il existe i tel que $\eta = \xi^i$, où $\xi^i = (\xi_1, \dots, 1 - \xi_i, \dots, \xi_N)$.

Et pour tout i , par définition de la réversibilité, on a :

$$\pi_N(\xi) \mathbb{P}(\xi^i \mid \xi) = \pi_N(\xi^i) \mathbb{P}(\xi \mid \xi^i). \quad (9.23)$$

Et comme, pour tout i , $\mathbb{P}(\xi^i \mid \xi) = \mathbb{P}(\xi \mid \xi^i) = 1/N$, cela implique que $\pi_N(\xi) = \pi_N(\xi^i)$, et donc que, pour tout ξ :

$$\pi_N(\xi) = \frac{1}{|\{0, 1\}^N|} = \frac{1}{2^N}. \quad (9.24)$$

Ce modèle a été introduit (dans sa description réduite) en 1907 par P. et T. Ehrenfest pour illustrer le paradoxe apparent entre la réversibilité des évolutions mécaniques (au niveau microscopique), et l'irréversibilité (macroscopique) des phénomènes thermodynamiques.

En effet, ici, bien que le processus soit réversible et que les états soient récurrents (la chaîne est irréductible), l'évolution macroscopique (au niveau de la loi du processus) est irréversible : on a, pour tout η dans $\{0, 1\}^N$

$$\lim_{t \rightarrow \infty} \sup_{\eta} \sum_{\xi} \left| \mathbb{P}(\xi(t) = \xi \mid \xi(0) = \eta) - \pi_N(\xi) \right| = 0. \quad (9.25)$$

9.2.1 Description réduite

On compte le nombre de particules dans l'urne U_1 au temps t .

On regarde donc la chaîne de Markov $X_N(t) = \sum_{i=1}^N \xi_i(t)$.

C'est une marche aléatoire sur $[0, \dots, N]$, dont les transitions sont

$$\begin{aligned} p_i &= \mathbb{P}(X(t+1) = i+1 \mid X(t) = i) = \frac{N-i}{N}, \\ q_i &= \mathbb{P}(X(t+1) = i-1 \mid X(t) = i) = \frac{i}{N}, \end{aligned} \quad (9.26)$$

pour tout $i = 0, \dots, N$.

En effet, pour passer de i à $i+1$ boules dans l'urne U_1 , on doit choisir une boule parmi les $N-1$ qui sont dans l'urne U_0 .

Dans cette description, on peut imaginer que l'on observe une particule qui se déplace entre 0 et N , qui saute à droite avec probabilité p_i , et à gauche avec probabilité q_i .

On a $p_i > q_i$ si $i < N/2$, et $q_i > p_i$ si $i > N/2$, avec égalité en $N/2$ ($p_{N/2} = q_{N/2} = 1/2$).

Remarque : on suppose donc ici que N est pair.

Il y a donc une force qui pousse la particule vers $N/2$, et on peut penser que la particule se déplace dans un puits centré en $N/2$, que l'on appelle état de basse énergie (l_N dans la terminologie du chapitre 4). Les bords du puits 0 et N sont donc les états de haute énergie (h_N dans la terminologie du chapitre 4).

La mesure d'équilibre de ce modèle est μ_N , donnée par

$$\mu_N(i) = \binom{N}{i} / 2^N. \quad (9.27)$$

C'est une mesure centrée autour de son maximum $N/2$, qui est donc l'état qui caractérise l'équilibre.

Ce processus est réversible par rapport à sa mesure d'équilibre.

En effet, on a :

$$\begin{aligned} \mu_N(i) p_i &= \frac{\binom{N}{i}}{2^N} \frac{N-i}{N} \\ &= \frac{1}{2^N N} \frac{N! (N-i)}{(N-i)! i!} \\ &= \frac{1}{2^N N} \frac{N! (i+1)}{(N-i-1)! (i+1)!} \\ &= \frac{\binom{N}{i+1}}{2^N} \frac{i+1}{N} \\ &= \mu_N(i+1) q_{i+1}. \end{aligned} \quad (9.28)$$

Mais, évidemment, dans cette description aussi, la loi du processus converge vers la mesure d'équilibre, et l'évolution macroscopique du système semble être irréversible. Si on part de zéro, la particule va se diriger vers le fond du puits, puis rester autour de $N/2$, qui est la région dans laquelle se concentre μ_N . Pour autant, les autres états n'ont pas une probabilité nulle sous la mesure invariante, et si l'on attend suffisamment longtemps, la particule va revenir en zéro (tous les états sont récurrents).

En effet, le théorème de Kac (2.9) nous dit que le temps de retour d'un état est inversement proportionnel à son poids sous la mesure invariante :

$$\mathbb{E} [T_{i \rightarrow i}] = \frac{1}{\mu_N(i)} \quad (9.29)$$

On a donc un premier élément de réponse au "paradoxe" : l'évolution n'est pas irréversible mais semble l'être, car plus on part d'un état qui se situe loin de l'équilibre (i.e. loin de $N/2$), plus l'espérance du temps de retour est grande. Et on ne peut donc observer un tel retour lors d'une expérience concrète.

Par exemple, avec $N = 20000$, et si la particule saute toutes les secondes, on a $\mathbb{E} [T_{N/2 \rightarrow N/2}] \approx 175 \text{ secondes}$, alors que $\mathbb{E} [T_{0 \rightarrow 0}] \approx 10^{6000}$ années ! (voir [31])

Dans la section suivante, on va montrer la métastabilité de la trajectoire de sortie de $N/2$ à 0 à l'aide du théorème 4.2 et des formules (9.1) et (9.12). On montre de la même façon que la trajectoire de $N/2$ à N a un comportement métastable. Ceci justifie le modèle d'Ehrenfest à 1/2 puits du chapitre 6, puisque le problème de sortie du puits est bien celui de remonter la pente (du profil énergétique). Pour étudier la convergence abrupte vers $N/2$ et la dernière sortie de $N/2$ à 0, on se reportera donc à l'étude du modèle à 1/2 puits du chapitre 6.

9.2.2 Métastabilité de la trajectoire de sortie

Comme dans la démonstration de théorème 5.6, les conditions (4.21) et (4.22) du théorème 4.2 sont satisfaites. Il reste à montrer l'existence des deux échelles (condition (4.19)). D'après (9.12), on a

$$\begin{aligned} \mathbb{E} [T_{0 \rightarrow N/2}] &= \sum_{k=0}^{N/2-1} \sum_{l=0}^k \frac{1}{p_l} \prod_{i=l+1}^k \frac{q_i}{p_i} \\ &= \sum_{k=0}^{N/2-1} \sum_{l=0}^k \frac{N}{N-l} \prod_{i=l+1}^k \frac{i}{N-i} \\ &= \sum_{k=0}^{N/2-1} \sum_{l=0}^k \frac{N}{N-l} \frac{N-l}{N-k} \frac{\binom{N}{l}}{\binom{N}{k}} \\ &= N \sum_{k=0}^{N/2-1} \frac{1}{N-k} \sum_{l=0}^k \frac{\binom{N}{l}}{\binom{N}{k}} \end{aligned} \quad (9.30)$$

On peut majorer le rapport des coefficients binomiaux, pour tout $l \leq k$:

$$\frac{\binom{N}{l}}{\binom{N}{k}} \leq \left(\frac{k}{N-k} \right)^{k-l}. \quad (9.31)$$

Et donc, comme $k < N/2 < N - k$, en majorant par la somme infinie :

$$\sum_{l=0}^k \frac{\binom{N}{l}}{\binom{N}{k}} \leq \sum_{l=0}^k \left(\frac{k}{N-k} \right)^{k-l} = \sum_{l=0}^k \left(\frac{k}{N-k} \right)^l \leq \sum_{l \geq 0} \left(\frac{k}{N-k} \right)^l = \frac{N-k}{N-2k} . \quad (9.32)$$

On a :

$$\mathbb{E} [T_{0 \rightarrow N/2}] \leq N \sum_{k=0}^{N/2-1} \frac{1}{N-2k} = \frac{N}{2} \sum_{k=1}^{N/2} \frac{1}{k} . \quad (9.33)$$

Et en utilisant l'approximation $\sum_{k=1}^x \frac{1}{k} \sim \log x$, on obtient, asymptotiquement, la majoration :

$$\mathbb{E} [T_{0 \rightarrow N/2}] \lesssim \frac{N}{2} \log \left(\frac{N}{2} \right) \quad (9.34)$$

Et d'après (9.1), on a

$$\begin{aligned} \mathbb{E} [T_{N/2 \rightarrow 0}] &= \sum_{k=1}^{N/2} \sum_{l=k}^N \frac{1}{q_l} \prod_{i=k}^{l-1} \frac{p_i}{q_i} \\ &= \sum_{k=1}^{N/2} \sum_{l=k}^N \frac{N}{l} \prod_{i=k}^{l-1} \frac{N-i}{i} \\ &= \sum_{k=1}^{N/2} \sum_{l=k}^N \frac{N}{l} \frac{l}{k} \frac{\binom{N}{l}}{\binom{N}{k}} \\ &= N \sum_{k=1}^{N/2} \frac{1}{k} \sum_{l=k}^N \frac{\binom{N}{l}}{\binom{N}{k}} \end{aligned} \quad (9.35)$$

En minorant par le premier terme de la somme, on obtient :

$$\mathbb{E} [T_{N/2 \rightarrow 0}] \geq \sum_{l=1}^N \binom{N}{l} = 2^N - 1 \quad (9.36)$$

On a donc deux états : $l_N = N/2$, et $h_N = 0$ tels que

$$\mathbb{E} [T_{h_N \rightarrow l_N}] \ll \mathbb{E} [T_{l_N \rightarrow h_N}] , \quad (9.37)$$

et d'après le théorème 4.2, la trajectoire de sortie de l'équilibre (de l_N vers h_N) est métastable.

C'est à dire que la variable $T_{N/2 \rightarrow 0} / \mathbb{E} [T_{N/2 \rightarrow 0}]$ converge en loi vers une variable exponentielle de paramètre 1.

Bibliographie

- [1] David Aldous. Random walks on finite groups and rapidly mixing Markov chains. In *Seminar on probability, XVII*, volume 986 of *Lecture Notes in Math.*, pages 243–297. Springer, Berlin, 1983.
- [2] David Aldous and Persi Diaconis. Shuffling cards and stopping times. *Amer. Math. Monthly*, 93(5) :333–348, 1986.
- [3] David Aldous and Persi Diaconis. Strong uniform times and finite random walks. *Adv. in Appl. Math.*, 8(1) :69–97, 1987.
- [4] Javiera Barrera, Béatrice Lachaud, and Bernard Ycart. Cut-off for n -tuples of exponentially converging processes. *Stochastic Process. Appl.*, 116(10) :1433–1446, 2006.
- [5] Dave Bayer and Persi Diaconis. Trailing the dovetail shuffle to its lair. *Ann. Appl. Probab.*, 2(2) :294–313, 1992.
- [6] Richard Bellman and Theodore Harris. Recurrence times for the Ehrenfest model. *Pacific J. Math.*, 1 :179–193, 1951.
- [7] Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.
- [8] Nicholas H. Bingham. Fluctuation theory for the Ehrenfest urn. *Adv. in Appl. Probab.*, 23(3) :598–611, 1991.
- [9] Anton Bovier. Metastability and ageing in stochastic dynamics. In *Dynamics and randomness II*, volume 10 of *Nonlinear Phenom. Complex Systems*, pages 17–79. Kluwer Acad. Publ., Dordrecht, 2004.
- [10] Anton Bovier. Metastability : a potential theoretic approach. In *International Congress of Mathematicians. Vol. III*, pages 499–518. Eur. Math. Soc., Zürich, 2006.
- [11] Krzysztof Burdzy and Wilfrid S. Kendall. Efficient Markovian couplings : examples and counterexamples. *Ann. Appl. Probab.*, 10(2) :362–409, 2000.
- [12] Marzio Cassandro, Antonio Galves, Enzo Olivieri, and Maria Eulália Vares. Metastable behavior of stochastic dynamics : a pathwise approach. *J. Statist. Phys.*, 35(5-6) :603–634, 1984.
- [13] Fan Rong K. Chung, Persi Diaconis, and Ronald L. Graham. Random walks arising in random number generation. *Ann. Probab.*, 15(3) :1148–1165, 1987.

- [14] Kai Lai Chung. *A course in probability theory*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, second edition, 1974. Probability and Mathematical Statistics, Vol. 21.
- [15] Frank den Hollander, Enzo Olivieri, and Elisabetta Scoppola. Metastability and nucleation for conservative dynamics. *J. Math. Phys.*, 41(3) :1424–1498, 2000. Probabilistic techniques in equilibrium and nonequilibrium statistical physics.
- [16] Persi Diaconis. *Group representations in probability and statistics*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11. Institute of Mathematical Statistics, Hayward, CA, 1988.
- [17] Persi Diaconis. The cutoff phenomenon in finite Markov chains. *Proc. Nat. Acad. Sci. U.S.A.*, 93(4) :1659–1664, 1996.
- [18] Persi Diaconis. Mathematical developments from the analysis of riffle shuffling. In *Groups, combinatorics & geometry (Durham, 2001)*, pages 73–97. World Sci. Publ., River Edge, NJ, 2003.
- [19] Persi Diaconis, James Allen Fill, and Jim Pitman. Analysis of top to random shuffles. *Combin. Probab. Comput.*, 1(2) :135–155, 1992.
- [20] Persi Diaconis, Ronald L. Graham, and John A. Morrison. Asymptotic analysis of a random walk on a hypercube with many dimensions. *Random Structures Algorithms*, 1(1) :51–72, 1990.
- [21] Persi Diaconis and Laurent Saloff-Coste. Separation cut-offs for birth and death chains. *Ann. Appl. Probab.*, 16(4) :2098–2122, 2006.
- [22] Persi Diaconis and Mehrdad Shahshahani. Time to reach stationarity in the Bernoulli-Laplace diffusion model. *SIAM J. Math. Anal.*, 18(1) :208–218, 1987.
- [23] Paul Ehrenfest and Tatiana Ehrenfest. *The conceptual foundations of the statistical approach in mechanics*. Dover Publications Inc., New York, english edition, 1990. Translated from the German by Michael J. Moravcsik, With a foreword by M. Kac and G. E. Uhlenbeck.
- [24] William Feller. *An introduction to probability theory and its applications. Vol. I*. Third edition. John Wiley & Sons Inc., New York, 1968.
- [25] Mark I. Freidlin and Alexander D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1984. Translated from the Russian by Joseph Szücs.
- [26] Sheldon Goldstein. Maximal coupling. *Z. Wahrsch. Verw. Gebiete*, 46(2) :193–204, 1978/79.
- [27] David Griffeath. A maximal coupling for Markov chains. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 31 :95–106, 1974/75.
- [28] David Griffeath. Uniform coupling of non-homogeneous Markov chains. *J. Appl. Probability*, 12(4) :753–762, 1975.

- [29] David Griffeath. Coupling methods for Markov processes. In *Studies in probability and ergodic theory*, volume 2 of *Adv. in Math. Suppl. Stud.*, pages 1–43. Academic Press, New York, 1978.
- [30] Paul G. Hoel, Sidney C. Port, and Charles J. Stone. *Introduction to stochastic processes*. Houghton Mifflin Co., Boston, Mass., 1972. The Houghton Mifflin Series in Statistics.
- [31] Mark Kac. Random walk and the theory of Brownian motion. *Amer. Math. Monthly*, 54 :369–391, 1947.
- [32] Samuel Karlin and James McGregor. Ehrenfest urn models. *J. Appl. Probability*, 2 :352–376, 1965.
- [33] Roman Kotecký and Enzo Olivieri. Droplet dynamics for asymmetric Ising model. *J. Statist. Phys.*, 70(5-6) :1121–1148, 1993.
- [34] Béatrice Lachaud. Cut-off and hitting times of a sample of Ornstein-Uhlenbeck processes and its average. *J. Appl. Probab.*, 42(4) :1069–1080, 2005.
- [35] Béatrice Lachaud and Bernard Ycart. Convergence times for parallel Markov chains. In *Positive systems*, volume 341 of *Lecture Notes in Control and Inform. Sci.*, pages 169–176. Springer, Berlin, 2006.
- [36] Gérard Letac and Lajos Takács. Random walks on an m -dimensional cube. *J. Reine Angew. Math.*, 310 :187–195, 1979.
- [37] Francesco Manzo, Francesca R. Nardi, Enzo Olivieri, and Elisabetta Scoppola. On the essential features of metastability : tunnelling time and critical configurations. *J. Statist. Phys.*, 115(1-2) :591–642, 2004.
- [38] Servet Martínez and Bernard Ycart. Decay rates and cutoff for convergence and hitting times of Markov chains with countably infinite state space. *Adv. in Appl. Probab.*, 33(1) :188–205, 2001.
- [39] Peter Matthews. Mixing rates for a random walk on the cube. *SIAM J. Algebraic Discrete Methods*, 8(4) :746–752, 1987.
- [40] E. Jordão Neves and Roberto H. Schonmann. Critical droplets and metastability for a Glauber dynamics at very low temperatures. *Comm. Math. Phys.*, 137(2) :209–230, 1991.
- [41] E. Jordão Neves and Roberto H. Schonmann. Behavior of droplets for a class of Glauber dynamics at very low temperature. *Probab. Theory Related Fields*, 91(3-4) :331–354, 1992.
- [42] Enzo Olivieri and Elisabetta Scoppola. Markov chains with exponentially small transition probabilities : first exit problem from a general domain. I. The reversible case. *J. Statist. Phys.*, 79(3-4) :613–647, 1995.
- [43] Enzo Olivieri and Elisabetta Scoppola. Markov chains with exponentially small transition probabilities : first exit problem from a general domain. II. The general case. *J. Statist. Phys.*, 84(5-6) :987–1041, 1996.

- [44] Enzo Olivieri and Elisabetta Scoppola. Metastability and typical exit paths in stochastic dynamics. In *European Congress of Mathematics, Vol. II (Budapest, 1996)*, volume 169 of *Progr. Math.*, pages 124–150. Birkhäuser, Basel, 1998.
- [45] Enzo Olivieri and Maria Eulália Vares. *Large deviations and metastability*, volume 100 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2005.
- [46] Jim Pitman. On coupling of Markov chains. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 35(4) :315–322, 1976.
- [47] Laurent Saloff-Coste. Lectures on finite Markov chains. In *Lectures on probability theory and statistics (Saint-Flour, 1996)*, volume 1665 of *Lecture Notes in Math.*, pages 301–413. Springer, Berlin, 1997.
- [48] Laurent Saloff-Coste. Random walks on finite groups. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pages 263–346. Springer, Berlin, 2004.
- [49] Roberto H. Schonmann. An approach to characterize metastability and critical droplets in stochastic Ising models. *Ann. Inst. H. Poincaré Phys. Théor.*, 55(2) :591–600, 1991. Multiscale phenomena (São Paulo, 1990).
- [50] Roberto H. Schonmann. The pattern of escape from metastability of a stochastic Ising model. *Comm. Math. Phys.*, 147(2) :231–240, 1992.
- [51] Elisabetta Scoppola. Metastability for Markov chains : a general procedure based on renormalization group ideas. In *Probability and phase transition (Cambridge, 1993)*, volume 420 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, pages 303–322. Kluwer Acad. Publ., Dordrecht, 1994.
- [52] Lajos Takács. On an urn problem of Paul and Tatiana Ehrenfest. *Math. Proc. Cambridge Philos. Soc.*, 86(1) :127–130, 1979.
- [53] Hermann Thorisson. On maximal and distributional coupling. *Ann. Probab.*, 14(3) :873–876, 1986.
- [54] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.
- [55] Bernard Ycart. Cutoff for samples of Markov chains. *ESAIM Probab. Statist.*, 3 :89–106 (electronic), 1999.
- [56] Bernard Ycart. Cutoff for Markov chains : some examples and applications. In *Complex systems (Santiago, 1998)*, volume 6 of *Nonlinear Phenom. Complex Systems*, pages 261–300. Kluwer Acad. Publ., Dordrecht, 2001.
- [57] Bernard Ycart. *Modèles et algorithmes markoviens*, volume 39 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 2002.