



HAL
open science

La distribution de la quasi-espèce pour une population finie

Joseba Dalmau

► **To cite this version:**

Joseba Dalmau. La distribution de la quasi-espèce pour une population finie. Probabilités [math.PR]. Université Paris Saclay (COMUE), 2016. Français. NNT : 2016SACLS452 . tel-01415001

HAL Id: tel-01415001

<https://theses.hal.science/tel-01415001>

Submitted on 12 Dec 2016

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.

NNT : 2016SACLS452

THÈSE DE DOCTORAT

de

L'UNIVERSITÉ PARIS-SACLAY

préparée à l'Université Paris-Sud

École doctorale de mathématiques Hadamard (EDMH, ED 574)

Établissement d'inscription : Université Paris-Sud

Laboratoire d'accueil : DMA (ENS), UMR 8553 CNRS

Spécialité de doctorat : Mathématiques aux interfaces

Joseba DALMAU

La distribution de la quasi-espèce pour une population finie

Thèse présentée et soutenue à Paris le 25 novembre 2016

Après avis des rapporteurs : MICHEL BENAÏM (Université de Neuchâtel)
NICOLAS CHAMPAGNAT (Inria Nancy)

Jury de soutenance :

RAPHAËL CERF	(École normale supérieure) Directeur de thèse
NICOLAS CHAMPAGNAT	(Inria Nancy) Rapporteur
LLOYD DEMETRIUS	(Harvard University) Examinateur
IRINA KOURKOVA	(Université Paris 6) Examinatrice
SYLVIE MÉLÉARD	(École Polytechnique) Présidente du Jury
AMANDINE VÉBER	(École Polytechnique) Examinatrice

**La distribution de la quasi-espèce
pour une population finie**

Titre : La distribution de la quasi-espèce pour une population finie.

Mots clé : quasi-espèce, Eigen, Moran, Wright–Fisher, Galton–Watson.

Résumé : Le concept de quasi-espèce, introduit par Manfred Eigen dans les années 70, décrit l'état d'équilibre d'une population subissant des forces de mutation et de sélection. La plupart des modèles classiques présentant un phénomène de quasi-espèce sont déterministes et considèrent une population de taille infinie. L'objectif de cette thèse est d'étudier plusieurs modèles stochastiques, dont la taille de la population est finie, et de montrer que le phénomène de la quasi-espèce est aussi présent dans ces modèles. Nous étudions en détail les modèles de Galton–Watson, Wright–Fisher et Moran. Nous confirmons que les trois modèles présentent un phénomène de quasi-espèce, et nous trouvons que la distribution de la quasi-espèce est la même pour les trois modèles, ainsi que pour le modèle original d'Eigen. De plus, nous décrivons explicitement la distribution de cette quasi-espèce pour le paysage à un pic, ainsi que pour des fonctions de fitness qui ne dépendent que de la distance de Hamming au génotype le plus adapté.

Title: The distribution of the quasispecies for a finite population

Keywords: quasispecies, Eigen, Moran, Wright–Fisher, Galton–Watson.

Abstract: Manfred Eigen introduced the concept of quasispecies in the early 70s, in order to describe the steady-state distribution of a population subject to mutation and selection forces. Most classical models showing a quasispecies phenomenon are deterministic and deal with a population of an infinite size. The aim of this thesis is to study several stochastic and finite-population models, and to show that a quasispecies phenomenon arises in these models too. We study in detail the Galton–Watson, Wright–Fisher and Moran models. We confirm that a quasispecies phenomenon is present in the three models, and we find that the distribution of this quasispecies is common to all three models as well as to Eigen's original model. Moreover, we describe explicitly the distribution of the quasispecies for the sharp peak landscape, as well as for fitness functions which depend only on the Hamming distance to the fittest genotype.

Remerciements

Tout d'abord, un énorme merci à **Raphaël Cerf**, sans qui cette thèse n'aurait jamais existé. Il a été le premier professeur que j'ai eu à mon arrivée en France. La facilité et le plaisir avec lesquels un étranger parlant quatre mots de français, comme je l'étais, peut assister à ses leçons et les comprendre, témoignent de sa qualité comme enseignant. Excellent professeur, il s'est avéré être un aussi formidable directeur de thèse. Toujours présent et disponible, il m'a engagé dans un sujet de thèse intéressant, stimulant, et surtout très beau. J'ai éprouvé un grand plaisir à travailler avec lui, il a su me transmettre tout son enthousiasme pour les mathématiques et je garderai toujours un très bon souvenir de nos discussions, qui sont allées bien au delà du domaine des mathématiques. Je ne peux guère m'empêcher d'admirer le dévouement et la ténacité avec lesquelles il a lu et relu mes travaux. Enfin, je tiens à le remercier notamment pour son hospitalité, ainsi que celle de **Katti, Benjamin, Nelly, Samuel** et **Juliette**; ils m'ont maintes fois accueilli chez eux, toujours de la façon la plus chaleureuse. Je garde un souvenir particulièrement cher de mes séjours à Cambridge et à Padoue.

Michel Benaïm et **Nicolas Champagnat** méritent aussi ma plus grande reconnaissance pour avoir accepté de rapporter cette thèse. Ils ont tous les deux fait un travail admirable en lisant avec attention ce manuscrit d'un bout à l'autre. Leurs remarques et commentaires ont été très enrichissants et m'ont aidé à améliorer le document.

Je remercie également **Lloyd Demetrius, Irina Kourkova, Sylvie Méléard** et **Amandine Véber** d'avoir fait l'effort de faire partie de mon jury de thèse. J'apprécie particulièrement la présence de **Lloyd Demetrius**, qui s'intéresse depuis des années au sujet de la quasi-espèce; mes travaux sur le processus de Galton-Watson sont en partie inspirés des siens. J'avais rencontré **Irina Kourkova** avant

même de commencer mes études en France, dans la réunion de présentation du M2 Probabilités et Applications à Jussieu ; je suis ravi de la retrouver dans mon jury de thèse. J'ai eu le plaisir d'avoir **Sylvie Méléard** et **Amandine Véber** comme professeurs dans le cadre du M2 Probabilités et Statistiques de l'Université Paris Sud, leurs cours ont certainement eu une influence sur mon goût pour la biologie mathématique ; j'ai appris avec enthousiasme leur participation à mon jury de thèse avec grand enthousiasme.

Je tiens ensuite à remercier **Paolo Dai Pra** et **Carlo Mariconda** de m'avoir permis de passer deux très belles semaines à Padoue, ce qui m'a offert un cadre très agréable pour travailler, avec comme résultat probablement le plus bel article que cette thèse ait donné. Je souhaite aussi remercier **Matthieu Jonckheere** de m'avoir invité à séjourner au BCAM à Bilbao ; ces deux semaines de discussions mathématiques constantes m'ont beaucoup appris.

Je ne peux qu'exprimer ma plus profonde gratitude à **Javier Duoandikoetxea**, qui m'a suivi pendant toute ma scolarité à l'Université du Pays Basque et qui m'a orienté vers la France. Ses cours d'analyse comptent sans doute parmi les plus beaux cours que j'ai suivis ; avec sa façon d'expliquer les mathématiques comme si d'un film il s'agissait, en arrivant aux théorèmes de façon inespéré, il a su captiver de nombreuses promotions d'élèves au Pays Basque.

Cette thèse n'aurait pas été possible sans l'aide précieuse de **Valérie Lavigne** et **Florence Rey**, qui m'ont aidé à me repérer dans la mer sombre de la bureaucratie académique à l'Université Paris-Sud, ni sans celle de **Bénédicte Auffray** et **Zaina Elmir**, sans qui je me serais sûrement noyé dans les eaux troubles de l'encore plus vaste océan de l'administration française.

Au cours de mon doctorat, j'ai eu l'occasion de rencontrer un bon nombre de collègues, co-bureaux et autres. Je voudrais d'abord remercier **Matthias**, mon grand frère de thèse, avec qui ça a toujours été un plaisir de discuter auprès d'un tableau, ou d'un verre. Je remercie l'accueil sympathique avec lequel je fus reçu par les doctorants et caïmans du passage vert au DMA : **Bastien**, toujours prêt à répondre à mes questions ; **Clément**, l'homme le moins frileux que j'ai connu ; **Igor**, organisateur infatigable des Bodégas ; **Ilaria**, qui m'a donné mes premières leçons d'italien ; et **Oriane**, qui est partie à Lyon et dont les sages conseils me manquent. Arrivèrent ensuite **Yichao**, que j'ai eu plusieurs fois le plaisir d'écouter jouer du piano, mais ce n'est pas pour autant qu'il a cessé de m'émerveiller ; **Shen**, qui tient toujours une conversation agréable et intéres-

sante ; **Guillaume**, avec qui rarement je réussis à manger à cause de nos différences culturelles concernant les horaires des repas ; et **Wei**, à qui je souhaite une très bonne réussite pour sa thèse.

Nous sommes bientôt arrivés à la fin, et je dois bien sûr remercier mes amis, qui m'ont soutenu pendant ces trois ans. Mila esker, **Naiara** eta **Zabalo**, egindako bisitengatik, eta elkarrekin pasatako opor ederrengatik. Eskerrik asko **Kobe**, fakultateko urte ederrengatik eta atzetik etorri den guztiagatik. Muchas gracias a **Abbas** por todas las cenas que hemos cocinado juntos, a **David** por haberme hablado de París y de los másteres de probabilidad, a **Malick** por haberme ayudado con el francés en mis primeros meses en Francia, y a **María**, por haberse acercado a nosotros cuando no conocíamos a nadie. Gracias por supuesto a **Rita**, por haberme aguantado durante nuestro exilio en el lejano KB. Muchas gracias también a **Simonsson**, por haber hecho tantas cosas útiles y por haberme invitado a dar esa charla en Gothenburgo. Mila esker **Eneko**, nire lehen lettre de motivation delakoekin laguntzeagatik, eta **Irantzu**, Parisera heldu nintzenez zure etxean hartu izanagatik. Je remercie **Paul** d'avoir fait son M2 le même année que moi. Gracias también a **Ale**, por todas las noches juntos en la Cité Universitaire, and of course to **Mudit**, who put up with me for several months, and has warmly hosted me in India. Muchas gracias a **Nico** por los innumerables cafés, cervezas, comidas... Grazie mille a tutta la banda dei rituali (e amici) : grazie **Camilla** per la tua gentilezza e per tutte le volte che hai pulito i piatti, **Dudo** per la tua motivazione inesauribile, et **Fred** pour les soirées jeux et le repas GoT. Grazie **Inès** per la tua ospitalità e per avere sempre risposto alle mie domande di biologia, **Jared** per avere sistemato et KB tante volte ; et **Louise** pour cette belle visite au Louvre. Grazie anche a **Marco** d'avermi accettato come coinquilino ; a **Martino** per le uscite in foresta, e le chiacchierate sempre così simpatiche ; y a **Marzia**, por toda su energía. Grazie oppure a **Fede**, **Giulia**, **Mattia Valerio** e **Vito** per tutte le serate passate insieme. Merci enfin à **Alice** de m'avoir accepté chez elle avec tant de sympathie.

Gracias sobre todo a **aita** y a **ama**, pues son sin lugar a duda las personas que más han contribuido a que mis estudios hayan llegado aquí. Gracias también a mi hermano **Mikel**.

Enfin, merci beaucoup à **Camille**, qui a lu toute ma thèse, d'un bout à l'autre, pour m'aider avec mon français. Mais qui surtout, m'a soutenu dans les moments de stress, et a su me rendre heureux jour après jour.

Table des matières

Table des matières	11
1 Introduction	19
1.1 État de l'art	20
1.2 Cadre	22
1.3 Le modèle d'Eigen	23
1.4 Le modèle de Galton–Watson	28
1.5 Le modèle de Wright–Fisher	30
1.6 Le modèle de Moran	33
1.7 Conclusion et perspectives	38
1.8 Publications	38
I Le modèle d'Eigen	41
2 Introduction	43
3 Généralités	47

3.1	Le modèle	47
3.2	Le modèle pour les classes de Hamming	49
3.3	Solutions stationnaires et stabilité	51
3.4	Régime asymptotique	57
4	Le paysage à un pic	59
4.1	Solutions stationnaires	59
4.2	La distribution de la quasi-espèce	61
4.3	Convergence des trajectoires	65
4.4	Convergence des systèmes	67
5	Le cas général	75
5.1	Solutions stationnaires	75
5.2	Deux formules pour la solution de (\mathcal{R})	80
5.3	Fitness constantes à partir d'un certain rang	83
5.4	Seuil d'erreur	89
5.5	Convergence des trajectoires	93
5.6	Convergence des systèmes	96
6	Conclusion	103
II	Le modèle de Galton–Watson	107
7	Le modèle de Galton–Watson	109

7.1 Introduction 109

7.2 Le modèle 111

7.3 Le processus d'occupation 112

7.4 Lien avec le modèle d'Eigen 114

7.5 La quasi-espèce pour Galton–Watson 116

7.6 Conclusion 119

III Le modèle de Wright–Fisher 121

8 Introduction 123

9 Modèle et résultats 127

9.1 Le modèle 127

9.2 Résultat principal 129

9.3 Lumping 131

9.4 Couplage et monotonie 133

9.5 Convergence vers un modèle de Moran 137

10 Bornes stochastiques 141

10.1 Processus inférieur et supérieur 141

10.2 Dynamique des processus bornants 143

10.3 Bornes sur la mesure invariante 146

10.4 Comparaison avec la master sequence 151

11 Chaînes de Markov	153
11.1 PGD pour la matrice de transition	153
11.2 Concentration autour de ρ^*	165
12 Fin de la démonstration	177
13 Conclusion	179
IV Le modèle de Moran	181
14 Introduction	183
15 Modèle et résultats	187
15.1 Le modèle	187
15.2 Résultat principal	190
15.3 Modèles normalisés	191
15.4 Lumping	194
15.5 Convergence vers le modèle d'Eigen	198
15.6 Heuristiques	202
16 Couplage et monotonie	203
16.1 Couplage	203
16.2 Monotonie	206
17 Bornes stochastiques	209

17.1	Première approche	209
17.1.1	Processus inférieur et supérieur	209
17.1.2	Dynamique des processus bornants	212
17.1.3	Bornes sur la mesure invariante	214
17.2	Le processus inférieur	218
17.3	Dynamique du processus inférieur	221
17.4	Le processus supérieur	223
17.5	Dynamique du processus supérieur	230
17.6	Bornes sur la mesure invariante	233
18	Réurrence et estimées	237
18.1	Le cas $k = 0$	240
18.2	L'étape de récurrence	249
18.3	Bornes stochastiques pour $(Z_n(k))_{n \geq 0}$	251
18.4	Chaînes de naissance et mort	257
19	Convergence	267
20	Cas neutre	275
20.1	Estimation du temps de découverte	276
20.2	Estimation du temps de persistance	283
21	Fin de la démonstration	289
21.1	Le cas sous-critique	290

21.2	Le cas sur-critique	290
22	Le principe de grandes déviations	293
22.1	Principe de grandes déviations	294
22.2	Outils	298
22.2.1	Transformée de Fenchel–Legendre	298
22.2.2	La log–Laplace	299
22.2.3	Entropie relative	300
22.2.4	Transformée de Fenchel–Legendre de la log–Laplace . . .	302
22.2.5	Transformée de Cramér	304
22.2.6	Le cumulants conditionnel	305
22.2.7	La transformée de Cramér généralisée	306
22.3	Bornes supplémentaires	307
22.3.1	Les probabilités C_x	307
22.3.2	Convergence uniforme de G et de ses dérivées	310
22.3.3	La fonction G^∞ et sa transformée de Fenchel–Legendre . .	312
22.3.4	La fonctionnelle d'action	314
22.4	Borne inférieure du PGD	319
22.5	Borne supérieure du PGD	329
23	Sortie d'un domaine	337
23.1	Perturbation aléatoire d'un système d'EDO	337
23.2	Résultats supplémentaires	338

Table des matières	17
23.3 Temps de sortie	340
24 Disparition de la quasi-espèce	345
24.1 Calcul de la fonctionnelle d'action	345
24.2 Calcul du quasi-potentiel	348
24.3 Trajectoire de disparition	351
25 Conclusion	357
A Lemmes sur les équations différentielles linéaires	359
B Lumping	363
Bibliographie	365
Index	373

Chapitre 1

Introduction

Dans les années 70, Manfred Eigen propose un modèle mathématique afin de décrire l'évolution d'une population de macromolécules prébiotiques [28]. Les macromolécules sont classées selon leur génotype, et l'évolution est guidée par deux forces principales : la sélection, qui permet à certains génotypes de se reproduire plus fréquemment que d'autres, et la mutation, qui permet, lors d'une reproduction, de changer de génotype. Les concentrations des différents génotypes évoluent selon un système d'équations différentielles.

Le modèle d'Eigen présente deux phénomènes d'importance. Le premier, un phénomène de transition de phase pour le taux de mutation : il existe un taux de mutation critique séparant un régime de désordre d'un régime ordonné, ce taux de mutation critique étant appelé le seuil d'erreur. Le deuxième, une distribution de quasi-espèce : dans le régime ordonné, la population à l'équilibre contient une concentration positive du génotype le plus adapté, ainsi qu'un nuage de mutants qui diffèrent du génotype principal de quelques mutations.

Ces phénomènes de seuil d'erreur et de quasi-espèce ont fait du modèle d'Eigen un modèle fondamental, non seulement pour modéliser des populations prébiotiques, mais aussi pour modéliser des populations de bactéries, de virus et même de cellules cancéreuses. Cependant, pour des génomes comme ceux des virus ou des bactéries, le nombre de génotypes qui pourraient potentiellement se former excède largement la taille de toute population viable. Le modèle d'Eigen est bâti avec les concentrations des différents génotypes, ce qui équivaut à considérer une population de taille infinie. Ce cadre ne permet donc pas

d'évaluer l'impact de la taille de la population dans le seuil d'erreur et dans la distribution de la quasi-espèce.

L'objectif de cette thèse est d'étudier différents modèles pour l'évolution d'une population finie, afin de comprendre l'impact de la taille de la population dans la distribution de la quasi-espèce. Les modèles étudiés sont trois des modèles mathématiques les plus célèbres en génétique des populations : le modèle de Galton–Watson, le modèle de Wright–Fisher, et le modèle de Moran. Dans un régime asymptotique adéquat, les trois modèles présentent un phénomène de transition de phase [10, 11, 19], sous une forme différente pour chaque modèle. Néanmoins, nous montrons que, quand une quasi-espèce est formée, la distribution de la quasi-espèce est la même dans les trois modèles, et c'est la distribution prédite dans le modèle d'Eigen. En outre, nous trouvons une formule explicite pour la distribution de la quasi-espèce.

Les parties II, III et IV de la thèse étudient les modèles de Galton–Watson, Wright–Fisher et Moran, respectivement. Les trois modèles sont intimement liés au modèle d'Eigen, et une étude détaillée du modèle d'Eigen, intéressante en soi, est aussi nécessaire pour une meilleure compréhension des autres modèles. Cette étude fait l'objet de la partie I. Dans le reste de cette introduction, nous présentons d'abord l'état de l'art du sujet, nous discutons ensuite les différents résultats obtenus au cours de la thèse, et nous finissons avec une brève conclusion.

1.1 État de l'art

Même si le concept était déjà présent dans l'article d'Eigen de 1971 [28], le terme de quasi-espèce ne parut que dans la série d'articles *The hypercycle*, écrite par Manfred Eigen et Peter Schuster [31, 32, 33]. Avant la série d'articles d'Eigen et Schuster, le système d'équations différentielles d'Eigen avait déjà attiré l'attention de plusieurs chercheurs. Ainsi, Thompson et McBride [78], et de manière indépendante Enns, Jones et Rangnekar [49], trouvèrent des solutions exactes pour le système d'Eigen. Une excellente synthèse du développement de la théorie de la quasi-espèce jusqu'à la fin des années 80 est réalisée par Eigen, McCaskill et Schuster dans [30].

Le traitement mathématique du modèle d'Eigen évolue ensuite grâce à l'intro-

duction des techniques de la mécanique statistique dans son étude. Les premiers travaux dans cette direction sont ceux de Leuthäusser [53, 54], qui établit un lien entre le modèle d'Eigen et un modèle d'Ising. L'interprétation du seuil d'erreur comme un phénomène de transition de phase [77] a incité plusieurs physiciens et mathématiciens à contribuer au développement de la théorie de la quasi-espèce, comme indiqué dans [4].

En même temps que les premiers développements théoriques sur la quasi-espèce, du côté expérimental, Esteban Domingo et ses collaborateurs [24, 25, 26], ainsi que John Holland et ses collaborateurs [44, 45], observent une variabilité génomique très importante dans des populations de virus d'ARN aussi bien *in vivo* que *in vitro*. Des observations similaires ont été faites par de nombreux autres chercheurs, notamment par Biebricher [6]. Le concept de quasi-espèce dépasse ainsi le cadre de l'évolution moléculaire et devient un objet central dans la compréhension des populations virales.

Une fois le lien entre le développement théorique du modèle de quasi-espèce et l'observation d'une grande variabilité dans des populations de virus d'ARN établi, il est naturel de s'interroger sur le rôle du seuil d'erreur dans des populations de virus. Le fait que de nombreux virus d'ARN ont des taux de mutation à peine en-dessous du seuil d'erreur [21, 27, 52, 79] fait largement consensus. Cette constatation a donné lieu à de nombreuses stratégies thérapeutiques, notamment contre le virus du VIH [2, 43, 55, 79] et d'autres virus [15, 80]. Ces stratégies sont basées sur l'idée de catastrophe d'erreur [29] ; il s'agit de monter artificiellement le taux de mutation du virus afin de traverser le seuil d'erreur, de sorte que l'information génétique pertinente est perdue après quelques générations, et le virus meurt. Il est aussi postulé qu'un comportement similaire guide l'évolution de certaines bactéries [8] et même des cellules cancéreuses [37, 75].

Deux des caractéristiques principales du modèle d'Eigen sont sa nature déterministe (il s'agit d'un système d'équations différentielles) et la considération d'une population infinie. Ces caractéristiques, bien que naturelles pour modéliser une population prébiotique, deviennent assez irréalistes à l'heure de modéliser des populations d'individus plus complexes comme les virus [85]. En effet, le nombre de génomes qu'un virus pourrait potentiellement former dépasse largement la taille de toute population viable. En outre, il est reconnu que des effets aléatoires jouent un rôle important dans l'évolution des populations d'une certaine complexité. Il est donc important pour la théorie de la quasi-espèce de développer et d'étudier des modèles à population finie, aléatoires, où

les concepts de seuil d'erreur et de quasi-espèce sont présents [30, 85].

Plusieurs travaux ont poursuivi la tâche de proposer des modèles aléatoires alternatifs au modèle d'Eigen. Un des premiers travaux dans ce sens est celui de Demetrius, Schuster et Sigmund [19], qui utilisent un processus de branchement avec un comportement asymptotique similaire au modèle d'Eigen. McCaskill [56] introduit un processus de Markov dont les taux de transition sont donnés par le système d'Eigen. Nowak et Schuster [61] se servent d'un processus de naissance et mort markovien pour donner une version stochastique du modèle d'Eigen dans le paysage à un pic. Alves et Fontanari [1] étudient la dépendance du seuil d'erreur vis-à-vis de la taille de la population dans le cas du paysage à un pic. Saakian, Deem et Hu [67] calculent la variance de la fitness moyenne dans un modèle à population finie pour approximer le modèle d'Eigen par le modèle à population finie. D'autres travaux ont aussi été réalisés par Deem, Muñoz et Park [63], Musso [60], Dixit, Srivastava et Vishnoi [20], ou Weinberger [81].

La plupart des travaux précédents utilisent le modèle d'Eigen comme point de départ et essaient de construire des modèles à population finie qui convergent vers le modèle d'Eigen en grande population. Nous explorons une voie alternative : nous partons des modèles mathématiques classiques en génétique des populations, et nous montrons que, dans un certain régime asymptotique, un phénomène de seuil d'erreur a lieu, et une quasi-espèce est créée. Nous découvrons ainsi des liens parfois inattendus entre les modèles classiques en génétique des populations et le modèle d'Eigen. Ceci va dans le sens des affirmations de Wilke [85] sur la compatibilité entre la théorie de la quasi-espèce et la génétique des populations classique. De plus, nous décrivons via une formule explicite la distribution de la quasi-espèce pour plusieurs paysages de fitness différents. Nous résumons nos résultats dans les sections suivantes.

1.2 Cadre

Afin de rendre l'exposition des résultats principaux plus claire, nous allons les introduire dans un cadre légèrement plus simple que dans le reste de la thèse. Ce cadre est commun à tous les modèles que nous étudions.

L'ensemble des génotypes est donné par l'hypercube $\{0, 1\}^\ell$ de dimension $\ell \geq 1$,

et est naturellement muni de la distance de Hamming :

$$\forall u, v \in \{0, 1\}^\ell \quad d_H(u, v) = \text{card} \{ 1 \leq l \leq \ell : u(l) \neq v(l) \}.$$

Nous considérons une fonction de fitness $f : \{0, 1\}^\ell \rightarrow \mathbb{R}^+$, et nous supposons que les mutations arrivent lors des reproductions, indépendamment avec probabilité $q \in]0, 1[$, i.e., le génotype u mute vers le génotype v avec probabilité

$$M(u, v) = q^{d_H(u, v)} (1 - q)^{\ell - d_H(u, v)}.$$

Nous allons surtout considérer deux types de fonctions de fitness :

Paysage à un pic. Il existe un génotype $w^* \in \{0, 1\}^\ell$, appelé master sequence, ayant fitness $\sigma > 1$. Tous les autres génotypes ont fitness 1. La fonction de fitness est donc donnée par

$$\forall u \in \{0, 1\}^\ell, \quad f(u) = \begin{cases} \sigma & \text{si } u = w^*, \\ 1 & \text{si } u \neq w^*. \end{cases}$$

Paysage de fitness dépendant de la classe de Hamming. Il existe un génotype privilégié $w^* \in \{0, 1\}^\ell$, appelé master sequence, ayant fitness maximale (strictement), et

$$\forall u, v \in \{0, 1\}^\ell \quad d_H(u, w^*) = d_H(v, w^*) \implies f(u) = f(v).$$

Tous les génotypes à une distance donnée d de la master sequence forment la classe de Hamming d .

1.3 Le modèle d'Eigen

Nous notons par $x_v(t)$ la concentration du génotype v dans la population au temps t . Dans le modèle d'Eigen, les concentrations $x_v(t)$ évoluent selon le système d'équations différentielles

$$(*) \quad x'_v(t) = \sum_{u \in \{0, 1\}^\ell} x_u(t) f(u) M(u, v) - x_v(t) \sum_{u \in \{0, 1\}^\ell} x_u(t) f(u), \quad v \in \{0, 1\}^\ell.$$

Ce système d'équations différentielles fut proposé par Eigen [28], qui l'étudia plus en détail dans le cas du paysage à un pic. Il étudia notamment le comportement du système quand la longueur des chaînes ℓ tend vers l'infini, et c'est

ainsi qu'il découvrit les phénomènes de seuil d'erreur et de quasi-espèce. Ces phénomènes apparaissent à l'équilibre, et différents chercheurs ont étudié les solutions stationnaires de (*).

Le système (*), sans considérer le régime asymptotique en chaîne longue, a été étudié par plusieurs auteurs [5, 48, 49, 76, 78]. Le résultat le plus significatif dans ces travaux est le suivant.

Théorème 1.3.1. *Le système d'Eigen (*) admet une unique solution stationnaire x^* . De plus, pour toute condition initiale, la solution du système d'Eigen converge vers x^* .*

Nous montrons ce résultat dans la section 3.3 à l'aide du théorème de Perron–Frobenius. Remarquons que dans le cas du système (*), il n'y a pas de formule explicite connue pour la solution stationnaire unique du théorème. Afin de retrouver les phénomènes de seuil d'erreur et de quasi-espèce, nous étudions ensuite le système (*) quand la longueur du génome tend vers l'infini. Le cas le plus simple possible est celui du paysage à un pic, que nous traitons dans le chapitre 4. Dans le chapitre 5, nous généralisons les résultats du chapitre 4 à des fonctions de fitness qui dépendent de la classe de Hamming. Mais pourquoi étudier des paysages de fitness qui dépendent de la classe de Hamming ? La distribution de la quasi-espèce, telle qu'Eigen l'avait décrite, est une distribution contenant une concentration positive de la master sequence, et un nuage de mutants proches de la master sequence (au sens de la distance de Hamming). Quand la longueur du génome ℓ tend vers l'infini, le nombre de génotypes à distance de Hamming 1 de la master sequence tend aussi vers l'infini, et si nous regardons la concentration d'un tel génotype, cette concentration tend vers 0. Or, si tous les génotypes à distance 1 de la master sequence partagent une même fitness, nous pouvons les regrouper et étudier la concentration de la classe 1 entière, qui cette fois-ci est positive. L'outil principal qui nous permet de regrouper les différents génotypes dans des classes de Hamming est donné par le lemme suivant.

Lemme 1.3.2. *Soient $b, c \in \{0, \dots, \ell\}$ et soit $u \in \{0, 1\}^\ell$ un génotype dans la classe de Hamming b . La somme*

$$\sum_{\substack{v \in \{0, 1\}^\ell \\ d_H(v, w^*) = c}} M(u, v)$$

ne dépend pas de u dans la classe de Hamming b , c'est une fonction de b et c seulement, que nous notons $M_H(b, c)$. Le coefficient $M_H(b, c)$ est donné par

$$M_H(b, c) = P(b + \text{Bin}(\ell - b, q) - \text{Bin}(b, q) = c).$$

Ce lemme nous permet de dériver un nouveau système d'équations différentielles d'après le système d'Eigen original (*). En effet, pour toute classe de Hamming $k \in \{0, \dots, \ell\}$, appelons $x_k(t)$ la somme des concentrations des génotypes dans la classe k :

$$x_k(t) = \sum_{\substack{u \in \{0,1\}^\ell \\ d_H(u, w^*) = k}} x_u(t).$$

Également, notons par $f_H(k)$ la valeur de la fonction de fitness, commune à tous les génotypes dans la classe k . Les concentrations $x_k(t)$ évoluent selon le système d'équations différentielles suivante :

$$(\mathcal{S}) \quad x'_k(t) = \sum_{i=0}^{\ell} x_i(t) f_H(i) M_H(i, k) - x_k(t) \sum_{i=0}^{\ell} x_i(t) f_H(i), \quad 0 \leq k \leq \ell.$$

L'objectif des chapitres 4 et 5 est d'étudier ce système quand la longueur des génotypes tend vers l'infini et le taux de mutation tend vers 0. Plus précisément, nous considérons le régime asymptotique

$$\ell \rightarrow \infty, \quad q \rightarrow 0, \quad \ell q \rightarrow a \in]0, +\infty[.$$

Nous obtenons alors un noyau de mutation limite donné par : pour tout $b, c \geq 0$,

$$M_\infty(b, c) = \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} M_H(b, c) = \begin{cases} e^{-a} \frac{a^{c-b}}{(c-b)!} & \text{si } b \leq c, \\ 0 & \text{si } b > c. \end{cases}$$

Nous obtenons aussi le système d'équations limite

$$(\mathcal{S}_\infty) \quad x'_k(t) = \sum_{i=0}^k x_i(t) f_H(i) e^{-a} \frac{a^{k-i}}{(k-i)!} - x_k(t) \sum_{i=0}^{\infty} x_i(t) f_H(i), \quad k \geq 0.$$

Nous nous intéressons principalement à trois aspects de ce système : ses solutions stationnaires, leur stabilité, et la relation avec le système (\mathcal{S}) . Dans le cas du paysage à un pic (chapitre 4), nous retrouvons le phénomène de transition de phase prédit par Eigen. En outre, nous trouvons une formule explicite pour la distribution de la quasi-espèce.

Proposition 1.3.3. *Si $\sigma e^{-a} \leq 1$, le système d'Eigen infini (\mathcal{S}_∞) n'a qu'une solution stationnaire :*

$$x_k = 0, \quad k \geq 0.$$

Si $\sigma e^{-\alpha} > 1$, le système d'Eigen infini a deux solutions stationnaires, la solution identiquement nulle et

$$\rho_k^* = (\sigma e^{-\alpha} - 1) \frac{\alpha^k}{k!} \sum_{i=1}^{\infty} \frac{i^k}{\sigma^i}, \quad k \geq 0.$$

La formule ci-dessus définit une distribution de probabilité sur les entiers positifs, que nous appelons la distribution de la quasi-espèce.

Définition 1.3.4. Nous disons qu'une variable aléatoire Y suit la distribution de la quasi-espèce de paramètres σ, α , notée $\mathcal{Q}(\sigma, \alpha)$, si

$$P(Y = k) = (\sigma e^{-\alpha} - 1) \frac{\alpha^k}{k!} \sum_{i=1}^{\infty} \frac{i^k}{\sigma^i}, \quad k \geq 0.$$

Nous regardons ensuite la convergence des trajectoires. Nous montrons qu'au-dessus du seuil d'erreur toutes les solutions de (\mathcal{S}_{∞}) convergent vers 0, et qu'en-dessous du seuil d'erreur, toute solution $(x(t))_{t \geq 0}$ du système (\mathcal{S}_{∞}) vérifiant $x_0(0) > 0$ converge vers ρ^* . Pour compléter le chapitre 4, nous montrons que, si les conditions initiales convergent, les solutions du système (\mathcal{S}) convergent vers les solutions du système infini (\mathcal{S}_{∞}) . Nous montrons aussi que le seul point stationnaire du système (\mathcal{S}) converge vers 0 si $\sigma e^{-\alpha} \leq 1$, et vers ρ^* si $\sigma e^{-\alpha} > 1$.

Nous généralisons ces résultats aux paysages de fitness dépendants de la classe de Hamming dans le chapitre 5. Nous donnons d'abord un résultat permettant de retrouver la fonction de fitness à partir de la distribution de la quasi-espèce, où par distribution de quasi-espèce nous entendons une solution stationnaire $(x_k^*)_{k \geq 0}$ non-négative du système (\mathcal{S}_{∞}) vérifiant $x_0^* > 0$ et $\sum_{k \geq 0} x_k^* = 1$.

Théorème 1.3.5. Supposons que $(x_k^*)_{k \geq 0}$ est une quasi-espèce. Alors,

$$\forall k \geq 1 \quad f_H(k) = \frac{f_H(0)}{x_k^*} \sum_{j=0}^k (-1)^j \frac{\alpha^j}{j!} x_{k-j}^*.$$

Ensuite nous obtenons les phénomènes de seuil d'erreur et de quasi-espèce pour des fonctions de fitness qui dépendent de la classe de Hamming.

Théorème 1.3.6. Nous avons :

- Si $f_H(0)e^{-a} > \limsup_{n \rightarrow \infty} f_H(n)$, alors il existe une distribution de quasi-espèce, donnée par

$$x_0^* = \left(\sum_{\substack{h \geq 0 \\ 0=i_0 < \dots < i_h}} \frac{a^{i_h} f_H(0)}{f_H(i_h)} \prod_{1 \leq t \leq h} \frac{f_H(i_t)}{(i_t - i_{t-1})!(f_H(0) - f_H(i_t))} \right)^{-1},$$

où le produit vide est égal à 1, et pour tout $k \geq 1$, nous avons

$$x_k^* = \frac{a^k}{f_H(k)} \times \sum_{\substack{1 \leq h \leq k \\ 0=i_0 < \dots < i_h=k}} \frac{\prod_{t=1}^h \frac{f_H(i_t)}{(i_t - i_{t-1})!(f_H(0) - f_H(i_t))}}{\sum_{\substack{h \geq 0 \\ 0=i_0 < \dots < i_h}} \frac{a^{i_h}}{f_H(i_h)} \prod_{1 \leq t \leq h} \frac{f_H(i_t)}{(i_t - i_{t-1})!(f_H(0) - f_H(i_t))}}.$$

- Si $f_H(0)e^{-a} < \liminf_{n \rightarrow \infty} f_H(n)$, alors il n'existe pas de distribution de quasi-espèce.

Nous voyons que le phénomène de transition de phase ne dépend que de la fitness de la master sequence et du comportement de la fonction de fitness à l'infini. Nous obtenons, comme pour le paysage à un pic, une formule explicite pour la distribution de la quasi-espèce. Cette formule est d'une grande richesse combinatoire, et peut être exprimée notamment à l'aide des coefficients up-down. Soit $\sigma = (\sigma(1), \dots, \sigma(n))$ une permutation de $\{1, \dots, n\}$. La signature de Niven associée à σ est la suite $(q_1, \dots, q_{n-1}) \in \{-1, +1\}^{n-1}$ telle que

$$\forall i \in \{1, \dots, n-1\} \quad q_i(\sigma(i+1) - \sigma(i)) > 0.$$

Définition 1.3.7. Soient $n \geq 1$, $0 \leq h < n$ et $0 = i_0 < i_1 < \dots < i_h < n$. Le coefficient up-down

$$\left\{ \begin{matrix} n \\ i_0, i_1, \dots, i_h \end{matrix} \right\}$$

est le nombre de permutations de $\{1, \dots, n\}$ ayant la signature de Niven (q_1, \dots, q_{n-1}) donnée par

$$\forall i \in \{1, \dots, n-1\} \quad q_i = \begin{cases} 1 & \text{si } i \in \{i_1, \dots, i_h\}, \\ -1 & \text{sinon.} \end{cases}$$

Pour $k \geq 1$, écrivons $y_k^* = x_k^*/x_0^*$. Nous avons l'expression suivante pour les y_k^* .

Théorème 1.3.8. *Pour tout $k \geq 1$, nous avons*

$$y_k^* = \frac{a^k}{k!} \left(\prod_{j=1}^k \frac{f_H(0)}{f_H(0) - f_H(j)} \right) \sum_{\substack{0 \leq h < k \\ 0=i_0 < \dots < i_h < k}} \left\{ \begin{matrix} k \\ i_0, \dots, i_h \end{matrix} \right\} \prod_{t=1}^h \frac{f_H(i_t)}{f_H(0)}.$$

Pour des fonctions de fitness constantes à partir d'un certain rang, nous pouvons même exprimer les y_k^* en fonction de la quasi-espèce pour le paysage à un pic.

Théorème 1.3.9. *Supposons qu'il existe un entier $N \geq 0$ tel que la fonction de fitness $f_H : \mathbb{N} \rightarrow \mathbb{R}^+$ vérifie*

$$f_H(N) \neq 1 \quad \text{et} \quad \forall k > N \quad f_H(k) = 1.$$

Alors, pour tout $k > N$,

$$y_k^* = \rho_k^* + \sum_{j=1}^N \frac{a^j}{j!} \rho_{k-j}^* \left(\frac{f_H(j)}{f_H(0) - f_H(j)} - \frac{1}{f_H(0) - 1} \right) \\ \times \left(1 + \sum_{h=1}^{j-1} \sum_{0=i_0 < \dots < i_h < j} \prod_{t=1}^h \binom{j - i_{t-1}}{j - i_t} \frac{f_H(i_t)}{f_H(0) - f_H(i_t)} \right).$$

Nous complétons le chapitre 5 avec les analogues des résultats du chapitre 4 sur la stabilité des solutions stationnaires du système (S_∞) et la convergence des trajectoires solutions de (S) vers les solutions de (S_∞) , ainsi que la convergence de la solution stationnaire de (S) vers celle de (S_∞) .

1.4 Le modèle de Galton–Watson

Proposé par Galton et Watson pour étudier l'extinction des noms de famille au sein de la noblesse anglaise [39], le modèle de Galton–Watson est l'un des premiers modèles mathématiques pour l'évolution d'une population finie. Le processus de Galton–Watson est mathématiquement très riche, et a donné lieu à de nombreuses recherches. D'excellents livres y sont consacrés, par exemple [3, 42, 47]. Le lien entre le processus de Galton–Watson et le modèle d'Eigen a été exploré par Demetrius, Schuster et Sigmund dans [19].

Le modèle de Galton–Watson possède de nombreuses versions. Nous considérons dans la partie II un processus de Galton–Watson à plusieurs types, avec mutation et sélection. Il s’agit d’un processus de Markov à temps discret

$$X_n = (X_n(0), \dots, X_n(\ell)), \quad n \geq 0.$$

La quantité $X_n(k)$ compte le nombre d’individus appartenant à la classe de Hamming k à la génération n . À chaque génération n , chaque individu donne naissance à un nombre aléatoire d’enfants, distribué selon une variable aléatoire de Poisson de paramètre la fitness de l’individu, indépendamment des autres individus et du passé du processus. Ensuite, chaque enfant mute selon le noyau de mutation M_H , et l’ensemble des enfants après mutation forme la génération $n + 1$.

Le processus de Galton–Watson est construit et étudié de façon formelle à l’aide des fonction génératrices. Pour $x \in \mathbb{N}^{\ell+1}$ et $0 \leq k \leq \ell$, nous définissons la probabilité $p^k(x)$ qu’un individu dans la classe k donne naissance à un ensemble d’enfants distribué selon le vecteur x :

$$p^k(x) = e^{-f_H(k)} f_H(k)^{|x|_1} \prod_{h=0}^{\ell} \frac{M_H(k, h)^{x(h)}}{x(h)!}.$$

Définissons, pour $k \in \{0, \dots, \ell\}$ et $s \in [0, 1]^{\ell+1}$,

$$g^k(s) = \sum_{x \in \mathbb{N}^{\ell+1}} p^k(x) \prod_{l=0}^{\ell} s(l)^{x(l)}.$$

Le mécanisme de transition du processus $(X_n)_{n \geq 0}$ est donné par : pour tout $n \geq 0$ et $x \in \mathbb{N}^{\ell+1}$, si $X_n = x$, alors X_{n+1} est la somme de $|x|_1$ vecteurs aléatoires indépendants ; dans cette somme, pour chaque $k \in \{0, \dots, \ell\}$, $x(k)$ vecteurs ont pour fonction génératrice g^k . L’état 0 est bien sûr un état absorbant. Quand le processus $(X_n)_{n \geq 0}$ n’est pas absorbé, nous disons qu’il survit.

Considérons la fonction de fitness du paysage à un pic :

$$\forall k \in \{0, \dots, \ell\}, \quad f_H(k) = \begin{cases} \sigma & \text{si } k = 0, \\ 1 & \text{si } k \geq 1. \end{cases}$$

Le résultat principal du chapitre 7 est le suivant.

Théorème 1.4.1. *Le processus de Galton–Watson $(X_n)_{n \geq 0}$ a une probabilité positive de survie. Conditionné à la survie :*

- Si $\sigma e^{-a} \leq 1$, alors

$$\forall k \geq 0 \quad \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \lim_{n \rightarrow \infty} \frac{X_n(k)}{X_n(0) + \dots + X_n(\ell)} = 0.$$

- Si $\sigma e^{-a} > 1$, alors

$$\forall k \geq 0 \quad \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \lim_{n \rightarrow \infty} \frac{X_n(k)}{X_n(0) + \dots + X_n(\ell)} = (\sigma e^{-a} - 1) \frac{a^k}{k!} \sum_{i \geq 1} \frac{i^k}{\sigma^i}.$$

La démonstration de ce théorème est faite en deux parties. D’abord, suivant les lignes de [19], nous montrons que la limite

$$\lim_{n \rightarrow \infty} \frac{X_n(k)}{X_n(0) + \dots + X_n(\ell)}$$

converge vers la seule solution stationnaire du système d’Eigen (fini) associé. Nous utilisons ensuite les résultats de la section 4.4 pour montrer la convergence de la deuxième limite. Nous complétons le chapitre 7 avec le résultat analogue pour des paysages de fitness dépendants de la classe de Hamming.

Remarquons que dans le modèle de Galton–Watson, quand nous conditionnons le processus à la survie, la population croît exponentiellement. La façon dont nous avons pris les limites dans le théorème, entraîne que, avant de prendre la deuxième limite, la taille de la population est déjà infinie. Le processus de Galton–Watson ne permet donc pas de mesurer l’impact de la taille de la population dans l’émergence d’une quasi-espèce. Il constitue tout de même un premier pas pour montrer la validité du modèle d’Eigen pour des populations finies, surtout s’il s’agit de populations dans leurs premières générations, car elles ont tendance à croître exponentiellement.

1.5 Le modèle de Wright–Fisher

Dans les années 30, Fisher [36] et Wright [86] proposent un modèle pour étudier la variation d’un allèle dans une population finie. Le modèle de Wright–Fisher

est probablement le modèle mathématique le plus célèbre en génétique des populations. Contrairement au modèle de Galton–Watson, il s’agit d’un modèle dans lequel la taille de la population reste constante au cours du temps.

Le modèle de Wright–Fisher original peut être étendu pour intégrer plusieurs types différents, la sélection et la mutation. Le processus de Wright–Fisher que nous considérons dans la partie III est une chaîne de Markov

$$X_n = (X_n(0), \dots, X_n(\ell)), \quad n \geq 0.$$

La quantité $X_n(k)$ compte le nombre d’individus appartenant à la classe de Hamming k à la génération n . Notons par $m \geq 1$ la taille de la population. Puisque la taille de la population m reste constante au cours du temps, le vecteur $(X_n)_{n \geq 0}$ est un élément de l’espace

$$\mathcal{P}_{\ell+1}^m = \{ (x(0), \dots, x(\ell)) \in \mathbb{N}^{\ell+1} : x(0) + \dots + x(\ell) = m \}.$$

Dans ce modèle, le mécanisme de transition d’une génération à la suivante se décompose en deux étapes. D’abord, nous échantillonnons avec remise m chromosomes de la population actuelle X_n . L’échantillonnage est fait à l’aide de la fonction de fitness f_H . Puis, chaque chromosome échantillonné mute selon la loi donnée par la matrice de mutation M_H . L’ensemble des m individus après mutation forme la génération $n + 1$. Il s’agit donc d’un modèle dans lequel les différentes générations ne se superposent pas. La matrice de transition p du processus de Wright–Fisher est donnée par

$$\forall x, x' \in \mathcal{P}_{\ell+1}^m \quad p(x, x') = \prod_{0 \leq h \leq \ell} \left(\frac{\sum_{k \in \{0, \dots, \ell\}} x(k) f_H(k) M_H(k, h)}{\sum_{0 \leq h \leq \ell} x(h) f_H(h)} \right)^{x'(h)}.$$

Le processus de Wright–Fisher est une chaîne de Markov ergodique, il admet donc une unique mesure de probabilité invariante. Nous nous intéressons aux quantités

$$\lim_{n \rightarrow \infty} E \left(\frac{X_n(k)}{m} \right),$$

qui sont bien définies grâce à l’ergodicité de la chaîne de Markov. Puisque nous avons un nouveau paramètre dans notre modèle, la taille de la population m , nous sommes amenés à considérer un nouveau régime asymptotique tenant compte de ce nouveau paramètre. Le régime asymptotique que nous regardons est

$$\begin{aligned} \ell &\rightarrow +\infty, & m &\rightarrow +\infty, & q &\rightarrow 0, \\ \ell q &\rightarrow a \in]0, +\infty[, & \frac{m}{\ell} &\rightarrow \alpha \in [0, +\infty], \end{aligned}$$

Nous ne considérons que le paysage à un pic. Notons par $I(p, t)$ la fonction de taux gouvernant le principe de grandes déviations de la loi binomiale de paramètre $p \in [0, 1]$, i.e.,

$$\forall t \in [0, 1] \quad I(p, t) = t \ln \frac{t}{p} + (1-t) \ln \frac{1-t}{1-p}.$$

Nous définissons, pour $a \in]0, +\infty[$,

$$\rho^*(a) = \begin{cases} \frac{\sigma e^{-a} - 1}{\sigma - 1} & \text{si } \sigma e^{-a} > 1, \\ 0 & \text{si } \sigma e^{-a} \leq 1, \end{cases}$$

$$\psi(a) = \inf_{l \in \mathbb{N}} \inf \left\{ \sum_{k=0}^{l-1} I\left(\frac{\sigma \rho_k}{(\sigma-1)\rho_k + 1}, \gamma_k\right) + \gamma_k I\left(e^{-a}, \frac{\rho_{k+1}}{\gamma_k}\right) : \right. \\ \left. \rho(0) = \rho^*(a), \rho_l = 0, \rho_k, \gamma_k \in [0, 1] \text{ pour } 0 \leq k \leq l \right\}.$$

Théorème 1.5.1. *Nous avons la dichotomie suivante :*

- Si $\alpha \psi(a) < \ln 2$, alors

$$\forall k \geq 0 \quad \lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \lim_{n \rightarrow \infty} \mathbb{E} \left(\frac{X_n(k)}{m} \right) = 0.$$

- Si $\alpha \psi(a) > \ln 2$, alors

$$\forall k \geq 0 \quad \lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \lim_{n \rightarrow \infty} \mathbb{E} \left(\frac{X_n(k)}{m} \right) = (\sigma e^{-a} - 1) \frac{a^k}{k!} \sum_{i \geq 1} \frac{i^k}{\sigma^i}.$$

De plus, dans les deux cas,

$$\forall k \geq 0 \quad \lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \lim_{n \rightarrow \infty} \text{Var} \left(\frac{X_n(k)}{m} \right) = 0.$$

L'idée d'utiliser le modèle de Wright-Fisher pour retrouver un phénomène de seuil d'erreur et une distribution de quasi-espèce vient de [11]. Dans [11], Cerf

démontre qu'il existe une courbe (la courbe $\alpha\psi(\alpha) = \ln 2$ du théorème) séparant un régime de désordre d'un régime ordonné. Dans le régime de désordre, la concentration de la master sequence tend vers 0, tandis que dans le régime ordonné, la concentration de la master sequence tend vers $(\sigma e^{-\alpha} - 1)/(\sigma - 1)$. Nous étendons le résultat de [11] à toutes les classes de Hamming $k \geq 0$. La démonstration repose sur des arguments de couplage et de monotonie, et utilise la théorie développée par Freidlin et Wentzell pour des perturbations aléatoires de systèmes dynamiques [38].

1.6 Le modèle de Moran

Dans la quatrième et dernière partie de la thèse, nous étudions le modèle de Moran. C'est une modification du modèle de Wright–Fisher, introduite par Moran [57] dans les années 50, qui permet de modéliser des générations qui se superposent. De même que dans le modèle de Wright–Fisher, dans le modèle de Moran la taille de la population est finie et reste constante au cours de l'évolution de la population.

Nous considérons un modèle de Moran à plusieurs types, avec sélection et mutation. Le processus de Moran est une chaîne de Markov à temps discret

$$X_n = (X_n(0), \dots, X_n(\ell)), \quad n \geq 0,$$

à valeurs dans $\mathcal{P}_{\ell+1}^m$. Soit $\lambda > f_H(0)$. Le mécanisme de transition pour passer de X_n à X_{n+1} est le suivant : un individu est choisi dans la population X_n uniformément au hasard, supposons qu'il appartienne à la classe k . Avec probabilité $1 - f_H(k)/\lambda$, rien ne se passe, et $X_{n+1} = X_n$. Avec probabilité $f_H(k)/\lambda$, l'individu choisi se reproduit, son enfant mute selon le noyau M_H , et remplace un individu dans la population X_n , choisi uniformément au hasard. Nous notons par $(e_i)_{0 \leq i \leq \ell}$ la base canonique de $\mathbb{R}^{\ell+1}$, indexée par $0, \dots, \ell$ au lieu de $1, \dots, \ell + 1$. Les seuls changements autorisés dans le processus de Moran sont du type

$$x \longrightarrow x - e_i + e_j, \quad 0 \leq i, j \leq \ell.$$

La matrice de transition du processus est donnée par

$$\forall x \in \mathcal{P}_{\ell+1}^m \quad \forall i, j \in \{0, \dots, \ell\} \quad i \neq j$$

$$p(x, x - e_i + e_j) = \frac{x(i)}{\lambda m^2} \sum_{h=0}^{\ell} x(h) f_H(h) M_H(h, j).$$

Les autres coefficients non diagonaux de la matrice de transition sont nuls. Les coefficients diagonaux sont choisis de sorte que la matrice soit stochastique, i.e., la somme de chaque ligne vaut 1. Le processus de Moran est une chaîne de Markov ergodique, il admet donc une unique mesure de probabilité invariante. Comme dans le processus de Wright–Fisher, nous nous intéressons aux quantités

$$\lim_{n \rightarrow \infty} \mathbb{E} \left(\frac{X_n(k)}{m} \right),$$

qui sont bien définies grâce à l’ergodicité de la chaîne de Markov. Nous considérons encore le régime asymptotique

$$\begin{aligned} \ell &\rightarrow +\infty, & m &\rightarrow +\infty, & q &\rightarrow 0, \\ \ell q &\rightarrow \alpha \in]0, +\infty[, & \frac{m}{\ell} &\rightarrow \alpha \in [0, +\infty]. \end{aligned}$$

Supposons d’abord que f_H est la fonction de fitness du paysage à un pic. Soit $\psi : \mathbb{R}^+ \rightarrow \mathbb{R}^+ \cup \{+\infty\}$ la fonction définie par

$$\forall \alpha < \ln \sigma \quad \psi(\alpha) = \frac{\sigma(1 - e^{-\alpha}) \ln \frac{\sigma(1 - e^{-\alpha})}{\sigma - 1} + \ln(\sigma e^{-\alpha})}{(1 - \sigma(1 - e^{-\alpha}))},$$

et $\psi(\alpha) = 0$ si $\alpha \geq \ln \sigma$. Nous avons le résultat suivant.

Théorème 1.6.1. *Nous avons la dichotomie suivante :*

- Si $\alpha \psi(\alpha) < \ln 2$, alors

$$\forall k \geq 0 \quad \lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha, \frac{m}{\ell} \rightarrow \alpha}} \lim_{n \rightarrow \infty} \mathbb{E} \left(\frac{X_n(k)}{m} \right) = 0.$$

- Si $\alpha \psi(\alpha) > \ln 2$, alors

$$\forall k \geq 0 \quad \lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha, \frac{m}{\ell} \rightarrow \alpha}} \lim_{n \rightarrow \infty} \mathbb{E} \left(\frac{X_n(k)}{m} \right) = (\sigma e^{-\alpha} - 1) \frac{\alpha^k}{k!} \sum_{i \geq 1} \frac{i^k}{\sigma^i}.$$

De plus, dans les deux cas,

$$\forall k \geq 0 \quad \lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha, \frac{m}{\ell} \rightarrow \alpha}} \lim_{n \rightarrow \infty} \text{Var} \left(\frac{X_n(k)}{m} \right) = 0.$$

L'idée d'utiliser le modèle de Moran pour retrouver un phénomène de seuil d'erreur et une distribution de quasi-espèce vient de [10]. Dans [10], Cerf démontre qu'il existe une courbe (la courbe $\alpha\psi(a) = \ln 2$ du théorème) séparant un régime de désordre d'un régime ordonné. Dans le régime de désordre, la concentration de la master sequence tend vers 0, tandis que dans le régime ordonné, la concentration de la master sequence tend vers $(\sigma e^{-a} - 1)/(\sigma - 1)$. Nous étendons le résultat de [10] à toutes les classes de Hamming $k \geq 0$. La démonstration est basée sur des arguments de couplage et de monotonie, ainsi que sur un argument de récurrence sur les coordonnées qui nous permet de traiter chaque coordonnée comme une chaîne de naissance et mort markovienne.

Nous considérons ensuite des paysages de fitness constantes à partir d'un certain rang $K \geq 0$, i.e., des fonctions de fitness vérifiant :

$$f_H(0) > \max_{1 \leq k \leq \ell} f_H(k), \quad f_H(K) \neq 1, \quad \forall k \in \{K+1, \dots, \ell\} \quad f_H(k) = 1.$$

Sous cette hypothèse, nous développons un principe de grandes déviations trajectorielle pour le processus de Moran. Soit $k \geq K$ et définissons le processus $(Y_t)_{t \geq 0}$ par :

$$Y_t = \frac{1}{m} (X_{[tm]}(0), \dots, X_{[tm]}(k)), \quad t \geq 0.$$

L'espace d'états du processus $(Y_t)_{t \geq 0}$ est :

$$\Delta^{k+1} = \{ y \in \{0, 1/m, \dots, 1\}^{k+1} : y_0 + \dots + y_k \leq 1 \}.$$

Puisque notre objectif est de faire tendre m vers l'infini, nous introduisons

$$\mathcal{D}^{k+1} = \{ x \in [0, 1]^{k+1} : x_0 + \dots + x_k \leq 1 \}.$$

Soit \mathcal{V} l'ensemble des vecteurs de \mathbb{R}^{k+1} associés aux possibles directions des sauts du processus $(Y_t)_{t \geq 0}$:

$$\mathcal{V} = \{ -e_i : 0 \leq i \leq k \} \cup \{ e_j : 0 \leq j \leq k \} \cup \{ -e_i + e_j : 0 \leq i, j \leq k, i \neq j \}.$$

Définissons les quantités suivantes : pour tout $y \in \mathcal{D}^{k+1}$ et $0 \leq i, j \leq k$,

$$C_y^\infty(-e_i) = \frac{1}{\lambda} y_i \left(\sum_{0 \leq h \leq k} y_h f_H(h) \left(1 - \sum_{h \leq l \leq k} e^{-a} \frac{a^{l-h}}{(l-h)!} \right) + 1 - \sum_{0 \leq h \leq k} y_h \right),$$

$$C_y^\infty(e_j) = \frac{1}{\lambda} \left(1 - \sum_{0 \leq l \leq k} y_l \right) \sum_{0 \leq h \leq j} y_h f_H(h) e^{-a} \frac{a^{j-h}}{(j-h)!},$$

$$C_y^\infty(-e_i + e_j) = \frac{1}{\lambda} y_i \sum_{0 \leq h \leq j} y_h f_H(h) e^{-a} \frac{a^{j-h}}{(j-h)!}.$$

Les quantités $C_y(v)$ correspondent aux probabilités de transition asymptotiques pour le processus $(Y_t)_{t \geq 0}$ quand ℓ tend vers l'infini, q tend vers 0 et ℓq tend vers a . Soit $G^\infty : \mathcal{D}^{k+1} \times \mathbb{R}^{k+1} \rightarrow \mathbb{R}$ la fonction définie par

$$G^\infty(y; z) = \ln \left(1 + \sum_{v \in \mathcal{V}} C_y^\infty(v) (e^{v \cdot z} - 1) \right),$$

où $v \cdot z$ représente le produit scalaire dans \mathbb{R}^{k+1} . Soit $H^\infty : \mathcal{D}^{k+1} \times \mathbb{R}^{k+1} \rightarrow \mathbb{R}^+$ la transformée de Fenchel–Legendre de la fonction $G^\infty(y, \cdot)$:

$$H^\infty(y; u) = \sup_{z \in \mathbb{R}^{k+1}} (u \cdot z - G^\infty(y; z)).$$

Soit $T > 0$ et considérons une trajectoire $\phi : [0, T] \rightarrow \mathcal{D}^{k+1}$. Nous définissons l'action $I_T(\phi)$ de la trajectoire ϕ par

$$I_T(\phi) = \int_0^T H^\infty(\phi(t), \phi'(t)) dt,$$

si la fonction ϕ est absolument continue et l'intégrale converge ; nous posons $I_T(\phi) = +\infty$ sinon. Nous notons par d la distance euclidienne dans \mathbb{R}^{k+1} , et nous notons par d_T la distance du supremum entre deux trajectoires sur l'intervalle $[0, T]$, i.e., pour $\phi, \psi : [0, T] \rightarrow \mathbb{R}^{k+1}$,

$$d_T(\phi, \psi) = \sup_{0 \leq t \leq T} d(\phi(t), \psi(t)).$$

Soit $\rho > 0$, nous définissons l'ensemble ρ -réduit de \mathcal{D}^{k+1} par

$$\mathcal{D}_\rho^{k+1} = \{ y \in \mathcal{D}^{k+1} : d(y, (\mathcal{D}^{k+1})^c) \geq \rho \},$$

et nous posons $\Delta_\rho^{k+1} = \Delta \cap \mathcal{D}_\rho^{k+1}$. Pour un point $y^0 \in \mathcal{D}_\rho^{k+1}$ et $\eta \geq 0$, notons par $\Phi_{y^0; T}(\eta)$ l'ensemble des trajectoires ϕ dans \mathcal{D}_ρ^{k+1} issues du point y^0 et dont l'action est inférieure ou égale à η , i.e.,

$$\Phi_{y^0; T}(\eta) = \{ \phi : [0, T] \rightarrow \mathcal{D}_\rho^{k+1} : \phi(0) = y^0 \text{ et } I_T(\phi) \leq \eta \}.$$

Soit T_ρ le temps de sortie du processus $(Y_t)_{t \geq 0}$ de l'ensemble \mathcal{D}_ρ^{k+1} , i.e.,

$$T_\rho = \inf \{ t \geq 0 : Y(t) \notin \mathcal{D}_\rho^{k+1} \}.$$

Nous considérons le processus $(Y^\rho(t), t \geq 0)$ arrêté au temps T_ρ :

$$\forall t \geq 0 \quad Y^\rho(t) = Y(t \wedge T_\rho).$$

Théorème 1.6.2 (PGD). *Nous avons le résultat suivant :*

Borne inférieure. *Pour tout $\delta, \gamma, \rho, \eta_0, T > 0$ avec $\rho > \delta$, asymptotiquement,*

$$P\left(d_T(Y, \phi) < \delta \mid Y(0) = y^0\right) \geq \exp(-m(I_T(\phi) + \gamma)),$$

uniformément sur tout $y^0 \in \Delta_\rho^{k+1}$, et sur tout $\phi \in \Phi_{y^0; T}(\eta_0)$.

Borne supérieure. *Pour tout $\delta, \gamma, \rho, \eta_0, T > 0$, asymptotiquement,*

$$P\left(d_T(Y^\rho, \Phi_{y^0; T}(\eta)) \geq \delta \mid Y^\rho(0) = y^0\right) \leq \exp(-m(\eta - \gamma)),$$

uniformément sur $y^0 \in \Delta_\rho^{k+1}$ et $\eta \leq \eta_0$.

Ce principe de grandes déviations est présenté et démontré dans le chapitre 22. Le résultat est inspiré de la théorie développée par Freidlin et Wentzell pour des perturbations aléatoires de systèmes dynamiques [38]. Des résultats de même nature peuvent être trouvés dans plusieurs travaux dus à Wentzell [82, 83, 84]. Notre situation diffère du cadre traité dans les références données par plusieurs aspects : nous considérons un régime asymptotique qui mêle plusieurs paramètres, au lieu d'un seul, et nous considérons un processus qui n'est pas markovien, mais seulement « asymptotiquement markovien ». De plus, notre processus est à valeurs dans un espace avec bord. Au vu de toutes ces particularités, nous avons opté pour démontrer le principe de grandes déviations dans le cas concret que nous traitons. Même si plusieurs détails changent, les idées de la démonstration restent les mêmes que dans les travaux de Wentzell. La borne inférieure des grandes déviations est démontrée à l'aide d'un changement exponentiel de mesure. La probabilité que nous voulons estimer est très grande sous la nouvelle mesure, qui a une densité sous forme exponentielle vis-à-vis de l'ancienne mesure, la difficulté est donc ramenée au contrôle de l'argument de l'exponentielle donnant la densité. Pour montrer la borne supérieure, nous approximons la trajectoire du processus par une trajectoire linéaire par morceaux, et nous montrons qu'il est très improbable que cette trajectoire soit loin des trajectoires de petite action.

Les seules trajectoires ϕ telles que l'action $I_T(\phi)$ vaut 0 sont les solutions du système d'Eigen limite (\mathcal{S}_∞) . Il s'ensuit que le processus $(Y_t)_{t \geq 0}$ suit les solutions du système (\mathcal{S}_∞) avec très grande probabilité. Nous exploitons ce fait dans le chapitre 23 pour estimer le temps moyen de sortie d'un domaine ouvert contenant la distribution de la quasi-espèce $(x_i^*)_{0 \leq i \leq k}$. Ce temps moyen est d'ordre exponentiel en la taille de la population m .

Pour finir, nous revenons au cas du paysage à un pic dans le chapitre 24. Dans ce cas, la dynamique du nombre de master sequences est "asymptotiquement markovienne", et nous pouvons appliquer le principe de grandes déviations en une dimension. Cela rend possible le calcul de la fonctionnelle d'action ainsi que du quasi-potentiel associé. Ceci nous permet de décrire la trajectoire de disparition de la quasi-espèce.

1.7 Conclusion et perspectives

Nous avons étudié les modèles de Galton–Watson, Wright–Fisher et Moran avec l'objectif de retrouver un phénomène de transition de phase et une distribution de quasi-espèce. Nous avons vu que les trois modèles présentent un phénomène de transition de phase, et que dans les modèles à population constante, Wright–Fisher et Moran, cette transition de phase fait intervenir la taille de la population. Même si la transition de phase se présente sous une forme différente en fonction du modèle, quand une quasi-espèce est créée, la distribution de cette quasi-espèce est la même pour les trois modèles, et coïncide aussi avec la distribution de la quasi-espèce pour le modèle d'Eigen.

L'obtention d'une formule explicite pour la distribution de la quasi-espèce dans le cas du paysage à un pic nous a menés à la recherche de formules explicites pour des paysages de fitness plus généraux. Ainsi, nous avons réussi à obtenir, pour le modèle d'Eigen, une formule explicite pour tout paysage de fitness dépendant de la classe de Hamming.

En ce qui concerne le modèle de Moran, nous espérons retrouver l'analogie du théorème 1.6.1 pour des paysages de fitness dépendants de la classe de Hamming et constants à partir d'un certain rang, à l'aide du principe de grandes déviations. Nous espérons aussi pouvoir faire de même pour le processus de Wright–Fisher.

1.8 Publications

Les résultats obtenus pendant le déroulement de cette thèse ont donné lieu à plusieurs articles de recherche. Les résultats du chapitre 5 concernant les dif-

férentes formules combinatoires pour la distribution de la quasi-espèce sur des fonctions de fitness qui dépendent de la classe de Hamming ont été publiés dans [13]. Le résultat principal pour le processus de Galton–Watson sur le paysage à un pic (chapitre 7) a été publié dans [18]. Le résultat principal de la partie III sur le modèle de Wright–Fisher a été publié dans l’article [17]. La note [16] montre la convergence du modèle de Moran vers le modèle d’Eigen (section 15.5), et le résultat principal concernant le modèle de Moran et le paysage à un pic, contenu dans les chapitres 15–21, a été publié dans [14]. Enfin, l’article [12] présente différentes représentations probabilistes de la distribution de la quasi-espèce. Les résultats de ce dernier article sont liés au sujet de cette thèse, cependant ils sont plus éloignés des modèles à population finies traités ici ; nous avons donc décidé de les omettre.

Première partie

Le modèle d'Eigen

Chapitre 2

Introduction

Selon la théorie initiée par Oparin [62] et Haldane [41] dans les années 20, les premières formes de vie sur terre ont dû être des macromolécules d'une grande simplicité. Avec cette considération en tête, Manfred Eigen se lance dans les années 70 dans la construction d'un cadre théorique afin de mieux comprendre les premières étapes de la vie terrestre [28]. Une petite partie de son travail porte sur la modélisation mathématique de l'évolution d'une population de macromolécules. La simplicité de ces macromolécules permet de décrire l'évolution d'une telle population par des réactions cinétiques et chimiques, à partir desquelles Eigen dérive un système d'équations différentielles. Il s'agit d'un système décrivant l'évolution d'une population sous deux forces principales : sélection et mutation.

Le système d'Eigen est l'un des premiers modèles biologiques à présenter un phénomène de transition de phase. En effet, quand les macromolécules sont longues, il existe un taux de mutation critique au-dessus duquel l'information génétique est perdue après quelques générations. Le taux de mutation critique est aussi appelé le seuil d'erreur. Le modèle d'Eigen présente un deuxième phénomène d'importance : la quasi-espèce. Pour des taux de mutation en-dessous du seuil d'erreur, une population à l'équilibre contient une concentration positive de la macromolécule la plus adaptée, la master sequence, ainsi qu'un nuage de mutants qui diffèrent de la master sequence de quelques mutations. Contrairement au cas d'une espèce « pure », il s'agit ici d'une population dans laquelle la master sequence peut être présente avec une concentration faible, et le nuage de mutants acquiert une énorme importance.



Manfred Eigen

L'efficacité des distributions de quasi-espèce pour décrire des populations de virus a attiré une grande attention sur le modèle d'Eigen [22, 23, 69], et la possible existence d'un seuil d'erreur pour des populations de virus a donné lieu à de nombreuses recherches pour induire une catastrophe d'erreur dans des populations de virus, afin de combattre diverses maladies virales [2, 15, 79]. La définition du modèle d'Eigen est remarquablement simple, cependant son analyse mathématique est ardue. Les deux questions principales qui se posent sont, la localisation du seuil d'erreur, et la description de la distribution de la quasi-espèce. Notamment, nous voudrions trouver la concentration de la master sequence et des différents mutants dans une population à l'équilibre.

Nous ne sommes pas les premiers à nous intéresser à ces questions. Un des premiers travaux sur le sujet est dû à Eigen lui-même, McCaskill et Schuster [30]. Dans ce travail, on s'aperçoit déjà que la description de la distribution de la quasi-espèce n'est pas en général une tâche simple. En effet, une telle description n'est possible qu'après avoir fait certaines approximations ou simplifications dans le modèle. Thompson et McBride [78], ainsi que Jones, Enns et Rangnekar [49] et Swetina et Schuster [76], réussissent à linéariser le système d'équations différentielles d'Eigen. Ils se servent ensuite de cette linéarisation pour caractériser la distribution de la quasi-espèce comme le vecteur propre associé à la plus grande valeur propre de la matrice du système linéarisé. Les solutions ainsi obtenues reçoivent le nom de solutions exactes. Des solutions

exactes pour des paysages de fitness particuliers ont aussi été obtenues par d'autres moyens par Saakian et Hu [68], Saakian [65], ou Saakian, Biebricher et Hu [66]. Également Novozhilov et Semenov [71, 72] et Bratus, Novozhilov et Semenov [7, 70] obtiennent des résultats plus concrets pour des choix particuliers de paysages de fitness et de noyaux de mutation.

Le contenu de cette première partie est une synthèse des résultats déjà connus pour le modèle d'Eigen, et de notre contribution au sujet. Nous commençons par introduire le modèle dans le chapitre 3, avec quelques considérations préliminaires. Nous montrons ensuite comment linéariser le système pour obtenir les premières représentations de la distribution de la quasi-espèce, suivant les lignes des travaux [49, 76, 78]. Finalement, nous présentons le régime asymptotique permettant de retrouver le seuil d'erreur. Ce régime asymptotique donne lieu à un système d'équations différentielles infini, qui sera l'objet d'étude principal des chapitres 4 et 5. Dans le chapitre 4 nous considérons le cas du paysage de fitness à un pic, c'est-à-dire, un seul génotype plus adapté dans un paysage neutre. Sous cette hypothèse, nous obtenons une formule explicite pour la distribution de la quasi-espèce. Nous justifions aussi le passage à la limite. Dans le chapitre 5, les résultats du chapitre 4 sont généralisés à des paysages de fitness plus généraux : ceux qui attribuent la même fitness à tous les génotypes qui diffèrent de la master sequence d'un même nombre de mutations.

Chapitre 3

Généralités

3.1 Le modèle

Le modèle d'Eigen a pour but de modéliser l'évolution au cours du temps d'une population de macromolécules. Nous allons d'abord introduire tous les éléments nécessaires pour pouvoir définir le modèle.

Les génotypes. Soit \mathcal{A} un alphabet fini de cardinal $\kappa \geq 2$. Nous identifions les macromolécules ou individus avec des chaînes de longueur $\ell \geq 1$ sur l'alphabet \mathcal{A} et nous appelons une telle chaîne un *génotype*. L'espace des génotypes est donc \mathcal{A}^ℓ . Des choix typiques pour \mathcal{A} sont l'ensemble des 20 aminoacides pour des protéines, $\{A, T, G, C\}$ pour des séquences d'ADN, ou $\{0, 1\}$ pour des séquences binaires.

La sélection. La sélection est donnée par un *paysage de fitness*, i.e., une fonction

$$f : \mathcal{A}^\ell \longrightarrow \mathbb{R}^+.$$

Pour une séquence donnée $u \in \mathcal{A}^\ell$, sa fitness $f(u)$ représente le taux auquel un individu ayant le génotype u se reproduit. Une fonction constante f correspond à un paysage de sélection neutre, où tous les individus se reproduisent au même taux. Le paysage non neutre le plus simple auquel nous pouvons penser est le *paysage à un pic*, i.e., tous les génotypes ont le même taux de reproduction sauf un génotype privilégié $w^* \in \mathcal{A}^\ell$, appelé la *master sequence*, qui a un taux de reproduction supérieur à tous les autres génotypes. Soit $\sigma > 1$, la fonction de

fitness f du paysage à un pic est donnée par

$$\forall \mathbf{u} \in \mathcal{A}^\ell, \quad f(\mathbf{u}) = \begin{cases} \sigma & \text{si } \mathbf{u} = \mathbf{w}^*, \\ 1 & \text{si } \mathbf{u} \neq \mathbf{w}^*. \end{cases}$$

Les mutations. Lors d'une reproduction, une nouvelle chaîne est créée par répllication du génome du parent, cependant, des erreurs peuvent arriver au cours de la répllication, c'est ce qu'on appelle les mutations. Nous supposons que les mutations arrivent indépendamment au hasard sur chaque site de la chaîne, avec probabilité $q \in [0, 1]$. Quand une mutation se produit sur un site particulier, la lettre est remplacée par une des $\kappa - 1$ lettres restantes dans l'alphabet, choisie uniformément au hasard. Nous pouvons coder le mécanisme de mutation dans un noyau $M : \mathcal{A}^\ell \times \mathcal{A}^\ell \rightarrow [0, 1]$ ainsi : pour tout $\mathbf{u}, \mathbf{v} \in \mathcal{A}^\ell$, la probabilité $M(\mathbf{u}, \mathbf{v})$ de transformer le génotype \mathbf{u} en \mathbf{v} par mutation est

$$M(\mathbf{u}, \mathbf{v}) = \prod_{i=1}^{\ell} \left(\frac{q}{\kappa - 1} 1_{\mathbf{u}(i) \neq \mathbf{v}(i)} + (1 - q) 1_{\mathbf{u}(i) = \mathbf{v}(i)} \right).$$

La distance de Hamming. L'espace des génotypes \mathcal{A}^ℓ est muni de la distance de Hamming d_H , qui mesure le nombre de digits différents entre deux chaînes, i.e., pour $\mathbf{u}, \mathbf{v} \in \mathcal{A}^\ell$

$$d_H(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^{\ell} 1_{\mathbf{u}(i) \neq \mathbf{v}(i)}.$$

Par exemple,

$$d_H(01101, 01011) = 2 \quad \text{et} \quad d_H(\text{AATGTT}, \text{CATGCC}) = 3.$$

Le noyau de mutation M peut être réécrit en utilisant la distance de Hamming : pour tout $\mathbf{u}, \mathbf{v} \in \mathcal{A}^\ell$,

$$M(\mathbf{u}, \mathbf{v}) = \left(\frac{q}{\kappa - 1} \right)^{d_H(\mathbf{u}, \mathbf{v})} (1 - q)^{\ell - d_H(\mathbf{u}, \mathbf{v})}.$$

Le modèle d'Eigen. Nous considérons une population infinie d'individus dont les génotypes sont décrits par les éléments de \mathcal{A}^ℓ , la composition de la population évolue au cours du temps. Pour $\mathbf{v} \in \mathcal{A}^\ell$, notons $x_{\mathbf{v}}(t)$ la concentration

d'individus ayant le génotype v dans la population au temps t . L'évolution des concentrations $x_v(t)$ est gouvernée par le système d'équations différentielles

$$x'_v(t) = \sum_{u \in \mathcal{A}^\ell} x_u(t) f(u) M(u, v) - x_v(t) \sum_{u \in \mathcal{A}^\ell} x_u(t) f(u), \quad v \in \mathcal{A}^\ell.$$

C'est le modèle d'Eigen. Le premier terme dans l'équation correspond à la création d'individus avec génotype v , le deuxième terme correspond à la destruction d'individus avec génotype v . Posons

$$\phi(t) = \sum_{u \in \mathcal{A}^\ell} x_u(t) f(u).$$

La quantité $\phi(t)$ est la *fitness moyenne* de la population au temps t , et elle correspond à la création totale de chaînes au temps t . Remarquons que si la concentration totale de la population au temps t est égale à 1, i.e.,

$$\sum_{v \in \mathcal{A}^\ell} x_v(t) = 1,$$

elle le sera aussi pour tout $s \geq t$. En effet, comme le noyau de mutation M est une matrice stochastique, en sommant sur toutes les équations du système d'Eigen, nous obtenons

$$\sum_{v \in \mathcal{A}^\ell} x'_v(t) = 0.$$

Nous allons toujours supposer que la concentration totale de la population est égale à 1.

Remarque 3.1.1. *Le modèle d'Eigen peut se généraliser naturellement dans le cadre suivant : \mathcal{E} est un ensemble fini, f est une fonction de \mathcal{E} dans \mathbb{R}^+ et M est une matrice stochastique sur \mathcal{E} .*

3.2 Le modèle pour les classes de Hamming

L'espace d'états du modèle d'Eigen est le simplexe de dimension $\kappa^\ell - 1$. C'est un espace peu adapté à notre étude à cause de sa grande dimension et de l'absence d'un ordre naturel. Afin de réduire la taille de l'espace d'états, nous allons regrouper les chaînes dans des classes de Hamming vis-à-vis d'une séquence de référence. Dans le cas du paysage de fitness à un pic, le choix de la séquence de référence est évident, nous choisissons la master sequence, w^* .

Classes de Hamming. Soit $l \in \{0, \dots, \ell\}$, nous disons que le génotype $u \in \mathcal{A}^\ell$ appartient à la classe de Hamming l s'il est à distance l de la master sequence, i.e., si $d_H(u, w^*) = l$.

Dans la section précédente, nous avons défini le modèle d'Eigen pour une fonction de fitness $f : \mathcal{A}^\ell \rightarrow \mathbb{R}^+$ quelconque. Nous ne considérons par la suite que des fonctions de fitness qui respectent les classes de Hamming, c'est à dire, des fonctions de fitness telles que, pour tout $u, v \in \mathcal{A}^\ell$,

$$d_H(u, w^*) = d_H(v, w^*) \Rightarrow f(u) = f(v).$$

Soit $f : \mathcal{A}^\ell \rightarrow \mathbb{R}^+$ une fonction de fitness vérifiant cette hypothèse. Nous pouvons alors définir une nouvelle fonction $f_H : \{0, \dots, \ell\} \rightarrow \mathbb{R}^+$ de la manière suivante : pour $l \in \{0, \dots, \ell\}$, nous posons $f_H(l)$ égale à la valeur de f commune à tout u dans la classe de Hamming l .

La fonction de fitness du paysage à un pic respecte les classes de Hamming. La fonction f_H associée est donnée par

$$\forall l \in \{0, \dots, \ell\}, \quad f_H(l) = \begin{cases} \sigma & \text{si } l = 0, \\ 1 & \text{si } 1 \leq l \leq \ell. \end{cases}$$

Le noyau de mutation M se factorise aussi à travers les classes de Hamming.

Lemme 3.2.1. Soient $b, c \in \{0, \dots, \ell\}$ et soit $u \in \mathcal{A}^\ell$ un génotype dans la classe de Hamming b . La somme

$$\sum_{\substack{v \in \mathcal{A}^\ell \\ d_H(v, w^*) = c}} M(u, v)$$

ne dépend pas de u dans la classe de Hamming b , c'est une fonction de b et c seulement, que nous notons $M_H(b, c)$. Le coefficient $M_H(b, c)$ est donné par

$$\sum_{\substack{0 \leq k \leq \ell - b \\ 0 \leq l \leq b \\ b + k - l = c}} \binom{\ell - b}{k} \binom{b}{l} q^k (1 - q)^{\ell - b - k} \left(\frac{q}{\kappa - 1}\right)^l \left(1 - \frac{q}{\kappa - 1}\right)^{b - l}.$$

Démonstration. Le génotype u étant dans la classe de Hamming b , il diffère exactement de b digits de la master sequence. Soient $l \in \{0, \dots, b\}$ et $k \in$

$\{0, \dots, \ell - b\}$. Parmi les b digits différents de la master sequence, la probabilité que exactement l digits mutent pour devenir égaux à ceux de la master sequence est

$$\binom{b}{l} \left(\frac{q}{\kappa - 1} \right)^l \left(1 - \frac{q}{\kappa - 1} \right)^{b-l}.$$

Également, parmi les $\ell - b$ digits égaux à la master sequence la probabilité que exactement k changent et deviennent différents de ceux de la master sequence est

$$\binom{\ell - b}{k} q^k (1 - q)^{\ell - k}.$$

Nous obtenons le résultat voulu en sommant les produits des deux probabilités, sur tous les l, k qui vérifient la contrainte $b + k - l = c$. \square

Remarque 3.2.2. Soient $b, c \in \{0, \dots, \ell\}$, soit X une variable aléatoire de distribution $\text{Bin}(b, q/(\kappa - 1))$ et soit Y une variable aléatoire de loi $\text{Bin}(\ell - b, q)$, avec X et Y indépendantes. Alors,

$$M_H(b, c) = P(b - X + Y = c).$$

Nous avons obtenu un nouvel espace d'états, l'espace des classes de Hamming $\{0, \dots, \ell\}$, une nouvelle fonction de fitness $f_H : \{0, \dots, \ell\} \rightarrow \mathbb{R}^+$, et un nouveau noyau de mutation $M_H : \{0, \dots, \ell\} \times \{0, \dots, \ell\} \rightarrow [0, 1]$. Pour $k \in \{0, \dots, \ell\}$, nous notons $x_k(t)$ la concentration des chaînes dans la classe de Hamming k au temps t . Les concentrations $x_k(t)$ obéissent au système d'équations

$$(S) \quad x'_k(t) = \sum_{i=0}^{\ell} x_i(t) f_H(i) M_H(i, k) - x_k(t) \sum_{i=0}^{\ell} x_i(t) f_H(i), \quad 0 \leq k \leq \ell.$$

Dans la section suivante, nous étudions un modèle d'Eigen plus général dont les résultats s'appliquent au système (S). Le reste du chapitre est consacré à étudier le système (S) dans un régime asymptotique particulier.

3.3 Solutions stationnaires et stabilité

Comme nous l'avons dit dans la remarque 3.1.1, le modèle d'Eigen peut être défini de manière plus générale pour un ensemble fini de génotypes \mathcal{E} , une fonction de fitness $f : \mathcal{E} \rightarrow \mathbb{R}^+$ et un noyau de mutation M sur \mathcal{E} . Nous faisons

ici un petit parenthèse pour étudier ce modèle généralisé. Nous supposons que l'ensemble \mathcal{E} a cardinal $N \geq 2$ et nous identifions \mathcal{E} avec l'ensemble $\{1, \dots, N\}$, nous parlons alors du i -ème génotype, pour $1 \leq i \leq N$. Le système d'équations d'Eigen général s'écrit ainsi :

$$(*) \quad x'_k(t) = \sum_{i=1}^N x_i(t)f(i)M(i, k) - x_k(t)\phi(t), \quad 1 \leq k \leq N,$$

où $\phi(t)$ représente la fitness moyenne de la population au temps t ,

$$\phi(t) = \sum_{i=1}^N x_i(t)f(i).$$

Soit \mathcal{S}^N le simplexe à N coordonnées :

$$\mathcal{S}^N = \{x \in [0, 1]^N : x_1 + \dots + x_N = 1\}.$$

Pour tout $x^0 \in \mathcal{S}^N$, la solution de (*) avec condition initiale $x(0) = x^0$ reste toujours dans \mathcal{S}^N , cela peut être vérifié en sommant sur toutes les équations du système (*).

Scaling. Le système (*) vérifie la propriété de *scaling* suivante : étant donnée une fonction de fitness $f : \{1, \dots, N\} \rightarrow \mathbb{R}^+$ et une constante $c > 0$ nous pouvons définir une nouvelle fonction de fitness $cf : \mathbb{N} \rightarrow \mathbb{R}^+$ par

$$\forall k \geq 0, \quad (cf)(k) = cf(k).$$

La trajectoire $(x(t), t \geq 0)$ est solution du système (*) associé à cf si et seulement si $(x(ct), t \geq 0)$ est solution du système (*) associé à f . Autrement dit, multiplier toutes les fitness par c correspond à accélérer le temps d'un facteur c . En particulier, les solutions stationnaires du système (*) sont invariantes par le changement $f \rightarrow cf$.

Nous définissons la matrice $(W(i, j))_{1 \leq i, j \leq N}$ par

$$\forall i, j \in \{1, \dots, N\}, \quad W(i, j) = f(i)M(i, j).$$

Nous faisons l'hypothèse suivante sur la matrice W :

Hypothèse. La matrice W a toutes ses composantes strictement positives. Sous cette hypothèse, le théorème de Perron–Frobenius peut s'appliquer à la matrice W :

Théorème 3.3.1 (Perron–Frobenius.). *Soit A une matrice carrée ayant tous ses coefficients strictement positifs. Il existe un nombre réel positif λ vérifiant les propriétés suivantes :*

- *La valeur λ est une valeur propre simple de A , et toute autre valeur propre de A a un module strictement inférieur à λ .*
- *Il existe un vecteur propre u de A , associé à λ , ayant toutes ses composantes strictement positives.*

La démonstration du théorème de Perron–Frobenius peut être trouvée dans [73]. Le théorème est vrai pour les vecteurs propres à gauche comme pour les vecteurs propres à droite. Rappelons que v est un *vecteur propre généralisé* de rang $k \geq 1$ de la matrice A associé à la valeur propre μ , si

$$(A - \mu I)^k v = 0 \quad \text{et} \quad (A - \mu I)^{k-1} v \neq 0.$$

Nous avons le résultat suivant :

Lemme 3.3.2. *Soit A une matrice carrée ayant tous ses coefficients strictement positifs. Tout vecteur propre généralisé de A ayant toutes ses composantes positives, est un multiple du vecteur propre de Perron–Frobenius.*

Démonstration. Soit v un vecteur propre généralisé (à droite) de rang $k \geq 1$ de A , associé à la valeur propre μ , ayant toutes ses composantes positives, avec au moins une composante non nulle. Soit λ la valeur propre de Perron–Frobenius de A et soit u le vecteur propre à gauche associé. Alors

$${}^t u (A - \mu I) = (\lambda - \mu) {}^t u.$$

Nous en déduisons que

$$0 = {}^t u (A - \mu I)^k v = (\lambda - \mu)^k ({}^t u) v.$$

Or, puisque u a toutes ses coordonnées strictement positives, ${}^t u v > 0$ et donc $\lambda = \mu$. Puisque λ est une valeur propre simple, le sous-espace propre généralisé associé à λ est de dimension 1, et nous concluons que v est multiple du vecteur propre à droite de Perron–Frobenius de A . \square

Soit λ la valeur propre associée à la matrice W par le théorème de Perron–Frobenius et soit u le vecteur propre unitaire de W ayant toutes ses composantes strictement positives. Nous avons le théorème suivant :

Théorème 3.3.3. *Le vecteur u est le seul point stationnaire du système d'équations (*) dans le simplexe \mathcal{S}^N . En outre, pour tout $x^0 \in \mathcal{S}^N$, la solution $(x(t))_{t \geq 0}$ de (*) avec condition initiale x^0 converge vers u quand t tend vers l'infini.*

Démonstration. Une solution stationnaire de (*), $x \in \mathcal{S}^N$, satisfait le système d'équations

$$0 = \sum_{i=1}^N x_i f(i) M(i, k) - x_k \phi(x), \quad 1 \leq k \leq N.$$

Ce système peut se réécrire sous forme matricielle : ${}^t x W = {}^t x \phi(x)$, avec,

$$\phi(x) = \sum_{i=1}^N x_i f(i) = \sum_{i,j=1}^N x_i f(i) M(i, j) = \sum_{i,j=1}^N x_i W(i, j).$$

Ainsi, si $x \in \mathcal{S}^N$ est une solution stationnaire de (*), alors x est un vecteur propre unitaire de W ayant toutes ses composantes positives et $\phi(x)$ est une valeur propre de W . L'unicité d'une telle solution découle du théorème de Perron–Frobenius. Pour démontrer l'existence, nous prenons λ la valeur propre de W donnée par le théorème de Perron–Frobenius, et u le vecteur propre unitaire associé. Nous remarquons que λ vérifie l'équation

$$\lambda = \sum_{i,j=1}^N u_i W(i, j),$$

et donc $x = u \in \mathcal{S}^N$ est bien une solution stationnaire de (*). Nous démontrons ensuite que toute solution de (*) converge vers u . Pour cela nous effectuons le changement de variables suivant :

$$y_k(t) = x_k(t) \exp \left(\int_0^t \phi(s) ds \right), \quad 1 \leq k \leq N.$$

Ce changement de variables rend le système linéaire, en effet,

$$y'_k(t) = \sum_{i=1}^N y_i(t) W(i, k), \quad 1 \leq k \leq N,$$

ce qui s'écrit sous forme matricielle ${}^t y' = {}^t y W$. Remarquons que pour retrouver $(x(t))_{t \geq 0}$ à partir de $(y(t))_{t \geq 0}$, il suffit de prendre

$$x_k(t) = \frac{y_k(t)}{y_1(t) + \cdots + y_N(t)}, \quad 1 \leq k \leq N.$$

En effet, cela se voit en sommant sur k dans la formule de changement de variables, ce qui donne

$$\sum_{k=1}^N y_k(t) = \exp\left(\int_0^t \phi(s) ds\right).$$

Il existe une matrice de changement de base P qui nous permet de passer de la matrice W à sa forme canonique de Jordan réelle :

$$PWP^{-1} = \widetilde{W} = \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & J & \\ 0 & & & \end{pmatrix},$$

où λ est la valeur propre donnée par le théorème de Perron–Frobenius et J une matrice de Jordan réelle. Remarquons que la première ligne de la matrice P , que nous appelons $P(1, \cdot)$, est un vecteur propre de W associé à λ , i.e.,

$$P(1, \cdot)W = P(1, \cdot)\lambda.$$

Nous faisons encore un changement de variables : nous posons ${}^t z = {}^t y P^{-1}$. Alors, z vérifie le système d'équations différentielles ${}^t z' = {}^t z \widetilde{W}$, dont la solution est donnée par

$$z(t) = \begin{pmatrix} e^{\lambda t} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & e^{Jt} & \\ 0 & & & \end{pmatrix} \begin{pmatrix} z_1(0) \\ z_2(0) \\ \vdots \\ z_N(0) \end{pmatrix},$$

où e^{Jt} est la matrice exponentielle définie par :

$$e^{Jt} = \sum_{n \geq 0} \frac{J^n t^n}{n!}.$$

Puisque toute valeur propre de la matrice J est en valeur absolue strictement plus petite que λ (par le théorème de Perron–Frobenius), nous avons pour la norme d'opérateur $\|\cdot\|$ associée à la norme euclidienne de \mathbb{R}^N :

$$\|e^{Jt}\| \leq e^{\|J\|t} < e^{\lambda t}.$$

Nous avons défini le vecteur z par ${}^t z = {}^t y P^{-1}$, ceci entraîne notamment que $z_1(0) \neq 0$. En effet, nous pouvons exprimer le vecteur $y(0)$ comme combinaison linéaire des lignes de la matrice P :

$${}^t y(0) = c_1 P(1, \cdot) + \cdots + c_N P(N, \cdot),$$

avec $c_1, \dots, c_N \in \mathbb{R}$. Nous en déduisons que

$$z_1(0) = {}^t y(0) P^{-1}(\cdot, 1) = c_1.$$

Or, les lignes de P sont des vecteurs propres généralisés de W , et d'après le lemme 3.3.2, le sous-espace engendré par $P(2, \cdot), \dots, P(N, \cdot)$ ne contient pas le simplexe \mathcal{S}^N . Comme $y(0) = x(0) \in \mathcal{S}^N$, nécessairement $c_1 \neq 0$. Puisque $z_1(0) \neq 0$,

$$\lim_{t \rightarrow \infty} \frac{z_k(t)}{z_1(t)} = 0, \quad 2 \leq k \leq N.$$

Vu que ${}^t y(t) = {}^t z(t) P$, il s'ensuit que, pour tout $k \in \{1, \dots, N\}$,

$$\begin{aligned} \lim_{t \rightarrow \infty} x_k(t) &= \lim_{t \rightarrow \infty} \frac{y_k(t)}{y_1(t) + \cdots + y_n(t)} \\ &= \lim_{t \rightarrow \infty} \frac{\sum_{1 \leq i \leq N} z_i(t) P(i, k)}{\sum_{1 \leq i, j \leq N} z_i(t) P(i, j)} = \frac{P(1, k)}{\sum_{j=1}^N P(1, j)} = u_k, \end{aligned}$$

où $u = (u_1, \dots, u_N)$ est le vecteur de Perron–Frobenius normalisé de W . C'est ce qu'il fallait démontrer. \square

Le résultat que nous venons de démontrer s'applique au système d'Eigen original défini dans la section 3.1, ainsi qu'au système d'Eigen pour les classes de Hamming défini dans la section 3.2, sous l'hypothèse que les fonctions de fitness associées soient strictement positives. Le théorème peut être généralisé à des matrices W avec des coefficients nuls sous l'hypothèse qu'elle soient irréductibles : une matrice carrée W est irréductible si pour tout pair d'indices i, j il existe un nombre naturel $m = m(i, j)$ tel que l'élément $W^m(i, j)$ est strictement positif. Les résultats du théorème 3.3.3 apparaissent déjà dans plusieurs articles dans la littérature sur le modèle d'Eigen, notamment dans [5, 48, 49, 78].

3.4 Régime asymptotique

Revenons au modèle d'Eigen pour les classes de Hamming introduit dans la section 3.2 :

$$(\mathcal{S}) \quad x'_k(t) = \sum_{i=0}^{\ell} x_i(t) f_H(i) M_H(i, k) - x_k(t) \sum_{i=0}^{\ell} x_i(t) f_H(i), \quad 0 \leq k \leq \ell.$$

Les résultats de la section précédente s'appliquent au système (\mathcal{S}) . Pour voir apparaître un phénomène de seuil d'erreur, nous étudions le système (\mathcal{S}) quand la longueur du génome ℓ tend vers l'infini et la probabilité de mutation q tend vers 0. Plus précisément, nous considérons le régime asymptotique suivant :

$$\ell \rightarrow \infty, \quad q \rightarrow 0, \quad \ell q \rightarrow a \in]0, +\infty[.$$

Dans la suite, quand nous disons qu'un résultat est vrai *asymptotiquement*, cela veut dire qu'il est vrai pour ℓ assez grand, q assez petit et ℓq assez proche de a .

Fonction de fitness limite. Pour que le régime asymptotique précédent ait du sens, nous supposons que, pour chaque $\ell \geq 1$, la fonction $f_H : \{0, \dots, \ell\} \rightarrow \mathbb{R}^+$ est la restriction à l'espace $\{0, \dots, \ell\}$ d'une fonction $f_\infty : \mathbb{N} \rightarrow \mathbb{R}^+$.

Le cas du paysage à un pic peut être étudié dans ce cadre en considérant la fonction $f_\infty : \mathbb{N} \rightarrow \mathbb{R}^+$ donnée par

$$\forall k \geq 0 \quad f_\infty(k) = \begin{cases} \sigma & \text{si } k = 0, \\ 1 & \text{si } k \geq 1. \end{cases}$$

Noyau de mutation limite. Soient $b, c \in \mathbb{N}$ et $\ell > b, c$. Soit X une variable aléatoire de loi binomiale de paramètres b et $q/(\kappa - 1)$, soit Y une variable aléatoire de loi binomiale de paramètres $\ell - b$ et q . Supposons que X et Y sont indépendantes. D'après la remarque 3.2.2,

$$M_H(b, c) = P(-X + Y = c - b).$$

Asymptotiquement, $P(X = 0)$ tend vers 1, et Y tend vers une variable aléatoire de loi de Poisson de paramètre a . Donc, pour tout $b, c \geq 0$,

$$\lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} M_H(b, c) = \begin{cases} e^{-a} \frac{a^{c-b}}{(c-b)!} & \text{si } b \leq c, \\ 0 & \text{si } b > c. \end{cases}$$

Nous notons M_∞ ce noyau limite.

Système d'Eigen limite. Nous nous intéressons donc au système d'équations limite suivant :

$$(\mathcal{S}_\infty) \quad x'_k(t) = \sum_{i=0}^k x_i(t) f_\infty(i) e^{-a} \frac{a^{k-i}}{(k-i)!} - x_k(t) \sum_{i=0}^{\infty} x_i(t) f_\infty(i), \quad k \geq 0.$$

Dans les sections suivantes, nous étudions les solutions stationnaires et la stabilité des trajectoires pour ce système, d'abord pour le paysage à un pic, puis pour des fonctions de fitness plus générales.

Chapitre 4

Le paysage à un pic

4.1 Solutions stationnaires

Nous étudions dans ce chapitre le système d'Eigen limite (\mathcal{S}_∞) introduit dans la section précédente, dans le cas du paysage à un pic, c'est-à-dire, quand la fonction $f_\infty : \mathbb{N} \rightarrow \mathbb{R}^+$ est donnée par :

$$\forall k \geq 0 \quad f_\infty(k) = \begin{cases} \sigma & \text{si } k = 0, \\ 1 & \text{si } k \geq 1, \end{cases}$$

où $\sigma > 1$. La fitness moyenne de la population dans ce cas peut être réécrite ainsi :

$$\phi(t) = \sum_{k=0}^{\infty} x_k(t) f_\infty(k) = (\sigma - 1)x_0(t) + 1.$$

De même, en remplaçant f_∞ par la fonction du paysage à un pic, le système d'Eigen infini se réécrit ainsi :

$$(\mathcal{S}_\infty^{sp}) \quad \begin{cases} x_0'(t) = \sigma x_0(t) e^{-a} - x_0(t) \phi(t), \\ x_k'(t) = \sigma x_0(t) e^{-a} \frac{a^k}{k!} + \sum_{i=1}^k x_i(t) e^{-a} \frac{a^{k-i}}{(k-i)!} - x_k(t) \phi(t), \quad k \geq 1. \end{cases}$$

Nous cherchons d'abord les solutions stationnaires du système. Nous avons le résultat suivant :

Proposition 4.1.1. *Si $\sigma e^{-a} \leq 1$, le système d'Eigen infini n'a qu'une solution stationnaire :*

$$x_k = 0, \quad k \geq 0.$$

Si $\sigma e^{-a} > 1$, le système d'Eigen infini a deux solutions stationnaires, la solution identiquement nulle et

$$\rho_k^* = (\sigma e^{-a} - 1) \frac{a^k}{k!} \sum_{i=1}^{\infty} \frac{i^k}{\sigma^i}, \quad k \geq 0.$$

Démonstration. Les solutions stationnaires du système vérifient

$$\begin{cases} 0 = \sigma x_0 e^{-a} - x_0 \phi(x) = x_0 (\sigma e^{-a} - \phi(x)), \\ 0 = \sigma x_0 e^{-a} \frac{a^k}{k!} + \sum_{i=1}^k x_i e^{-a} \frac{a^{k-i}}{(k-i)!} - x_k \phi(x), \quad k \geq 1. \end{cases}$$

Supposons d'abord que $\sigma e^{-a} \leq 1$. Puisque $\phi(x) = (\sigma - 1)x_0 + 1 \geq 1$ nous concluons, d'après la première équation, que dans ce cas $x_0 = 0$. Nous vérifions par récurrence que $x_k = 0$ pour tout $k \geq 1$. Supposons ensuite que $\sigma e^{-a} > 1$. D'après la première équation, nous concluons que soit $x_0 = 0$, soit $\phi(x) = \sigma e^{-a}$, et dans ce cas, en remplaçant $\phi(x) = (\sigma - 1)x_0 + 1$, nous avons

$$x_0 = \frac{\sigma e^{-a} - 1}{\sigma - 1}.$$

Nous pouvons maintenant résoudre le reste des équations par récurrence. Pour la condition initiale $x_0 = 0$, il est clair que $x_k = 0$ pour tout $k \geq 1$. Pour la condition initiale $x_0 = (\sigma e^{-a} - 1)/(\sigma - 1)$, nous remplaçons $\phi(x) = \sigma e^{-a}$ dans le système et nous obtenons

$$0 = \sigma x_0 \frac{a^k}{k!} + \sum_{i=1}^k x_i \frac{a^{k-i}}{(k-i)!} - \sigma x_k, \quad k \geq 1.$$

Nous allons résoudre cette relation de récurrence par la méthode des fonctions génératrices. Posons

$$g(X) = \sum_{k=0}^{\infty} x_k X^k.$$

Alors

$$g(X)e^{aX} = \sum_{k=0}^{\infty} \left(\sum_{i=0}^k x_i \frac{a^{k-i}}{(k-i)!} \right) X^k.$$

Or, d'après la relation de récurrence,

$$\sum_{i=0}^k x_i \frac{a^{k-i}}{(k-i)!} = \sigma x_k - (\sigma - 1)x_0 \frac{a^k}{k!}.$$

Nous en déduisons que

$$g(X)e^{aX} = \sigma g(X) - (\sigma - 1)x_0 e^{aX}.$$

Et en remplaçant $x_0 = (\sigma e^{-a} - 1)/(\sigma - 1)$, nous obtenons

$$\begin{aligned} g(X) &= (\sigma e^{-a} - 1) \frac{e^{aX}/\sigma}{1 - e^{aX}/\sigma} = (\sigma e^{-a} - 1) \sum_{i=1}^{\infty} \frac{e^{aiX}}{\sigma^i} \\ &= (\sigma e^{-a} - 1) \sum_{i=1}^{\infty} \frac{1}{\sigma^i} \sum_{k=0}^{\infty} \frac{(ai)^k}{k!} X^k = (\sigma e^{-a} - 1) \sum_{k=0}^{\infty} \frac{a^k}{k!} \left(\sum_{i=1}^{\infty} \frac{i^k}{\sigma^i} \right) X^k. \end{aligned}$$

Nous concluons que

$$x_k = (\sigma e^{-a} - 1) \frac{a^k}{k!} \sum_{i=1}^{\infty} \frac{i^k}{\sigma^i}, \quad k \geq 0,$$

comme voulu. □

Avant de continuer avec l'étude du système (S_{∞}^{sp}) , nous allons faire une petite parenthèse pour analyser le point stationnaire $(\rho_k^*)_{k \geq 0}$.

4.2 La distribution de la quasi-espèce

Dans cette section, nous supposons que $\sigma e^{-a} > 1$, et nous notons par $(\rho_k^*)_{k \geq 0}$ le point stationnaire non trivial du système limite d'Eigen (S_{∞}^{sp}) obtenu dans la proposition précédente. Dans la démonstration, nous avons obtenu la relation

$$g(X) = \sum_{k=0}^{\infty} \rho_k^* X^k = (\sigma e^{-a} - 1) \frac{e^{aX}/\sigma}{1 - e^{aX}/\sigma}.$$

En particulier, $g(1) = 1$, et

$$\sum_{k \geq 0} \rho_k^* = 1.$$

Nous pouvons donc définir une distribution de probabilité à partir du point stationnaire $(\rho_k^*)_{k \geq 0}$.

Définition 4.2.1. Nous disons qu'une variable aléatoire Y suit la loi de la quasi-espèce de paramètres σ , α , notée $\mathcal{Q}(\sigma, \alpha)$, si

$$P(Y = k) = (\sigma e^{-\alpha} - 1) \frac{\alpha^k}{k!} \sum_{i=1}^{\infty} \frac{i^k}{\sigma^i}, \quad k \geq 0.$$

Calculons l'espérance et la variance de la distribution $\mathcal{Q}(\sigma, \alpha)$. Réécrivons la fonction génératrice sous la forme

$$g(X) = \sum_{k \geq 0} \rho_k^* X^k = (\sigma e^{-\alpha} - 1) \frac{1}{\sigma e^{-\alpha X} - 1}.$$

Les deux premières dérivées de g sont

$$g'(X) = (\sigma e^{-\alpha} - 1) \frac{\sigma \alpha e^{-\alpha X}}{(\sigma e^{-\alpha X} - 1)^2},$$

$$g''(X) = \frac{(\sigma e^{-\alpha} - 1)(\sigma \alpha^2 e^{-\alpha X})(\sigma e^{-\alpha X} + 1)}{(\sigma e^{-\alpha X} - 1)^3}.$$

Nous obtenons alors

$$E(Y) = \sum_{k \geq 0} k \rho_k^* = g'(1) = \frac{\sigma \alpha e^{-\alpha}}{\sigma e^{-\alpha} - 1},$$

$$E(Y^2) = \sum_{k \geq 0} k^2 \rho_k^* = g''(1) + g'(1) = \frac{\sigma \alpha e^{-\alpha}}{(\sigma e^{-\alpha} - 1)^2} ((\alpha + 1)\sigma e^{-\alpha} + \alpha - 1),$$

$$\text{Var}(Y) = E(Y^2) - E(Y)^2 = \frac{\sigma \alpha e^{-\alpha}(\sigma e^{-\alpha} + \alpha - 1)}{(\sigma e^{-\alpha} - 1)^2}.$$

Les figures 4.1 et 4.2 représentent graphiquement la fréquence de la master sequence et des 10 premières classes de Hamming en fonction de α , pour $\sigma = 5$ et $\sigma = 1000000$.

Formules alternatives. Nous avons donné l'expression suivante pour la distribution de la quasi-espèce :

$$\rho_k^* = (\sigma e^{-\alpha} - 1) \frac{\alpha^k}{k!} \sum_{i \geq 1} \frac{i^k}{\sigma^i}, \quad k \geq 0.$$

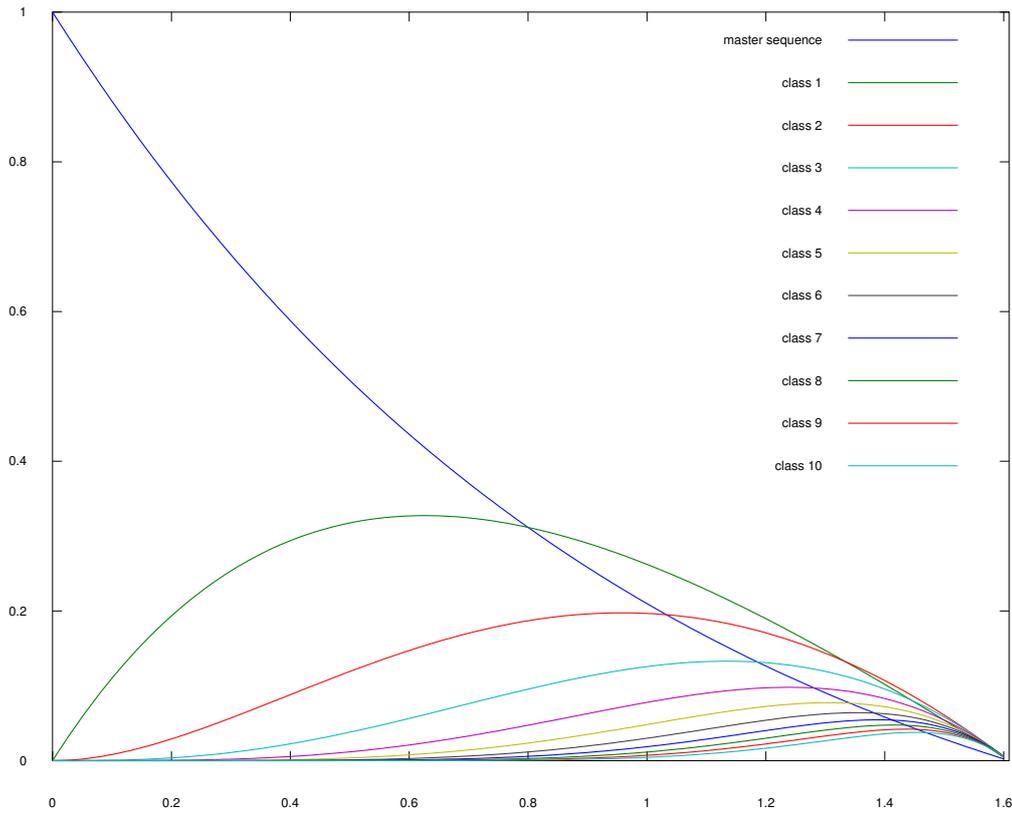


FIGURE 4.1 – Fréquence de w^* et des 10 premières classes pour $\sigma = 5$.

Nous pouvons réécrire cette formule à l'aide du polylogarithme, ou fonction de Jonquière, qui est défini pour tout $s \in \mathbb{R}$ et $|z| < 1$ par

$$\text{Li}_s(z) = \sum_{i=1}^{\infty} \frac{z^i}{i^s}.$$

Avec cette notation,

$$\rho_k^* = (\sigma e^{-a} - 1) \frac{a^k}{k!} \text{Li}_{-k}\left(\frac{1}{\sigma}\right), \quad k \geq 0.$$

Posons maintenant $s = 1/\sigma$, nous avons la relation

$$\sum_{i \geq 1} s^i = \frac{s}{1-s}.$$

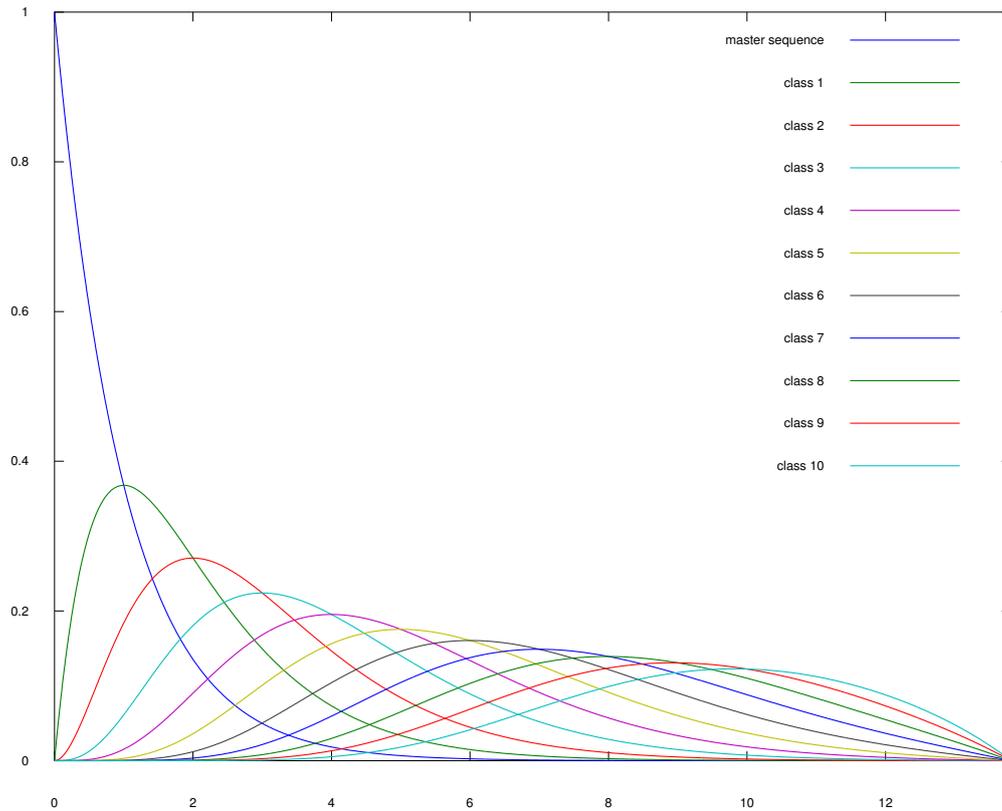


FIGURE 4.2 – Fréquence de w^* et des 10 premières classes pour $\sigma = 10^6$.

Nous dérivons et nous multiplions par s les deux côtés successivement pour obtenir :

$$\begin{aligned} \sum_{i \geq 1} i s^i &= \frac{s}{(1-s)^2}, \\ \sum_{i \geq 1} i^2 s^i &= \frac{s}{(1-s)^3} (1+s), \\ \sum_{i \geq 1} i^3 s^i &= \frac{s}{(1-s)^4} (1+4s+s^2), \\ \sum_{i \geq 1} i^4 s^i &= \frac{s}{(1-s)^5} (1+11s+11s^2+s^3). \end{aligned}$$

Les nombres apparaissant dans les polynômes de droite sont les nombres eulériens. Pour $k \geq 1$ et $0 \leq i < k$, le nombre eulérien $E(k, i)$ est le nombre de permutations de $\{1, \dots, k\}$ ayant exactement i montées, c'est-à-dire, i éléments

supérieurs à l'élément précédent. En général, nous avons l'identité

$$\sum_{i \geq 1} i^k s^i = \frac{s}{(1-s)^{k+1}} \sum_{i=0}^{k-1} E(k, i) s^i, \quad k \geq 1,$$

ce qui donne, pour tout $k \geq 0$:

$$\rho_k^* = (\sigma e^{-a} - 1) \frac{a^k}{k!} \times \frac{1}{(\sigma - 1)^{k+1}} \sum_{i=0}^{k-1} E(k, i) \sigma^{k-i}.$$

4.3 Convergence des trajectoires

Après cette parenthèse sur la distribution de la quasi-espèce, nous reprenons l'étude du système d'Eigen limite :

$$(S_{\infty}^{sp}) \quad \begin{cases} x_0'(t) = \sigma x_0(t) e^{-a} - x_0(t) \phi(t), \\ x_k'(t) = \sigma x_0(t) \frac{a^k}{k!} e^{-a} + \sum_{i=1}^k x_i(t) \frac{a^{k-i}}{(k-i)!} e^{-a} - x_k(t) \phi(t), \quad k \geq 1. \end{cases}$$

Nous nous intéressons dans cette section au comportement asymptotique des solutions du système (S_{∞}^{sp}) . Nous avons le résultat suivant.

Théorème 4.3.1. *Soit $(x(t), t \geq 0)$ une solution de (S_{∞}^{sp}) .*

- Si $\sigma e^{-a} \leq 1$, alors

$$\lim_{t \rightarrow \infty} x_k(t) = 0, \quad k \geq 0.$$

- Si $\sigma e^{-a} > 1$,

– soit $x_0(0) = 0$ et alors $x_0(t) = 0$ pour tout $t \geq 0$ et

$$\lim_{t \rightarrow \infty} x_k(t) = 0, \quad k \geq 1,$$

– soit $x_0(0) > 0$ et alors

$$\lim_{t \rightarrow \infty} x_k(t) = \rho_k^* = (\sigma e^{-a} - 1) \frac{a^k}{k!} \sum_{i \geq 1} \frac{i^k}{\sigma^i}, \quad k \geq 0.$$

Démonstration. Nous allons étudier seulement le cas $\sigma e^{-\alpha} > 1$, $x_0(0) > 0$ (les démonstrations dans les autres cas sont similaires ou plus simples). Commençons par étudier la première coordonnée $x_0(t)$. Rappelons que, puisque nous travaillons avec le paysage à un pic, la fitness moyenne se réécrit sous la forme $\phi(t) = (\sigma - 1)x_0(t) + 1$. Nous devons donc étudier l'équation différentielle

$$x_0'(t) = x_0(t)(\sigma e^{-\alpha} - (\sigma - 1)x_0(t) - 1).$$

L'équation

$$x_0(\sigma e^{-\alpha} - (\sigma - 1)x_0 - 1) = 0$$

a pour racines les points 0 et ρ_0^* . Évidemment, si $x_0(0) = \rho_0^*$, nous avons $x_0(t) = \rho_0^*$ pour tout $t \geq 0$. Supposons donc que $x_0(0) \neq \rho_0^*$. Grâce à l'équation différentielle, nous voyons que la dérivée de la trajectoire $x_0(t)$ est positive sur l'ensemble $\{t : x_0(t) \in]0, \rho_0^*[\}$, et négative sur l'ensemble $\{t : x_0(t) \in]\rho_0^*, 1[\}$. Supposons que $x_0(0)$ est dans $]0, \rho_0^*[$ (la preuve pour le cas $x_0(0) \in]\rho_0^*, 1[$ est similaire). Les solutions de l'équation différentielle sont continues et ne peuvent pas s'intersecter, ce qui implique que

$$\forall t \geq 0 \quad x_0(t) \in]0, \rho_0^*[.$$

En particulier, $x_0'(t) > 0$ pour tout $t \geq 0$ et $x_0(t)$ est une trajectoire strictement croissante. De plus, elle est bornée par ρ_0^* , donc elle a une limite

$$\tilde{x}_0 = \lim_{t \rightarrow \infty} x_0(t).$$

Nous allons montrer que $\tilde{x}_0 = \rho_0^*$. Par le théorème des accroissements finis, pour tout $t \geq 0$, il existe ξ dans $]t, t + 1[$ tel que

$$x_0(t + 1) - x_0(t) = x_0'(t + 1) - x_0'(t) = x_0'(t) + x_0'(\xi).$$

Nous faisons tendre t vers $l'infini$ et nous obtenons

$$\lim_{t \rightarrow \infty} x_0'(\xi) = 0.$$

Par ailleurs,

$$x_0'(\xi) = x_0(\xi)(\sigma e^{-\alpha} - (\sigma - 1)x_0(\xi) - 1),$$

d'où, en passant à la limite,

$$\tilde{x}_0(\sigma e^{-\alpha} - (\sigma - 1)\tilde{x}_0 - 1) = 0.$$

Puisque $\tilde{x}_0 > x_0(0) > 0$, nécessairement $\tilde{x}_0 = \rho_0^*$. Remarquons que pour toute condition initiale, la trajectoire $(x_0(t))_{t \geq 0}$ est Lipschitz.

Nous faisons ensuite la preuve par récurrence. Soit $k \geq 1$ et supposons que le résultat est démontré jusqu'au rang $k - 1$:

$$\lim_{t \rightarrow \infty} (x_0(t), \dots, x_{k-1}(t)) = (\rho_0^*, \dots, \rho_{k-1}^*).$$

Supposons de plus que les trajectoires $(x_0(t))_{t \geq 0}, \dots, (x_{k-1}(t))_{t \geq 0}$ sont Lipschitz. La k -ième équation du système d'Eigen infini est :

$$x'_k(t) = x_0(t)\sigma e^{-a} \frac{a^k}{k!} + \sum_{i=1}^k x_i(t) e^{-a} \frac{a^{k-i}}{(k-i)!} - x_k(t)((\sigma - 1)x_0(t) + 1).$$

Posons

$$\begin{aligned} \alpha(t) &= x_0(t) e^{-a} \frac{a^k}{k!} + \sum_{i=1}^{k-1} x_i(t) e^{-a} \frac{a^{k-i}}{(k-i)!}, \\ \beta(t) &= (\sigma - 1)x_0(t) + 1 - e^{-a}. \end{aligned}$$

Alors $x'_k(t) = \alpha(t) - \beta(t)x_k(t)$, où α et β sont des fonctions Lipschitz, positives et convergentes quant t tend vers l'infini. En particulier $(x_k(t))_{t \geq 0}$ est Lipschitz. Les limites de $\alpha(t)$ et $\beta(t)$, que nous notons α^* et β^* , vérifient

$$\frac{\alpha^*}{\beta^*} = \frac{1}{(\sigma - 1)\rho_0^* + 1 - e^{-a}} \left(\rho_0^* \sigma e^{-a} \frac{a^k}{k!} + \sum_{i=1}^{k-1} \rho_i^* \frac{a^{k-i}}{(k-i)!} \right).$$

En remplaçant ρ_0^* par sa valeur dans le dénominateur, nous constatons que α^*/β^* est donné par la même formule que ρ_k^* dans la relation de récurrence de la distribution de la quasi-espèce. Nous concluons (annexe A) que

$$\lim_{t \rightarrow \infty} x_k(t) = \frac{\alpha^*}{\beta^*} = \rho_k^*,$$

comme voulu. □

4.4 Convergence des systèmes

Considérons le système d'Eigen pour les classes de Hamming avec le paysage à un pic (\mathcal{S}^{sp}) donné par : pour $k \in \{0, \dots, \ell\}$,

$$x'_k(t) = x_0(t)\sigma M_H(0, k) + \sum_{i=1}^{\ell} x_i(t) M_H(i, k) - x_k(t)((\sigma - 1)x_0(t) + 1).$$

Nous avons démontré dans la section 3.3 que ce système d'équations différentielles n'admet qu'une seule solution stationnaire $u = (u_k)_{0 \leq k \leq \ell}$, et que toute solution du système (S^{sp}) converge vers u , indépendamment de la condition initiale. Quand nous regardons le régime asymptotique

$$\ell \rightarrow \infty, \quad q \rightarrow 0, \quad \ell q \rightarrow \alpha \in]0, +\infty[,$$

nous obtenons le système d'équations limite (S_{∞}^{sp}) . Dans les sections 4.1 et 4.3, nous avons vu que le système d'équations (S_{∞}^{sp}) n'admet qu'un seul point stationnaire, 0, quand $\sigma e^{-\alpha} \leq 1$, et admet deux points stationnaires, 0 et $\rho^* = (\rho_k^*)_{k \geq 0}$, quand $\sigma e^{-\alpha} > 1$. De plus, quand $\sigma e^{-\alpha} \leq 1$, toutes les solutions du système infini convergent vers 0, et quand $\sigma e^{-\alpha} > 1$, toutes les solutions du système infini convergent : vers 0 si la condition initiale vérifie $x_0(0) = 0$, et vers ρ^* si la condition initiale vérifie $x_0(0) > 0$. Mais dans quel sens le système d'Eigen pour les classes de Hamming (S^{sp}) converge-t-il vers le système infini (S_{∞}^{sp}) ? Dans cette section, nous démontrons que les solutions du système d'Eigen pour les classes de Hamming convergent vers les solutions du système infini. Nous démontrons aussi que la solution stationnaire du système fini converge vers une solution stationnaire du système infini. Nous allons d'abord montrer deux lemmes qui nous seront utiles par la suite. Nous définissons une application

$$\mathcal{M}_H : \{0, \dots, \ell\} \times [0, 1]^\ell \rightarrow \{0, \dots, \ell\}$$

en posant

$$\forall b \in \{0, \dots, \ell\} \quad \forall u_1, \dots, u_\ell \in [0, 1]^\ell$$

$$\mathcal{M}_H(b, u_1, \dots, u_\ell) = b - \sum_{k=1}^b 1_{u_k < q/(\kappa-1)} + \sum_{k=b+1}^{\ell} 1_{u_k > 1-q}.$$

L'application \mathcal{M}_H est définie de sorte que, si U_1, \dots, U_ℓ sont des variables aléatoires indépendantes de loi uniforme sur l'intervalle $[0, 1]$, alors pour tout b dans $\{0, \dots, \ell\}$, la loi de $\mathcal{M}_H(b, U_1, \dots, U_\ell)$ est donnée par la ligne de la matrice M_H associée à b , i.e.,

$$\forall c \in \{0, \dots, \ell\} \quad P(\mathcal{M}_H(b, U_1, \dots, U_\ell) = c) = M_H(b, c).$$

Lemme 4.4.1. *Si $q \in [0, 1 - 1/\kappa]$, l'application \mathcal{M}_H est croissante vis-à-vis de la classe de Hamming, i.e.,*

$$\forall b, c \in \{0, \dots, \ell\} \quad \forall u_1, \dots, u_\ell \in [0, 1]$$

$$b \leq c \implies \mathcal{M}_H(b, u_1, \dots, u_\ell) \leq \mathcal{M}_H(c, u_1, \dots, u_\ell).$$

Démonstration. Soient $0 \leq b \leq c \leq \ell$. Le résultat est évidemment vrai dans le cas $b = c$. Supposons donc $b < c$. Nous avons, par définition de l'application \mathcal{M}_H ,

$$\mathcal{M}_H(c, u_1, \dots, u_\ell) - \mathcal{M}_H(b, u_1, \dots, u_\ell) = c - b - \sum_{k=b+1}^c \left(1_{u_k > 1-q} + 1_{u_k < q/(\kappa-1)} \right).$$

Puisque $q \in [0, 1 - 1/\kappa]$, dans chaque paire d'indicatrices dans la somme, les deux ne peuvent pas être strictement positives simultanément. Donc la somme vaut au plus $c - b$, ce qui démontre le lemme. \square

Lemme 4.4.2. *Pour tout $k \geq 0$ et $i \geq 1$, la probabilité de mutation $M_H(k+i, k)$ tend vers 0 asymptotiquement, uniformément sur $i \geq 1$.*

Démonstration. Soient U_1, \dots, U_ℓ des variables aléatoires indépendantes de loi uniforme sur l'intervalle $[0, 1]$. Par le lemme précédent, nous avons, pour $k \geq 0$ et $i \geq 1$,

$$\begin{aligned} M_H(k+i, k) &\leq \sum_{l=0}^k M_H(k+i, l) = P(\mathcal{M}_H(k+i, U_1, \dots, U_\ell) \leq k) \\ &\leq P(\mathcal{M}_H(k+1, U_1, \dots, U_\ell) \leq k) = \sum_{l=0}^k M_H(k+1, l). \end{aligned}$$

Cette dernière somme ne dépend plus de $i \geq 1$ et tend vers 0 asymptotiquement. \square

Théorème 4.4.3. *Soit $(x(t))_{t \geq 0}$ une solution du système d'Eigen pour les classes de Hamming (\mathcal{S}^{sp}) et soit $(x^\infty(t))_{t \geq 0}$ une solution du système d'Eigen infini (\mathcal{S}_∞^{sp}). Supposons que l'on a la convergence des conditions initiales :*

$$\forall k \geq 0, \quad \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} x_k(0) = x_k^\infty(0).$$

Alors pour tout $T > 0$ et pour tout $k \geq 0$,

$$\lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \sup_{0 \leq t \leq T} |x_k(t) - x_k^\infty(t)| = 0.$$

Démonstration. Nous allons démontrer le théorème par récurrence. Soit $\varepsilon > 0$ et montrons d'abord qu'asymptotiquement,

$$\sup_{0 \leq t \leq T} |x_0(t) - x_0^\infty(t)| < \varepsilon.$$

Nous avons, pour tout $t \geq 0$,

$$x_0(t) = x_0(0) + \int_0^t \left(x_0(s) \sigma M_H(0,0) + \sum_{k=1}^{\ell} x_k(s) M_H(k,0) - x_0(s) ((\sigma - 1)x_0(s) + 1) \right) ds.$$

De même, pour tout $t \geq 0$,

$$x_0^\infty(t) = x_0^\infty(0) + \int_0^t \left(x_0^\infty(s) \sigma M_\infty(0,0) - x_0^\infty(s) ((\sigma - 1)x_0^\infty(s) + 1) \right) ds.$$

Soit $\delta > 0$. Asymptotiquement, par le lemme 4.4.2, pour tout $s \geq 0$,

$$\sum_{k=1}^{\ell} x_k(s) M_H(k,0) \leq \sum_{k=1}^{\ell} x_k(s) \delta \leq \delta.$$

Nous avons aussi la convergence des conditions initiales, d'où, asymptotiquement,

$$|x_0(0) - x_0^\infty(0)| \leq \delta.$$

Donc, pour tout $t \in [0, T]$,

$$|x_0(t) - x_0^\infty(t)| \leq \delta(1 + T) + \int_0^t \left(|x_0(s) \sigma M_H(0,0) - x_0^\infty(s) \sigma M_\infty(0,0)| + (\sigma - 1) |x_0(s)^2 - x_0^\infty(s)^2| + |x_0(s) - x_0^\infty(s)| \right) ds.$$

Or, asymptotiquement, $|M_H(0,0) - M_\infty(0,0)| < \delta$, et

$$|x_0(s)^2 - x_0^\infty(s)^2| = |(x_0(s) + x_0^\infty(s))(x_0(s) - x_0^\infty(s))| \leq 2|x_0(s) - x_0^\infty(s)|.$$

Nous en déduisons que, pour tout $t \in [0, T]$,

$$|x_0(t) - x_0^\infty(t)| \leq \delta(1 + (1 + \sigma)T) + (3\sigma - 1) \int_0^t |x_0(s) - x_0^\infty(s)| ds.$$

Par le lemme de Gronwall, pour tout $t \in [0, T]$,

$$|x_0(t) - x_0^\infty(t)| \leq \delta(1 + (1 + \sigma)T) e^{(3\sigma - 1)t} \leq \delta(1 + (1 + \sigma)T) e^{(3\sigma - 1)T}.$$

Nous choisissons $\delta < \varepsilon(1 + (1 + \sigma)T)^{-1}e^{-(3\sigma-1)T}$ et nous concluons, qu'asymptotiquement,

$$\sup_{0 \leq t \leq T} |x_0(t) - x_0^\infty(t)| < \varepsilon.$$

Soit maintenant $k \geq 1$ et supposons que le résultat du théorème est vrai pour tout $0 \leq i \leq k-1$. Soit $\varepsilon > 0$ et montrons qu'alors, asymptotiquement,

$$\sup_{0 \leq t \leq T} |x_k(t) - x_k^\infty(t)| < \varepsilon.$$

Nous avons, pour tout $t \geq 0$,

$$x_k(t) = x_k(0) + \int_0^t \left(x_0(s)\sigma M_H(0, k) + \sum_{i=1}^{\ell} x_i(s)M_H(i, k) - x_k(s)((\sigma-1)x_0(s) + 1) \right) ds.$$

De même, pour tout $t \geq 0$,

$$x_k^\infty(t) = x_k^\infty(0) + \int_0^t \left(x_0^\infty(s)\sigma M_\infty(0, k) + \sum_{i=1}^k x_i^\infty(s)M_\infty(i, k) - x_k^\infty(s)((\sigma-1)x_0^\infty(s) + 1) \right) ds.$$

Soit $\delta > 0$. Asymptotiquement

$$\begin{aligned} \forall i \in \{0, \dots, k\}, & \quad |M_H(i, k) - M_\infty(i, k)| < \delta, \\ \forall i \in \{k+1, \dots, \ell\}, & \quad M_H(i, k) < \delta, \\ \forall s \in [0, T], & \quad \max_{0 \leq i < k} |x_i(s) - x_i^\infty(s)| < \delta, \\ & \quad |x_k(0) - x_k^\infty(0)| < \delta. \end{aligned}$$

Nous en déduisons qu'asymptotiquement, pour tout $t \in [0, T]$,

$$|x_k(t) - x_k^\infty(t)| \leq \delta C_1 + C_2 \int_0^t |x_k(s) - x_k^\infty(s)| ds,$$

où C_1, C_2 sont des constantes positives qui ne dépendent ni de ε ni de δ . Nous concluons à l'aide du lemme de Gronwall, en choisissant $\delta < \varepsilon C_1^{-1}e^{-C_2 t}$. \square

Théorème 4.4.4. Soit $u = (u_k)_{0 \leq k \leq \ell}$ la solution stationnaire unique du système d'Eigen pour les classes de Hamming (\mathcal{S}^{sp}).

- Si $\sigma e^{-a} \leq 1$,

$$\forall k \geq 0, \quad \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} u_k = 0.$$

- Si $\sigma e^{-a} > 1$,

$$\forall k \geq 0, \quad \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} u_k = x_k^* = (\sigma e^{-a} - 1) \frac{a^k}{k!} \sum_{i \geq 1} \frac{i^k}{\sigma^i}.$$

Démonstration. Définissons la matrice $(W(i, j), 0 \leq i, j \leq \ell)$ par :

$$\forall i, j \in \{0, \dots, \ell\}, \quad W(i, j) = f_H(i) M_H(i, j).$$

Le vecteur unitaire u vérifie l'équation

$$\phi u_k = \sum_{i=0}^{\ell} u_i W(i, k), \quad 0 \leq k \leq \ell,$$

ou encore sous forme matricielle, $\phi^t u = {}^t u W$, où

$$\phi = \sum_{i=0}^{\ell} u_i f_H(i) = (\sigma - 1)u_0 + 1.$$

En particulier, $\phi \in]1, \sigma[$. La valeur ϕ est aussi la valeur propre maximale de la matrice W (théorème 3.3.3). Quitte à extraire une sous-suite, nous pouvons supposer l'existence des limites suivantes :

$$\phi^* = \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \phi, \quad u_k^* = \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} u_k, \quad k \geq 0.$$

D'après la première équation du système $\phi^t u = {}^t u W$, nous avons

$$u_0 \sigma M_H(0, 0) < \phi u_0 < u_0 \sigma M_H(0, 0) + \max_{1 \leq i \leq \ell} M_H(i, 0).$$

Puisque $\phi > 1$, et qu'en plus $M_H(0, 0)$ converge vers e^{-a} , nous concluons que $\phi^* \geq \max\{1, \sigma e^{-a}\}$. Remarquons que

$$\phi = (\sigma - 1)u_0 + 1.$$

En passant à la limite dans les deux formules précédentes,

$$\phi^* = (\sigma - 1)u_0^* + 1, \quad \phi^* u_0^* = u_0^* \sigma e^{-a}.$$

Puisque $\phi^* \geq \max\{1, \sigma e^{-a}\}$, nous concluons que :

- Si $\sigma e^{-a} \leq 1$, alors $\phi^* = 1$ et $u_0^* = 0$.
- Si $\sigma e^{-a} > 1$, alors

$$\phi^* = \sigma e^{-a} \quad \text{et} \quad u_0^* = \frac{\sigma e^{-a} - 1}{\sigma - 1}.$$

De même, d'après la k -ième équation du système $\phi u^t = u^t W$, nous avons

$$u_0 \sigma M_H(0, k) + \sum_{i=1}^k u_i M_H(i, k) < \phi u_k < u_0 \sigma M_H(0, k) + \sum_{i=1}^k u_i M_H(i, k) + \max_{k < i \leq \ell} M_H(i, k).$$

En passant à la limite, nous obtenons la relation de récurrence

$$u_k^* \sigma e^{-a} = u_0^* \sigma e^{-a} \frac{a^k}{k!} + \sum_{i=1}^k u_i^* e^{-a} \frac{a^{k-i}}{(k-i)!}, \quad k \geq 1.$$

Nous concluons que si $\sigma e^{-a} \leq 1$, alors $u_k^* = 0$ pour tout $k \geq 0$, et si $\sigma e^{-a} > 1$, alors $u_k^* = \rho_k^*$, pour tout $k \geq 0$. □

Nous avons donc montré les convergences suivantes :

- Pour $\sigma e^{-a} \leq 1$, soient $x(t)$ et $x^\infty(t)$, des solutions du système d'Eigen fini et infini respectivement. Si l'on a la convergence des conditions initiales $x(0)$ vers $x^\infty(0)$, alors pour tout $k \geq 0$,

$$\begin{array}{ccc} x_k(t) & \xrightarrow{t \rightarrow \infty} & u_k \\ \downarrow \begin{array}{l} \ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a \end{array} & & \downarrow \begin{array}{l} \ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a \end{array} \\ x_k^\infty(t) & \xrightarrow{t \rightarrow \infty} & 0 \end{array}$$

- Pour $\sigma e^{-a} > 1$, soient $x(t)$ et $x^\infty(t)$, des solutions du système d'Eigen fini et infini respectivement, avec $x_0^\infty(0) > 0$. Si l'on a la convergence des conditions

initiales $x(0)$ vers $x^\infty(0)$, alors pour tout $k \geq 0$,

$$\begin{array}{ccc}
 x_k(t) & \xrightarrow{t \rightarrow \infty} & u_k \\
 \downarrow \begin{array}{l} \ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a \end{array} & & \downarrow \begin{array}{l} \ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a \end{array} \\
 x_k^\infty(t) & \xrightarrow{t \rightarrow \infty} & \rho_k^*
 \end{array}$$

Dans les sections qui suivent, nous étendons les résultats obtenus pour le paysage à un pic à des fonctions de fitness $f_\infty : \mathbb{N} \rightarrow \mathbb{R}^+$ plus générales.

Chapitre 5

Le cas général

5.1 Solutions stationnaires : premières considérations

Dans cette section, nous notons la fonction de fitness f_∞ simplement par f , pour alléger les notations. Soit $f : \mathbb{N} \rightarrow \mathbb{R}^+$. Nous considérons le système d'équations différentielles

$$(S_\infty) \quad x'_k(t) = \sum_{i=0}^k x_i(t) f(i) e^{-a} \frac{a^{k-i}}{(k-i)!} - x_k(t) \sum_{i=0}^{\infty} x_i(t) f(i), \quad k \geq 0,$$

et nous nous intéressons aux solutions stationnaires du système. Nous présentons d'abord quelques remarques et lemmes qui justifient notre analyse ultérieure. Nous établissons aussi un résultat qui permet de calculer le paysage de fitness qui régit le système (S_∞) , lorsque l'on connaît les concentrations des différentes classes à l'équilibre.

Scaling. Remarquons d'abord que, tout comme le système d'Eigen général introduit dans la section 3.3, le système d'équations (S_∞) vérifie la propriété de scaling. C'est-à-dire qu'un changement dans la fonction de fitness du type $f \rightarrow cf$, pour $c > 0$, équivaut à une accélération du temps $t \rightarrow ct$. En particulier, les solutions stationnaires du système sont les mêmes pour tous les multiples d'une même fonction de fitness cf , $c > 0$.

Puisque nous voulons trouver les solutions stationnaires du système d'équa-

tions différentielles précédent, nous cherchons à résoudre le système d'équations

$$(\mathcal{S}_f) \quad 0 = \sum_{i=0}^k x_i f(i) e^{-a} \frac{a^{k-i}}{(k-i)!} - x_k \sum_{i=0}^{\infty} x_i f(i), \quad k \geq 0.$$

Comme précédemment, nous dénotons par $\phi(x)$ la fitness moyenne :

$$\phi(x) = \sum_{i=0}^{\infty} x_i f(i).$$

Nous supposons que la fonction de fitness a sa valeur maximale en 0, i.e., nous faisons l'hypothèse suivante :

Hypothèse \mathcal{H} . Nous supposons que la fonction f vérifie $f(0) > f(k)$, $k \geq 1$.

Nous disons que $(x_k)_{k \geq 0}$ est une distribution de quasi-espèce autour de 0 si elle est solution du système (\mathcal{S}_f) , si tous les x_k sont positifs ou nuls, et si elle vérifie de plus

$$x_0 > 0 \quad \text{et} \quad \sum_{k \geq 0} x_k = 1.$$

Notre objectif principal est de trouver des solutions stationnaires de (\mathcal{S}_f) qui sont des distributions de quasi-espèce autour de 0, pour des fonctions de fitness vérifiant l'hypothèse \mathcal{H} . L'hypothèse \mathcal{H} est justifiée pour les raisons suivantes. D'abord, la classe de Hamming 0 correspond à la master sequence (le génotype le plus adapté). Ensuite, toute solution stationnaire (non négative et unitaire) de (\mathcal{S}_g) pour une fonction de fitness g quelconque est une distribution de quasi-espèce autour de 0 pour une fonction de fitness f qui vérifie l'hypothèse \mathcal{H} . Nous détaillons ce dernier point dans les deux lemmes ci-dessous. Soit $g : \mathbb{N} \rightarrow \mathbb{R}^+$ une fonction de fitness quelconque (ne vérifiant pas forcément l'hypothèse \mathcal{H}).

Définition 5.1.1. Soit $K \geq 0$, nous appelons $(x_k)_{k \geq 0}$ une distribution de quasi-espèce autour de K associée à g si elle est une solution de (\mathcal{S}_g) , telle que tous les x_k sont positifs ou nuls et

$$x_0 = \dots = x_{K-1} = 0 < x_K \quad \text{et} \quad \sum_{k \geq K} x_k = 1.$$

Lemme 5.1.2. Soit $K \geq 0$, et définissons $f : \mathbb{N} \rightarrow \mathbb{R}^+$ par

$$\forall k \geq 0 \quad f(k) = g(K + k).$$

La suite $(x_k)_{k \geq 0}$ est une distribution de quasi-espèce autour de K associée à g si et seulement si la suite $(x_{k+i})_{i \geq 0}$ est une distribution de quasi-espèce autour de 0 associée à f .

Démonstration. Soit $(x_k)_{k \geq 0}$ une distribution de quasi-espèce autour de K associée à g . Puisque $x_0 = \dots = x_{K-1} = 0$, pour tout $k \geq K$, nous avons

$$0 = \sum_{i=K}^k x_i g(i) e^{-a} \frac{a^{k-i}}{(k-i)!} - x_k \sum_{i \geq K} x_i g(i).$$

Nous posons $j = k - K$ et $h = i - K$ dans la formule précédente et nous voyons que, pour tout $j \geq 0$,

$$\begin{aligned} 0 &= \sum_{h=0}^j x_{K+h} g(K+h) e^{-a} \frac{a^{j-h}}{(j-h)!} - x_{K+j} \sum_{h \geq 0} x_{K+h} g(K+h) \\ &= \sum_{h=0}^j x_{K+h} f(h) e^{-a} \frac{a^{j-h}}{(j-h)!} - x_{K+j} \sum_{h \geq 0} x_{K+h} f(h). \end{aligned}$$

Par conséquent, la suite $(x_{K+j})_{j \geq 0}$ est une distribution de quasi-espèce autour de 0 associée à f . L'implication réciproque se démontre de manière similaire. \square

Lemme 5.1.3. Soit $K \geq 1$ tel que

$$g(K) > \max_{0 \leq k < K} g(k).$$

Alors, pour k dans $\{0, \dots, K-1\}$, il n'existe pas de distribution de quasi-espèce autour de k associée à g .

Démonstration. Supposons que $(x_k)_{k \geq 0}$ est une solution de (S_g) . Soit k dans $\{0, \dots, K-1\}$ et supposons de plus que $x_0 = \dots = x_{k-1} = 0$ et $x_k > 0$. Nous allons montrer que si $x_{k+1} \geq 0, \dots, x_{K-1} \geq 0$, alors nécessairement $x_K < 0$. D'un côté, si nous écrivons la k -ième équation de (S_g) , nous en déduisons que $\phi(x) = g(k)e^{-a}$. D'un autre côté, si nous écrivons la K -ième équation du système (S_g) , nous voyons que

$$x_K = \frac{1}{\phi(x) - g(K)e^{-a}} \sum_{j=k}^{K-1} x_j g(j) e^{-a} \frac{a^{K-j}}{(K-j)!}.$$

Vu que $g(k) < g(K)$, si $x_k > 0$ et $x_{k+1}, \dots, x_{K-1} \geq 0$, nécessairement $x_K < 0$. Nous en déduisons qu'il n'existe pas de distribution de quasi-espèce autour de k associée à g . \square

Les lemmes précédents justifient l'hypothèse \mathcal{H} sur la fonction de fitness f , et motivent aussi la recherche des distributions de quasi-espèce autour de 0 associées à f . Dorénavant, nous supposons toujours que f vérifie l'hypothèse \mathcal{H} et nous ne nous intéresserons qu'aux distributions de quasi-espèce autour de 0. Si $(x_k)_{k \geq 0}$ est une distribution de quasi-espèce autour de 0 associée à f , et s'il n'y a pas de confusion possible, nous dirons simplement que $(x_k)_{k \geq 0}$ est une quasi-espèce. Ainsi, une quasi-espèce est une solution de (\mathcal{S}_f) telle que

$$x_0 > 0, \quad \forall k \geq 0 \quad x_k \geq 0, \quad \sum_{k \geq 0} x_k = 1.$$

Il est important de noter que, si $(x_k)_{k \geq 0}$ est une quasi-espèce, d'après la première équation du système (\mathcal{S}_f) , nous avons $\phi(x) = f(0)e^{-a}$. Cela entraîne notamment que toutes les classes ont une concentration strictement positive dans la quasi-espèce. En effet, pour $k \geq 1$, d'après la k -ième équation de (\mathcal{S}_f) ,

$$x_k = \frac{1}{\phi(x) - f(k)e^{-a}} \sum_{i=0}^{k-1} x_i f(i) e^{-a} \frac{a^{k-i}}{(k-i)!}.$$

Puisque $x_0 > 0$, que $\phi(x) = f(0)e^{-a}$ et que par l'hypothèse \mathcal{H} , nous avons $f(0) > f(k)$, nous en déduisons que $x_k > 0$ pour tout $k \geq 1$. Cette remarque est particulièrement importante, car elle nous permet d'exprimer les fitness des différentes classes en fonction des concentrations dans la quasi-espèce :

Théorème 5.1.4. *Supposons que $(x_k)_{k \geq 0}$ est une quasi-espèce. Alors,*

$$\forall k \geq 1 \quad f(k) = \frac{f(0)}{x_k} \sum_{j=0}^k (-1)^j \frac{a^j}{j!} x_{k-j}.$$

L'intérêt de ce résultat réside dans ses applications potentielles. En effet, dans les expériences pratiques, nous pouvons mesurer les concentrations des différents génotypes et nous essayons d'en déduire le paysage de fitness. Les techniques actuelles permettent même de séquencer des populations de virus *in vivo*, et le modèle de quasi-espèce est un des outils principaux pour inférer les paysages de fitness à partir de données expérimentales [69].

Démonstration. Nous faisons la preuve par récurrence. Puisque $(x_k)_{k \geq 0}$ est une quasi-espèce, en particulier, $x_0 > 0$ et $\phi(x) = f(0)e^{-a}$. Après une petite manipulation algébrique, nous obtenons

$$f(k) = \frac{1}{x_k} \left(x_k f(0) - \sum_{j=0}^{k-1} x_j f(j) \frac{a^{k-j}}{(k-j)!} \right), \quad k \geq 1.$$

En particulier, pour $k = 1$,

$$f(1) = \frac{f(0)}{x_1}(x_1 - ax_0).$$

Donc, la formule est vraie pour $k = 1$. Fixons maintenant $k > 1$ et supposons que la formule est vraie jusqu'au rang $k - 1$. En remplaçant $f(1), \dots, f(k - 1)$ dans la formule ci-dessus, nous obtenons

$$f(k) = \frac{f(0)}{x_k} \left(x_k - \sum_{j=0}^{k-1} \sum_{h=0}^j (-1)^h x_{j-h} \frac{a^h}{h!} \frac{a^{k-j}}{(k-j)!} \right).$$

Fixons $i \in \{1, \dots, k\}$ et cherchons le coefficient de x_{k-i} dans l'expression ci-dessus ; il est égal à

$$- \sum_{\substack{0 \leq h \leq j < k \\ j-h=k-i}} (-1)^h \frac{a^h}{h!} \frac{a^{k-j}}{(k-j)!} = - \sum_{0 \leq h < i} (-1)^h \frac{a^h}{h!} \frac{a^{i-h}}{(i-h)!} = (-1)^i \frac{a^i}{i!}.$$

Le théorème est donc démontré. \square

Revenons maintenant à notre objectif principal : chercher des quasi-espèces associées à une fonction de fitness f connue. Nous cherchons donc une formule d'inversion de la formule précédente. Puisque nous nous intéressons aux solutions du système (S_f) vérifiant $x_0 > 0$, nous remplaçons $\phi(x)$ par $f(0)e^{-a}$ dans (S_f) et nous obtenons une relation de récurrence pour la suite $(x_k)_{k \geq 0}$. Nous cherchons d'abord à résoudre la relation de récurrence avec condition initiale égale à 1 :

$$(R) \quad \begin{aligned} y_0 &= 1, \\ y_k &= \frac{1}{f(0) - f(k)} \sum_{j=0}^{k-1} y_j f(j) \frac{a^{k-j}}{(k-j)!}, \quad k \geq 1. \end{aligned}$$

Une fois que nous aurons la solution de la relation de récurrence (R) , nous pourrions décider de l'existence d'une quasi-espèce grâce au lemme suivant, et, le cas échéant, retrouver la quasi-espèce à partir de la solution de (R) .

Lemme 5.1.5. *Soit $(y_k)_{k \geq 0}$ la solution de la relation de récurrence (R) .*

• *Si la série de terme général y_k converge, il existe une unique quasi-espèce $(x_k)_{k \geq 0}$ autour de 0, donnée par :*

$$x_0 = \left(\sum_{k \geq 0} y_k \right)^{-1}, \quad x_k = x_0 y_k, \quad k \geq 1.$$

• *Si la série de terme général y_k diverge, il n'existe pas de quasi-espèce autour de 0.*

Démonstration. La première assertion du lemme est évidente. Pour démontrer la deuxième, il est suffisant de remarquer que, si $(x_k)_{k \geq 0}$ est une distribution de quasi-espèce, alors la suite $(y_k)_{k \geq 0}$ définie par la relation $y_k = x_k/x_0$, $k \geq 0$, vérifie la relation de récurrence (\mathcal{R}) , ce qui entraîne la convergence de la série de terme général y_k . \square

5.2 Deux formules pour la solution de (\mathcal{R})

Nous donnons ici deux formules explicites différentes pour la suite $(y_k)_{k \geq 0}$. La première formule exprime les valeurs $(y_k)_{k \geq 1}$ en fonction des fitness et des coefficients multinomiaux.

Théorème 5.2.1. *Pour tout $k \geq 1$,*

$$y_k = \frac{a^k f(0)}{k! f(k)} \times \sum_{\substack{1 \leq h \leq k \\ 0 = i_0 < \dots < i_h = k}} \frac{k!}{(i_1 - i_0)!(i_2 - i_1)! \dots (i_h - i_{h-1})!} \prod_{t=1}^h \frac{f(i_t)}{f(0) - f(i_t)}.$$

Démonstration. D'après la relation de récurrence (\mathcal{R}) , nous avons :

$$y_k (f(0) - f(k)) = \sum_{j=0}^{k-1} y_j f(j) \frac{a^{k-j}}{(k-j)!}, \quad k \geq 1.$$

Faisons les changements de variable suivants :

$$\begin{aligned} z_0 &= y_0, & g(0) &= f(0), \\ z_j &= \frac{y_j}{a^j} (f(0) - f(j)), & g(j) &= \frac{f(j)}{f(0) - f(j)}, \quad j \geq 1. \end{aligned}$$

La relation de récurrence se réécrit alors

$$z_k = \sum_{j=0}^{k-1} z_j \frac{g(j)}{(k-j)!}, \quad k \geq 1.$$

En itérant cette formule, nous obtenons, pour tout $k \geq 1$,

$$z_k = z_0 g(0) \left(\frac{1}{k!} + \sum_{h=1}^{k-1} \sum_{1 \leq i_1 < \dots < i_h < k} \frac{1}{i_1!(i_2 - i_1)! \dots (k - i_h)!} \prod_{t=1}^h g(i_t) \right).$$

En multipliant tout par $g(k)$, la formule se réécrit ainsi :

$$z_k g(k) = z_0 g(0) \sum_{h=1}^k \sum_{1 \leq i_1 < \dots < i_h = k} \frac{1}{i_1!(i_2 - i_1)! \dots (i_h - i_{h-1})!} \prod_{t=1}^h g(i_t).$$

Nous remplaçons z_0, z_k et $g(0), \dots, g(k-1)$ par leur valeur et nous obtenons la formule voulue. \square

Corollaire 5.2.2. *Il existe une quasi-espèce autour de 0 si et seulement si :*

$$\sum_{h \geq 1} \sum_{1 \leq i_1 < \dots < i_h} \frac{a^{i_h}}{f(i_h)} \prod_{1 \leq t \leq h} \frac{f(i_t)}{(i_t - i_{t-1})!(f(0) - f(i_t))} < +\infty.$$

Démonstration. Nous sommes sur y_k , pour $k \geq 1$, dans la formule du théorème et nous échangeons l'ordre des sommes pour obtenir la formule du corollaire. Nous concluons grâce au lemme 5.1.5. \square

Coefficients up-down. La suite $(y_k)_{k \geq 1}$ s'exprime en fonction des coefficients up-down [74], que nous définissons ici. Soit $\sigma = (\sigma(1), \dots, \sigma(n))$ une permutation de $\{1, \dots, n\}$. La signature de Niven associée à σ est la suite $(q_1, \dots, q_{n-1}) \in \{-1, +1\}^{n-1}$ telle que,

$$\forall i \in \{1, \dots, n-1\} \quad q_i(\sigma(i+1) - \sigma(i)) > 0.$$

Autrement dit, q_i vaut $+1$ si $\sigma(i+1) > \sigma(i)$ et q_i vaut -1 si $\sigma(i+1) < \sigma(i)$. Par exemple, la signature de Niven associée à la permutation $(3, 1, 5, 6, 2, 4)$ est $(-1, +1, +1, -1, +1)$. Étant donnée une signature de Niven, nous pouvons nous demander combien il existe de permutations ayant cette signature. Ces nombres sont donnés par les coefficients up-down.

Définition 5.2.3. *Soient $n \geq 1, 0 \leq h < n$ et $0 = i_0 < i_1 < \dots < i_h < n$. Le coefficient up-down*

$$\left\{ \begin{matrix} n \\ i_0, i_1, \dots, i_h \end{matrix} \right\}$$

est le nombre de permutations de $\{1, \dots, n\}$ ayant la signature de Niven (q_1, \dots, q_{n-1}) donnée par

$$\forall i \in \{1, \dots, n-1\} \quad q_i = \begin{cases} 1 & \text{si } i \in \{i_1, \dots, i_h\}, \\ -1 & \text{sinon.} \end{cases}$$

Théorème 5.2.4. *Pour tout $k \geq 1$, nous avons*

$$y_k = \frac{a^k}{k!} \left(\prod_{j=1}^k \frac{f(0)}{f(0) - f(j)} \right) \sum_{\substack{0 \leq h < k \\ 0 = i_0 < \dots < i_h < k}} \left\{ \begin{matrix} k \\ i_0, \dots, i_h \end{matrix} \right\} \prod_{t=1}^h \frac{f(i_t)}{f(0)}.$$

Démonstration. Nous reprenons la formule obtenue dans le théorème précédent, en utilisant $\prod_{j=1}^k (f(0) - f(j))$ comme dénominateur commun, nous obtenons ainsi

$$y_k = \frac{a^k}{k!} \frac{f(0)/f(k)}{\prod_{1 \leq j \leq k} (f(0) - f(j))} \times \sum_{\substack{1 \leq h \leq k \\ 1 \leq i_1 < \dots < i_h = k}} \frac{k!}{i_1!(i_2 - i_1)! \dots (i_h - i_{h-1})!} \prod_{t=1}^h f(i_t) \prod_{\substack{1 \leq j \leq k \\ j \neq i_1, \dots, i_h}} (f(0) - f(j)).$$

L'expression dans la deuxième ligne est un polynôme homogène de degré k dans les variables $f(0), \dots, f(k)$. À chaque $h \in \{1, \dots, k\}$ et chaque $1 \leq i_1 < \dots < i_h = k$ correspond le monôme $f(0)^{k-h} f(i_1) \dots f(i_h)$. Nous calculons le coefficient de chacun de ces monômes et nous concluons que

$$y_k = \frac{a^k}{k!} \frac{f(0)/f(k)}{\prod_{1 \leq j \leq k} (f(0) - f(j))} \sum_{\substack{1 \leq h \leq k \\ 1 \leq i_1 < \dots < i_h = k}} f(0)^{k-h} f(i_1) \dots f(i_h) \times \sum_{\substack{1 \leq t \leq h \\ 1 \leq j_1 < \dots < j_t = h}} (-1)^{h-t} \frac{k!}{i_{j_1}!(i_{j_2} - i_{j_1})! \dots (i_{j_t} - i_{j_{t-1}})!}.$$

Or, d'après [9], nous savons que

$$\left\{ \begin{matrix} k \\ 0, i_1, \dots, i_{h-1} \end{matrix} \right\} = \sum_{\substack{1 \leq t \leq h \\ 1 \leq j_1 < \dots < j_t = h}} (-1)^{h-t} \frac{k!}{i_{j_1}!(i_{j_2} - i_{j_1})! \dots (i_{j_t} - i_{j_{t-1}})!}.$$

Nous remplaçons dans la formule précédente, et puisque $i_h = k$ nous obtenons :

$$\begin{aligned} y_k &= \frac{a^k}{k!} \frac{f(0)/f(k)}{\prod_{1 \leq j \leq k} (f(0) - f(j))} \sum_{\substack{1 \leq h \leq k \\ 1 \leq i_1 < \dots < i_h = k}} f(0)^{k-h} \prod_{t=1}^h f(i_t) \left\{ \begin{matrix} k \\ 0, i_1, \dots, i_{h-1} \end{matrix} \right\} \\ &= \frac{a^k}{k!} \prod_{1 \leq j \leq k} \frac{f(0)}{f(0) - f(j)} \sum_{\substack{1 \leq h < k \\ 0 = i_0 < \dots < i_h < k}} \prod_{t=1}^h \frac{f(i_t)}{f(0)} \left\{ \begin{matrix} k \\ i_0, \dots, i_h \end{matrix} \right\}. \end{aligned}$$

Ce qui conclut la démonstration du théorème. □

5.3 Fonctions de fitness constantes à partir d'un certain rang

Nous considérons dans cette section des fonctions f constantes à partir d'une certaine valeur N . Nous montrons que, pour de telles fonctions, la suite $(y_k)_{k > N}$ s'exprime en fonction des valeurs $(q_k)_{k \geq 0}$ de la quasi-espèce pour le paysage à un pic, données par

$$q_k = (f(0)/c - 1) \frac{a^k}{k!} \sum_{i \geq 1} \frac{i^k}{(f(0)/c)^i}.$$

Rappelons que $(q_k)_{k \geq 0}$ est la solution de la relation de récurrence (\mathcal{R}) pour la fonction $(f(0), c, c, \dots)$, où $0 < c < f(0)$.

Théorème 5.3.1. *Supposons qu'il existe un entier $N \geq 0$ et une constante $c > 0$ tels que la fonction $f : \mathbb{N} \rightarrow \mathbb{R}^+$ vérifie*

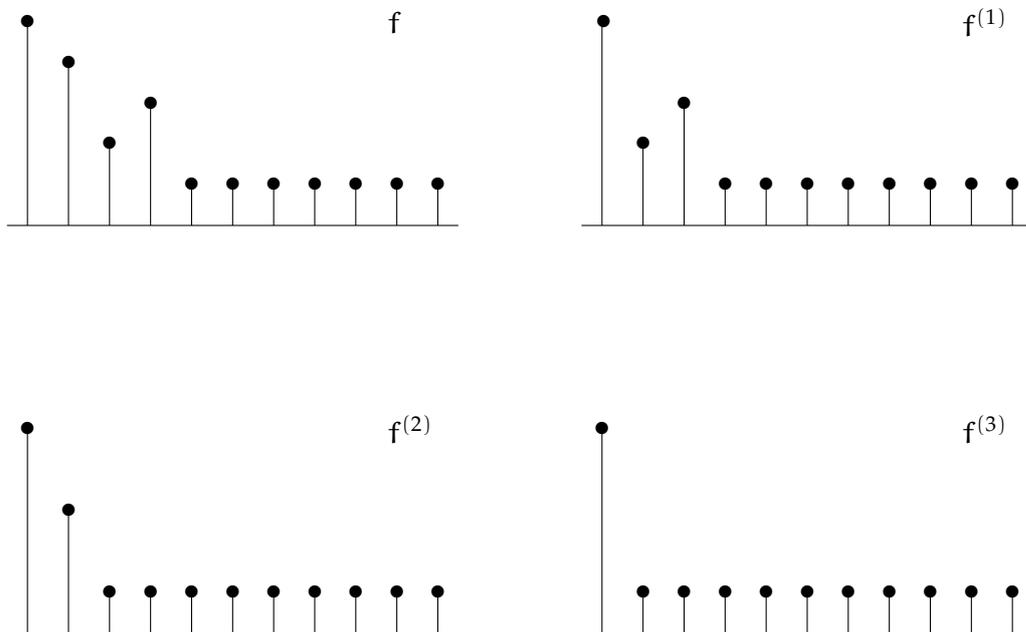
$$f(N) \neq c \quad \text{et} \quad \forall k > N \quad f(k) = c.$$

Alors, pour tout $k > N$,

$$\begin{aligned} y_k &= q_k + \sum_{j=1}^N \frac{a^j}{j!} q_{k-j} \left(\frac{f(j)}{f(0) - f(j)} - \frac{c}{f(0) - c} \right) \\ &\quad \times \left(1 + \sum_{h=1}^{j-1} \sum_{0=i_0 < \dots < i_h < j} \prod_{t=1}^h \binom{j - i_{t-1}}{j - i_t} \frac{f(i_t)}{f(0) - f(i_t)} \right). \end{aligned}$$

Avant de procéder à la preuve, nous donnons quelques résultats intermédiaires. Nous introduisons d'abord quelques notations. Pour un paysage de fitness f et $k \geq 0$, nous définissons le paysage de fitness $f^{(k)}$ obtenu en décalant de k vers la gauche les fitness des classes et en conservant la fitness de la classe 0, ce qui donne donc

$$\forall j \geq 0 \quad f^{(k)}(j) = \begin{cases} f(0) & \text{si } j = 0, \\ f(k+j) & \text{si } j \geq 1. \end{cases}$$



Pour un paysage de fitness f , et $k \geq 0$, nous appelons $(y_k(f))_{k \geq 1}$ la solution de la relation de récurrence (\mathcal{R}) pour le paysage de fitness f . Nous allons commencer par établir le lemme suivant, qui exprime la valeur $y_k(f)$ en fonction de $y_{k-1}(f^{(1)}), \dots, y_1(f^{(k-1)})$.

Lemme 5.3.2. *Pour tout $k \geq 2$, nous avons :*

$$y_k(f) = \frac{\alpha^k}{k!} \frac{f(0)}{f(0) - f(k)} + \sum_{j=1}^{k-1} \frac{\alpha^j}{j!} \frac{f(j)}{f(0) - f(j)} y_{k-j}(f^{(j)}).$$

Démonstration. Nous partons de l'identité obtenue dans le théorème 5.2.1,

$$y_k(f) = \frac{\alpha^k}{k!} \frac{f(0)}{f(0) - f(k)} \left(1 + \sum_{\substack{1 \leq h < k \\ 1 \leq i_1 < \dots < i_h < k}} \frac{k!}{i_1!(i_2 - i_1)! \dots (k - i_h)!} \prod_{t=1}^h \frac{f(i_t)}{f(0) - f(i_t)} \right),$$

et nous décomposons la somme selon la valeur du premier indice :

$$\begin{aligned} & \sum_{\substack{1 \leq h < k \\ 1 \leq i_1 < \dots < i_h < k}} \frac{k!}{i_1!(i_2 - i_1)! \dots (k - i_h)!} \prod_{t=1}^h \frac{f(i_t)}{f(0) - f(i_t)} = \sum_{i_1=1}^{k-1} \frac{f(i_1)}{f(0) - f(i_1)} \\ & \times \left(\frac{k!}{i_1!(k - i_1)!} + \sum_{\substack{2 \leq h \leq k - i_1 \\ i_1 < i_2 < \dots < i_h < k}} \frac{k!}{i_1!(i_2 - i_1)! \dots (k - i_h)!} \prod_{t=2}^h \frac{f(i_t)}{f(0) - f(i_t)} \right). \end{aligned}$$

Nous faisons les changements d'indices suivants :

$$\begin{aligned} j &= i_1, & h' &= h - 1, & t' &= t - 1, \\ i'_1 &= i_2 - i_1, & \dots, & & i'_{h'} &= i_h - i_1. \end{aligned}$$

Remarquons qu'en particulier nous avons $f(i_{t'}) = f(i_{t'} + i_1) = f^{(j)}(i'_{t'})$. L'expression précédente devient :

$$\begin{aligned} & \sum_{j=1}^{k-1} \frac{k!}{j!(k-j)!} \frac{f(j)}{f(0) - f(j)} \\ & \times \left(1 + \sum_{\substack{1 \leq h' < k-j \\ 1 \leq i'_1 < \dots < i'_{h'} < k-j}} \frac{(k-j)!}{i'_1!(i'_2 - i'_1)! \dots (k-j - i'_{h'})!} \prod_{t'=1}^{h'} \frac{f^{(j)}(i'_{t'})}{f^{(j)}(0) - f^{(j)}(i'_{t'})} \right). \end{aligned}$$

Vu que $f(0) = f^{(j)}(0)$ et $f(k) = f^{(j)}(k-j)$ pour tout $j \in \{1, \dots, k-1\}$, en enlevant les ' des indices, nous trouvons

$$\begin{aligned} y_k(f) &= \frac{\alpha^k}{k!} \frac{f(0)}{f(0) - f(k)} + \sum_{j=1}^{k-1} \frac{\alpha^j}{j!} \frac{f(j)}{f(0) - f(j)} \times \frac{\alpha^{k-j}}{(k-j)!} \frac{f^{(j)}(0)}{f^{(j)}(0) - f^{(j)}(k-j)} \\ & \times \left(1 + \sum_{\substack{1 \leq h < k-j \\ 1 \leq i_1 < \dots < i_h < k-j}} \frac{(k-j)!}{i_1!(i_2 - i_1)! \dots (k-j - i_h)!} \prod_{t=1}^h \frac{f^{(j)}(i_t)}{f^{(j)}(0) - f^{(j)}(i_t)} \right). \end{aligned}$$

Or, par le théorème 5.2.1,

$$y_{k-j}(f^{(j)}) = \frac{\alpha^{k-j}}{(k-j)!} \frac{f^{(j)}(0)}{f^{(j)}(0) - f^{(j)}(k-j)} \times \left(1 + \sum_{h=1}^{k-j} \sum_{1 \leq i_1 < \dots < i_h < k-j} \frac{(k-j)!}{i_1!(i_2 - i_1)! \dots (k-j - i_h)!} \prod_{t=1}^h \frac{f^{(j)}(i_t)}{f^{(j)}(0) - f^{(j)}(i_t)} \right).$$

Nous concluons en remplaçant dans la formule précédente. \square

Soit $f : \mathbb{N} \rightarrow \mathbb{R}^+$ une fonction de fitness telle qu'il existe $N \geq 0$ et une constante $c > 0$ vérifiant

$$f(N) \neq c \quad \text{et} \quad \forall k > N \quad f(k) = c.$$

Supposons que $(y_k(f))_{k \geq 0}$ est une solution de la relation de récurrence (\mathcal{R}) pour la fonction f . Nous voulons montrer que, pour tout $k > N$,

$$y_k(f) = q_k + \sum_{j=1}^N \frac{\alpha^j}{j!} q_{k-j} \left(\frac{f(j)}{f(0) - f(j)} - \frac{c}{f(0) - c} \right) \times \left(1 + \sum_{h=1}^{j-1} \sum_{0=i_0 < \dots < i_h < j} \prod_{t=1}^h \binom{j - i_{t-1}}{j - i_t} \frac{f(i_t)}{f(0) - f(i_t)} \right),$$

où $(q_k)_{k \geq 0}$ est la solution de la relation de récurrence (\mathcal{R}) pour le paysage de fitness à un pic $(f(0), c, c, \dots)$, i.e.,

$$q_k = (f(0)/c - 1) \frac{\alpha^k}{k!} \sum_{i \geq 1} \frac{i^k}{(f(0)/c)^i}, \quad k \geq 1.$$

Avant de faire la preuve du théorème 5.3.1, nous introduisons une notation afin de réécrire plus simplement la formule à démontrer. Pour une fonction de fitness f et $l \geq 1$ nous notons

$$C_l(f) = 1 + \sum_{h=1}^{l-1} \sum_{0=i_0 < i_1 < \dots < i_h < l} \prod_{t=1}^h \binom{l - i_{t-1}}{l - i_t} \frac{f(i_t)}{f(0) - f(i_t)}.$$

Lemme 5.3.3. *Les coefficients C_i , $i \geq 2$, vérifient la relation de récurrence*

$$C_i(f) = 1 + \sum_{j=1}^{i-1} \binom{i}{j} \frac{f(j)}{f(0) - f(j)} C_{i-j}(f^{(j)}).$$

Démonstration. Soit $i \geq 2$. Pour $j \in \{1, \dots, i-1\}$,

$$C_{i-j}(f^{(j)}) = 1 + \sum_{h=1}^{i-j-1} \sum_{0=l_0 < l_1 < \dots < l_h < i-j} \prod_{t=1}^h \binom{i-j-l_{t-1}}{i-j-l_t} \frac{f(j+l_t)}{f(0)-f(j+l_t)}.$$

Nous remplaçons $C_{i-1}(f^{(1)}), \dots, C_1(f^{(i-1)})$ dans la formule précédente, nous faisons le changement d'indices suivants :

$$h' = h + 1, \quad j = i_1, \quad j + l_1 = i_2, \quad \dots, \quad j + l_h = i_{h'},$$

nous échangeons l'ordre des sommes et nous obtenons la formule voulue pour $C_i(f)$. \square

Preuve du théorème 5.3.1. Nous allons faire la preuve par récurrence sur N . Supposons d'abord que $N = 1$ et soit $k \geq 2$. Alors toutes les fonctions de fitness $f^{(j)}$, $j \geq 1$, sont égales à la fonction de fitness du paysage à un pic. Donc, d'après le lemme 5.3.2 :

$$y_k(f) = \frac{a^k}{k!} \frac{f(0)}{f(0)-c} q_0 + a \frac{f(1)}{f(0)-f(1)} q_{k-1} + \sum_{j=2}^{k-1} \frac{a^j}{j!} \frac{c}{f(0)-c} q_{k-j}.$$

Or, la séquence $(q_k)_{k \geq 0}$ vérifie la relation de récurrence (\mathcal{R}) pour la fonction $(f(0), c, c, \dots)$, i.e.,

$$\forall k \geq 1 \quad q_k = \frac{1}{f(0)-c} \left(\frac{a^k}{k!} f(0) + c \sum_{j=1}^{k-1} \frac{a^j}{j!} q_{k-j} \right).$$

Il s'ensuit que

$$y_k(f) = q_k + a \left(\frac{f(1)}{f(0)-f(1)} - \frac{c}{f(0)-c} \right) q_{k-1}.$$

Le cas $N = 1$ est donc établi. Soit maintenant $N \geq 2$ et supposons que la formule du théorème 5.3.1 est vraie jusqu'au rang $N-1$. Soit $k > N$. D'une part, pour tout $j \geq N$, la fonction de fitness $f^{(j)}$ est égale à la fonction du paysage à un pic, donc $y_{k-j}(f^{(j)}) = q_{k-j}$ pour tout $j \in \{1, \dots, N\}$. D'autre part, $f(N+1) = \dots =$

$f(k) = c$. Nous avons alors, par le lemme 5.3.2,

$$\begin{aligned}
 y_k(f) &= \frac{a^k}{k!} \frac{f(0)}{f(0) - f(k)} + \sum_{j=1}^{k-1} \frac{a^j}{j!} \frac{f(j)}{f(0) - f(j)} y_{k-j}(f^{(j)}) \\
 &= \frac{1}{f(0) - c} \left(\frac{a^k}{k!} f(0) + c \sum_{j=1}^{k-1} \frac{a^j}{j!} q_{k-j} \right) - \frac{c}{f(0) - c} \sum_{j=1}^N \frac{a^j}{j!} q_{k-j} \\
 &\quad + \sum_{j=1}^N \frac{a^j}{j!} \frac{f(j)}{f(0) - f(j)} y_{k-j}(f^{(j)}) \\
 &= q_k - \frac{c}{f(0) - c} \sum_{j=1}^N \frac{a^j}{j!} q_{k-j} + \sum_{j=1}^N \frac{a^j}{j!} \frac{f(j)}{f(0) - f(j)} y_{k-j}(f^{(j)}).
 \end{aligned}$$

Par hypothèse de récurrence, pour tout $j \in \{1, \dots, N\}$, nous avons

$$\begin{aligned}
 y_{k-j}(f^{(j)}) &= q_{k-j} + \sum_{l=1}^{N-j} \frac{a^l}{l!} q_{k-j-l} \left(\frac{f^{(j)}(l)}{f^{(j)}(0) - f^{(j)}(l)} - \frac{c}{f^{(j)}(0) - c} \right) C_l(f^{(j)}) \\
 &= q_{k-j} + \sum_{l=1}^{N-j} \frac{a^l}{l!} q_{k-j-l} \left(\frac{f(j+l)}{f(0) - f(j+l)} - \frac{c}{f(0) - c} \right) C_l(f^{(j)}).
 \end{aligned}$$

Nous remplaçons $y_{k-N}(f^{(N)}), \dots, y_{k-1}(f^{(1)})$ dans la formule donnant $y_k(f)$ et nous obtenons

$$\begin{aligned}
 y_k(f) &= q_k - \frac{c}{f(0) - c} \sum_{j=1}^N \frac{a^j}{j!} q_{k-j} + \sum_{j=1}^N \frac{a^j}{j!} \frac{f(j)}{f(0) - f(j)} \\
 &\quad \times \left(q_{k-j} + \sum_{l=1}^{N-j} \frac{a^l}{l!} q_{k-j-l} \left(\frac{f(j+l)}{f(0) - f(j+l)} - \frac{c}{f(0) - c} \right) C_l(f^{(j)}) \right).
 \end{aligned}$$

Fixons $i \in \{1, \dots, N\}$. Le coefficient de q_{k-i} dans le développement de $y_k(f)$

sera donc égal à :

$$\begin{aligned} & \frac{\alpha^i}{i!} \left(\frac{f(i)}{f(0) - f(i)} - \frac{c}{f(0) - c} \right) \\ & + \sum_{\substack{1 \leq j \leq N \\ 1 \leq l \leq N-j \\ j+l=i}} \frac{\alpha^j}{j!} \frac{f(j)}{f(0) - f(j)} \frac{\alpha^l}{l!} \left(\frac{f(j+l)}{f(0) - f(j+l)} - \frac{c}{f(0) - c} \right) C_l(f^{(j)}) \\ & = \frac{\alpha^i}{i!} \left(\frac{f(i)}{f(0) - f(i)} - \frac{c}{f(0) - c} \right) \left(1 + \sum_{j=1}^{i-1} \binom{i}{j} \frac{f(j)}{f(0) - f(j)} C_{i-j}(f^{(j)}) \right). \end{aligned}$$

Nous concluons grâce au lemme 5.3.3. \square

Remarque 5.3.4. *Le résultat reste vrai pour des fonctions de fitness qui s'annulent à partir d'une certaine valeur $N \geq 1$. Le seul changement à faire est de remplacer c par 0 dans la formule et la suite $(q_k)_{k \geq 0}$ par la solution de (\mathcal{R}) associée à la fonction de fitness $(f(0), 0, 0, \dots)$, i. e., par $p_k = \alpha^k/k!$.*

5.4 Seuil d'erreur

Nous donnons ici une condition assurant l'existence de la quasi-espèce. Rappelons que nous disons que $(x_k)_{k \geq 0}$ est une distribution de quasi-espèce autour de 0, si, en plus de satisfaire le système (\mathcal{S}_f) , x_0 est strictement positif et la somme des x_k vaut 1.

Corollaire 5.4.1. *Nous avons :*

- Si $f(0)e^{-\alpha} > \limsup_{n \rightarrow \infty} f(n)$, alors la série de terme général y_k converge, et il existe une quasi-espèce autour de 0.
- Si $f(0)e^{-\alpha} < \liminf_{n \rightarrow \infty} f(n)$, alors la série de terme général y_k diverge, et il n'existe pas de quasi-espèce autour de 0.

D'après ce corollaire, nous voyons que le phénomène de seuil d'erreur observé par Eigen est présent pour les fonctions de fitness f qui vérifient l'hypothèse \mathcal{H} . De plus, le seuil d'erreur ne dépend que de la fitness $f(0)$ de la master sequence, du paramètre α , et du comportement limite de la fonction de fitness. Pour $\alpha > 0$ et f une fonction de fitness vérifiant l'hypothèse \mathcal{H} ainsi que $f(0)e^{-\alpha} > \limsup_{n \rightarrow \infty} f(n)$, nous pouvons donc définir une distribution de probabilité :

Définition 5.4.2. Nous disons qu'une variable aléatoire Y suit la loi de la quasi-espèce généralisée de paramètres f et α , notée $\mathcal{QG}(f, \alpha)$, si

$$P(Y = k) = \frac{\alpha^k f(0)}{f(k)} \times \sum_{\substack{1 \leq h \leq k \\ 1 \leq i_1 < \dots < i_h = k}} \frac{\prod_{t=1}^h \frac{f(i_t)}{(i_t - i_{t-1})!(f(0) - f(i_t))}}{\sum_{h \geq 1} \frac{\alpha^{i_h}}{f(i_h)} \prod_{1 \leq t \leq h} \frac{f(i_t)}{(i_t - i_{t-1})!(f(0) - f(i_t))}}.$$

Nous démontrons deux lemmes avant de faire la preuve du corollaire. Soient $f, g : \mathbb{N} \rightarrow \mathbb{R}_+$ deux fonctions de fitness vérifiant $f(0) = g(0)$ et

$$\forall k \geq 1 \quad f(k) \geq g(k).$$

Alors, pour tout $k \geq 1$, nous avons

$$\frac{1}{f(0) - f(k)} \sum_{j=0}^{k-1} y_j f(j) \frac{\alpha^{k-j}}{(k-j)!} \geq \frac{1}{g(0) - g(k)} \sum_{j=0}^{k-1} y_j g(j) \frac{\alpha^{k-j}}{(k-j)!}.$$

Si nous définissons $(y_k(f))_{k \geq 0}$ et $(y_k(g))_{k \geq 0}$ comme les solutions des relations de récurrence (\mathcal{R}) correspondant aux fonctions f et g , nous voyons par récurrence que, pour tout $k \geq 0$, $y_k(f) \geq y_k(g)$. Nous avons obtenu le lemme suivant :

Lemme 5.4.3. Soient $f, g : \mathbb{N} \rightarrow \mathbb{R}^+$ deux fonctions de fitness vérifiant $f(0) = g(0)$ et pour tout $k \geq 1$, $f(k) \geq g(k)$. Alors, pour tout $k \geq 0$, $y_k(f) \geq y_k(g)$.

Soit maintenant $N \geq 1$, $c \geq 0$, $\sigma > c$ et $g_N : \mathbb{N} \rightarrow \mathbb{R}^+$ la fonction de fitness définie par :

$$\forall k \geq 0, \quad g_N(k) = \begin{cases} \sigma & \text{si } k = 0, \\ 0 & \text{si } 1 \leq k \leq N, \\ c & \text{si } N + 1 \leq k. \end{cases}$$

Nous avons le résultat suivant :

Lemme 5.4.4. La série de terme général $y_k(g_N)$ converge si et seulement si $\sigma e^{-\alpha} > c$.

Démonstration. Nous savons que le résultat est vrai pour la fonction de fitness du paysage à un pic, i.e., pour $N = 0$. Par le lemme de comparaison 5.4.3, si $\sigma e^{-a} > c$, la série de terme général $y_k(g_N)$ converge. Il nous reste à voir l'implication réciproque. Supposons donc que $\sigma e^{-a} \leq c$. Par le lemme 5.1.5, nous savons que la convergence de la série de terme général $y_k(g_N)$ est équivalente à l'existence d'une quasi-espèce pour la fonction de fitness g_N . Montrons qu'une telle quasi-espèce ne peut pas exister si $\sigma e^{-a} \leq c$. Par l'absurde, supposons qu'une telle quasi-espèce $(x_k)_{k \geq 0}$ existe. Elle vérifie alors :

$$\begin{aligned} x_0 &> 0, \quad \sum_{k \geq 0} x_k = 1, \\ \phi &= \sigma x_0 + c \sum_{k > N} x_k, \\ 0 &= x_0(\sigma e^{-a} - \phi), \\ 0 &= x_0 \sigma \frac{a^k}{k!} e^{-a} - x_k \phi, \quad 1 \leq k \leq N. \end{aligned}$$

En particulier $\phi = \sigma e^{-a}$ et $x_k = x_0 a^k / k!$ pour tout $k \in \{1, \dots, N\}$. Il s'ensuit que

$$\sigma e^{-a} = \phi = \sigma x_0 + c(1 - (x_0 + \dots + x_N)) = \left(\sigma - c \sum_{k=0}^N \frac{a^k}{k!} \right) x_0 + c.$$

Posons

$$t_N(a) = \sum_{k=0}^N \frac{a^k}{k!}.$$

Nous concluons que x_0 est donné par la formule

$$x_0 = \frac{\sigma e^{-a} - c}{\sigma - c t_N(a)}.$$

Notons par a^* la seule solution positive de l'équation $c t_N(a) = \sigma$. Puisque l'expression obtenue pour x_0 n'est pas strictement positive quand a appartient à $[\ln(\sigma/c), a^*]$, une quasi-espèce ne peut pas exister quand a est dans cet intervalle. Si par contre $a \geq a^*$, nous avons

$$x_0 + \dots + x_N = t_N(a) x_0 = \frac{\sigma e^{-a} t_N(a) - c t_N(a)}{\sigma - c t_N(a)}.$$

Or, $t_N(a) < e^a$, ce qui implique que l'expression précédente est strictement plus grande que 1. Une quasi-espèce ne peut donc pas exister si $a \geq a^*$ non plus. \square

Démonstration du corollaire. D'après le lemme 5.1.5, afin de démontrer le corollaire 5.4.1, il suffit de montrer que si $f(0)e^{-a} > \limsup_{n \rightarrow \infty} f(n)$, la série de terme général $y_k(f)$ converge, et diverge si $f(0)e^{-a} < \liminf_{n \rightarrow \infty} f(n)$. Nous montrons d'abord que si $f(0)e^{-a} > \limsup_{n \rightarrow \infty} f(n)$ la série de terme général $y_k(f)$ converge. Pour cela supposons d'abord que la fonction de fitness f est constante égale à c à partir d'une certaine valeur $N \geq 0$. Nous pouvons alors appliquer le théorème 5.3.1, et nous avons donc

$$\forall k > N, \quad y_k(f) = q_k + \sum_{j=1}^N \frac{a^j}{j!} q_{k-j} \left(\frac{f(j)}{f(0) - f(j)} - \frac{c}{f(0) - c} \right) C_j(f).$$

Il s'ensuit que

$$\begin{aligned} \sum_{k \geq 0} y_k(f) &= \sum_{k=0}^N y_k(f) + \sum_{k > N} q_k \\ &\quad + \sum_{j=1}^N \frac{a^j}{j!} \left(\sum_{k > N} q_{k-j} \right) \left(\frac{f(j)}{f(0) - f(j)} - \frac{c}{f(0) - c} \right) C_j(f). \end{aligned}$$

Or, la série de terme général q_k converge si $f(0)e^{-a} > c$. Quand la fonction $f : \mathbb{N} \rightarrow \mathbb{R}^+$ n'est pas constante à partir d'un certain rang, nous posons

$$c^\infty = \limsup_{n \rightarrow \infty} f(n).$$

Soit $\varepsilon > 0$, il existe $N \geq 0$ tel que pour tout $k > N$, $f(k) < c^\infty + \varepsilon$. Définissons la fonction $f^N : \mathbb{N} \rightarrow \mathbb{R}_+$ par :

$$\forall k \geq 0 \quad f^N(k) = \begin{cases} f(k) & \text{si } 0 \leq k \leq N, \\ c^\infty + \varepsilon & \text{si } k > N, \end{cases}$$

et soit $(y_k(f^N))_{k \geq 0}$ la solution de la relation de récurrence (\mathcal{R}) correspondant à f^N . Pour ε suffisamment petit, $f^N(0)e^{-a} > c^\infty + \varepsilon$. Puisque f^N est constante égale à $c^\infty + \varepsilon$ à partir de la valeur N , la série de terme général $y_k(f^N)$ converge. D'après le lemme de comparaison 5.4.3, il en est de même de la série de terme général $y_k(f)$. Nous montrons ensuite que si $f(0)e^{-a} < \liminf_{n \rightarrow \infty} f(n)$, alors la série de terme général y_k diverge. Définissons la fonction $f_N : \mathbb{N} \rightarrow \mathbb{R}^+$ par :

$$\forall k \geq 0 \quad f_N(k) = \begin{cases} f(0) & \text{si } k = 0, \\ 0 & \text{si } 1 \leq k \leq N, \\ c^\infty - \varepsilon & \text{si } k > N, \end{cases}$$

et soit $(y_k(f_N))_{k \geq 0}$ la solution de la relation de récurrence (\mathcal{R}) correspondant à f_N . Pour ε suffisamment petit, $f^N(0)e^{-a} < c^\infty - \varepsilon$. Au vu du lemme 5.4.4, la série de terme général $y_k(f^N)$ diverge. D'après le lemme de comparaison 5.4.3, il en est de même de la série de terme général $y_k(f)$. \square

5.5 Convergence des trajectoires

Soit $f : \mathbb{N} \rightarrow \mathbb{R}^+$. Nous considérons le système d'Eigen infini :

$$(\mathcal{S}_\infty) \quad x'_k(t) = \sum_{i=0}^k x_i(t)f(i)e^{-a} \frac{a^{k-i}}{(k-i)!} - x_k(t) \sum_{i=0}^{\infty} x_i(t)f(i), \quad k \geq 0.$$

Comme dans la section précédente, nous supposons que $f(0) > f(k)$ pour tout $k \geq 1$. Nous supposons de plus que la fonction f converge et nous appelons $c \geq 0$ sa limite. Nous créons une séquence $(k_i)_{i \geq 0}$ de manière récurrente : nous posons $k_0 = 0$ si $f(0)e^{-a} > c$ et $k_0 = \infty$ sinon. Pour $i \geq 1$, nous définissons k_i par :

$$k_i = \inf\{k > k_{i-1} : f(k)e^{-a} > c \text{ et } f(k) > f(k+j), j \geq 1\}.$$

Au vu du corollaire 5.4.1 et des lemmes 5.1.2, 5.1.3, le système d'équations (\mathcal{S}_∞) a autant de solutions stationnaires qu'il y a d'éléments finis dans la séquence $(k_i)_{i \geq 1}$. Rappelons que nous ne considérons que des solutions stationnaires $(x_k)_{k \geq 0}$ vérifiant $\sum_{k \geq 0} x_k = 1$ (à l'exception de la solution identiquement nulle). À chacune de ces valeurs k_i , nous associons la solution stationnaire $(x_k^i)_{k \geq 0}$ correspondante, qui vérifie :

$$x_0^i = \dots = x_{k_i-1}^i = 0 < x_{k_i}^i.$$

Bien sûr, si $k_i = \infty$, la solution stationnaire associée est la solution identiquement nulle. Nous avons le résultat suivant :

Théorème 5.5.1. *Soit $(x(t))_{t \geq 0}$ une solution du système (\mathcal{S}_∞) . Pour tout $k \geq 0$ et pour tout $i \geq 1$,*

$$\lim_{t \rightarrow \infty} x_k(t) = x_k^i$$

si et seulement si la condition initiale vérifie

$$x_0(0) = \dots = x_{k_i-1}(0) = 0 \quad \text{et} \quad \max_{k_{i-1} < k \leq k_i} x_k(0) > 0.$$

Démonstration. Nous allons seulement démontrer que si $k_0 = 0$ et $x_0(0) > 0$, alors pour tout $k \geq 0$,

$$\lim_{t \rightarrow \infty} x_k(t) = x_k^* = x_k^0,$$

où $(x_k^*)_{k \geq 0}$ est la solution stationnaire de (S_∞) associée à 0, i.e., la distribution de quasi-espèce autour de 0 associée à f . Le reste du théorème se démontre de façon très similaire. Remarquons d'abord que si $x_0(0) > 0$, alors $x_0(t) > 0$ pour tout $t \geq 0$. En effet, puisque $\phi(t) \leq f(0)$,

$$x_0'(t) = x_0(t)f(0)e^{-a} - x_0(t)\phi(t) \geq x_0(t)f(0)(e^{-a} - 1).$$

Ceci entraîne que

$$x_0(t) \geq x_0(0)e^{-f(0)(1-e^{-a})t} > 0 \quad t \geq 0.$$

Nous pouvons donc faire le changement de variables suivant : pour tout $k \geq 0$, nous posons $y_k(t) = x_k(t)/x_0(t)$. En dérivant, nous obtenons le système d'équations différentielles

$$y_k'(t) = \sum_{i=0}^k y_i(t)f(i)e^{-a} \frac{a^{k-i}}{(k-i)!} - y_k(t)f(0)e^{-a}, \quad k \geq 1.$$

Le changement de variables nous a permis de rendre le système linéaire. Nous allons montrer par récurrence que pour tout $k \geq 0$, $y_k(t)$ converge vers y_k^* quand t tend vers l'infini, où $(y_k^*)_{k \geq 0}$ est la solution de la relation de récurrence (\mathcal{R}) de la section précédente. Pour $k = 0$ le résultat est évident, vu que $y_0(t) = 1$ pour tout $t \geq 0$. Soit $k \geq 1$ et supposons maintenant que $y_i(t)$ tend vers y_i^* pour tout $i \in \{0, \dots, k-1\}$. Nous avons

$$y_k'(t) = \sum_{i=0}^{k-1} y_i(t)f(i)e^{-a} \frac{a^{k-i}}{(k-i)!} - (f(0) - f(k))e^{-a}y_k(t).$$

Nous en déduisons que (annexe A) :

$$\lim_{t \rightarrow \infty} y_k(t) = \frac{1}{f(0) - f(k)} \sum_{i=0}^{k-1} y_i^* f(i) \frac{a^{k-i}}{(k-i)!} = y_k^*.$$

Ceci termine la récurrence. Il nous reste à montrer que

$$\lim_{t \rightarrow \infty} x_0(t) = x_0^*.$$

Nous avons :

$$x_0(t) = \left(\sum_{k=0}^{\infty} y_k(t) \right)^{-1} \quad \text{et} \quad x_0^* = \left(\sum_{k=0}^{\infty} y_k^* \right)^{-1}.$$

Nous montrons d'abord la convergence dans le cas des fonctions de fitness constantes à partir d'un certain rang, et nous étendons ensuite le résultat au cas général. Supposons donc qu'il existe $N \geq 0$ tel que la fonction fitness est constante égale à c pour tout $n > N$. Dans ce cas, La fitness moyenne $\phi(t)$ s'exprime en fonction de $x_0(t), \dots, x_N(t)$:

$$\phi(t) = \sum_{0 \leq k \leq N} x_k(t)(f(k) - c) + c = x_0(t) \sum_{0 \leq k \leq N} y_k(t)(f(k) - c) + c.$$

Il en est de même de la fitness moyenne limite ϕ^* :

$$\phi^* = \sum_{0 \leq k \leq N} x_k^*(f(k) - c) + c = x_0^* \sum_{0 \leq k \leq N} y_k^*(f(k) - c) + c.$$

Or, $\phi^* = f(0)e^{-a}$. Nous en déduisons que

$$\sum_{0 \leq k \leq N} y_k^*(f(k) - c) = \frac{f(0)e^{-a} - c}{x_0^*} > 0.$$

Posons :

$$\begin{aligned} \alpha(t) &= \sum_{0 \leq k \leq N} y_k(t)(f(k) - c), \\ \alpha^* &= \sum_{0 \leq k \leq N} y_k^*(f(k) - c), \\ \beta &= f(0)e^{-a} - c. \end{aligned}$$

L'équation différentielle pour $x_0(t)$ se réécrit ainsi :

$$x_0'(t) = x_0(t)(\beta - x_0(t)\alpha(t)).$$

Nous allons démontrer que $x_0(t)$ converge vers $\beta/\alpha^* = x_0^*$. Soit $\varepsilon > 0$ tel que $\alpha^* - \varepsilon > 0$, et soit $T \geq 0$ assez grand pour que

$$\forall t \geq T \quad |\alpha(t) - \alpha^*| < \varepsilon.$$

Alors, pour tout $t \geq T$, la dérivée $x_0'(t)$ est strictement positive sur $]0, \beta/(\alpha^* + \varepsilon)[$ et strictement négative sur $] \beta/(\alpha^* - \varepsilon), 1]$. Nous en déduisons qu'il existe $T_1 > T$ tel que

$$\forall t \geq T_1 \quad \frac{\beta}{\alpha^* + \varepsilon} \leq x_0(t) \leq \frac{\beta}{\alpha^* - \varepsilon}.$$

Nous faisons tendre ε vers 0 et nous obtenons la convergence de $x_0(t)$ vers x_0^* . Cela entraîne notamment la convergence :

$$\lim_{t \rightarrow \infty} \sum_{k \geq 0} y_k(t) = \sum_{k \geq 0} y_k^*.$$

Si f n'est pas constante à partir d'un certain rang, nous choisissons $\varepsilon > 0$ tel que $f(0)e^{-a} > c + \varepsilon$ et $N \geq 0$ tel que

$$\forall n > N \quad f(n) < c + \varepsilon.$$

Soit $f_N : \mathbb{N} \rightarrow \mathbb{R}^+$ la fonction définie par :

$$\forall n \geq 0 \quad f_N(n) = \begin{cases} f(n) & \text{si } n \leq N, \\ c + \varepsilon & \text{si } n > N. \end{cases}$$

Considérons le système d'équations différentielles

$$z_k'(t) = \sum_{0 \leq i \leq N} z_i(t) f_N(i) e^{-a} \frac{a^{k-i}}{(k-i)!} - z_k(t) f_N(0) e^{-a}.$$

Vu que $f \leq f_N$ et que $f(0) = f_N(0)$, si $y_k(0) \leq z_k(0)$ pour tout $k \geq 0$, nous avons, grâce au lemme A.2,

$$\forall k \geq 0 \quad \forall t \geq 0 \quad y_k(t) \leq z_k(t).$$

De plus, puisque f_N est constante à partir d'un certain rang, la série de terme général $z_k(t)$ converge vers la série de terme général z_k^* . Nous concluons qu'il en est de même de la série de terme général $y_k(t)$, comme voulu. \square

5.6 Convergence des systèmes

Considérons le système d'Eigen pour les classes de Hamming :

$$(S) \quad x_k'(t) = \sum_{i=0}^{\ell} x_i(t) f(i) M_H(i, k) - x_k(t) \phi(t), \quad 0 \leq k \leq \ell.$$

Sous l'hypothèse que la fonction f est strictement positive, nous avons démontré que le système d'équations différentielles (S) n'admet qu'une seule solution stationnaire $u = (u_k)_{0 \leq k \leq \ell}$, et que toutes les solutions du système convergent vers u . Quand nous regardons le régime asymptotique

$$\ell \rightarrow \infty, \quad q \rightarrow 0, \quad \ell q \rightarrow a \in]0, +\infty[,$$

nous obtenons le système d'équations limite (S_∞) . Nous supposons, comme dans la section précédente, que $f(0) > f(k)$ pour tout $k \geq 1$. Nous supposons de plus que la fonction f est strictement positive et qu'elle a une limite en $+\infty$, que nous notons $c \geq 0$. Nous montrons maintenant que, sous ces hypothèses, les solutions du système d'Eigen pour les classes de Hamming (S) convergent vers les solutions du système infini (S_∞) . Nous démontrons aussi que la solution stationnaire du système fini converge vers une solution stationnaire du système infini.

Théorème 5.6.1. *Soit $(x(t))_{t \geq 0}$ une solution du système d'Eigen pour les classes de Hamming (S) et soit $(x^\infty(t))_{t \geq 0}$ une solution du système d'Eigen infini (S_∞) . Supposons que l'on a la convergence des conditions initiales :*

$$\forall k \geq 0, \quad \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} x_k(0) = x_k^\infty(0).$$

Alors, pour tout $T > 0$ et pour tout $k \geq 0$,

$$\lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \sup_{0 \leq t \leq T} |x_k(t) - x_k^\infty(t)| = 0.$$

Démonstration. Soient $\varepsilon, \delta > 0$ et soit N assez grand pour que

$$\forall n \geq N, \quad |f(n) - c| < \delta.$$

Nous allons montrer que, pour tout $n \geq N$, asymptotiquement,

$$\sum_{k=0}^n |x_k(t) - x_k^\infty(t)| < \varepsilon.$$

Nous avons, pour tout $k \geq 0$ et $t \geq 0$,

$$x_k(t) = x_k(0) + \int_0^t \left(\sum_{i=0}^{\ell} x_i(s) f(i) M_H(i, k) - x_k(s) \phi(s) \right) ds,$$

$$x_k^\infty(t) = x_k^\infty(0) + \int_0^t \left(\sum_{i=0}^k x_i^\infty(s) f(i) M_\infty(i, k) - x_k^\infty(s) \phi^\infty(s) \right) ds.$$

Soit $n \geq N$. Asymptotiquement, pour tout $k \in \{1, \dots, n\}$,

$$\begin{aligned} \forall i \in \{0, \dots, k\}, & \quad |M_H(i, k) - M_\infty(i, k)| < \delta, \\ \forall i \in \{k+1, \dots, \ell\}, & \quad M_H(i, k) < \delta, \\ & \quad |x_k(0) - x_k^\infty(0)| < \delta. \end{aligned}$$

De plus, pour tout $t \geq 0$,

$$\begin{aligned} \left| \phi(t) - \sum_{k=0}^N x_k(t)(f(k) - c) - c \right| &= \left| \sum_{k=N+1}^{\ell} x_k(t)(f(k) - c) \right| < \delta, \\ \left| \phi^\infty(t) - \sum_{k=0}^N x_k^\infty(t)(f(k) - c) - c \right| &= \left| \sum_{k \geq N+1} x_k^\infty(t)(f(k) - c) \right| < \delta. \end{aligned}$$

Nous en déduisons que, pour tout $t \in [0, T]$,

$$\begin{aligned} |x_k(t) - x_k^\infty(t)| &\leq |x_k(0) - x_k^\infty(0)| + \\ &\quad \sum_{i=0}^k \int_0^t |x_i(s)f(i)M_H(i, k) - x_i^\infty(s)f(i)M_\infty(i, k)| ds + \\ &\quad \sum_{i=k+1}^{\ell} \int_0^t x_i(s)f(i)M_H(i, k) ds + \int_0^t |x_k(s)\phi(s) - x_k^\infty(s)\phi^\infty(s)| ds \\ &\leq \delta + \sum_{i=0}^k f(i)M_\infty(i, k) \int_0^t |x_i(s) - x_i^\infty(s)| ds + f(0)T\delta + 2f(0)T\delta + \\ &\quad \sum_{i=0}^N |f(i) - c| \int_0^t |x_k(s)x_i(s) - x_k^\infty(s)x_i^\infty(s)| ds + c \int_0^t |x_k(s) - x_k^\infty(s)| ds + 2T\delta. \end{aligned}$$

Or,

$$\begin{aligned} &|x_k(s)x_i(s) - x_k^\infty(s)x_i^\infty(s)| \\ &= |x_k(s)x_i(s) - x_k(s)x_i^\infty(s) + x_k(s)x_i^\infty(s) - x_k^\infty(s)x_i^\infty(s)| \\ &\leq |x_k(s) - x_k^\infty(s)| + |x_i(s) - x_i^\infty(s)|. \end{aligned}$$

En appelant C le maximum des $|f(i) - c|$ pour $1 \leq i \leq N$, il vient

$$\begin{aligned} |x_k(t) - x_k^\infty(t)| &\leq \delta(1 + 2(f(0) + 1)T) + \\ &\quad \sum_{i=0}^k f(0)M_\infty(i, k) \int_0^t |x_i(s) - x_i^\infty(s)| ds + \\ &\quad (C(N + 1) + c) \int_0^t |x_k(s) - x_k^\infty(s)| ds + C \sum_{i=0}^N \int_0^t |x_i(s) - x_i^\infty(s)| ds. \end{aligned}$$

Nous sommions pour $0 \leq k \leq n$ et nous obtenons

$$\begin{aligned} \sum_{k=0}^n |x_k(t) - x_k^\infty(t)| &\leq \delta(n+1)(1+2(f(0)+1)T) + \\ &\sum_{i=0}^n f(0) \left(\sum_{k=i}^n M_\infty(i,k) \right) \int_0^t |x_i(s) - x_i^\infty(s)| ds + \\ &\sum_{k=0}^n (C(N+1)+c) \int_0^t |x_k(s) - x_k^\infty(s)| ds + \\ &(n+1)C \sum_{i=0}^n \int_0^t |x_i(s) - x_i^\infty(s)| ds. \end{aligned}$$

Nous concluons que

$$\sum_{k=0}^n |x_k(t) - x_k^\infty(t)| \leq \delta C_1 + C_2 \int_0^t \left(\sum_{k=0}^n |x_k(s) - x_k^\infty(s)| \right) ds,$$

où C_1, C_2 sont des constantes positives qui ne dépendent ni de ℓ ni de q . Nous concluons à l'aide du lemme de Gronwall, en choisissant $\delta < \varepsilon C_1^{-1} e^{-C_2 t}$. \square

Proposition 5.6.2. Soit $u = (u_k)_{0 \leq k \leq \ell}$ la solution stationnaire unique du système d'Eigen pour les classes de Hamming (\mathcal{S}) .

- Si $f(0)e^{-a} \leq c$,

$$\forall k \geq 0, \quad \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} u_k = 0.$$

- Si $f(0)e^{-a} > c$,

$$\forall k \geq 0, \quad \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} u_k = x_k^*,$$

où $(x_k^*)_{k \geq 0}$ est la quasi-espèce autour de 0 associée à la fonction f .

Démonstration. Définissons la matrice $(W(i, j), 0 \leq i, j \leq \ell)$ par :

$$\forall i, j \in \{0, \dots, \ell\}, \quad W(i, j) = f(i)M_H(i, j).$$

Le vecteur u vérifie l'équation

$$\phi u_k = \sum_{i=0}^{\ell} u_i W(i, k), \quad 0 \leq k \leq \ell,$$

où

$$\phi = \sum_{i=0}^{\ell} u_i f(i),$$

est aussi la valeur propre maximale de la matrice W . En particulier, $\phi \in]0, \sigma[$. Quitte à extraire une sous-suite, nous pouvons supposer l'existence des limites suivantes :

$$\phi^* = \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \phi, \quad u_k^* = \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} u_k, \quad k \geq 0.$$

D'après la k -ième équation du système $\phi u^T = u^T W$, nous avons que

$$\sum_{i=0}^k u_i f(i) M_H(i, k) < \phi u_k < \sum_{i=0}^k u_i f(i) M_H(i, k) + f(0) \max_{k < i \leq \ell} M_H(i, k).$$

En particulier, d'après l'inégalité de droite avec $k = 0$, nous obtenons en divisant par u_0 et en passant à la limite que $\phi^* \geq f(0)e^{-a}$. En passant à la limite nous obtenons le système d'équations

$$\phi^* u_k^* = \sum_{i=0}^k u_i^* f(i) e^{-a} \frac{a^{k-i}}{(k-i)!}, \quad k \geq 0.$$

En particulier $\phi^* u_0^* = u_0^* f(0) e^{-a}$. Puisque la somme des composantes du vecteur $(u_k)_{k \geq 0}$ vaut 1, la suite $(u_k^*)_{k \geq 0}$ vérifie

$$\sum_{k \geq 0} u_k^* \leq 1.$$

Nous avons montré dans la section 5.1 que ce système n'admet que la solution triviale $u_k^* = 0$ quand $f(0)e^{-a} < c$. Par contre, si $f(0)e^{-a} > c$, nous voyons que forcément $u_0^* > 0$: en effet, si K est le premier indice $k \geq 0$ tel que $u_k^* > 0$, nous déduisons des inégalités ci-dessus que

$$\phi^* \leq f(K) e^{-a} < f(0) e^{-a},$$

en contradiction avec le fait que $\phi^* \geq f(0) e^{-a}$. De même, si $u_k^* = 0$ pour tout $k \geq 0$, alors en prenant N assez grand tel que, pour tout $n \geq N$, $|f(n) - c| < \varepsilon$, nous en déduisons que

$$\phi = \sum_{k=0}^{\ell} u_k f(k) \leq \sum_{k=0}^N u_k f(k) + (c + \varepsilon) \left(1 - \sum_{k=0}^N u_k \right).$$

Ceci entraîne que $\phi^* \leq c + \varepsilon$, ce qui, pour ε assez petit, est en contradiction avec $\phi^* \geq f(0)e^{-a}$. Donc $u_0^* > 0$ et $\phi^* = f(0)e^{-a}$. En sommant sur $k \geq 0$ dans le système d'équations, nous voyons qu'on a bien

$$\phi^* \sum_{k \geq 0} u_k^* = \sum_{k \geq 0} u_k^* f(k) = \phi^*.$$

Donc $(u_k^*)_{k \geq 0}$ est unitaire, et c'est bien la distribution de quasi-espèce autour de 0 associée à la fonction de fitness f . \square

Nous avons donc montré les convergences suivantes :

- Pour $f(0)e^{-a} \leq c$, soient $x(t)$ et $x^\infty(t)$, des solutions du système d'Eigen fini et infini respectivement. Si l'on a la convergence des conditions initiales, alors, pour tout $k \geq 0$,

$$\begin{array}{ccc} x_k(t) & \xrightarrow{t \rightarrow \infty} & u_k \\ \downarrow \begin{array}{l} \ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a \end{array} & & \downarrow \begin{array}{l} \ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a \end{array} \\ x_k^\infty(t) & \xrightarrow{t \rightarrow \infty} & 0 \end{array}$$

- Pour $f(0)e^{-a} > c$, soient $x(t)$ et $x^\infty(t)$, des solutions du système d'Eigen fini et infini respectivement, avec $x_0^\infty(0) > 0$. Si l'on a la convergence des conditions initiales, alors, pour tout $k \geq 0$,

$$\begin{array}{ccc} x_k(t) & \xrightarrow{t \rightarrow \infty} & u_k \\ \downarrow \begin{array}{l} \ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a \end{array} & & \downarrow \begin{array}{l} \ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a \end{array} \\ x_k^\infty(t) & \xrightarrow{t \rightarrow \infty} & x_k^* \end{array}$$

Chapitre 6

Conclusion

Nous avons présenté et étudié en détail le modèle d'Eigen. Les caractéristiques les plus remarquables de ce modèle sont, d'une part, le phénomène de seuil d'erreur pour le taux de mutation, et d'autre part, la distribution de quasi-espèce pour des taux de mutation sous-critiques.

Même si l'objectif principal du modèle d'Eigen est de décrire l'évolution d'une population prébiotique, le seuil d'erreur pour le taux de mutation et la distribution de quasi-espèce sont des phénomènes présents dans des populations plus complexes. Les prédictions du modèle d'Eigen s'appliquent particulièrement bien aux populations de virus [22, 23]. Cependant, les virus sont bien plus complexes que les organismes prébiotiques, notamment le nombre de génotypes possibles dépasse largement la taille de toute population viable [30, 85]. Ceci est une caractéristique essentielle des populations de virus que le modèle d'Eigen ne prend pas du tout en compte (la taille de la population est infinie alors que la longueur du génome ne l'est pas). C'est ce point qui a valu le plus de critiques à l'application du modèle d'Eigen en virologie [46]. En outre, la description de l'évolution d'une population basée sur des équations issues de la cinétique chimique n'est plus possible pour des individus d'une certaine complexité.

Il est donc souhaitable d'étudier des modèles à population finie, aléatoires, afin de retrouver un phénomène de seuil d'erreur pour le taux de mutation et une distribution de quasi-espèce. Dans ce but, dans les parties suivantes, nous allons étudier les trois modèles mathématiques les plus célèbres en génétique des

populations : le modèle de Galton–Watson, le modèle de Wright–Fisher, et le modèle de Moran. Nous allons considérer, pour ces trois modèles, le même cadre que celui que nous avons utilisé pour le modèle d’Eigen, c’est-à-dire, la même structure des génotypes, et le même mécanisme de reproduction et de mutation. Également, nous allons nous intéresser au régime asymptotique

$$\ell \rightarrow \infty, \quad q \rightarrow 0, \quad \ell q \rightarrow a.$$

Nous ferons simultanément tendre la taille de la population vers l’infini afin de mettre en évidence un phénomène de seuil d’erreur et une distribution de quasi-espèce.

Systemes d’équations différentielles

Dans la suite du texte, nous faisons souvent référence aux différents systèmes d’équations différentielles que nous avons introduits dans cette première partie. Afin de faciliter la tâche du lecteur, nous listons tous ces systèmes ici. Le système d’Eigen dans la forme la plus générale est

$$(*) \quad x'_k(t) = \sum_{i=1}^N x_i(t) f(i) M(i, k) - x_k(t) \phi(t), \quad 1 \leq k \leq N,$$

où les génotypes sont numérotés de 1 à N , $f : \{1, \dots, N\} \rightarrow \mathbb{R}^+$ est une fonction de fitness quelconque et $(M(i, k))_{1 \leq i, k \leq N}$ est une matrice stochastique. Le système original d’Eigen est :

$$x'_v(t) = \sum_{u \in \mathcal{A}^\ell} x_u(t) f(u) M(u, v) - x_v(t) \sum_{u \in \mathcal{A}^\ell} x_u(t) f(u), \quad v \in \mathcal{A}^\ell,$$

où $f : \mathcal{A}^\ell \rightarrow \mathbb{R}^+$ est une fonction quelconque et le noyau de mutation est donné par $M(u, v) = (q/(\kappa - 1))^{d_H(u, v)} (1 - q)^{\ell - d_H(u, v)}$. S’il existe une master sequence et la fonction de fitness a la même valeur sur tous les génomes à la même distance de la master sequence, nous obtenons le système sur les classes de Hamming,

$$(\mathcal{S}) \quad x'_k(t) = \sum_{i=0}^{\ell} x_i(t) f_H(i) M_H(i, k) - x_k(t) \sum_{i=0}^{\ell} x_i(t) f_H(i), \quad 0 \leq k \leq \ell.$$

Ici, $f_H : \{0, \dots, \ell\} \rightarrow \mathbb{R}^+$ représente la fonction de fitness induite sur les classes et le noyau de mutation sur les classes est donné par

$$M_H(i, k) = P(i - \text{Bin}(i, 1/(\kappa - 1)) + \text{Bin}(\ell - i, q) = k).$$

Quand la longueur des chaînes tend vers l'infini, la probabilité de mutation tend vers 0 et ℓq tend vers α , le système (S) converge vers un système infini. Dans le cas du paysage à un pic, ce système infini est

$$(\mathcal{S}_\infty^{\text{sp}}) \quad \begin{cases} x'_0(t) = \sigma x_0(t) e^{-\alpha} - x_0(t) \phi(t), \\ x'_k(t) = \sigma x_0(t) e^{-\alpha} \frac{\alpha^k}{k!} + \sum_{i=1}^k x_i(t) e^{-\alpha} \frac{\alpha^{k-i}}{(k-i)!} - x_k(t) \phi(t), \quad k \geq 1, \end{cases}$$

où $\sigma > 1$ représente l'avantage sélectif de la master sequence, et $\phi(t) = (\sigma - 1)x_0(t) + 1$ est la fitness moyenne de la population au temps t . Dans le cas général des fonctions de fitness qui dépendent de la classe de Hamming, le système (S) converge vers

$$(\mathcal{S}_\infty) \quad x'_k(t) = \sum_{i=0}^k x_i(t) f_H(i) e^{-\alpha} \frac{\alpha^{k-i}}{(k-i)!} - x_k(t) \sum_{i=0}^{\infty} x_i(t) f_H(i), \quad k \geq 0,$$

où cette fois-ci f_H est une fonction de \mathbb{N} dans \mathbb{R}^+ .

Deuxième partie

Le modèle de Galton–Watson

Chapitre 7

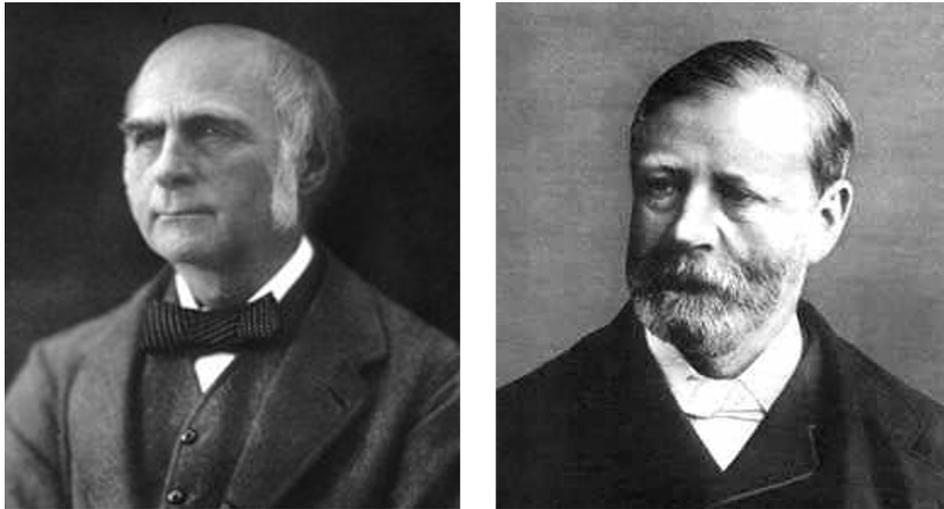
Le modèle de Galton–Watson

7.1 Introduction

Le modèle de Galton–Watson est l’un des plus célèbres modèles pour décrire l’évolution au cours du temps d’une population finie. Afin d’étudier l’extinction des noms de famille au sein de la noblesse anglaise, le modèle fut proposé par Francis Galton dans les années 1870, et étudié ensuite en détail par Henry William Watson [39]. Le modèle fut longtemps oublié jusqu’à sa reprise par Fisher [35], et Haldane [40], qui, l’appliquèrent avec succès à des problèmes en génétique des populations, comme par exemple, l’évolution d’un nouveau mutant dans une population.

Le modèle de Galton–Watson considère une population évoluant génération par génération. À chaque génération, chaque individu présent dans la population donne naissance à un nombre aléatoire d’individus, indépendamment des autres individus présents dans la population, et indépendamment du passé du processus. La loi décrivant le nombre de naissances par individu est fixée et commune à tous les individus et générations. Il s’agit d’un modèle avec une taille de population variable et dont les générations ne se superposent pas.

Dans le modèle de Galton–Watson initial, tous les individus sont identiques. Or, le modèle se généralise facilement pour intégrer des types d’individus différents, il est alors appelé processus de Galton–Watson à plusieurs types [42]. Ainsi, le modèle de Galton–Watson est rendu compatible avec le cadre du mo-



Francis Galton et Henry William Watson

dèle d'Eigen : un nombre fini de génotypes, une fonction de fitness et un noyau de mutation.

L'idée d'utiliser un processus de Galton–Watson comme version stochastique du modèle d'Eigen a déjà été mise en oeuvre avec succès par Demetrius, Schuster et Sigmund [19], puis généralisée par Antoneli, Bosco, Castro et Janini avec trois types différents de mutation : neutres, délétères, et avantageuses. Nous allons nous placer plutôt dans le cadre original de [19] : Les auteurs de [19] considèrent un nombre fini de génotypes avec une fonction de fitness et un noyau de mutation générales, et étudient la distribution des types quand la génération du processus de Galton–Watson tend vers l'infini. Ils montrent que cette distribution est la même que la seule solution stationnaire du système d'Eigen associé.

Notre objectif est de retrouver les formules explicites pour les différentes distributions de quasi-espèce, comme dans la partie I, pour cela, il est nécessaire de faire un passage à la limite adéquat car nous ne pouvons pas appliquer le résultat de [19] directement. Nous allons d'abord définir le modèle de Galton–Watson à plusieurs types, et introduire l'outil principal pour son étude : les fonctions génératrices. Nous regrouperons ensuite les différents génotypes dans des classes de Hamming, et nous construirons un nouveau processus de Galton–Watson, cette fois-ci avec un nombre de types bien plus petit. Nous montrerons aussi le lien avec le modèle d'Eigen, et pour finir, nous retrouverons les formules pour les différentes distributions de quasi-espèce.

7.2 Le modèle

Puisque notre objectif est de retrouver un phénomène de seuil d'erreur et une distribution de quasi-espèce pour le modèle de Galton–Watson, nous reprenons le cadre introduit dans la section 3.1 pour le modèle d'Eigen. Nous considérons donc des individus dont les génotypes sont des chaînes de longueur $\ell \geq 1$ sur un alphabet fini \mathcal{A} de cardinal $\kappa \geq 2$. Nous nous donnons également une fonction de fitness $f : \mathcal{A}^\ell \rightarrow \mathbb{R}^+$, et nous supposons que les mutations arrivent indépendamment au hasard sur chaque site de la chaîne avec probabilité $q \in]0, 1[$. La matrice de mutation $(M(u, v), u, v \in \mathcal{A}^\ell)$ est donc donnée par :

$$M(u, v) = \left(\frac{q}{\kappa - 1} \right)^{d_H(u, v)} (1 - q)^{\ell - d_H(u, v)}, \quad u, v \in \mathcal{A}^\ell.$$

Comme dans la partie précédente, $d_H(u, v)$ représente la distance de Hamming entre u et v , i.e., le nombre de digits où les deux chaînes diffèrent. Le processus de Galton–Watson à plusieurs types est une chaîne de Markov à valeurs dans \mathbb{N}^{κ^ℓ} ,

$$X_n = (X_n(u), u \in \mathcal{A}^\ell), \quad n \geq 0.$$

Pour chaque $u \in \mathcal{A}^\ell$ et $n \geq 0$, la quantité $X_n(u)$ représente le nombre de copies de u présentes dans la population à la génération n . À chaque génération, chaque individu de la population a un nombre d'enfants distribué selon une variable aléatoire de Poisson de paramètre sa fitness, indépendamment des autres individus de la population et du passé du processus. Puis chaque enfant mute selon le noyau de mutation M . La nouvelle génération est formée par l'ensemble des enfants après mutation.

Fonctions génératrices. Le processus de Galton–Watson décrit ci-dessus est classiquement étudié en utilisant des fonctions génératrices, qui servent aussi à décrire formellement le mécanisme de transition du processus. Soit $u \in \mathcal{A}^\ell$ et soit $g^u : [-1, 1]^{\mathcal{A}^\ell} \rightarrow \mathbb{R}$ la fonction définie par :

$$\forall s \in [-1, 1]^{\mathcal{A}^\ell} \quad g^u(s) = \sum_{r \in \mathbb{N}^{\mathcal{A}^\ell}} p^u(r) \prod_{v \in \mathcal{A}^\ell} s(v)^{r(v)},$$

où $p^u(r)$ est la probabilité qu'un individu avec génotype u ait $r(v)$ enfants avec génotype v , pour chaque $v \in \mathcal{A}^\ell$; cette probabilité est donnée par une multinomiale dont le paramètre est régi par une loi de Poisson :

$$\forall r \in \mathbb{N}^{\mathcal{A}^\ell} \quad p^u(r) = e^{-f(u)} f(u)^{|r|_1} \prod_{v \in \mathcal{A}^\ell} \frac{M(u, v)^{r(v)}}{r(v)!}.$$

Ici $\|\cdot\|_1$ représente la norme 1 du vecteur r , i.e., la somme de ses éléments. Pour une population X_0 formée par un seul individu $u \in \mathcal{A}^\ell$, X_1 est un vecteur aléatoire ayant par fonction génératrice g^u . En général, pour tout $n \geq 0$, si $X_n = r \in \mathbb{N}^{\mathcal{A}^\ell}$, alors X_{n+1} est la somme de $\|r\|_1$ vecteurs aléatoires indépendants ; dans cette somme, pour chaque $u \in \mathcal{A}^\ell$, $r(u)$ vecteurs ont pour fonction génératrice g^u . Le vecteur nul 0 est absorbant.

Remarque 7.2.1. *Le processus de Galton–Watson à plusieurs types peut être défini de manière plus générale pour n’importe quel ensemble fini ou dénombrable \mathcal{E} , fonction de fitness $f : \mathcal{E} \rightarrow \mathbb{R}_+$, et matrice stochastique $M : \mathcal{E} \times \mathcal{E} \rightarrow [0, 1]$.*

7.3 Le processus d’occupation

Comme pour le modèle d’Eigen, notre but est d’étudier le processus de Galton–Watson dans le régime asymptotique

$$\ell \rightarrow \infty \quad q \rightarrow 0 \quad \ell q \rightarrow a \in]0, \infty[,$$

afin de faire apparaître un phénomène de seuil d’erreur et une distribution de quasi-espèce. Nous supposons donc qu’il existe un génotype $w^* \in \mathcal{A}^\ell$ de fitness maximale, et que la fonction de fitness $f : \mathcal{A}^\ell \rightarrow \mathbb{R}^+$ respecte les classes de Hamming vis à vis de w^* . Le génotype w^* est appelé master sequence. Ainsi, nous supposons qu’il existe une fonction $f_H : \mathbb{N} \rightarrow \mathbb{R}^+$ vérifiant l’hypothèse \mathcal{H} de la partie précédente ($f_H(0) > f(k)$ pour tout $k \geq 1$) telle que

$$\forall u \in \mathcal{A}^\ell, \quad f(u) = f_H(d_H(u, w^*)).$$

Nous construisons maintenant un processus d’occupation

$$(O_n)_{n \geq 0} = (O_n(0), \dots, O_n(\ell))_{n \geq 0},$$

à partir du processus de Galton–Watson $(X_n)_{n \geq 0}$. La quantité $O_n(l)$ représente le nombre d’individus dans la population X_n qui sont à distance de Hamming l de la master sequence. Pour ce faire, nous utilisons la technique classique de *lumping* (annexe B). Notons par \mathcal{C}_l la classe de Hamming l , i.e., l’ensemble de séquences de \mathcal{A}^ℓ qui sont à distance l de la master sequence. Soit $\mathcal{O} : \mathbb{N}^{\mathcal{A}^\ell} \rightarrow \mathbb{N}^{\ell+1}$ l’application qui à chaque population $r \in \mathbb{N}^{\mathcal{A}^\ell}$ associe la distribution d’occupation correspondante :

$$\forall r \in \mathbb{N}^{\mathcal{A}^\ell} \quad \forall l \in \{0, \dots, \ell\} \quad \mathcal{O}(r)(l) = \sum_{u \in \mathcal{C}_l} r(u).$$

Le processus d'occupation $(O_n)_{n \geq 0}$ est défini par :

$$\forall n \geq 0 \quad O_n = \mathcal{O}(X_n).$$

Nous allons vérifier que le processus d'occupation est aussi un processus de Galton–Watson. Soient $k \in \{0, \dots, \ell\}$, $\mathbf{u} \in \mathcal{C}_k$, et $\mathbf{o} \in \mathbb{N}^{\ell+1}$. Nous avons

$$\begin{aligned} \sum_{\substack{\mathbf{r} \in \mathbb{N}^{\mathcal{A}^\ell} \\ \mathcal{O}(\mathbf{r}) = \mathbf{o}}} p^{\mathbf{u}}(\mathbf{r}) &= \sum_{\substack{\mathbf{r} \in \mathbb{N}^{\mathcal{A}^\ell} \\ \mathcal{O}(\mathbf{r}) = \mathbf{o}}} e^{-f(\mathbf{u})} f(\mathbf{u})^{|\mathbf{r}|_1} \prod_{\mathbf{v} \in \mathcal{A}^\ell} \frac{M(\mathbf{u}, \mathbf{v})^{r(\mathbf{v})}}{r(\mathbf{v})!} \\ &= e^{-f_H(k)} f_H(k)^{|\mathbf{o}|_1} \sum_{\substack{\mathbf{r} \in \mathbb{N}^{\mathcal{A}^\ell} \\ \mathcal{O}(\mathbf{r}) = \mathbf{o}}} \prod_{\mathbf{v} \in \mathcal{A}^\ell} \frac{M(\mathbf{u}, \mathbf{v})^{r(\mathbf{v})}}{r(\mathbf{v})!}. \end{aligned}$$

Puisque la matrice de mutation respecte les classes de Hamming (lemme 3.2.1), nous pouvons décomposer la somme et le produit à l'aide des classes de Hamming :

$$\sum_{\substack{\mathbf{r} \in \mathbb{N}^{\mathcal{A}^\ell} \\ \mathcal{O}(\mathbf{r}) = \mathbf{o}}} \prod_{\mathbf{v} \in \mathcal{A}^\ell} \frac{M(\mathbf{u}, \mathbf{v})^{r(\mathbf{v})}}{r(\mathbf{v})!} = \prod_{l=0}^{\ell} \left(\sum_{\substack{\mathbf{t} \in \mathbb{N}^{\mathcal{C}_l} \\ |\mathbf{t}|_1 = \mathbf{o}(l)}} \prod_{\mathbf{v} \in \mathcal{C}_l} \frac{M(\mathbf{u}, \mathbf{v})^{t(\mathbf{v})}}{t(\mathbf{v})!} \right) = \prod_{l=0}^{\ell} \frac{M_H(k, l)^{\mathbf{o}(l)}}{\mathbf{o}(l)!}.$$

Soit

$$p^k(\mathbf{o}) = \sum_{\substack{\mathbf{r} \in \mathbb{N}^{\mathcal{A}^\ell} \\ \mathcal{O}(\mathbf{r}) = \mathbf{o}}} p^{\mathbf{u}}(\mathbf{r}) = e^{-f_H(k)} f_H(k)^{|\mathbf{o}|_1} \prod_{l=0}^{\ell} \frac{M_H(k, l)^{\mathbf{o}(l)}}{\mathbf{o}(l)!}.$$

Puisque cette expression ne dépend de \mathbf{u} qu'à travers k , la somme

$$\sum_{\substack{\mathbf{r}' \in \mathbb{N}^{\ell+1} \\ \mathcal{O}(\mathbf{r}') = \mathbf{o}'}} P(X_{n+1} = \mathbf{r}' | X_n = \mathbf{r})$$

ne dépend de \mathbf{r} qu'à travers $\mathbf{o} = \mathcal{O}(\mathbf{r})$. Donc, d'après le théorème de lumping dans l'annexe B, le processus $(O_n)_{n \geq 0}$ est une chaîne de Markov. Définissons, pour $k \in \{0, \dots, \ell\}$ et $\mathbf{s} \in [0, 1]^{\ell+1}$,

$$g^k(\mathbf{s}) = \sum_{\mathbf{o} \in \mathbb{N}^{\ell+1}} p^k(\mathbf{o}) \prod_{l=0}^{\ell} s(l)^{\mathbf{o}(l)}.$$

Le processus $(O_n)_{n \geq 0}$ est un processus de Galton–Watson à $\ell + 1$ types, le mécanisme de transition étant donné par : pour tout $n \geq 0$ et $\mathbf{o} \in \mathbb{N}^{\ell+1}$, si $O_n = \mathbf{o}$,

alors O_{n+1} est la somme de $|o|_1$ vecteurs aléatoires indépendants ; dans cette somme, pour chaque $k \in \{0, \dots, \ell\}$, $o(k)$ vecteurs ont pour fonction génératrice g^k .

Dans les sections qui suivent nous étudions le processus d'occupation $(O_n)_{n \geq 0}$ dans le régime asymptotique

$$\ell \rightarrow \infty \quad q \rightarrow 0 \quad \ell q \rightarrow a \in]0, \infty[,$$

et nous montrons comment le phénomène de seuil d'erreur et la distribution de quasi-espèce apparaissent. Avant de passer à l'étude asymptotique du modèle, nous étudions le processus de Galton–Watson à plusieurs types, dans un cadre plus général, afin de révéler le lien entre le modèle d'Eigen et le modèle de Galton–Watson.

7.4 Lien avec le modèle d'Eigen

Nous considérons ici le modèle d'Eigen généralisé et le processus de Galton–Watson associé (remarque 7.2.1) pour présenter le lien entre les deux. Considérons donc un ensemble fini de génotypes \mathcal{E} , une fonction de fitness $f : \mathcal{E} \rightarrow \mathbb{R}^+$ et un noyau de mutation $M : \mathcal{E} \times \mathcal{E} \rightarrow [0, 1]$ (le noyau M est une matrice stochastique). Comme dans la section 3.3 de la première partie, nous identifions \mathcal{E} avec l'ensemble $\{1, \dots, N\}$, où $N \geq 2$ est le cardinal de \mathcal{E} , et nous parlons de i -ème génotype. Nous rappelons que le système d'Eigen s'écrit ainsi :

$$(*) \quad x'_k(t) = \sum_{i=1}^N x_i(t) f(i) M(i, k) - x_k(t) \sum_{i=1}^N x_i(t) f(i), \quad 1 \leq k \leq N.$$

Les solutions stationnaires de (*) et leur stabilité ont été étudiées dans la section 3.3 de la partie 1. Nous pouvons également définir un processus de Galton–Watson $(Z_n)_{n \geq 0}$ à partir de l'ensemble de génotypes $\{1, \dots, N\}$, la fonction de fitness f et le noyau de mutation M . En effet, pour chaque $k \in \{1, \dots, N\}$, définissons la fonction $g^k : [-1, 1] \rightarrow \mathbb{R}$ par :

$$\forall s \in [-1, 1]^N, \quad g^k(s) = \sum_{r \in \mathbb{N}^N} p^k(s) \prod_{1 \leq i \leq N} s(i)^{r(i)},$$

où $p^k(s)$ est la probabilité qu'un individu avec génotype k ait $r(i)$ enfants avec génotype i , pour $1 \leq i \leq N$, i.e.,

$$p^k(r) = e^{-f(k)} f(k)^{|r|_1} \prod_{1 \leq i \leq N} \frac{M(k, i)^{r(i)}}{r(i)!}.$$

Pour une population Z_0 formée par un seul individu ayant le génotype k , Z_1 est un vecteur aléatoire ayant pour fonction génératrice g^k . En général, pour tout $n \geq 0$, si $Z_n = r \in \mathbb{N}^N$, alors Z_{n+1} est la somme de $|r|_1$ vecteurs aléatoires indépendants ; dans cette somme, pour chaque $k \in \{1, \dots, N\}$, $r(k)$ vecteurs ont pour fonction génératrice g^k . Le vecteur nul 0 est absorbant.

Nous étudions maintenant le processus $(Z_n)_{n \geq 0}$ à l'aide de la théorie classique des processus de branchement [42]. L'étude du processus de Galton–Watson $(Z_n)_{n \geq 0}$ repose en grande partie sur l'étude de la matrice moyenne associée. La *matrice moyenne* W de $(Z_n)_{n \geq 0}$ a pour coefficients les espérances $W(i, j)$ du nombre d'individus avec génotype j dans la première génération, lorsque le processus commence avec un individu du type i , pour $1 \leq i, j \leq N$. La matrice moyenne est donc donnée par :

$$\forall i, j \in \{1, \dots, N\}, \quad W(i, j) = f(i)M(i, j).$$

Remarquons ici qu'il s'agit de la matrice W qui est au cœur de l'étude du système d'Eigen (*) (cf. la section 3.3 de la partie I). Comme pour le modèle d'Eigen, nous supposons que la matrice W vérifie l'hypothèse suivante :

Hypothèse. La matrice W a toutes ses composantes strictement positives.

Sous cette hypothèse, le théorème de Perron–Frobenius s'applique à la matrice W : il existe une unique valeur propre maximale λ et un unique vecteur propre positif et unitaire u associé à λ . D'après la théorie générale des processus de Galton–Watson à plusieurs types (chapitre II, théorèmes 7.1 et 9.2 dans [42]), nous avons :

Théorème 7.4.1. *Si $\lambda \leq 1$, la population s'éteint presque sûrement. Si $\lambda > 1$, la population a une probabilité positive de survie, et conditionnée à la survie, nous avons :*

$$\lim_{n \rightarrow \infty} \frac{Z_n(k)}{Z_n(1) + \dots + Z_n(N)} = u_k, \quad 1 \leq k \leq N.$$

La valeur propre λ représente la fitness moyenne d'une population distribuée

selon le vecteur u :

$$\lambda = \sum_{1 \leq k \leq N} u_k f(k).$$

Le théorème reste vrai pour des matrices W avec des éléments positifs ou nuls, sous l'hypothèse que la matrice W soit primitive.

Scaling. Soit $c > 0$ et considérons l'effet du changement $f \rightarrow cf$ dans la fonction de fitness. Appelons \widetilde{W} la matrice moyenne du processus $(Z_n)_{n \geq 0}$ associé à la fonction de fitness cf et soient $\widetilde{\lambda}$ et \widetilde{u} la valeur propre et le vecteur propre associés à \widetilde{W} par le théorème de Perron–Frobenius. Nous avons $\widetilde{W} = cW$, ce qui entraîne notamment que $\widetilde{\lambda} = c\lambda$ et $\widetilde{u} = u$. Nous en déduisons donc que pour toute fonction de fitness f , en prenant $c > 1/\lambda$, le processus $(Z_n)_{n \geq 0}$ associé à la fonction de fitness cf a une probabilité positive de survie, et conditionné à la survie

$$\lim_{n \rightarrow \infty} \frac{Z_n(k)}{Z_n(1) + \cdots + Z_n(N)} = u_k, \quad 1 \leq k \leq N.$$

Corollaire 7.4.2. Soit $(x_t)_{t \geq 0}$ une solution du système d'Eigen (*) associé à la fonction de fitness f . Soit $(Z_n)_{n \geq 0}$ le processus associé à la fonction de fitness cf , avec $c > 1/\lambda$, conditionné à la survie. Alors $(x(t))_{t \geq 0}$ et $(Z_n)_{n \geq 0}$ ont le même comportement asymptotique :

$$\lim_{n \rightarrow \infty} \frac{Z_n(k)}{Z_n(1) + \cdots + Z_n(N)} = \lim_{t \rightarrow \infty} x_k(t) = u_k, \quad 1 \leq k \leq N.$$

7.5 La quasi-espèce pour le processus de Galton–Watson

Les résultats de la section précédente s'appliquent au processus de Galton–Watson $(X_n)_{n \geq 0}$ ainsi qu'au processus d'occupation $(O_n)_{n \geq 0}$, sous l'hypothèse que la matrice moyenne associée soit primitive. En particulier, les résultats de la section précédente s'appliquent à la fonction de fitness du paysage à un pic $f_{sp} : \{0, \dots, \ell\} \rightarrow \mathbb{R}^+$, donnée par :

$$f_{sp}(k) = \begin{cases} \sigma & \text{si } k = 0, \\ 1 & \text{si } k \geq 1. \end{cases}$$

Le paramètre $\sigma > 1$ représente l'avantage adaptatif de la master sequence. Soit W_{sp} la matrice moyenne du processus $(O_n)_{n \geq 0}$ pour le paysage à un pic, soient

λ_{sp} et u_{sp} la valeur propre et le vecteur propre associés à W_{sp} par le théorème de Perron–Frobenius. Rappelons que la valeur propre λ_{sp} représente la fitness moyenne d’une population distribuée selon le vecteur unitaire u_{sp} :

$$\lambda_{sp} = \sum_{0 \leq k \leq \ell} u_{sp}(k) f_{sp}(k) = \sigma u_{sp}(0) + 1 - u_{sp}(0).$$

Notamment $\lambda_{sp} > 1$. D’après cette remarque et la proposition 4.4.4 de la partie I, nous avons le résultat suivant :

Théorème 7.5.1. *Le processus d’occupation $(O_n)_{n \geq 0}$ a une probabilité positive de survie. Conditionné à la survie :*

- si $\sigma e^{-\alpha} \leq 1$, alors

$$\forall k \geq 0 \quad \lim_{\ell \rightarrow \infty, q \rightarrow 0} \lim_{\ell q \rightarrow \alpha} \lim_{n \rightarrow \infty} \frac{O_n(k)}{O_n(0) + \dots + O_n(\ell)} = 0.$$

- si $\sigma e^{-\alpha} > 1$, alors

$$\forall k \geq 0 \quad \lim_{\ell \rightarrow \infty, q \rightarrow 0} \lim_{\ell q \rightarrow \alpha} \lim_{n \rightarrow \infty} \frac{O_n(k)}{O_n(0) + \dots + O_n(\ell)} = (\sigma e^{-\alpha} - 1) \frac{\alpha^k}{k!} \sum_{i \geq 1} \frac{i^k}{\sigma^i}.$$

Nous voyons apparaître, dans le terme de droite de cette formule, la concentration de la k -ième classe de Hamming dans la distribution de la quasi-espèce $\mathcal{Q}(\sigma, \alpha)$ (cf. la section 4.2).

Corollaire 7.5.2. *Le processus original $(X_n)_{n \geq 0}$ a une probabilité positive de survie. Conditionné à la survie :*

- si $\sigma e^{-\alpha} \leq 1$, alors

$$\forall k \geq 0 \quad \lim_{\ell \rightarrow \infty, q \rightarrow 0} \lim_{\ell q \rightarrow \alpha} \lim_{n \rightarrow \infty} \frac{1}{|X_n|_1} \sum_{u \in \mathcal{C}_k} X_n(u) = 0.$$

- si $\sigma e^{-\alpha} > 1$, alors

$$\forall k \geq 0 \quad \lim_{\ell \rightarrow \infty, q \rightarrow 0} \lim_{\ell q \rightarrow \alpha} \lim_{n \rightarrow \infty} \frac{1}{|X_n|_1} \sum_{u \in \mathcal{C}_k} X_n(u) = (\sigma e^{-\alpha} - 1) \frac{\alpha^k}{k!} \sum_{i \geq 1} \frac{i^k}{\sigma^i}.$$

Pour des fonctions de fitness $f_H : \mathbb{N} \rightarrow \mathbb{R}^+$ différentes de la fonction du paysage à un pic, outre l’hypothèse \mathcal{H} nous supposons que :

- La matrice moyenne W du processus $(Z_n)_{n \geq 0}$ associée à la fonction f_H est primitive.
- La valeur propre λ associée à W par le théorème de Perron–Frobenius est strictement supérieure à 1.
- La fonction de fitness $f_H(n)$ converge quand n tend vers l’infini ; nous notons par L la limite.

Remarquons que, d’après la discussion sur le *scaling* dans la section précédente, si $\lambda \leq 1$ nous pouvons faire le changement $f_H \rightarrow cf_H$ dans la fitness, avec $c > 1/\lambda$. Sous ces hypothèses, d’après la proposition 5.6.2 de la partie I, nous avons le résultat suivant :

Théorème 7.5.3. *Le processus d’occupation $(O_n)_{n \geq 0}$ a une probabilité positive de survie. Conditionné à la survie :*

- si $f_H(0)e^{-\alpha} \leq L$, alors

$$\forall k \geq 0 \quad \lim_{\ell \rightarrow \infty, q \rightarrow 0} \lim_{\ell q \rightarrow \alpha} \lim_{n \rightarrow \infty} \frac{O_n(k)}{O_n(0) + \dots + O_n(\ell)} = 0.$$

- si $f_H(0)e^{-\alpha} > L$, alors

$$\forall k \geq 0 \quad \lim_{\ell \rightarrow \infty, q \rightarrow 0} \lim_{\ell q \rightarrow \alpha} \lim_{n \rightarrow \infty} \frac{O_n(k)}{O_n(0) + \dots + O_n(\ell)} = x_k^*,$$

où x_k^* est la concentration de la k -ième classe de Hamming dans la distribution de la quasi-espèce généralisée $\mathcal{QG}(f_H, \alpha)$ (cf. la définition 5.4.2).

Corollaire 7.5.4. *Le processus original $(X_n)_{n \geq 0}$ a une probabilité positive de survie. Conditionné à la survie :*

- si $f_H(0)e^{-\alpha} \leq L$, alors

$$\forall k \geq 0 \quad \lim_{\ell \rightarrow \infty, q \rightarrow 0} \lim_{\ell q \rightarrow \alpha} \lim_{n \rightarrow \infty} \frac{1}{|X_n|_1} \sum_{u \in \mathcal{C}_k} X_n(u) = 0.$$

- si $f_H(0)e^{-\alpha} > L$, alors

$$\forall k \geq 0 \quad \lim_{\ell \rightarrow \infty, q \rightarrow 0} \lim_{\ell q \rightarrow \alpha} \lim_{n \rightarrow \infty} \frac{1}{|X_n|_1} \sum_{u \in \mathcal{C}_k} X_n(u) = x_k^*.$$

7.6 Conclusion

Nous avons considéré un processus de Galton–Watson à plusieurs types dans un cadre analogue au cadre du modèle d’Eigen, et nous avons réussi à retrouver le phénomène de seuil d’erreur et les différentes distributions de quasi–espèce, comme dans la partie I. Même si les résultats obtenus marquent un premier pas pour justifier l’utilisation des concepts de seuil d’erreur et de quasi–espèce pour des populations finies, les résultats sont peu surprenants et en partie insatisfaisants.

Peu surprenants car l’objet central dans l’étude du processus de Galton–Watson est la matrice $W = fM$, avec sa valeur propre maximale et vecteur propre associé. Ces objets jouaient déjà un rôle centrale dans le cas du modèle d’Eigen. Insatisfaisants, car nous ne pouvons pas mesurer l’impact de la taille de la population avec ce modèle. En effet, la croissance exponentielle de la population fait que la taille de la population dépasse le nombre fini de génotypes après quelques générations, ce qui était justement ce que nous voulions éviter (cf. le chapitre 6).

Pour surmonter cette difficulté, nous considérons ensuite les modèle de Wright–Fisher et le modèle de Moran, qui, à la différence du processus de Galton–Watson, sont des modèles avec une taille de population constante. Cela nous permettra de mieux comprendre le rôle joué par la taille de la population.

Troisième partie

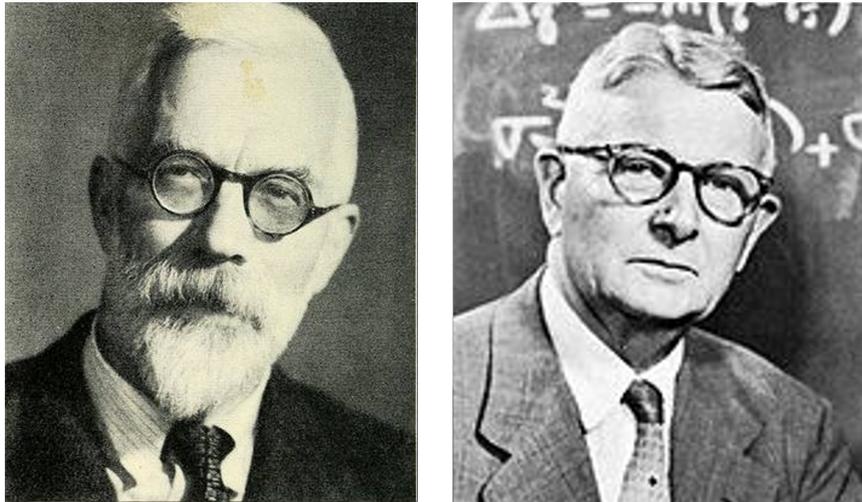
Le modèle de Wright–Fisher

Chapitre 8

Introduction

Imaginons maintenant une population finie d'individus haploïdes, et intéressons nous à un gène concret de son génome, qui peut présenter deux allèles a ou A . Le modèle proposé indépendamment par Ronald Fisher [36] et Sewall Wright [86] étudie la variation de la proportion des allèles quand la population évolue. Dans le modèle de Wright–Fisher, la taille de la population, m , est constante au cours du temps. Pour passer d'une génération à la suivante, nous tirons m individus au hasard, avec remise, de la population courante, que nous appelons les *enfants*, et qui forment la nouvelle génération. Il s'agit donc d'une évolution où les différentes générations ne se superposent pas. Pour introduire la sélection dans le modèle, le tirage au hasard est fait en fonction de la fitness de chaque individu, et pour introduire les mutations, chaque enfant mute selon le noyau de mutation voulu. Le modèle de Wright–Fisher s'étend naturellement à une population avec un nombre fini quelconque de types.

Suivant les lignes proposées par Cerf dans [11], nous allons considérer le modèle de Wright–Fisher dans le même cadre que le modèle d'Eigen : un nombre fini de types avec une fonction de fitness et un noyau de mutation. Le modèle de Wright–Fisher décrit ci-dessus donne naissance à une chaîne de Markov. Sous des hypothèses raisonnables sur le paysage de fitness et le noyau de mutation, la chaîne de Markov de Wright–Fisher est ergodique, et nous pouvons donc nous intéresser à la distribution des types quand le nombre de générations tend vers l'infini (l'analogie des solutions stationnaires du système d'équations différentielles d'Eigen). L'étude du processus de Wright–Fisher est nettement plus difficile que celle du processus de Galton–Watson, car la population constante



Ronald Fisher et Sewall Wright

introduit un couplage entre les différents individus. Nous allons donc nous concentrer sur le cas le plus simple : celui du paysage à un pic.

Pour retrouver un phénomène de seuil d'erreur et une distribution de quasi-espèce, il ne suffit pas cette fois-ci de considérer le régime asymptotique du modèle d'Eigen et du modèle de Galton–Watson (cf. la section 3.4). En effet, la présence d'un nouveau paramètre, la taille de la population, nous oblige à concilier le régime asymptotique du modèle d'Eigen et la limite des grandes populations ; la façon la plus naturelle de procéder est de considérer que la longueur du génome et la taille de la population sont du même ordre. Ainsi, Cerf obtient une courbe critique séparant un régime de désordre et un régime de quasi-espèce [11]. Il obtient aussi la concentration de la master sequence dans le régime de quasi-espèce, qui coïncide avec la concentration de la master sequence pour le modèle d'Eigen associé. L'objectif de cette partie est de montrer qu'il en est de même pour les concentrations des autres classes de Hamming.

Nous pouvons distinguer deux comportements radicalement différents dans le processus de Wright–Fisher : quand la master sequence est présente dans la population et quand elle ne l'est pas. En effet, si la master sequence n'est pas présente dans la population, la population est neutre, et le processus de Wright–Fisher dédie son temps à explorer l'espace d'états de manière aléatoire. Le temps moyen pour retrouver la master sequence, le *temps de découverte*, est alors de l'ordre de la taille de l'espace d'états. Inversement, si la master sequence est présente dans la population, et si sa fitness est assez grande, elle

met très longtemps à disparaître de la population, le temps moyen pour cette disparition est appelé le *temps de persistance*. La comparaison entre le temps de découverte et le temps de persistance donne lieu à la courbe critique mentionnée. Pour retrouver les concentrations des différentes classes dans la distribution de la quasi-espèce, il suffit donc de se concentrer sur la région sur-critique, et sur la dynamique du processus quand la master sequence est présente dans la population.

Dans le premier chapitre, nous présentons le processus de Wright–Fisher, nous énonçons le résultat principal de cette partie, et nous donnons quelques résultats de [11] utiles pour la suite. Une des principales difficultés est que la taille de l'espace d'états croît avec la longueur des genotypes (il y a de plus en plus de classes de Hamming). Pour surmonter cette difficulté, nous bornons ensuite la chaîne de Markov de Wright–Fisher par d'autres chaînes de Markov plus simples, dont l'espace d'états ne croît pas, même si la longueur du génome grandit. Le chapitre final porte sur l'étude de ces chaînes de Markov simplifiées. Dans le régime asymptotique que nous allons considérer, les trajectoires de ces chaînes de Markov se concentrent sur des trajectoires déterministes, données par un système dynamique. Le système dynamique en question est une version discrète du système d'équations d'Eigen, et les points fixes du système dynamique coïncident avec les solutions stationnaires du système d'Eigen. Nous avons donc deux points fixes, 0 et ρ^* , où ρ^* représente la distribution de la quasi-espèce :

$$\forall k \geq 0 \quad \rho_k^* = (\sigma e^{-a} - 1) \frac{a^k}{k!} \sum_{i \geq 1} \frac{i^k}{\sigma^i}.$$

De plus, le point fixe ρ^* attire toutes les trajectoires avec une condition initiale ayant la coordonnée 0 strictement positive. L'enjeu est de montrer que la mesure invariante des chaînes de Markov simplifiées se concentre autour de ρ^* . Le problème est que la mesure invariante est une fonction de toute la trajectoire du processus ; les chaînes étant ergodiques, elles sont ramenées près du point fixe 0 une infinité de fois et elles pourraient y rester longtemps. Nous nous inspirons de la théorie de Freidlin et Wentzell pour des perturbations aléatoires de systèmes dynamiques [38] pour contrôler les excursions des chaînes de Markov en dehors d'un voisinage de ρ^* , et nous montrons ainsi que la longueur et la fréquence de ces excursions deviennent asymptotiquement négligeables.

Chapitre 9

Modèle et résultats

Dans ce chapitre, nous introduisons le modèle de Wright–Fisher qui sera l’objet de notre étude. Ensuite nous énonçons le résultat principal. Pour finir nous rappelons quelques outils de [11] dont nous aurons besoin au cours du texte.

9.1 Le modèle

Puisque notre objectif est de retrouver un phénomène de seuil d’erreur et une distribution de quasi–espèce pour le modèle de Wright–Fisher, nous reprenons le cadre introduit dans la section 3.1 pour le modèle d’Eigen. Nous considérons donc des individus dont les génotypes sont des chaînes de longueur $\ell \geq 1$ sur un alphabet fini \mathcal{A} de cardinal $\kappa \geq 2$. Nous nous donnons également une fonction de fitness $f : \mathcal{A}^\ell \rightarrow \mathbb{R}^+$, et nous supposons que les mutations arrivent indépendamment au hasard sur chaque site de la chaîne avec probabilité $q \in]0, 1[$. La matrice de mutation $(M(u, v), u, v \in \mathcal{A}^\ell)$ est donc donnée par :

$$M(u, v) = \left(\frac{q}{\kappa - 1} \right)^{d_H(u, v)} (1 - q)^{\ell - d_H(u, v)}, \quad u, v \in \mathcal{A}^\ell.$$

Comme précédemment, $d_H(u, v)$ représente la distance de Hamming entre u et v , i.e., le nombre de digits où les deux chaînes diffèrent. La taille de la population est finie, nous la notons par $m \geq 1$. La taille de la population reste constante au cours de l’évolution. Une population x est donc un élément de l’espace $(\mathcal{A}^\ell)^m$.

Nous construisons une fonction de sélection $F : \mathcal{A} \times (\mathcal{A}^\ell)^m \rightarrow [0, 1]$ à partir de f comme suit :

$$\forall u \in \mathcal{A}^\ell \quad \forall x \in (\mathcal{A}^\ell)^m \quad F(u, x) = \frac{f(u)}{f(x(1)) + \cdots + f(x(m))} \sum_{1 \leq i \leq m} 1_{x(i)=u}$$

Dans ce modèle, le mécanisme de transition d'une génération à la suivante se décompose en deux pas. D'abord, nous échantillonnons avec remise m chromosomes de la population actuelle. L'échantillonnage est fait à l'aide de la fonction de sélection F . Puis, chaque chromosome échantillonné mute selon la loi donnée par la matrice de mutation. Le modèle de Wright–Fisher est une chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ d'espace d'états $(\mathcal{A}^\ell)^m$, et de matrice de transition

$$\forall n \in \mathbb{N} \quad \forall x, y \in (\mathcal{A}^\ell)^m$$

$$P(X_{n+1} = y \mid X_n = x) = \prod_{1 \leq i \leq m} \left(\sum_{u \in \mathcal{A}^\ell} F(u, x) M(u, y(i)) \right).$$

Ergodicité. Si la fonction de fitness $f : \mathcal{A}^\ell \rightarrow \mathbb{R}^+$ est strictement positive, le processus que nous venons de décrire est ergodique. Nous voulons retrouver une distribution de quasi-espèce pour le modèle de Wright–Fisher. Comme nous avons bien observé pour le modèle d'Eigen dans la partie I que la distribution de quasi-espèce apparaît dans un état d'équilibre, nous allons supposer que toutes les fitness sont strictement positives.

Hypothèse. La fonction de fitness $f : \mathcal{A}^\ell \rightarrow \mathbb{R}^+$ est strictement positive.

Sous cette hypothèse, le processus de Wright–Fisher est ergodique.

Régime asymptotique. Le but des sections qui suivent est d'étudier le modèle de Wright–Fisher dans le régime asymptotique

$$\begin{aligned} \ell &\rightarrow +\infty, & m &\rightarrow +\infty, & q &\rightarrow 0, \\ \ell q &\rightarrow a \in]0, +\infty[, & \frac{m}{\ell} &\rightarrow \alpha \in [0, +\infty[, \end{aligned}$$

afin de montrer l'existence d'un phénomène de seuil d'erreur et d'une distribution de quasi-espèce.

9.2 Résultat principal

Nous ne considérons que le paysage à un pic. Soit $\sigma > 1$. Il existe une séquence, notée w^* et appelée master sequence, telle que la fonction de fitness est donnée par

$$\forall u \in \mathcal{A}^\ell \quad f(u) = \begin{cases} 1 & \text{si } u \neq w^*, \\ \sigma & \text{si } u = w^*. \end{cases}$$

Soit x une population dans $(\mathcal{A}^\ell)^m$. Fixons $k \geq 0$ un entier. Nous nous intéressons au nombre $N^k(x)$ de chromosomes dans la population x à distance k de la master sequence :

$$N^k(x) = \text{card} \{ i : 1 \leq i \leq m, d_H(x(i), w^*) = k \}.$$

Notons par $I(p, t)$ la fonction de taux gouvernant le principe de grandes déviations de la loi binomiale de paramètre $p \in [0, 1]$, i.e.,

$$\forall t \in [0, 1] \quad I(p, t) = t \ln \frac{t}{p} + (1-t) \ln \frac{1-t}{1-p}.$$

Nous définissons, pour $a \in]0, +\infty[$,

$$\rho^*(a) = \begin{cases} \frac{\sigma e^{-a} - 1}{\sigma - 1} & \text{si } \sigma e^{-a} > 1, \\ 0 & \text{si } \sigma e^{-a} \leq 1, \end{cases}$$

$$\psi(a) = \inf_{l \in \mathbb{N}} \inf \left\{ \sum_{k=0}^{l-1} I\left(\frac{\sigma \rho_k}{(\sigma - 1)\rho_k + 1}, \gamma_k\right) + \gamma_l I\left(e^{-a}, \frac{\rho_{k+1}}{\gamma_k}\right) : \right. \\ \left. \rho(0) = \rho^*(a), \rho_l = 0, \rho_k, \gamma_k \in [0, 1] \text{ pour } 0 \leq k \leq l \right\}.$$

Théorème 9.2.1. *Nous avons la dichotomie suivante :*

- Si $\alpha \psi(a) < \ln \kappa$, alors

$$\forall k \geq 0 \quad \lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \lim_{t \rightarrow \infty} E\left(\frac{1}{m} N^k(X_t)\right) = 0.$$

- Si $\alpha\psi(a) > \ln \kappa$, alors

$$\forall k \geq 0 \quad \lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \lim_{t \rightarrow \infty} \mathbb{E} \left(\frac{1}{m} N^k(X_t) \right) = \rho_k^* = (\sigma e^{-a} - 1) \frac{a^k}{k!} \sum_{i \geq 1} \frac{i^k}{\sigma^i}.$$

De plus, dans les deux cas,

$$\forall k \geq 0 \quad \lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \lim_{t \rightarrow \infty} \text{Var} \left(\frac{1}{m} N^k(X_t) \right) = 0.$$

Donnons d'abord quelques idées de preuve pour le résultat principal. Le processus de Wright–Fisher $(X_n)_{n \geq 0}$ n'est pas très maniable, car son espace d'états est énorme et n'est muni d'aucun ordre naturel. Dans le même esprit que pour le processus de Galton–Watson 7.3, le premier pas sera donc de réduire l'espace d'états du processus et d'en dériver un processus plus simple, le processus d'occupation $(O_n)_{n \geq 0}$. Le processus d'occupation ne garde que l'information pertinente, i.e., le nombre d'individus dans chaque classe de Hamming. Le processus d'occupation est encore une chaîne de Markov, qui a la vertu d'être plus simple que le processus original, et dont l'espace d'états est muni d'un ordre naturel. C'est l'existence de cet ordre qui nous permettra de comparer le processus d'occupation avec des chaînes de Markov encore plus simples.

Nous fixons ensuite $K \geq 0$ et nous étudions la concentration des individus dans la K -ième classe de Hamming. Pour cela, nous démarrons le processus d'une distribution d'occupation n'ayant aucun individu dans les classes $0, \dots, K$, et nous estimons le temps moyen que le processus met à trouver un individu dans les classes $0, \dots, K$, c'est ce que nous appelons le temps de découverte. Également, nous démarrons le processus d'une distribution ayant des individus dans les classes $0, \dots, K$, et nous estimons le temps moyen qu'il faut pour que tous les individus dans les classes $0, \dots, K$ disparaissent, c'est ce que nous appelons le temps de persistance. Asymptotiquement, quand $\alpha\psi(a) < \ln \kappa$, le temps de persistance devient négligeable par rapport au temps de découverte, tandis que quand $\alpha\psi(a) > \ln \kappa$, c'est le temps de découverte qui devient négligeable par rapport au temps de persistance. Ce fait suffit déjà à démontrer la première assertion du théorème 9.2.1. Pour le cas $K = 0$, ceci a été démontré dans [11]. Pour le cas général $K \geq 0$, le traitement est le même que pour le cas $K = 0$, et la démonstration n'apporte aucune contribution à la compréhension du modèle. Nous allons donc admettre ce fait et nous allons nous concentrer sur le cas plus intéressant $\alpha\psi(a) > \ln \kappa$.

Pour traiter le cas $\alpha\psi(a) > \ln \kappa$, nous construisons un couplage, afin de comparer le processus d'occupation avec des processus plus simples. Les processus simplifiés seront des chaînes de Markov qui répliqueront la dynamique des classes de Hamming $0, \dots, K$ seulement. Ces chaînes de Markov peuvent être vues comme des perturbations aléatoires d'un même système dynamique avec deux points fixes : un point fixe instable en 0 et un point fixe stable en ρ^* . Nous utiliserons la théorie développée par Freidlin et Wentzell [38], ainsi que quelques estimations utiles de [11], pour montrer que les processus perturbés passent la plupart de leur temps près du point fixe stable ρ^* , ce qui permettra de montrer que les mesures de probabilité invariantes des processus perturbés se concentrent autour de ρ^* .

9.3 Lumping

L'espace d'états du processus de Wright–Fisher $(\mathcal{A}^\ell)^m$ n'est pas toujours très maniable. Il y a deux façons naturelles de le réduire. L'idée, dans les deux cas, est de ne plus regarder chaque génotype mais seulement les classes de Hamming. Nous construisons ainsi deux processus. Le premier processus remplace le génotype de chaque individu par la classe de Hamming auquel il appartient, nous l'appelons *processus de distance*. Le deuxième processus compte le nombre d'individus dans chaque classe de Hamming, nous l'appelons *processus d'occupation*. L'outil clé pour cette construction est le lemme de *lumping* pour des chaînes de Markov (cf. l'annexe B). Nous formalisons ces idées dans les lignes qui suivent. Soit $H : \mathcal{A}^\ell \rightarrow \{0, \dots, \ell\}$ l'application qui à chaque génotype u associe sa classe de Hamming vis-à-vis de la master sequence :

$$\forall u \in \mathcal{A}^\ell \quad H(u) = d_H(u, w^*).$$

La fonction de fitness $f : \mathcal{A}^\ell \rightarrow \mathbb{R}^+$ du paysage à un pic respecte les classes de Hamming. Nous pouvons donc définir une nouvelle fonction de fitness $f_H : \{0, \dots, \ell\} \rightarrow \mathbb{R}^+$ en posant $f_H(k)$ égale à la valeur commune de $f(u)$ pour u dans la classe de Hamming k :

$$\forall k \in \{0, \dots, \ell\} \quad f_H(k) = \begin{cases} \sigma & \text{si } k = 0, \\ 1 & \text{si } k \geq 1. \end{cases}$$

D'après le lemme 3.2.1, le noyau de mutation respecte aussi les classes de Hamming. Nous en dérivons un noyau de mutation $(M_H(b, c), 0 \leq b, c \leq \ell)$, donné

par :

$$M_H(b, c) = \sum_{\substack{0 \leq k \leq \ell - b \\ 0 \leq l \leq b \\ b - l + k = c}} \binom{\ell - b}{k} \binom{b}{l} q^k (1 - q)^{\ell - b - k} \left(\frac{q}{\kappa - 1} \right)^l \left(1 - \frac{q}{\kappa - 1} \right)^{b - l}.$$

Processus de distance. Nous définissons ensuite une fonction vectorielle $\mathbb{H} : (\mathcal{A}^\ell)^m \rightarrow \{0, \dots, \ell\}^m$ par :

$$\forall x = \begin{pmatrix} x(1) \\ \vdots \\ x(m) \end{pmatrix} \in (\mathcal{A}^\ell)^m \quad \mathbb{H}(x) = \begin{pmatrix} H(x(1)) \\ \vdots \\ H(x(m)) \end{pmatrix}.$$

Nous définissons le processus de distance $(D_n)_{n \geq 0}$ par

$$\forall n \geq 0 \quad D_n = \mathbb{H}(X_n).$$

Le processus de distance est un processus de Markov, ce qui se voit d'après le théorème de *lumping* B.1 et la proposition suivante :

Proposition 9.3.1. *La matrice de transition p du modèle de Wright–Fisher $(X_n)_{n \geq 0}$ vérifie :*

$$\forall e \in \{0, \dots, \ell\}^m \quad \forall x, y \in (\mathcal{A}^\ell)^m \\ \mathbb{H}(x) = \mathbb{H}(y) \implies \sum_{\substack{z \in (\mathcal{A}^\ell)^m \\ \mathbb{H}(z) = e}} p(x, z) = \sum_{\substack{z \in (\mathcal{A}^\ell)^m \\ \mathbb{H}(z) = e}} p(y, z).$$

La démonstration est similaire à celle de la section 7.3. Nous concluons grâce au théorème de *lumping* B.1 que le processus de distance $(D_n)_{n \geq 0}$ est un processus de Markov. L'espace d'états du processus de distance est $\{0, \dots, \ell\}^m$ et la matrice de transition p_H est donnée par :

$$\forall d, e \in \{0, \dots, \ell\}^m \\ p_H(d, e) = \prod_{1 \leq i \leq m} \left(\sum_{1 \leq j \leq m} \frac{f_H(d(j)) M_H(d(j), e(i))}{f_H(d(1)) + \dots + f_H(d(m))} \right).$$

Processus d'occupation.

Notons $\mathcal{P}_{\ell+1}^m$ l'ensemble des partitions ordonnées de l'entier m en au plus $\ell + 1$ parties :

$$\mathcal{P}_{\ell+1}^m = \{ (o(0), \dots, o(\ell)) \in \mathbb{N}^{\ell+1} : o(0) + \dots + o(\ell) = m \}.$$

Nous interprétons ces partitions comme des distributions d'occupation. La partition $(o(0), \dots, o(\ell))$ correspond à une population qui a exactement $o(l)$ chromosomes à distance l de la master sequence, pour $l \in \{0, \dots, \ell\}$. Soit \mathcal{O} l'application qui à chaque population $x \in (\mathcal{A}^\ell)^m$ associe la distribution d'occupation correspondante :

$$\forall k \in \{0, \dots, \ell\} \quad \mathcal{O}(x)(k) = \sum_{1 \leq i \leq m} 1_{H(x(i))=k}.$$

Nous définissons le processus d'occupation $(O_n)_{n \geq 0}$ par :

$$\forall n \geq 0 \quad O_n = \mathcal{O}(X_n).$$

Le processus d'occupation peut également être obtenu à partir du processus distance. En effet, nous définissons l'application $\mathcal{O}_H : (\mathcal{A}^\ell)^m \rightarrow \mathcal{P}_{\ell+1}^m$ par :

$$\forall d \in (\mathcal{A}^\ell)^m \quad \forall k \in \{0, \dots, \ell\} \quad \mathcal{O}_H(d)(k) = \sum_{1 \leq i \leq m} 1_{d(i)=k}.$$

Nous avons donc,

$$\forall x \in (\mathcal{A}^\ell)^m \quad \mathcal{O}(x) = \mathcal{O}_H(\mathbb{H}(x)).$$

La démonstration de ce résultat est similaire à celle de la section 7.3. Nous pouvons donc appliquer le théorème de *lumping* B.1 pour conclure que le processus $(O_n)_{n \geq 0}$ est bien un processus de Markov. L'espace d'états du processus de Markov $(O_n)_{n \geq 0}$ est $\mathcal{P}_{\ell+1}^m$, et la matrice de transition p_O est donnée par :

$$\forall o, o' \in \mathcal{P}_{\ell+1}^m \quad p_O(o, o') = \prod_{0 \leq h \leq \ell} \left(\frac{\sum_{k \in \{0, \dots, \ell\}} o(k) f_H(k) M_H(k, h)}{\sum_{0 \leq h \leq \ell} o(h) f_H(h)} \right)^{o'(h)}.$$

9.4 Couplage et monotonie

Cette section est un résumé de la section 5 de [11]. Nous reproduisons sans démonstration les points les plus importants pour la suite du texte.

Nous construisons tous nos processus dans un grand espace de probabilité (Ω, \mathcal{F}, P) ayant les variables aléatoires indépendantes suivantes, toutes avec loi uniforme sur l'intervalle $[0, 1]$:

$$U_n^{i,j}, \quad n \geq 1, \quad 1 \leq i \leq m, \quad 1 \leq j \leq \ell,$$

$$S_n^i, \quad n \geq 1, \quad 1 \leq i \leq m.$$

Nous posons

$$\forall n \geq 1 \quad R_n = \begin{pmatrix} S_n^1 & U_n^{1,1} & \dots & U_n^{1,\ell} \\ \vdots & \vdots & \ddots & \vdots \\ S_n^m & U_n^{m,1} & \dots & U_n^{m,\ell} \end{pmatrix}.$$

La matrice aléatoire R_n sert à accomplir le n -ième pas de la chaîne de Markov $(X_n)_{n \geq 0}$. Par construction, la suite $(R_n)_{n \geq 1}$ est une suite de matrices aléatoires indépendantes et identiquement distribuées à valeurs dans l'espace \mathcal{R} des matrices de taille $m \times (\ell + 1)$ avec coefficients dans $[0, 1]$.

Mutation. Pour coupler le mécanisme de mutation, nous définissons une application

$$\mathcal{M}_H : \{0, \dots, \ell\} \times [0, 1]^\ell \rightarrow \{0, \dots, \ell\}.$$

Nous posons

$$\forall b \in \{0, \dots, \ell\} \quad \forall u_1, \dots, u_\ell \in [0, 1]^\ell$$

$$\mathcal{M}_H(b, u_1, \dots, u_\ell) = b - \sum_{k=1}^b 1_{u_k < q/(\kappa-1)} + \sum_{k=b+1}^{\ell} 1_{u_k > 1-q}.$$

L'application \mathcal{M}_H est définie de sorte que si U_1, \dots, U_ℓ sont des variables aléatoires indépendantes et uniformes sur l'intervalle $[0, 1]$, alors pour tout $b \in \{0, \dots, \ell\}$, la loi de $\mathcal{M}_H(b, U_1, \dots, U_\ell)$ est donnée par la ligne de la matrice M_H associée à b , i.e.,

$$\forall c \in \{0, \dots, \ell\} \quad P(\mathcal{M}_H(b, U_1, \dots, U_\ell) = c) = M_H(b, c).$$

Sélection pour le processus distance. Nous définissons une application

$$\mathcal{S}_H : \{0, \dots, \ell\}^m \times [0, 1] \rightarrow \{1, \dots, m\}.$$

Pour $d \in \{0, \dots, \ell\}^m$ et pour $s \in [0, 1[$, nous définissons $\mathcal{S}_H(d, s) = i$ où i est le seul indice dans $\{1, \dots, m\}$ tel que

$$\frac{f_H(d(1)) + \dots + f_H(d(i-1))}{f_H(d(1)) + \dots + f_H(d(m))} \leq s < \frac{f_H(d(1)) + \dots + f_H(d(i))}{f_H(d(1)) + \dots + f_H(d(m))}.$$

L'application \mathcal{S}_H est définie de sorte que si S est une variable aléatoire de loi uniforme sur l'intervalle $[0, 1]$, alors pour tout $d \in \{0, \dots, \ell\}^m$, la loi de $\mathcal{S}_H(d, S)$ est donnée par

$$\forall i \in \{1, \dots, m\} \quad P(\mathcal{S}_H(d, S) = i) = \frac{f_H(d(i))}{f_H(d(1)) + \dots + f_H(d(m))}.$$

Couplage pour le processus distance. Nous définissons une application

$$\Psi_H : \{0, \dots, \ell\}^m \times \mathcal{R} \rightarrow \{0, \dots, \ell\}^m.$$

pour coupler le processus distance avec différentes conditions initiales. L'application Ψ_H est définie par

$$\forall r \in \mathcal{R} \quad \forall d \in \{0, \dots, \ell\}^m$$

$$\Psi_H(d, r) = \begin{pmatrix} \mathcal{M}_H(d(\mathcal{S}_H(d, r(1, 1))), r(1, 2), \dots, r(1, \ell + 1)) \\ \vdots \\ \mathcal{M}_H(d(\mathcal{S}_H(d, r(m, 1))), r(m, 2), \dots, r(m, \ell + 1)) \end{pmatrix}.$$

Soit $d \in \{0, \dots, \ell\}^m$ le point de départ du processus. Nous construisons le processus distance $(D_n)_{n \geq 0}$ en prenant $D_0 = d$ et

$$\forall n \geq 1 \quad D_n = \Psi_H(D_{n-1}, R_n).$$

Le processus $(D_n)_{n \geq 0}$ est une chaîne de Markov issue de d et de matrice de transition p_H .

Sélection pour le processus d'occupation. Nous définissons une application $\mathcal{S}_O : \mathcal{P}_{\ell+1}^m \times [0, 1] \rightarrow \{0, \dots, \ell\}$. Pour $o \in \mathcal{P}_{\ell+1}^m$ et $s \in [0, 1[$, nous définissons $\mathcal{S}_O(o, s) = l$ où l est le seul indice dans $\{0, \dots, \ell\}$ tel que

$$\frac{o(0)f_H(0) + \dots + o(l-1)f_H(l-1)}{o(0)f_H(0) + \dots + o(\ell)f_H(\ell)} \leq s < \frac{o(0)f_H(0) + \dots + o(l)f_H(l)}{o(0)f_H(0) + \dots + o(\ell)f_H(\ell)}.$$

L'application \mathcal{S}_O est définie de sorte que si S est une variable aléatoire uniforme sur l'intervalle $[0, 1]$, alors pour tout $o \in \mathcal{P}_{\ell+1}^m$, la loi de $\mathcal{S}_O(o, S)$ est donnée par

$$\forall l \in \{0, \dots, \ell\} \quad P(\mathcal{S}_O(o, S) = l) = \frac{o(l)f_H(l)}{o(0)f_H(0) + \dots + o(\ell)f_H(\ell)}.$$

Couplage pour le processus d'occupation. Nous définissons une application $\Psi_O : \mathcal{P}_{\ell+1}^m \times \mathcal{R} \rightarrow \mathcal{P}_{\ell+1}^m$ pour coupler le processus d'occupation avec différentes conditions initiales. Nous définissons l'application Ψ_O comme suit :

$$\forall r \in \mathcal{R} \quad \forall o \in \mathcal{P}_{\ell+1}^m$$

$$\Psi_O(o, r) = \mathcal{O}_H \left(\begin{array}{c} \mathcal{M}_H(d(\mathcal{S}_H(d, r(1, 1))), r(1, 2), \dots, r(1, \ell + 1)) \\ \vdots \\ \mathcal{M}_H(d(\mathcal{S}_H(d, r(m, 1))), r(m, 2), \dots, r(m, \ell + 1)) \end{array} \right).$$

Soit $o \in \mathcal{P}_{\ell+1}^m$ le point de départ du processus. Nous construisons le processus d'occupation $(O_n)_{n \geq 0}$ en prenant $O_0 = o$ et

$$\forall n \geq 1 \quad O_n = \Psi_O(O_{n-1}, R_n).$$

Le processus $(O_n)_{n \geq 0}$ est une chaîne de Markov issue de o et de matrice de transition p_O .

Monotonie. L'espace $\{0, \dots, \ell\}^m$ est naturellement muni d'un ordre partiel :

$$d \leq e \iff \forall i \in \{1, \dots, m\} \quad d(i) \leq e(i).$$

Par le lemme 4.4.1, l'application \mathcal{M}_H est croissante vis-à-vis de la classe de Hamming, i.e.,

$$\forall b, c \in \{0, \dots, \ell\} \quad \forall u_1, \dots, u_\ell \in [0, 1]$$

$$b \leq c \implies \mathcal{M}_H(b, u_1, \dots, u_\ell) \leq \mathcal{M}_H(c, u_1, \dots, u_\ell).$$

Lemme 9.4.1. Dans le cas neutre $\sigma = 1$, l'application \mathcal{S}_H est croissante vis-à-vis de la classe de Hamming, i.e.,

$$\forall d, e \in \{0, \dots, \ell\}^m \quad \forall s \in [0, 1] \quad d \leq e \implies d(\mathcal{S}_H(d, s)) \leq e(\mathcal{S}_H(e, s)).$$

Lemme 9.4.2. Dans le cas neutre $\sigma = 1$, l'application Ψ_H est croissante vis-à-vis des distances, i.e.,

$$\forall d, e \in \{0, \dots, \ell\}^m \quad \forall r \in \mathcal{R}, \quad d \leq e \implies \Psi_H(d, r) \leq \Psi_H(e, r).$$

Corollaire 9.4.3. Dans le cas neutre $\sigma = 1$, le processus distance $(D_n)_{n \geq 0}$ est monotone.

Nous définissons un ordre \preceq dans l'espace $\mathcal{P}_{\ell+1}^m$ comme suit. Soient o, o' deux distributions d'occupation. Nous disons que o est inférieure ou égale à o' , et nous écrivons $o \preceq o'$, si

$$\forall l \leq \ell \quad o(0) + \dots + o(l) \leq o'(0) + \dots + o'(l).$$

L'application \mathcal{S}_O est décroissante vis-à-vis de la distribution d'occupation, i.e.,

$$\forall o, o' \in \mathcal{P}_{\ell+1}^m \quad \forall s \in [0, 1] \quad o \preceq o' \Rightarrow \mathcal{S}_O(o, s) \geq \mathcal{S}_O(o', s).$$

Lemme 9.4.4. *L'application Ψ_O est croissante vis-à-vis de la distribution d'occupation, i.e.,*

$$\forall o, o' \in \mathcal{P}_{\ell+1}^m \quad \forall r \in \mathcal{R} \quad o \preceq o' \Rightarrow \Psi_O(o, r) \preceq \Psi_O(o', r).$$

Corollaire 9.4.5. *Le processus d'occupation $(O_n)_{n \geq 0}$ est monotone.*

9.5 Convergence vers un modèle de Moran

Afin de donner un argument heuristique de notre résultat principal, nous faisons ici une petite parenthèse pour donner un résultat du type loi des grands nombres pour le processus de Wright–Fisher. Nous le faisons dans un cadre plus général que celui que nous avons introduit ci-dessus. Nous considérons l'ensemble de génotypes $\{1, \dots, N\}$, une fonction de fitness $f : \{1, \dots, N\} \rightarrow \mathbb{R}^+$, et un noyau de mutation $(M(i, j), 1 \leq i, j \leq N)$. Nous notons par m la taille de la population. À chaque population $x \in \{1, \dots, N\}^m$, nous pouvons associer une distribution d'occupation $z_x = (z_x(1), \dots, z_x(N))$ par :

$$\forall k \geq 0 \quad z_x(k) = \sum_{i=1}^m 1_{x(i)=k},$$

et une fitness moyenne $\phi(x)$ par :

$$\phi(x) = \sum_{k=1}^N f(k)z_x(k).$$

Le processus de Wright–Fisher est la chaîne de Markov $(X_n)_{n \geq 0}$ à valeurs dans $\{1, \dots, N\}^m$ et de matrice de transition p donnée par :

$$\forall x, y \in \{1, \dots, N\}^m \quad p(x, y) = \prod_{1 \leq i \leq m} \left(\sum_{k=1}^N \frac{f(k)z_x(k)}{\phi(x)} M(k, i) \right)^{y(i)}.$$

Soit $(Z_n)_{n \geq 0} = (z_{X_n}/m)_{n \geq 0}$ le processus d'occupation normalisé associé au processus de Wright–Fisher. Soit \mathcal{S}^N le simplexe à N coordonnées :

$$\mathcal{S}^N = \{ \mathbf{u} \in [0, 1]^N : u_0 + \dots + u_N = 1 \},$$

et définissons l'application $F : \mathcal{S}^N \rightarrow \mathcal{S}^N$ par :

$$\forall \mathbf{u} \in \mathcal{S}^N \quad \forall k \in \{1, \dots, N\} \quad F_k(\mathbf{u}) = \sum_{i=1}^N \frac{f(i)u_i}{\sum_{1 \leq h \leq N} f(h)u_h} M(i, k).$$

Nous avons le résultat suivant :

Théorème 9.5.1. *Soit $\delta > 0$. Pour tout $\mathbf{u} \in \mathcal{S}^N$:*

$$\lim_{m \rightarrow \infty} \mathbb{P} \left(|Z_1 - F(\mathbf{u})| > \delta \mid Z_0 = \frac{\lfloor \mathbf{u}m \rfloor}{m} \right) = 0.$$

Nous en déduisons immédiatement le résultat suivant.

Corollaire 9.5.2. *Soient $\delta, T > 0$. Pour tout $\mathbf{u} \in \mathcal{S}^N$:*

$$\lim_{m \rightarrow \infty} \mathbb{P} \left(\sup_{0 \leq n \leq T} |Z_n - z(n)| > \delta \mid Z_0 = \frac{\lfloor \mathbf{u}m \rfloor}{m} \right) = 0,$$

où $(z(n))_{n \geq 0}$ est la trajectoire donnée par le système dynamique $z(0) = \mathbf{u}$ et $z(n) = F(z(n-1))$, $n \geq 1$.

Le système dynamique apparaissant dans ce résultat est une version discrète du modèle d'Eigen, et avait déjà été proposé par Moran comme modèle d'évolution d'une population avec sélection et mutation [58, 59]. Nous procédons maintenant à la démonstration du théorème.

Démonstration. Soient $\delta > 0$ et $\mathbf{u} \in \mathcal{S}^N$. Nous avons :

$$\begin{aligned} \mathbb{P}(|Z_1 - F(\mathbf{u})| > \delta) &= \mathbb{P} \left(\sum_{k=1}^N |Z_1(k) - F_k(\mathbf{u})|^2 > \delta^2 \right) \\ &\leq \sum_{k=1}^N \mathbb{P}(|Z_1(k) - F_k(\mathbf{u})| > \delta/\sqrt{N}) \\ &\leq \mathbb{P}(|Z_1(k) - \mathbb{E}(Z_1(k))| > \delta/2\sqrt{N}) + \mathbb{P}(|\mathbb{E}(Z_1(k)) - F_k(\mathbf{u})| > \delta/2\sqrt{N}) \end{aligned}$$

Or, la variable aléatoire $Z_1(k)$ est donnée par :

$$Z_1(k) = \frac{1}{m} \sum_{i=1}^m 1_{X_1(i)=k}.$$

Les variables aléatoires $Y_i = 1_{X_1(i)=k}$, $1 \leq i \leq m$ sont des variables aléatoires indépendantes et identiquement distribuées, de loi de Bernoulli de paramètre $p_k = F(\lfloor um \rfloor / m)$. Nous avons donc $E(Z_1(K)) = p_k$ et $\text{Var}(Z_1(k)) = p_k(1 - p_k)/m$. Quand m tend vers l'infini, la première des probabilités de droite dans la formule ci-dessus tend vers 0 grâce à l'inégalité de Chebyshev. La deuxième tend aussi vers 0 car p_k converge vers $F_k(u)$ quand m tend vers l'infini. \square

Chapitre 10

Bornes stochastiques

Dans cette section, nous construisons un processus inférieur $(O_n^\ell)_{n \geq 0}$ et un processus supérieur $(O_n^{K+1})_{n \geq 0}$ pour borner stochastiquement le processus d'occupation $(O_n)_{n \geq 0}$, i.e., pour tout $o \in \mathcal{P}_{\ell+1}^m$, si $O_0^\ell = O_0 = O_0^{K+1} = o$, alors

$$\forall n \geq 0 \quad O_n^\ell \preceq O_n \preceq O_n^{K+1}.$$

Les processus $(O_n^\ell)_{n \geq 0}$ et $(O_n^{K+1})_{n \geq 0}$ auront la vertu d'être plus simples que le processus $(O_n)_{n \geq 0}$, tout en gardant l'information pertinente sur la dynamique des coordonnées $0, \dots, K$.

10.1 Processus inférieur et supérieur

Décrivons de manière informelle la dynamique du processus inférieur. Soit $K \leq \ell$. Lorsque la master sequence n'est pas présente dans la population, le processus $(O_n^\ell)_{n \geq 0}$ évolue exactement comme le processus originel $(O_n)_{n \geq 0}$. Dès qu'une master sequence apparaît, tous les chromosomes dans les classes $K+1, \dots, \ell$ sont envoyés dans la classe de Hamming ℓ . Tant que la master sequence est présente dans la population, tout chromosome qui mute vers les classes $K+1, \dots, \ell$ est directement envoyé à la classe ℓ . La dynamique du processus supérieur est similaire, sauf que nous remplaçons la classe de Hamming ℓ par la classe $K+1$. Nous formalisons cette construction dans le reste de la section. Nous définissons deux applications $\pi_\ell, \pi_{K+1} : \mathcal{P}_{\ell+1}^m \rightarrow \mathcal{P}_{\ell+1}^m$ en posant pour

$\mathbf{o} \in \mathcal{P}_{\ell+1}^m$,

$$\begin{aligned}\pi_\ell(\mathbf{o}) &= (\mathbf{o}(0), \dots, \mathbf{o}(K), 0, \dots, 0, m - (\mathbf{o}(0) + \dots + \mathbf{o}(K))), \\ \pi_{K+1}(\mathbf{o}) &= (\mathbf{o}(0), \dots, \mathbf{o}(K), m - (\mathbf{o}(0) + \dots + \mathbf{o}(K)), 0, \dots, 0).\end{aligned}$$

Évidemment,

$$\forall \mathbf{o} \in \mathcal{P}_{\ell+1}^m \quad \pi_\ell(\mathbf{o}) \preceq \mathbf{o} \preceq \pi_{K+1}(\mathbf{o}).$$

Notons par \mathcal{W}^* l'ensemble des distributions d'occupation ayant au moins une master sequence, c'est-à-dire,

$$\mathcal{W}^* = \{ \mathbf{o} \in \mathcal{P}_{\ell+1}^m : \mathbf{o}(0) \geq 1 \},$$

et soit \mathcal{N} l'ensemble des distributions d'occupation n'ayant aucune master sequence, c'est-à-dire,

$$\mathcal{N} = \{ \mathbf{o} \in \mathcal{P}_{\ell+1}^m : \mathbf{o}(0) = 0 \}.$$

Définissons :

$$\begin{aligned}\mathbf{o}_{\text{enter}}^\ell &= (1, 0, \dots, 0, m - 1), & \mathbf{o}_{\text{exit}}^\ell &= (0, \dots, 0, m), \\ \mathbf{o}_{\text{enter}}^{K+1} &= (m, 0, \dots, 0), & \mathbf{o}_{\text{exit}}^{K+1} &= (0, m, 0, \dots, 0).\end{aligned}$$

Les distributions d'occupation $\mathbf{o}_{\text{enter}}^\ell$ et $\mathbf{o}_{\text{exit}}^\ell$ sont les minima absolus des ensembles \mathcal{W}^* et \mathcal{N} . Les distributions $\mathbf{o}_{\text{enter}}^{K+1}$ et $\mathbf{o}_{\text{exit}}^{K+1}$ sont les maxima absolus des ensembles \mathcal{W}^* et \mathcal{N} . Soit $\Psi_{\mathbf{O}}$ l'application de couplage définie dans la section 9.4. Nous définissons une application inférieure $\Psi_{\mathbf{O}}^\ell$ en posant pour $\mathbf{o} \in \mathcal{P}_{\ell+1}^m$ et $\mathbf{r} \in \mathcal{R}$,

$$\Psi_{\mathbf{O}}^\ell(\mathbf{o}, \mathbf{r}) = \begin{cases} \Psi_{\mathbf{O}}(\mathbf{o}, \mathbf{r}) & \text{si } \mathbf{o} \in \mathcal{N} \text{ et } \Psi_{\mathbf{O}}(\mathbf{o}, \mathbf{r}) \notin \mathcal{W}^* \\ \mathbf{o}_{\text{enter}}^\ell & \text{si } \mathbf{o} \in \mathcal{N} \text{ et } \Psi_{\mathbf{O}}(\mathbf{o}, \mathbf{r}) \in \mathcal{W}^* \\ \pi_\ell(\Psi_{\mathbf{O}}(\pi_\ell(\mathbf{o}), \mathbf{r})) & \text{si } \mathbf{o} \in \mathcal{W}^* \text{ et } \Psi_{\mathbf{O}}(\pi_\ell(\mathbf{o}), \mathbf{r}) \notin \mathcal{N} \\ \mathbf{o}_{\text{exit}}^\ell & \text{si } \mathbf{o} \in \mathcal{W}^* \text{ et } \Psi_{\mathbf{O}}(\pi_\ell(\mathbf{o}), \mathbf{r}) \in \mathcal{N} \end{cases}$$

Également, nous définissons une application supérieure $\Psi_{\mathbf{O}}^{K+1}$ en posant pour $\mathbf{o} \in \mathcal{P}_{\ell+1}^m$ et $\mathbf{r} \in \mathcal{R}$,

$$\Psi_{\mathbf{O}}^{K+1}(\mathbf{o}, \mathbf{r}) = \begin{cases} \Psi_{\mathbf{O}}(\mathbf{o}, \mathbf{r}) & \text{si } \mathbf{o} \in \mathcal{N} \text{ et } \Psi_{\mathbf{O}}(\mathbf{o}, \mathbf{r}) \notin \mathcal{W}^* \\ \mathbf{o}_{\text{enter}}^{K+1} & \text{si } \mathbf{o} \in \mathcal{N} \text{ et } \Psi_{\mathbf{O}}(\mathbf{o}, \mathbf{r}) \in \mathcal{W}^* \\ \pi_{K+1}(\Psi_{\mathbf{O}}(\pi_{K+1}(\mathbf{o}), \mathbf{r})) & \text{si } \mathbf{o} \in \mathcal{W}^* \text{ et } \Psi_{\mathbf{O}}(\pi_{K+1}(\mathbf{o}), \mathbf{r}) \notin \mathcal{N} \\ \mathbf{o}_{\text{exit}}^{K+1} & \text{si } \mathbf{o} \in \mathcal{W}^* \text{ et } \Psi_{\mathbf{O}}(\pi_{K+1}(\mathbf{o}), \mathbf{r}) \in \mathcal{N} \end{cases}$$

Le lemme 9.4.4 entraîne que l'application Ψ_O^ℓ reste en-dessous de l'application Ψ_O et l'application Ψ_O^{K+1} reste au-dessus de l'application Ψ_O au sens suivant :

$$\forall o \in \mathcal{P}_{\ell+1}^m \quad \forall r \in \mathcal{R} \quad \Psi_O^\ell(o, r) \preceq \Psi_O(o, r) \preceq \Psi_O^{K+1}(o, r).$$

Nous définissons un processus inférieur $(O_n^\ell)_{n \geq 0}$ et un processus supérieur $(O_n^{K+1})_{n \geq 0}$ à l'aide d'une suite i.i.d. $(R_n)_{n \geq 1}$ et des applications Ψ_O^ℓ et Ψ_O^{K+1} . Soit $o \in \mathcal{P}_{\ell+1}^m$ le point de départ des processus. Nous posons $O_0^\ell = O_0^{K+1} = o$ et

$$\forall n \geq 1 \quad O_n^\ell = \Psi_O^\ell(O_{n-1}^\ell, R_n), \quad O_n^{K+1} = \Psi_O^{K+1}(O_{n-1}^{K+1}, R_n).$$

Proposition 10.1.1. *Supposons que les processus $(O_n^\ell)_{n \geq 0}$, $(O_n)_{n \geq 0}$ et $(O_n^{K+1})_{n \geq 0}$, sont issus d'une même distribution d'occupation o . Nous avons alors*

$$\forall n \geq 0 \quad O_n^\ell \preceq O_n \preceq O_n^{K+1}.$$

Démonstration. Nous allons montrer l'inégalité par récurrence sur $n \in \mathbb{N}$. Pour $n = 0$ nous avons bien $O_0 = O_0^\ell = O_0^{K+1}(0) = o$. Supposons que l'inégalité est vraie au temps $n \in \mathbb{N}$, c'est-à-dire que $O_n^\ell \preceq O_n \preceq O_n^{K+1}$. Nous avons

$$\begin{aligned} O_{n+1}^\ell &= \Psi_O^\ell(O_n^\ell, R_{n+1}), \\ O_{n+1} &= \Psi_O(O_n, R_{n+1}), \\ O_{n+1}^{K+1} &= \Psi_O^{K+1}(O_n^{K+1}, R_{n+1}). \end{aligned}$$

Par le lemme 9.4.4 et l'hypothèse de récurrence nous obtenons

$$\Psi_O(O_n^\ell, R_{n+1}) \preceq \Psi_O(O_n, R_{n+1}) \preceq \Psi_O(O_n^{K+1}, R_{n+1}).$$

L'application Ψ_O^ℓ est en-dessous de l'application Ψ_O , et l'application Ψ_O^{K+1} est au-dessus de l'application Ψ_O , donc,

$$\begin{aligned} \Psi_O^\ell(O_n^\ell, R_{n+1}) &\preceq \Psi_O(O_n^\ell, R_{n+1}), \\ \Psi_O(O_n^{K+1}, R_{n+1}) &\preceq \Psi_O^{K+1}(O_n^{K+1}, R_{n+1}). \end{aligned}$$

Les inégalités précédentes donnent bien $O_{n+1}^\ell \preceq O_{n+1} \preceq O_{n+1}^{K+1}$. □

10.2 Dynamique des processus bornants

Nous étudions ici les dynamiques des processus $(O_n^\ell)_{n \geq 0}$ et $(O_n^{K+1})_{n \geq 0}$ dans \mathcal{W}^* . Les calculs sont les mêmes pour les deux processus. Soit θ un symbole

représentant $K + 1$ ou ℓ et notons par $(O_n^\theta)_{n \geq 0}$ le processus correspondant. Pour le processus $(O_n^\theta)_{n \geq 0}$, les états dans l'ensemble

$$\mathcal{T}^\theta = \{ o \in \mathcal{P}_{\ell+1}^m : o(0) \geq 1 \text{ et } o(0) + \dots + o(K) + o(\theta) < m \}$$

sont transitoires, et les états dans $\mathcal{N} \cup (\mathcal{W}^* \setminus \mathcal{T}^\theta)$ forment une classe de récurrence. Nous observons le mécanisme de transition du processus restreint à l'ensemble $\mathcal{W}^* \setminus \mathcal{T}^\theta$. Puisque

$$\mathcal{W}^* \setminus \mathcal{T}^\theta = \{ o \in \mathcal{P}_{\ell+1}^m : o(0) \geq 1 \text{ et } o(0) + \dots + o(K) + o(\theta) = m \},$$

un état dans $\mathcal{W}^* \setminus \mathcal{T}^\theta$ est totalement déterminé par les nombres d'occupation des classes $0, \dots, K$. Ainsi, dès que le processus $(O_n^\theta)_{n \geq 0}$ est issu d'un point de $\mathcal{W}^* \setminus \mathcal{T}^\theta$, la dynamique de $(O_n^\theta(0), \dots, O_n^\theta(K))_{n \geq 0}$ est markovienne jusqu'au temps de sortie de $\mathcal{W}^* \setminus \mathcal{T}^\theta$. Notons $(Z_n^\theta)_{n \geq 0}$ la chaîne de Markov à espace d'états

$$\mathbb{D}^{K+1} = \{ z \in \mathbb{N}^{K+1} : z_0 + \dots + z_K \leq m \},$$

ayant pour matrice de transition : pour $z, z' \in \mathbb{D}^{K+1}$,

- Si $z_0, z'_0 \geq 1$,

$$\forall n \geq 0 \quad P(Z_{n+1}^\theta = z' \mid Z_n^\theta = z) = P\left((O_{n+1}^\theta(0), \dots, O_{n+1}^\theta(K)) = z' \mid (O_n^\theta(0), \dots, O_n^\theta(K)) = z \right).$$

- Si $z = (0, \dots, 0)$,

$$\forall n \geq 0 \quad P(Z_{n+1}^\theta = z^\theta \mid Z_n^\theta = (0, \dots, 0)) = 1$$

où $z^\ell = (1, 0, \dots, 0)$ et $z^{K+1} = (m, 0, \dots, 0)$.

- Si $z_0 \geq 1$ et $z' = (0, \dots, 0)$,

$$\forall n \geq 0 \quad P(Z_{n+1}^\theta = (0, \dots, 0) \mid Z_n^\theta = z) = \sum_{z': z'_0=0} P\left((O_{n+1}^\theta(0), \dots, O_{n+1}^\theta(K)) = z' \mid (O_n^\theta(0), \dots, O_n^\theta(K)) = z \right).$$

Les autres coefficients non diagonaux de la matrice de transition sont nuls. Les coefficients diagonaux sont choisis de sorte que la matrice soit stochastique, i.e., la somme de chaque ligne vaut 1. Notons par $p^\theta(z, z')$ la probabilité de transition précédente et calculons sa valeur. Nous introduisons d'abord quelques

notations. Pour tout $d \geq 1$ et tout $v \in \mathbb{R}^d$ nous notons par $|v|_1$ la norme 1 du vecteur v :

$$|v|_1 = |v_1| + \cdots + |v_d|.$$

Pour une matrice carrée $M \in \mathbb{R}^{d^2}$, et $i \in \{1, \dots, d\}$, nous notons par $M(i, \cdot)$ ou $M_{i, \cdot}$ la i -ème ligne de M , et par $M(\cdot, i)$ ou $M_{\cdot, i}$ la i -ème colonne de M . Nous notons aussi par $|M|_1$ la norme 1 de M dans \mathbb{R}^{d^2} :

$$|M|_1 = \sum_{i,j=1}^d |M_{ij}|.$$

Nous disons qu'un vecteur $s \in \mathbb{D}^{K+1}$ est compatible avec un autre vecteur $z \in \mathbb{D}^{K+1}$, et nous écrivons $s \sim z$, si

$$\begin{aligned} z(i) = 0 &\Rightarrow s(i) = 0 & \forall i \in \{0, \dots, K\}, \\ |z|_1 = m &\Rightarrow |s|_1 = m. \end{aligned}$$

Nous disons qu'une matrice $b \in \mathbb{N}^{(K+1)^2}$ est compatible avec les vecteurs $s, z' \in \mathbb{D}^{K+1}$ et nous écrivons $b \sim (s, z')$ si

$$\forall i \in \{0, \dots, K\} \quad |b(i, \cdot)|_1 \leq s_i \quad \text{and} \quad |b(\cdot, i)|_1 \leq z'_i.$$

Finalement nous définissons, pour $i \in \{0, \dots, K\} \cup \{\theta\}$, le vecteur $M_H(i) \in \mathbb{D}^{K+1}$ par :

$$M_H(i) = (M_H(i, 0), \dots, M_H(i, K)).$$

Soient $z, z' \in \mathbb{D}^{K+1}$ tels que $z_0, z'_0 \geq 1$. Ensuite nous utilisons la définition du mécanisme de transition de $(O_n^\theta)_{n \geq 0}$ pour calculer la valeur de $p^\theta(z, z')$:

$$p^\theta(z, z') = \sum_{s \sim z} \sum_{b \sim (s, z')} p^\theta(z, s, b, z'),$$

où $p^\theta(z, s, b, z')$ est la probabilité que, sachant $Z_n^\theta = z$:

- pour $i \in \{0, \dots, K\}$, s_i individus de la classe i sont sélectionnés, et $m - |s|_1$ individus de la classe θ sont sélectionnés. La probabilité de cet événement est donnée par :

$$\frac{m!}{s_0! \cdots s_K! (m - |s|_1)!} \times \frac{(\sigma z_0)^{s_0} z_1^{s_1} \cdots z_K^{s_K} (m - |z|_1)^{m - |s|_1}}{((\sigma - 1)z_0 + m)^m}.$$

• pour $i, j \in \{0, \dots, K\}$, b_{ij} individus mutent de la classe i vers la classe j , et $s_i - |b(i, \cdot)|_1$ individus mutent de la classe i vers la classe θ . Pour $i \in \{0, \dots, K\}$, la probabilité de cet événement est donnée par :

$$\frac{s_i!}{b_{i0}! \cdots b_{iK}!(s_i - |b(i, \cdot)|_1)!} \times M_H(i, 0)^{b_{i0}} \cdots M_H(i, K)^{b_{iK}} (1 - |M_H(i)|_1)^{s_i - |b(i, \cdot)|_1},$$

• pour $j \in \{0, \dots, K\}$, $z'_j - |b(\cdot, j)|_1$ individus mutent de la classe de Hamming θ vers la classe j , et $m - |s|_1 - |z'|_1 + |b|_1$ individus de la classe θ ne mutent vers aucune des classes $\{0, \dots, K\}$. La probabilité de cet événement est donnée par :

$$\frac{(m - |s|_1)!}{(z'_0 - |b(\cdot, 0)|_1)! \cdots (z'_K - |b(\cdot, K)|_1)!(m - |s|_1 - |z'|_1 + |b|_1)!} \\ \times M_H(\theta, 0)^{z'_0 - |b(\cdot, 0)|_1} \cdots M_H(\theta, K)^{z'_K - |b(\cdot, K)|_1} (1 - |M_H(\theta)|_1)^{m - |s|_1 - |z'|_1 + |b|_1}.$$

Enfin,

$$p^\theta(z, s, b, z') = \frac{m!}{s_0! \cdots s_K!(m - |s|_1)!} \times \frac{(\sigma z_0)^{s_0} z_1^{s_1} \cdots z_K^{s_K} (m - |z|_1)^{m - |s|_1}}{((\sigma - 1)z_0 + m)^m} \times \\ \prod_{i=0}^K \frac{s_i!}{b_{i0}! \cdots b_{iK}!(s_i - |b_{i \cdot}|_1)!} \times M_H(i, 0)^{b_{i0}} \cdots M_H(i, K)^{b_{iK}} (1 - |M_H(i)|_1)^{s_i - |b_{i \cdot}|_1} \\ \times \frac{(m - |s|_1)!}{(z'_0 - |b_{\cdot 0}|_1)! \cdots (z'_K - |b_{\cdot K}|_1)!(m - |s|_1 - |z'|_1 + |b|_1)!} \\ \times M_H(\theta, 0)^{z'_0 - |b_{\cdot 0}|_1} \cdots M_H(\theta, K)^{z'_K - |b_{\cdot K}|_1} (1 - |M_H(\theta)|_1)^{m - |s|_1 - |z'|_1 + |b|_1}.$$

10.3 Bornes sur la mesure invariante

Notons par μ_O^ℓ , μ_O , μ_O^{K+1} les mesures de probabilité invariantes des processus $(O_n^\ell)_{n \geq 0}$, $(O_n)_{n \geq 0}$, $(O_n^{K+1})_{n \geq 0}$. Soit ν_K la mesure image de μ_O par l'application

$$o \in \mathcal{P}_{\ell+1}^m \mapsto \frac{o(0) + \cdots + o(K)}{m} \in [0, 1].$$

Pour toute fonction $g : [0, 1] \rightarrow \mathbb{R}$,

$$\begin{aligned} \int_{[0,1]} g \, d\nu_K &= \int_{\mathcal{P}_{\ell+1}^m} g \left(\frac{o(0) + \cdots + o(K)}{m} \right) d\mu_O(o) \\ &= \lim_{n \rightarrow \infty} \mathbb{E} \left(g \left(\frac{O_n(0) + \cdots + O_n(K)}{m} \right) \right). \end{aligned}$$

Fixons $g : [0, 1] \rightarrow \mathbb{R}$ une fonction croissante telle que $g(0) = 0$. Nous avons (grâce à la proposition 10.1.1) les inégalités suivantes. Pour tout $n \geq 0$,

$$\begin{aligned} g \left(\frac{O_n^\ell(0) + \cdots + O_n^\ell(K)}{m} \right) \\ \leq g \left(\frac{O_n(0) + \cdots + O_n(K)}{m} \right) \leq \\ g \left(\frac{O_n^{K+1}(0) + \cdots + O_n^{K+1}(K)}{m} \right). \end{aligned}$$

Nous prenons l'espérance, nous envoyons n vers l' ∞ , et nous obtenons

$$\begin{aligned} \int_{\mathcal{P}_{\ell+1}^m} g \left(\frac{o(0) + \cdots + o(K)}{m} \right) d\mu_O^\ell(o) \\ \leq \int_{[0,1]} g \, d\nu_K \leq \\ \int_{\mathcal{P}_{\ell+1}^m} g \left(\frac{o(0) + \cdots + o(K)}{m} \right) d\mu_O^{K+1}(o). \end{aligned}$$

Nous cherchons ensuite des estimations sur les intégrales ci-dessus. La stratégie est la même pour les intégrales inférieure et supérieure. Soit θ un symbole représentant $K + 1$ ou ℓ . Nous étudions la mesure de probabilité invariante μ_O^θ . Pour le processus $(O_n^\theta)_{n \geq 0}$, les états dans \mathcal{T}^θ sont transitoires et les populations dans $\mathcal{N} \cup (\mathcal{W}^* \setminus \mathcal{T}^\theta)$ forment une classe de récurrence. Nous appliquons ensuite un résultat de renouvellement, que nous donnons ci-dessous, afin d'exprimer la mesure de probabilité invariante μ^θ en fonction du processus $(Z_n^\theta)_{n \geq 0}$. Soit $(X_n)_{n \geq 0}$ une chaîne de Markov discrète à valeurs dans l'espace d'états fini \mathcal{E} . Supposons que $(X_n)_{n \geq 0}$ est ergodique et soit μ sa mesure de probabilité invariante.

Proposition 10.3.1. *Soient \mathcal{W}^* un sous-ensemble de \mathcal{E} et e un point dans $\mathcal{E} \setminus \mathcal{W}^*$. Soit f une fonction de \mathcal{E} dans \mathbb{R} . Soient*

$$\tau^* = \inf \{ n \geq 0 : X_n \in \mathcal{W}^* \}, \quad \tau = \inf \{ n \geq \tau^* : X_n = e \}.$$

Nous avons

$$\int_{\mathcal{E}} f(x) d\mu(x) = \frac{1}{\mathbb{E}(\tau | X_0 = e)} \mathbb{E} \left(\sum_{n=0}^{\tau-1} f(X_n) \mid X_0 = e \right).$$

Démonstration. Définissons deux suites de temps d'arrêt, $(\tau_k^*)_{k \geq 1}$ et $(\tau_k)_{k \geq 0}$. Posons $\tau_0 = 0$ et

$$\begin{aligned} \tau_1^* &= \inf \{ n \geq 0 : X_n \in \mathcal{W}^* \}, & \tau_1 &= \inf \{ n \geq \tau_1^* : X_n = e \}, \\ \vdots & & \vdots & \\ \tau_k^* &= \inf \{ n \geq \tau_{k-1}^* : X_n \in \mathcal{W}^* \}, & \tau_k &= \inf \{ n \geq \tau_k^* : X_n = e \}, \\ \vdots & & \vdots & \end{aligned}$$

Par la propriété de Markov forte, les excursions

$$(X_n, \tau_k \leq n \leq \tau_{k+1}), \quad k \geq 1,$$

sont indépendantes et identiquement distribuées. En particulier, la suite

$$(\tau_{k+1} - \tau_k)_{k \geq 1}$$

est une suite de variables aléatoires i.i.d.. La loi de chacune des variables aléatoires $(\tau_{k+1} - \tau_k)$ est la même que la loi du temps d'arrêt τ_1 quand le processus $(X_n)_{n \geq 0}$ est issu de e . Pour $k \geq 1$, nous décomposons τ_k comme la somme

$$\tau_k = \tau_1 + \sum_{h=1}^{k-1} (\tau_{h+1} - \tau_h).$$

Puisque l'espace d'états \mathcal{E} est fini et que la chaîne de Markov $(X_n)_{n \geq 0}$ est ergodique, le temps d'arrêt τ_1 est fini avec probabilité 1, et donc intégrable. La loi des grands nombres classique entraîne que

$$\lim_{k \rightarrow \infty} \frac{\tau_k}{k} = \mathbb{E}_e(\tau_1) \quad \text{avec probabilité 1.}$$

Si le processus $(X_n)_{n \geq 0}$ est issu de e , les temps d'arrêt τ_1^* , τ_1 satisfont $\tau_1^* \geq 1$, $\tau_1 \geq 2$, en particulier, l'espérance $\mathbb{E}_e(\tau_1)$ est strictement positive, d'où

$$\lim_{k \rightarrow \infty} \tau_k = +\infty \quad \text{avec probabilité 1.}$$

Nous définissons ensuite

$$\forall n \geq 0 \quad K(n) = \max \{ k \geq 0 : \tau_k \leq n \}.$$

Nous voyons que, avec probabilité 1, $K(n)$ est fini pour tout $n \geq 0$. Par définition de $K(n)$, nous avons

$$\forall n \geq 0 \quad \tau_{K(n)} \leq n < \tau_{K(n)+1},$$

De plus, comme $\tau_k < +\infty$ pour tout k , nous avons

$$\lim_{n \rightarrow \infty} K(n) = +\infty \quad \text{avec probabilité 1.}$$

L'inégalité précédente peut alors s'écrire sous la forme :

$$\frac{\tau_{K(n)}}{K(n)} \leq \frac{n}{K(n)} < \frac{\tau_{K(n)+1}}{K(n)+1} \times \frac{K(n)+1}{K(n)}.$$

Nous faisons tendre n vers l' ∞ et nous concluons que

$$\lim_{n \rightarrow \infty} \frac{K(n)}{n} = \frac{1}{E_e(\tau_1)} \quad \text{avec probabilité 1.}$$

Nous supposons que le processus $(X_n)_{n \geq 0}$ est issu de e . Soit f une fonction de \mathcal{E} dans \mathbb{R} . Par le théorème ergodique pour les chaînes de Markov, nous avons

$$\lim_{n \rightarrow \infty} E_e(f(X_n)) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{s=0}^n f(X_s).$$

Nous décomposons la dernière somme comme suit :

$$\sum_{s=0}^n f(X_s) = \sum_{k=1}^{K(n)} \sum_{s=\tau_{k-1}}^{\tau_k-1} f(X_s) + \sum_{s=\tau_{K(n)}}^n f(X_s),$$

La somme

$$N_k = \sum_{s=\tau_{k-1}}^{\tau_k-1} f(X_s)$$

est une fonction déterministe de l'excursion $(X_n, \tau_{k-1} \leq n \leq \tau_k)$, donc les variables aléatoires $(N_k, k \geq 1)$ sont i.i.d.. Nous appliquons la loi des grands nombres classique et nous obtenons

$$\lim_{n \rightarrow \infty} \frac{1}{K(n)} \sum_{k=1}^{K(n)} N_k = E_e(N_1) \quad \text{avec probabilité 1.}$$

Nous écrivons

$$\frac{1}{n} \sum_{s=0}^n f(X_s) = \frac{K(n)}{n} \times \frac{1}{K(n)} \sum_{k=1}^{K(n)} N_k + \frac{1}{n} \sum_{s=\tau_{K(n)}}^n f(X_s).$$

Nous avons

$$\begin{aligned} \left| \frac{1}{n} \sum_{s=\tau_{K(n)}}^n f(X_s) \right| &\leq \left(\sup_{[0,1]} |f| \right) \frac{1}{n} (\tau_{K(n)+1} - \tau_{K(n)}) \\ &\leq \left(\sup_{[0,1]} |f| \right) \frac{K(n)}{n} \left(\frac{K(n)+1}{K(n)} \frac{\tau_{K(n)+1}}{K(n)+1} - \frac{\tau_{K(n)}}{K(n)} \right), \end{aligned}$$

qui tend vers 0 avec probabilité 1 quand n tend vers $l'\infty$. Nous concluons que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{s=0}^n f(X_s) = \frac{E_e(N_1)}{E_e(\tau_1)} \quad \text{avec probabilité 1.}$$

Ceci entraîne le résultat voulu. □

Nous appliquons le résultat de renouvellement à la chaîne de Markov $(O_n^\theta)_{n \geq 0}$ restreinte à $\mathcal{N} \cup (\mathcal{W}^* \setminus \mathcal{J}^\theta)$, à l'ensemble $\mathcal{W}^* \setminus \mathcal{J}^\theta$, à la distribution d'occupation o_{exit}^θ et à la fonction $o \mapsto g((o(0) + \dots + o(K))/m)$. Soient

$$\tau^* = \inf \{ n \geq 0 : O_n^\theta \in \mathcal{W}^* \}, \quad \tau = \inf \{ n \geq \tau^* : O_n^\theta = o_{\text{exit}}^\theta \}.$$

Nous avons alors

$$\int_{\mathcal{P}_{\ell+1}^m} g\left(\frac{o(0) + \dots + o(K)}{m}\right) d\mu_O^\theta(o) = \frac{E\left(\sum_{n=0}^{\tau} g\left(\frac{O_n^\theta(0) + \dots + O_n^\theta(K)}{m}\right) \middle| O_0^\theta = o_{\text{exit}}^\theta\right)}{E(\tau | O_0^\theta = o_{\text{exit}}^\theta)}.$$

Tant que le processus $(O_n^\theta)_{n \geq 0}$ est dans $\mathcal{W}^* \setminus \mathcal{J}^\theta$, la dynamique des coordonnées $(O_n^\theta(0), \dots, O_n^\theta(K))_{n \geq 0}$ est la même que celle de la chaîne de Markov $(Z_n^\theta)_{n \geq 0}$ définie à la fin de la section précédente. Supposons que $(Z_n^\theta)_{n \geq 0}$ est issue de z^θ , où z^θ est le point de \mathbb{D}^{K+1} donné par

$$z^\ell = (1, 0, \dots, 0), \quad z^{K+1} = (m, 0, \dots, 0).$$

Soit τ_0 le premier temps où la coordonnée 0 du processus $(Z_n^\theta)_{n \geq 0}$ s'annule, i.e.,

$$\tau_0 = \inf \{ n \geq 0 : Z_n^\theta(0) = 0 \}.$$

Comme le processus $(O_n^\theta)_{n \geq 0}$ entre toujours dans $\mathcal{W}^* \setminus \mathcal{T}^\theta$ par le point o_{enter}^θ , la loi du temps τ_0 est la même que la loi de $\tau - \tau^*$ pour le processus $(O_n^\theta)_{n \geq 0}$ issu de o_{exit}^θ . Nous en déduisons que $((O_n^\theta(0), \dots, O_n^\theta(K)), \tau^* \leq n \leq \tau)$ et $(Z_n^\theta, 0 \leq n \leq \tau_0)$ ont même loi. Il s'ensuit que

$$\mathbb{E}(\tau - \tau^* \mid O_0^\theta = o_{\text{exit}}^\theta) = \mathbb{E}(\tau_0 \mid Z_0^\theta = z^\theta),$$

$$\mathbb{E} \left(\sum_{n=\tau^*}^{\tau} g \left(\frac{O_n^\theta(0) + \dots + O_n^\theta(K)}{m} \right) \mid O_0^\theta = o_{\text{exit}}^\theta \right) = \\ \mathbb{E} \left(\sum_{n=0}^{\tau_0} g \left(\frac{Z_n^\theta(0) + \dots + Z_n^\theta(K)}{m} \right) \mid Z_0^\theta = z^\theta \right).$$

Nous pouvons alors réécrire la formule pour la mesure de probabilité invariante μ_O^θ de la manière suivante :

$$\int_{\mathcal{P}_{\ell+1}^m} g \left(\frac{o(0) + \dots + o(K)}{m} \right) d\mu_O^\theta(o) = \\ \frac{\mathbb{E} \left(\sum_{n=0}^{\tau^*} g \left(\frac{O_n^\theta(0) + \dots + O_n^\theta(K)}{m} \right) \mid O_0^\theta = o_{\text{exit}}^\theta \right)}{\mathbb{E}(\tau^* \mid O_0^\theta = o_{\text{exit}}^\theta) + \mathbb{E}(\tau_0 \mid Z_0^\theta = z^\theta)} \\ + \frac{\mathbb{E} \left(\sum_{n=0}^{\tau_0} g \left(\frac{Z_n^\theta(0) + \dots + Z_n^\theta(K)}{m} \right) \mid Z_0^\theta = z^\theta \right)}{\mathbb{E}(\tau^* \mid O_0^\theta = o_{\text{exit}}^\theta) + \mathbb{E}(\tau_0 \mid Z_0^\theta = z^\theta)}.$$

Nous essayerons ensuite d'estimer chacun des termes du membre de droite de cette formule.

10.4 Comparaison avec la master sequence

Les processus que nous avons construits sont stochastiquement dominés par les processus d'occupation inférieur et supérieur $(O_n^\ell)_{n \geq 0}$, $(O_n^{K+1})_{n \geq 0}$ avec $K = 0$.

Appelons $(\Theta_n^\ell)_{n \geq 0}$ et $(\Theta_n^1)_{n \geq 0}$ les processus d'occupation inférieur et supérieur avec $K = 0$. Pour un processus stochastique $(X_n)_{n \geq 0}$ à valeurs dans $\mathcal{P}_{\ell+1}^m$, nous définissons le temps de sortie de l'ensemble \mathcal{W}^* pour le processus $(X_n)_{n \geq 0}$ comme suit :

$$\tau(X) = \inf \{ n \geq 0 : X_n \in \mathcal{N} \}.$$

Soit $o \in \mathcal{W}^*$ le point de départ des processus $(O_n^\ell)_{n \geq 0}$, $(O_n^{K+1})_{n \geq 0}$, $(\Theta_n^\ell)_{n \geq 0}$ et $(\Theta_n^1)_{n \geq 0}$. Nous avons

$$\begin{aligned} \forall n \in \{0, \dots, \tau(\Theta^\ell)\} & \quad \Theta_n^\ell \preceq O_n^\ell \preceq O_n^{K+1} \preceq \Theta_n^1, \\ \forall n \in \{0, \dots, \tau(O^{K+1})\} & \quad O_n^{K+1} \preceq \Theta_n^1. \end{aligned}$$

En particulier, pour tout $o \in \mathcal{W}^*$,

$$\begin{aligned} \mathbb{E}(\tau(\Theta^\ell) \mid \Theta_0^\ell = o) & \leq \mathbb{E}(\tau(O^\ell) \mid \Theta_0^\ell = o) \\ & \leq \mathbb{E}(\tau(O^{K+1}) \mid \Theta_0^\ell = o) \leq \mathbb{E}(\tau(\Theta^1) \mid \Theta_0^\ell = o). \end{aligned}$$

De plus, la relation entre les processus Θ et O est naturellement conservée par les chaînes de Markov correspondantes. Soient $(\tilde{Z}_n^\ell)_{n \geq 0}$ et $(\tilde{Z}_n^1)_{n \geq 0}$ les chaînes de Markov obtenues à partir des processus $(\Theta_n^\ell)_{n \geq 0}$ et $(\Theta_n^1)_{n \geq 0}$. Les chaînes de Markov $(\tilde{Z}_n^\ell)_{n \geq 0}$, $(\tilde{Z}_n^1)_{n \geq 0}$ répliquent la dynamique des processus $(\Theta_n^\ell)_{n \geq 0}$, $(\Theta_n^1)_{n \geq 0}$ dans \mathcal{W}^* , de même que les chaînes $(Z_n^\ell)_{n \geq 0}$, $(Z_n^{K+1})_{n \geq 0}$ répliquent la dynamique des processus $(O_n^\ell)_{n \geq 0}$, $(O_n^{K+1})_{n \geq 0}$ dans \mathcal{W}^* . Nous pouvons donc étendre le couplage existant entre les différents processus d'occupation aux chaînes de Markov $(\tilde{Z}_n^\ell)_{n \geq 0}$, $(\tilde{Z}_n^1)_{n \geq 0}$, $(Z_n^\ell)_{n \geq 0}$, $(Z_n^{K+1})_{n \geq 0}$. Pour un processus stochastique $(X_n)_{n \geq 0}$ à valeurs dans \mathbb{D}^{K+1} , nous définissons le temps de sortie de l'ensemble $\{z \in \mathbb{D}^{K+1} : z_0 \geq 1\}$ pour le processus $(X_n)_{n \geq 0}$ comme suit :

$$\tau(X) = \inf \{ n \geq 0 : X_n(0) = 0 \}.$$

Nous obtenons les relations suivantes :

$$\begin{aligned} \forall n \in \{0, \dots, \tau(\tilde{Z}^\ell)\} & \quad \tilde{Z}_n^\ell \leq Z_n^\ell(0) \leq Z_n^{K+1}(0) \leq \tilde{Z}_n^1, \\ \forall n \in \{0, \dots, \tau(Z^{K+1})\} & \quad Z_n^{K+1}(0) \leq \tilde{Z}_n^1. \end{aligned}$$

Les chaînes de Markov $(\tilde{Z}_n^\ell)_{n \geq 0}$ et $(\tilde{Z}_n^1)_{n \geq 0}$ ont été étudiées en détail dans [11]. Grâce aux inégalités précédentes, la plupart des estimations asymptotiques développés dans [11] pour les chaînes de Markov $(\tilde{Z}_n^\ell)_{n \geq 0}$ et $(\tilde{Z}_n^1)_{n \geq 0}$ restent valables pour les processus $(Z_n^\ell(0))_{n \geq 0}$ et $(Z_n^{K+1}(0))_{n \geq 0}$.

Chapitre 11

Chaînes de Markov

Nous étudions dans cette section les chaînes de Markov $(Z_n^\ell)_{n \geq 0}$ et $(Z_n^{K+1})_{n \geq 0}$. Les calculs sont les mêmes pour les deux processus, soit donc θ un symbole représentant ℓ ou $K + 1$ et étudions la chaîne de Markov $(Z_n^\theta)_{n \geq 0}$. Nous allons faire des estimations dans le régime asymptotique

$$\ell \rightarrow +\infty, \quad m \rightarrow +\infty, \quad q \rightarrow 0, \quad \ell q \rightarrow a \in]0, +\infty[.$$

Nous dirons qu'une propriété est vraie asymptotiquement si elle est vraie pour ℓ, m assez grands, q assez petit et ℓq assez proche de a .

11.1 Grandes déviations pour la matrice de transition

Nous définissons l'ensemble $\mathcal{D}_K \subset \mathbb{R}^{K+1}$ par

$$\mathcal{D}_K = \left\{ r \in \mathbb{R}^{K+1} : r_0 \geq 0, \dots, r_K \geq 0 \text{ and } r_0 + \dots + r_K \leq 1 \right\}.$$

Pour $p, t \in \mathcal{D}_K$ nous définissons $I_K(p, t)$ comme suit :

$$I_K(p, t) = \sum_{k=0}^K t_k \ln \frac{t_k}{p_k} + (1 - |t|_1) \ln \frac{1 - |t|_1}{1 - |p|_1}.$$

Par convention, $0 \ln(0/0) = 0$. L'application $I(p, \cdot)$ est la fonction de taux gouvernant les grandes déviations d'une distribution multinomiale de paramètres

n et p_0, \dots, p_K . Nous avons l'estimation suivante pour les coefficients multinomiaux.

Lemme 11.1.1. Soient $n \geq N \geq 1$ et $i_1, \dots, i_N \in \mathbb{N}$ tels que $i_1 + \dots + i_N = n$. Nous avons

$$\left| \ln \frac{n!}{i_1! \cdots i_N!} + \sum_{k=1}^N i_k \ln \frac{i_k}{n} \right| \leq N \ln n + 2N.$$

Démonstration. Soit, pour $n \in \mathbb{N}$, $\phi(n) = \ln n! - n \ln n + n$. Nous avons

$$\begin{aligned} \ln \frac{n!}{i_1! \cdots i_N!} &= n \ln n - n + \phi(n) - \sum_{k=1}^N (i_k \ln i_k - i_k + \phi(i_k)) \\ &= - \sum_{k=1}^N i_k \ln \frac{i_k}{n} + \phi(n) - \sum_{k=1}^N \phi(i_k). \end{aligned}$$

Puisque $1 \leq \phi(n) \leq \ln n + 2$ pour tout $n \geq 1$, nous concluons que

$$1 - N(\ln n + 2) \leq \phi(n) - \sum_{k=1}^N \phi(i_k) \leq \ln n + 2 - N,$$

ce qui entraîne l'inégalité cherchée. \square

Nous définissons une fonction $h : \mathcal{D}_K \rightarrow \mathcal{D}_K$ en prenant, pour tout $r \in \mathcal{D}_K$,

$$h(r) = \frac{1}{(\sigma - 1)r_0 + 1} (\sigma r_0, r_1, \dots, r_K).$$

Nous définissons aussi une fonction $I_\ell : \mathcal{D}_K \times \mathcal{D}_K \times [0, 1]^{(K+1)^2} \times \mathcal{D}_K \rightarrow [0, +\infty]$ en prenant, pour $r, \xi, t \in \mathcal{D}_K$ et $\beta \in [0, 1]^{(K+1)^2}$,

$$\begin{aligned} I_\ell(r, \xi, \beta, t) &= I_K(h(r), \xi) + \sum_{k=0}^K \xi_k I_K(M_H(k), \xi_k^{-1} \beta(k, \cdot)) \\ &\quad + (1 - |\xi|_1) I_K(M_H(\theta), (1 - |\xi|_1)^{-1} (t_0 - |\beta(\cdot, 0)|_1, \dots, t_K - |\beta(\cdot, K)|_1)). \end{aligned}$$

Nous prenons le logarithme dans l'expression de p^θ que nous avons donnée à la fin de la section 10.2. Nous avons, pour tout $z, z', s \in \mathbb{D}^{K+1}$ et $\mathbf{b} \in \mathbb{N}^{(K+1)^2}$,

$$\begin{aligned} \ln p^\theta(z, s, \mathbf{b}, z') &= -m I_K \left(h \left(\frac{z}{m}, \frac{s}{m} \right) - \sum_{k=0}^K s_k I_K (M_H(k), s_k^{-1} \mathbf{b}(k, \cdot)) \right. \\ &\quad \left. - (m - |s|_1) I_K \left(M_H(\theta), (m - |s|_1)^{-1} (z'_0 - |\mathbf{b}(\cdot, 0)|_1, \dots, z'_K - |\mathbf{b}(\cdot, K)|_1) \right) \right) \\ &\quad + \Phi(z, s, \mathbf{b}, z') = -m I_\ell \left(\frac{z}{m}, \frac{s}{m}, \frac{\mathbf{b}}{m}, \frac{z'}{m} \right) + \Phi(z, s, \mathbf{b}, z'). \end{aligned}$$

Le terme d'erreur $\Phi(z, s, \mathbf{b}, z')$ satisfait, pour tout $m \geq 1$,

$$\forall z, z', s \in \mathbb{D}^{K+1} \quad \forall \mathbf{b} \in \mathbb{N}^{(K+1)^2} \quad |\Phi(z, s, \mathbf{b}, z')| \leq C(K)(\ln m + 1),$$

où $C(K)$ est une constante qui dépend de K mais ne dépend pas de m . Rappelons qu'asymptotiquement, nous avons pour tout $i, j \geq 0$,

$$\lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} M_H(i, j) = \begin{cases} e^{-a} \frac{a^{j-i}}{(j-i)!} & \text{si } i \leq j, \\ 0 & \text{si } i > j. \end{cases}$$

Définissons, pour tout $i, j \in \mathbb{N}$,

$$M_\infty(i, j) = \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} M_H(i, j),$$

et pour $k \in \{0, \dots, K\} \cup \{\theta\}$, soit $M_\infty(k)$ le vecteur de \mathcal{D}_K donné par

$$M_\infty(k) = (M_\infty(k, 0), \dots, M_\infty(k, K)).$$

Pour $\mathbf{t} \in \mathcal{D}_K$, nous appelons $\mathcal{B}(\mathbf{t})$ le sous-ensemble de $[0, 1]^{(K+1)^2}$ des matrices triangulaires supérieures β telles que $\beta(0, k) + \dots + \beta(K, k) = t_k$ pour tout $k \in \{0, \dots, K\}$, i.e.,

$$\mathcal{B}(\mathbf{t}) = \{ \beta \in [0, 1]^{(K+1)^2} : \beta(i, j) = 0, i > j \text{ et } |\beta(\cdot, k)|_1 = t_k, 0 \leq k \leq K \}.$$

Pour $r, \xi, \mathbf{t} \in \mathcal{D}_K$ et $\beta \in [0, 1]^{(K+1)^2}$, nous obtenons

$$\lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} I_\ell(r, \xi, \beta, \mathbf{t}) = \begin{cases} I_K(r, \xi, \beta, \mathbf{t}) & \text{si } \beta \in \mathcal{B}(\mathbf{t}), \\ +\infty & \text{sinon,} \end{cases}$$

où la fonction $I_K(r, \xi, \beta, t)$ est donnée par

$$I_K(r, \xi, \beta, t) = I_K(h(r), \xi) + \sum_{k=0}^K \xi_k I_K(M_\infty(k), \xi_k^{-1} \beta(k, \cdot)).$$

Nous définissons une fonction $V_1 : \mathcal{D}_K \times \mathcal{D}_K \rightarrow [0, \infty]$ en prenant, pour $r, t \in \mathcal{D}_K$,

$$V_1(r, t) = \inf \{ I_K(r, \xi, \beta, t) : \xi \in \mathcal{D}_K, \beta \in \mathcal{B}(t) \}.$$

Proposition 11.1.2. *Les probabilités de transition de la chaîne de Markov $(Z_n^\theta)_{n \geq 0}$ vérifient le principe de grandes déviations gouverné par V_1 : pour tout sous-ensemble U, U' de \mathcal{D}_K et pour tout $\rho \in U$, nous avons, pour tout $n \geq 0$,*

$$\begin{aligned} & - \inf \{ V_1(\rho, t) : t \in U' \} \\ & \leq \liminf_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \frac{1}{m} \ln P(Z_{n+1}^\theta \in mU' \mid Z_n^\theta = \lfloor \rho m \rfloor) \\ & \leq \limsup_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \frac{1}{m} \ln \sup_{z \in mU} P(Z_{n+1}^\theta \in mU' \mid Z_n^\theta = z) \\ & \leq - \inf \{ V_1(r, t) : r \in \bar{U}, t \in \bar{U}' \}, \end{aligned}$$

où $\lfloor \rho m \rfloor = (\lfloor \rho_0 m \rfloor, \dots, \lfloor \rho_K m \rfloor)$.

Démonstration. Nous commençons par démontrer la borne supérieure. Soient U, U' deux sous-ensembles de \mathcal{D}_K et soit $z \in mU$. Pour tout $n \geq 0$,

$$\begin{aligned} P(Z_{n+1}^\theta \in mU' \mid Z_n^\theta = z) &= \sum_{z' \in mU' \cap \mathbb{D}^{K+1}} p^\theta(z, z') \\ &= \sum_{z' \in mU' \cap \mathbb{D}^{K+1}} \sum_{s \sim z} \sum_{b \sim (s, z')} p^\theta(z, s, b, z'). \end{aligned}$$

Grâce aux estimations développées pour le coefficient p^θ , nous avons, pour tout $m \geq 1$,

$$\begin{aligned} & \sup_{z \in mU} P(Z_{n+1}^\theta \in mU' \mid Z_n^\theta = z) \\ & \leq m^{C(K)} \max \{ p^\theta(z, s, b, z') : z \in mU, s \sim z, z' \in mU', b \sim (s, z') \} \\ & \leq m^{C(K)} \exp \left(-m \min \left\{ I_\ell \left(\frac{z}{m}, \frac{s}{m}, \frac{b}{m}, \frac{z'}{m} \right) : \right. \right. \\ & \quad \left. \left. z \in mU, s \sim z, z' \in mU', b \sim (s, z') \right\} \right), \end{aligned}$$

où $C(K)$ est une constante qui dépend de K mais ne dépend pas de m . Pour chaque $m \geq 1$, soient $z_m, s_m, z'_m \in \mathbb{D}^{K+1}$, $b_m \in \{0, \dots, m\}^{(K+1)^2}$ trois termes qui réalisent le minimum précédent. Les ensembles \mathcal{D}_K et $[0, 1]$ étant compacts, quitte à extraire une sous-suite, nous pouvons considérer que quand $m \rightarrow \infty$,

$$\frac{z_m}{m} \rightarrow \rho, \quad \frac{s_m}{m} \rightarrow \xi, \quad \frac{b_m}{m} \rightarrow \beta, \quad \frac{z'_m}{m} \rightarrow t.$$

Nous observons ensuite l'expression

$$\limsup_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha}} -I_\ell \left(\frac{z_m}{m}, \frac{s_m}{m}, \frac{b_m}{m}, \frac{z'_m}{m} \right).$$

Si β n'est pas une matrice triangulaire supérieure, ou si $|\beta(\cdot, j)|_1 \neq t_j$ pour un des $j \in \{0, \dots, K\}$, la limite est $-\infty$. Nous n'avons besoin de considérer que le cas où β est triangulaire supérieure et $|\beta(\cdot, j)|_1 = t_j$ pour tout $j \in \{0, \dots, K\}$. Dans ce cas, nous avons

$$\begin{aligned} \limsup_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha}} -I_\ell \left(\frac{z_m}{m}, \frac{s_m}{m}, \frac{b_m}{m}, \frac{z'_m}{m} \right) \\ \leq -I_K(h(\rho), \xi) - \sum_{k=0}^K \xi_k I_K(M_\infty(k), s_k^{-1} \beta(k, \cdot)). \end{aligned}$$

Ce qui entraîne la borne supérieure du principe de grandes déviations :

$$\begin{aligned} \limsup_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha}} \frac{1}{m} \ln \sup_{z \in mU} P(Z_{n+1}^\theta \in mU' \mid Z_n^\theta = z) \\ \leq -\inf \{ I_K(r, \xi, \beta, t) : r \in \bar{U}, \xi \in \mathcal{D}_K, \beta \in \mathcal{B}(t), t \in \bar{U}' \}. \end{aligned}$$

Montrons ensuite la borne inférieure. Pour tout $d \geq 1$ et $v \in \mathbb{R}^d$, nous définis-

sons $[v] = ([v_0], \dots, [v_d])$. Soient $\xi, t \in \mathcal{D}_K$ et soit $\beta \in \mathcal{B}(t)$. Nous avons

$$\begin{aligned} P(Z_{n+1}^\theta = [tm] \mid Z_n^\theta = [\rho m]) &\geq p^\theta([\rho m], [\xi m], [\beta m], [tm]) \\ &\geq m^{-C(K)} \exp\left(-mI_\ell\left(\frac{[\rho m]}{m}, \frac{[\xi m]}{m}, \frac{[\beta m]}{m}, \frac{[tm]}{m}\right)\right) \\ &\geq m^{-C(K)} \exp\left(-mI_K\left(h\left(\frac{[\rho m]}{m}\right), \frac{[\xi m]}{m}\right)\right) \\ &\quad - \sum_{k=0}^K \xi_k I_K\left(M_H(k), \frac{[\beta(k, \cdot)m]}{[\xi_k m]}\right) - (m - |[\xi m]_1|) \\ &\quad \times I_K\left(M_H(\theta), \frac{[t_0 m] - |[\beta(\cdot, 0)m]_1|, \dots, [t_K m] - |[\beta(\cdot, K)m]_1|}{m - |[\xi m]_1|}\right). \end{aligned}$$

Où, à nouveau, $C(K)$ est une constante qui dépend de K mais ne dépend pas de m . Nous prenons le logarithme et nous envoyons m et ℓ vers $l'\infty$ et q vers 0. Nous obtenons,

$$\liminf_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \frac{1}{m} \ln P(Z_{n+1}^\theta = [tm] \mid Z_n^\theta = [\rho m]) \geq -I_K(\rho, \xi, \beta, t).$$

De plus, si $t \in \mathcal{U}$, pour m assez grand, $[tm]$ appartient à $m\mathcal{U}$. Donc,

$$\liminf_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \frac{1}{m} \ln P(Z_{n+1}^\theta \in m\mathcal{U} \mid Z_n^\theta = [\rho m]) \geq -I_K(\rho, \xi, \beta, t).$$

Nous optimisons sur ξ, β, t et nous obtenons la borne inférieure du principe de grandes déviations :

$$\begin{aligned} \liminf_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \frac{1}{m} \ln P(Z_{n+1}^\theta \in m\mathcal{U} \mid Z_n^\theta = [\rho m]) \\ \geq -\inf \{ I_K(\rho, \xi, \beta, t) : \xi \in \mathcal{D}_K, \beta \in \mathcal{B}(t), t \in \mathcal{U} \}. \end{aligned}$$

□

Une démonstration similaire montre que les probabilités de transition de l pas vérifient aussi un principe de grandes déviations. Pour $l \geq 1$, nous définissons

une fonction V_l sur $\mathcal{D}_K \times \mathcal{D}_K$ ainsi :

$$V_l(r, t) = \inf \left\{ \sum_{k=0}^{l-1} I_K(\rho_k, \xi_k, \beta_k, \rho_{k+1}) : \right. \\ \left. \rho_0 = r, \rho_l = t, \rho_k, \xi_k \in \mathcal{D}_K, \beta_k \in \mathcal{B}(t) \text{ pour } 0 \leq k < l \right\}.$$

Corollaire 11.1.3. *Pour $l \geq 1$, la probabilité de transition de l pas de $(Z_n^\theta)_{n \geq 0}$ vérifie le principe de grandes déviations gouverné par V_l : pour tout sous-ensemble $\mathcal{U}, \mathcal{U}'$ de \mathcal{D}_K , pour tout $\rho \in \mathcal{U}$, nous avons, pour tout $n \geq 0$,*

$$\begin{aligned} & - \inf \{ V_l(\rho, t) : t \in \mathcal{U}' \} \\ & \leq \liminf_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \frac{1}{m} \ln P(Z_{n+1}^\theta \in m\mathcal{U}' \mid Z_n^\theta = \lfloor \rho m \rfloor) \\ & \leq \limsup_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} \frac{1}{m} \ln \sup_{z \in m\mathcal{U}} P(Z_{n+1}^\theta \in m\mathcal{U}' \mid Z_n^\theta = z) \\ & \leq - \inf \{ V_l(r, t) : r \in \bar{\mathcal{U}}, t \in \bar{\mathcal{U}}' \}. \end{aligned}$$

Nous cherchons ensuite les zéros de la fonction de taux $I_K(r, \xi, \beta, t)$. Nous avons que

$$I_K(r, \xi, \beta, t) = 0 \iff \xi = h(r) \text{ et } \frac{1}{\xi_k} \beta(k, \cdot) = M_\infty(k) \text{ pour } 0 \leq k \leq K.$$

Définissons une fonction $F = (F_0, \dots, F_K) : \mathcal{D}_K \rightarrow \mathcal{D}_K$ en prenant, pour $r \in \mathcal{D}_K$ et $k \in \{0, \dots, K\}$,

$$F_k(r) = \sum_{i=0}^k h_i(r) e^{-a} \frac{a^{k-i}}{(k-i)!}.$$

La chaîne de Markov $(Z_n^\theta)_{n \geq 0}$ peut être considérée comme une perturbation du système dynamique associé à l'application F :

$$z^0 \in \mathcal{D}_K, \quad \forall n \geq 1 \quad z^n = F(z^{n-1}).$$

Nous rappelons que, pour $r \in \mathcal{D}_K$,

$$h(r) = \frac{1}{(\sigma - 1)r_0 + 1} (\sigma r_0, r_1, \dots, r_K).$$

Nous pouvons donc réécrire, pour $k \in \{0, \dots, K\}$,

$$F_k(\mathbf{r}) = \frac{e^{-\alpha}}{(\sigma - 1)r_0 + 1} \left(\frac{\alpha^k}{k!} \sigma r_0 + \sum_{i=1}^k \frac{\alpha^{k-i}}{(k-i)!} r_i \right).$$

Soit ρ^* le point de \mathcal{D}_K donné par :

$$\forall k \in \{0, \dots, K\} \quad \rho_k^* = (\sigma e^{-\alpha} - 1) \frac{\alpha^k}{k!} \sum_{i \geq 1} \frac{i^k}{\sigma^i}.$$

Proposition 11.1.4. *Nous avons la dichotomie suivante :*

- si $\sigma e^{-\alpha} \leq 1$, la fonction F admet un seul point fixe, 0, et $(z^n)_{n \geq 0}$ converge vers 0.
- si $\sigma e^{-\alpha} > 1$, la fonction F admet deux points fixes, 0 et ρ^* . Si $z_0^0 = 0$, la suite $(z^n)_{n \in \mathbb{N}}$ converge vers 0, et si $z_0^0 > 0$, la suite $(z^n)_{n \in \mathbb{N}}$ converge vers ρ^* .

Démonstration. Pour $k \in \{0, \dots, K\}$, la fonction $F_k(\mathbf{r})$ est une fonction des paramètres r_0, \dots, r_k seulement. Nous pouvons donc résoudre par récurrence le système d'équations

$$F_k(\mathbf{r}) = r_k \quad 0 \leq k \leq K.$$

En effet, pour $k = 0$ nous avons

$$F_0(\mathbf{r}) = \frac{\sigma e^{-\alpha} r_0}{(\sigma - 1)r_0 + 1}.$$

L'équation $F_0(\mathbf{r}) = r_0$ a deux solutions : $r_0 = 0$ et $r_0 = \rho_0^*$. Pour k dans $\{1, \dots, K\}$, nous avons

$$F_k(\mathbf{r}) = \frac{e^{-\alpha}}{(\sigma - 1)r_0 + 1} \left(\frac{\alpha^k}{k!} \sigma r_0 + \sum_{i=1}^k \frac{\alpha^{k-i}}{(k-i)!} r_i \right).$$

Donc, $F_k(\mathbf{r}) = r_k$ si et seulement si

$$r_k = \frac{e^{-\alpha}}{(\sigma - 1)r_0 + 1 - e^{-\alpha}} \left(\frac{\alpha^k}{k!} \sigma r_0 + \sum_{i=1}^{k-1} \frac{\alpha^{k-i}}{(k-i)!} r_i \right).$$

Nous obtenons une relation de récurrence. Pour la condition initiale $r_0 = 0$, la solution est $r_k = 0$ pour tout $k \in \{0, \dots, K\}$, et pour la condition initiale $r_0 = \rho_0^*$, la solution est $r_k = \rho_k^*$ pour tout $k \in \{0, \dots, K\}$, comme nous pouvons le voir dans la section 4.1.

Il ne reste qu'à montrer la convergence. Nous ne la montrerons que dans le cas $\sigma e^{-\alpha} > 1, z_0^0 > 0$. Les preuves des cas restants sont similaires ou plus simples. Nous allons faire la preuve par récurrence sur les coordonnées. Puisque la fonction $F_0(r) = \sigma e^{-\alpha} r_0 / ((\sigma - 1)r_0 + 1)$ est croissante et concave en r_0 , la suite $(z^n(0))_{n \geq 0}$ est monotone et converge vers ρ_0^* . Soit $k \in \{1, \dots, K\}$ et supposons que

$$\lim_{n \rightarrow \infty} (z_0^n, \dots, z_{k-1}^n) = (\rho_0^*, \dots, \rho_{k-1}^*).$$

Soit $\varepsilon > 0$. Nous définissons deux fonctions $\underline{F}, \bar{F} : [0, 1] \rightarrow [0, 1]$ en posant, pour $\rho \in [0, 1]$,

$$\underline{F}(\rho) = \frac{e^{-\alpha}}{(\sigma - 1)(\rho_0^* + \varepsilon) + 1} \left(\frac{a^k}{k!} \sigma (\rho_0^* - \varepsilon) + \sum_{i=1}^{k-1} \frac{a^{k-i}}{(k-i)!} (\rho_i^* - \varepsilon) + \rho \right),$$

$$\bar{F}(\rho) = \frac{e^{-\alpha}}{(\sigma - 1)(\rho_0^* - \varepsilon) + 1} \left(\frac{a^k}{k!} \sigma (\rho_0^* + \varepsilon) + \sum_{i=1}^{k-1} \frac{a^{k-i}}{(k-i)!} (\rho_i^* + \varepsilon) + \rho \right).$$

Par l'hypothèse de récurrence, il existe $N \in \mathbb{N}$ tel que pour tout $n \geq N$ et $i \in \{0, \dots, k-1\}$, $|\rho_i^n - \rho_i^*| < \varepsilon$. Nous avons donc, pour tout $n \geq N$ et pour tout $\rho \in [0, 1]$,

$$\underline{F}(\rho) \leq F_k(\rho_0^n, \dots, \rho_{k-1}^n, \rho) \leq \bar{F}(\rho).$$

Nous définissons deux suites, $(\underline{z}^n)_{n \geq N}$ et $(\bar{z}^n)_{n \geq N}$, en posant $\underline{z}^N = \bar{z}^N = z_k^N$ et pour $n > N$

$$\underline{z}^n = \underline{F}(\underline{z}^{n-1}), \quad \bar{z}^n = \bar{F}(\bar{z}^{n-1}).$$

Ainsi, pour tout $n \geq N$, nous avons $\underline{z}^n \leq z_k^n \leq \bar{z}^n$. Puisque $\underline{F}(\rho)$ et $\bar{F}(\rho)$ sont des polynômes de premier degré, et que pour ε assez petit leurs coefficients principaux sont plus petits que 1, les séquences $(\underline{z}^n)_{n \geq N}$ et $(\bar{z}^n)_{n \geq N}$ convergent vers les points fixes des fonctions \underline{F} et \bar{F} , qui sont donnés par :

$$\underline{\rho}_k^* = \frac{e^{-\alpha}}{(\sigma - 1)(\rho_0^* + \varepsilon) + 1 - e^{-\alpha}} \left(\frac{a^k}{k!} \sigma (\rho_0^* - \varepsilon) + \sum_{i=1}^{k-1} \frac{a^{k-i}}{(k-i)!} (\rho_i^* - \varepsilon) \right),$$

$$\bar{\rho}_k^* = \frac{e^{-\alpha}}{(\sigma - 1)(\rho_0^* - \varepsilon) + 1 - e^{-\alpha}} \left(\frac{a^k}{k!} \sigma (\rho_0^* + \varepsilon) + \sum_{i=1}^{k-1} \frac{a^{k-i}}{(k-i)!} (\rho_i^* + \varepsilon) \right).$$

Nous faisons tendre ε vers 0 et nous voyons que

$$\lim_{n \rightarrow \infty} z_k^n = \frac{e^{-\alpha}}{(\sigma - 1)\rho_0^* + 1 - e^{-\alpha}} \left(\frac{a^k}{k!} \sigma \rho_0^* + \sum_{i=1}^{k-1} \frac{a^{k-i}}{(k-i)!} \rho_i^* \right) = \rho_k^*,$$

ce qui conclut le pas de récurrence. □

Nous étudions ensuite la fonction de taux V , que nous définissons en posant, pour $s, t \in \mathcal{D}_K$,

$$V(s, t) = \inf_{l \geq 1} V_l(s, t) = \inf_{l \geq 1} \inf \left\{ \sum_{k=0}^{l-1} I_K(\rho_k, \xi_k, \beta_k, \rho_{k+1}) : \right. \\ \left. \rho_0 = s, \rho_l = t, \rho_k, \xi_k \in \mathcal{D}_K, \beta_k \in \mathcal{B}(t_k) \text{ pour } 0 \leq k \leq l \right\}.$$

Lemme 11.1.5. *Supposons que $\sigma e^{-a} > 1$. Pour $s, t \in \mathcal{D}_K$, nous avons $V(s, t) = 0$ si et seulement si*

- *il existe $l \geq 1$ tel que $t = F^l(s)$,*
- *ou $s(0) = 0$ et $t = (0, \dots, 0)$,*
- *ou $s(0) > 0$ et $t = \rho^*$.*

Démonstration. Soient s, t deux éléments de l'ensemble \mathcal{D}_K tels que $V(s, t) = 0$. Pour chaque $n \geq 1$, soit $(\rho_0^n, \xi_0^n, \beta_0^n, \dots, \rho_{l(n)}^n)$ une suite de longueur $l(n)$ telle que $\rho_0^n = s, \rho_{l(n)}^n = t$ et

$$\sum_{i=0}^{l(n)-1} I_K(\rho_i^n, \xi_i^n, \beta_i^n, \rho_{i+1}^n) \leq \frac{1}{n}.$$

Si la suite $(l(n))_{n \geq 1}$ est bornée, nous pouvons extraire une sous-suite

$$(\rho_0^{\phi(n)}, \xi_0^{\phi(n)}, \beta_0^{\phi(n)}, \dots, \rho_{l(\phi(n))}^{\phi(n)})$$

telle que $l(\phi(n)) = l$ ne dépend pas de n , et pour tout $k \in \{0, \dots, l-1\}$, les limites suivantes existent :

$$\lim_{n \rightarrow \infty} \rho_k^{\phi(n)} = \rho_k, \quad \lim_{n \rightarrow \infty} \xi_k^{\phi(n)} = \xi_k, \quad \lim_{n \rightarrow \infty} \beta_k^{\phi(n)} = \beta_k.$$

L'application I_K étant semi-continue inférieurement,

$$\forall k \in \{0, \dots, l-1\} \quad I_K(\rho_k, \xi_k, \beta_k, \rho_{k+1}) = 0.$$

Il s'ensuit que

$$\forall k \in \{0, \dots, l-1\} \quad \rho_k = F^k(\rho_0).$$

Puisque $\rho_0 = s$ et $\rho_l = t$, nous concluons que $t = F^l(s)$. Supposons ensuite que la suite $(l(n))_{n \geq 0}$ n'est pas bornée. Par le procédé diagonal de Cantor, nous pouvons extraire une sous-suite

$$(\rho_0^{\phi(n)}, \xi_0^{\phi(n)}, \beta_0^{\phi(n)}, \dots, \rho_{l(\phi(n))}^{\phi(n)})$$

telle que, pour tout $k \geq 0$, les limites suivantes existent :

$$\lim_{n \rightarrow \infty} \rho_k^{\phi(n)} = \rho_k, \quad \lim_{n \rightarrow \infty} \xi_k^{\phi(n)} = \xi_k, \quad \lim_{n \rightarrow \infty} \beta_k^{\phi(n)} = \beta_k.$$

L'application I_K est semi-continue inférieurement, nous avons donc

$$\forall k \geq 0 \quad I_K(\rho_k, \xi_k, \beta_k, \rho_{k+1}) = 0.$$

Il s'ensuit que

$$\forall k \geq 0 \quad \rho_k = F^k(\rho_0).$$

Supposons d'abord que $s(0) > 0$, et montrons que alors $t = \rho^*$. Soit $\beta^* \in [0, 1]^{(K+1)^2}$ la matrice triangulaire supérieure donnée par :

$$\forall k \in \{0, \dots, K\} \quad \beta^*(k, \cdot) = \rho^*(k)(0, \dots, 0, M_\infty(k)).$$

Nous avons $I(\rho^*, F(\rho^*), \beta^*, \rho^*) = 0$. Par continuité, il existe un voisinage U de ρ^* tel que

$$\forall \rho \in U \quad V_1(\rho^*, \rho) < I_K(\rho^*, F(\rho^*), \beta^*, \rho) < \varepsilon.$$

Comme $s(0) > 0$, la suite $(F^n(s))_{n \geq 0}$ converge vers ρ^* , et il existe $h \geq 1$ tel que $F^h(s) \in U$. En particulier,

$$\lim_{n \rightarrow \infty} \rho_h^{\phi(n)} = \rho_h = F^h(s) \in U,$$

d'où, pour n assez grand, $\rho_h^{\phi(n)}$ est dans U , ce qui entraîne

$$V(\rho^*, t) \leq V_1(\rho^*, \rho_h^{\phi(n)}) + V(\rho_h^{\phi(n)}, t) < \varepsilon + \frac{1}{n}.$$

Nous faisons $n \rightarrow \infty$ et $\varepsilon \rightarrow 0$ et nous obtenons $V(\rho^*, t) = 0$. Soient $\delta, \gamma > 0$ tels que $\delta < \gamma/\sigma$ et soit

$$V =]\rho^*(0) - \delta, \rho^*(0) + \delta[\times \prod_{k=1}^K]\rho^*(k) - \gamma, \rho^*(k) + \gamma[.$$

Soit α l'infimum

$$\alpha = \inf \{ I_K(\rho_0, \xi_0, \beta_0, \rho_1) : \rho_0 \in \bar{V}, \xi_0 \in \mathcal{D}_K, \beta_0 \in [0, 1]^{(K+1)^2}, \rho_1 \notin V \}.$$

L'ensemble $\bar{V} \times \mathcal{D}_K \times [0, 1]^{(K+1)^2} \times (\mathcal{D}_K \setminus V)$ est compact; par semi-continuité inférieure, il existe $(\rho_0^*, \xi_0^*, \beta_0^*, \rho_1^*) \in \bar{V} \times \mathcal{D}_K \times [0, 1]^{(K+1)^2} \times (\mathcal{D}_K \setminus V)$ tel que

$$\alpha = I_K(\rho_0^*, \xi_0^*, \beta_0^*, \rho_1^*).$$

Soit $\rho \in \bar{V}$ et montrons que $F(\rho) \in V$. La fonction F_0 est croissante et continue, donc

$$F_0(\rho) \in [F_0(\rho^*(0) - \delta), F_0(\rho^*(0) + \delta)].$$

De plus, nous avons

$$\rho^*(0) - \delta < F(\rho^*(0) - \delta) \leq F(\rho^*(0) + \delta) < \rho^*(0) + \delta.$$

Donc, $F_0(\rho) \in]\rho^*(0) - \delta, \rho^*(0) + \delta[$. Soit $k \in \{1, \dots, K\}$, nous avons

$$\begin{aligned} \frac{e^{-\alpha}}{(\sigma-1)(\rho^*(0) + \delta) + 1} \left(\frac{\alpha^k}{k!} \sigma(\rho^*(0) - \delta) + \sum_{i=1}^k \frac{\alpha^{k-i}}{(k-i)!} (\rho^*(i) - \gamma) \right) \\ < F_k(\rho) < \\ \frac{e^{-\alpha}}{(\sigma-1)(\rho^*(0) - \delta) + 1} \left(\frac{\alpha^k}{k!} \sigma(\rho^*(0) + \delta) + \sum_{i=1}^k \frac{\alpha^{k-i}}{(k-i)!} (\rho^*(i) + \gamma) \right). \end{aligned}$$

Puisque $\rho^*(k)$ vérifie la relation de récurrence

$$\rho^*(k) = \frac{e^{-\alpha}}{(\sigma-1)\rho^*(0) + 1} \left(\frac{\alpha^k}{k!} \sigma \rho^*(0) + \sum_{i=1}^k \frac{\alpha^{k-i}}{(k-i)!} \rho^*(i) \right),$$

et que $\delta < \gamma/\sigma$, les inégalités précédentes se mettent sous la forme

$$\begin{aligned} \frac{(\sigma-1)\rho^*(0) + 1}{(\sigma-1)(\rho^*(0) + \delta) + 1} \rho^*(k) - \frac{\gamma}{(\sigma-1)(\rho^*(0) + \delta) + 1} \\ \leq F_k(\rho) \leq \\ \frac{(\sigma-1)\rho^*(0) + 1}{(\sigma-1)(\rho^*(0) - \delta) + 1} \rho^*(k) + \frac{\gamma}{(\sigma-1)(\rho^*(0) - \delta) + 1}. \end{aligned}$$

Ce qui entraîne que, pour δ assez petit,

$$\rho^*(k) - \gamma < F_k(\rho) < \rho^*(k) + \gamma.$$

Enfin, $F(\bar{V}) \subset V$. Nous concluons que $\rho_1^* \neq F(\rho^*(0))$ et $\alpha > 0$. C'est-à-dire que toute suite $(\rho_0, \xi_0, \beta_0, \dots, \rho_l)$ telle que

$$\rho_0 \in V \quad \text{et} \quad \sum_{k=0}^l I_K(\rho_k, \xi_k, \beta_k, \rho_{k+1}) < \alpha$$

reste dans V . En particulier, tout point t satisfaisant $V(\rho^*, t) = 0$ appartient à V . Nous faisons $\delta, \gamma \rightarrow 0$ et nous obtenons que $t = \rho^*$.

Supposons finalement que $s(0) = 0$ et montrons que $t = (0, \dots, 0)$. Si $\rho \in \mathcal{D}_K$ est tel que $\rho(0) > 0$, pour tout $\xi \in \mathcal{D}_K$ et $\beta \in [0, 1]^{(K+1)^2}$, nous avons $I_K(s, \xi, \beta, \rho) = +\infty$. Donc, toute suite $(\rho_0, \xi_0, \beta_0, \dots, \rho_l)$ vérifiant $\rho_0 = s$ et

$$\sum_{k=0}^{l-1} I_K(\rho_k, \xi_k, \beta_k, \rho_{k+1}) < +\infty$$

doit aussi vérifier $\rho_k(0) = 0$ pour tout $k \in \{0, \dots, l\}$. Nous concluons que, nécessairement, $t(0) = 0$. Il nous reste à montrer que $t(1) = \dots = t(k) = 0$. Soit $\varepsilon > 0$. Nous avons $I_K(0, 0, 0, 0) = 0$. Par continuité, il existe un voisinage U de $(0, \dots, 0)$ tel que

$$\forall \rho \in U \quad V_1(0, \rho) < I_K(0, 0, 0, \rho) < \varepsilon.$$

Comme $s(0) = 0$, d'après la proposition précédente, la suite $(F^n(s))_{n \geq 0}$ tend vers $(0, \dots, 0)$, et il existe $h \geq 1$ tel que $F^h(s) \in U$. En particulier,

$$\lim_{n \rightarrow \infty} \rho_h^{\phi(n)} = \rho_h = F^h(s) \in U,$$

d'où, pour n assez grand, $\rho_h^{\phi(n)}$ est dans U , ce qui entraîne

$$V(0, t) \leq V_1(0, \rho_h^{\phi(n)}) + V(\rho_h^{\phi(n)}, t) \leq \varepsilon + \frac{1}{n}.$$

Nous faisons tendre n vers l'infini et ε vers 0 et nous obtenons $V(0, t) = 0$. Le résultat $t = 0$ découle d'un argument similaire à celui qui nous a permis de montrer que $t = \rho^*$ dans le cas $s(0) > 0$. \square

11.2 Concentration autour de ρ^*

Dans cette section, nous montrons que, quand $\sigma e^{-\alpha} > 1$, asymptotiquement, la chaîne de Markov $(Z_n^\theta)_{n \geq 0}$ se concentre autour du point ρ^* . Expliquons d'abord

l'idée de la preuve. La chaîne de Markov $(Z_n^\theta)_{n \geq 0}$ est une perturbation du système dynamique associé à l'application F . L'application F a deux points fixes, $(0, \dots, 0)$ et ρ^* . Le bassin d'attraction de $(0, \dots, 0)$ est l'ensemble $D = \{r \in \mathcal{D}_K : r(0) = 0\}$, et le bassin d'attraction de ρ^* est l'ensemble $\mathcal{D}_K \setminus D$. D'abord nous estimons le temps de sortie d'un voisinage de l'ensemble D ; nous comparons la chaîne de Markov $(Z_n^\theta)_{n \geq 0}$ avec les chaînes de Markov dans [11] et nous utilisons les estimations développées dans [11] pour conclure que le processus $(Z_n^\theta)_{n \geq 0}$ quitte le voisinage de l'ensemble D dans un nombre logarithmique de pas avec une probabilité raisonnable. Puis nous estimons le temps que le processus $(Z_n^\theta)_{n \geq 0}$ passe en dehors des voisinages de D et ρ^* ; le processus tend à suivre les trajectoires discrètes données par le système dynamique, nous verrons donc que, avec très grande probabilité, en un nombre fini de pas, le processus rentre dans un voisinage de D ou de ρ^* . Ces estimations, avec le principe de grandes déviations pour la matrice de transition de $(Z_n^\theta)_{n \geq 0}$ que nous avons démontré dans la section précédente, suffiront pour montrer que la mesure de probabilité invariante du processus $(Z_n^\theta)_{n \geq 0}$ converge vers la masse de Dirac en ρ^* .

Pour simplifier la notation, nous enlevons l'indice θ dans la suite, et nous le mettons là où il est absolument nécessaire. De plus, nous notons par P_z et E_z les probabilités et espérances lorsque l'état initial est $z \in \mathbb{D}^{K+1}$. Tout au long de cette section, nous supposons que $\sigma e^{-\alpha} > 1$.

Définissons

$$D_\delta = \{r \in \mathcal{D}_K : 0 < r(0) < \delta\}.$$

Lemme 11.2.1. *Pour tout $\delta > 0$, il existe $c > 0$, dépendant de δ , tel que, asymptotiquement, pour tout $z_0 \in \mathbb{D}^{K+1}$ tel que $z_0(0) \geq 1$, nous avons*

$$P_{z_0}(Z_1(0) > 0, \dots, Z_{\lfloor c \ln m \rfloor - 1}(0) > 0, Z_{\lfloor c \ln m \rfloor} \in m(\mathcal{D}_K \setminus \overline{D_\delta})) \geq \frac{1}{m^{c \ln m}}.$$

Démonstration. Soit $(\tilde{Z}_t^\ell)_{t \geq 0}$ la chaîne de Markov inférieure dans [11]. Soit τ le premier temps où le processus $(\tilde{Z}_t^\ell)_{t \geq 0}$ s'annule :

$$\tau = \inf \{t \geq 0 : \tilde{Z}_t = 0\}.$$

D'après les remarques dans la section 10.4, nous avons

$$\begin{aligned} & P_{z_0}(Z_1(0) > 0, \dots, Z_{\lfloor c \ln m \rfloor - 1}(0) > 0, Z_{\lfloor c \ln m \rfloor} \in m(\mathcal{D}_K \setminus \overline{D_\delta})) = \\ & P_{z_0}(Z_1(0) > 0, \dots, Z_{\lfloor c \ln m \rfloor - 1}(0) > 0, Z_{\lfloor c \ln m \rfloor} \in m(\mathcal{D}_K \setminus \overline{D_\delta}), \tau > \lfloor c \ln m \rfloor) \\ & \geq P_1(\tilde{Z}_1 > 0, \dots, \tilde{Z}_{\lfloor c \ln m \rfloor - 1} > 0, \tilde{Z}_{\lfloor c \ln m \rfloor} > m(\rho^*(0) - \delta)). \end{aligned}$$

Grâce au lemme 7.7 dans [11], cette dernière probabilité est bornée inférieurement par $1/m^{c \ln m}$, ce qui donne le résultat voulu. \square

Ensuite nous estimons la longueur d'une excursion du processus $(Z_n^\theta)_{n \geq 0}$ en dehors des voisinages de D et ρ^* . Pour $\rho \in \mathcal{D}_K$ et $\delta > 0$, nous définissons le δ -voisinage de ρ par

$$U(\rho, \delta) = \{r \in \mathcal{D}_K : |r - \rho| < \delta\}.$$

Lemme 11.2.2. *Pour tout $\delta > 0$, il existe $h \geq 1$ et $c > 0$, dépendant de δ , tels que, asymptotiquement, pour tout $r \in \mathcal{D}_K$ avec $r_0 \geq \delta$, nous avons*

$$P_{\lfloor r_m \rfloor}(Z_1(0) > 0, \dots, Z_{h-1}(0) > 0, Z_h \in mU(\rho^*, \delta)) \geq 1 - \exp(-cm).$$

Démonstration. Soit $\delta > 0$ et définissons l'ensemble \mathcal{K} par

$$\mathcal{K} = \{r \in \mathcal{D} : r_0 \geq \delta\}.$$

Pour tout $r \in \mathcal{K}$ il existe un entier positif $h_r \geq 0$ tel que $F^{h_r}(r) \in U(\rho^*, \delta/4)$. Par continuité de l'application F , pour tout $r \in \mathcal{K}$, il existe aussi des réels positifs $\delta_0^r, \dots, \delta_{h_r}^r$ tels que $\delta_0^r, \dots, \delta_{h_r}^r < \delta/2$ et

$$\forall k \in \{0, \dots, h_r\} \quad F(U(F^{k-1}(r), \delta_{k-1}^r)) \subset U(F^k(r), \delta_k^r/2).$$

La famille $\{U(r, \delta_0^r) : r \in \mathcal{K}\}$ forme un recouvrement par des ouverts de l'ensemble \mathcal{K} ; puisque \mathcal{K} est un ensemble compact, nous pouvons extraire un sous-recouvrement fini, i.e., il existe $N \in \mathbb{N}$ et $r_1, \dots, r_N \in \mathcal{K}$ tels que

$$\mathcal{K} \subset U_0 = \bigcup_{n=1}^N U(r_n, \delta_0^{r_n}).$$

Soit $h = \max\{h_{r_i} : 1 \leq i \leq N\}$. Pour $n \in \{1, \dots, N\}$, nous prenons des réels positifs $\delta_{h_{r_n}+1}^{r_n}, \dots, \delta_h^{r_n}$ vérifiant, comme avant,

$$\forall k \in \{h_{r_n} + 1, \dots, h\} \quad F(U(F^{k-1}(r_n), \delta_{k-1}^{r_n})) \subset U(F^k(r_n), \delta_k^{r_n}/2).$$

Définissons

$$\forall k \in \{1, \dots, h-1\} \quad \mathcal{U}_k = \bigcup_{n=1}^N \mathcal{U}(F^k(r_n), \delta_k^{r_n}).$$

Nous avons, pour tout $r \in \mathcal{K}$,

$$P_{\lfloor mr \rfloor}(Z_1(0) > 0, \dots, Z_{h-1}(0) > 0, Z_h \in m\mathcal{U}(\rho^*, \delta)) \geq P_{\lfloor mr \rfloor}(\forall k \in \{1, \dots, h\} \quad Z_k \in m\mathcal{U}_k).$$

Passons au complémentaire :

$$\begin{aligned} P_{\lfloor mr \rfloor}(\exists k \in \{1, \dots, h-1\} \quad Z_k(0) = 0 \text{ or } Z_h \notin m\mathcal{U}(\rho^*, \delta)) \\ \leq P_{\lfloor mr \rfloor}(\exists k \in \{1, \dots, h\} \quad Z_k \notin m\mathcal{U}_k) \\ \leq \sum_{1 \leq k \leq h} P_{\lfloor mr \rfloor}(Z_1 \in m\mathcal{U}_1, \dots, Z_{k-1} \in m\mathcal{U}_{k-1}, Z_k \notin m\mathcal{U}_k) \\ \leq \sum_{1 \leq k \leq h} \sum_{z \in m\mathcal{U}_{k-1}} P_{\lfloor mr \rfloor}(Z_{k-1} = z, Z_k \notin m\mathcal{U}_k) \\ \leq \sum_{1 \leq k \leq h} \sum_{z \in m\mathcal{U}_{k-1}} P_z(Z_k \notin m\mathcal{U}_k) P_{\lfloor mr \rfloor}(Z_{k-1} = z) \\ \leq \sum_{1 \leq k \leq h} \max \{ P_z(Z_1 \notin m\mathcal{U}_k) : z \in m\mathcal{U}_{k-1} \}. \end{aligned}$$

Le principe de grandes déviations de la proposition 11.1.2 entraîne que, pour tout $k \in \{1, \dots, h\}$,

$$\begin{aligned} \limsup_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha}} \frac{1}{m} \ln \max_{z \in m\mathcal{U}_{k-1}} P_z(Z_1 \notin m\mathcal{U}_k) \leq \\ - \inf \{ I(r, \xi, \beta, t) : r \in \overline{\mathcal{U}_{k-1}}, \xi \in \mathcal{D}, \beta \in \mathcal{B}(t), t \notin \mathcal{U}_k \}. \end{aligned}$$

Pour tout $r \in \overline{\mathcal{U}_{k-1}}$, nous avons $F(r) \in \mathcal{U}_k$, l'infimum précédent est donc strictement positif. Puisque h est fixé, nous concluons que

$$\limsup_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha}} \frac{1}{m} \ln P_{\lfloor mr \rfloor}(\exists k \in \{1, \dots, h-1\} \quad Z_k = 0 \text{ or } Z_h \notin m\mathcal{U}(\rho^*, \delta)) < 0,$$

ce qui entraîne le résultat voulu. □

Corollaire 11.2.3. Soit $\delta > 0$. Il existe $h \geq 1$, $c \geq 0$, dépendants de δ , tels que, asymptotiquement, pour tout $r \in \mathcal{D}_K \setminus (\overline{D}_\delta \cup \mathcal{U}(\rho^*, \delta))$ et pour tout $n \geq 0$, nous avons

$$P_{\lfloor rm \rfloor}(Z_t \in \mathcal{D}_K \setminus (\overline{D}_\delta \cup \mathcal{U}(\rho^*, \delta)) \text{ pour } 0 \leq t \leq n) \leq \exp\left(-cm \left\lfloor \frac{n}{h} \right\rfloor\right).$$

Démonstration. Soit $\delta > 0$. Soit $C = \mathcal{D}_K \setminus (\overline{D}_\delta \cup \mathcal{U}(\rho^*, \delta))$. Soit $r \in C$, et soient $h \geq 1$ et $c > 0$ comme dans le lemme précédent. Nous divisons l'intervalle $\{0, \dots, n\}$ en sous-intervalles de longueur h , et nous utilisons l'estimée du lemme 11.2.2 sur chaque sous-intervalle. Soit $i \geq 0$, grâce à la propriété de Markov,

$$\begin{aligned} P_{\lfloor rm \rfloor}(Z_t \in C \text{ pour } 0 \leq t \leq (i+1)h) &= \sum_{z \in C} P_{\lfloor rm \rfloor}(Z_t \in C \text{ pour } 0 \leq t \leq (i+1)h, Z_{ih} = z) \\ &= \sum_{z \in C} P_{\lfloor rm \rfloor}(Z_t \in C \text{ pour } 0 \leq t \leq ih, Z_{ih} = z) \\ &\quad \times P_z(Z_t \in C \text{ pour } 0 \leq t \leq h) \\ &\leq \sum_{z \in C} P_{\lfloor rm \rfloor}(Z_t \in C \text{ pour } 0 \leq t \leq ih, Z_{ih} = z) \\ &\quad \times P_z(Z_1(0) > 0, \dots, Z_{h-1}(0) > 0, Z_h \notin m\mathcal{U}(\rho^*, \delta)) \\ &\leq P_{\lfloor rm \rfloor}(Z_t \in C \text{ pour } 0 \leq t \leq ih) \exp(-cm). \end{aligned}$$

Nous itérons cette inégalité et nous obtenons :

$$\forall i \geq 0 \quad P_{\lfloor rm \rfloor}(Z_t \in C \text{ pour } 0 \leq t \leq ih) \leq \exp(-cmi).$$

Nous finissons en prenant $i = \lfloor n/h \rfloor$. □

Soit $g : [0, 1] \rightarrow [0, 1]$ une fonction continue, croissante et nulle en 0.

Proposition 11.2.4. Pour tout $z^0 \in \mathbb{D}^{K+1}$ avec $z_0^0 \geq 1$, nous avons

$$\lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha}} \frac{\mathbb{E}\left(\sum_{n=0}^{\tau_0} g\left(\frac{Z_n(0) + \dots + Z_n(K)}{m} \mid Z_0 = z^0\right)\right)}{\mathbb{E}(\tau_0 \mid Z_0 = z^0)} = g(\rho_0^* + \dots + \rho_K^*).$$

Démonstration. Soit $\varepsilon > 0$. La fonction g étant continue, il existe $\delta > 0$ tel que

$$\forall \rho \in \mathcal{U}(\rho^*, 2\delta) \quad |g(\rho_0 + \dots + \rho_K) - g(\rho_0^* + \dots + \rho_K^*)| < \varepsilon.$$

Nous définissons une suite de temps d'arrêt pour contrôler les excursions de la chaîne de Markov $(Z_n)_{n \geq 0}$ en dehors de $U(\rho^*, \delta)$. Prenons $T_0 = 0$ et

$$\begin{aligned} T_1^* &= \inf\{n \geq 0 : \frac{Z_n}{m} \in U(\rho^*, \delta)\} & T_1 &= \inf\{n \geq T_1^* : \frac{Z_n}{m} \notin U(\rho^*, 2\delta)\} \\ &\vdots & &\vdots \\ T_k^* &= \inf\{n \geq T_{k-1}^* : \frac{Z_n}{m} \in U(\rho^*, \delta)\} & T_k &= \inf\{n \geq T_k^* : \frac{Z_n}{m} \notin U(\rho^*, 2\delta)\} \\ &\vdots & &\vdots \end{aligned}$$

Nous avons donc

$$\begin{aligned} \sum_{n=0}^{\tau_0} g\left(\frac{Z_n(0) + \dots + Z_n(K)}{m}\right) - g(\rho_0^* + \dots + \rho_K^*)\tau_0 = \\ \sum_{k \geq 1} \sum_{n=T_{k-1}^* \wedge \tau_0}^{T_k^* \wedge \tau_0 - 1} \left(g\left(\frac{Z_n(0) + \dots + Z_n(K)}{m}\right) - g(\rho_0^* + \dots + \rho_K^*) \right) \\ + \sum_{k \geq 1} \sum_{n=T_k^* \wedge \tau_0}^{T_k \wedge \tau_0 - 1} \left(g\left(\frac{Z_n(0) + \dots + Z_n(K)}{m}\right) - g(\rho_0^* + \dots + \rho_K^*) \right). \end{aligned}$$

En prenant des valeurs absolues,

$$\begin{aligned} \left| \sum_{n=0}^{\tau_0} g\left(\frac{Z_n(0) + \dots + Z_n(K)}{m}\right) - g(\rho_0^* + \dots + \rho_K^*)\tau_0 \right| \leq \\ 2g(1) \sum_{k \geq 1} (T_k^* \wedge \tau_0 - T_{k-1} \wedge \tau_0) + \varepsilon\tau_0. \end{aligned}$$

Il nous reste à contrôler la somme. Définissons, pour $n \geq 0$,

$$N(n) = \max \{ k \geq 1 : T_{k-1} < n \}.$$

Nous réécrivons la somme comme suit :

$$\sum_{k \geq 1} (T_k^* \wedge \tau_0 - T_{k-1} \wedge \tau_0) = \sum_{k=1}^{N(\tau_0)} (T_k^* \wedge \tau_0 - T_{k-1}).$$

Soit $\eta > 0$ et définissons le temps t_m^η par :

$$t_m^\eta = \exp(m(\tilde{V}(\rho^*, 0) + \eta)),$$

où \tilde{V} est la fonction de taux gouvernant la dynamique de la master sequence (cf. [11]), i.e., pour tout $x, y \in [0, 1]$,

$$\tilde{V}(x, y) = \inf_{l \geq 1} \inf \left\{ \sum_{k=0}^{l-1} I_0(\rho_k, \xi_k, \beta_k, \rho_{k+1}) : \right. \\ \left. \rho_0 = s, \rho_l = t, \rho_k, \xi_k \in \mathcal{D}_0, \beta_k \in \mathcal{B}(t_k) \text{ pour } 0 \leq k \leq l \right\}.$$

Nous décomposons la somme précédente ainsi :

$$\sum_{k=1}^{N(\tau_0)} (T_k^* \wedge \tau_0 - T_{k-1}) \leq 1_{\tau_0 > t_m^n} \tau_0 + 1_{\tau_0 \leq t_m^n} \sum_{k=1}^{N(\tau_0)} (T_k^* \wedge \tau_0 - T_{k-1}).$$

Soit $z^0 \in \mathbb{D}^{K+1}$ avec $z_0^0 \geq 1$. Puisque les estimées sont les mêmes pour toute condition initiale, nous n'écrivons plus le point de départ dans la suite : les probabilités et espérances correspondent toutes au point initial z^0 , sauf mention contraire. Nous prenons l'espérance dans les inégalités précédentes et nous obtenons

$$\left| \mathbb{E} \left(\sum_{n=0}^{\tau_0} g \left(\frac{Z_n(0) + \dots + Z_n(K)}{m} \right) \right) - g(\rho_0^* + \dots + \rho_K^*) \mathbb{E}(\tau_0) \right| \leq \\ \varepsilon \mathbb{E}(\tau_0) + 2g(1) \mathbb{E}(1_{\tau_0 > t_m^n} \tau_0) + 2g(1) \mathbb{E} \left(1_{\tau_0 \leq t_m^n} \sum_{k=1}^{N(\tau_0)} (T_k^* \wedge \tau_0 - T_{k-1}) \right).$$

Nous savons, grâce aux estimées de la section 7.3 de [11], que

$$\lim_{m \rightarrow \infty} \mathbb{E}(1_{\tau_0 > t_m^n} \tau_0) = 0.$$

Pour contrôler le dernier terme nous donnons d'abord une borne supérieure pour N .

Lemme 11.2.5. *Il existe $c > 0$, dépendant de δ , tel que, asymptotiquement,*

$$\forall k, n \geq 0 \quad \mathbb{P}(N(n) > k) \leq \frac{n^k}{k!} \exp(-cmk).$$

Démonstration. Pour $k \geq 0$, nous définissons

$$S_k = \sup \{ T_k^* \leq n < T_k : Z_n \in m\mathcal{U}(\rho^*, \delta) \}.$$

Pour $k, n \geq 0$, nous avons :

$$P(N(n) > k) = P(T_k < n) = \sum_{t^* \leq s < t < n} P(T_k^* = t^*, S_k = s, T_k = t).$$

Soient $h \geq 1$ et $c \geq 0$ comme dans le corollaire 11.2.3. Pour t^* et s fixés, nous séparons la somme sur t en deux morceaux :

$$\sum_{t: s < t < n} P(T_k^* = t^*, S_k = s, T_k = t) = \sum_{t: t > s+h+1} \dots + \sum_{t: s < t \leq s+h+1} \dots$$

Nous étudions d'abord la première somme, quand $t > s + h + 1$. Conditionnons par rapport à l'état du processus au temps $s + 1$:

$$\begin{aligned} \sum_{t: t > s+h+1} \dots &= \sum_{\substack{t: t > s+h+1 \\ z \in mU(\rho^*, 2\delta) \setminus mU(\rho^*, \delta)}} P(T_k^* = t^*, S_k = s, Z_{s+1} = z, T_k = t) \\ &\leq \sum_{\substack{t: t > s+h+1 \\ z \in mU(\rho^*, 2\delta) \setminus mU(\rho^*, \delta)}} P\left(\begin{array}{l} T_k^* = t^*, Z_{s+1} = z, Z_t \notin mU(\rho^*, 2\delta) \\ Z_{s+1}, \dots, Z_{t-1} \in mU(\rho^*, 2\delta) \setminus mU(\delta) \end{array} \right) = \\ &\sum_{\substack{t: t > s+h+1 \\ z \in mU(\rho^*, 2\delta) \setminus mU(\rho^*, \delta)}} P_z(Z_{s+1}, \dots, Z_{t-1} \in mU(\rho^*, 2\delta) \setminus mU(\rho^*, \delta), Z_t \notin mU(\rho^*, 2\delta)) \\ &\quad \times P(T_k^* = t^*, Z_{s+1} = z). \end{aligned}$$

Pour $0 \leq s < n$ et $t > s + h + 1$,

$$\begin{aligned} &P_z(Z_{s+1}, \dots, Z_{t-1} \in mU(\rho^*, 2\delta) \setminus mU(\rho^*, \delta), Z_t \notin mU(\rho^*, 2\delta)) \\ &\leq P_z(Z_r \in \mathcal{D}_K \setminus (\overline{D}_\delta \cup U(\rho^*, \delta)) \text{ pour } s+1 \leq r \leq t-1) \\ &\leq \exp\left(-cm \left\lfloor \frac{t-s-2}{h} \right\rfloor\right). \end{aligned}$$

Donc,

$$\sum_{t: t > s+h+1} \dots \leq \left(\sum_{t \geq h} \exp\left(-cm \left\lfloor \frac{t}{h} \right\rfloor\right) \right) P(T_k^* = t^*).$$

Nous contrôlons ensuite la deuxième somme. Conditionnons selon l'état du processus au temps s :

$$\begin{aligned} \sum_{t: s < t \leq s+h+1} \dots &= \sum_{\substack{t: s < t \leq s+h+1 \\ z \in mU(\delta)}} P(T_k^* = t^*, S_k = s, Z_s = z, T_k = t) \\ &\leq \sum_{\substack{t: s < t \leq s+h+1 \\ z \in mU(\delta)}} P_z(Z_t \notin mU(\rho^*, 2\delta)) P(T_k^* = t^*, Z_s = z). \end{aligned}$$

Grâce au principe de grandes déviations du corollaire 11.1.3, pour tout $t \in \{1, \dots, h+1\}$ et pour tout $z \in mU(\rho^*, \delta)$,

$$\begin{aligned} \limsup_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha}} \frac{1}{m} \ln P_z(Z_t \notin mU(\rho^*, 2\delta)) \\ \leq -\inf \{ V_t(\rho, r) : \rho \in \overline{\rho^*, U(\delta)}, r \notin U(\rho^*, 2\delta) \}. \end{aligned}$$

Par compacité, l'infimum est atteint, soient $\bar{\rho}$ et \bar{r} deux éléments pour lesquels l'infimum est atteint. Comme $F(\bar{\rho}) \in U(\rho^*, \delta)$, \bar{r} n'est pas un itéré de ρ par F , et l'infimum précédent est nécessairement positif. Nous concluons qu'il existe $c' > 0$, dépendant de δ , tel que, asymptotiquement,

$$\sum_{t:s < t \leq s+h+1} \dots \leq \exp(-c'm)P(T_k^* = t^*).$$

Prenons $c'' > 0$ tel que, asymptotiquement,

$$\sum_{t>h} \exp\left(-cm \left\lfloor \frac{t}{h} \right\rfloor\right) + \exp(-c'm) \leq \exp(-c''m).$$

Nous avons donc, pour tout $n, k \geq 0$

$$\begin{aligned} P(T_k < n) &\leq \sum_{t^* \leq s < n} \exp(-c''m)P(T_k^* = t^*) \\ &\leq \sum_{s < n} \exp(-c''m)P(T_k^* \leq s) \leq \sum_{s < n} \exp(-c''m)P(T_{k-1} \leq s). \end{aligned}$$

Nous itérons cette inégalité et nous obtenons

$$P(T_k < n) \leq \sum_{0 \leq n_0 < \dots < n_{k-1} < n} \exp(-c''mk) \leq \frac{n^k}{k!} \exp(-c''mk).$$

Ce qui conclut la démonstration. □

Nous estimons ensuite le terme

$$E \left(\mathbf{1}_{\tau_0 \leq t_m^n} \sum_{k=1}^{N(\tau_0)} (T_k^* \wedge \tau_0 - T_{k-1}) \right).$$

Par l'inégalité de Cauchy–Schwarz, nous avons

$$\begin{aligned} \mathbb{E} \left(\mathbf{1}_{\tau_0 \leq t_m^n} \sum_{k=1}^{N(\tau_0)} (T_k^* \wedge \tau_0 - T_{k-1}) \right) &= \sum_{k \geq 1} \mathbb{E}(\mathbf{1}_{\tau_0 \leq t_m^n} \mathbf{1}_{k \leq N(\tau_0)} (T_k^* \wedge \tau_0 - T_{k-1})) \\ &\leq \sum_{k \geq 0} \mathbb{P}(\tau_0 \leq t_m^n, N(\tau_0) \geq k)^{1/2} \mathbb{E}(\mathbf{1}_{k \leq N(\tau_0)} (T_k^* \wedge \tau_0 - T_{k-1})^2)^{1/2} \\ &\leq \sum_{k \geq 1} \mathbb{P}(N(t_m^n) \geq k)^{1/2} \mathbb{E}(\mathbf{1}_{k \leq N(\tau_0)} (T_k^* \wedge \tau_0 - T_{k-1})^2)^{1/2}. \end{aligned}$$

Si $1 \leq k \leq N(\tau_0)$, alors $T_{k-1} < \tau_0$ et $Z_{T_{k-1}}(0) > 0$. Grâce à la propriété de Markov,

$$\begin{aligned} &\mathbb{E}(\mathbf{1}_{k \leq N(\tau_0)} (T_k^* \wedge \tau_0 - T_{k-1})^2) \\ &= \sum_{z \in \mathbb{D}^{K+1}; z_0 \geq 1} \mathbb{E}(\mathbf{1}_{k \leq N(\tau_0)} (T_k^* \wedge \tau_0 - T_{k-1})^2 \mid Z_{T_{k-1}} = z) \times \mathbb{P}(Z_{T_{k-1}} = z) \\ &\leq \sum_{z \in \mathbb{D}^{K+1}; z_0 \geq 1} \mathbb{E}_z((T_1^* \wedge \tau_0)^2) \mathbb{P}(Z_{T_{k-1}} = z). \end{aligned}$$

Maintenant nous cherchons une borne supérieure pour le temps $T_1^* \wedge \tau_0$, si la chaîne est issue de $z \in \mathbb{D}^{K+1}$ avec $z_0 \geq 1$.

Lemme 11.2.6. *Supposons que $\sigma e^{-\alpha} > 1$. Pour tout $\delta > 0$, il existe $h \geq 1$, $c > 0$, dépendants de δ , tels que, asymptotiquement, pour $z \in \mathbb{D}^{K+1}$ avec $z_0 \geq 1$,*

$$\mathbb{P}_z(Z_{\lfloor c \ln m \rfloor + h} \in m\mathcal{U}(\rho^*, \delta)) \geq \frac{1/2}{m^{c \ln m}}.$$

Démonstration. Grâce au lemme 11.2.1, il existe $c > 0$, tel que, asymptotiquement, pour tout $z \in \mathbb{D}^{K+1}$ avec $z_0 \geq 1$,

$$\mathbb{P}_z(Z_{\lfloor c \ln m \rfloor}(0) > \delta m) \geq \frac{1}{m^{c \ln m}}.$$

De même, grâce au lemme 11.2.2, il existe $h \geq 1$ et $c' > 0$, tels que, asymptotiquement, pour tout $z' \in \mathbb{D}^{K+1}$ avec $z'_0 \geq \delta m$,

$$\mathbb{P}_{z'}(Z_h \in m\mathcal{U}(\rho^*, \delta)) \geq 1 - \exp(-c'm).$$

Nous avons donc

$$\begin{aligned} &\mathbb{P}_z(Z_{\lfloor c \ln m \rfloor + h} \in m\mathcal{U}(\rho^*, \delta)) \\ &\geq \mathbb{P}_z(Z_{\lfloor c \ln m \rfloor}(0) \geq \delta m, Z_{\lfloor c \ln m \rfloor + h} \in m\mathcal{U}(\rho^*, \delta)) = \\ &\sum_{z': z'_0 \geq \delta m} \mathbb{P}_{z'}(Z_h \in m\mathcal{U}(\rho^*, \delta)) \mathbb{P}_z(Z_{\lfloor c \ln m \rfloor} = z') \geq (1 - \exp(-c'm)) \frac{1}{m^{c \ln m}}, \end{aligned}$$

ce qui entraîne le résultat voulu. \square

Corollaire 11.2.7. *Pour tout $\delta > 0$, il existe $h \geq 1$, $c > 0$, dépendants de δ , tels que, asymptotiquement, pour tout $z \in \mathbb{D}^{K+1}$ avec $z_0 \geq 1$,*

$$\forall n \geq 0 \quad P_z(T_1^* \wedge \tau_0 \geq n(\lfloor c \ln m \rfloor + h)) \leq \left(1 - \frac{1/2}{m^{c \ln m}}\right)^n.$$

La preuve est similaire à celle du corollaire 11.2.3. Grâce au corollaire précédent, asymptotiquement, pour tout $z \in \mathbb{D}^{K+1}$ avec $z_0 \geq 1$,

$$E_z((T_1^* \wedge \tau_0)^2) = \sum_{k \geq 1} P_z(T_1^* \wedge \tau_0 \geq \sqrt{k}) \leq \sum_{k \geq 1} \left(1 - \frac{1/2}{m^{c \ln m}}\right)^{\lfloor \frac{\sqrt{k}}{\lfloor c \ln m \rfloor + k} \rfloor}.$$

Posons

$$\alpha = 1 - \frac{1/2}{m^{c \ln m}}, \quad t = \lfloor c \ln m \rfloor + h.$$

Nous avons :

$$\begin{aligned} \sum_{k \geq 1} \alpha^{\lfloor \sqrt{k}/t \rfloor} &\leq \sum_{k \geq 1} \alpha^{\sqrt{k}/t-1} \leq \int_0^\infty \alpha^{\sqrt{x}/t-1} dx \\ &= \frac{t^2}{\alpha(\ln \alpha)^2} \int_0^\infty e^{-\sqrt{y}} dy = \frac{2t^2}{\alpha(\ln \alpha)^2}. \end{aligned}$$

Donc, asymptotiquement, pour tout $z \in \mathbb{D}^{K+1}$ avec $z_0 \geq 1$,

$$E_z((T_1^* \wedge \tau_0)^2) \leq m^{3c \ln m}.$$

Alors, pour tout $k \geq 1$,

$$E(1_{k \leq N(\tau_0)}(T_k^* \wedge \tau_0 - T_{k-1})^2) \leq m^{3c \ln m}.$$

Nous en déduisons l'inégalité suivante :

$$\begin{aligned} E\left(1_{\tau_0 \leq t_m^n} \sum_{k=0}^{N(\tau_0)} (T_k^* \wedge \tau_0 - T_{k-1})\right) &\leq \sum_{k \geq 0} (m^{3c \ln m})^{1/2} P(N(t_m^n) > k)^{1/2} \\ &\leq m^{3c \ln m} \left(t_m^n \exp(-cm/3) + \sum_{k \geq t_m^n \exp(-cm/3)} \left(\frac{(t_m^n)^k}{k!} \exp(-cmk) \right)^{1/2} \right) \\ &\leq m^{3c \ln m} \left(t_m^n \exp(-cm/3) + \sum_{k \geq 0} \exp\left(\frac{k}{2} - cm\frac{k}{3}\right) \right). \end{aligned}$$

La dernière inégalité a été obtenue grâce au fait que $k! > (k/e)^k$. Nous choisissons η tel que $0 < \eta < c/3$. D'après l'inégalité précédente :

$$\limsup_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha}} \frac{1}{m} \ln \mathbb{E} \left(1_{\tau_0 \leq t_m^\eta} \sum_{k=1}^{N(\tau_0)} (\tau_k^* \wedge \tau_0 - \tau_{k-1}) \right) \leq \max \left(\tilde{V}(\rho^*(0), 0) + \eta - \frac{c}{3} \right) < \tilde{V}(\rho^*(0), 0),$$

où \tilde{V} est la fonction de taux gouvernant la dynamique de la master sequence. Ceci entraîne le résultat voulu. \square

Chapitre 12

Fin de la démonstration

La première affirmation dans le théorème 9.2.1 a été démontrée dans [11] pour le cas de la master sequence, $K = 0$. La démonstration pour le cas $K \geq 1$ ne requiert pas de nouvelles idées ni ne nous aide à mieux comprendre le modèle ; c'est une généralisation directe du cas $K = 0$. Nous nous concentrons donc sur la deuxième affirmation du théorème 9.2.1. Supposons que $\alpha\psi(a) > \ln \kappa$. Comme nous pouvons voir dans [11], et grâce aux résultats de comparaison de la section 10.4, nous avons les estimations suivantes : $\forall a \in]0, +\infty[, \forall \alpha \in [0, +\infty[$,

$$\lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \frac{1}{m} \ln E(\tau_0 | Z_0^\theta = z_{\text{enter}}^\theta) = \tilde{V}(\rho^*, 0),$$
$$\limsup_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \frac{1}{\ell} \ln E(\tau^* | O_0^\theta = o_{\text{exit}}^\theta) \leq \ln \kappa.$$

Donc, puisque nous étudions le cas $\alpha\psi(a) > \ln \kappa$,

$$\lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \frac{E(\tau_0 | Z_0^\theta = z_{\text{enter}}^\theta)}{E(\tau^* | O_0^\theta = o_{\text{exit}}^\theta)} = +\infty.$$

D'une part, vu que g est une fonction bornée, nous avons que

$$\lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \frac{E\left(\sum_{n=0}^{\tau^*-1} g\left(\frac{|\pi(O_n^\theta)|_1}{m}\right) \middle| O_0^\theta = o_{\text{exit}}^\theta\right)}{E(\tau^* | O_0^\theta = o_{\text{exit}}^\theta) + E(\tau_0 | Z_0^\theta = z_{\text{enter}}^\theta)} = 0.$$

D'autre part, d'après la proposition 11.2.4, nous avons aussi

$$\lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha, \frac{m}{\ell} \rightarrow \alpha}} \frac{\mathbb{E} \left(\sum_{n=0}^{\tau_0-1} g \left(\frac{|Z_n^\theta|_1}{m} \right) \middle| Z_0^\theta = z_{\text{enter}}^\theta \right)}{\mathbb{E}(\tau^* | O_0^\theta = o_{\text{exit}}^\theta) + \mathbb{E}(\tau_0 | Z_0^\theta = z_{\text{enter}}^\theta)} = g(\rho_0^* + \dots + \rho_K^*).$$

Nous revenons à la formule de la fin de la section 10.3, et nous voyons que ces deux limites suffisent à achever la démonstration du théorème 9.2.1.

Chapitre 13

Conclusion

Nous avons montré que dans le cas du paysage à un pic, le modèle de Wright–Fisher présente un phénomène de seuil d’erreur et une distribution de quasi–espèce. De plus, la distribution de la quasi–espèce pour le modèle de Wright–Fisher est la même que pour le modèle d’Eigen, $\mathcal{Q}(\sigma, \alpha)$.

L’étude du processus de Wright–Fisher met en évidence l’impact de la taille de la population dans la quasi–espèce. Nous voyons que la taille de la population joue un rôle important dans l’émergence d’une quasi–espèce. En effet, le modèle d’Eigen prédit la création d’une quasi–espèce dès que la condition $\sigma e^{-\alpha} > 1$ est satisfaite. Dans le modèle de Wright–Fisher, la condition $\sigma e^{-\alpha} > 1$ est nécessaire mais n’est pas suffisante pour la création d’une quasi–espèce, il faut en plus que la taille de la population soit assez grande. Cette condition supplémentaire vient du fait que, dans le processus de Wright–Fisher, l’avantage sélectif de la master sequence doit faire face à la dérive génétique, une dimension qui était totalement absente dans le modèle d’Eigen.

Quand la taille de la population est suffisamment grande, et qu’une quasi–espèce se crée, la distribution de cette quasi–espèce est la même que dans le modèle d’Eigen. Ceci montre que les prédictions du modèle d’Eigen sont très robustes, et que les phénomènes de seuil d’erreur et de quasi–espèce restent valables pour des populations finies, tout en faisant attention à ce que la taille de la population ne soit pas trop petite.

Le modèle de Wright–Fisher est un modèle dans lequel les générations ne se su-

perposent pas. Cela est parfois utile pour modéliser par exemple des individus avec des cycles de vie très précis. Cependant, nous nous intéressons souvent à des situations où les différentes générations cohabitent ensemble pendant longtemps, et où les reproductions arrivent de manière plus aléatoire. L'étude de cette différence qualitative dans la reproduction, et sa relation avec le phénomène de seuil d'erreur et de quasi-espèce sera l'objet de la partie suivante.

Quatrième partie

Le modèle de Moran

Chapitre 14

Introduction

En 1958, Patrick Moran modifie le processus de Wright–Fisher afin de permettre des générations qui se superposent [57]. Tout comme dans le modèle original de Fisher et Wright, le modèle de Moran considère une population de m gènes avec deux allèles possibles, a et A . Par contre, cette fois-ci, à chaque pas de temps un seul individu est sélectionné pour se reproduire, et son enfant remplace un individu pris uniformément au hasard dans la population courante. La sélection est faite en fonction de la fitness des individus, et l’enfant produit peut muter avant de remplacer quelqu’un dans la population.

Le modèle de Moran se généralise facilement à un nombre fini quelconque de types. Nous allons utiliser le modèle de Moran avec le cadre introduit par Eigen, afin de retrouver un phénomène de seuil d’erreur et de quasi-espèce. Tout comme dans le modèle de Wright–Fisher, le processus de Moran est une chaîne de Markov ergodique, et nous nous intéressons à la concentration des différentes classes de Hamming quand le nombre de générations tend vers l’infini. Également, la taille de la population constante introduit un couplage entre les différents individus qui rend l’étude du processus de Moran bien plus difficile que l’étude du modèle de Galton–Watson.

Nous allons considérer le même régime asymptotique que pour le processus de Wright–Fisher, comme dans [10]. Ainsi, Cerf obtient une courbe critique séparant un régime de désordre d’un régime de quasi-espèce [10]. Il obtient aussi la concentration de la master sequence dans le régime de quasi-espèce, qui coïncide avec la concentration de la master sequence pour le modèle d’Eigen associé.



Patrick Moran

L'objectif de cette partie est de montrer qu'il en est de même pour les concentrations des autres classes de Hamming.

Nous pouvons distinguer deux comportements radicalement différents dans le processus de Moran : quand la master sequence est présente dans la population et quand elle ne l'est pas. En effet, si la master sequence n'est pas présente dans la population, la population est neutre, et le processus de Moran dédie son temps à explorer l'espace d'états de manière aléatoire. Le temps moyen pour retrouver la master sequence, le *temps de découverte*, est alors de l'ordre de la taille de l'espace d'états. Inversement, si la master sequence est présente dans la population, et si sa fitness est assez grande, elle met très longtemps à disparaître de la population, le temps moyen pour cette disparition est appelé le *temps de persistance*. La comparaison entre le temps de découverte et le temps de persistance donne lieu à la courbe critique mentionnée. Pour retrouver les concentrations des différentes classes dans la distribution de la quasi-espèce, il suffit donc d'étudier la région sur-critique et la dynamique du processus quand la master sequence est présente dans la population.

Dans le premier chapitre, nous présentons le processus de Moran, nous énon-

çons le résultat principal de cette partie, et nous donnons quelques résultats de [10] utiles pour la suite. Comme dans le modèle de Wright–Fisher, une des principales difficultés est que la taille de l’espace d’états croît avec la longueur des génotypes (il y a de plus en plus de classes de Hamming). Pour surmonter cette difficulté, nous bornons ensuite la chaîne de Markov de Moran par d’autres chaînes de Markov plus simples, dont l’espace d’états ne croît pas, même si la longueur du génome grandit (chapitre 17). Ces chaînes de Markov simplifiées sont étudiées au chapitre 18. Dans le régime asymptotique que nous allons considérer, les trajectoires de ces chaînes de Markov se concentrent sur des trajectoires déterministes, données par un système d’équations différentielles. Le système d’équations en question s’obtient à partir du système d’équations d’Eigen par un changement de temps, et a donc les mêmes solutions stationnaires que le système d’Eigen. Nous avons donc deux solutions stationnaires, 0 et ρ^* , où ρ^* représente la la distribution de la quasi-espèce :

$$\forall k \geq 0 \quad \rho_k^* = (\sigma e^{-\alpha} - 1) \frac{\alpha^k}{k!} \sum_{i \geq 1} \frac{i^k}{\sigma^i}.$$

De plus, la solution ρ^* attire toutes les trajectoires avec une condition initiale ayant la coordonnée 0 strictement positive. L’objectif est de montrer que la mesure invariante des chaînes de Markov simplifiées se concentre autour de ρ^* (chapitre 19). Le problème est que la mesure invariante est une fonction de toute la trajectoire du processus ; les chaînes étant ergodiques, elles sont ramenées près du point fixe 0 une infinité de fois et elles pourraient y rester longtemps.

Même si la problématique et les résultats sont très similaires au cas du modèle de Wright–Fisher, le traitement est différent. En effet, le processus de Wright–Fisher converge vers un système dynamique à temps discret, tandis que le processus de Moran converge vers une équation différentielle. Le système d’équations limite a une forme triangulaire, c’est-à-dire, qu’asymptotiquement, la dynamique de chaque coordonnée ne dépend que d’elle-même et des coordonnées précédentes. De plus, cette dynamique est une dynamique de naissance et mort. Ainsi, nous réussissons à construire des chaînes de naissance et mort markoviennes pour borner la dynamique des différentes coordonnées, que nous pouvons ensuite étudier grâce à des formules explicites déjà existantes dans la littérature.

Chapitre 15

Modèle et résultats

Dans ce chapitre, nous introduisons le modèle de Moran qui sera l'objet de notre étude. Ensuite nous énonçons le résultat principaux de cette partie. Pour finir, nous donnons les outils principales dont nous aurons besoin.

15.1 Le modèle

Puisque notre objectif est de retrouver un phénomène de seuil d'erreur et une distribution de quasi-espèce pour le modèle de Moran, nous reprenons le cadre introduit dans la section 3.1 pour le modèle d'Eigen. Nous considérons donc des individus dont les génotypes sont des chaînes de longueur $\ell \geq 1$ sur un alphabet fini \mathcal{A} de cardinal $\kappa \geq 2$. Nous nous donnons également une fonction de fitness $f : \mathcal{A}^\ell \rightarrow \mathbb{R}^+$, et nous supposons que les mutations arrivent indépendamment au hasard sur chaque site de la chaîne avec probabilité $q \in]0, 1[$. La matrice de mutation $(M(u, v), u, v \in \mathcal{A}^\ell)$ est donc donnée par :

$$M(u, v) = \left(\frac{q}{\kappa - 1} \right)^{d_H(u, v)} (1 - q)^{\ell - d_H(u, v)}, \quad u, v \in \mathcal{A}^\ell.$$

Comme précédemment, $d_H(u, v)$ représente distance de Hamming entre u et v , i.e., le nombre de digits où les deux chaînes diffèrent. La taille de la population est finie, et nous la notons par $m \geq 1$. La taille de la population reste constante au cours du temps.

Dynamique. De façon informelle, le modèle de Moran évolue ainsi : les m individus dans la population se reproduisent dans de temps aléatoires, plus ou moins souvent, en fonction de leur fitness. Deux reproductions ne peuvent pas avoir lieu au même instant. Lorsqu'un individu se reproduit, il crée un enfant qui mute selon le noyau de mutation M . Finalement, l'enfant remplace un individu choisi uniformément au hasard dans la population. Il s'agit bien d'une dynamique markovienne, c'est à dire, qu'à chaque instant, l'évolution future du processus ne dépend du passé qu'à travers l'état courant de la population. Nous utiliserons donc un processus de Markov pour décrire le modèle de Moran.

Transformations possibles. Dans le modèle de Moran, les seules transformations possibles dans une population consistent à remplacer un chromosome de la population par un nouveau chromosome. Pour une population $x \in (\mathcal{A}^\ell)^m$, et pour $j \in \{1, \dots, m\}$, $u \in \mathcal{A}^\ell$, nous notons par $x(j \leftarrow u)$ la population x dans laquelle le j -ème chromosome $x(j)$ est remplacé par u :

$$x(j \leftarrow u) = \begin{pmatrix} x(1) \\ \vdots \\ x(j-1) \\ u \\ x(j+1) \\ \vdots \\ x(m) \end{pmatrix}.$$

Le modèle de Moran classique peut être défini aussi bien en temps continu qu'en temps discret.

Temps discret. Le modèle de Moran en temps discret est une chaîne de Markov $(X_n)_{n \geq 0}$, d'espace d'états $(\mathcal{A}^\ell)^m$, et matrice de transition p donnée par :

$$\forall x \in (\mathcal{A}^\ell)^m \quad \forall j \in \{1, \dots, m\} \quad \forall u \in \mathcal{A}^\ell \setminus \{x(j)\} \\ p(x, x(j \leftarrow u)) = \frac{1}{\lambda m^2} \sum_{1 \leq i \leq m} f(x(i)) M(x(i), u),$$

où λ est un nombre réel supérieur au maximum de la fonction fitness f :

$$\lambda \geq \max\{f(u) : u \in \mathcal{A}^\ell\}.$$

Les autres coefficients non diagonaux de la matrice de transition sont nuls. Les coefficients diagonaux sont choisis de sorte que la matrice soit stochastique,

i.e., la somme de chaque ligne vaut 1. Le mécanisme de transition pour passer de X_n à X_{n+1} est le suivant : le i -ème individu est choisi dans la population X_n uniformément au hasard. Avec probabilité $1 - f(X_n(i))/\lambda$ rien ne se passe, et $X_{n+1} = X_n$. Avec probabilité $f(X_n(i))/\lambda$ l'individu choisi se reproduit, son enfant mute selon le noyau M , et remplace un individu dans la population X_n , choisi uniformément au hasard.

Temps continu. Le modèle de Moran en temps continu est un processus de Markov $(X_t)_{t \geq 0}$, d'espace d'états $(\mathcal{A}^\ell)^m$, et générateur infinitésimal donné par : pour toute fonction $\varphi : (\mathcal{A}^\ell)^m \rightarrow \mathbb{R}$ et pour tout $x \in (\mathcal{A}^\ell)^m$,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \left(\mathbb{E}(\varphi(X_t) | X_0 = x) - \varphi(x) \right) = \sum_{1 \leq i, j \leq m} \sum_{u \in \mathcal{A}^\ell} f(x(i)) M(x(i), u) \left(\varphi(x(j \leftarrow u)) - \varphi(x) \right)$$

La dynamique du processus $(X_t)_{t \geq 0}$ est gouvernée par une horloge exponentielle de paramètre λm^2 . Quand l'horloge sonne, la population change selon le même mécanisme de transition que la chaîne de Markov $(X_n)_{n \geq 0}$. Dans le modèle en temps continu, le même générateur es obtenu si l'on suppose que chaque individu $x(i)$ dans la population x a une horloge exponentielle associée, de paramètre $f(x(i))m$, et que les horloges sont indépendantes. Lorsqu'une horloge sonne, l'individu correspondant se reproduit, mute selon le noyau M et remplace un individu choisi uniformément au hasard dans la population x .

Ergodicité. Si la fonction de fitness $f : \mathcal{A}^\ell \rightarrow \mathbb{R}^+$ est strictement positive, les processus que nous venons de décrire sont ergodiques. Puisque nous voulons retrouver une distribution de quasi-espèce pour le modèle de Moran, et que comme nous avons bien observé pour le modèle d'Eigen dans la partie I, la distribution de quasi-espèce apparaît dans un état d'équilibre, nous allons supposer que toutes les fitness sont strictement positives.

Hypothèse. La fonction de fitness $f : \mathcal{A}^\ell \rightarrow \mathbb{R}^+$ est strictement positive.

Sous cette hypothèse, le processus de Moran en temps discret, et le processus de Moran en temps continu sont des processus ergodiques. De plus, les deux processus ont la même mesure invariante.

Régime asymptotique. Le but des sections qui suivent est d'étudier ces modèles

dans le régime asymptotique

$$\begin{aligned} \ell &\rightarrow +\infty, & m &\rightarrow +\infty, & q &\rightarrow 0, \\ \ell q &\rightarrow a \in]0, +\infty[, & \frac{m}{\ell} &\rightarrow \alpha \in [0, +\infty], \end{aligned}$$

afin de montrer l'existence d'un phénomène de seuil d'erreur et d'une distribution de quasi-espèce.

15.2 Résultat principal

Nous ne considérons que le paysage à un pic. Soit $\sigma > 1$. Il existe une séquence, notée w^* et appelée master sequence, telle que la fonction de fitness est donnée par

$$\forall u \in \mathcal{A}^\ell \quad f(u) = \begin{cases} \sigma & \text{si } u = w^*, \\ 1 & \text{si } u \neq w^*. \end{cases}$$

Soit x une population dans $(\mathcal{A}^\ell)^m$. Fixons $k \geq 0$ un entier. Nous nous intéressons au nombre $N^k(x)$ de chromosomes dans la population x à distance de Hamming k de la master sequence :

$$N^k(x) = \text{card} \{ i : 1 \leq i \leq m, d_H(x(i), w^*) = k \}.$$

Soit $\psi : \mathbb{R}^+ \rightarrow \mathbb{R}^+ \cup \{+\infty\}$ la fonction définie par

$$\forall a < \ln \sigma \quad \psi(a) = \frac{\sigma(1 - e^{-a}) \ln \frac{\sigma(1 - e^{-a})}{\sigma - 1} + \ln(\sigma e^{-a})}{(1 - \sigma(1 - e^{-a}))}$$

et $\psi(a) = 0$ si $a \geq \ln \sigma$. Nous allons démontrer le résultat suivant :

Théorème 15.2.1. *Nous avons la dichotomie suivante :*

- Si $\alpha \psi(a) < \ln \kappa$, alors

$$\forall k \geq 0 \quad \lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \lim_{t \rightarrow \infty} E \left(\frac{1}{m} N^k(X_t) \right) = 0.$$

- Si $\alpha \psi(a) > \ln \kappa$, alors

$$\forall k \geq 0 \quad \lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \lim_{t \rightarrow \infty} E \left(\frac{1}{m} N^k(X_t) \right) = \rho_k^* = (\sigma e^{-a} - 1) \frac{a^k}{k!} \sum_{i \geq 1} \frac{i^k}{\sigma^i}.$$

De plus, dans les deux cas,

$$\forall k \geq 0 \quad \lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha, \frac{m}{\ell} \rightarrow \alpha}} \lim_{t \rightarrow \infty} \text{Var} \left(\frac{1}{m} N^k(X_t) \right) = 0.$$

15.3 Modèles normalisés

Dans les modèles de Moran discret et continu décrits dans la section 15.1, le taux de reproduction total dans la population varie en fonction de la composition de la population. Introduisons la *fitness moyenne* $\phi(x)$ d'une population $x \in (\mathcal{A}^\ell)^m$:

$$\phi(x) = \frac{f(x(1)) + \cdots + f(x(m))}{m}.$$

Les reproductions arrivent plus souvent dans une population avec une fitness moyenne forte que dans une population avec une fitness moyenne faible. Or, nous pouvons également définir des processus de Moran *normalisés*, en imposant que le taux de reproduction total de toute population soit constant et égal à 1, ce qui peut être utile. En temps discret, le processus de Moran normalisé est la chaîne de Markov $(X_n)_{n \geq 0}$ à valeurs dans $(\mathcal{A}^\ell)^m$ de matrice de transition

$$\forall x \in (\mathcal{A}^\ell)^m \quad \forall j \in \{1, \dots, m\} \quad \forall u \in \mathcal{A}^\ell \setminus \{x(j)\} \\ p(x, x(j \leftarrow u)) = \frac{1}{m^2 \phi(x)} \sum_{1 \leq i \leq m} f(x(i)) M(x(i), u).$$

Les autres coefficients non diagonaux de la matrice de transition sont nuls. Les coefficients diagonaux sont choisis de sorte que la matrice soit stochastique, i.e., la somme de chaque ligne vaut 1. En temps continu, le processus de Moran normalisé est le processus de Markov $(X_t)_{t \geq 0}$ à valeurs dans $(\mathcal{A}^\ell)^m$ et ayant générateur infinitésimal L donné par : pour toute fonction $\varphi : (\mathcal{A}^\ell)^m \rightarrow \mathbb{R}$ et pour tout $x \in (\mathcal{A}^\ell)^m$,

$$L\varphi(x) = \sum_{1 \leq i, j \leq m} \sum_{u \in \mathcal{A}^\ell} \frac{f(x(i)) M(x(i), u)}{m \phi(x)} (\varphi(x(j \leftarrow u)) - \varphi(x)).$$

Tout comme pour le modèle de Moran classique, le processus de Moran discret normalisé et le processus de Moran continu normalisé ont la même mesure de probabilité invariante. Nous montrons maintenant que le théorème 15.2.1 est

vrai pour le processus original (i.e. non normalisé) si et seulement si il est vrai pour le processus normalisé. Mais d'abord nous énonçons un théorème plus fort que le théorème principal 15.2.1. Appelons $(X_n^\alpha)_{n \geq 0}$ le processus de Moran discret original et $(X_n^\beta)_{n \geq 0}$ le processus de Moran discret normalisé. Soit θ un symbole représentant α ou β . Appelons également μ^θ la mesure de probabilité invariante du processus $(X_n^\theta)_{n \geq 0}$. Pour $k \geq 0$ notons par ν_k^θ la mesure image de μ^θ par l'application

$$x \in (\mathcal{A}^\ell)^m \mapsto \frac{1}{m} N^k(x) \in [0, 1].$$

Posons, pour $k \geq 0$:

$$\rho_k^* = (\sigma e^{-\alpha} - 1) \frac{\alpha^k}{k!} \sum_{i \geq 1} \frac{i^k}{\sigma^i}.$$

Nous avons le résultat suivant :

Théorème 15.3.1. *Nous supposons que*

$$\ell \rightarrow +\infty, \quad m \rightarrow +\infty, \quad q \rightarrow 0,$$

et que

$$\ell q \rightarrow \alpha \in]0, +\infty[, \quad \frac{m}{\ell} \rightarrow \alpha \in [0, +\infty].$$

Soit $\psi : \mathbb{R}^+ \rightarrow \mathbb{R}^+ \cup \{+\infty\}$ la fonction définie dans la section précédente. Nous avons :

- Si $\alpha \psi(\alpha) < \ln \kappa$, alors pour tout $k \geq 0$, la mesure ν_k^θ converge vers la masse de Dirac au point 0 :

$$\forall \varepsilon > 0 \quad \nu_k^\theta([0, \varepsilon]) \rightarrow 1.$$

- Si $\alpha \psi(\alpha) > \ln \kappa$, alors pour tout $k \geq 0$, la mesure ν_k^θ converge vers la masse de Dirac au point ρ_k^* :

$$\forall \varepsilon > 0 \quad \nu_k^\theta([\rho_k^* - \varepsilon, \rho_k^* + \varepsilon]) \rightarrow 1.$$

Le reste de cette partie est dédié à démontrer ce théorème, plutôt que le résultat principal 15.2.1. Ce théorème avec $\theta = \alpha$ entraîne le théorème principal 15.2.1. Nous allons démontrer ce théorème pour le modèle de Moran normalisé $(X_n^\beta)_{n \geq 0}$, car le modèle de Moran normalisé nous permet d'utiliser des arguments de couplage et de monotonie qui rendent la démonstration plus élémentaire que dans le cas général. Nous montrons maintenant que le théorème 15.3.1 est vrai pour le modèle de Moran original, si et seulement si il est vrai pour le modèle de Moran normalisé.

Proposition 15.3.2. *Le théorème 15.3.1 est vrai pour le modèle de Moran originale (i.e. avec $\theta = \alpha$), si et seulement si il est vrai pour le modèle de Moran normalisé (i.e. avec $\theta = \beta$).*

Démonstration. Pour θ égal à α ou β , appelons p^θ la matrice de transition du processus $(X_n^\theta)_{n \geq 0}$. La mesure de probabilité μ^β est la seule solution du système d'équations

$$\forall x \in (\mathcal{A}^\ell)^m \quad \mu^\beta(x) = \sum_{y \in (\mathcal{A}^\ell)^m} \mu^\beta(y) p^\beta(y, x).$$

Nous pouvons réécrire ce système ainsi :

$$\forall x \in (\mathcal{A}^\ell)^m \quad \mu^\beta(x) \sum_{\substack{y \in (\mathcal{A}^\ell)^m \\ y \neq x}} p^\beta(x, y) = \sum_{\substack{y \in (\mathcal{A}^\ell)^m \\ y \neq x}} \mu^\beta(y) p^\beta(y, x).$$

Les matrices de transition p^β et p^α vérifient la relation :

$$\forall x, y \in (\mathcal{A}^\ell)^m, \quad x \neq y, \quad p^\alpha(x, y) = \frac{\phi(x)}{\lambda} p^\beta(x, y).$$

D'où,

$$\forall x \in (\mathcal{A}^\ell)^m \quad \frac{\mu^\beta(x)\lambda}{\phi(x)} \sum_{\substack{y \in (\mathcal{A}^\ell)^m \\ y \neq x}} p^\alpha(x, y) = \sum_{\substack{y \in (\mathcal{A}^\ell)^m \\ y \neq x}} \frac{\mu^\beta(y)\lambda}{\phi(y)} p^\alpha(y, x).$$

L'unicité de la mesure de probabilité invariante μ^α entraîne que

$$\forall x \in (\mathcal{A}^\ell)^m \quad \mu^\alpha(x) = \frac{\mu^\beta(x)/\phi(x)}{\sum_{y \in (\mathcal{A}^\ell)^m} \mu^\beta(y)/\phi(y)}.$$

Puisque nous travaillons avec le paysage à un pic, la fitness moyenne $\phi(x)$ peut se réécrire ainsi :

$$\phi(x) = 1 + (\sigma - 1) \frac{1}{m} N^0(x).$$

Pour $\theta = \alpha, \beta$ et $k \geq 0$ nous notons par ν_k^θ l'image de la mesure μ^θ par l'application

$$x \in (\mathcal{A}^\ell)^m \longmapsto \frac{1}{m} N^k(x) \in [0, 1].$$

Nous pouvons donc réécrire

$$\begin{aligned} \sum_{\mathbf{y} \in (\mathcal{A}^\ell)^m} \frac{\mu^\beta(\mathbf{y})}{\Phi(\mathbf{y})} &= \sum_{\mathbf{y} \in (\mathcal{A}^\ell)^m} \frac{\mu^\beta(\mathbf{y})}{1 + (\sigma - 1)N^0(\mathbf{y})/m} \\ &= \int_{[0,1]^{k+1}} (1 + (\sigma - 1)t_0)^{-1} d\nu_0^\beta(t_0). \end{aligned}$$

Nous posons, pour tout $k \geq 0$,

$$\forall t \in [0, 1]^k \quad \beta(t) = (1 + (\sigma - 1)t_0)^{-1}.$$

Alors, pour $k \geq 0$ et $g : [0, 1] \rightarrow \mathbb{R}$ une fonction quelconque, nous avons

$$\begin{aligned} \int_{[0,1]} g d\nu_k^\alpha &= \lim_{n \rightarrow \infty} \mathbb{E} \left(g \left(\frac{1}{m} N^k(X_n^\alpha) \right) \right) = \sum_{x \in (\mathcal{A}^\ell)^m} g \left(\frac{1}{m} N^k(x) \right) \mu^\alpha(x) = \\ &= \frac{\sum_{x \in (\mathcal{A}^\ell)^m} g \left(\frac{N^k(x)}{m} \right) \frac{\mu^\beta(x)}{\Phi(x)}}{\sum_{\mathbf{y} \in (\mathcal{A}^\ell)^m} \frac{\mu^\beta(\mathbf{y})}{\Phi(\mathbf{y})}} = \frac{\int g(t) \beta(t) d\nu_k^\beta(t)}{\int \beta(t) d\nu_0^\beta(t)} \end{aligned}$$

Nous considérons maintenant le régime asymptotique du théorème 15.3.1. Par le théorème 15.3.1, sauf sur la courbe critique $\alpha\psi(a) = \ln \kappa$, les mesures de probabilité ν_k^β convergent vers des masses de Dirac. En vu du dernier calcul, il en est de même de les mesures de probabilité ν_k^α , et les masses de Dirac coïncident. Un argument similaire montre l'implication réciproque. \square

Dorénavant, nous ne travaillons que avec le processus de Moran normalisé.

15.4 Lumping

De même que pour le modèle de Wright–Fisher dans la partie précédente, nous voudrions réduire l'espace d'états du processus de Moran, $(\mathcal{A}^\ell)^m$, afin de le rendre plus maniable. Dans le même esprit que pour le processus de Wright–Fisher dans la section 9.3, nous dérivons ici un processus de distance et un processus d'occupation à partir du processus de Moran normalisé. Même si les détails sont similaires à ceux de la section 9.3, nous les incluons ici car

cela nous permet de fixer quelques notations importantes pour la suite. Soit $H : \mathcal{A}^\ell \rightarrow \{0, \dots, \ell\}$ l'application qui à chaque génotype u l'associe sa classe de Hamming vis-à-vis de la master sequence :

$$\forall u \in \mathcal{A}^\ell \quad H(u) = d_H(u, w^*).$$

La fonction de fitness $f : \mathcal{A}^\ell \rightarrow \mathbb{R}^+$ du paysage à un pic respecte les classes de Hamming. Nous pouvons donc définir une nouvelle fonction de fitness $f_H : \{0, \dots, \ell\} \rightarrow \mathbb{R}^+$ en posant $f_H(k)$ égale à la valeur commune de $f(u)$ pour u dans la classe de Hamming k :

$$\forall k \in \{0, \dots, \ell\} \quad f_H(k) = \begin{cases} \sigma & \text{si } k = 0, \\ 1 & \text{si } k \geq 1. \end{cases}$$

D'après le lemme 3.2.1, le noyau de mutation respecte aussi les classe de Hamming. Nous en dérivons un noyau de mutation $(M_H(b, c), 0 \leq b, c \leq \ell)$, donné par :

$$M_H(b, c) = \sum_{\substack{0 \leq k \leq \ell - b \\ 0 \leq l \leq b \\ b - l + k = c}} \binom{\ell - b}{k} \binom{b}{l} q^k (1 - q)^{\ell - b - k} \left(\frac{q}{\kappa - 1}\right)^l \left(1 - \frac{q}{\kappa - 1}\right)^{b - l}.$$

Processus de distance. Nous définissons ensuite une fonction vectorielle $\mathbb{H} : (\mathcal{A}^\ell)^m \rightarrow \{0, \dots, \ell\}^m$ par :

$$\forall x = \begin{pmatrix} x(1) \\ \vdots \\ x(m) \end{pmatrix} \in (\mathcal{A}^\ell)^m \quad \mathbb{H}(x) = \begin{pmatrix} H(x(1)) \\ \vdots \\ H(x(m)) \end{pmatrix}.$$

Nous définissons le processus de distance $(D_n)_{n \geq 0}$ par

$$\forall n \geq 0 \quad D_n = \mathbb{H}(X_n).$$

Le processus de distance est un processus de Markov, ce qui se voit d'après le théorème de *lumping* B.1 et la proposition suivante :

Proposition 15.4.1. *La matrice de transition p du processus de Moran $(X_n)_{n \geq 0}$ vérifie :*

$$\forall e \in \{0, \dots, \ell\}^m \quad \forall x, y \in (\mathcal{A}^\ell)^m \\ \mathbb{H}(x) = \mathbb{H}(y) \implies \sum_{\substack{z \in (\mathcal{A}^\ell)^m \\ \mathbb{H}(z) = e}} p(x, z) = \sum_{\substack{z \in (\mathcal{A}^\ell)^m \\ \mathbb{H}(z) = e}} p(y, z).$$

La démonstration est similaire à celle de la section 7.3. Nous concluons grâce au théorème de *lumping* B.1 que le processus de distance $(D_n)_{n \geq 0}$ est un processus de Markov. L'espace d'états du processus de distance est $\{0, \dots, \ell\}^m$ et la matrice de transition p_H est donnée par :

$$\forall d \in \{0, \dots, \ell\}^m \quad \forall j \in \{1, \dots, m\} \quad \forall b \in \{0, \dots, \ell\} \setminus \{d(j)\}$$

$$p_H(d, d(j \leftarrow b)) = \frac{1}{m \sum_{1 \leq i \leq m} f_H(d(i))} \sum_{1 \leq i \leq m} f_H(d(i)) M_H(d(i), b).$$

Processus d'occupation. Nous définissons l'ensemble $\mathcal{P}_{\ell+1}^m$ des partitions ordonnées de l'entier m dans au plus $\ell + 1$ parties :

$$\mathcal{P}_{\ell+1}^m = \{ (o(0), \dots, o(\ell)) \in \mathbb{N}^{\ell+1} : o(0) + \dots + o(\ell) = m \}.$$

Nous regardons ces partitions comme distributions d'occupation. La partition $(o(0), \dots, o(\ell))$ correspond à une population où $o(k)$ individus appartiennent à la classe de Hamming k . Soit \mathcal{O} l'application qu'à chaque population $x \in (\mathcal{A}^\ell)^m$ associe la distribution d'occupation correspondante :

$$\forall k \in \{0, \dots, \ell\} \quad \mathcal{O}(x)(k) = \sum_{1 \leq i \leq m} 1_{H(x(i))=k}.$$

Nous définissons le processus d'occupation $(O_n)_{n \geq 0}$ par :

$$\forall n \geq 0 \quad O_n = \mathcal{O}(X_n).$$

Le processus d'occupation peut également être obtenu à partir du processus distance. En effet, nous définissons l'application $\mathcal{O}_H : (\mathcal{A}^\ell)^m \rightarrow \mathcal{P}_{\ell+1}^m$ par :

$$\forall d \in (\mathcal{A}^\ell)^m \quad \forall k \in \{0, \dots, \ell\} \quad \mathcal{O}_H(d)(k) = \sum_{1 \leq i \leq m} 1_{d(i)=k}.$$

Nous avons donc,

$$\forall x \in (\mathcal{A}^\ell)^m \quad \mathcal{O}(x) = \mathcal{O}_H(\mathbb{H}(x)).$$

Les seuls changements acceptés pour le processus $(O_n)_{n \geq 0}$ sont les changements du type

$$o \rightarrow o - e_i + e_j, \quad 0 \leq i, j \leq \ell,$$

où $(e_k)_{0 \leq k \leq \ell}$ représente la base canonique de $\mathbb{R}^{\ell+1}$, indexée par $0, \dots, \ell$ au lieu de $1, \dots, \ell + 1$. Pour une distribution d'occupation $o \in \mathcal{P}_{\ell+1}^m$, et $i, j \in \{0, \dots, \ell\}$ nous utilisons souvent la notation

$$o(i \rightarrow j) = o - e_i + e_j.$$

Proposition 15.4.2. *La matrice de transition p_H du processus de distance $(D_n)_{n \geq 0}$ vérifie :*

$$\forall o \in \mathcal{P}_{\ell+1}^m \quad \forall d, e \in \{0, \dots, \ell\}^m \\ \mathcal{O}_H(d) = \mathcal{O}_H(e) \implies \sum_{\substack{f \in \{0, \dots, \ell\}^m \\ \mathcal{O}_H(f) = o}} p_H(d, f) = \sum_{\substack{f \in \{0, \dots, \ell\}^m \\ \mathcal{O}_H(f) = o}} p_H(e, f).$$

La démonstration de ce résultat est similaire à celle de la section 7.3. Nous pouvons donc appliquer le théorème de *lumping* B.1 pour conclure que le processus $(O_n)_{n \geq 0}$ est bien un processus de Markov. L'espace d'états du processus de Markov $(O_n)_{n \geq 0}$ est $\mathcal{P}_{\ell+1}^m$, et la matrice de transition p_O est donnée par :

$$\forall o \in \mathcal{P}_{\ell+1}^m \quad \forall i, j \in \{0, \dots, \ell\} \quad i \neq j \\ p_O(o, o - e_i + e_j) = \frac{1}{(\sigma - 1)o(0) + m} o(i) \sum_{h=0}^{\ell} o(h) f_H(h) M_H(h, j).$$

Mesures de probabilité invariantes Fixons $k \geq 0$. Notre but est d'estimer la loi ν_k décrivant la fraction de chromosomes qui sont dans la classe de Hamming k dans une population à l'équilibre. La mesure de probabilité ν_k est une mesure de probabilité sur l'intervalle $[0, 1]$ qui satisfait les identités suivantes. Pour toute fonction $g : [0, 1] \rightarrow \mathbb{R}$,

$$\int_{[0,1]} g d\nu_k = \lim_{n \rightarrow \infty} \mathbb{E} \left(g \left(\frac{1}{m} N^k(X_n) \right) \right) = \int_{(\mathcal{A}^\ell)^m} g \left(\frac{1}{m} N^k(x) \right) d\mu(x),$$

où μ est la mesure de probabilité invariante du processus $(X_n)_{n \geq 0}$ et $N^k(x)$ est le nombre de chromosomes dans la population x qui sont à distance k de la master sequence :

$$N^k(x) = \text{card} \{ i : 1 \leq i \leq m, d_H(x(i), w^*) = k \}.$$

Le nombre $N^k(x)$ peut être exprimé en fonction de $\mathbb{H}(x)$:

$$\forall x \in (\mathcal{A}^\ell)^m \quad N^k(x) = N_H^k(\mathbb{H}(x)),$$

où N_H^k est la fonction donnée par

$$\forall d \in \{0, \dots, \ell\}^m \quad N_H^k(d) = \text{card} \{ i : 1 \leq i \leq m, d(i) = k \}.$$

Soit μ_H la mesure de probabilité invariante du processus $(D_n)_{n \geq 0}$. Nous avons pour $d \in \{0, \dots, \ell\}^m$,

$$\begin{aligned} \mu_H(d) &= \lim_{n \rightarrow \infty} P(D_n = d) = \lim_{n \rightarrow \infty} P(\mathbb{H}(X_n) = d) \\ &= \lim_{n \rightarrow \infty} P(X_n \in \mathbb{H}^{-1}(d)) = \mu(\mathbb{H}^{-1}(d)). \end{aligned}$$

Il s'ensuit que, pour toute fonction $g : [0, 1] \rightarrow \mathbb{R}$,

$$\begin{aligned} \int_{[0,1]} g \, d\nu_k &= \int_{(\mathcal{A}^\ell)^m} g\left(\frac{1}{m} N^k(x)\right) d\mu(x) \\ &= \int_{(\mathcal{A}^\ell)^m} g\left(\frac{1}{m} N_H^k(\mathbb{H}(x))\right) d\mu(x) \\ &= \int_{\{0, \dots, \ell\}^m} g\left(\frac{1}{m} N_H^k(d)\right) d\mu_H(d). \end{aligned}$$

De manière analogue, la mesure de probabilité invariante μ_O de la chaîne de Markov $(O_n)_{n \geq 0}$ satisfait pour toute fonction $g : [0, 1] \rightarrow \mathbb{R}$ l'identité

$$\int_{[0,1]} g \, d\nu_k = \int_{\mathcal{P}_{\ell+1}^m} g\left(\frac{o(k)}{m}\right) d\mu_O(o).$$

15.5 Convergence vers le modèle d'Eigen

Afin de donner un argument heuristique de notre résultat principal, nous donnons ici un résultat du type loi des grandes nombres pour le processus de Moran. Nous le faisons dans un cadre plus général que celui introduit ci-dessus. Nous considérons l'ensemble de génotypes $\{1, \dots, N\}$, une fonction de fitness $f : \{1, \dots, N\} \rightarrow \mathbb{R}^+$, et un noyau de mutation $(M(i, j))_{1 \leq i, j \leq N}$. Notons par m la taille de la population. Le processus d'occupation de Moran associé au cadre que nous venons d'introduire est une chaîne de Markov $(X_n)_{n \geq 0}$ à valeurs dans

$$\mathcal{P}_N^m = \{z \in \mathbb{N}^N : z(1) + \dots + z(N) = m\},$$

et de matrice de transition p donnée par :

$$\begin{aligned} \forall z \in \mathcal{P}_N^m \quad \forall i, j \in \{1, \dots, N\} \quad i \neq j \\ p(z, z - e_i + e_j) = \frac{z(i)}{\lambda m^2} \sum_{h=1}^N z(h) f(h) M(h, j), \end{aligned}$$

où $\lambda \geq \max\{f(h) : 1 \leq h \leq N\}$ et $(e_i)_{1 \leq i \leq N}$ représente la base canonique de \mathbb{R}^N .

Rappelons que dans ce contexte, le modèle d'Eigen est donné par le système d'équations différentielles :

$$(S) \quad x'_k(t) = \sum_{1 \leq i \leq N} x_i(t)f(i)M(i, k) - x_k(t) \sum_{1 \leq i \leq N} x_i(t)f(i), \quad 1 \leq k \leq N.$$

Soit S^N le simplexe à N coordonnées :

$$S^N = \{x \in [0, 1]^N : x_1 + \dots + x_N = 1\}.$$

Nous allons montrer le résultat suivant :

Théorème 15.5.1. *Soit $(X_n)_{n \geq 0}$ le processus d'occupation du modèle de Moran décrit ci-dessus. Supposons qu'on a la convergence des conditions initiales vers $x^0 \in S^N$:*

$$\lim_{m \rightarrow \infty} \frac{1}{m} X_0 = x^0,$$

et soit $(x(t))_{t \geq 0}$ la solution du système d'équations différentielles (S) avec condition initiale $x(0) = x^0$. Alors, pour tout $\delta, T > 0$ nous avons

$$\lim_{m \rightarrow \infty} P\left(\sup_{0 \leq t \leq T} \left| \frac{1}{m} X_{\lfloor \lambda m t \rfloor} - x(t) \right| > \delta\right) = 0.$$

Ce théorème résulte d'une application directe du théorème 4.7 dans [51]. Afin de montrer ce résultat, nous allons d'abord présenter le théorème 4.7 dans [51], et nous montrerons ensuite que toutes les hypothèses nécessaires pour appliquer le théorème sont vérifiées dans notre cas.

Convergence d'une famille de chaînes de Markov. Soit $d \geq 1$ et soit E un sous ensemble borélien de \mathbb{R}^d . Soit $((X_n^m)_{n \geq 0}, m \geq 1)$ une séquence de chaînes de Markov à valeurs dans $E_m \subset E$ et matrices de transition $(p^m(x, y))_{x, y \in E_m}$. Soit $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$ et considérons le système d'équations différentielles

$$x'_i(t) = F_i(x(t)), \quad 1 \leq i \leq d.$$

Le théorème 4.7 dans [51] donne une série de conditions suffisantes sous lesquelles la suite de chaînes de Markov $(X^m)_{m \geq 1}$ converge vers une solution du système d'équations différentielles ci-dessus. Le résultat original dans [51] est formulé pour le cas plus général des chaînes de Markov à espace d'état

continu. Nous modifions la notation de [51] afin de la rendre plus adaptée à notre cas particulier. Le théorème 4.7 dans [51] peut être appliqué si les hypothèses suivantes sont vérifiées. Il existe des suites de nombres positives $(\alpha_m)_{m \geq 1}$ et $(\varepsilon_m)_{m \geq 1}$ telles que

1. $\lim_{m \rightarrow \infty} \alpha_m = \infty$ et $\lim_{m \rightarrow \infty} \varepsilon_m = 0$.
2. $\sup_{m \geq 1} \sup_{x \in E_m} \alpha_m \sum_{y \in E_m} |y - x| p^m(x, y) < \infty$.
3. $\lim_{m \rightarrow \infty} \sup_{x \in E_m} \alpha_m \sum_{y \in E_m: |y-x| > \varepsilon_m} |y - x| p^m(x, y) = 0$.

Nous définissons, pour $m \geq 1$, $F^m(x) = \alpha_m \sum_{y \in E_m} (y - x) p^m(x, y)$.

4. $\lim_{m \rightarrow \infty} \sup_{x \in E_m} |F^m(x) - F(x)| = 0$.
5. Il existe une constante M telle que

$$\forall x, y \in E, \quad |F(x) - F(y)| \leq M|x - y|.$$

Théorème 15.5.2 (Kurtz). *Si les conditions 1–5 sont satisfaites, et que l'on a de plus la convergence des conditions initiales*

$$\lim_{m \rightarrow \infty} X_0^m = x^0,$$

alors, pour tout $\delta, T > 0$, nous avons

$$\lim_{m \rightarrow \infty} P \left(\sup_{0 \leq t \leq T} |X_{[\alpha_m t]}^m - x(t)| > \delta \right) = 0.$$

Démonstration du théorème 15.5.1. Soit $(X_n)_{n \geq 0}$ le processus d'occupation associé au modèle de Moran introduit au début de la section. Nous voudrions appliquer le théorème 15.5.2 à la suite de chaînes de Markov $((X_n/m)_{n \geq 0}, m \geq 1)$. Nous avons juste à trouver les bonnes suites $(\alpha_m)_{m \geq 1}$ et $(\varepsilon_m)_{m \geq 1}$ et à vérifier que les conditions 1–5 sont satisfaites dans notre cas. Pour $m \geq 1$, soit $\alpha_m = \lambda m$ et $\varepsilon_m = 2/m$. Les suites $(\alpha_m)_{m \geq 1}$ et $(\varepsilon_m)_{m \geq 1}$ vérifient la condition 1. Pour voir qu'elles vérifient également la condition 2, nous remarquons que, pour $z \in \mathcal{P}_N^m$,

$$\sum_{z' \in \mathcal{P}_N^m} \left| \frac{z}{m} - \frac{z'}{m} \right| p(o, z') = \sum_{1 \leq i, j \leq N} \left| -\frac{e_i}{m} + \frac{e_j}{m} \right| p(z, z - e_i + e_j) \leq \frac{\sqrt{2}}{m}.$$

D'où

$$\sup_{m \geq 1} \sup_{z \in \mathcal{P}_N^m} \alpha_m \sum_{z' \in \mathcal{P}_N^m} \left| \frac{z'}{m} - \frac{z}{m} \right| p(z, z') \leq \lambda \sqrt{2} < \infty,$$

et la condition 2 est donc vérifiée. Puisque $p(z, z') > 0$ si et seulement si $|z - z'| \leq \sqrt{2}$, pour tout $m \geq 1$ et $z \in \mathcal{P}_{\ell+1}^m$, nous avons

$$\sum_{z' \in \mathcal{P}_N^m : |z' - z| > m \varepsilon_m} \left| \frac{z'}{m} - \frac{z}{m} \right| p(z, z') = 0.$$

Nous en déduisons que la condition 3 est aussi satisfaite. Soit $F : \mathbb{R}^N \rightarrow \mathbb{R}^N$ la fonction définie par

$$\forall k \in \{1, \dots, N\} \quad \forall x \in \mathbb{R}^N$$

$$F_k(x) = \sum_{1 \leq i \leq N} x_i f(i) M(i, k) - x_k \sum_{0 \leq j \leq \ell} x_j f(j).$$

Toutes les dérivées partielles de F étant bornées sur le simplexe \mathcal{S}^N , F est une fonction Lipschitzienne sur \mathcal{S}^N , i.e., la condition 5 est vérifiée. Pour finir, nous calculons, pour $m \geq 1$ et $x \in \mathcal{P}_N^m/m$, la valeur de $F^m(x)$. Par définition,

$$F^m(x) = \lambda m \sum_{z \in \mathcal{P}_N^m} \left(\frac{z}{m} - x \right) p(mx, z) = \lambda \sum_{1 \leq i, j \leq N} (-e_i + e_j) p(mx, mx - e_i + e_j).$$

Donc, pour $k \in \{1, \dots, N\}$, nous avons

$$F_k^m(x) = \lambda \sum_{i: i \neq k} p(mx, mx - e_i + e_k) - \lambda \sum_{i: i \neq k} p(mx, mx - e_k + e_i)$$

$$= \sum_{i: i \neq k} x_i \sum_{1 \leq j \leq N} x_j f(j) M(j, k) - \sum_{i: i \neq k} x_k \sum_{1 \leq j \leq N} x_j f(j) M(j, i).$$

Puisque $x_1 + \dots + x_N = 1$ et que la matrice M est stochastique,

$$F_k^m(x) = (1 - x_k) \sum_{1 \leq j \leq N} x_j f(j) M(j, k) - x_k \sum_{1 \leq j \leq N} x_j f(j) (1 - M(j, i))$$

$$= \sum_{1 \leq j \leq N} x_j f(j) M(j, k) - x_k \sum_{1 \leq j \leq N} s_j f(j).$$

Donc, la fonction F^m coïncide avec la fonction F sur l'ensemble \mathcal{P}_N^m/m , ce qui entraîne que la condition 4 est satisfaite. Puisque les 5 conditions sont vérifiées, nous pouvons appliquer le théorème de 15.5.2 à la suite de chaînes de Markov $((\mathcal{X}_n/m)_{n \geq 0}, m \geq 1)$ et nous obtenons le résultat voulu. \square

15.6 Heuristiques

Définissons, pour $k \geq 0$

$$\rho_k = \lim_{n \rightarrow \infty} \mathbb{E} \left(\frac{1}{m} N^k(X_n) \right) = \lim_{n \rightarrow \infty} \mathbb{E} \left(\frac{1}{m} O_n(k) \right).$$

Nous nous intéressons au comportement des quantités ρ_k quand l'on considère le régime asymptotique :

$$\begin{aligned} \ell &\rightarrow +\infty, & m &\rightarrow +\infty, & q &\rightarrow 0, \\ \ell q &\rightarrow a \in]0, +\infty[, & \frac{m}{\ell} &\rightarrow \alpha \in [0, +\infty[. \end{aligned}$$

Le résultat de la section précédente montre que le modèle de Moran converge vers le modèle d'Eigen, dans la limite de grande population. Comme nous avons vu dans la partie I, quand

$$\ell \rightarrow +\infty \quad q \rightarrow 0 \quad \ell q \rightarrow a \in]0, +\infty[,$$

le modèle d'Eigen converge vers un système d'équations limite (S_∞) avec les caractéristiques suivantes :

- Si $\sigma e^{-a} < 1$ le seul point stationnaire du système (S_∞) est $(0, 0, \dots)$ et toutes les solutions $(x(t))_{t \geq 0}$ de (S_∞) convergent vers $(0, 0, \dots)$ quand t tend vers l'infini.
- Si $\sigma e^{-a} > 1$ le système (S_∞) admet deux points stationnaires, 0 et le point $(\rho_k^*)_{k \geq 0}$ du théorème 15.3.1. Toute solution $(x(t))_{t \geq 0}$ de (S_∞) vérifiant $x_0(0) > 0$ converge vers $(\rho_k^*)_{k \geq 0}$ quand t tend vers l'infini.

Nous espérons donc que quand $\sigma e^{-a} < 1$, les quantités ρ_k convergent vers 0. Si $\sigma e^{-a} > 1$, le comportement asymptotique du processus $(O_n)_{n \geq 0}$ est plus délicat. Si la taille de la population est assez grande, ce qui est exprimé dans la condition $\alpha \psi(a) > \ln \kappa$ sur l'intervalle $[0, \ln \sigma[$, la master sequence est trouvée assez souvent et le système déterministe guide le processus $(O_n)_{n \geq 0}$ vers $(\rho_k^*)_{k \geq 0}$ assez souvent, dans ce cas les quantités ρ_k convergent vers ρ_k^* . Par contre, si la taille de la population n'est pas assez grande, la master sequence n'est pas trouvée assez souvent, et le processus $(O_n)_{n \geq 0}$ n'est pas porté près du point stationnaire $(\rho_k^*)_{k \geq 0}$ assez souvent. Il passe plutôt son temps à explorer l'espace d'états loin de la master sequence. Dans ce cas les quantités ρ_k convergent vers 0. Nous formalisons ces idées dans le prochain chapitre.

Chapitre 16

Couplage et monotonie

Cette section est un résumé de la section 7 de [10], nous reproduisons les points les plus importants pour la suite du texte, mais nous ne donnons pas les démonstrations des résultats.

16.1 Couplage

Nous construisons tous nos processus dans un grand espace de probabilité. Considérons un espace de probabilité (Ω, \mathcal{F}, P) ayant les variables aléatoires indépendantes suivantes :

- une suite de variables aléatoires J_n , $n \geq 1$, de loi uniforme à valeurs dans l'ensemble $\{1, \dots, m\}$.
- une famille de variables aléatoires $U_{n,l}$, $n \geq 1$, $1 \leq l \leq \ell$, de loi uniforme sur l'intervalle $[0, 1]$.
- une suite de variables aléatoires S_n , $n \geq 1$, de loi uniforme sur l'intervalle $[0, 1]$.

Nous posons

$$\forall n \geq 1 \quad R_n = (S_n, J_n, U_{n,1}, \dots, U_{n,\ell}).$$

Le vecteur aléatoire R_n sert à accomplir le n -ième pas de la chaîne de Markov $(X_n)_{n \geq 0}$. Par construction, la suite $(R_n)_{n \geq 1}$ est une suite de vecteurs aléatoires indépendants et identiquement distribués à valeurs dans

$$\mathcal{R} = [0, 1] \times \{1, \dots, m\} \times [0, 1]^\ell.$$

Nous définissons d'abord deux applications \mathcal{M}_H et \mathcal{S}_H pour coupler les mécanismes de mutation et de sélection.

Mutation. Nous définissons une application

$$\mathcal{M}_H : \{0, \dots, \ell\} \times [0, 1]^\ell \rightarrow \{0, \dots, \ell\}$$

en posant

$$\forall b \in \{0, \dots, \ell\} \quad \forall u_1, \dots, u_\ell \in [0, 1]^\ell$$

$$\mathcal{M}_H(b, u_1, \dots, u_\ell) = b - \sum_{k=1}^b 1_{u_k < q/(\kappa-1)} + \sum_{k=b+1}^{\ell} 1_{u_k > 1-q}.$$

L'application \mathcal{M}_H est définie de sorte que si U_1, \dots, U_ℓ sont des variables aléatoires indépendantes et uniformes sur l'intervalle $[0, 1]$, alors pour tout $b \in \{0, \dots, \ell\}$, la loi de $\mathcal{M}_H(b, U_1, \dots, U_\ell)$ est donnée par la ligne de la matrice M_H associée à b , i.e.,

$$\forall c \in \{0, \dots, \ell\} \quad P(\mathcal{M}_H(b, U_1, \dots, U_\ell) = c) = M_H(b, c).$$

Sélection pour le processus distance. Nous définissons une application

$$\mathcal{S}_H : \{0, \dots, \ell\}^m \times [0, 1] \rightarrow \{1, \dots, m\}.$$

Pour $d \in \{0, \dots, \ell\}^m$ et pour $s \in [0, 1[$, nous définissons $\mathcal{S}_H(d, s) = i$ où i est le seul indice dans $\{1, \dots, m\}$ tel que

$$\frac{f_H(d(1)) + \dots + f_H(d(i-1))}{f_H(d(1)) + \dots + f_H(d(m))} \leq s < \frac{f_H(d(1)) + \dots + f_H(d(i))}{f_H(d(1)) + \dots + f_H(d(m))}.$$

L'application \mathcal{S}_H est définie de sorte que si S est une variable aléatoire de loi uniforme sur l'intervalle $[0, 1]$, alors pour tout $d \in \{0, \dots, \ell\}^m$, la loi de $\mathcal{S}_H(d, S)$ est donnée par

$$\forall i \in \{1, \dots, m\} \quad P(\mathcal{S}_H(d, S) = i) = \frac{f_H(d(i))}{f_H(d(1)) + \dots + f_H(d(m))}.$$

Couplage pour le processus distance. Nous définissons une application

$$\Phi_H : \{0, \dots, \ell\}^m \times \mathcal{R} \rightarrow \{0, \dots, \ell\}^m$$

pour coupler le processus distance avec différents fonctions de fitness et probabilités de mutation. L'application Φ_H est définie par

$$\forall r = (s, j, u_1, \dots, u_\ell) \in \mathcal{R}, \quad \forall d \in \{0, \dots, \ell\}^m, \\ \Phi_H(d, r) = d(j \leftarrow \mathcal{M}_H(d(\mathcal{S}_H(d, s)), u_1, \dots, u_\ell)).$$

Soit $d \in \{0, \dots, \ell\}^m$ le point de départ du processus. Nous construisons le processus distance $(D_n)_{n \geq 0}$ en prenant $D_0 = d$ et

$$\forall n \geq 1 \quad D_n = \Phi_H(D_{n-1}, R_n).$$

Le processus $(D_n)_{n \geq 0}$ est une chaîne de Markov issue de d et de matrice de transition p_H .

Sélection pour le processus d'occupation. Nous définissons une application

$$\mathcal{S}_O : \mathcal{P}_{\ell+1}^m \times [0, 1] \rightarrow \{0, \dots, \ell\}.$$

Pour $o \in \mathcal{P}_{\ell+1}^m$ et $s \in [0, 1[$, nous définissons $\mathcal{S}_O(o, s) = l$ où l est le seul indice dans $\{0, \dots, \ell\}$ tel que

$$\frac{o(0)f_H(0) + \dots + o(l-1)f_H(l-1)}{o(0)f_H(0) + \dots + o(\ell)f_H(\ell)} \leq s < \frac{o(0)f_H(0) + \dots + o(l)f_H(l)}{o(0)f_H(0) + \dots + o(\ell)f_H(\ell)}.$$

L'application \mathcal{S}_O est définie de sorte que si S est une variable aléatoire uniforme sur l'intervalle $[0, 1]$, alors pour tout $o \in \mathcal{P}_{\ell+1}^m$, la loi de $\mathcal{S}_O(o, S)$ est donnée par

$$\forall l \in \{0, \dots, \ell\} \quad P(\mathcal{S}_O(o, S) = l) = \frac{o(l)f_H(l)}{o(0)f_H(0) + \dots + o(\ell)f_H(\ell)}.$$

Couplage pour le processus d'occupation

Nous définissons une application

$$\Phi_O : \mathcal{P}_{\ell+1}^m \times \mathcal{R} \rightarrow \mathcal{P}_{\ell+1}^m$$

pour coupler le processus d'occupation avec différentes conditions initiales et différents paramètres σ et q . Nous définissons l'application Φ_O comme suit.

Soit $r = (s, j, u_1, \dots, u_\ell) \in \mathcal{R}$. Soit $o \in \mathcal{P}_{\ell+1}^m$, prenons $l = \mathcal{S}_O(o, s)$ et soit k le seul indice dans $\{0, \dots, \ell\}$ tel que

$$o(0) + \dots + o(k-1) < j \leq o(0) + \dots + o(k).$$

Nous définissons l'application Φ_O par

$$\Phi_O(o, r) = o(k \rightarrow \mathcal{M}_H(l, u_1, \dots, u_\ell)).$$

Soit $o \in \mathcal{P}_{\ell+1}^m$ le point de départ du processus. Nous construisons le processus d'occupation $(O_n)_{n \geq 0}$ en posant $O_0 = o$ et

$$\forall n \geq 1 \quad O_n = \Phi_O(O_{n-1}, R_n).$$

Le processus $(O_n)_{n \geq 0}$ est une chaîne de Markov issue de o et de matrice de transition p_O .

16.2 Monotonie

L'espace $\{0, \dots, \ell\}^m$ est naturellement muni d'un ordre partiel :

$$d \leq e \iff \forall i \in \{1, \dots, m\} \quad d(i) \leq e(i).$$

Lemme 16.2.1. *L'application \mathcal{M}_H est croissante vis-à-vis de la classe de Hamming, i.e.,*

$$\begin{aligned} \forall b, c \in \{0, \dots, \ell\} \quad \forall u_1, \dots, u_\ell \in [0, 1] \\ b \leq c \implies \mathcal{M}_H(b, u_1, \dots, u_\ell) \leq \mathcal{M}_H(c, u_1, \dots, u_\ell). \end{aligned}$$

Lemme 16.2.2. *Dans le cas neutre $\sigma = 1$, l'application \mathcal{S}_H est croissante vis-à-vis de la classe de Hamming, i.e.,*

$$\begin{aligned} \forall d, e \in \{0, \dots, \ell\}^m \quad \forall s \in [0, 1] \\ d \leq e \implies d(\mathcal{S}_H(d, s)) \leq e(\mathcal{S}_H(e, s)). \end{aligned}$$

Lemme 16.2.3. *Dans le cas neutre $\sigma = 1$, l'application Φ_H est croissante vis-à-vis des distances, i.e.,*

$$\forall d, e \in \{0, \dots, \ell\}^m \quad \forall r \in \mathcal{R}, \quad d \leq e \implies \Phi_H(d, r) \leq \Phi_H(e, r).$$

Nous définissons un ordre dans l'espace \preceq dans $\mathcal{P}_{\ell+1}^m$ comme suit. Soient $\mathbf{o} = (o(0), \dots, o(\ell))$ et $\mathbf{o}' = (o'(0), \dots, o'(\ell))$ dans $\mathcal{P}_{\ell+1}^m$. Nous disons que \mathbf{o} est inférieur ou égale à \mathbf{o}' , et nous écrivons $\mathbf{o} \preceq \mathbf{o}'$, si

$$\forall l \leq \ell \quad o(0) + \dots + o(l) \leq o'(0) + \dots + o'(l).$$

Lemme 16.2.4. *Dans le cas du paysage à un pic, l'application \mathcal{S}_O est décroissante vis-à-vis de la distribution d'occupation, i.e.,*

$$\forall \mathbf{o}, \mathbf{o}' \in \mathcal{P}_{\ell+1}^m \quad \forall s \in [0, 1] \quad \mathbf{o} \preceq \mathbf{o}' \quad \Rightarrow \quad \mathcal{S}_O(\mathbf{o}, s) \geq \mathcal{S}_O(\mathbf{o}', s).$$

Lemme 16.2.5. *Dans le cas du paysage à un pic, l'application Φ_O est croissante vis-à-vis de la distribution d'occupation, i.e.,*

$$\forall \mathbf{o}, \mathbf{o}' \in \mathcal{P}_{\ell+1}^m \quad \forall r \in \mathcal{R} \quad \mathbf{o} \preceq \mathbf{o}' \quad \Rightarrow \quad \Phi_O(\mathbf{o}, r) \preceq \Phi_O(\mathbf{o}', r).$$

Corollaire 16.2.6. *Dans le cas neutre $\sigma = 1$, le processus distance $(D_n)_{n \geq 0}$ est monotone.*

Corollaire 16.2.7. *Dans le cas du paysage à un pic, le processus d'occupation $(O_n)_{n \geq 0}$ est monotone.*

Chapitre 17

Bornes stochastiques

L'objectif de cette section est de construire un processus inférieur $(O_n^\ell)_{n \geq 0}$ et un processus supérieur $(O_n^{K+1})_{n \geq 0}$ satisfaisant

$$\forall n \geq 0 \quad O_n^\ell \preceq O_n \preceq O_n^{K+1}.$$

Les processus $(O_n^\ell)_{n \geq 0}$ et $(O_n^{K+1})_{n \geq 0}$ auront la vertu d'être plus simples que le processus $(O_n)_{n \geq 0}$, tout en gardant l'information pertinente sur la dynamique des coordonnées $0, \dots, K$.

17.1 Première approche

Afin de construire les processus inférieur et supérieur, nous imitons ici l'approche suivie dans le cas du processus de Wright–Fisher (chapitre 10).

17.1.1 Processus inférieur et supérieur

Nous allons construire un processus inférieur $(O_n^\ell)_{n \geq 0}$ et un processus supérieur $(O_n^{K+1})_{n \geq 0}$ satisfaisant

$$\forall n \geq 0 \quad O_n^\ell \preceq O_n \preceq O_n^{K+1}.$$

Décrivons de manière informelle la dynamique du processus inférieur. Tant qu'il n'y a aucun individu dans les classes de Hamming $0, \dots, K$, le processus $(O_n^\ell)_{n \geq 0}$ évolue exactement comme le processus original $(O_n)_{n \geq 0}$. Dès que nous trouvons un individu dans une des classes $0, \dots, K$, tous les chromosomes dans les classes $K + 1, \dots, \ell$ sont envoyés dans la classe de Hamming ℓ . Tant qu'il y a des individus dans les classes $0, \dots, K$, si un chromosome mute vers les classes $K + 1, \dots, \ell$ il est directement envoyé dans la classe ℓ . La dynamique du processus supérieur est similaire, sauf que nous remplaçons la classe de Hamming ℓ par la classe $K + 1$. Définissons deux applications $\pi_\ell, \pi_{K+1} : \mathcal{P}_{\ell+1}^m \rightarrow \mathcal{P}_{\ell+1}^m$ en posant pour $o \in \mathcal{P}_{\ell+1}^m$ et $l \in \{0, \dots, \ell\}$

$$\pi_\ell(o)(l) = \begin{cases} o(l) & \text{si } 0 \leq l \leq K, \\ 0 & \text{si } K < l < \ell, \\ m - (o(0) + \dots + o(K)) & \text{si } l = \ell, \end{cases}$$

et de façon similaire,

$$\pi_{K+1}(o)(l) = \begin{cases} o(l) & \text{si } 0 \leq l \leq K, \\ m - (o(0) + \dots + o(K)) & \text{si } l = K + 1, \\ 0 & \text{si } K + 2 \leq l \leq \ell. \end{cases}$$

Évidemment,

$$\forall o \in \mathcal{P}_{\ell+1}^m \quad \pi_\ell(o) \preceq o \preceq \pi_{K+1}(o).$$

Notons par \mathcal{W}_K^* l'ensemble des distributions d'occupation ayant au moins un individu dans une des classes de Hamming $0, \dots, K$, c'est à dire,

$$\mathcal{W}_K^* = \{ o \in \mathcal{P}_{\ell+1}^m : o(0) + \dots + o(K) \geq 1 \},$$

et soit \mathcal{N}_K l'ensemble des distributions d'occupation n'ayant aucun individu dans les classes de Hamming $0, \dots, K$, c'est à dire,

$$\mathcal{N}_K = \{ o \in \mathcal{P}_{\ell+1}^m : o(0) + \dots + o(K) = 0 \}.$$

Soit Φ_O l'application de couplage définie dans la section 16.1. Nous définissons une application inférieure Φ_O^ℓ en prenant pour $o \in \mathcal{P}_{\ell+1}^m$ et $r \in \mathcal{R}$,

$$\Phi_O^\ell(o, r) = \begin{cases} \Phi_O(o, r) & \text{si } o \in \mathcal{N}_K \text{ et } \Phi_O(o, r) \notin \mathcal{W}_K^*, \\ \pi_\ell(\Phi_O(o, r)) & \text{si } o \in \mathcal{N}_K \text{ et } \Phi_O(o, r) \in \mathcal{W}_K^*, \\ \pi_\ell(\Phi_O(\pi_\ell(o), r)) & \text{si } o \in \mathcal{W}_K^*. \end{cases}$$

Également, nous définissons une application supérieure Φ_O^{K+1} en prenant pour $o \in \mathcal{P}_{\ell+1}^m$ et $r \in \mathcal{R}$,

$$\Phi_O^{K+1}(o, r) = \begin{cases} \Phi_O(o, r) & \text{si } o \in \mathcal{N}_K \text{ et } \Phi_O(o, r) \notin \mathcal{W}_K^*, \\ \pi_{K+1}(\Phi_O(o, r)) & \text{si } o \in \mathcal{N}_K \text{ et } \Phi_O(o, r) \in \mathcal{W}_K^*, \\ \pi_{K+1}(\Phi_O(\pi_{K+1}(o), r)) & \text{si } o \in \mathcal{W}_K^*. \end{cases}$$

Grâce au lemme 16.2.5 nous voyons que l'application Φ_O^ℓ reste en-dessous de l'application Φ_O et l'application Φ_O^{K+1} reste au-dessus de l'application Φ_O au sens de l'ordre de $\mathcal{P}_{\ell+1}^m$, i.e.,

$$\forall r \in \mathcal{R} \quad \forall o \in \mathcal{P}_{\ell+1}^m \quad \Phi_O^\ell(o, r) \preceq \Phi_O(o, r) \preceq \Phi_O^{K+1}(o, r).$$

Nous définissons un processus inférieur $(O_n^\ell)_{n \geq 0}$ et un processus supérieur $(O_n^{K+1})_{n \geq 0}$ à l'aide d'une suite i.i.d. $(R_n)_{n \geq 1}$ et des applications Φ_O^ℓ et Φ_O^{K+1} . Soit $o \in \mathcal{P}_{\ell+1}^m$ le point de départ du processus. Nous prenons $O_0^\ell = O_0^{K+1} = o$ et

$$\forall n \geq 1 \quad O_n^\ell = \Phi_O^\ell(O_{n-1}^\ell, R_n), \quad O_n^{K+1} = \Phi_O^{K+1}(O_{n-1}^{K+1}, R_n).$$

Proposition 17.1.1. *Supposons que les processus $(O_n^\ell)_{n \geq 0}$, $(O_n)_{n \geq 0}$, $(O_n^{K+1})_{n \geq 0}$, sont issus d'une même distribution d'occupation o . Nous avons alors*

$$\forall n \geq 0 \quad O_n^\ell \preceq O_n \preceq O_n^{K+1}.$$

Démonstration. Nous allons montrer l'inégalité par récurrence sur $n \in \mathbb{N}$. Pour $n = 0$ nous avons bien $O_0 = O_0^\ell = O_0^{K+1} = o$. Supposons que l'inégalité est vraie au temps $t = n \in \mathbb{N}$, c'est à dire que $O_n^\ell \preceq O_n \preceq O_n^{K+1}$. Nous avons

$$\begin{aligned} O_{n+1}^\ell &= \Phi_O^\ell(O_n^\ell, R_{n+1}), \\ O_{n+1} &= \Phi_O(O_n, R_{n+1}), \\ O_{n+1}^{K+1} &= \Phi_O^{K+1}(O_n^{K+1}, R_{n+1}). \end{aligned}$$

Par le lemme 16.2.5 et l'hypothèse de récurrence nous obtenons

$$\Phi_O(O_n^\ell, R_{n+1}) \preceq \Phi_O(O_n, R_{n+1}) \preceq \Phi_O(O_n^{K+1}, R_{n+1}).$$

L'application Φ_O^ℓ est en-dessous de l'application Φ_O et l'application Φ_O^{K+1} est au-dessus de l'application Φ_O , donc,

$$\begin{aligned} \Phi_O^\ell(O_n^\ell, R_{n+1}) &\preceq \Phi_O(O_n^\ell, R_{n+1}), \\ \Phi_O(O_n^{K+1}, R_{n+1}) &\preceq \Phi_O^{K+1}(O_n^{K+1}, R_{n+1}). \end{aligned}$$

Nous combinons les inégalités précédentes et nous obtenons

$$O_{n+1}^\ell \preceq O_{n+1} \preceq O_{n+1}^{K+1}.$$

□

17.1.2 Dynamique des processus bornants

Ensuite nous étudions les dynamiques des processus $(O_n^\ell)_{n \geq 0}$ et $(O_n^{K+1})_{n \geq 0}$ dans \mathcal{W}_K^* . Les calculs sont les mêmes pour les deux processus. Soit θ un symbole représentant $K+1$ ou ℓ et notons par $(O_n^\theta)_{n \geq 0}$ le processus correspondant. Pour le processus $(O_n^\theta)_{n \geq 0}$, les états dans l'ensemble

$$\mathcal{T}_K^\theta = \{ o \in \mathcal{P}_{\ell+1}^m : o(0) + \dots + o(K) \geq 1 \text{ et } o(0) + \dots + o(K) + o(\theta) < m \}$$

sont transitoires, et les états dans $\mathcal{N}_K \cup (\mathcal{W}_K^* \setminus \mathcal{T}_K^\theta)$ forment une classe de récurrence. Nous observons le mécanisme de transition du processus restreint à $\mathcal{W}_K^* \setminus \mathcal{T}_K^\theta$. Puisque

$$\mathcal{W}_K^* \setminus \mathcal{T}_K^\theta = \{ o \in \mathcal{P}_{\ell+1}^m : o(0) + \dots + o(K) \geq 1 \text{ et } o(0) + \dots + o(K) + o(\theta) = m \},$$

un état dans $\mathcal{W}_K^* \setminus \mathcal{T}_K^\theta$ est totalement déterminé par les nombres d'occupation des classes $0, \dots, K$. Pour $i \in \{0, \dots, K\}$, soit o_i^θ la distribution d'occupation ayant un chromosome dans la classe de Hamming i et $m-1$ chromosomes dans la classe θ :

$$\forall l \in \{0, \dots, \ell\} \quad o_i^\theta(l) = \begin{cases} 1 & \text{si } l = i, \\ m-1 & \text{si } l = \theta, \\ 0 & \text{sinon.} \end{cases}$$

Le processus $(O_n^\theta)_{n \geq 0}$ entre toujours dans l'ensemble $\mathcal{W}_K^* \setminus \mathcal{T}_K^\theta$ par un des états o_i^θ , $i \in \{0, \dots, K\}$. Nous notons par $(e_i)_{0 \leq i \leq K}$ la base canonique de \mathbb{R}^{K+1} indexée par $0, \dots, K$ au lieu de $1, \dots, K+1$. Les seules transitions possibles pour les classes de Hamming $\{0, \dots, K\}$ du processus $(O_n^\theta)_{n \geq 0}$ issu d'un point de $\mathcal{W}_K^* \setminus \mathcal{T}_K^\theta$, sont

$$(o(0), \dots, o(K)) \longrightarrow (o(0), \dots, o(K)) - e_i \quad \text{si } 1 \leq o(i), \quad 0 \leq i \leq K,$$

$$(o(0), \dots, o(K)) \longrightarrow (o(0), \dots, o(K)) + e_i \quad \text{si } o(0) + \dots + o(K) < m, \quad 0 \leq i \leq K,$$

$$(o(0), \dots, o(K)) \longrightarrow (o(0), \dots, o(K)) - e_i + e_j \quad \text{si } 1 \leq o(i), \quad 0 \leq i, j \leq K, \quad i \neq j.$$

Soit o_{exit}^θ la distribution d'occupation ayant m chromosomes dans la classe de Hamming θ :

$$\forall l \in \{0, \dots, \ell\} \quad o_{\text{exit}}^\theta(l) = \begin{cases} m & \text{si } l = \theta, \\ 0 & \text{sinon.} \end{cases}$$

Le processus $(O_n^\theta)_{n \geq 0}$ sort toujours de l'ensemble $\mathcal{W}_K^* \setminus \mathcal{J}_K^\theta$ par l'état o_{exit}^θ . Ainsi, dès que le processus $(O_n^\theta)_{n \geq 0}$ est issu d'un point de $\mathcal{W}_K^* \setminus \mathcal{J}_K^\theta$, la dynamique de $(O_n^\theta(0), \dots, O_n^\theta(K))_{n \geq 0}$ est celle d'une chaîne de Markov, jusqu'au temps de sortie de $\mathcal{W}_K^* \setminus \mathcal{J}_K^\theta$. Soit $(Z_n^\theta)_{n \geq 0}$ une chaîne de Markov à valeurs dans

$$\mathbb{D}^{K+1} = \{z \in \mathbb{N}^{K+1} : z_0 + \dots + z_K \leq m\},$$

ayant pour matrice de transition :

- Pour $0 \leq i \leq K$ et $0 < z_0 + \dots + z_K < m$,

$$\begin{aligned} P(Z_{t+1}^\theta = z + e_i \mid Z_t^\theta = z) &= \frac{m - \sum_{l=0}^K z_l}{m((\sigma - 1)z_0 + m)} \\ &\times \left(\sigma z_0 M_H(0, i) + \sum_{l=1}^K z_l M_H(l, i) + \left(m - \sum_{l=0}^K z_l \right) M_H(\theta, i) \right). \end{aligned}$$

- Pour $0 \leq i \leq K$ et $1 \leq z_i$,

$$\begin{aligned} P(Z_{t+1}^\theta = z - e_i \mid Z_t^\theta = z) &= \\ &\frac{z_i}{m((\sigma - 1)z_0 + m)} \times \left(\sigma z_0 \left(1 - \sum_{h=0}^K M_H(0, h) \right) \right. \\ &\left. + \sum_{l=1}^K z_l \left(1 - \sum_{h=0}^K M_H(l, h) \right) + \left(m - \sum_{l=0}^K z_l \right) \left(1 - \sum_{h=0}^K M_H(\theta, h) \right) \right). \end{aligned}$$

- Pour $0 \leq i, j \leq K$, $i \neq j$ et $1 \leq z_i$,

$$\begin{aligned} P(Z_{t+1}^\theta = z - e_i + e_j \mid Z_t^\theta = z) &= \frac{z_i}{m((\sigma - 1)z_0 + m)} \\ &\times \left(\sigma z_0 M_H(0, j) + \sum_{l=1}^K z_l M_H(l, j) + \left(m - \sum_{l=0}^K z_l \right) M_H(\theta, j) \right). \end{aligned}$$

Nous avons intentionnellement laissé indéfinies les probabilités de transition partant de l'état $(0, \dots, 0)$. Elles ne jouent aucun rôle pour répliquer la dynamique du processus $(O_n^\theta)_{n \geq 0}$ dans $\mathcal{W}_K^* \setminus \mathcal{T}_K^\theta$. Nous préciserons les probabilités de transition partant de l'état $(0, \dots, 0)$ un peu plus tard, en les choisissant de la manière la plus opportune. Les autres coefficients non diagonaux de la matrice de transition sont nuls. Les coefficients diagonaux sont choisis de sorte que la matrice soit stochastique, i.e., la somme de chaque ligne vaut 1.

17.1.3 Bornes sur la mesure invariante

Notons par μ_O^ℓ , μ_O , μ_O^{K+1} les mesures de probabilité invariantes des processus $(O_n^\ell)_{n \geq 0}$, $(O_n)_{n \geq 0}$, $(O_n^{K+1})_{n \geq 0}$. Soit ν_K la mesure image de μ_O par l'application

$$o \in \mathcal{P}_{\ell+1}^m \mapsto \frac{o(0) + \dots + o(K)}{m} \in [0, 1].$$

Pour toute fonction $f : [0, 1] \rightarrow \mathbb{R}$,

$$\begin{aligned} \int_{[0,1]} f \, d\nu_K &= \int_{\mathcal{P}_{\ell+1}^m} f\left(\frac{o(0) + \dots + o(K)}{m}\right) d\mu_O(o) \\ &= \lim_{n \rightarrow \infty} \mathbb{E}\left(f\left(\frac{O_n(0) + \dots + O_n(K)}{m}\right)\right). \end{aligned}$$

Fixons $f : [0, 1] \rightarrow \mathbb{R}$ une fonction croissante telle que $f(0) = 0$. Nous avons —grâce à la proposition 17.1.1— les inégalités suivantes. Pour tout $n \geq 0$,

$$\begin{aligned} f\left(\frac{O_n^\ell(0) + \dots + O_n^\ell(K)}{m}\right) \\ \leq f\left(\frac{O_n(0) + \dots + O_n(K)}{m}\right) \leq \\ f\left(\frac{O_n^{K+1}(0) + \dots + O_n^{K+1}(K)}{m}\right). \end{aligned}$$

Nous prenons l'espérance, nous envoyons n vers l' ∞ , et nous obtenons

$$\begin{aligned} \int_{\mathcal{P}_{\ell+1}^m} f\left(\frac{o(0) + \cdots + o(K)}{m}\right) d\mu_O^\ell(o) \\ \leq \int_{[0,1]} f d\nu_K \leq \\ \int_{\mathcal{P}_{\ell+1}^m} f\left(\frac{o(0) + \cdots + o(K)}{m}\right) d\mu_O^{K+1}(o). \end{aligned}$$

Nous cherchons ensuite des estimées sur les intégrales ci-dessus. La stratégie est la même pour les intégrales inférieure et supérieure. Soit θ un symbole représentant $K+1$ ou ℓ . Nous étudions la mesure de probabilité invariante μ_O^θ . Pour le processus $(O_n^\theta)_{n \geq 0}$, les états dans \mathcal{T}_K^θ sont transitoires et les populations dans $\mathcal{N}_K \cup (\mathcal{W}_K^* \setminus \mathcal{T}_K^\theta)$ forment une classe de récurrence. Nous appliquons le résultat de renouvellement (proposition 10.3.1) au processus $(O_n^\theta)_{n \geq 0}$ restreint à $\mathcal{N}_K \cup (\mathcal{W}_K^* \setminus \mathcal{T}_K^\theta)$, à l'ensemble $\mathcal{W}_K^* \setminus \mathcal{T}_K^\theta$, à la distribution d'occupation o_{exit}^θ et à la fonction $o \mapsto f((o(0) + \cdots + o(K))/m)$. Soient

$$\begin{aligned} \tau_K^* &= \inf \{ n \geq 0 : O_n^\theta \in \mathcal{W}_K^* \}, \\ \tau_K &= \inf \{ n \geq \tau_K^* : O_n^\theta = o_{\text{exit}}^\theta \}. \end{aligned}$$

Nous avons alors

$$\begin{aligned} \int_{\mathcal{P}_{\ell+1}^m} f\left(\frac{o(0) + \cdots + o(K)}{m}\right) d\mu_O^\theta(o) = \\ \frac{\mathbb{E}\left(\sum_{n=\tau_K^*}^{\tau_K-1} f\left(\frac{O_n^\theta(0) + \cdots + O_n^\theta(K)}{m}\right) \mid O_0^\theta = o_{\text{exit}}^\theta\right)}{\mathbb{E}(\tau_K \mid O_0^\theta = o_{\text{exit}}^\theta)}. \end{aligned}$$

Soit π_O la distribution d'entrée du processus $(O_n^\theta)_{n \geq 0}$ dans \mathcal{W}_K^* lorsque $O_0^\theta = o_{\text{exit}}^\theta$, i.e.,

$$\forall o \in \mathcal{P}_{\ell+1}^m \quad \pi_O(o) = \mathbb{P}(O_{\tau_K^*}^\theta = o \mid O_0^\theta = o_{\text{exit}}^\theta).$$

Puisque le support de π_O est $\{o_i : 0 \leq i \leq K\}$, nous pouvons définir une distribution π sur \mathbb{D}^{K+1} comme suit :

$$\begin{aligned} \pi(e_i) &= \pi_O(o_i) && \text{pour } 0 \leq i \leq K, \\ \pi(z) &= 0 && \text{pour } z \neq e_0, \dots, e_K. \end{aligned}$$

Lorsque le processus $(O_n^\theta)_{n \geq 0}$ est dans l'ensemble $\mathcal{W}_K^* \setminus \mathcal{T}_K^\theta$, la dynamique du processus $(O_n^\theta(0), \dots, O_n^\theta(K))_{n \geq 0}$ est la même que celle de la chaîne de Markov $(Z_n^\theta)_{n \geq 0}$ définie à la fin de la section 17.1.2. Supposons que $(Z_n^\theta)_{n \geq 0}$ a pour distribution initiale la distribution π , et notons par P_π et E_π les probabilités et espérances par rapport à cette condition initiale. Soit τ_0 le temps d'atteinte de l'origine 0 dans \mathbb{D}^{K+1} , i.e.,

$$\tau_0 = \inf \{ n \geq 0 : Z_n^\theta = 0 \}.$$

La loi du temps τ_0 est la même que la loi de $\tau_K - \tau_K^*$, i.e., la loi du temps de sortie de $\mathcal{W}_K^* \setminus \mathcal{T}_K^\theta$ après τ_K^* . Nous en déduisons que le processus $(Z_n^\theta, 0 \leq n \leq \tau_0 - 1)$ et $((O_n^\theta(0), \dots, O_n^\theta(K)), \tau_K^* \leq n \leq \tau_K - 1)$ ont même loi. Donc,

$$\begin{aligned} \mathbb{E} \left(\sum_{n=\tau_K^*}^{\tau_K-1} f \left(\frac{O_n^\theta(0) + \dots + O_n^\theta(K)}{m} \right) \mid O_0^\theta = o_{\text{exit}}^\theta \right) = \\ \mathbb{E}_\pi \left(\sum_{n=0}^{\tau_0-1} f \left(\frac{Z_n^\theta(0) + \dots + Z_n^\theta(K)}{m} \right) \right). \end{aligned}$$

De plus, grâce à la propriété de Markov,

$$\mathbb{E}(\tau_K - \tau_K^* \mid O_0^\theta = o_{\text{exit}}^\theta) = \mathbb{E}_\pi(\tau_0).$$

La formule pour la mesure de probabilité invariante μ_O^θ se met alors sous la forme

$$\begin{aligned} \int_{\mathcal{P}_{\ell+1}^m} f \left(\frac{o(0) + \dots + o(K)}{m} \right) d\mu_O^\theta(o) = \\ \frac{\mathbb{E}_\pi \left(\sum_{n=0}^{\tau_0-1} f \left(\frac{Z_n^\theta(0) + \dots + Z_n^\theta(K)}{m} \right) \right)}{\mathbb{E}(\tau_K^* \mid O_0^\theta = o_{\text{exit}}^\theta) + \mathbb{E}_\pi(\tau_0)}. \end{aligned}$$

Pour pouvoir réécrire cette formule en fonction de la mesure de probabilité invariante du processus $(Z_n^\theta)_{n \geq 0}$, nous allons appliquer le résultat de renouvellement (proposition 10.3.1) au processus $(Z_n^\theta)_{n \geq 0}$, à l'ensemble $\mathbb{D}^{K+1} \setminus \{0\}$, au point 0 et à l'application $z \mapsto f((z_0 + \dots + z_K)/m)$. Définissons

$$\begin{aligned} \tau_1 &= \inf \{ n \geq 0 : Z_n^\theta \neq 0 \}, \\ \tau'_0 &= \inf \{ n \geq \tau_1 : Z_n^\theta = 0 \}, \end{aligned}$$

et soit ν^θ la mesure de probabilité invariante du processus $(Z_n^\theta)_{n \geq 0}$. Nous avons

$$\int_{\mathbb{D}^{K+1}} f\left(\frac{z_0 + \dots + z_K}{m}\right) d\nu^\theta(z) = \frac{\mathbb{E}\left(\sum_{n=\tau_1}^{\tau'_0-1} f\left(\frac{Z_n^\theta(0) + \dots + Z_n^\theta(K)}{m}\right) \mid Z_0^\theta = 0\right)}{\mathbb{E}(\tau'_0 \mid Z_0^\theta = 0)}.$$

Nous n'avons pas encore choisi les probabilités de transition de la chaîne de Markov $(Z_n^\theta)_{n \geq 0}$ partant de 0. Un choix commode est le suivant :

$$\forall z \in \mathbb{D}^{K+1} \quad \mathbb{P}(Z_{n+1}^\theta = z \mid Z_n^\theta = 0) = \pi(z).$$

Avec ce choix, vu que $\pi(0) = 0$, grâce à la propriété de Markov, nous pouvons écrire la formule pour la mesure de probabilité invariante ν^θ de la manière suivante :

$$\int_{\mathbb{D}^{K+1}} f\left(\frac{z_0 + \dots + z_K}{m}\right) d\nu^\theta(z) = \frac{\mathbb{E}_\pi\left(\sum_{n=0}^{\tau_0-1} f\left(\frac{Z_n^\theta(0) + \dots + Z_n^\theta(K)}{m}\right) ds\right)}{1 + \mathbb{E}_\pi(\tau_0)}.$$

Nous pouvons finalement exprimer la mesure de probabilité invariante μ_O^θ en fonction de ν^θ :

$$\int_{\mathcal{P}_{\ell+1}^m} f\left(\frac{o(0) + \dots + o(K)}{m}\right) d\mu_O^\theta(o) = \frac{1 + \mathbb{E}_\pi(\tau_0)}{\mathbb{E}(\tau_K^* \mid O_0^\theta = o_{\text{exit}}^\theta) + \mathbb{E}_\pi(\tau_0)} \int_{\mathbb{D}^{K+1}} f\left(\frac{z_0 + \dots + z_K}{m}\right) d\nu^\theta(z). \quad (\mathcal{J}_1)$$

Il s'agit maintenant d'estimer chacun des termes de droite dans l'identité \mathcal{J}_1 . L'enjeu principal est d'estimer l'intégrale. Le processus $(Z_n^\theta)_{n \geq 0}$ est une chaîne de Markov, *dépendante de la densité*, c'est-à-dire que, quand la taille de la population tend vers l'infini, la dynamique de la chaîne de Markov approche celle d'une trajectoire déterministe. Cette trajectoire déterministe est une solution (à un changement de temps près, dû à notre choix du modèle de Moran normalisé) du système d'Eigen (cf. la section 15.5). Asymptotiquement, le modèle d'Eigen acquiert une forme triangulaire :

$$x'_k(t) = \sum_{0 \leq h \leq k} x_h(t) f_H(h) M_H(h, k) - x_k(t) \left((\sigma - 1)x_0(t) + 1 \right), \quad k \geq 0.$$

Il s'est avéré assez difficile de travailler avec les processus $(Z_n^\theta)_{n \geq 0}$ définis ci-dessus, car il faut considérer toutes les coordonnées en même temps. C'est pour cela que, inspirés de la forme triangulaire du système, nous avons développé des nouveaux processus inférieur et supérieur, avec une propriété supplémentaire : la coordonnée k du processus ne dépend pas de la dynamique des coordonnées supérieures à k . Ceci nous permettra d'estimer l'intégrale ci-dessus par un argument de récurrence sur les coordonnées. Nous allons tout de même nous servir plus tard de l'identité \mathcal{J}_1 , pour montrer que, dans le cas $\sigma e^{-\alpha} < 1$, la mesure ν^θ tend vers la masse de Dirac en 0. Dans le reste de la section, nous construisons les processus inférieur et supérieur raffinés afin de pouvoir mieux traiter le cas $\sigma e^{-\alpha} > 1$.

17.2 Le processus inférieur

Nous construisons d'abord le processus inférieur $(O_n^\ell)_{n \geq 0}$. Pour commencer, décrivons de manière informelle la dynamique du processus inférieur. Lorsque la master sequence n'est pas présente dans la population, le processus $(O_n^\ell)_{n \geq 0}$ évolue exactement comme le processus original $(O_n)_{n \geq 0}$. Dès qu'une master sequence apparaît, tous les chromosomes dans les classes $K+1, \dots, \ell$ sont envoyés dans la classe de Hamming ℓ . Tant que la master sequence est présente dans la population, tout chromosome qui mute vers les classes $K+1, \dots, \ell$ est directement envoyé dans la classe ℓ . De plus, tout chromosome qui mute d'une classe de Hamming vers une classe inférieure est aussi envoyé dans la classe ℓ . Pour rendre rigoureuse cette construction, nous allons d'abord modifier l'application Φ_O définie dans la section 16 et construire une application Φ_O^ℓ . Nous réalisons ensuite le couplage grâce à une nouvelle application

$$\underline{\Phi}_O : \mathcal{P}_{\ell+1}^m \times \mathcal{R} \longrightarrow \mathcal{P}_{\ell+1}^m.$$

Soit $r = (s, i, j, u_1, \dots, u_\ell) \in \mathcal{R}$. Soit $o \in \mathcal{P}_{\ell+1}^m$. Prenons $l = S_O(o, s)$ et soit k le seul indice dans $\{0, \dots, \ell\}$ tel que

$$o(0) + \dots + o(k-1) < j \leq o(0) + \dots + o(k).$$

L'application $\underline{\Phi}_O$ est définie par :

$$\underline{\Phi}_O(o, r) = \begin{cases} o(k \rightarrow \ell) & \text{si } \mathcal{M}_H(l, u_1, \dots, u_\ell) < l, \\ o(k \rightarrow \mathcal{M}_H(l, u_1, \dots, u_\ell)) & \text{sinon.} \end{cases}$$

À partir de cette construction, il est clair que

$$\forall r \in \mathcal{R} \quad \forall o \in \mathcal{P}_{\ell+1}^m \quad \underline{\Phi}_O(o, r) \preceq \Phi_O(o, r).$$

Soit $\pi_\ell : \mathcal{P}_{\ell+1}^m \rightarrow \mathcal{P}_{\ell+1}^m$ l'application définie par : pour tout $o \in \mathcal{P}_{\ell+1}^m$ et $l \in \{0, \dots, \ell\}$,

$$\pi_\ell(o)(l) = \begin{cases} o(l) & \text{si } 0 \leq l \leq K, \\ 0 & \text{si } K < l < \ell, \\ m - (o(0) + \dots + o(K)) & \text{si } l = \ell. \end{cases}$$

Cette application vérifie

$$\forall o \in \mathcal{P}_{\ell+1}^m \quad \pi_\ell(o) \preceq o.$$

Notons \mathcal{W}^* l'ensemble des distributions d'occupation ayant au moins une master sequence, i.e.,

$$\mathcal{W}^* = \{ o \in \mathcal{P}_{\ell+1}^m : o(0) \geq 1 \},$$

et notons \mathcal{N} l'ensemble des distributions d'occupation n'ayant aucune master sequence, i.e.,

$$\mathcal{N} = \{ o \in \mathcal{P}_{\ell+1}^m : o(0) = 0 \}.$$

Définissons

$$o_{\text{exit}}^\ell = (0, \dots, 0, m), \quad o_{\text{enter}}^\ell = (1, 0, \dots, 0, m-1).$$

Nous définissons une application inférieure Φ_O^ℓ en prenant pour $o \in \mathcal{P}_{\ell+1}^m$ et $r \in \mathcal{R}$,

$$\Phi_O^\ell(o, r) = \begin{cases} \Phi_O(o, r) & \text{si } o \in \mathcal{N} \text{ et } \Phi_O(o, r) \notin \mathcal{W}^* \\ o_{\text{enter}}^\ell & \text{si } o \in \mathcal{N} \text{ et } \Phi_O(o, r) \in \mathcal{W}^* \\ \pi_\ell(\underline{\Phi}_O(\pi_\ell(o), r)) & \text{si } o \in \mathcal{W}^* \text{ et } \underline{\Phi}_O(\pi_\ell(o), r) \notin \mathcal{N} \\ o_{\text{exit}}^\ell & \text{si } o \in \mathcal{W}^* \text{ et } \underline{\Phi}_O(\pi_\ell(o), r) \in \mathcal{N} \end{cases}$$

Vérifions que l'application Φ_O^ℓ reste en-dessous de l'application Φ_O au sens de l'ordre de $\mathcal{P}_{\ell+1}^m$.

Proposition 17.2.1. *Pour tout $r \in \mathcal{R}$ et pour tout $o \in \mathcal{P}_{\ell+1}^m$,*

$$\Phi_O^\ell(o, r) \preceq \Phi_O(o, r).$$

Démonstration. Soient $r \in \mathcal{R}$ et $o \in \mathcal{P}_{\ell+1}^m$. Nous considérons les cas suivants :

- Si $o \in \mathcal{N}$ et $\Phi_O(o, r) \notin \mathcal{W}^*$, alors

$$\Phi_O^\ell(o, r) = \Phi_O(o, r).$$

- Si $o \in \mathcal{N}$ et $\Phi_O(o, r) \in \mathcal{W}^*$, nous avons

$$\Phi_O^\ell(o, r) = o_{\text{enter}}^\ell \quad \text{et} \quad \Phi_O(o, r)(0) = 1,$$

nous déduisons l'inégalité voulue du fait suivant :

$$\forall o \in \mathcal{P}_{\ell+1}^m \text{ avec } o(0) = 1, \quad o_{\text{enter}}^\ell \preceq o.$$

- Si $o \in \mathcal{W}^*$ et $\underline{\Phi}_O(\pi_\ell(o), r) \notin \mathcal{N}$, comme l'application $\underline{\Phi}_O$ reste en dessous de l'application Φ_O ,

$$\underline{\Phi}_O(\pi_\ell(o), r) \preceq \Phi_O(\pi_\ell(o), r).$$

Comme de plus $\pi_\ell(o) \preceq o$, et que Φ_O est monotone, nous concluons que

$$\pi_\ell\left(\underline{\Phi}_O(\pi_\ell(o), r)\right) \preceq \underline{\Phi}_O(\pi_\ell(o), r) \preceq \Phi_O(\pi_\ell(o), r) \preceq \Phi_O(o, r).$$

- Si $o \in \mathcal{W}^*$ et $\underline{\Phi}_O(\pi_\ell(o), r) \in \mathcal{N}$, nous avons $\Phi_O^\ell(o) = o_{\text{exit}}^\ell$ et nous remarquons que pour tout $o \in \mathcal{P}_{\ell+1}^m$, $o_{\text{exit}}^\ell \preceq o$.

Nous concluons donc que pour tout $o \in \mathcal{P}_{\ell+1}^m$ et pour tout $r \in \mathcal{R}$,

$$\Phi_O^\ell(o, r) \preceq \Phi_O(o, r).$$

□

Nous définissons un processus inférieur $(O_t^\ell)_{t \geq 0}$ à l'aide d'une suite de vecteurs aléatoires i.i.d. $(R_n)_{n \geq 1}$ et de l'application Φ_O^ℓ . Soit $o \in \mathcal{P}_{\ell+1}^m$ le point de départ du processus. Nous prenons $O_0^\ell = o$ et

$$\forall n \geq 1 \quad O_n^\ell = \Phi_O^\ell(O_{n-1}^\ell, R_n).$$

Proposition 17.2.2. *Supposons que les processus $(O_n^\ell)_{n \geq 0}$, $(O_n)_{n \geq 0}$, sont issus d'une même distribution d'occupation o . Nous avons alors*

$$\forall n \geq 0 \quad O_n^\ell \preceq O_n.$$

La démonstration est similaire à celle de la proposition 17.1.1.

17.3 Dynamique du processus inférieur

Nous étudions maintenant la dynamique du processus inférieur $(O_n^\ell)_{n \geq 0}$ dans \mathcal{W}^* . Pour le processus $(O_n^\ell)_{n \geq 0}$, les états de l'ensemble

$$\mathcal{T}^\ell = \{ o \in \mathcal{P}_{\ell+1}^m : o(0) \geq 1 \text{ et } o(0) + \dots + o(K) + o(\ell) < m \}$$

sont transitoires, et les états dans $\mathcal{N} \cup (\mathcal{W}^* \setminus \mathcal{T}^\ell)$ forment une classe de récurrence. Nous observons le mécanisme de transition du processus restreinte à $\mathcal{W}^* \setminus \mathcal{T}^\ell$. Puisque

$$\mathcal{W}^* \setminus \mathcal{T}^\ell = \{ o \in \mathcal{P}_{\ell+1}^m : o(0) \geq 1 \text{ et } o(0) + \dots + o(K) + o(\ell) = m \},$$

un état dans $\mathcal{W}^* \setminus \mathcal{T}^\ell$ est totalement déterminé par les nombres d'occupation des classes $0, \dots, K$. Le processus $(O_n^\ell)_{n \geq 0}$ entre toujours dans l'ensemble $\mathcal{W}^* \setminus \mathcal{T}^\ell$ par l'état o_{enter}^ℓ . Pour $i \in \{0, \dots, K\}$, notons par $(e_i)_{0 \leq i \leq K}$ la base canonique de \mathbb{R}^{K+1} indexée par $0, \dots, K$ au lieu de $1, \dots, K+1$. Les seules transitions possibles pour les classes de Hamming $0, \dots, K$ du processus $(O_n^\ell)_{n \geq 0}$ issu d'un point de $\mathcal{W}^* \setminus \mathcal{T}^\ell$ sont

$$(o(0), \dots, o(K)) \longrightarrow (o(0), \dots, o(K)) - e_i \quad \text{si } 1 \leq o(i), \quad 0 \leq i \leq K,$$

$$(o(0), \dots, o(K)) \longrightarrow (o(0), \dots, o(K)) + e_i \quad \text{si } o(0) + \dots + o(K) \leq m - 1, \quad 0 \leq i \leq K,$$

$$(o(0), \dots, o(K)) \longrightarrow (o(0), \dots, o(K)) - e_i + e_j \quad \text{si } 1 \leq o(i), \quad 0 \leq i, j \leq K, \quad i \neq j.$$

Le processus $(O_n^\ell)_{n \geq 0}$ sort toujours de l'ensemble $\mathcal{W}^* \setminus \mathcal{T}^\ell$ par l'état o_{exit}^ℓ . Si le processus $(O_n^\ell)_{n \geq 0}$ est issu d'un point de $\mathcal{W}^* \setminus \mathcal{T}^\ell$, la dynamique du processus $(O_n^\ell(0), \dots, O_n^\ell(K))_{n \geq 0}$ est celle d'une chaîne de Markov, jusqu'au temps de sortie de $\mathcal{W}^* \setminus \mathcal{T}^\ell$. Soit $(Z_n^\ell)_{n \geq 0}$ une chaîne de Markov à valeurs dans

$$\mathbb{D}^{K+1} = \{ z \in \mathbb{N}^{K+1} : z_0 + \dots + z_K \leq m \},$$

issue du point $z^\ell = (1, 0, \dots, 0)$, ayant pour matrice de transition :

- Pour $0 \leq i \leq K$ et $0 < z_0 + \dots + z_K < m$,

$$p(z, z + e_i) = \frac{m - \sum_{l=0}^K z_l}{m((\sigma - 1)z_0 + m)} \times \left(\sigma z_0 M_H(0, i) + \sum_{l=1}^i z_l M_H(l, i) \right).$$

- Pour $0 \leq i \leq K$ et $1 \leq z_i$,

$$p(z, z - e_i) = \frac{z_i}{m((\sigma - 1)z_0 + m)} \times \left(\sigma z_0 \left(1 - \sum_{h=0}^K M_H(0, h) \right) + \sum_{l=1}^K z_l \left(1 - \sum_{h=l}^K M_H(l, h) \right) + m - \sum_{l=0}^K z_l \right).$$

- Pour $0 \leq i, j \leq K$, $i \neq j$ et $1 \leq z_i$,

$$p(z, z - e_i + e_j) = \frac{z_i}{m((\sigma - 1)z_0 + m)} \times \left(\sigma z_0 M_H(0, j) + \sum_{l=1}^j z_l M_H(l, j) \right).$$

Les autres coefficients non diagonaux de la matrice de transition sont nuls. Les coefficients diagonaux sont choisis de sorte que la matrice soit stochastique, i.e., la somme de chaque ligne vaut 1.

Notre but est de répliquer la dynamique du processus $(O_n^\ell(0), \dots, O_n^\ell(K))_{n \geq 0}$ dans l'ensemble $\mathcal{W}^* \setminus \mathcal{T}^\ell$. Dans cette dynamique, les probabilités de transition

$$\begin{aligned} p(z, z') & \text{ avec } z_0 = 0, \\ p(z, z') & \text{ avec } z_0 = 1 \text{ et } z'_0 = 0. \end{aligned}$$

ne jouent aucun rôle. Tant que nous ne modifions pas les autres probabilités de transition, nous pouvons choisir ces probabilités comme nous voulons. Nous allons le faire de la manière la plus opportune. Si la coordonnée 0 passe de 1 à 0, nous obligeons le processus à sauter directement sur $(0, \dots, 0)$. De plus, quand le processus est en $(0, \dots, 0)$ il saute vers $(1, 0, \dots, 0)$ avec probabilité 1. Ainsi, nous obtenons une nouvelle matrice de transition p^ℓ :

- Pour $z, z' \in \mathbb{D}^{K+1}$ avec $z_0 = 1$ et $z'_0 = 0$,

$$\begin{aligned} p^\ell(z, (0, \dots, 0)) &= p(z, z - e_0) + \sum_{i=1}^K p(z, z - e_0 + e_i), \\ p^\ell(z, z') &= 0 \quad \text{si } z' \neq (0, \dots, 0). \end{aligned}$$

- Pour $z, z' \in \mathbb{D}^{K+1}$ avec $z_0 = 0$,

$$\begin{aligned} p^\ell(z, (1, 0, \dots, 0)) &= 1, \\ p^\ell(z, z') &= 0 \quad \text{si } z' \neq (1, 0, \dots, 0). \end{aligned}$$

et $p^\ell(z, z') = p(z, z')$ pour les $z, z' \in \mathbb{D}^{K+1}$ restants.

17.4 Le processus supérieur

Nous construisons maintenant un processus supérieur $(O_n^{K+1})_{n \geq 0}$. Pour commencer, décrivons de manière informelle la dynamique du processus supérieur. Lorsque la master sequence n'est pas présente dans la population, le processus $(O_n^{K+1})_{n \geq 0}$ évolue exactement comme le processus original $(O_n)_{n \geq 0}$. Dès qu'une master sequence apparaît, tous les chromosomes dans les classes $K+1, \dots, \ell$ sont envoyés dans la classe de Hamming $K+1$. Lorsque la master sequence est présente dans la population, tout chromosome qui mute vers les classes $K+1, \dots, \ell$ est directement envoyé dans la classe $K+1$. De plus, pour tout $c < b$ la probabilité de mutation de la classe de Hamming b vers la classe de Hamming c est fixée égale à $M_H(c+1, c)$. Pour rendre rigoureuse cette construction, nous allons d'abord modifier les probabilités de mutation $(M_H(b, c), 0 \leq b, c \leq \ell)$ pour définir les probabilités de mutations associées au processus $(O_n^{K+1})_{n \geq 0}$. Nous définissons pour tout $b \in \{0, \dots, K+1\}$ et $c \in \{0, \dots, \ell\}$,

$$M_H^{K+1}(b, c) = \begin{cases} M_H(c+1, c) & \text{si } 0 \leq c < b \\ M_H(b, c) & \text{si } b \leq c \leq K \\ 0 & \text{si } c \in \{K+2, \dots, \ell\} \end{cases}$$

Le coefficient $M_H^{K+1}(b, K+1)$ est ajusté pour que la somme de ligne fasse 1, i.e., nous posons, pour $b \in \{0, \dots, K+1\}$

$$\begin{aligned} M_H^{K+1}(b, K+1) &= 1 - \sum_{h=0}^K M_H^{K+1}(b, h) = \\ &= 1 - \sum_{h=0}^{b-1} M_H(h+1, h) - \sum_{h=b}^K M_H(b, h). \end{aligned}$$

Pour tout $b \in \{K+2, \dots, \ell\}$,

$$\forall c \in \{0, \dots, \ell\} \quad M_H^{K+1}(b, c) = M_H(b, c).$$

Nous devons vérifier que $(M_H^{K+1}(b, c), 0 \leq b, c \leq \ell)$ est une matrice stochastique, i.e., que toutes les entrées de la matrice sont positives ou nulles et que la somme de chaque ligne vaut 1. Puisque $(M_H(b, c), 0 \leq b, c \leq \ell)$ est déjà une matrice stochastique, ces conditions sont satisfaites pour les lignes $M_H^{K+1}(b, \cdot)$, $b \in \{K+2, \dots, \ell\}$. En ce qui concerne les $K+2$ premières lignes, il reste à vérifier que le coefficient $M_H^{K+1}(b, K+1)$ est positif ou nul, i.e., que pour $b \in \{0, \dots, K+1\}$,

$$\sum_{h=0}^{b-1} M_H(h+1, h) + \sum_{h=b}^K M_H(b, h) \leq 1.$$

Nous étudions le régime asymptotique

$$\begin{aligned} \ell \rightarrow +\infty, \quad m \rightarrow +\infty, \quad q \rightarrow 0, \\ \ell q \rightarrow a, \quad \frac{m}{\ell} \rightarrow \alpha. \end{aligned}$$

Il suffit donc de vérifier l'inégalité précédente pour ℓ, m assez grands et q assez petit. Nous avons les limites suivantes pour les probabilités de mutation :

$$\forall b, c \geq 0 \quad \lim_{\substack{\ell, m \rightarrow \infty \\ q \rightarrow 0, \ell q \rightarrow a}} M_H(b, c) = \begin{cases} 0 & \text{si } 0 \leq c < b, \\ \frac{a^{c-b}}{(c-b)!} e^{-a} & \text{si } 0 \leq b \leq c. \end{cases}$$

Il s'ensuit que

$$\begin{aligned} \lim_{\substack{\ell, m \rightarrow \infty \\ q \rightarrow 0, \ell q \rightarrow a}} \sum_{h=0}^{b-1} M_H(h+1, h) + \sum_{h=b}^K M_H(b, h) = \\ \sum_{h=b}^K \frac{a^{h-b}}{(h-b)!} e^{-a} = \sum_{k=0}^{K-b} \frac{a^k}{k!} e^{-a} < 1. \end{aligned}$$

Nous en déduisons que, pour ℓ, m assez grands et q assez petit, la matrice $(M_H^{K+1}(b, c), 0 \leq b, c \leq \ell)$ est bien une matrice stochastique. Nous construisons alors des applications

$$\mathcal{M}'_H, \mathcal{M}_H^{K+1} : \{0, \dots, \ell\} \times [0, 1] \longrightarrow \{0, \dots, \ell\}$$

pour coupler les mécanismes de mutation des processus $(O_n)_{n \geq 0}$, $(O_n^{K+1})_{n \geq 0}$, ce qui permettra de les comparer. Soient $b \in \{0, \dots, \ell\}$ et $u \in [0, 1]$. Nous définissons $\mathcal{M}'_H(b, u)$ comme le seul indice $c \in \{0, \dots, \ell\}$ satisfaisant

$$M_H(b, 0) + \dots + M_H(b, c-1) < u \leq M_H(b, 0) + \dots + M_H(b, c).$$

Également, nous définissons $\mathcal{M}_H^{K+1}(b, u)$ comme le seul indice $c \in \{0, \dots, \ell\}$ satisfaisant

$$M_H^{K+1}(b, 0) + \dots + M_H^{K+1}(b, c-1) < u \leq M_H^{K+1}(b, 0) + \dots + M_H^{K+1}(b, c).$$

Lemme 17.4.1. *L'application \mathcal{M}'_H est plus grande que l'application \mathcal{M}_H^{K+1} , i.e.,*

$$\forall b \in \{0, \dots, \ell\} \quad \forall u \in [0, 1] \quad \mathcal{M}'_H(b, u) \geq \mathcal{M}_H^{K+1}(b, u).$$

Démonstration. La construction de la matrice $(M_H^{K+1}(b, c), 0 \leq b, c \leq \ell)$ entraîne que, pour tout $b, c \in \{0, \dots, \ell\}$, nous avons

$$\sum_{h=0}^c M_H(b, h) \leq \sum_{h=0}^c M_H^{K+1}(b, h).$$

Cette inégalité entraîne le résultat voulu. □

Nous avons en plus le résultat suivant :

Lemme 17.4.2. *L'application \mathcal{M}'_H est croissante par rapport au premier argument, i.e.,*

$$\forall b, c \in \{0, \dots, \ell\} \quad \forall u \in [0, 1] \quad b \leq c \Rightarrow \mathcal{M}'_H(b, u) \leq \mathcal{M}'_H(c, u).$$

Démonstration. Considérons l'application

$$\mathcal{M}_H : \{0, \dots, \ell\} \times [0, 1]^\ell \rightarrow \{0, \dots, \ell\}$$

définie dans la section 16 par

$$\forall b \in \{0, \dots, \ell\} \quad \forall u_1, \dots, u_\ell \in [0, 1]^\ell$$

$$\mathcal{M}_H(b, u_1, \dots, u_\ell) = b - \sum_{k=1}^b 1_{u_k < p/\kappa} + \sum_{k=b+1}^{\ell} 1_{u_k > 1-p(1-1/\kappa)}.$$

L'intérêt de cette application réside dans le fait suivant : si U_1, \dots, U_ℓ sont des variables aléatoires indépendantes de loi uniforme dans l'intervalle $[0, 1]$, alors pour tout $b \in \{0, \dots, \ell\}$, la loi de $\mathcal{M}_H(b, U_1, \dots, U_\ell)$ est celle de la ligne de la matrice de mutation M_H associée à b , i.e.,

$$\forall c \in \{0, \dots, \ell\} \quad P(\mathcal{M}_H(b, U_1, \dots, U_\ell) = c) = M_H(b, c).$$

De plus, par le lemme 4.4.1, l'application \mathcal{M}_H est croissante par rapport au premier argument, i.e., pour tout $b, c \in \{0, \dots, \ell\}$ et $u_1, \dots, u_\ell \in [0, 1]$,

$$b \leq c \Rightarrow \mathcal{M}_H(b, u_1, \dots, u_\ell) \leq \mathcal{M}_H(c, u_1, \dots, u_\ell).$$

Soient $a, b, h \in \{0, \dots, \ell\}$ avec $a \leq b$ et soient U_1, \dots, U_ℓ des variables aléatoires indépendantes de loi uniforme sur l'intervalle $[0, 1]$. Alors, par les propriétés de l'application \mathcal{M}_H ,

$$\begin{aligned} M_H(a, 0) + \dots + M_H(a, h) &= P(\mathcal{M}_H(a, U_1, \dots, U_\ell) \leq h) \geq \\ &P(\mathcal{M}_H(b, U_1, \dots, U_\ell) \leq h) = M_H(b, 0) + \dots + M_H(b, h). \end{aligned}$$

Nous obtenons ainsi le résultat voulu. \square

Définissons

$$\mathcal{R}' = [0, 1] \times \{0, \dots, \ell\} \times [0, 1].$$

Nous construisons maintenant des applications de couplage

$$\Phi'_O, \bar{\Phi}_O : \mathcal{P}_{\ell+1}^m \times \mathcal{R}' \longrightarrow \mathcal{P}_{\ell+1}^m.$$

Soit $r = (s, j, u) \in \mathcal{R}'$. Soit $o \in \mathcal{P}_{\ell+1}^m$, prenons $l = \mathcal{S}_O(o, s)$ et soit k le seul indice dans $\{0, \dots, \ell\}$ satisfaisant

$$o(0) + \dots + o(k-1) < j \leq o(0) + \dots + o(k).$$

Nous définissons les applications Φ'_O et $\bar{\Phi}_O$ par :

$$\begin{aligned} \Phi'_O(o, r) &= o(k \rightarrow \mathcal{M}'_H(l, u)), \\ \bar{\Phi}_O(o, r) &= o(k \rightarrow \mathcal{M}^{k+1}_H(l, u)). \end{aligned}$$

Le lemme 17.4.1 entraîne que

$$\forall r \in \mathcal{R}' \quad \forall o \in \mathcal{P}_{\ell+1}^m \quad \Phi'_O(o, r) \preceq \bar{\Phi}_O(o, r).$$

Nous supposons que notre espace de probabilité (Ω, \mathcal{F}, P) contient en plus une suite i.i.d. de vecteurs aléatoires $(R'_n)_{n \geq 1}$ à valeurs dans \mathcal{R}' , indépendante des variables aléatoires définies au début de la section 16. Nous construisons le processus $(O_n)_{n \geq 0}$ à l'aide d'une suite i.i.d. $(R'_n)_{n \geq 1}$ à valeurs dans \mathcal{R}' . Soit $o \in \mathcal{P}_{\ell+1}^m$ le point de départ du processus, nous posons $O_0 = o$ et

$$\forall n \geq 1 \quad O_n = \Phi'_O(O_{n-1}, R'_n).$$

La monotonie du processus $(O_n)_{n \geq 0}$ découle du résultat suivant :

Lemme 17.4.3. *L'application de couplage Φ'_O est croissante par rapport à la distribution d'occupation, i.e.,*

$$\forall o, o' \in \mathcal{P}_{\ell+1}^m \quad \forall r \in \mathcal{R}' \quad o \preceq o' \quad \Rightarrow \quad \Phi'_O(o, r) \preceq \Phi'_O(o', r).$$

Démonstration. Soit $r = (s, j, u) \in \mathcal{R}'$ et soient $o, o' \in \mathcal{P}_{\ell+1}^m$ tels que $o \preceq o'$. Prenons $l = \mathcal{S}_O(o, s)$, $l' = \mathcal{S}_O(o', s)$ et soient k, k' les seuls indices dans $\{0, \dots, \ell\}$ satisfaisant

$$\begin{aligned} o(0) + \dots + o(k-1) &< j \leq o(0) + \dots + o(k), \\ o'(0) + \dots + o'(k'-1) &< j \leq o'(0) + \dots + o'(k'). \end{aligned}$$

Puisque $o \preceq o'$, nous avons $k \geq k'$. posons aussi

$$b = \mathcal{M}'_H(l, u), \quad b' = \mathcal{M}'_H(l', u).$$

Par le lemme 16.2.4, $l \geq l'$, donc, grâce au lemme 17.4.2, $b \geq b'$. Nous comparons finalement

$$\Phi'_O(o, r) = o(k \rightarrow b), \quad \Phi'_O(o', r) = o'(k' \rightarrow b').$$

à l'aide du lemme 17.4.4 ci-dessous.

Lemme 17.4.4. *Soient $o, o' \in \mathcal{P}_{\ell+1}^m$, $k, k', b, b' \in \{0, \dots, \ell\}$. Si $o \preceq o'$, $k \geq k'$ et $b \geq b'$, alors*

$$o(k \rightarrow b) \preceq o'(k' \rightarrow b').$$

Démonstration. Soit $h \in \{0, \dots, \ell\}$. Nous avons

$$o(k \rightarrow b)(0) + \dots + o(k \rightarrow b)(h) = o(0) + \dots + o(h) - 1_{k \leq h} + 1_{b \leq h}.$$

Comme $o \preceq o'$, nous avons $o(0) + \dots + o(h) \leq o'(0) + \dots + o'(h)$. De plus $b \geq b'$, d'où $1_{b \leq h} \leq 1_{b' \leq h}$. Le problème est dans la fonction $1_{k \leq h}$. Nous considérons les cas suivants :

- $k \leq h$. Alors

$$\begin{aligned} o(0) + \dots + o(h) - 1_{k \leq h} + 1_{b \leq h} &\leq o'(0) + \dots + o'(h) - 1 + 1_{b \leq h} \\ &\leq o'(0) + \dots + o'(h) - 1_{k' \leq h} + 1_{b' \leq h}. \end{aligned}$$

- $k' \leq h < k$. Nous avons par définition de k, k' ,

$$o(0) + \dots + o(h) < j \leq o'(0) + \dots + o'(h)$$

d'où

$$o(0) + \cdots + o(h) \leq o'(0) + \cdots + o'(h) - 1.$$

Nous en déduisons que

$$o(0) + \cdots + o(h) + 1_{b \leq h} \leq o'(0) + \cdots + o'(h) - 1_{k' \leq h} + 1_{b' \leq h}.$$

• $h < k'$. Alors

$$o(0) + \cdots + o(h) + 1_{b \leq h} \leq o'(0) + \cdots + o'(h) + 1_{b' \leq h}.$$

Dans chacun des cas nous avons

$$o(k \rightarrow b)(0) + \cdots + o(k \rightarrow b)(h) \leq o'(k' \rightarrow b')(0) + \cdots + o'(k' \rightarrow b')(h).$$

□

nous concluons que $\Phi'_O(o, r) \preceq \Phi'_O(o', r)$ comme annoncé. □

Nous construisons ensuite le processus supérieur $(O_n^{K+1})_{n \geq 0}$. Définissons l'application $\pi_{K+1} : \mathcal{P}_{\ell+1}^m \rightarrow \mathcal{P}_{\ell+1}^m$ par : pour tout $o \in \mathcal{P}_{\ell+1}^m$ et $l \in \{0, \dots, \ell\}$

$$\pi_{K+1}(o)(l) = \begin{cases} o(l) & \text{si } 0 \leq l \leq K, \\ m - (o(0) + \cdots + o(K)) & \text{si } l = K + 1, \\ 0 & \text{si } K + 2 \leq l \leq \ell. \end{cases}$$

Cette application vérifie

$$\forall o \in \mathcal{P}_{\ell+1}^m \quad o \preceq \pi_{K+1}(o).$$

Notons \mathcal{W}^* l'ensemble des distributions d'occupation ayant au moins une master sequence, i.e.,

$$\mathcal{W}^* = \{ o \in \mathcal{P}_{\ell+1}^m : o(0) \geq 1 \},$$

et notons \mathcal{N} l'ensemble des distributions d'occupation n'ayant aucune master sequence, i.e.,

$$\mathcal{N} = \{ o \in \mathcal{P}_{\ell+1}^m : o(0) = 0 \}.$$

Posons

$$o_{\text{exit}}^{K+1} = (0, m, 0, \dots, 0), \quad o_{\text{enter}}^{K+1} = (1, m - 1, 0, \dots, 0).$$

Nous définissons une application supérieure Φ_O^{K+1} en posant pour $o \in \mathcal{P}_{\ell+1}^m$ et $r \in \mathcal{R}'$,

$$\Phi_O^{K+1}(o, r) = \begin{cases} \Phi'_O(o, r) & \text{si } o \in \mathcal{N} \text{ et } \Phi'_O(o, r) \notin \mathcal{W}^* \\ o_{\text{enter}}^{K+1} & \text{si } o \in \mathcal{N} \text{ et } \Phi'_O(o, r) \in \mathcal{W}^* \\ \pi_{K+1}(\overline{\Phi}_O(\pi_{K+1}(o), r)) & \text{si } o \in \mathcal{W}^* \text{ et } \overline{\Phi}_O(\pi_{K+1}(o), r) \notin \mathcal{N} \\ o_{\text{exit}}^{K+1} & \text{si } o \in \mathcal{W}^* \text{ et } \overline{\Phi}_O(\pi_{K+1}(o), r) \in \mathcal{N} \end{cases}$$

Vérifions que l'application Φ_O^{K+1} reste au-dessus de l'application Φ'_O au sens de l'ordre de $\mathcal{P}_{\ell+1}^m$.

Proposition 17.4.5. *Pour tout $r \in \mathcal{R}'$ et pour tout $o \in \mathcal{P}_{\ell+1}^m$,*

$$\Phi'_O(o, r) \preceq \Phi_O^{K+1}(o, r).$$

Démonstration. Soient $r \in \mathcal{R}'$ et $o \in \mathcal{P}_{\ell+1}^m$. Nous regardons les cas suivants

- Si $o \in \mathcal{N}$ et $\Phi'_O(o, r) \notin \mathcal{W}^*$, alors

$$\Phi'_O(o, r) = \Phi_O^{K+1}(o, r).$$

- Si $o \in \mathcal{N}$ et $\Phi'_O(o, r) \in \mathcal{W}^*$, nous avons

$$\Phi'_O(o, r)(0) = 1 \quad \text{et} \quad \Phi_O^{K+1}(o, r) = o_{\text{enter}}^{K+1}.$$

Nous déduisons l'inégalité voulue du fait suivant :

$$\forall o \in \mathcal{P}_{\ell+1}^m \text{ avec } o(0) = 1, \quad o \preceq o_{\text{enter}}^{K+1}.$$

- Si $o \in \mathcal{W}^*$, et $\overline{\Phi}_O(\pi_{K+1}(o), r) \notin \mathcal{N}$, comme l'application $\overline{\Phi}_O$ reste au dessus de l'application Φ'_O , nous avons

$$\Phi'_O(\pi_{K+1}(o), r) \preceq \overline{\Phi}_O(\pi_{K+1}(o), r).$$

Comme de plus $\pi_{K+1}(o) \preceq o$, et que Φ_O est monotone, nous concluons que

$$\Phi'_O(o, r) \preceq \pi_{K+1}(\overline{\Phi}_O(\pi_{K+1}(o), r)).$$

- Si $o \in \mathcal{W}^*$, et $\overline{\Phi}_O(\pi_{K+1}(o), r) \in \mathcal{N}$, nous montrons comme dans le point précédent que

$$\Phi'_O(o, r) \preceq \overline{\Phi}_O(\pi_{K+1}(o), r),$$

ce qui entraîne que $\Phi'_O(o, r) \in \mathcal{N}$. De plus, $\Phi_O^{K+1}(o, r) = o_{\text{exit}}^{K+1}$ et nous concluons en remarquant que pour tout $o \in \mathcal{N}$, $o \preceq o_{\text{exit}}^{K+1}$.

Enfin, pour tout $o \in \mathcal{P}_{\ell+1}^m$ et pour tout $r \in \mathcal{R}'$,

$$\Phi'_O(o, r) \preceq \Phi_O^{K+1}(o, r).$$

□

Nous définissons un processus supérieur $(O_n^{K+1})_{n \geq 0}$ à l'aide d'une suite i.i.d. $(R'_n)_{n \geq 1} \in \mathcal{R}'$ et de l'application Φ_O^{K+1} . Soit $o \in \mathcal{P}_{\ell+1}^m$ le point de départ du processus. Nous prenons $O_0^{K+1} = o$ et

$$\forall n \geq 1 \quad O_n^{K+1} = \Phi_O^{K+1}(O_{n-1}^{K+1}, R'_n).$$

Proposition 17.4.6. *Supposons que les processus $(O_n)_{n \geq 0}$ et $(O_n^{K+1})_{n \geq 0}$ sont issus d'une même distribution d'occupation o . Nous avons alors*

$$\forall n \geq 0 \quad O_n \preceq O_n^{K+1}.$$

La démonstration est la même que celle de la proposition 17.1.1.

17.5 Dynamique du processus supérieur

Ensuite nous étudions la dynamique du processus supérieur $(O_n^{K+1})_{n \geq 0}$ dans \mathcal{W}^* . Pour le processus $(O_n^{K+1})_{n \geq 0}$, les états

$$\mathcal{J}^{K+1} = \{ o \in \mathcal{P}_{\ell+1}^m : o(0) \geq 1 \text{ et } o(0) + \dots + o(K) + o(K+1) < m \}$$

sont transitoires, et les états dans $\mathcal{N} \cup (\mathcal{W}^* \setminus \mathcal{J}^{K+1})$ forment une classe de récurrence. Nous étudions le mécanisme de transition du processus restreint à $\mathcal{W}^* \setminus \mathcal{J}^{K+1}$. Puisque

$$\mathcal{W}^* \setminus \mathcal{J}^{K+1} = \{ o \in \mathcal{P}_{\ell+1}^m : o(0) \geq 1 \text{ et } o(0) + \dots + o(K) + o(K+1) = m \},$$

un état dans $\mathcal{W}^* \setminus \mathcal{J}^{K+1}$ est totalement déterminé par les nombres d'occupation des classes $0, \dots, K$. Le processus $(O_n^{K+1})_{n \geq 0}$ entre toujours dans l'ensemble

$\mathcal{W}^* \setminus \mathcal{T}^{K+1}$ par l'état o_{enter}^{K+1} . Les seules transitions possibles pour les classes de Hamming $\{0, \dots, K\}$ du processus $(O_t^{K+1})_{t \geq 0}$ issu d'un point de $\mathcal{W}^* \setminus \mathcal{T}^{K+1}$ sont

$$(o(0), \dots, o(K)) \longrightarrow (o(0), \dots, o(K)) - e_i \quad \text{si } 1 \leq o(i), \quad 0 \leq i \leq K,$$

$$(o(0), \dots, o(K)) \longrightarrow (o(0), \dots, o(K)) + e_i \quad \text{si } o(0) + \dots + o(K) \leq m - 1, \quad 0 \leq i \leq K,$$

$$(o(0), \dots, o(K)) \longrightarrow (o(0), \dots, o(K)) - e_i + e_j \quad \text{si } 1 \leq o(i), \quad 0 \leq i, j \leq K, \quad i \neq j.$$

Le processus $(O_n^{K+1})_{n \geq 0}$ sort toujours de l'ensemble $\mathcal{W}^* \setminus \mathcal{T}^{K+1}$ par l'état o_{exit}^{K+1} . Si le processus $(O_n^{K+1})_{n \geq 0}$ est issu d'un point de $\mathcal{W}^* \setminus \mathcal{T}^{K+1}$, jusqu'au temps de sortie de $\mathcal{W}^* \setminus \mathcal{T}^{K+1}$, la dynamique de $(O_n^{K+1}(0), \dots, O_n^{K+1}(K))_{n \geq 0}$ est celle d'une chaîne de Markov. Notons par $(Z_n^{K+1})_{n \geq 0}$ une chaîne de Markov à valeurs dans

$$\mathbb{D}^{K+1} = \{z \in \mathbb{N}^{K+1} : z_0 + \dots + z_K \leq m\},$$

issue du point $z^{K+1} = (1, m - 1, 0, \dots, 0)$, ayant pour matrice de transition :

- Pour $0 \leq i \leq K$ et $0 < z_0 + \dots + z_K < m$,

$$p(z, z + e_i) = \frac{m - \sum_{l=0}^K z_l}{m((\sigma - 1)z_0 + m)} \times \left(\sigma z_0 M_H^{K+1}(0, i) + \sum_{l=1}^K z_l M_H^{K+1}(l, i) + \left(m - \sum_{l=0}^K z_l \right) M_H^{K+1}(K + 1, i) \right).$$

- Pour $0 \leq i \leq K$ et $1 \leq z_i$,

$$p(z, z - e_i) = \frac{z_i}{m((\sigma - 1)z_0 + m)} \times \left(\sigma z_0 \left(1 - \sum_{h=0}^K M_H^{K+1}(0, h) \right) + \sum_{l=1}^K z_l \left(1 - \sum_{h=0}^K M_H^{K+1}(l, h) \right) + \left(m - \sum_{l=0}^K z_l \right) \left(1 - \sum_{h=0}^K M_H^{K+1}(K + 1, h) \right) \right).$$

- Pour $0 \leq i, j \leq K$, $i \neq j$ et $1 \leq z_i$,

$$p(z, z - e_i + e_j) = \frac{z_i}{m((\sigma - 1)z_0 + m)} \times \left(\sigma z_0 M_H^{K+1}(0, j) + \sum_{l=1}^K z_l M_H^{K+1}(l, j) + \left(m - \sum_{l=0}^K z_l \right) M_H^{K+1}(K+1, j) \right).$$

Les autres coefficients non diagonaux de la matrice de transition sont nuls. Les coefficients diagonaux sont choisis de sorte que la matrice soit stochastique, i.e., la somme de chaque ligne vaut 1.

Notre but est de répliquer la dynamique de $(O_n^{K+1}(0), \dots, O_n^{K+1}(K))_{n \geq 0}$ dans l'ensemble $\mathcal{W}^* \setminus \mathcal{T}^{K+1}$. Dans cette dynamique, les probabilités de transition

$$\begin{aligned} p(z, z') & \text{ avec } z_0 = 0, \\ p(z, z') & \text{ avec } z_0 = 1 \text{ et } z'_0 = 0. \end{aligned}$$

ne jouent aucun rôle. Si nous ne modifions pas les autres probabilités de transition, nous pouvons choisir ces probabilités à notre guise, sans changer la dynamique dans $\mathcal{W}^* \setminus \mathcal{T}^{K+1}$. Nous allons le faire de la manière suivante, si la coordonnée 0 veut passer de 1 à 0, nous obligeons le processus à sauter automatiquement vers $(0, \dots, 0)$. Puis, quand le processus est en $(0, \dots, 0)$ il saute vers $(1, m-1, \dots, 0)$ avec probabilité 1. Ainsi, nous obtenons une nouvelle matrice de transition p^{K+1}

- Pour $z \in \mathbb{D}^{K+1}$ tel que $z_0 = 1$,

$$\begin{aligned} p^{K+1}(z, (0, \dots, 0)) &= p(z, z - e_0) + \sum_{i=1}^K p(z, z - e_0 + e_i), \\ p^{K+1}(z, z') &= 0 \quad \text{si } z' \neq (0, \dots, 0). \end{aligned}$$

- Pour $z \in \mathbb{D}^{K+1}$ tel que $z_0 = 0$,

$$\begin{aligned} p^{K+1}(z, (1, m-1, \dots, 0)) &= 1, \\ p^{K+1}(z, z') &= 0 \quad \text{si } z' \neq (1, m-1, \dots, 0), \end{aligned}$$

et $P^{K+1}(z, z') = p(z, z')$ pour les $z, z' \in \mathbb{D}^{K+1}$ restants.

17.6 Bornes sur la mesure invariante

Notons par μ_O^ℓ , μ_O , μ_O^{K+1} les mesures de probabilité invariantes des processus $(O_n^\ell)_{n \geq 0}$, $(O_n)_{n \geq 0}$, $(O_n^{K+1})_{n \geq 0}$. Soit ν_K la mesure image de μ_O par l'application

$$o \in \mathcal{P}_{\ell+1}^m \mapsto \frac{1}{m} (o(0) + \cdots + o(K)) \in [0, 1].$$

Pour toute fonction $f : [0, 1] \rightarrow \mathbb{R}$,

$$\begin{aligned} \int_{[0,1]} f \, d\nu_K &= \int_{\mathcal{P}_{\ell+1}^m} f\left(\frac{o(0) + \cdots + o(K)}{m}\right) d\mu_O(o) \\ &= \lim_{n \rightarrow \infty} \mathbb{E}\left(f\left(\frac{O_n(0) + \cdots + O_n(K)}{m}\right)\right). \end{aligned}$$

Fixons $f : [0, 1] \rightarrow \mathbb{R}$ une fonction croissante telle que $f(0) = 0$. Grâce à la proposition 17.2.2 nous avons l'inégalité suivante :

$$\forall n \geq 0 \quad f\left(\frac{O_n^\ell(0) + \cdots + O_n^\ell(K)}{m}\right) \leq f\left(\frac{O_n(0) + \cdots + O_n(K)}{m}\right).$$

Puis, grâce à la proposition 17.4.6,

$$\forall n \geq 0 \quad f\left(\frac{O_n(0) + \cdots + O_n(K)}{m}\right) \leq f\left(\frac{O_n^{K+1}(0) + \cdots + O_n^{K+1}(K)}{m}\right).$$

Nous prenons les espérances et nous envoyons n vers $l' \infty$, et nous obtenons

$$\begin{aligned} \int_{\mathcal{P}_{\ell+1}^m} f\left(\frac{o(0) + \cdots + o(K)}{m}\right) d\mu_O^\ell(o) \\ \leq \int_{[0,1]} f \, d\nu_K \leq \\ \int_{\mathcal{P}_{\ell+1}^m} f\left(\frac{o(0) + \cdots + o(K)}{m}\right) d\mu_O^{K+1}(o). \end{aligned}$$

Nous cherchons ensuite des estimées sur les intégrales ci-dessus. La stratégie est la même pour les intégrales inférieure et supérieure. Soit θ un symbole représentant $K + 1$ ou ℓ et étudions la mesure de probabilité invariante μ_O^θ . Pour le processus $(O_n^\theta)_{n \geq 0}$, les états dans \mathcal{T}^θ sont transitoires et les états dans $\mathcal{N} \cup (\mathcal{W}^* \setminus \mathcal{T}^\theta)$ forment une classe de récurrence. Nous appliquons le résultat de renouvellement de la proposition 10.3.1 au processus $(O_n^\theta)_{n \geq 0}$ restreint à

$\mathcal{N} \cup (\mathcal{W}^* \setminus \mathcal{T}^\theta)$, à l'ensemble $\mathcal{W}^* \setminus \mathcal{T}^\theta$, à la distribution d'occupation $\mathbf{o}_{\text{exit}}^\theta$ et à la fonction $\mathbf{o} \mapsto f((\mathbf{o}(0) + \dots + \mathbf{o}(K))/m)$. Soient

$$\begin{aligned}\tau^* &= \inf \{ n \geq 0 : O_n^\theta \in \mathcal{W}^* \}, \\ \tau &= \inf \{ n \geq \tau^* : O_n^\theta = \mathbf{o}_{\text{exit}}^\theta \}.\end{aligned}$$

Nous avons alors

$$\begin{aligned}\int_{\mathcal{P}_{\ell+1}^m} f\left(\frac{\mathbf{o}(0) + \dots + \mathbf{o}(K)}{m}\right) d\mu_{\mathbf{O}}^\theta(\mathbf{o}) &= \\ \frac{\mathbb{E}\left(\sum_{n=0}^{\tau-1} f\left(\frac{O_n^\theta(0) + \dots + O_n^\theta(K)}{m}\right) \mid O_0^\theta = \mathbf{o}_{\text{exit}}^\theta\right)}{\mathbb{E}(\tau \mid O_0^\theta = \mathbf{o}_{\text{exit}}^\theta)} &= \\ \frac{\mathbb{E}\left(\sum_{n=0}^{\tau^*-1} f\left(\frac{O_n^\theta(0) + \dots + O_n^\theta(K)}{m}\right) \mid O_0^\theta = \mathbf{o}_{\text{exit}}^\theta\right)}{\mathbb{E}(\tau \mid O_0^\theta = \mathbf{o}_{\text{exit}}^\theta)} &+ \\ + \frac{\mathbb{E}\left(\sum_{n=\tau^*}^{\tau-1} f\left(\frac{O_n^\theta(0) + \dots + O_n^\theta(K)}{m}\right) \mid O_0^\theta = \mathbf{o}_{\text{exit}}^\theta\right)}{\mathbb{E}(\tau \mid O_0^\theta = \mathbf{o}_{\text{exit}}^\theta)}.\end{aligned}$$

Tant que le processus $(O_n^\theta)_{n \geq 0}$ est dans $\mathcal{W}^* \setminus \mathcal{T}^\theta$, la dynamique des coordonnées $(O_n^\theta(0), \dots, O_n^\theta(K))_{n \geq 0}$ est la même que celle de la chaîne de Markov $(Z_n^\theta)_{n \geq 0}$ définie à la fin des sections 17.3 et 17.5. Supposons que $(Z_n^\theta)_{n \geq 0}$ est issue de z^θ , où z^θ est le point de \mathbb{D}^{K+1} donné par

$$z^\ell = (1, 0, \dots, 0) \quad z^{K+1} = (1, m-1, 0, \dots, 0).$$

Soit τ_0 le premier temps où la coordonnée 0 s'annule, i.e.

$$\tau_0 = \inf \{ n \geq 0 : Z_n^\theta(0) = 0 \}.$$

Comme le processus $(O_n^\theta)_{n \geq 0}$ entre toujours dans $\mathcal{W}^* \setminus \mathcal{T}^\theta$ par le point $\mathbf{o}_{\text{enter}}^\theta$, La loi du temps τ_0 est la même que la loi de $\tau - \tau^*$ pour le processus $(O_n^\theta)_{n \geq 0}$ issu de $\mathbf{o}_{\text{exit}}^\theta$. Nous en déduisons que $((O_n^\theta(0), \dots, O_n^\theta(K)), \tau^* \leq n \leq \tau)$ et $(Z_n^\theta, 0 \leq n \leq \tau_0)$ ont même loi. Il s'en suit que,

$$\mathbb{E}(\tau - \tau^* \mid O_0^\theta = \mathbf{o}_{\text{exit}}^\theta) = \mathbb{E}(\tau_0 \mid Z_0^\theta = z^\theta),$$

$$\begin{aligned} \mathbb{E} \left(\sum_{n=\tau^*}^{\tau-1} f \left(\frac{O_n^\theta(0) + \cdots + O_n^\theta(K)}{m} \right) \middle| O_0^\theta = o_{\text{exit}}^\theta \right) = \\ \mathbb{E} \left(\sum_{n=0}^{\tau_0-1} f \left(\frac{Z_n^\theta(0) + \cdots + Z_n^\theta(K)}{m} \right) \middle| Z_0^\theta = z^\theta \right). \end{aligned}$$

Nous pouvons alors réécrire la formule pour la mesure de probabilité invariante μ_O^θ de la manière suivante :

$$\begin{aligned} \int_{\mathcal{P}_{\ell+1}^m} f \left(\frac{o(0) + \cdots + o(K)}{m} \right) d\mu_O^\theta(o) = \\ \frac{\mathbb{E} \left(\sum_{n=0}^{\tau^*-1} f \left(\frac{O_n^\theta(0) + \cdots + O_n^\theta(K)}{m} \right) \middle| O_0^\theta = o_{\text{exit}}^\theta \right)}{\mathbb{E}(\tau^* | O_0^\theta = o_{\text{exit}}^\theta) + \mathbb{E}(\tau_0 | Z_0^\theta = z^\theta)} \\ + \frac{\mathbb{E} \left(\sum_{n=0}^{\tau_0-1} f \left(\frac{Z_n^\theta(0) + \cdots + Z_n^\theta(K)}{m} \right) \middle| Z_0^\theta = z^\theta \right)}{\mathbb{E}(\tau^* | O_0^\theta = o_{\text{exit}}^\theta) + \mathbb{E}(\tau_0 | Z_0^\theta = z^\theta)}. \end{aligned}$$

Le processus $(Z_n^\theta)_{n \geq 0}$ entre toujours dans l'ensemble $\{z \in \mathbb{D}^{K+1} : z_0 \geq 1\}$ par l'état z^θ , et il sort toujours de l'ensemble $\{z \in \mathbb{D}^{K+1} : z_0 \geq 1\}$ par l'état $(0, \dots, 0)$. Afin de réécrire la formule précédente en fonction de la mesure de probabilité invariante du processus $(Z_n^\theta)_{n \geq 0}$, nous appliquons le résultat de renouvellement de la proposition 10.3.1 au processus $(Z_n^\theta)_{n \geq 0}$, à l'ensemble $\{z \in \mathbb{D}^{K+1} : z_0 \geq 1\}$, au point 0 et à l'application $z \mapsto f((z_0 + \cdots + z_K)/m)$. Posons

$$\begin{aligned} \tau_1 &= \inf \{ n \geq 0 : Z_n^\theta(0) \geq 1 \}, \\ \tau'_0 &= \inf \{ n \geq \tau_1 : Z_n^\theta = 0 \}, \end{aligned}$$

et soit ν^θ la mesure de probabilité invariante du processus $(Z_n^\theta)_{n \geq 0}$. Nous avons

$$\begin{aligned} \int_{\mathbb{D}^{K+1}} f \left(\frac{z_0 + \cdots + z_K}{m} \right) d\nu^\theta(z) = \\ \frac{\mathbb{E} \left(\sum_{n=0}^{\tau'_0-1} f \left(\frac{Z_n^\theta(0) + \cdots + Z_n^\theta(K)}{m} \right) \middle| Z_0^\theta = 0 \right)}{\mathbb{E}(\tau'_0 | Z_0^\theta = 0)}. \end{aligned}$$

Puisque $p(0, z^\theta) = 1$, nous pouvons réécrire, grâce à la propriété de Markov,

$$\int_{\mathbb{D}^{K+1}} f\left(\frac{z_0 + \cdots + z_K}{m}\right) d\nu^\theta(z) = \frac{\mathbb{E}\left(\sum_{n=0}^{\tau_0-1} f\left(\frac{Z_n^\theta(0) + \cdots + Z_n^\theta(K)}{m}\right) \mid Z_0^\theta = z^\theta\right)}{1 + \mathbb{E}(\tau_0 \mid Z_0^\theta = z^\theta)}.$$

Nous pouvons finalement exprimer la mesure de probabilité invariante μ_O^θ en fonction de ν^θ :

$$\begin{aligned} \int_{\mathcal{P}_{\ell+1}^m} f\left(\frac{o(0) + \cdots + o(K)}{m}\right) d\mu_O^\theta(o) = & \quad (\mathcal{J}_2) \\ & \frac{\mathbb{E}\left(\sum_{n=0}^{\tau^*-1} f\left(\frac{O_n^\theta(0) + \cdots + O_n^\theta(K)}{m}\right) \mid O_0^\theta = o_{\text{exit}}^\theta\right)}{\mathbb{E}(\tau^* \mid O_0^\theta = o_{\text{exit}}^\theta) + \mathbb{E}(\tau_0 \mid Z_0^\theta = z^\theta)} \\ & + \frac{1 + \mathbb{E}(\tau_0 \mid Z_0^\theta = z^\theta)}{\mathbb{E}(\tau^* \mid O_0^\theta = o_{\text{exit}}^\theta) + \mathbb{E}(\tau_0 \mid Z_0^\theta = z^\theta)} \int_{\mathbb{D}^{K+1}} f\left(\frac{z_0 + \cdots + z_K}{m}\right) d\nu^\theta(z). \end{aligned}$$

Nous essayerons ensuite d'estimer chacun des termes du membre de droite de l'identité (\mathcal{J}_2) .

Chapitre 18

Réurrence et estimées

Dans ce chapitre nous estimons la valeur de l'intégrale

$$\int_{\mathbb{D}^{K+1}} f\left(\frac{z_0 + \cdots + z_K}{m}\right) d\nu^\theta(z).$$

Nous supposons tout le long de ce chapitre que $\sigma e^{-a} > 1$ (le cas $\sigma e^{-a} < 1$ sera traité de manière différente). Soit ρ le point de $[0, 1]^{K+1}$ donné par :

$$\rho_k = (\sigma e^{-a} - 1) \frac{a^k}{k!} \sum_{i \geq 0} \frac{i^k}{\sigma^i} \quad 0 \leq k \leq K.$$

Soit $f : [0, 1] \rightarrow \mathbb{R}$ une fonction continue, croissante et nulle en 0. Nous considérons le régime asymptotique

$$\begin{aligned} \ell &\rightarrow +\infty, & m &\rightarrow +\infty, & q &\rightarrow 0, \\ \ell q &\rightarrow a, & \frac{m}{\ell} &\rightarrow \alpha. \end{aligned}$$

Notre but est de montrer que dans ce régime nous avons à la limite

$$\lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \int_{\mathbb{D}^{K+1}} f\left(\frac{z_0 + \cdots + z_K}{m}\right) d\nu^\theta(z) = f(\rho_0 + \cdots + \rho_K).$$

Pour montrer la convergence de la mesure de probabilité invariante ν^θ , nous cherchons d'abord des estimations sur la chaîne de Markov $(Z_n^\theta)_{n \geq 0}$. Pour alléger la notation nous enlevons le θ et nous écrivons simplement $(Z_n)_{n \geq 0}$. Définissons pour $0 \leq k \leq K$ et pour $\delta > 0$,

$$U_k(\delta) = \left\{ z \in \mathbb{D}^{K+1} : \left| \frac{z_i}{m} - \rho_i \right| < \delta, 0 \leq i \leq k \right\}.$$

Définissons aussi pour $0 \leq k \leq K$ les temps d'arrêt

$$\begin{aligned}\tau_k^* &= \inf \{ n \geq 0 : Z_n \in U_k(\delta) \}, \\ \tau_k &= \inf \{ n \geq 0 : Z_n \notin U_k(2\delta) \}.\end{aligned}$$

Nous allons montrer le résultat suivant :

Théorème 18.0.1. *Soit $\delta > 0$. Pour tout $0 \leq k \leq K$, il existe des réels strictement positifs $\alpha_k, \alpha'_k, \beta_k, \beta'_k$ (qui dépendent de δ), tels que pour ℓ, m suffisamment grands et q suffisamment petit nous avons :*

- Pour tout $z \in \mathbb{D}^{K+1}$,

$$P(\tau_k^* \geq m^{\alpha_k} \mid Z_0 = z) \leq \exp(-\alpha'_k m).$$

- Pour tout $z \in U_k(\delta)$,

$$P(\tau_k \leq \exp(\beta_k m) \mid Z_0 = z) \leq \exp(-\beta'_k m).$$

Nous allons faire la preuve par récurrence sur k . Décrivons tout d'abord la stratégie de la preuve. Par définition de $U_k(\delta)$, nous avons

$$U_0(\delta) \supseteq U_1(\delta) \supseteq \dots \supseteq U_K(\delta).$$

De plus, pour décider si le processus $(Z_n)_{n \geq 0}$ est dans $U_k(\delta)$ il suffit de regarder les coordonnées $Z_n(0), \dots, Z_n(k)$. Par construction, la dynamique de la coordonnée $Z_n(k)$ ne dépend pas des coordonnées $k+1, \dots, K$. Le cas $k=0$ se réduit à étudier des chaînes de naissance et mort très similaires à celles étudiées dans [10]. Supposons le résultat démontré au rang $k-1$. Nous construisons des chaînes de naissance et mort pour borner stochastiquement la dynamique du processus $(Z_n(k))_{n \geq 0}$ dans l'ensemble $U_{k-1}(\delta)$. Ces chaînes de naissance et mort peuvent être étudiées comme celles du cas $k=0$. Pour rendre rigoureuse cette stratégie nous allons d'abord vérifier que pour tout $0 \leq k \leq K$, le processus $(Z_n(k))_{n \geq 0}$ ne dépend pas des coordonnées $k+1, \dots, K$. Prenons donc k dans $\{0, \dots, K\}$, z un point de \mathbb{D}^{K+1} , et calculons les expressions

$$\begin{aligned}P(Z_{n+1}(k) = z_k + 1 \mid Z_n = z) &= p(z, z + e_k) + \sum_{i:i \neq k} p(z, z + e_k - e_i), \\ P(Z_{n+1}(k) = z_k - 1 \mid Z_n = z) &= p(z, z - e_k) + \sum_{i:i \neq k} p(z, z - e_k + e_i).\end{aligned}$$

Pour le processus inférieur $(Z_n^\ell)_{n \geq 0}$ nous avons

$$p^\ell(z, z + e_k) + \sum_{i:i \neq k} p^\ell(z, z + e_k - e_i) = \frac{m - z_k}{m((\sigma - 1)z_0 + m)} \times \left(\sigma z_0 M_H(0, k) + \sum_{l=1}^k z_l M_H(l, k) \right),$$

$$p^\ell(z, z - e_k) + \sum_{i:i \neq k} p^\ell(z, z - e_k + e_i) = \frac{z_k}{m((\sigma - 1)z_0 + m)} \times \left(\sigma z_0 (1 - M_H(0, k)) + \sum_{l=1}^k z_l (1 - M_H(l, k)) + m - \sum_{l=0}^k z_l \right).$$

Ces quantités ne dépendent pas de z_{k+1}, \dots, z_K . Pour le processus supérieur $(Z_n^{K+1})_{n \geq 0}$ nous avons

$$p^{K+1}(z, z + e_k) + \sum_{i:i \neq k} p^{K+1}(z, z + e_k - e_i) = \frac{m - z_k}{m((\sigma - 1)z_0 + m)} \times \left(\sigma z_0 M_H(0, k) + \sum_{l=1}^k z_l M_H(l, k) + \left(m - \sum_{l=0}^k z_l \right) M_H(k + 1, k) \right),$$

$$p^{K+1}(z, z - e_k) + \sum_{i:i \neq k} p^{K+1}(z, z - e_k + e_i) = \frac{z_k}{m((\sigma - 1)z_0 + m)} \left(\sigma z_0 (1 - M_H(0, k)) + \sum_{l=1}^k z_l (1 - M_H(l, k)) + \left(m - \sum_{l=0}^k z_l \right) (1 - M_H(k + 1, k)) \right).$$

Comme dans le cas du processus inférieur, les deux expressions ne dépendent pas des coordonnées z_{k+1}, \dots, z_K . Définissons pour $k \in \{0, \dots, K\}$,

$$M_H^\ell(\ell, k) = 0, \quad M_H^{K+1}(K + 1, k) = M_H(k + 1, k).$$

Avec ces nouvelles valeurs nous pouvons écrire les probabilités de transition

précédentes pour le processus $(Z_n^\theta)_{n \geq 0}$ de la manière suivante :

$$p^\theta(z, z + e_k) + \sum_{i: i \neq k} p^\theta(z, z + e_k - e_i) = \frac{m - z_k}{m((\sigma - 1)z_0 + m)} \\ \times \left(\sigma z_0 M_H(0, k) + \sum_{l=1}^k z_l M_H(l, k) + \left(m - \sum_{l=0}^k z_l \right) M_H^\theta(\theta, k) \right),$$

$$p^\theta(z, z - e_k) + \sum_{i: i \neq k} p^\theta(z, z - e_k + e_i) = \\ \frac{z_k}{m((\sigma - 1)z_0 + m)} \left(\sigma z_0 (1 - M_H(0, k)) + \sum_{l=1}^k z_l (1 - M_H(l, k)) \right. \\ \left. + \left(m - \sum_{l=0}^k z_l \right) (1 - M_H^\theta(\theta, k)) \right).$$

Nous examinons d'abord le cas $k = 0$. Nous faisons ensuite le pas de récurrence.

18.1 Le cas $k = 0$

Nous regardons ici le cas $k = 0$. Le processus $(Z_n(0))_{n \geq 0}$ est une chaîne de naissance et mort à valeurs dans l'espace $\{0, \dots, m\}$ et de probabilités de transition

$$P(Z_{n+1}(0) = 1 \mid Z_n(0) = 0) = 1,$$

$$P(Z_{n+1}(0) = i + 1 \mid Z_n(0) = i) = \frac{m - i}{m((\sigma - 1)i + m)} \\ \times \left(\sigma i M_H(0, 0) + (m - i) M_H^\theta(\theta, 0) \right), \quad 0 < i < m,$$

$$P(Z_{n+1}(0) = i - 1 \mid Z_n(0) = i) = \frac{i}{m((\sigma - 1)i + m)} \\ \times \left(\sigma i (1 - M_H(0, 0)) + (m - i) (1 - M_H^\theta(\theta, 0)) \right), \quad 0 < i \leq m.$$

Nous donnons ensuite quelques formules explicites pour les chaînes de naissance et mort. Ces formules sont classiques, elles peuvent être trouvées en appliquant les formules générales pour les chaînes de Markov à espace d'états fini

([38]). Soit $(X_n)_{n \geq 0}$ une chaîne de naissance et mort à valeurs dans l'espace fini $\{0, \dots, m\}$ de probabilités de transition

$$\begin{aligned} P(X_{n+1} = i + 1 \mid X_n = i) &= \delta_i, & 0 \leq i < m, \\ P(X_{n+1} = i - 1 \mid X_n = i) &= \gamma_i, & 0 < i \leq m. \end{aligned}$$

Nous définissons

$$\pi(0) = 1, \quad \pi(i) = \frac{\delta_1 \cdots \delta_i}{\gamma_1 \cdots \gamma_i}, \quad 0 < i < m.$$

Pour tout sous-ensemble E de $\{0, \dots, m\}$, nous définissons le temps $\tau(E)$ d'entrée dans E ,

$$\tau(E) = \inf \{n \geq 0 : X_n \in E\}.$$

Soit a un point dans $\{1, \dots, m\}$. Nous avons la formule suivante pour le temps d'atteinte de a en partant de 0 :

$$E(\tau(\{a\}) \mid X_0 = 0) = \sum_{i=0}^{a-1} \sum_{j=i}^{a-1} \frac{1}{\delta_i} \frac{\pi(i)}{\pi(j)}.$$

Soit maintenant a un point dans $\{0, \dots, m-1\}$. Nous avons la formule suivante pour le temps d'atteinte de a en partant de m :

$$E(\tau(\{a\}) \mid X_0 = m) = \sum_{i=a+1}^m \sum_{j=a}^{i-1} \frac{1}{\delta_i} \frac{\pi(i)}{\pi(j)}.$$

Soient $a < i < b$ trois points dans $\{0, \dots, m\}$. Nous avons les formules suivantes pour le point d'entrée dans $\{a, b\}$:

$$\begin{aligned} P(X_{\tau(\{a,b\})} = a \mid X_0 = i) &= \frac{\sum_{j=i}^{b-1} \frac{1}{\pi(j)}}{\sum_{j=a}^{b-1} \frac{1}{\pi(j)}}, \\ P(X_{\tau(\{a,b\})} = b \mid X_0 = i) &= \frac{\sum_{j=a}^{i-1} \frac{1}{\pi(j)}}{\sum_{j=a}^{b-1} \frac{1}{\pi(j)}}. \end{aligned}$$

Nous allons utiliser ces formules pour étudier les chaînes de naissance et mort $(Z_n^{K+1}(0))_{n \geq 0}$ et $(Z_n^\ell(0))_{n \geq 0}$. Le processus $(Z_n^{K+1}(0))_{n \geq 0}$ est la chaîne de naissance et mort supérieure étudiée dans [10]. Le processus $(Z_n^\ell(0))_{n \geq 0}$ est une chaîne de naissance et mort encore plus simple que la chaîne de naissance et mort inférieure considérée dans [10]. En tout cas, l'analyse des chaînes de naissance et mort de [10] reste valable pour le processus $(Z_n(0))_{n \geq 0}$. Nous reproduisons les points les plus importants de cette analyse. Pour alléger la notation nous prenons $(X_n)_{n \geq 0} = (Z_n^\theta(0))_{n \geq 0}$. Le processus $(X_n)_{n \geq 0}$ est une chaîne de naissance et mort à valeurs dans l'espace $\{0, \dots, m\}$ de probabilités de transition

$$\begin{aligned} \delta_0 &= 1, \\ \delta_i &= \frac{\sigma i(m-i)M_H(0,0) + (m-i)^2 M_H^\theta(\theta,0)}{m(\sigma i + m - i)}, & 1 \leq i \leq m-1, \\ \gamma_i &= \frac{\sigma i^2(1 - M_H(0,0)) + i(m-i)(1 - M_H^\theta(\theta,0))}{m(\sigma i + m - i)}, & 1 \leq i \leq m. \end{aligned}$$

Nous allons d'abord estimer les produits $\pi(i)$. Commençons par étudier le rapport δ_i/γ_i . Nous avons

$$\forall i \in \{1, \dots, m-1\} \quad \frac{\delta_i}{\gamma_i} = \phi\left(M_H(0,0), M_H(\theta,0), \frac{i}{m}\right),$$

où $\phi :]0, 1[\times]0, 1[\times]0, 1[\rightarrow]0, +\infty[$ est la fonction donnée par

$$\phi(\beta, \varepsilon, \rho) = \frac{(1-\rho)(\sigma\beta\rho + (1-\rho)\varepsilon)}{\rho(\sigma(1-\beta)\rho + (1-\rho)(1-\varepsilon))}.$$

Pour étudier les produits $\pi(i)$ il suffit de savoir si la valeur de ϕ est plus grande ou plus petite que 1. L'équation $\phi(\beta, \varepsilon, \rho) = 1$ se met sous la forme

$$(\sigma-1)\rho^2 + (1-\sigma\beta+\varepsilon)\rho - \varepsilon = 0.$$

Cette équation a une racine positive, donnée par

$$\rho(\beta, \varepsilon) = \frac{1}{2(\sigma-1)} \left(\sigma\beta - 1 - \varepsilon + \sqrt{(\sigma\beta - 1 - \varepsilon)^2 + 4\varepsilon(\sigma-1)} \right).$$

Nous avons donc

$$\begin{aligned} \phi(\beta, \varepsilon, \rho) &> 1 & \text{si } \rho < \rho(\beta, \varepsilon), \\ \phi(\beta, \varepsilon, \rho) &< 1 & \text{si } \rho > \rho(\beta, \varepsilon). \end{aligned}$$

Cela entraîne que

$$\begin{aligned} 1 \leq i \leq j \leq \lfloor \rho(\beta, \varepsilon)m \rfloor &\implies \pi(i) \leq \pi(j), \\ \lfloor \rho(\beta, \varepsilon)m \rfloor \leq i \leq j \leq m &\implies \pi(i) \geq \pi(j). \end{aligned}$$

Nous en déduisons que le produit $\pi(i)$ est maximal pour $i = \lfloor \rho(\beta, \varepsilon)m \rfloor$:

$$\max_{1 \leq i \leq m} \pi(i) = \pi(\lfloor \rho(\beta, \varepsilon)m \rfloor).$$

Nous avons par ailleurs

$$\begin{aligned} M_H(0, 0) &= (1 - q)^\ell, \\ M_H^{K+1}(K + 1, 0) &= (1 - q)^{\ell-1} \frac{q}{\kappa - 1}, \quad M_H^\ell(\ell, 0) = 0. \end{aligned}$$

Nous supposons que

$$\ell \rightarrow +\infty, \quad m \rightarrow +\infty, \quad q \rightarrow 0,$$

avec de plus

$$\ell q \rightarrow a \in]0, +\infty[.$$

Nous avons alors

$$\begin{aligned} \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} M_H(0, 0) &= \exp(-a), \\ \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} M_H^{K+1}(K + 1, 0) &= \lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} M_H^\ell(\ell, 0) = 0. \end{aligned}$$

Le résultat suivant, donne des estimées de grandes déviations pour les produits $\pi(i)$.

Proposition 18.1.1. *Soit $a \in]0, +\infty[$. Pour $\rho \in [0, 1]$, nous avons*

$$\lim_{\substack{\ell, m \rightarrow \infty \\ q \rightarrow 0, \ell q \rightarrow a}} \frac{1}{m} \ln \pi(\lfloor \rho m \rfloor) = \int_0^\rho \ln \phi(\exp(-a), 0, s) ds.$$

La preuve détaillée se trouve dans [10]. Nous définissons, pour $\sigma e^{-a} > 1$,

$$\rho_0 = \rho(\exp(-a), 0) = \frac{\sigma e^{-a} - 1}{\sigma - 1}.$$

Comme $\phi(e^{-a}, 0, s) > 1$ pour $s < \rho_0$ et $\phi(e^{-a}, 0, s) < 1$ pour $s > \rho_0$, la fonction

$$\rho \mapsto \int_0^\rho \ln \phi(e^{-a}, 0, s) ds$$

est croissante sur $]0, \rho_0[$, décroissante sur $] \rho_0, 1[$, et maximale pour $\rho = \rho_0$. Le point ρ_0 est la limite du point $\rho(M_H(0, 0), M_H(\theta, 0))$ quand $\ell, m \rightarrow \infty$, $q \rightarrow 0$ et $\ell q \rightarrow a$. Nous utilisons les résultats précédents pour obtenir des estimations pertinentes sur la chaîne de naissance et mort $(X_n)_{n \geq 0}$. Nous cherchons d'abord des estimées sur le temps d'entrée dans un voisinage de ρ_0 . Définissons pour $\delta > 0$

$$V_0(\delta) = \left\{ x \in \{0, \dots, m\} : \left| \frac{x}{m} - \rho_0 \right| < \delta \right\}.$$

Soit $\tau(V_0(\delta))$ le temps d'entrée dans $V_0(\delta)$,

$$\tau(V_0(\delta)) = \inf \{ n \geq 0 : X_n \in V_0(\delta) \}.$$

Nous allons démontrer le résultat suivant.

Proposition 18.1.2. *Soit $\delta > 0$. Il existe des réels strictement positifs α_0, α'_0 (qui dépendent de δ), tels que pour ℓ, m suffisamment grands et q suffisamment petit,*

$$\forall x \in \{0, \dots, m\} \quad P(\tau(V_0(\delta)) \geq m^{\alpha_0} \mid X_0 = x) \leq \exp(-\alpha'_0 m).$$

Démonstration. Par monotonie il suffit de faire la preuve pour $x = 0$ et $x = m$. Soient b et c les extrémités de l'intervalle $V_0(\delta)$,

$$b = \lfloor (\rho_0 - \delta)m \rfloor + 1, \quad c = \lfloor (\rho_0 + \delta)m \rfloor.$$

Grâce aux formules données précédemment nous avons

$$E(\tau(V_0(\delta)) \mid X_0 = 0) = E(\tau(\{b\}) \mid X_0 = 0) = \sum_{i=0}^{b-1} \sum_{j=i}^{b-1} \frac{1}{\delta_i} \frac{\pi(i)}{\pi(j)}.$$

Pour ℓ, m assez grands et q assez petit, le point $\tilde{\rho} = \rho(M_H(0, 0), M_H(\theta, 0))$ est dans $V_0(\delta)$ et de plus $b \leq \tilde{\rho} \leq c$. Comme les produits $\pi(i)$ sont croissants sur $\{1, \dots, \tilde{\rho} - 1\}$, nous voyons que

$$1 \leq i \leq j \leq b - 1 \implies \frac{\pi(i)}{\pi(j)} \leq 1.$$

De plus, pour $m \geq 1$,

$$\forall i \in \{0, \dots, m - 1\} \quad \delta_i \geq \frac{M_H(0, 0)}{m^2}.$$

Il s'ensuit que,

$$E(\tau(V_0(\delta)) \mid X_0 = 0) \leq \frac{1}{M_H(0,0)} m^4.$$

Soit $\alpha > 4$. Par l'inégalité de Markov,

$$P(\tau(V_0(\delta)) > m^\alpha \mid X_0 = 0) \leq \frac{1}{M_H(0,0)} m^{-(\alpha-4)}.$$

Nous allons maintenant estimer pour $k \geq 1$

$$P(\tau(V_0(\delta)) > km^\alpha \mid X_0 = 0).$$

Nous décomposons selon les états possibles au temps $(k-1)m^\alpha$:

$$\begin{aligned} P(\tau(V_0(\delta)) \geq km^\alpha \mid X_0 = 0) &= \\ &\sum_{x < b} P(X_{(k-1)m^\alpha} = x, \tau(V_0(\delta)) \geq (k-1)m^\alpha, \tau(V_0(\delta)) \geq km^\alpha \mid X_0 = 0) \\ &= \sum_{x < b} P(X_{(k-1)m^\alpha} = x, \tau(V_0(\delta)) \geq (k-1)m^\alpha \mid X_0 = 0) \\ &\quad \times P(\tau(V_0(\delta)) \geq km^\alpha \mid X_0 = 0, X_{(k-1)m^\alpha} = x, \tau(V_0(\delta)) \geq (k-1)m^\alpha). \end{aligned}$$

Or, par la propriété de Markov, et grâce la monotonie du modèle,

$$\begin{aligned} P(\tau(V_0(\delta)) \geq km^\alpha \mid X_0 = 0, X_{(k-1)m^\alpha} = x, \tau(V_0(\delta)) \geq (k-1)m^\alpha) \\ = P(\tau(V_0(\delta)) \geq km^\alpha - (k-1)m^\alpha \mid X_0 = x) \\ \leq P(\tau(V_0(\delta)) \geq m^\alpha \mid X_0 = 0) \leq \frac{1}{M_H(0,0)} m^{-(\alpha-4)}. \end{aligned}$$

Donc, pour tout $k \geq 1$,

$$P(\tau(V_0(\delta)) \geq km^\alpha \mid X_0 = 0) \leq \frac{1}{M_H(0,0)} m^{-(\alpha-4)} P(\tau(V_0(\delta)) \geq (k-1)m^\alpha \mid X_0 = 0).$$

Nous itérons ce procédé pour les temps $(k-2)m^\alpha, \dots, 2m^\alpha, m^\alpha$ et nous obtenons

$$\begin{aligned} P(\tau(V_0(\delta)) \geq km^\alpha \mid X_0 = 0) &\leq \left(\frac{1}{M_H(0,0)} m^{-(\alpha-4)} \right)^k \\ &= \exp \left(-k((\alpha-4) \ln m + \ln M_H(0,0)) \right). \end{aligned}$$

Choisissons donc $\alpha_0 > 5$ et soit m_0 suffisamment grand pour avoir

$$(\alpha_0 - 4) \ln m_0 + \ln M_H(0, 0) > 0.$$

Nous obtenons le résultat voulu en prenant $\alpha'_0 = (\alpha_0 - 4) \ln m_0 + \ln M_H(0, 0)$ et $k = m$. La preuve pour l'état initial $X_0 = m$ est similaire. \square

Nous cherchons ensuite des estimées sur le temps de sortie d'un voisinage de ρ_0 . Soit $\delta > 0$. Soit $\tau(V_0(2\delta)^c)$ le temps de sortie de l'ensemble $V_0(2\delta)$,

$$\tau(V_0(2\delta)^c) = \inf \{ n \geq 0 : X_n \notin V_0(2\delta) \}.$$

Nous allons démontrer le résultat suivant :

Proposition 18.1.3. *Soit $\delta > 0$. Il existe des réels strictement positifs β_0, β'_0 (qui dépendent de δ), tels que pour ℓ, m suffisamment grands et q suffisamment petit*

$$\forall x \in V_0(\delta) \quad P(\tau(V_0(2\delta)^c) < \exp(\beta_0 m) \mid X_0 = x) \leq \exp(-\beta'_0 m).$$

Démonstration. Soit $t > 0$ et regardons d'abord la valeur de

$$P(\tau(V_0(2\delta)^c) \leq t \mid X_0 = x).$$

Soit θ le dernier temps avant $\tau(V_0(2\delta)^c)$ où le processus $(X_n)_{n \geq 0}$ est dans $V_0(\delta)$, i.e.,

$$\theta = \max \{ n < \tau(V_0(2\delta)^c) : X_n \in V_0(\delta) \}.$$

Soient b et c les extrémités de l'intervalle $V_0(\delta)$,

$$b = \lfloor (\rho_0 - \delta)m \rfloor + 1, \quad c = \lfloor (\rho_0 + \delta)m \rfloor.$$

Également, soient b' et c' les extrémités de l'intervalle $V_0(2\delta)$,

$$b' = \lfloor (\rho_0 - 2\delta)m \rfloor + 1, \quad c' = \lfloor (\rho_0 + 2\delta)m \rfloor.$$

Pour ℓ, m assez grands et q assez petit nous avons

$$\begin{aligned} P(\tau(V_0(2\delta)^c) \leq t \mid X_0 = x) &= \sum_{n < t} P(\theta = n, \tau(V_0(2\delta)^c) \leq t \mid X_0 = x) \\ &= \sum_{n < t} \left(P(\theta = n, X_{n+1} = b - 1, \tau(V_0(2\delta)^c) \leq t \mid X_0 = x) \right. \\ &\quad \left. + P(\theta = n, X_{n+1} = c + 1, \tau(V_0(2\delta)^c) \leq t \mid X_0 = x) \right). \end{aligned}$$

Considérons le premier terme dans la parenthèse. Par la propriété de Markov, nous avons

$$\begin{aligned}
& \mathbb{P}(\theta = n, X_{n+1} = b - 1, \tau(V_0(2\delta)^c) \leq t \mid X_0 = x) \\
&= \mathbb{P}\left(X_n = b, X_{n+1} = b - 1, \tau(V_0(2\delta)^c) \leq t \mid X_0 = x\right) \\
&\leq \mathbb{P}\left(X_r \notin V_0(\delta) \text{ pour } r \leq \tau(V_0(2\delta)^c) \mid X_0 = b - 1\right) \\
&\leq \mathbb{P}\left(X_{\tau(V_0(2\delta)^c \cup \{b\})} \in V_0(2\delta)^c \mid X_0 = b - 1\right) = \\
&\quad \mathbb{P}(X_{\tau(\{b'-1, b\})} = b' - 1 \mid X_0 = b - 1).
\end{aligned}$$

Nous utilisons la formule donnée au début de la section pour écrire

$$\mathbb{P}(X_{\tau(\{b'-1, b\})} = b' - 1 \mid X_0 = b - 1) = \frac{1}{\frac{\pi(b-1)}{\sum_{i=b'-1}^{b-1} \frac{1}{\pi(i)}}}.$$

Donc,

$$\mathbb{P}(X_{\tau(\{b'-1, b\})} = b' - 1 \mid X_0 = b - 1) \leq \frac{\pi(b' - 1)}{\pi(b - 1)}.$$

Soit $\varepsilon > 0$ et soient ℓ, m assez grands et q assez petit de façon que

$$\begin{aligned}
& \left| \frac{1}{m} \ln \pi(b - 1) - \int_0^{\rho_0 - \delta} \ln \phi(e^{-a}, 0, s) ds \right| < \frac{\varepsilon}{2}, \\
& \left| \frac{1}{m} \ln \pi(b' - 1) - \int_0^{\rho_0 - 2\delta} \ln \phi(e^{-a}, 0, s) ds \right| < \frac{\varepsilon}{2}.
\end{aligned}$$

Nous avons alors

$$\begin{aligned}
\frac{\pi(b' - 1)}{\pi(b - 1)} &= \exp\left(m\left(\frac{1}{m} \ln \pi(b' - 1) - \frac{1}{m} \ln \pi(b - 1)\right)\right) \leq \\
&\quad \exp\left(m\left(-\int_{\rho_0 - 2\delta}^{\rho_0 - \delta} \ln \phi(e^{-a}, 0, s) ds + \varepsilon\right)\right).
\end{aligned}$$

La fonction

$$\rho \mapsto \int_0^{\rho} \ln \phi(e^{-a}, 0, s) ds$$

est croissante sur $]0, \rho_0[$. Donc,

$$\int_{\rho_0-2\delta}^{\rho_0-\delta} \ln \phi(e^{-\alpha}, 0, s) ds > 0.$$

Choisissons ε tel que

$$\varepsilon < \int_{\rho_0-2\delta}^{\rho_0-\delta} \ln \phi(e^{-\alpha}, 0, s) ds.$$

Alors si nous prenons

$$\beta_1 = \int_{\rho_0-2\delta}^{\rho_0-\delta} \ln \phi(e^{-\alpha}, 0, s) ds - \varepsilon,$$

nous avons $\beta_1 > 0$ et

$$P(\theta = n, X_{n+1} = b - 1, \tau(V_0(2\delta)^c) \leq t \mid X_0 = x) \leq \exp(-\beta_1 m).$$

Nous faisons un travail similaire pour le terme

$$P(\theta = n, X_{n+1} = c + 1, \tau(V_0(2\delta)^c) \leq t \mid X_0 = x)$$

et nous obtenons un $\beta_2 > 0$ tel que

$$P(\theta = n, X_{n+1} = c + 1, \tau(V_0(2\delta)^c) \leq t \mid X_0 = x) \leq \exp(-\beta_2 m).$$

Il s'ensuit que

$$P(\tau(V_0(2\delta)^c) \leq t \mid X_0 = x) \leq t \left(\exp(-\beta_1 m) + \exp(-\beta_2 m) \right).$$

Soit $\beta_0 < \min(\beta_1, \beta_2)$. Nous avons donc

$$P(\tau(V_0(2\delta)^c) \leq \exp(\beta_0 m) \mid X_0 = x) \leq 2 \exp(-\min(\beta_1 - \beta_0, \beta_2 - \beta_0)m).$$

En choisissant, pour m_0 suffisamment grand,

$$\beta'_0 = \min(\beta_1 - \beta_0, \beta_2 - \beta_0) - \frac{1}{m_0} \ln 2$$

nous obtenons le résultat voulu. □

18.2 L'étape de récurrence

Soit $k \in \{1, \dots, K\}$. Dans cette section nous supposons que le résultat du théorème 18.0.1 est vrai au rang $k - 1$ et nous démontrons que le résultat reste vrai au rang k . Soient $\delta, \delta' > 0$. Rappelons que pour $k \in \{0, \dots, K\}$ le sous-ensemble $U_k(\delta)$ de \mathbb{D}^{K+1} est défini par

$$U_k(\delta) = \left\{ z \in \mathbb{D}^{K+1} : \left| \frac{z_i}{m} - \rho_i \right| < \delta \quad 0 \leq i \leq k \right\}.$$

Posons

$$\begin{aligned} \tau_{k-1}^* &= \inf \{ n \geq 0 : Z_n \in U_{k-1}(\delta') \}, \\ \tau_{k-1} &= \inf \{ n \geq 0 : Z_n \notin U_{k-1}(2\delta') \}. \end{aligned}$$

Soient aussi

$$\begin{aligned} \tau_k^* &= \inf \{ n \geq 0 : Z_n \in U_k(\delta) \}, \\ \tau_k &= \inf \{ n \geq 0 : Z_n \notin U_k(2\delta) \}. \end{aligned}$$

Nous allons montrer qu'il existe $\alpha_k, \alpha'_k, \beta_k, \beta'_k$ des réels strictement positifs tels que, pour ℓ, m suffisamment grands et q suffisamment petit,

$$\begin{aligned} \forall z \in \mathbb{D}^{K+1} \quad & P(\tau_k^* \geq m^{\alpha_k} \mid Z_0 = z) \leq \exp(-\alpha'_k m), \\ \forall z \in U_k(\delta) \quad & P(\tau_k \leq \exp(\beta_k m) \mid Z_0 = z) \leq \exp(-\beta'_k m). \end{aligned}$$

Par l'hypothèse de récurrence, il existe $\alpha_{k-1}, \alpha'_{k-1}, \beta_{k-1}, \beta'_{k-1}$ des réels strictement positifs tels que, pour ℓ, m suffisamment grands et q suffisamment petit,

$$\begin{aligned} \forall z \in \mathbb{D}^{K+1} \quad & P(\tau_{k-1}^* \geq m^{\alpha_{k-1}} \mid Z_0 = z) \leq \exp(-\alpha'_{k-1} m), \\ \forall z \in U_{k-1}(\delta') \quad & P(\tau_{k-1} \leq \exp(\beta_{k-1} m) \mid Z_0 = z) \leq \exp(-\beta'_{k-1} m). \end{aligned}$$

Soient z^0 un point dans \mathbb{D}^{K+1} et $\alpha > \alpha_{k-1}$. Nous avons

$$\begin{aligned} P(\tau_k^* \geq m^\alpha \mid Z_0 = z^0) &= P(\tau_{k-1}^* \geq m^{\alpha_{k-1}}, \tau_k^* \geq m^\alpha \mid Z_0 = z^0) \\ &\quad + P(\tau_{k-1}^* < m^{\alpha_{k-1}}, \tau_k^* \geq m^\alpha \mid Z_0 = z^0). \end{aligned}$$

Or, par l'hypothèse de récurrence, le premier terme dans la somme est majoré par la quantité $\exp(-\alpha'_{k-1} m)$. Il reste à étudier le deuxième terme. Par la pro-

priété de Markov, nous avons

$$\begin{aligned}
& P(\tau_{k-1}^* < m^{\alpha_{k-1}}, \tau_k^* \geq m^\alpha \mid Z_0 = z^0) \\
&= \sum_{\substack{n < m^{\alpha_{k-1}} \\ z \in \mathcal{U}_{k-1}(\delta')}} P(\tau_{k-1}^* = n, Z_n = z, \tau_k^* \geq m^\alpha \mid Z_0 = z^0) \\
&= \sum_{\substack{n < m^{\alpha_{k-1}} \\ z \in \mathcal{U}_{k-1}(\delta')}} P(\tau_{k-1}^* = n, Z_n = z \mid Z_0 = z^0) P(\tau_k^* \geq m^\alpha - n \mid Z_0 = z).
\end{aligned}$$

Soit m suffisamment grand pour que $m^\alpha - m^{\alpha_{k-1}} < \exp(\beta_{k-1}m)$. Alors, pour $n < m^{\alpha_{k-1}}$,

$$\begin{aligned}
P(\tau_k^* \geq m^\alpha - n \mid Z_0 = z) &\leq P(\tau_k^* \geq m^\alpha - m^{\alpha_{k-1}} \mid Z_0 = z) \\
&= P(\tau_{k-1} \leq \exp(\beta_{k-1}m), \tau_k^* \geq m^\alpha - m^{\alpha_{k-1}} \mid Z_0 = z) \\
&\quad + P(\tau_{k-1} > \exp(\beta_{k-1}m), \tau_k^* \geq m^\alpha - m^{\alpha_{k-1}} \mid Z_0 = z).
\end{aligned}$$

Toujours par l'hypothèse de récurrence, le premier terme dans la somme est majoré par $\exp(-\beta'_{k-1}m)$. Pour le deuxième terme nous avons

$$\begin{aligned}
& P(\tau_{k-1} > \exp(\beta_{k-1}m), \tau_k^* \geq m^\alpha - m^{\alpha_{k-1}} \mid Z_0 = z) \\
&\leq P(\tau_{k-1} \geq m^\alpha - m^{\alpha_{k-1}}, \tau_k^* \geq m^\alpha - m^{\alpha_{k-1}} \mid Z_0 = z) \\
&\leq P(\tau_k^* \geq m^\alpha - m^{\alpha_{k-1}} \mid \tau_{k-1} \geq m^\alpha - m^{\alpha_{k-1}}, Z_0 = z) \\
&= P(\tau_k^* \geq m^\alpha - m^{\alpha_{k-1}} \mid Z_0 = z, Z_n \in \mathcal{U}_{k-1}(2\delta'), \quad 0 \leq n \leq m^\alpha - m^{\alpha_{k-1}}).
\end{aligned}$$

Puisque

$$\mathcal{U}_k(\delta) = \left\{ z \in \mathbb{D}^{k+1} : \left| \frac{z_i}{m} - \rho_i \right| < \delta \quad 0 \leq i \leq k \right\},$$

afin de montrer le résultat, il nous suffira d'étudier le processus $(Z_n(k))_{n \geq 0}$ lorsque la chaîne de Markov $(Z_n)_{n \geq 0}$ est dans $\mathcal{U}_{k-1}(2\delta')$. C'est également le cas pour le temps de sortie d'un voisinage de (ρ_0, \dots, ρ_k) . En effet, prenons z^0 un point dans $\mathcal{U}_k(\delta)$ et $\beta < \beta_{k-1}$, nous avons

$$\begin{aligned}
P(\tau_k \leq \exp(\beta m) \mid Z_0 = z^0) &= \\
& P(\tau_{k-1} \leq \exp(\beta_{k-1}m), \tau_k \leq \exp(\beta m) \mid Z_0 = z^0) \\
&\quad + P(\tau_{k-1} > \exp(\beta_{k-1}m), \tau_k \leq \exp(\beta m) \mid Z_0 = z^0).
\end{aligned}$$

Or, par l'hypothèse de récurrence, le premier terme dans la somme est borné par la quantité $\exp(-\beta'_{k-1}m)$. Pour le deuxième terme, nous avons

$$\begin{aligned} & \mathbb{P}(\tau_{k-1} > \exp(\beta_{k-1}m), \tau_k \leq \exp(\beta m) \mid Z_0 = z^0) \\ & \leq \mathbb{P}(\tau_{k-1} > \exp(\beta m), \tau_k \leq \exp(\beta m) \mid Z_0 = z^0) \\ & \leq \mathbb{P}(\tau_k \leq \exp(\beta m) \mid Z_0 = z^0, \tau_{k-1} > \exp(\beta m)) \\ & = \mathbb{P}(\tau_k \leq \exp(\beta m) \mid Z_0 = z^0, Z_n \in \mathcal{U}_{k-1}(2\delta') \quad 0 \leq n \leq \exp(\beta m)). \end{aligned}$$

Dans ce cas aussi, il suffit d'étudier le processus $(Z_n(k))_{n \geq 0}$ lorsque la chaîne de Markov $(Z_n)_{n \geq 0}$ est dans $\mathcal{U}_{k-1}(2\delta')$. Plus précisément, nous devons montrer le résultat suivant.

Proposition 18.2.1. *Soient $\delta, \delta' > 0$. Il existe $\alpha_k, \alpha'_k, \beta_k, \beta'_k$ des réels strictement positifs tels que, pour ℓ, m suffisamment grands et q, δ' suffisamment petits,*

- Pour tout z dans $\mathcal{U}_{k-1}(\delta')$,

$$\mathbb{P}(\tau_k^* \geq m^{\alpha_k} \mid Z_0 = z, Z_n \in \mathcal{U}_{k-1}(\delta'), \quad 0 \leq n \leq m^{\alpha_k}) \leq e^{-\alpha'_k m}.$$

- Pour tout z dans $\mathcal{U}_{k-1}(\delta') \cap \mathcal{U}_k(\delta)$,

$$\mathbb{P}(\tau_k \leq e^{\beta_k m} \mid Z_0 = z, Z_n \in \mathcal{U}_{k-1}(\delta'), \quad 0 \leq n \leq e^{\beta_k m}) \leq e^{-\beta'_k m}.$$

18.3 Bornes stochastiques pour $(Z_n(k))_{n \geq 0}$

Si nous conditionnons par rapport au processus $(Z_n(0), \dots, Z_n(k-1))_{n \geq 0}$, le processus $(Z_n(k))_{n \geq 0}$ de la k -ième coordonnée peut être vu comme un processus de naissance et mort dont les probabilités de transition dépendent du temps.

Nous allons construire un processus $(X_n)_{n \geq 0}$ dont la loi conditionnelle sachant les trajectoires $(Z_n(0))_{n \geq 0}, \dots, (Z_n(k-1))_{n \geq 0}$ est la même que celle de la coordonnée $(Z_n(k))_{t \geq 0}$, et nous allons réaliser un couplage entre ce processus $(X_n)_{n \geq 0}$ et deux chaînes de naissance et mort, une chaîne inférieure $(X_n^L)_{n \geq 0}$ et une chaîne supérieure $(X_n^U)_{n \geq 0}$. Le couplage n'est opérationnel que lorsque la chaîne $(Z_n)_{n \geq 0}$ est dans $\mathcal{U}_{k-1}(\delta')$. Mais justement, pour obtenir les estimation voulues, il suffit de regarder la chaîne $(Z_n)_{n \geq 0}$ dans $\mathcal{U}_{k-1}(\delta')$. Nous définissons d'abord des applications $\delta_i, \gamma_i : [0, 1]^k \times \{0, \dots, m\} \rightarrow \{0, \dots, m\}$, pour $i \in \{0, \dots, K\}$, par les formules suivantes. Soit $x = (x_0, \dots, x_{k-1}) \in [0, 1]^k$

- Nous posons $\delta_m(x) = 0$ et pour $0 \leq i < m$,

$$\delta_i(x) = \frac{1 - i/m}{(\sigma - 1)x_0 + 1} \left(\sigma x_0 M_H(0, k) + \sum_{l=1}^{k-1} x_l M_H(l, k) + \frac{i}{m} M_H(k, k) + \left(1 - \sum_{l=0}^{k-1} x_l - \frac{i}{m} \right) M_H(\theta, k) \right).$$

- Nous posons $\gamma_0(x) = 0$ et pour $0 < i \leq m$,

$$\gamma_i(x) = \frac{i/m}{(\sigma - 1)x_0 + 1} \left(\sigma x_0 (1 - M_H(0, k)) + \sum_{l=1}^{k-1} x_l (1 - M_H(l, k)) + \frac{i}{m} (1 - M_H(k, k)) + \left(1 - \sum_{l=0}^{k-1} x_l - \frac{i}{m} \right) (1 - M_H(\theta, k)) \right).$$

Les probabilités de transition pour le processus $(Z_n^\theta(k))_{n \geq 0}$ sont données par : pour $z = (z_0, \dots, z_k) \in \mathbb{D}^{k+1} \setminus \{0\}$ et $i = z_k$,

$$\begin{aligned} P(Z_{t+1}(k) = i + 1 \mid Z_t = z) &= \delta_i \left(\frac{z_0}{m}, \dots, \frac{z_{k-1}}{m} \right), \\ P(Z_{t+1}(k) = i - 1 \mid Z_t = z) &= \gamma_i \left(\frac{z_0}{m}, \dots, \frac{z_{k-1}}{m} \right). \end{aligned}$$

Il est difficile de réaliser un couplage en utilisant le processus $(Z_n^\theta(k))_{n \geq 0}$. Afin de réaliser notre couplage, nous allons construire un processus $(X_n)_{n \geq 0}$ qui a même loi que $(Z_n^\theta(k))_{n \geq 0}$, conditionnellement à $(Z_n(0), \dots, Z_n(k-1))_{n \geq 0}$. Nous définissons une application

$$C : [0, 1]^k \times \{0, \dots, m\} \times [0, 1] \longrightarrow [0, 1]$$

en posant pour $x \in [0, 1]^k$, $i \in \{0, \dots, m\}$ et $u \in [0, 1]$,

$$C(x, i, u) = i - 1_{u < \gamma_i(x)} + 1_{u > 1 - \delta_i(x)}.$$

L'application C est définie de sorte que si U est une variable aléatoire de loi uniforme sur l'intervalle $[0, 1]$, alors pour $x \in [0, 1]^k$ et $i \in \{0, \dots, m\}$,

$$P(C(x, i, U) = i + 1) = \delta_i(x), \quad P(C(x, i, U) = i - 1) = \gamma_i(x).$$

Lemme 18.3.1. *Pour m suffisamment grand l'application C est croissante par rapport à l'argument i :*

$$\begin{aligned} \forall x \in [0, 1]^k \quad \forall i, j \in \{0, \dots, m\} \quad \forall u \in [0, 1] \\ i \leq j \implies C(x, i, u) \leq C(x, j, u). \end{aligned}$$

Démonstration. Il suffit de montrer le résultat pour $j = i + 1$. Nous avons

$$C(x, i + 1, u) - C(x, i, u) = 1 - (1_{u < \gamma_{i+1}(x)} - 1_{u < \gamma_i(x)}) + (1_{u > 1 - \delta_{i+1}(x)} - 1_{u > 1 - \delta_i(x)}).$$

Cette quantité est négative si et seulement si

$$\gamma_i(x) < u < \gamma_{i+1}(x) \quad \text{et} \quad 1 - \delta_i(x) < u < 1 - \delta_{i+1}(x).$$

Ceci n'est possible que si $\delta_i(x) + \gamma_{i+1}(x) > 1$. Cependant, si nous posons $y \in [0, 1]$ et $i = \lfloor ym \rfloor$, nous voyons que

$$\begin{aligned} \lim_{m \rightarrow \infty} \delta_i(x) + \gamma_{i+1}(x) &= \\ &= \frac{1}{(\sigma - 1)x_0 + 1} \left(\sigma x_0 \left((1 - y)M_H(0, k) + y(1 - M_H(0, k)) \right) \right. \\ &\quad + \sum_{l=1}^{k-1} x_l \left((1 - y)M_H(l, k) + y(1 - M_H(l, k)) \right) \\ &\quad + y \left((1 - y)M_H(k, k) + y(1 - M_H(k, k)) \right) \\ &\quad \left. + \left(1 - \sum_{l=0}^{k-1} x_l - y \right) \left((1 - y)M_H(\theta, k) + y(1 - M_H(\theta, k)) \right) \right). \end{aligned}$$

Or, si $u \in]0, 1[$ et $v \in [0, 1]$, nous avons

$$u(1 - v) + (1 - u)v < 1 - v + v = 1.$$

La limite précédente est alors strictement plus petite que 1. Donc, pour m assez grand,

$$\delta_i(x) + \gamma_{i+1}(x) \leq 1.$$

□

Soit $(U_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes et identiquement distribuées, de loi uniforme sur l'intervalle $[0, 1]$. Nous définissons un processus $(X_n)_{n \geq 0}$ à valeurs dans $\{0, \dots, m\}$ à l'aide du processus $(Z_n)_{n \geq 0}$ et de la suite $(U_n)_{n \geq 1}$. Soit $z \in \mathbb{D}^{k+1} \setminus \{0\}$ le point de départ du processus $(Z_n)_{n \geq 0}$. Nous posons $X_0 = z_k$ et

$$\forall n \geq 1 \quad X_n = C\left(\frac{Z_{n-1}(0)}{m}, \dots, \frac{Z_{n-1}(k-1)}{m}, X_{n-1}, U_n\right).$$

Conditionnellement aux trajectoires $(Z_n(0))_{n \geq 0}, \dots, (Z_n(k-1))_{n \geq 0}$, les processus $(X_n)_{n \geq 0}$ et $(Z_n(k))_{k \geq 0}$ ont même loi.

Nous construisons ensuite une chaîne de naissance et mort inférieure $(X_n^L)_{n \geq 0}$ et une chaîne de naissance et mort supérieure $(X_n^U)_{n \geq 0}$ pour borner stochastiquement le processus $(X_n)_{n \geq 0}$ lorsque $(Z_n)_{n \geq 0}$ est dans $U_{k-1}(\delta')$. Pour $\delta' > 0$, soit $W_{k-1}(\delta')$ le sous-ensemble de $[0, 1]^k$ défini par

$$W_{k-1}(\delta') = \{x \in [0, 1]^k : |x_i - \rho_i| < \delta', 0 \leq i < k\}.$$

Remarquons que

$$W_{k-1}(\delta') \cap \frac{\mathbb{Z}^k}{m} = U_{k-1}(\delta').$$

Soit $(X_n^L)_{n \geq 0}$ une chaîne de naissance et mort à valeurs dans $\{0, \dots, m\}$ de probabilités de transition

$$\begin{aligned} \delta_i^L &= \min\{\delta_i(x) : x \in W_{k-1}(\delta')\}, & 0 \leq i \leq m, \\ \gamma_i^L &= \max\{\gamma_i(x) : x \in W_{k-1}(\delta')\}, & 0 \leq i \leq m. \end{aligned}$$

Également, soit $(X_n^U)_{n \geq 0}$ la chaîne de naissance et mort à valeurs dans l'ensemble $\{0, \dots, m\}$ avec probabilités de transition

$$\begin{aligned} \delta_i^U &= \max\{\delta_i(x) : x \in W_{k-1}(\delta')\} & 0 \leq i \leq m, \\ \gamma_i^U &= \min\{\gamma_i(x) : x \in W_{k-1}(\delta')\} & 0 \leq i \leq m. \end{aligned}$$

Vérifions que les processus $(X_n^L)_{n \geq 0}$ et $(X_n^U)_{n \geq 0}$ sont bien définis, i.e.,

$$\forall i \in \{1, \dots, m-1\} \quad \delta_i^L + \gamma_i^L \leq 1 \quad \text{et} \quad \delta_i^U + \gamma_i^U \leq 1.$$

Soit $x_i \in W_{k-1}(\delta')$ un point qui maximise la fonction $\gamma_i(x)$. Nous avons

$$\delta_i^L + \gamma_i^L = \delta_i^L + \gamma_i(x_i) \leq \delta_i(x_i) + \gamma_i(x_i) \leq 1.$$

De manière analogue, si $x'_i \in W_{k-1}(\delta')$ est un point qui maximise la fonction $\delta_i(x)$, nous avons

$$\delta_i^U + \gamma_i^U = \delta_i(x'_i) + \gamma_i^U \leq \delta_i(x'_i) + \gamma_i(x'_i) \leq 1.$$

Ainsi les chaînes de naissance et mort $(X_n^L)_{n \geq 0}$, $(X_n^U)_{n \geq 0}$ sont bien définies. Nous construisons ensuite un couplage pour pouvoir comparer les processus $(X_n)_{n \geq 0}$, $(X_n^L)_{n \geq 0}$, $(X_n^U)_{n \geq 0}$. Nous définissons des applications

$$C^L, C^U : \{0, \dots, m\} \times [0, 1] \longrightarrow [0, 1]$$

pour coupler les mécanismes de transition des processus $(X_n^L)_{n \geq 0}$ et $(X_n^U)_{n \geq 0}$. Soient i dans $\{0, \dots, m\}$ et u dans $[0, 1]$. Nous posons

$$\begin{aligned} C^L(i, u) &= i - 1_{u < \gamma_i^L} + 1_{u > 1 - \delta_i^L}, \\ C^U(i, u) &= i - 1_{u < \gamma_i^U} + 1_{u > 1 - \delta_i^U}. \end{aligned}$$

Les applications C^L, C^U sont définies de sorte que si U est une variable aléatoire de loi uniforme sur l'intervalle $[0, 1]$, alors

$$\begin{aligned} P(C^L(i, U) = i + 1) &= \delta_i^L & \text{et} & \quad P(C^L(i, U) = i - 1) = \gamma_i^L, \\ P(C^U(i, U) = i + 1) &= \delta_i^U & \text{et} & \quad P(C^U(i, U) = i - 1) = \gamma_i^U. \end{aligned}$$

Par construction des probabilités de transition $\delta_i^L, \gamma_i^L, \delta_i^U, \gamma_i^U$, nous voyons que l'application C^L est en-dessous de l'application C et que l'application C^U est au-dessus de l'application C au sens suivant :

$$\forall x \in W_{k-1}(\delta') \quad \forall i \in \{0, \dots, m\} \quad \forall u \in [0, 1] \\ C^L(i, u) \leq C(x, i, u) \leq C^U(i, u).$$

Nous définissons les processus $(X_n^L)_{n \geq 0}, (X_n^U)_{n \geq 0}$ à l'aide de la suite de variables aléatoires $(U_n)_{n \geq 1}$ comme suit. Soit $a \in \{0, \dots, m\}$ le point de départ des processus. Nous prenons $X_0^L = X_0^U = a$ et

$$\forall n \geq 1 \quad X_n^L = C(X_{n-1}^L, U_n), \quad X_n^U = C(X_{n-1}^U, U_n).$$

Soit τ le temps de sortie de l'ensemble $U_{k-1}(\delta')$:

$$\tau = \inf\{n \geq 0 : Z_n \notin U_{k-1}(\delta')\}.$$

Proposition 18.3.2. *Pour tout $z \in U_{k-1}(\delta')$, si $Z_0 = z$ et $X_0^L = X_0 = X_0^U = z_k$, alors*

$$\forall n \in [0, \tau], \quad X_n^L \leq X_n \leq X_n^U.$$

Démonstration. Nous allons montrer l'inégalité par récurrence sur $n \in \mathbb{N}$. Pour $n = 0$ nous avons bien $X_0 = X_0^L = X_0^U = z_k$. Supposons que l'inégalité est vraie au temps $n < \tau$, c'est à dire que $X_n^L \leq X_n \leq X_n^U$ et que $Z_n(0)/m, \dots, Z_n(k-1)/m$ sont tous dans $W_{k-1}(\delta')$. Nous avons

$$\begin{aligned} X_{n+1}^L &= C^L(X_n^L, U_{n+1}), \\ X_{n+1} &= C\left(\frac{Z_n(0)}{m}, \dots, \frac{Z_n(k-1)}{m}, X_n, U_{n+1}\right), \\ X_{n+1}^U &= C^U(X_n^U, U_{n+1}). \end{aligned}$$

Par le lemme 18.3.1 et l'hypothèse de récurrence, nous obtenons

$$\begin{aligned} C\left(\frac{Z_n(0)}{m}, \dots, \frac{Z_n(k-1)}{m}, X_n^L, U_{n+1}\right) \\ \leq C\left(\frac{Z_n(0)}{m}, \dots, \frac{Z_n(k-1)}{m}, X_n, U_{n+1}\right) \leq \\ C\left(\frac{Z_n(0)}{m}, \dots, \frac{Z_n(k-1)}{m}, X_n^U, U_{n+1}\right). \end{aligned}$$

L'application C^L est en-dessous de l'application C et l'application C^U est au-dessus de l'application C , donc,

$$\begin{aligned} C^L(X_n^L, U_{n+1}) &\leq C\left(\frac{Z_n(0)}{m}, \dots, \frac{Z_n(k-1)}{m}, X_n^L, U_{n+1}\right), \\ C^U(X_n^U, U_{n+1}) &\geq C\left(\frac{Z_n(0)}{m}, \dots, \frac{Z_n(k-1)}{m}, X_n^U, U_{n+1}\right). \end{aligned}$$

Nous combinons les inégalités précédentes et nous obtenons ainsi l'inégalité au temps $n + 1$:

$$X_{n+1}^L \leq X_{n+1} \leq X_{n+1}^U.$$

□

Soit $\delta > 0$ et définissons

$$V_k(\delta) = \left\{ i \in \{0, \dots, m\} : \left| \frac{i}{m} - \rho_k \right| < \delta \right\}.$$

Soient aussi

$$\begin{aligned} \tau^L(V_k(\delta)) &= \inf \{ n \geq 0 : X_n^L \in V_k(\delta) \}, \\ \tau^L(V_k(2\delta)^c) &= \inf \{ n \geq 0 : X_n^L \notin V_k(2\delta) \}, \\ \tau^U(V_k(\delta)) &= \inf \{ n \geq 0 : X_n^U \in V_k(\delta) \}, \\ \tau^U(V_k(2\delta)^c) &= \inf \{ n \geq 0 : X_n^U \notin V_k(2\delta) \}. \end{aligned}$$

Soit $\delta > 0$. Nous allons utiliser le résultat suivant pour finir la démonstration du théorème 18.0.1.

Proposition 18.3.3. *Soit $\delta > 0$. Il existe $\alpha_k, \alpha'_k, \beta_k, \beta'_k$ des réels strictement positifs (dépendant de δ), tels que pour ℓ, m suffisamment grands et q, δ' suffisamment petits*

- Pour tout $i \in \{0, \dots, m\}$,

$$\begin{aligned} P(\tau^L(V_k(\delta)) \geq m^{\alpha_k} \mid X_0^L = i) &\leq \exp(-\alpha'_k m), \\ P(\tau^U(V_k(\delta)) \geq m^{\alpha_k} \mid X_0^U = i) &\leq \exp(-\alpha'_k m). \end{aligned}$$

- Pour tout $i \in V_k(\delta)$,

$$\begin{aligned} P(\tau^L(V_k(2\delta)^c) \leq \exp(\beta_k m) \mid X_0^L = i) &\leq \exp(-\beta'_k m), \\ P(\tau^U(V_k(2\delta)^c) \leq \exp(\beta_k m) \mid X_0^U = i) &\leq \exp(-\beta'_k m). \end{aligned}$$

La preuve est la même pour les deux processus, par commodité nous allons travailler avec le processus $(X_n^L)_{n \geq 0}$.

18.4 Chaînes de naissance et mort

Nous étudions ici la dynamique du processus $(X_n^L)_{n \geq 0}$. Pour alléger l'écriture nous enlevons le L des notations dans toute cette section. Cherchons d'abord les points x de $W_{k-1}(\delta')$ qui réalisent les valeurs δ_i^L et γ_i^L . Comme

$$\forall l \in \{1, \dots, k-1\} \quad M_H(l, k) \geq M_H(\theta, k),$$

nous voyons que $\delta_i(x_0, \dots, x_{k-1})$ est croissante en les variables x_1, \dots, x_{k-1} , et $\gamma_i(x_0, \dots, x_{k-1})$ est décroissante en x_1, \dots, x_{k-1} . Il s'ensuit que pour tout $i \in \{0, \dots, m\}$,

$$\begin{aligned} \delta_i &= \min_{x_0: |x_0 - \rho_0| \leq \delta'} \delta_i(x_0, \rho_1 - \delta', \dots, \rho_{k-1} - \delta'), \\ \gamma_i &= \max_{x_0: |x_0 - \rho_0| \leq \delta'} \gamma_i(x_0, \rho_1 - \delta', \dots, \rho_{k-1} - \delta'). \end{aligned}$$

Pour étudier le comportement par rapport à x_0 , nous dérivons :

$$\begin{aligned} \frac{\partial \delta_i(x_0, \dots, x_{k-1})}{\partial x_0} &= \frac{1 - i/m}{((\sigma - 1)x_0 + 1)^2} \left(\sigma (M_H(0, k) - M_H(\theta, k)) \right. \\ &\quad \left. - (\sigma - 1) \sum_{l=1}^{k-1} x_l (M_H(l, k) - M_H(\theta, k)) - (\sigma - 1) \frac{i}{m} (M_H(k, k) - M_H(\theta, k)) \right), \end{aligned}$$

$$\frac{\partial \gamma_i(x_0, \dots, x_{k-1})}{\partial x_0} = \frac{i/m}{((\sigma-1)x_0 + 1)^2} \left(-\sigma \left(M_H(0, k) - M_H(\theta, k) \right) + \right. \\ \left. (\sigma-1) \sum_{l=1}^{k-1} x_l \left(M_H(l, k) - M_H(\theta, k) \right) + (\sigma-1) \frac{i}{m} \left(M_H(k, k) - M_H(\theta, k) \right) \right).$$

Nous voyons que le signe des dérivées partielles de $\delta_i(x)$ et $\gamma_i(x)$ par rapport à x_0 ne dépend pas de x_0 . Les variables x_1, \dots, x_{k-1} étant fixées, les fonctions $\delta_i(x)$ et $\gamma_i(x)$ sont monotones vis à vis de x_0 . De plus, les dérivées partielles de $\delta_i(x)$ et $\gamma_i(x)$ par rapport à x_0 ont des signes opposés et

$$\frac{\partial \delta_i(x_0, \dots, x_{k-1})}{\partial x_0} = 0 \iff \frac{\partial \gamma_i(x_0, \dots, x_{k-1})}{\partial x_0} = 0 \\ \iff \sigma \left(M_H(0, k) - M_H(\theta, k) \right) - (\sigma-1) \sum_{l=1}^{k-1} x_l \left(M_H(l, k) - M_H(\theta, k) \right) \\ - (\sigma-1) \frac{i}{m} \left(M_H(k, k) - M_H(\theta, k) \right) = 0.$$

Nous supposons que

$$\ell \rightarrow +\infty, \quad m \rightarrow +\infty, \quad q \rightarrow 0,$$

de sorte que

$$\ell q \rightarrow a \in]0, +\infty[.$$

Nous avons alors

$$\lim_{\substack{\ell \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a}} M_H(l, k) = \begin{cases} \frac{a^{k-l}}{(k-l)!} e^{-a} & \text{si } l \leq k, \\ 0 & \text{si } l = \theta. \end{cases}$$

Pour ℓ assez grand et q assez petit, le coefficient $M_H(l, k) - M_H(\theta, k)$ est strictement positif. Comme l'équation ci-dessus est linéaire en i , nous obtenons la conclusion suivante. Il existe une valeur $i^* \in \{0, \dots, m\}$, qui dépend de $m, \rho_1, \dots, \rho_{k-1}, \delta'$, telle que :

• Si $0 \leq i \leq i^*$, la fonction $\delta_i(x_0, \rho_1 - \delta', \dots, \rho_{k-1} - \delta')$ est décroissante en x_0 , la fonction $\gamma_i(x_0, \rho_1 - \delta', \dots, \rho_{k-1} - \delta')$ est croissante en x_0 , et

$$\delta_i = \delta_i(\rho_0 + \delta', \rho_1 - \delta', \dots, \rho_{k-1} - \delta'), \\ \gamma_i = \gamma_i(\rho_0 + \delta', \rho_1 - \delta', \dots, \rho_{k-1} - \delta').$$

De plus, la fonction $\phi(\beta, \varepsilon, x, \rho)$ est continue et croissante en les variables β, ε . Soit $\psi :]0, 1]^{k+1} \times [0, 1[\times]0, 1[\longrightarrow]0, +\infty[$ la fonction donnée par :

$$\forall \beta \in]0, 1]^{k+1} \quad \forall \varepsilon \in [0, 1[\quad \forall \rho \in]0, 1[$$

$$\psi(\beta, \varepsilon, \rho) = \begin{cases} \phi\left(\beta, \varepsilon, \rho_0 + \delta', \rho_1 - \delta', \dots, \rho_{k-1} - \delta', \rho\right) & \text{si } \rho \leq i^*/m, \\ \phi\left(\beta, \varepsilon, \rho_0 - \delta', \rho_1 - \delta', \dots, \rho_{k-1} - \delta', \rho\right) & \text{si } \rho > i^*/m. \end{cases}$$

Le résultat suivant donne des estimées de grandes déviations pour les produits $\pi(i)$.

Proposition 18.4.1. Soit $a \in]0, +\infty[$. Pour $\rho \in [0, 1]$, nous avons

$$\lim_{\substack{\ell, m \rightarrow \infty \\ q \rightarrow 0, \ell q \rightarrow a}} \frac{1}{m} \ln \pi(\lfloor \rho m \rfloor) = \int_0^\rho \ln \psi\left(e^{-a} \frac{a^k}{k!}, \dots, e^{-a}, 0, s\right) ds.$$

Démonstration. Soit $\rho \in [0, 1]$. Pour $m \geq 1$, nous avons

$$\frac{1}{m} \ln \pi(\lfloor \rho m \rfloor) = \frac{1}{m} \sum_{i=1}^{\lfloor \rho m \rfloor} \ln \psi\left(M_H(0, k), \dots, M_H(k, k), M_H(\theta, k), \frac{i}{m}\right).$$

Soit $\varepsilon > 0$ tel que

$$\varepsilon < \min_{0 \leq l \leq k} e^{-a} \frac{a^l}{l!}.$$

Pour ℓ, m suffisamment grands et q suffisamment petit, nous avons

$$\left| M_H(l, k) - e^{-a} \frac{a^{k-l}}{(k-l)!} \right| < \varepsilon \quad 0 \leq l \leq k,$$

$$0 < M_H(\theta, k) < \varepsilon.$$

De plus, comme l'application ϕ est monotone vis-à-vis des deux premières variables,

$$\begin{aligned} \frac{1}{m} \sum_{i=1}^{\lfloor \rho m \rfloor} \ln \psi\left(e^{-a} \frac{a^k}{k!} - \varepsilon, \dots, e^{-a} - \varepsilon, 0, \frac{i}{m}\right) \\ \leq \frac{1}{m} \ln \pi(\lfloor \rho m \rfloor) \leq \\ \frac{1}{m} \sum_{i=1}^{\lfloor \rho m \rfloor} \ln \psi\left(e^{-a} \frac{a^k}{k!} + \varepsilon, \dots, e^{-a} + \varepsilon, \varepsilon, \frac{i}{m}\right). \end{aligned}$$

Nous reconnaissons des sommes de Riemann. En faisant tendre ℓ, m vers $l' \infty$ et q vers 0, nous obtenons

$$\liminf_{\substack{\ell, m \rightarrow \infty \\ q \rightarrow 0, \ell q \rightarrow a}} \frac{1}{m} \ln \pi(\lfloor \rho m \rfloor) \geq \int_0^\rho \ln \psi \left(e^{-a} \frac{a^k}{k!} - \varepsilon, \dots, e^{-a} - \varepsilon, 0, s \right) ds,$$

$$\limsup_{\substack{\ell, m \rightarrow \infty \\ q \rightarrow 0, \ell q \rightarrow a}} \frac{1}{m} \ln \pi(\lfloor \rho m \rfloor) \leq \int_0^\rho \ln \psi \left(e^{-a} \frac{a^k}{k!} + \varepsilon, \dots, e^{-a} + \varepsilon, \varepsilon, s \right) ds.$$

Nous envoyons ε vers 0 et nous obtenons le résultat cherché. \square

Définissons

$$\rho^- = \min \left\{ \rho \left(e^{-a} \frac{a^k}{k!}, \dots, e^{-a}, 0, \rho_0 + \delta', \rho_1 - \delta', \dots, \rho_{k-1} - \delta' \right), \right. \\ \left. \rho \left(e^{-a} \frac{a^k}{k!}, \dots, e^{-a}, 0, \rho_0 - \delta', \rho_1 - \delta', \dots, \rho_{k-1} - \delta' \right) \right\},$$

$$\rho^+ = \max \left\{ \rho \left(e^{-a} \frac{a^k}{k!}, \dots, e^{-a}, 0, \rho_0 + \delta', \rho_1 - \delta', \dots, \rho_{k-1} - \delta' \right), \right. \\ \left. \rho \left(e^{-a} \frac{a^k}{k!}, \dots, e^{-a}, 0, \rho_0 - \delta', \rho_1 - \delta', \dots, \rho_{k-1} - \delta' \right) \right\}.$$

Nous avons alors

$$\psi \left(e^{-a} \frac{a^k}{k!}, \dots, e^{-a}, 0, s \right) > 1 \quad \text{pour } s < \rho^-,$$

$$\psi \left(e^{-a} \frac{a^k}{k!}, \dots, e^{-a}, 0, s \right) < 1 \quad \text{pour } s > \rho^+,$$

et la fonction

$$\rho \mapsto \int_0^\rho \ln \psi \left(e^{-a} \frac{a^k}{k!}, \dots, e^{-a}, 0, s \right) ds$$

est croissante sur $]0, \rho^-[$ et décroissante sur $]\rho^+, 1[$. De plus, lorsque δ' tend vers 0, les points ρ^- et ρ^+ convergent vers ρ_k :

$$\lim_{\delta' \rightarrow 0} \rho^- = \lim_{\delta' \rightarrow 0} \rho^+ = \rho_k.$$

Définissons

$$r^- = \min \left\{ \rho(M_H(0, k), \dots, M_H(k, k), M_H(\theta, k), \rho_0 + \delta', \rho_1 - \delta', \dots, \rho_{k-1} - \delta'), \right. \\ \left. \rho(M_H(0, k), \dots, M_H(k, k), M_H(\theta, k), \rho_0 - \delta', \rho_1 - \delta', \dots, \rho_{k-1} - \delta') \right\},$$

$$r^+ = \max \left\{ \rho(M_H(0, k), \dots, M_H(k, k), M_H(\theta, k), \rho_0 + \delta', \rho_1 - \delta', \dots, \rho_{k-1} - \delta'), \right. \\ \left. \rho(M_H(0, k), \dots, M_H(k, k), M_H(\theta, k), \rho_0 - \delta', \rho_1 - \delta', \dots, \rho_{k-1} - \delta') \right\}.$$

Nous avons donc

$$\begin{aligned} 1 \leq i \leq j \leq r^- m &\implies \pi(i) \leq \pi(j), \\ r^+ m \leq i \leq j \leq m &\implies \pi(i) \geq \pi(j), \end{aligned}$$

et la situation est plus délicate entre $r^- m$ et $r^+ m$. Or, quand $\ell, m \rightarrow \infty, q \rightarrow 0$ et $\ell q \rightarrow a$, nous avons

$$\lim_{\substack{\ell, m \rightarrow \infty \\ q \rightarrow 0, \ell q \rightarrow a}} r^- = \rho^-, \quad \lim_{\substack{\ell, m \rightarrow \infty \\ q \rightarrow 0, \ell q \rightarrow a}} r^+ = \rho^+,$$

Nous cherchons ensuite des estimées sur le temps d'entrée dans un voisinage de ρ_k . Prenons $\delta > 0$ et définissons

$$V_k(\delta) = \left\{ x \in \{0, \dots, m\} : \left| \frac{x}{m} - \rho_k \right| < \delta \right\}.$$

Soit $\tau(V_k(\delta))$ le temps d'entrée dans $V_k(\delta)$,

$$\tau(V_k(\delta)) = \inf \{ n \geq 0 : X_n \in V_k(\delta) \}.$$

Nous allons démontrer le résultat suivant

Proposition 18.4.2. *Soit $\delta > 0$. Ils existe réels strictement positifs α_k, α'_k (qui dépendent de δ) tels que pour ℓ, m suffisamment grands et q, δ' suffisamment petits,*

$$\forall x \in \{0, \dots, m\} \quad P(\tau(V_k(\delta)) \geq m^{\alpha_k} \mid X_0 = x) \leq \exp(-\alpha'_k m).$$

Démonstration. Par monotonie il suffit de faire la preuve pour $x = 0$ et $x = m$. Soient b et c les extrémités de l'intervalle $V_k(\delta)$,

$$b = \lfloor (\rho_k - \delta)m \rfloor + 1, \quad c = \lfloor (\rho_k + \delta)m \rfloor.$$

Grâce aux formules de la section 18.1, nous avons

$$E(\tau(V_k(\delta)) \mid X_0 = 0) = E(\tau(\{b\}) \mid X_0 = 0) = \sum_{i=0}^{b-1} \sum_{j=i}^{b-1} \frac{1}{\delta_i} \frac{\pi(i)}{\pi(j)}.$$

Pour ℓ, m suffisamment grands et q, δ' suffisamment petits, le point r^- est dans $V_k(\delta)$, et aussi $b \leq r^-$. Comme les produits $\pi(i)$ sont croissants sur l'intervalle $\{1, \dots, \lfloor r^- m \rfloor\}$, nous obtenons que si $1 \leq i \leq j \leq b - 1$, alors

$$\frac{\pi(i)}{\pi(j)} \leq 1.$$

De plus, pour $m \geq 2$,

$$\forall i \in \{0, \dots, m-1\} \quad \delta_i \geq \frac{\sigma(\rho_0 - \delta') M_H(0, k)}{m((\sigma - 1)(\rho_0 + \delta) + 1)0} \geq \frac{c}{m},$$

pour une constante c strictement positive qui dépend de k, δ' mais pas de m . Il s'ensuit que

$$E(\tau(V_k(\delta)) \mid X_0 = 0) \leq cm^4.$$

Soit $\alpha > 4$, l'inégalité de Markov entraîne

$$P(\tau(V_k(\delta)) \geq m^\alpha \mid X_0 = 0) \leq cm^{-(\alpha-4)}.$$

Nous développons l'expression

$$P(\tau(V_k(\delta)) \geq m^{\alpha+1} \mid X_0 = 0)$$

comme dans la preuve de la proposition 18.1.2 et nous obtenons

$$\begin{aligned} P(\tau(V_k(\delta)) \geq m^{\alpha+1} \mid X_0 = 0) &\leq \left(cm^{-(\alpha-4)} \right)^m \\ &= \exp\left(-m((\alpha-4) \ln m - \ln c) \right). \end{aligned}$$

Choisissons $\alpha_k > 5$ et soit m_0 suffisamment grand pour avoir

$$(\alpha_k - 4) \ln m_0 - \ln c > 0.$$

Nous choisissons $\alpha'_k = (\alpha_k - 4) \ln m_0 - \ln c$ et nous obtenons le résultat voulu. La preuve pour l'état initial $X_0 = m$ est similaire. \square

Nous cherchons ensuite des estimées sur le temps de sortie d'un voisinage de ρ_k . Soit $\delta > 0$. Soit $\tau(V_k(2\delta)^c)$ le temps de sortie de l'ensemble $V_k(2\delta)$:

$$\tau(V_k(2\delta)^c) = \inf \{ n \geq 0 : X_n \notin V \}.$$

Nous allons démontrer le résultat suivant :

Proposition 18.4.3. Soit $\delta > 0$. Il existe réels strictement positifs β_k, β'_k (qui dépendent de δ) tels que pour ℓ, m suffisamment grands et q, δ' suffisamment petits

$$\forall x \in V_k(\delta) \quad P(\tau(V_k(2\delta)^c) < \exp(\beta_k m) \mid X_0 = x) \leq \exp(-\beta'_k m).$$

Démonstration. Soit $t > 0$ et regardons d'abord la valeur de

$$P(\tau(V_k(2\delta)^c) \leq t \mid X_0 = x).$$

Soit θ le dernier temps avant $\tau(V_k(2\delta)^c)$ où le processus $(X_n)_{n \geq 0}$ est dans $V_k(\delta)$, i.e.,

$$\theta = \max \{ n < \tau(V_k(2\delta)^c) : X_n \in V_k(\delta) \}.$$

Soient b et c les extrémités de l'intervalle $V_k(\delta)$,

$$b = \lfloor (\rho_k - \delta)m \rfloor + 1, \quad c = \lfloor (\rho_k + \delta)m \rfloor.$$

Également, soient b' et c' les extrémités de l'intervalle $V_k(2\delta)$,

$$b' = \lfloor (\rho_k - 2\delta)m \rfloor + 1, \quad c' = \lfloor (\rho_k + 2\delta)m \rfloor.$$

Pour ℓ, m assez grands et q assez petit, nous avons

$$\begin{aligned} P(\tau(V_k(2\delta)^c) \leq t \mid X_0 = x) &= \sum_{n < t} P(\theta = n, \tau(V_k(2\delta)^c) \leq t \mid X_0 = x) \\ &= \sum_{n < t} \left(P(\theta = n, X_{n+1} = b - 1, \tau(V_k(2\delta)^c) \leq t \mid X_0 = x) \right. \\ &\quad \left. + P(\theta = n, X_{n+1} = c + 1, \tau(V_k(2\delta)^c) \leq t \mid X_0 = x) \right). \end{aligned}$$

Développons le premier terme dans la parenthèse. Par la propriété de Markov, nous avons

$$\begin{aligned} &P(\theta = n, X_{n+1} = b - 1, \tau(V_k(2\delta)^c) \leq t \mid X_0 = x) \\ &= P\left(X_n = b, X_{n+1} = b - 1, \tau(V_k(2\delta)^c) \leq t, \right. \\ &\quad \left. X_r \notin V_k(\delta) \text{ pour } n < r \leq \tau(V_k(2\delta)^c) \mid X_0 = x \right) \\ &\leq P(X_r \notin V_k(\delta) \text{ pour } r \leq \tau(V_k(2\delta)^c), \tau(V_k(2\delta)^c) \leq t - n - 1 \mid X_0 = b - 1) \\ &\leq P(X_{\tau(V_k(2\delta)^c) \cup \{b\}} \in V_k(2\delta)^c \mid X_0 = b - 1) \\ &= P(X_{\tau(\{b'-1, b\})} = b' - 1 \mid X_0 = b - 1). \end{aligned}$$

Nous utilisons maintenant la formule donnée dans la section 18.1 pour écrire

$$P(X_{\tau(\{b'-1, b\})} = b' - 1 \mid X_0 = b - 1) = \frac{1}{\frac{\pi(b-1)}{\sum_{i=b'-1}^{b-1} \frac{1}{\pi(i)}}}.$$

Nous en déduisons que

$$P(X_{\tau(\{b'-1, b\})} = b' - 1 \mid X_0 = b - 1) \leq \frac{\pi(b' - 1)}{\pi(b - 1)}.$$

Soit $\varepsilon > 0$ et soient ℓ, m assez grands et q assez petit de sorte que

$$\left| \frac{1}{m} \ln \pi(b - 1) - \int_0^{\rho_k - \delta} \ln \psi \left(e^{-a} \frac{a^k}{k!}, \dots, e^{-a}, 0, s \right) ds \right| < \frac{\varepsilon}{2},$$

$$\left| \frac{1}{m} \ln \pi(b' - 1) - \int_0^{\rho_k - 2\delta} \ln \psi \left(e^{-a} \frac{a^k}{k!}, \dots, e^{-a}, 0, s \right) ds \right| < \frac{\varepsilon}{2},$$

Nous avons donc

$$\frac{\pi(b' - 1)}{\pi(b - 1)} = \exp \left(m \left(\frac{1}{m} \ln \pi(b' - 1) - \frac{1}{m} \ln \pi(b - 1) \right) \right) \leq$$

$$\exp \left(m \left(- \int_{\rho_k - 2\delta}^{\rho_k - \delta} \ln \psi \left(e^{-a} \frac{a^k}{k!}, \dots, e^{-a}, 0, s \right) ds + \varepsilon \right) \right).$$

La fonction

$$\rho \mapsto \int_0^{\rho} \ln \psi \left(e^{-a} \frac{a^k}{k!}, \dots, e^{-a}, 0, s \right) ds$$

est croissante sur $]0, \rho^-[$. Or, ρ^- converge vers ρ_k lorsque δ' tend vers 0. En prenant δ' assez petit, nous avons $\rho_k - \delta < \rho^-$, donc

$$\int_{\rho_k - 2\delta}^{\rho_k - \delta} \ln \psi \left(e^{-a} \frac{a^k}{k!}, \dots, e^{-a}, 0, s \right) ds > 0.$$

Choisissons ε tel que

$$\varepsilon < \int_{\rho_k - 2\delta}^{\rho_k - \delta} \ln \psi \left(e^{-a} \frac{a^k}{k!}, \dots, e^{-a}, 0, s \right) ds.$$

Alors si nous prenons

$$\beta_1 = \int_{\rho_k - 2\delta}^{\rho_k - \delta} \ln \psi \left(e^{-a} \frac{a^k}{k!}, \dots, e^{-a}, 0, s \right) ds - \varepsilon,$$

nous avons $\beta_1 > 0$ et

$$P(\theta = n, X_{n+1} = b - 1, \tau(V_k(2\delta)^c) \leq t \mid X_0 = x) \leq \exp(-\beta_1 m).$$

Nous faisons un développement similaire pour le terme

$$P(\theta = n, X_{n+1} = c + 1, \tau(V_k(2\delta)^c) \leq t \mid X_0 = x)$$

et nous obtenons un $\beta_2 > 0$ tel que

$$P(\theta = n, X_{n+1} = c + 1, \tau(V_k(2\delta)^c) \leq t \mid X_0 = x) \leq \exp(-\beta_2 m).$$

Finalement,

$$P(\tau(V_k(2\delta)^c) \leq t \mid X_0 = x) \leq t \left(\exp(-\beta_1 m) + \exp(-\beta_2 m) \right).$$

Soit $\beta_k < \min(\beta_1, \beta_2)$. Nous avons donc

$$P(\tau_k(V(2\delta)^c) \leq \exp(\beta_k m) \mid X_0 = x) \leq 2 \exp(-\min(\beta_1 - \beta_k, \beta_2 - \beta_k)m).$$

En choisissant, pour m_0 suffisamment grand,

$$\beta'_k = \min(\beta_1 - \beta_k, \beta_2 - \beta_k) - \frac{1}{m_0} \ln 2$$

nous obtenons le résultat voulu. □

Chapitre 19

Convergence

Dans cette section nous montrons, dans le cas $\sigma e^{-\alpha} > 1$, la convergence de la mesure de probabilité invariante ν^θ vers la masse de Dirac en ρ . Soit $\sigma e^{-\alpha} > 1$. Soit ρ le point de $[0, 1]^{K+1}$ donné par :

$$\rho_k = (\sigma e^{-\alpha} - 1) \frac{\alpha^k}{k!} \sum_{i \geq 1} \frac{i^k}{\sigma^i}.$$

Nous considérons le régime asymptotique

$$\begin{aligned} \ell \rightarrow +\infty, \quad m \rightarrow +\infty, \quad q \rightarrow 0, \\ \ell q \rightarrow \alpha, \quad \frac{m}{\ell} \rightarrow \alpha. \end{aligned}$$

Théorème 19.0.4. *Pour toute fonction $f : [0, 1] \rightarrow \mathbb{R}$ continue, croissante et nulle en 0, nous avons*

$$\lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha, \frac{m}{\ell} \rightarrow \alpha}} \int_{\mathbb{D}^{K+1}} f\left(\frac{z_0 + \dots + z_K}{m}\right) d\nu^\theta(z) = f(\rho_0 + \dots + \rho_K).$$

Démonstration. Soit $\delta > 0$ et définissons

$$U_K(\delta) = \left\{ z \in \mathbb{D}^{K+1} : \left| \frac{z_k}{m} - \rho_k \right| < \delta, \quad 0 \leq k \leq K \right\}.$$

Nous définissons deux suites de temps d'arrêt $(T_n)_{n \geq 0}$ et $(T_n^*)_{n \geq 1}$. Posons $T_0 = 0$

et soient

$$\begin{aligned} T_1^* &= \inf \{ n \geq 0 : Z_n \in U_K(\delta) \}, & T_1 &= \inf \{ n \geq T_1^* : Z_n \notin U_K(2\delta) \}, \\ \vdots & & \vdots & \\ T_i^* &= \inf \{ n \geq T_{i-1} : Z_n \in U_K(\delta) \}, & T_i &= \inf \{ n \geq T_i^* : Z_n \notin U_K(2\delta) \}, \\ \vdots & & \vdots & \end{aligned}$$

D'après le théorème ergodique pour les chaînes de Markov, nous avons

$$\int_{\mathbb{D}^{K+1}} f\left(\frac{z_0 + \cdots + z_K}{m}\right) d\nu^\theta(z) = \lim_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left(\sum_{i=0}^{t-1} f\left(\frac{Z_i(0) + \cdots + Z_i(K)}{m}\right) \right).$$

Soit $t \geq 0$. Nous décomposons cette dernière somme à l'aide des suites de temps d'arrêt $(T_i)_{i \geq 0}$ et $(T_i^*)_{i \geq 1}$:

$$\begin{aligned} \sum_{n=0}^t f\left(\frac{Z_n(0) + \cdots + Z_n(K)}{m}\right) &= \\ & \sum_{i \geq 1} \sum_{n=T_{i-1} \wedge t}^{T_i^* \wedge t-1} f\left(\frac{Z_n(0) + \cdots + Z_n(K)}{m}\right) \\ & \quad + \sum_{i \geq 1} \sum_{n=T_i^* \wedge t}^{T_i \wedge t-1} f\left(\frac{Z_n(0) + \cdots + Z_n(K)}{m}\right), \end{aligned}$$

où le symbole \wedge représente le minimum de deux quantités. La fonction f est continue, croissante et nulle en 0. Soit $\varepsilon > 0$ et choisissons δ suffisamment petit pour que

$$\forall z \in U_K(\delta) \quad \left| f\left(\frac{z_0 + \cdots + z_K}{m}\right) - f(\rho_0 + \cdots + \rho_K) \right| < \frac{\varepsilon}{2}.$$

Nous avons alors

$$\begin{aligned} \left| \mathbb{E} \left(\sum_{n=0}^{t-1} f\left(\frac{Z_n(0) + \cdots + Z_n(K)}{m}\right) \right) - t f(\rho_0 + \cdots + \rho_K) \right| &\leq \\ & \sum_{i \geq 1} 2f(1) \mathbb{E}(T_i^* \wedge t - T_{i-1} \wedge t) + \frac{t\varepsilon}{2}. \end{aligned}$$

Nous étudions ensuite l'expression

$$\sum_{i \geq 1} (T_i^* \wedge t - T_{i-1} \wedge t).$$

Définissons

$$N(t) = \max \{ i \geq 0 : T_i \leq t \}.$$

nous avons

$$\sum_{i \geq 1} (T_i^* \wedge t - T_{i-1} \wedge t) \leq \sum_{i=1}^{N(t)} (T_i^* - T_{i-1}) + (t - T_{N(t)}).$$

Nous étudions ensuite la variable aléatoire $N(t)$. Soient $j \in \mathbb{N}$ et $z \in \mathbb{D}^{k+1}$. Nous cherchons des estimées sur la valeur de

$$P(N(j \exp(\alpha m)/2) \geq j \mid Z_0 = z).$$

Par définition de $N(t)$, nous avons

$$N(t) \geq i \iff T_i \leq t.$$

De plus,

$$T_i = T_1 + (T_2 - T_1) + \dots + (T_i - T_{i-1}).$$

Définissons pour $i \geq 1$,

$$\begin{aligned} Y_i &= T_i - T_{i-1}, \\ Y_i^* &= T_i - T_i^*. \end{aligned}$$

Nous définissons aussi pour $j \geq 1$,

$$S_j = \sum_{i=1}^j Y_i.$$

Nous voyons donc que

$$P(N(j \exp(\alpha m)/2) \geq j \mid Z_0 = z) = P(S_j \leq j \exp(\alpha m)/2 \mid Z_0 = z).$$

Par le théorème 18.0.1, il existe des réels strictement positifs α et α' tels que, pour tout $i \geq 1$,

$$\begin{aligned} P(Y_i \leq \exp(\alpha m) \mid Z_0 = z) &\leq P(Y_i^* \leq \exp(\alpha m) \mid Z_0 = z) \\ &= \sum_{z' \in \mathcal{U}_k(\delta)} P(Y_i^* \leq \exp(\alpha m) \mid Z_0 = z, Z_{T_i^*} = z') P(Z_{T_i^*} = z' \mid Z_0 = z) \\ &= \sum_{z' \in \mathcal{U}_k(\delta)} P(T_1^* \leq \exp(\alpha m) \mid Z_0 = z') P(Z_{T_1^*} = z' \mid Z_0 = z) \\ &\leq \exp(-\alpha' m). \end{aligned}$$

Définissons les variables aléatoires de loi Bernoulli suivantes :

$$\forall i \geq 1, \quad \varepsilon_i = 1_{Y_i^* \leq \exp(\alpha m)}.$$

Nous voyons que si $S_n \leq n \exp(\alpha m)/2$, alors il existe au moins $n/2$ indices dans $\{1, \dots, n\}$ tels que $Y_i^* < Y_i \leq \exp(\alpha m)$. Donc,

$$S_j \leq \frac{j \exp(\alpha m)}{2} \implies \varepsilon_1 + \dots + \varepsilon_j \geq \frac{j}{2}.$$

Ainsi,

$$\begin{aligned} P\left(N(j \exp(\alpha m)/2) \geq j \mid Z_0 = z\right) &= \\ P\left(S_j \leq j \exp(\alpha m)/2 \mid Z_0 = z\right) &\leq \\ P\left(\varepsilon_1 + \dots + \varepsilon_j \geq j/2\right). & \end{aligned}$$

Par l'inégalité de Chebyshev exponentielle nous avons, pour $\lambda \geq 0$

$$P\left(\varepsilon_1 + \dots + \varepsilon_j \geq j/2\right) \leq \exp\left(-\lambda/2 + \ln E\left(\exp(\lambda \varepsilon_1/j) \cdots \exp(\lambda \varepsilon_j/j)\right)\right).$$

Or, comme les variables aléatoires $\varepsilon_1, \dots, \varepsilon_{j-1}$ sont mesurables par rapport à la filtration générée par $(Z_n, 0 \leq n \leq \tau_j^*)$,

$$\begin{aligned} E\left(\exp(\lambda \varepsilon_1/j) \cdots \exp(\lambda \varepsilon_j/j)\right) &= \\ E\left(E\left(\exp(\lambda \varepsilon_1/j) \cdots \exp(\lambda \varepsilon_j/j) \mid Z_n, 0 \leq n \leq \tau_j^*\right)\right) &= \\ E\left(\exp(\lambda \varepsilon_1/j) \cdots \exp(\lambda \varepsilon_{j-1}/j) E\left(\exp(\lambda \varepsilon_j/j) \mid Z_n, 0 \leq n \leq \tau_j^*\right)\right). & \end{aligned}$$

Par la propriété de Markov forte, nous avons pour l'espérance conditionnelle

$$E\left(\exp(\lambda \varepsilon_j/j) \mid Z_n, 0 \leq n \leq \tau_j^*\right) = E\left(\exp(\lambda \varepsilon_1/j) \mid Z_0 = Z_{\tau_j^*}\right).$$

Or, pour tout z' dans $\mathcal{U}_K(\delta)$,

$$E\left(\exp(\lambda \varepsilon_1/j) \mid Z_0 = z'\right) \leq \exp\left(-\alpha' m + \frac{\lambda}{j}\right) + 1 - \exp(-\alpha' m),$$

d'où,

$$E\left(\exp(\lambda \varepsilon_1/n) \mid Z_0 = Z_{\tau_j^*}\right) \leq \exp\left(-\alpha' m + \frac{\lambda}{j}\right) + 1 - \exp(-\alpha' m).$$

Nous itérons ce procédé et nous obtenons

$$E(\exp(\lambda\varepsilon_1/j) \cdots \exp(\lambda\varepsilon_j/j)) \leq \left(\exp\left(-\alpha'm + \frac{\lambda}{j}\right) + 1 - \exp(-\alpha'm) \right)^j.$$

Nous faisons le changement de variable $\lambda \rightarrow j\lambda$ pour obtenir

$$\begin{aligned} P(\varepsilon_1 + \cdots + \varepsilon_j \geq j/2) \\ \leq \exp\left(-j\left(\lambda/2 - \ln\left(\exp(-\alpha'm + \lambda) + 1 - \exp(-\alpha'm)\right)\right)\right) \leq \\ \exp(-j\Lambda^*(1/2)), \end{aligned}$$

où $\Lambda^*(t)$ est la transformée de Cramér de la loi de Bernoulli de paramètre $p = \exp(-\alpha'm)$:

$$\Lambda^*(t) = \sup_{\lambda \geq 0} \left(\lambda t - \ln E(\exp(\lambda \text{Ber}(p))) \right).$$

La fonction $\Lambda^*(t)$ se met sous la forme

$$\Lambda^*(t) = t \ln \frac{t}{p} + (1-t) \ln \frac{1-t}{1-p}.$$

Nous voyons donc que dans notre cas,

$$\Lambda^*(1/2) = \frac{1}{2} \ln \frac{\exp(\alpha'm)}{2} + \frac{1}{2} \ln \frac{\exp(\alpha'm)}{2(\exp(\alpha'm) - 1)} \geq c(m),$$

où $c(m)$ est une constante positive qui dépend de m mais pas de j . Il s'ensuit que, pour m suffisamment grand,

$$\forall j \geq 1 \quad P\left(N(j \exp(\alpha m)/2) \geq j \mid Z_0 = z\right) \leq \exp(-jc(m)),$$

où pour m suffisamment grand, $c(m)$ est une constante positive qui dépend de m mais pas de j . Soit $t \geq 0$. Nous cherchons une borne supérieure pour l'espérance

$$E\left(\sum_{k=1}^{N(t)} (T_k^* - T_{k-1}) + (t - T_{N(t)})\right).$$

Puisque la somme dans la parenthèse ne peut pas dépasser la valeur t , nous avons pour $j \geq 1$

$$\begin{aligned} \mathbb{E} \left(\sum_{k=1}^{N(t)} (T_k^* - T_{k-1}) + (t - T_{N(t)}) \right) &\leq \\ &\mathbb{E} \left(\left(\sum_{k=1}^{N(t)} (T_k^* - T_{k-1}) + (t - T_{N(t)}) \right) 1_{N(t) < j} \right) + tP(N(t) \geq j). \end{aligned}$$

Soit

$$j_t = \min \left\{ j \in \mathbb{N} : t \leq \frac{j \exp(\alpha m)}{2} \right\}.$$

Grâce à l'étude de la variable aléatoire $N(j \exp(\alpha m)/2)$, en prenant $j = j_t$, nous voyons que le deuxième terme est borné par

$$\frac{j_t}{2} e^{\alpha m} P(N(j_t e^{\alpha m}/2) \geq j_t) \leq \frac{j_t}{2} \exp(\alpha m - j_t c(m)),$$

qui tend vers 0 quand t tend vers $l'\infty$. Par ailleurs, pour le premier terme, nous avons grâce au théorème 18.0.1

$$\begin{aligned} \mathbb{E} \left(\left(\sum_{k=1}^{N(t)} (T_k^* - T_{k-1}) + (t - T_{N(t)}) \right) 1_{N(t) < j_t} \right) \\ \leq \mathbb{E} \left(\sum_{k=1}^{j_t+1} (T_k^* - T_{k-1}) \right) = \sum_{k=1}^{j_t+1} \mathbb{E}(T_k^* - T_{k-1}) \leq (j_t + 1)m^\alpha. \end{aligned}$$

Finalement, nous combinons les inégalités précédentes et nous obtenons, pour m suffisamment grand, et pour tout $t > 0$,

$$\begin{aligned} \frac{1}{t} \mathbb{E} \left(\sum_{n \geq 1} (T_n^* \wedge t - T_{n-1} \wedge t) \right) \\ \leq \frac{2}{j_t \exp(\alpha m)} \left(\frac{j_t + 1}{2} \exp(\alpha m - j_t c(m)) + (j_t + 1)m^\alpha \right). \end{aligned}$$

Quand t tend vers $l'\infty$, cette expression tend vers $2m^\alpha / \exp(\alpha m)$, qui tend vers 0 avec m . Donc, pour m assez grand, il existe $t_m > 0$ tel que pour tout $t \geq t_m$, nous avons

$$\left| \frac{1}{t} \mathbb{E} \left(\sum_{n=0}^t f \left(\frac{Z_n(0) + \dots + Z_n(K)}{m} \right) \right) - f(\rho_0 + \dots + \rho_K) \right| < \varepsilon.$$

Et donc,

$$\lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \int_{\mathbb{D}^{K+1}} f\left(\frac{z_0 + \cdots + z_K}{m}\right) d\nu^\theta(z) = f(\rho_0 + \cdots + \rho_K).$$

□

Chapitre 20

Cas neutre

Reprenons la formule (\mathcal{J}_2) de la fin de la section 17.6. Rappelons que le temps de découverte de la master sequence, τ^* , est défini par :

$$\tau^* = \inf \{ n \geq 0 : O_n^\theta \in \mathcal{W}^* \} = \inf \{ n \geq 0 : O_n \in \mathcal{W}^* \}.$$

Le temps τ^* a été étudié en détail dans [10], qui a obtenu le résultat suivant (propositions 10.3 et 10.6 dans [10]) :

Théorème 20.0.5. *Pour tout $o \in \mathcal{P}_{\ell+1}^m \setminus \mathcal{W}^*$, nous avons :*

$$\lim_{\substack{\ell, m \rightarrow \infty \\ q \rightarrow 0, \ell q \rightarrow \alpha}} \frac{1}{\ell} \ln E(\tau^* | O_0 = o) = \ln \kappa.$$

Dans le cas $\sigma e^{-\alpha} > 1$, ce théorème et les estimations obtenues dans le chapitre précédent suffisent pour déduire la convergence :

$$\int_{\mathcal{P}_{\ell+1}^m} f\left(\frac{o(0) + \dots + o(K)}{m}\right) d\mu_O^\theta(o) \longrightarrow \rho_0 + \dots + \rho_K.$$

Pour compléter notre théorème principal, il nous resterait à montrer que dans le cas $\sigma e^{-\alpha} < 1$, nous avons la convergence

$$\int_{\mathcal{P}_{\ell+1}^m} f\left(\frac{o(0) + \dots + o(K)}{m}\right) d\mu_O^\theta(o) \longrightarrow 0.$$

Puisque la fonction $f(o(0) + \dots + o(K))/m$ ne vaut pas 0 sur l'ensemble des distributions neutres \mathcal{N} , des estimations supplémentaires sont nécessaires : il suffit de montrer que le temps passé par le processus $(O_n^\theta)_{n \geq 0}$ dans l'ensemble

$$\{o \in \mathcal{P}_{\ell+1}^m : o(0) = 0 \text{ et } o(1) + \dots + o(K) \geq 1\},$$

est négligeable par rapport au temps passé par le processus $(O_n^\theta)_{n \geq 0}$ ailleurs dans la phase neutre. Pour ce faire, nous allons nous servir de l'identité (J_1) obtenue à la fin de la section 17.1.3 :

$$\int_{\mathcal{P}_{\ell+1}^m} f\left(\frac{o(0) + \dots + o(K)}{m}\right) d\mu_O^\theta(o) = \tag{J_1} \\ \frac{1 + E_\pi(\tau_0)}{E(\tau_K^* | O_0^\theta = o_{\text{exit}}^\theta) + E_\pi(\tau_0)} \int_{\mathbb{D}^{K+1}} f\left(\frac{z_0 + \dots + z_K}{m}\right) d\nu^\theta(z).$$

Nous rappelons que les temps τ_K^* et τ_0 ci-dessus sont définis par :

$$\tau_K^* = \inf \{n \geq 0 : O_n^\theta \in \mathcal{W}_K^*\}, \\ \tau_0 = \inf \{n \geq 0 : Z_n^\theta = 0\},$$

où cette fois-ci, les chaînes de Markov $(Z_n^\theta)_{n \geq 0}$ sont celles construites par la première approche, dans la section 17.1.1. Le reste de ce chapitre est dédié à estimer le temps de découverte τ_K^* et le temps de persistance τ_0 .

20.1 Estimation du temps de découverte

Commençons par estimer $E(\tau_K^* | O_0^\theta = o_{\text{exit}}^\theta)$. Soit o une distribution d'occupation n'ayant aucun chromosome dans les classes $0, \dots, K$. Issus de l'état o , les trois processus $(O_n)_{n \geq 0}$, $(O_n^\ell)_{n \geq 0}$, $(O_n^{K+1})_{n \geq 0}$ ont la même dynamique jusqu'au temps τ_K^* . Il suffit donc d'estimer l'espérance de τ_K^* pour le processus d'occupation original $(O_n)_{n \geq 0}$. Nous allons montrer le résultat suivant :

Proposition 20.1.1. *Pour tout $o \in \mathcal{N}_K$, nous avons*

$$\lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha, \frac{m}{\ell} \rightarrow \alpha}} \frac{1}{\ell} \ln E(\tau_K^* | O_0 = o) = \ln \kappa.$$

Démonstration. Soit τ^* le temps de découverte de la master sequence, i.e.,

$$\tau^* = \inf \{ n \geq 0 : O_n(0) \geq 1 \}.$$

Évidemment, $\tau^* \geq \tau_K^*$. Grâce au théorème 20.0.5, nous déduisons la borne supérieure pour le temps de découverte τ_K^* ; pour tout $o \in \mathcal{N}_K$,

$$\limsup_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha, \frac{m}{\ell} \rightarrow \alpha}} \frac{1}{\ell} \ln E(\tau_K^* \mid O_0 = o) \leq \ln \kappa.$$

Il ne reste qu'à montrer la borne inférieure. Pour ce faire, nous allons suivre la stratégie employée dans [10] pour estimer le temps de découverte de la master sequence. Nous allons travailler avec le processus distance $(D_n)_{n \geq 0}$. Par abus de notation, nous posons

$$\begin{aligned} \mathcal{N}_K &= \{ d \in \{0, \dots, \ell\}^m : d(i) > K, 1 \leq i \leq m \}, \\ \mathcal{W}_K^* &= \{ d \in \{0, \dots, \ell\}^m : \exists i \in \{1, \dots, m\} \text{ tel que } d(i) \leq K \}. \end{aligned}$$

Soit

$$\tau_K^* = \inf \{ t \geq 0 : D_t \in \mathcal{W}_K^* \}.$$

La loi du temps de découverte τ_K^* est la même pour le processus distance et pour le processus d'occupation. Pour $b \in \{0, \dots, \ell\}$, nous notons par $(b)^m$ le vecteur colonne dont toutes les entrées sont b :

$$(b)^m = \begin{pmatrix} b \\ \vdots \\ b \end{pmatrix}.$$

Nous ne travaillons que dans le paysage à un pic. Ainsi, par le corollaire 16.2.6, le processus distance $(D_n)_{n \geq 0}$ restreint à \mathcal{N}_K est monotone, et nous avons

$$\forall d \in \mathcal{N}_K \quad E(\tau_K^* \mid D_0 = d) \geq E(\tau_K^* \mid D_0 = (K+1)^m).$$

Nous supposons que le processus distance est issu du point $(K+1)^m$ et nous allons estimer l'espérance

$$E(\tau_K^* \mid D_0 = (K+1)^m).$$

Nous commençons par donner quelques résultats extraits de [10]. Notons par \mathcal{B} la loi binomiale de paramètres ℓ et $1 - 1/\kappa$.

Lemme 20.1.2. Pour $b \leq \ell/2$, nous avons

$$\frac{1}{\kappa^\ell} \left(\frac{\ell}{2b} \right)^b \leq \mathcal{B}(b) \leq \frac{\ell^b}{\kappa^{\ell-b}}.$$

Démonstration. Soit $b \leq \ell/2$. Alors

$$\mathcal{B}(b) = \binom{\ell}{b} \left(1 - \frac{1}{\kappa}\right)^b \left(\frac{1}{\kappa}\right)^{\ell-b} \geq \binom{\ell}{b} \frac{1}{\kappa^\ell} \geq \frac{(\ell-b)^b}{b^b} \frac{1}{\kappa^\ell} \geq \left(\frac{\ell}{2b}\right)^b \frac{1}{\kappa^\ell}.$$

La majoration de $\mathcal{B}(b)$ est immédiate. □

Lemme 20.1.3. Pour $\rho \in [0, 1]$, nous avons

$$\lim_{\ell \rightarrow \infty} \frac{1}{\ell} \ln \mathcal{B}(\lfloor \rho \ell \rfloor) = -(1-\rho) \ln(\kappa(1-\rho)) - \rho \ln \frac{\kappa\rho}{\kappa-1}.$$

Démonstration. Nous écrivons

$$\begin{aligned} \ln \mathcal{B}(\lfloor \rho \ell \rfloor) &= \ln \frac{\ell \cdots (\ell - \lfloor \rho \ell \rfloor + 1)}{1 \cdots \lfloor \rho \ell \rfloor} + (\ell - \lfloor \rho \ell \rfloor) \ln \frac{1}{\kappa} + \lfloor \rho \ell \rfloor \ln \left(1 - \frac{1}{\kappa}\right) \\ &= \sum_{k=0}^{\lfloor \rho \ell \rfloor - 1} \ln \left(1 - \frac{k}{\ell}\right) - \sum_{k=1}^{\lfloor \rho \ell \rfloor} \ln \frac{k}{\ell} + (\ell - \lfloor \rho \ell \rfloor) \ln \frac{1}{\kappa} + \lfloor \rho \ell \rfloor \ln \left(1 - \frac{1}{\kappa}\right). \end{aligned}$$

Nous reconnaissons des sommes de Riemann pour les fonctions $\ln(1-x)$ et $\ln x$, donc

$$\lim_{\ell \rightarrow \infty} \frac{1}{\ell} \ln \mathcal{B}(\lfloor \rho \ell \rfloor) = \int_0^\rho \ln \frac{1-x}{x} dx + (1-\rho) \ln \frac{1}{\kappa} + \rho \ln \left(1 - \frac{1}{\kappa}\right).$$

Nous finissons la preuve en faisant l'intégration. □

Le lemme suivant est une variation du lemme 10.15 de [10]. La démonstration est la même et nous l'omettons.

Lemme 20.1.4. Pour $b \in \{K+1, \dots, \ell\}$, nous avons

$$\forall n \geq 0 \quad P(\tau_K^* \leq n \mid D_0 = (b)^m) \leq nm \frac{\mathcal{B}(0) + \dots + \mathcal{B}(K)}{\mathcal{B}(b)}.$$

Nous démontrons ensuite la borne inférieure pour le temps d'atteinte τ_K^* :

$$\liminf_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha, \frac{m}{\ell} \rightarrow \alpha}} \frac{1}{\ell} \ln \mathbb{E}(\tau_K^* \mid D_0 = (K+1)^m) \geq \ln \kappa.$$

Soit \mathcal{E} l'événement

$$\mathcal{E} = \{ \forall n \leq m\ell^{3/4} \quad \forall l \leq \ln \ell \quad U_{n,l} > p/\kappa \}.$$

Si l'événement \mathcal{E} se produit, alors jusqu'au temps $m\ell^{3/4}$, aucun événement de mutation dans le processus $(D_n)_{n \geq 0}$ ne peut créer un chromosome qui soit dans une des classes $0, \dots, K$. En effet, sur \mathcal{E} ,

$$\begin{aligned} \forall b \in \{K+1, \dots, \ell\} \quad \forall n \leq m\ell^{3/4} \\ \mathcal{M}_H(b, U_{n,1}, \dots, U_{n,\ell}) &\geq \mathcal{M}_H(K+1, U_{n,1}, \dots, U_{n,\ell}) \\ &\geq K+1 + \sum_{l=K+2}^{\ell} 1_{U_{n,l} > 1-p(1-1/\kappa)} \geq K+1. \end{aligned}$$

Donc, sur l'événement \mathcal{E} , nous avons $\tau_K^* \geq m\ell^{3/4}$. La probabilité de \mathcal{E} est

$$P(\mathcal{E}) = \left(1 - \frac{p}{\kappa}\right)^{m\ell^{3/4} \ln \ell}.$$

Soit $\varepsilon > 0$. Nous supposons que le processus est issu de $(K+1)^m$ et nous estimons la probabilité

$$\begin{aligned} P(\tau_K^* > \kappa^{\ell(1-\varepsilon)}) &\geq P(\tau_K^* > \kappa^{\ell(1-\varepsilon)}, \mathcal{E}) \\ &\geq P(\forall t \in \{m\ell^{3/4}, \dots, \kappa^{\ell(1-\varepsilon)}\} \quad D_t \in \mathcal{N}_K, \mathcal{E}) \\ &= \sum_{d \in \mathcal{N}_K} P(\forall t \in \{m\ell^{3/4}, \dots, \kappa^{\ell(1-\varepsilon)}\} \quad D_t \in \mathcal{N}_K, D_{m\ell^{3/4}} = d, \mathcal{E}) \\ &\geq \sum_{d \geq (\ln \ell)^m} P(\forall t \in \{m\ell^{3/4}, \dots, \kappa^{\ell(1-\varepsilon)}\} \quad D_t \in \mathcal{N}_K \mid D_{m\ell^{3/4}} = d, \mathcal{E}) \\ &\quad \times P(D_{m\ell^{3/4}} = d, \mathcal{E}). \end{aligned} \tag{\Sigma_1}$$

Nous utilisons la propriété de Markov, et nous obtenons

$$\begin{aligned} P(\forall t \in \{m\ell^{3/4}, \dots, \kappa^{\ell(1-\varepsilon)}\} \quad D_t \in \mathcal{N}_K \mid D_{m\ell^{3/4}} = d, \mathcal{E}) \\ = P(\forall t \in \{0, \dots, \kappa^{\ell(1-\varepsilon)} - m\ell^{3/4}\} \quad D_t \in \mathcal{N}_K \mid D_0 = d) \\ = P(\tau_K^* > \kappa^{\ell(1-\varepsilon)} - m\ell^{3/4} \mid D_0 = d) \geq P(\tau_K^* > \kappa^{\ell(1-\varepsilon)} \mid D_0 = d). \end{aligned}$$

Dans le cas neutre, le processus distance est monotone (corollaire 16.2.6). Donc, pour tout $d \geq (\ln \ell)^m$,

$$P(\tau_K^* > \kappa^{\ell(1-\varepsilon)} \mid D_0 = d) \geq P(\tau_K^* > \kappa^{\ell(1-\varepsilon)} \mid D_0 = (\ln \ell)^m).$$

Ainsi l'inégalité (Σ_1) devient

$$P(\tau_K^* > \kappa^{\ell(1-\varepsilon)}) \geq P(\tau_K^* > \kappa^{\ell(1-\varepsilon)} \mid D_0 = (\ln \ell)^m) P(D_{m\ell^{3/4}} \geq (\ln \ell)^m, \mathcal{E}). \quad (\circ)$$

Grâce à la monotonie du processus distance, nous avons

$$P(D_{m\ell^{3/4}} \geq (\ln \ell)^m, \mathcal{E} \mid D_0 = (K+1)^m) \geq P(D_{m\ell^{3/4}} \geq (\ln \ell)^m, \mathcal{E} \mid D_0 = (1)^m).$$

Nous avons aussi l'estimée suivante (section 10.5 de [10]) :

$$P(D_{m\ell^{3/4}} \geq (\ln \ell)^m, \mathcal{E} \mid D_0 = (1)^m) \geq \left(1 - m \exp\left(-\frac{1}{3}(\ln \ell)^2\right)\right) \left(1 - \frac{p}{\kappa}\right)^{m\ell^{3/4} \ln \ell}. \quad (\nabla)$$

Nous estimons ensuite

$$P(\tau_K^* > \kappa^{\ell(1-\varepsilon)} \mid D_0 = (\ln \ell)^m).$$

Soit $\varepsilon' > 0$ et soit $\ell_\kappa = \ell(1 - 1/\kappa)$. Nous avons

$$\begin{aligned} & P(\tau_K^* > \kappa^{\ell(1-\varepsilon)} \mid D_0 = (\ln \ell)^m) \\ &= P(\tau_K^* > m\ell^2, D_t \in \mathcal{N}_K \text{ pour } m\ell^2 \leq t \leq \kappa^{\ell(1-\varepsilon)} \mid D_0 = (\ln \ell)^m) \\ &= \sum_{d \in \mathcal{N}_K} P(\tau_K^* > m\ell^2, D_{m\ell^2} = d \mid D_0 = (\ln \ell)^m) \\ &\geq \sum_{d \geq (\ell_\kappa(1-\varepsilon'))^m} P(D_t \in \mathcal{N}_K \text{ pour } m\ell^2 \leq t \leq \kappa^{\ell(1-\varepsilon)} \mid \tau_K^* > m\ell^2, D_{m\ell^2} = d) \\ &\quad \times P(\tau_K^* > m\ell^2, D_{m\ell^2} = d \mid D_0 = (\ln \ell)^m). \quad (\Sigma_3) \end{aligned}$$

Par la propriété de Markov, et grâce à la monotonie du processus $(D_t)_{t \geq 0}$, nous

avons pour $d \geq (\ell_\kappa(1 - \varepsilon'))^m$,

$$\begin{aligned} & \mathbb{P}\left(D_t \in \mathcal{N}_\kappa \text{ pour } m\ell^2 \leq t \leq \kappa^{\ell(1-\varepsilon)} \mid \tau_\kappa^* > m\ell^2, D_{m\ell^2} = d\right) \\ &= \mathbb{P}\left(\forall t \in \{0, \dots, \kappa^{\ell(1-\varepsilon)} - m\ell^2\} \quad D_t \in \mathcal{N}_\kappa \mid D_0 = d\right) \\ &= \mathbb{P}\left(\tau_\kappa^* > \kappa^{\ell(1-\varepsilon)} - m\ell^2 \mid D_0 = d\right) \geq \mathbb{P}\left(\tau_\kappa^* > \kappa^{\ell(1-\varepsilon)} \mid D_0 = d\right) \\ &\geq \mathbb{P}\left(\tau_\kappa^* > \kappa^{\ell(1-\varepsilon)} \mid D_0 = (\ell_\kappa(1 - \varepsilon'))^m\right). \end{aligned}$$

La formule (Σ_3) devient donc

$$\begin{aligned} & \mathbb{P}\left(\tau_\kappa^* > \kappa^{\ell(1-\varepsilon)} \mid D_0 = (\ln \ell)^m\right) \geq \mathbb{P}\left(\tau_\kappa^* > \kappa^{\ell(1-\varepsilon)} \mid D_0 = (\ell_\kappa(1 - \varepsilon'))^m\right) \\ &\quad \times \mathbb{P}\left(\tau_\kappa^* > m\ell^2, D_{m\ell^2} \geq (\ell_\kappa(1 - \varepsilon'))^m \mid D_0 = (\ln \ell)^m\right). \quad (\heartsuit) \end{aligned}$$

Estimation de $\mathbb{P}(\tau_\kappa^* > m\ell^2, D_{m\ell^2} \geq (\ell_\kappa(1 - \varepsilon'))^m \mid D_0 = (\ln \ell)^m)$. Nous écrivons

$$\begin{aligned} & \mathbb{P}\left(\tau_\kappa^* > m\ell^2, D_{m\ell^2} \geq (\ell_\kappa(1 - \varepsilon'))^m \mid D_0 = (\ln \ell)^m\right) \geq \quad (\natural) \\ & \mathbb{P}\left(D_{m\ell^2} \geq (\ell_\kappa(1 - \varepsilon'))^m \mid D_0 = (\ln \ell)^m\right) - \mathbb{P}\left(\tau_\kappa^* \leq m\ell^2 \mid D_0 = (\ln \ell)^m\right). \end{aligned}$$

Pour contrôler le dernier terme, nous utilisons l'inégalité du lemme 20.1.4 avec $n = m\ell^2$ et $b = \ln \ell$:

$$\mathbb{P}\left(\tau_\kappa^* \leq m\ell^2 \mid D_0 = (\ln \ell)^m\right) \leq (m\ell)^2 \frac{\mathcal{B}(0) + \dots + \mathcal{B}(K)}{\mathcal{B}(\ln \ell)}.$$

Par le lemme 20.1.2, nous avons

$$\frac{\mathcal{B}(0) + \dots + \mathcal{B}(K)}{\mathcal{B}(\ln \ell)} \leq \frac{1 - (\ell\kappa)^{K+1}}{1 - \ell\kappa} \left(\frac{2 \ln \ell}{\ell}\right)^{\ln \ell},$$

d'où

$$\mathbb{P}\left(\tau_\kappa^* \leq m\ell^2 \mid D_0 = (\ln \ell)^m\right) \leq (m\ell)^2 \frac{1 - (\ell\kappa)^{K+1}}{1 - \ell\kappa} \left(\frac{2 \ln \ell}{\ell}\right)^{\ln \ell}. \quad (\flat)$$

Pour l'autre terme nous utilisons une estimée de la section 10.5 de [10] : il existe une constante $c(\varepsilon') > 0$ telle que pour ℓ suffisamment grand, nous avons

$$\mathbb{P}\left(D_{m\ell^2} \geq (\ell_\kappa(1 - \varepsilon'))^m \mid D_0 = (\ln \ell)^m\right) \geq 1 - m \exp\left(-\frac{1}{2}c(\varepsilon')\ell\right).$$

Cette inégalité et les estimées (a) et (b) entraînent que

$$\begin{aligned} \mathbb{P}\left(\tau_K^* > m\ell^2, D_{m\ell^2} \geq (\ell_\kappa(1-\varepsilon'))^m \mid D_0 = (\ln \ell)^m\right) &\geq \quad (\clubsuit) \\ &1 - m \exp\left(-\frac{1}{2}c(\varepsilon')\ell\right) - (m\ell)^2 \frac{1 - (\ell_\kappa)^{K+1}}{1 - \ell_\kappa} \left(\frac{2 \ln \ell}{\ell}\right)^{\ln \ell}. \end{aligned}$$

Estimation de $\mathbb{P}(\tau_K^* > \kappa^{\ell(1-\varepsilon)} \mid D_0 = (\ell_\kappa(1-\varepsilon'))^m)$. Nous utilisons d'abord l'inégalité du lemme 20.1.4 avec $n = \kappa^{\ell(1-\varepsilon)}$ et $b = \ell_\kappa(1-\varepsilon')$:

$$\mathbb{P}\left(\tau_K^* \leq \kappa^{\ell(1-\varepsilon)} \mid D_0 = (\ell_\kappa(1-\varepsilon'))^m\right) \leq \kappa^{\ell(1-\varepsilon)} m \frac{\mathcal{B}(0) + \dots + \mathcal{B}(K)}{\mathcal{B}(\ell_\kappa(1-\varepsilon'))}.$$

Pour ε' assez petit, par les estimées de grandes déviations du lemme 20.1.3, nous voyons qu'il existe $c(\varepsilon, \varepsilon') > 0$ tel que, pour ℓ suffisamment grand,

$$\mathbb{P}\left(\tau_K^* \leq \kappa^{\ell(1-\varepsilon)} \mid D_0 = (\ell_\kappa(1-\varepsilon'))^m\right) \leq \exp(-c(\varepsilon, \varepsilon')\ell). \quad (\spadesuit)$$

Nous injectons les inégalités (\clubsuit) et (\spadesuit) dans (\heartsuit) et nous obtenons

$$\begin{aligned} \mathbb{P}\left(\tau_K^* > \kappa^{\ell(1-\varepsilon)} \mid D_0 = (\ln \ell)^m\right) &\geq \left(1 - \exp(-c(\varepsilon, \varepsilon')\ell)\right) \quad (\triangle) \\ &\left(1 - m \exp\left(-\frac{1}{2}c(\varepsilon')\ell\right) - (m\ell)^2 \frac{1 - (\ell_\kappa)^{K+1}}{1 - \ell_\kappa} \left(\frac{2 \ln \ell}{\ell}\right)^{\ln \ell}\right). \end{aligned}$$

À l'aide des inégalités (\circ), (∇), (\triangle), nous concluons que, pour ℓ suffisamment grand,

$$\begin{aligned} \mathbb{P}\left(\tau_K^* > \kappa^{\ell(1-\varepsilon)} \mid D_0 = (K+1)^m\right) &\geq \\ &\left(1 - m \exp\left(-\frac{1}{3}(\ln \ell)^2\right)\right) \left(1 - \frac{p}{\kappa}\right)^{m\ell^{3/4} \ln \ell} \left(1 - \exp(-c(\varepsilon, \varepsilon')\ell)\right) \\ &\times \left(1 - m \exp\left(-\frac{1}{2}c(\varepsilon')\ell\right) - (m\ell)^2 \frac{1 - (\ell_\kappa)^{K+1}}{1 - \ell_\kappa} \left(\frac{2 \ln \ell}{\ell}\right)^{\ln \ell}\right). \end{aligned}$$

De plus, grâce à l'inégalité de Markov,

$$\mathbb{E}\left(\tau_K^* \mid D_0 = (K+1)^m\right) \geq \kappa^{\ell(1-\varepsilon)} \mathbb{P}\left(\tau_K^* \geq \kappa^{\ell(1-\varepsilon)} \mid D_0 = (K+1)^m\right).$$

Nous en déduisons que

$$\liminf_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \frac{1}{\ell} \ln \mathbb{E}\left(\tau^* \mid D_0 = (K+1)^m\right) \geq (1-\varepsilon) \ln \kappa.$$

Nous envoyons ε vers 0 et nous obtenons la borne inférieure voulue. \square

20.2 Estimation du temps de persistance

Nous allons maintenant estimer l'espérance $E_\pi(\tau_0)$. Rappelons que le temps τ_0 est défini par :

$$\tau_0 = \inf \{ n \geq 0 : Z_n^\theta = 0 \},$$

où $(Z_n^\theta)_{n \geq 0}$ sont les chaînes de naissance et mort construites par notre première approche dans la section 17.1.3. Si le processus $(O_n^\theta)_{n \geq 0}$ est issu de la distribution d'occupation o_{exit}^θ , et le processus $(Z_n^\theta)_{n \geq 0}$ a pour distribution initiale la distribution π , alors le processus $((O_t^\theta(0), \dots, O_t^\theta(K)), \tau_K^* \leq t \leq \tau_K)$ et $(Z_t^\theta, 0 \leq t \leq \tau_0)$ ont même loi. En particulier, nous avons

$$E_\pi(\tau_0) = E(\tau_K - \tau_K^* | O_0^\theta = o_{\text{exit}}^\theta).$$

Pour tout sous-ensemble E de $\mathcal{P}_{\ell+1}^m$, nous définissons le temps $\tau(E)$ d'entrée dans l'ensemble E :

$$\tau(E) = \inf \{ t \geq 0 : O_t^\theta \in E \}.$$

La chaîne de Markov $(O_n^\theta)_{n \geq 0}$ entre toujours dans l'ensemble \mathcal{W}^* par un des états $(o_i^\theta, 0 \leq i \leq K)$. Grâce à la monotonie du processus $(O_n^\theta)_{n \geq 0}$, nous avons

$$E(\tau(\mathcal{N}_K) | O_0^\theta = o_K^\theta) \leq E(\tau_K - \tau_K^* | O_0^\theta = o_{\text{exit}}^\theta) \leq E(\tau(\mathcal{N}_K) | O_0^\theta = o_0^\theta).$$

Soit $\psi :]0, +\infty[\rightarrow [0, +\infty[$ la fonction définie par $\psi(a) = 0$ si $a \geq \ln \sigma$ et

$$\forall a < \ln \sigma \quad \psi(a) = \frac{\sigma(1 - e^{-a}) \ln \frac{\sigma(1 - e^{-a})}{\sigma - 1} + \ln(\sigma e^{-a})}{(1 - \sigma(1 - e^{-a}))}.$$

Proposition 20.2.1. *Pour tout $i \in \{0, \dots, K\}$, nous avons*

$$\lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha, \frac{m}{\ell} \rightarrow \alpha}} \frac{1}{m} \ln E(\tau(\mathcal{N}_K) | O_0^\theta = o_i^\theta) = \psi(\alpha).$$

Démonstration. Soit $o' = (1, m-1, 0, \dots, 0)$. Par la proposition 17.1.1, nous avons pour tout $i \in \{0, \dots, K\}$

$$E(\tau(\mathcal{N}_K) | O_0^\ell = o_K^\ell) \leq E(\tau(\mathcal{N}_K) | O_0^\theta = o_i^\theta) \leq E(\tau(\mathcal{N}_K) | O_0^{K+1} = o').$$

Nous allons montrer d'abord la borne inférieure :

$$\liminf_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha, \frac{m}{\ell} \rightarrow \alpha}} \frac{1}{m} \ln E(\tau(\mathcal{N}_K) | O_0^\ell = o_K^\ell) \geq \psi(\alpha).$$

En effet nous avons

$$\begin{aligned} \mathbb{E}(\tau(\mathcal{N}_K) | O_0^\ell = o_K^\ell) &\geq \\ &\mathbb{E}(\tau(\mathcal{N}_K) | O_0^\ell = o_K^\ell, O_1^\ell = o_0^\ell) \mathbb{P}(O_1^\ell = o_0^\ell | O_0^\ell = o_K^\ell). \end{aligned}$$

D'une part, nous avons grâce à la propriété de Markov et à la monotonie du processus

$$\mathbb{E}(\tau(\mathcal{N}_K) | O_0^\ell = o_K^\ell, O_1^\ell = o_0^\ell) \geq 1 + \mathbb{E}(\tau(\mathcal{N}_K) | O_0^\ell = o_0^\ell).$$

Puisque $\tau(\mathcal{N}_K) \geq \tau(\mathcal{N})$, nous avons grâce à l'estimée sur le temps $\tau(\mathcal{N})$ (corollaire 9.2 de [10]),

$$\begin{aligned} \liminf_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \frac{1}{m} \ln \mathbb{E}(\tau(\mathcal{N}_K) | O_0^\ell = o_0^\ell) &\geq \\ \liminf_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \frac{1}{m} \ln \mathbb{E}(\tau(\mathcal{N}) | O_0^\ell = o_0^\ell) &\geq \psi(a). \end{aligned}$$

D'autre part, nous avons

$$\begin{aligned} \mathbb{P}(O_1^\ell = o_0^\ell | O_0^\ell = o_K^\ell) &= \frac{1}{m^2} (M_H(K, 0) + (m-1)M_H(\ell, 0)) \\ &\geq \frac{1}{m^2} M_H(K, 0) = \frac{1}{m^2} \left(1 - p \left(1 - \frac{1}{\kappa}\right)\right)^{\ell-K} \left(\frac{p}{\kappa}\right)^K. \end{aligned}$$

Ceci entraîne que

$$\liminf_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \frac{1}{m} \ln \mathbb{P}(O_1^\ell = o_0^\ell | O_0^\ell = o_K^\ell) \geq 0.$$

Nous combinons les inégalités précédentes et nous obtenons le résultat voulu :

$$\liminf_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \frac{1}{m} \ln \mathbb{E}(\tau(\mathcal{N}_K) | O_0^\ell = o_K^\ell) \geq \psi(a).$$

Nous démontrons ensuite la borne supérieure

$$\limsup_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow a, \frac{m}{\ell} \rightarrow \alpha}} \frac{1}{m} \ln \mathbb{E}(\tau(\mathcal{N}_K) | O_0^{K+1} = o') \leq \psi(a).$$

Pour alléger la notation, nous désignons par P_o et E_o les probabilités et espérances pour le processus $(O_n^{K+1})_{n \geq 0}$ issu de la distribution d'occupation o . Soit o'' la distribution donnée par $o'' = (0, m, 0, \dots, 0)$. Puisque $\tau(\mathcal{N}_K) \geq \tau(\mathcal{N})$, nous avons

$$E_{o'}(\tau(\mathcal{N}_K)) = E_{o'}(\tau(\mathcal{N})) + E_{o'}(\tau(\mathcal{N}_K) - \tau(\mathcal{N})).$$

Or, grâce à la propriété de Markov forte et à la monotonie du processus

$$\begin{aligned} E_{o'}(\tau(\mathcal{N}_K) - \tau(\mathcal{N})) &= \sum_{o \in \mathcal{N}} E_o(\tau(\mathcal{N}_K)) P_{o'}(\tau(\mathcal{N}) = o) \\ &\leq \sum_{o \in \mathcal{N}} E_{o''}(\tau(\mathcal{N}_K)) P_{o'}(\tau(\mathcal{N}) = o) = E_{o''}(\tau(\mathcal{N}_K)). \end{aligned}$$

Nous développons cette espérance :

$$E_{o''}(\tau(\mathcal{N}_K)) = E_{o''}(\tau(\mathcal{W}^* \cup \mathcal{N}_K)) + E_{o''}(\tau(\mathcal{N}_K) - \tau(\mathcal{W}^* \cup \mathcal{N}_K)).$$

Or,

$$\begin{aligned} E_{o''}(\tau(\mathcal{N}_K) - \tau(\mathcal{W}^* \cup \mathcal{N}_K)) &= \sum_{o \in \mathcal{W}^*} E_o(\tau(\mathcal{N}_K)) P_{o''}(\tau(\mathcal{W}^* \cup \mathcal{N}_K) = o) \\ &\leq \sum_{o \in \mathcal{W}^*} E_{o'}(\tau(\mathcal{N}_K)) P_{o''}(\tau(\mathcal{W}^* \cup \mathcal{N}_K) = o) \\ &= E_{o'}(\tau(\mathcal{N}_K)) P_{o''}(\tau(\mathcal{W}^*) < \tau(\mathcal{N}_K)). \end{aligned}$$

Donc,

$$E_{o''}(\tau(\mathcal{N}_K)) \leq E_{o''}(\tau(\mathcal{W}^* \cup \mathcal{N}_K)) + E_{o'}(\tau(\mathcal{N}_K)) P_{o''}(\tau(\mathcal{W}^*) < \tau(\mathcal{N}_K)).$$

Nous en déduisons que

$$E_{o'}(\tau(\mathcal{N}_K)) \leq \frac{1}{P_{o''}(\tau(\mathcal{N}_K) < \tau(\mathcal{W}^*))} \left(E_{o''}(\tau(\mathcal{W}^* \cup \mathcal{N}_K)) + E_{o'}(\tau(\mathcal{N}_K)) P_{o''}(\tau(\mathcal{W}^*) < \tau(\mathcal{N}_K)) \right).$$

Ensuite nous estimons les trois termes du membre de droite de cette formule. Pour la première espérance nous utilisons l'estimée pour le temps de persistance de la master sequence (corollaire 9.2 de [10]) et nous obtenons :

$$\limsup_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha, \frac{m}{\ell} \rightarrow \alpha}} \frac{1}{m} \ln E_{o'}(\tau(\mathcal{N})) \leq \psi(\alpha).$$

Pour borner inférieurement la probabilité du dénominateur, nous utilisons les estimées du temps de découverte τ_K^* . Sur l'événement

$$\mathcal{E} = \left\{ \forall n \leq m\ell^{3/4} \quad \forall l \leq \ln \ell \quad U_{n,l} > p/\kappa \right\},$$

si $D_{m\ell^{3/4}} \geq (\ln \ell)^m$, nous avons nécessairement $\tau(\mathcal{N}_K) < \tau(\mathcal{W}^*)$. D'où

$$\begin{aligned} P_{o''}(\tau(\mathcal{N}_K) < \tau(\mathcal{W}^*)) &\geq P(D_{m\ell^{3/4}} \geq (\ln \ell)^m, \mathcal{E} \mid D_0 = (1)^m) \geq \\ &\left(1 - m \exp\left(-\frac{1}{3}(\ln \ell)^2\right)\right) \left(1 - \frac{p}{\kappa}\right)^{m\ell^{3/4} \ln \ell}. \end{aligned}$$

Il ne reste qu'à estimer l'espérance $E_{o''}(\tau(\mathcal{W}^* \cup \mathcal{N}_K))$. Estimons d'abord pour $n \geq 0$ la probabilité $P_{o''}(\tau(\mathcal{W}^* \cup \mathcal{N}_K) > n)$. Nous avons,

$$\begin{aligned} P_{o''}(\tau(\mathcal{W}^* \cup \mathcal{N}_K) > n) &= \\ &\sum_{o \in \mathcal{W}_K^* \setminus \mathcal{W}^*} P_{o''}(\tau(\mathcal{W}^* \cup \mathcal{N}_K) > n, O_{n-1}^{K+1} = o, \tau(\mathcal{W}^* \cup \mathcal{N}_K) > n-1) \\ &= \sum_{o \in \mathcal{W}_K^* \setminus \mathcal{W}^*} P_{o''}(\tau(\mathcal{W}^* \cup \mathcal{N}_K) > n \mid O_{n-1}^{K+1} = o, \tau(\mathcal{W}^* \cup \mathcal{N}_K) > n-1) \\ &\quad \times P_{o''}(O_{n-1}^{K+1} = o, \tau(\mathcal{W}^* \cup \mathcal{N}_K) > n-1). \end{aligned}$$

Grâce à la propriété de Markov et à la monotonie du processus, nous avons pour tout $o \in \mathcal{W}_K^* \setminus \mathcal{W}^*$,

$$\begin{aligned} P_{o''}(\tau(\mathcal{W}^* \cup \mathcal{N}_K) > n \mid O_{n-1}^{K+1} = o, \tau(\mathcal{W}^* \cup \mathcal{N}_K) > n-1) &= \\ P_o(\tau(\mathcal{W}^* \cup \mathcal{N}_K) > 1) &= 1 - P_o(\tau(\mathcal{W}^* \cup \mathcal{N}_K) = 1) \\ \leq 1 - P_o(\tau(\mathcal{W}^*) = 1) &\leq 1 - P(\tau(\mathcal{W}^*) = 1 \mid O_0^{K+1} = o_K^{K+1}) \\ &\leq 1 - \frac{M_H(K, 0)}{m^2}. \end{aligned}$$

D'où

$$P_{o''}(\tau(\mathcal{W}^* \cup \mathcal{N}_K) > n) \leq \left(1 - \frac{M_H(K, 0)}{m^2}\right) P_{o''}(\tau(\mathcal{W}^* \cup \mathcal{N}_K) > n-1).$$

Nous itérons cette inégalité et nous obtenons

$$P_{o''}(\tau(\mathcal{W}^* \cup \mathcal{N}_K) > n) \leq \left(1 - \frac{M_H(K, 0)}{m^2}\right)^n.$$

Enfin,

$$\begin{aligned} E_{o''}(\tau(\mathcal{W}^* \cup \mathcal{N}_K)) &= \sum_{n \geq 0} P_{o''}(\tau(\mathcal{W}^* \cup \mathcal{N}_K) > n) \\ &\leq \sum_{n \geq 0} \left(1 - \frac{M_H(K, 0)}{m^2}\right)^n = \frac{m^2}{M_H(K, 0)} = \frac{m^2}{\left(1 - p\left(1 - \frac{1}{K}\right)\right)^{\ell-K} \left(\frac{p}{K}\right)^K}. \end{aligned}$$

Nous combinons les inégalités précédentes et nous obtenons le résultat voulu. \square

Chapitre 21

Fin de la démonstration du résultat principal

Nous utilisons ici les estimations développées dans les trois chapitres précédents, pour finir la démonstration du théorème 15.3.1 :

Soit $\psi :]0, +\infty[\rightarrow [0, +\infty[$ la fonction définie par $\psi(a) = 0$ si $a \geq \ln \sigma$ et

$$\forall a < \ln \sigma \quad \psi(a) = \frac{\sigma(1 - e^{-a}) \ln \frac{\sigma(1 - e^{-a})}{\sigma - 1} + \ln(\sigma e^{-a})}{(1 - \sigma(1 - e^{-a}))}.$$

Théorème 21.0.2. *Nous supposons que*

$$\ell \rightarrow +\infty, \quad m \rightarrow +\infty, \quad q \rightarrow 0,$$

et que

$$\ell q \rightarrow \alpha \in]0, +\infty[, \quad \frac{m}{\ell} \rightarrow \alpha \in [0, +\infty].$$

Nous avons la dichotomie suivante :

- Si $\alpha \psi(a) < \ln \kappa$, alors pour tout $k \geq 0$, la mesure ν_k^θ converge vers la masse de Dirac au point 0 :

$$\forall \varepsilon > 0 \quad \nu_k^\theta([0, \varepsilon]) \rightarrow 1.$$

- Si $\alpha \psi(a) > \ln \kappa$, alors pour tout $k \geq 0$, la mesure ν_k^θ converge vers la masse de Dirac au point ρ_k^* :

$$\forall \varepsilon > 0 \quad \nu_k^\theta([\rho_k^* - \varepsilon, \rho_k^* + \varepsilon]) \rightarrow 1.$$

Nous traitons d'abord le cas sous-critique et finalement le cas sur-critique.

21.1 Le cas sous-critique

Nous traitons le cas sous-critique à l'aide de l'identité (J_1) obtenue à la fin de la section 17.1.3 :

$$\int_{\mathcal{P}_{\ell+1}^m} f\left(\frac{o(0) + \dots + o(K)}{m}\right) d\mu_O^\theta(o) = \tag{J_1}$$

$$\frac{1 + E_\pi(\tau_0)}{E(\tau_K^* | O_0^\theta = o_{\text{exit}}^\theta) + E_\pi(\tau_0)} \int_{\mathbb{D}^{K+1}} f\left(\frac{z_0 + \dots + z_K}{m}\right) d\nu^\theta(z).$$

D'après les asymptotiques obtenues dans le chapitre 20 (proposition 20.1.1 et proposition 20.2.1), nous concluons que pour $\alpha \in [0, +\infty[$ ou $\alpha = +\infty$,

$$\lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha, \frac{m}{\ell} \rightarrow \alpha}} \frac{E_\pi(\tau_0)}{E(\tau_K^* | O_0^\theta = o_{\text{exit}}^\theta)} = \begin{cases} 0 & \text{if } \alpha \psi(\alpha) < \ln \kappa \\ +\infty & \text{if } \alpha \psi(\alpha) > \ln \kappa \end{cases}$$

En particulier, si $\alpha \phi(\alpha) < \ln \kappa$, nous avons

$$\lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha, \frac{m}{\ell} \rightarrow \alpha}} \int_{\mathcal{P}_{\ell+1}^m} f\left(\frac{o(0) + \dots + o(K)}{m}\right) d\mu_O^\theta(o) = 0.$$

21.2 Le cas sur-critique

Nous traitons le cas sous-critique à l'aide de l'identité (J_2) obtenue à la fin de la section 17.6 :

$$\int_{\mathcal{P}_{\ell+1}^m} f\left(\frac{o(0) + \dots + o(K)}{m}\right) d\mu_O^\theta(o) = \tag{J_2}$$

$$\frac{E\left(\sum_{n=0}^{\tau^*-1} f\left(\frac{O_n^\theta(0) + \dots + O_n^\theta(K)}{m}\right) \middle| O_0^\theta = o_{\text{exit}}^\theta\right)}{E(\tau^* | O_0^\theta = o_{\text{exit}}^\theta) + E(\tau_0 | Z_0^\theta = z^\theta)}$$

$$+ \frac{1 + E(\tau_0 | Z_0^\theta = z^\theta)}{E(\tau^* | O_0^\theta = o_{\text{exit}}^\theta) + E(\tau_0 | Z_0^\theta = z^\theta)} \int_{\mathbb{D}^{K+1}} f\left(\frac{z_0 + \dots + z_K}{m}\right) d\nu^\theta(z).$$

Les temps τ^* et τ_0 coïncident avec le temps de découverte de la master sequence et le temps de persistance de la master sequence de [10]. En particulier, nous avons les estimées suivantes :

$$\lim_{\substack{\ell, m \rightarrow \infty \\ q \rightarrow 0, \ell q \rightarrow \alpha}} \frac{1}{\ell} \ln E(\tau^* | O_0^\theta = o_{\text{exit}}^\theta) = \ln \kappa,$$

$$\lim_{\substack{\ell, m \rightarrow \infty \\ q \rightarrow 0, \ell q \rightarrow \alpha}} \frac{1}{m} \ln E(\tau_0 | Z_0^\theta = z^\theta) = \psi(\alpha),$$

Nous en déduisons que

$$\lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha, \frac{m}{\ell} \rightarrow \alpha}} \frac{E(\tau_0 | Z_0^\theta = z^\theta)}{E(\tau^* | O_0^\theta = o_{\text{exit}}^\theta)} = \begin{cases} 0 & \text{si } \alpha \psi(\alpha) < \ln \kappa, \\ +\infty & \text{si } \alpha \psi(\alpha) > \ln \kappa. \end{cases}$$

Ceci achève la démonstration du théorème principal 15.3.1.

Chapitre 22

Le principe de grandes déviations

Le résultat du chapitre précédent ne concerne que le paysage à un pic. Nous voudrions maintenant étudier le modèle de Moran pour des paysages de fitness plus généraux. Une des propriétés clé dans l'étude du modèle de Moran avec le paysage à un pic est la monotonie du processus. Or, cette monotonie est perdue lorsque l'on considère un paysage de fitness plus général. Soient $\sigma > \eta > 1$ et considérons par exemple la fonction de fitness donnée par :

$$\forall \mathbf{u} \in \mathcal{A}^\ell \quad f(\mathbf{u}) = \begin{cases} \sigma & \text{si } \mathbf{u} = \mathbf{w}^*, \\ \eta & \text{si } d(\mathbf{u}, \mathbf{w}^*) = 1, \\ 1 & \text{sinon.} \end{cases}$$

Soient $\mathbf{o}, \mathbf{o}' \in \mathcal{P}_{\ell+1}^m$ les distributions d'occupation suivantes :

$$\mathbf{o} = (1, m-1, 0, \dots, 0), \quad \mathbf{o}' = (1, 0, \dots, 0, m-1).$$

Nous avons $\mathbf{o} \succeq \mathbf{o}'$, cependant, en faisant un pas du processus d'occupation, la première population a plus de chances d'éliminer la master sequence que la deuxième. Il n'est donc pas possible de construire un couplage pour garantir la monotonie du processus d'occupation. Nous sommes donc obligés de développer des méthodes plus robustes afin d'étudier des fonctions de fitness plus générales. Au vu de nos résultats pour le modèle d'Eigen dans la partie I, nous cherchons maintenant à étudier le modèle de Moran pour des fonctions de fitness dépendantes de la classe de Hamming, et constantes à partir d'un certain rang. Nous supposons tout au long de ce chapitre que la fonction de fitness $f : \mathcal{A}^\ell \rightarrow \mathbb{R}^+$ est une fonction strictement positive vérifiant :

- Il existe une séquence, notée w^* et appelée master sequence, de fitness maximale, i.e.,

$$f(w^*) > \max \{ f(u) : u \in \mathcal{A}^\ell \setminus \{w^*\} \}.$$

- La fonction de fitness f respecte les classes de Hamming vis-à-vis de la master sequence, i.e., pour tout $u, v \in \mathcal{A}^\ell$,

$$d_H(u, w^*) = d_H(v, w^*) \implies f(u) = f(v).$$

- Il existe $K \geq 0$ tel que, pour des génotypes dans des classes de Hamming supérieures à K , la fitness vaut 1, i.e., pour tout $u \in \mathcal{A}^\ell$,

$$d_H(u, w^*) > K \implies f(u) = 1.$$

Dans le chapitre précédent, nous avons travaillé presque exclusivement avec le processus de Moran normalisé, car le processus de Moran normalisé était monotone, alors que le processus original ne l'était pas. Or, dans ce chapitre, nous n'avons plus d'espoir d'avoir un processus monotone, nous choisissons donc de travailler directement avec le processus de Moran original, ce qui est, à priori, un choix plus naturel. Tout au long de ce chapitre, $(X_n)_{n \geq 0}$ représente le processus de Moran original, et $(D_n)_{n \geq 0}$, $(O_n)_{n \geq 0}$ sont les processus de distance et d'occupation dérivés du processus original $(X_n)_{n \geq 0}$.

22.1 Principe de grandes déviations

L'objectif de cette section est de montrer un principe de grandes déviations trajectorien pour le processus d'occupation $(O_n)_{n \geq 0}$ associé au processus de Moran non normalisé. Le processus d'occupation $(O_n)_{n \geq 0}$ est une chaîne de Markov à valeurs dans $\mathcal{P}_{\ell+1}^m$ et de matrice de transition p_O donnée par :

$$\forall o \in \mathcal{P}_{\ell+1}^m \quad \forall i, j \in \{0, \dots, \ell\} \quad i \neq j$$

$$p_O(o, o - e_i + e_j) = \frac{1}{\lambda m} o(i) \sum_{h=0}^{\ell} o(h) f_H(h) M_H(h, j).$$

Soit $\tau = 1/m$. Soit $k \geq K$ et définissons les processus \mathcal{X}, \mathcal{Y} par :

$$\begin{aligned} \mathcal{X}(t) &= \frac{1}{m} O_{\lfloor t/\tau \rfloor}, \quad t \geq 0, \\ \mathcal{Y}(t) &= (\mathcal{X}_0(t), \dots, \mathcal{X}_k(t)), \quad t \geq 0. \end{aligned}$$

Les espaces d'états des processus \mathcal{X} et Y sont respectivement :

$$\begin{aligned}\Sigma^{\ell+1} &= \{x \in \{0, 1/m, \dots, 1\}^{\ell+1} : x_0 + \dots + x_\ell = 1\}, \\ \Delta^{k+1} &= \{x \in \{0, 1/m, \dots, 1\}^{k+1} : x_0 + \dots + x_k \leq 1\}.\end{aligned}$$

Puisque notre objectif est de faire tendre m vers l'infini, nous posons :

$$\begin{aligned}\mathcal{S}^{\ell+1} &= \{x \in [0, 1]^{\ell+1} : x_0 + \dots + x_\ell = 1\}, \\ \mathcal{D}^{k+1} &= \{x \in [0, 1]^{k+1} : x_0 + \dots + x_k \leq 1\}.\end{aligned}$$

Remarquons que $\Sigma^{\ell+1}$ et Δ^{k+1} sont des sous-ensembles de $\mathcal{S}^{\ell+1}$ et \mathcal{D}^{k+1} . Soit $\pi_k : \mathcal{S}^{\ell+1} \rightarrow \mathcal{D}^{k+1}$ l'application de projection définie par :

$$\forall x \in \mathcal{S}^{\ell+1} \quad \pi_k(x) = (x_0, \dots, x_k).$$

La dynamique du processus Y n'est pas markovienne, car pour déterminer les sauts du processus Y nous avons besoin de connaître des coordonnées du processus \mathcal{X} que nous n'avons pas incluses dans la définition de Y . Notons par $(e_i)_{0 \leq i \leq k}$ la base canonique de \mathbb{R}^{k+1} indexée par $0, \dots, k$ au lieu de $1, \dots, k+1$. Soit \mathcal{V} l'ensemble des vecteurs de \mathbb{R}^{k+1} associés aux possibles directions des sauts du processus Y :

$$\mathcal{V} = \{-e_i : 0 \leq i \leq k\} \cup \{e_j : 0 \leq j \leq k\} \cup \{-e_i + e_j : 0 \leq i, j \leq k, i \neq j\}.$$

Nous définissons les quantités suivantes : pour tout $x \in \mathcal{S}^{\ell+1}$ et $0 \leq i, j \leq k$,

$$\begin{aligned}C_x(-e_i) &= \frac{1}{\lambda} x_i \sum_{0 \leq h \leq \ell} x_h f_H(h) \left(1 - \sum_{0 \leq l \leq k} M_H(h, l)\right), \\ C_x(e_j) &= \frac{1}{\lambda} \left(1 - \sum_{0 \leq l \leq k} x_l\right) \sum_{0 \leq h \leq \ell} x_h f_H(h) M_H(h, j), \\ C_x(-e_i + e_j) &= \frac{1}{\lambda} x_i \sum_{0 \leq h \leq \ell} x_h f_H(h) M_H(h, j).\end{aligned}$$

Les processus \mathcal{X} et Y sont des processus càdlàg qui sautent dans des instants multiples de τ . Les quantités que nous venons de définir correspondent aux probabilités de transition du processus Y sachant le processus \mathcal{X} , c'est-à-dire : pour tout $x \in \Sigma^{\ell+1}$ et $v \in \mathcal{V}$,

$$P(Y(\tau) = \pi_k(x) + v/m \mid \mathcal{X}_0 = x) = C_x(v),$$

Rappelons que quand

$$\ell \rightarrow +\infty \quad q \rightarrow 0 \quad \ell q \rightarrow a \in]0, +\infty[,$$

les probabilités de mutation $M_H(i, j)$ tendent vers :

$$\forall i, j \geq 0 \quad M_H(i, j) \rightarrow M_\infty(i, j) = \begin{cases} e^{-a} \frac{a^{j-i}}{(j-i)!} & \text{si } i \leq j, \\ 0 & \text{si } i > j. \end{cases}$$

Puisque la fonction f_H est constante et égale à 1 à partir de K , et que nous avons pris $k \geq K$, les quantités C_x convergent vers les quantités $C_{\pi_k(x)}^\infty$ données par : pour tout $y \in \mathcal{D}^{k+1}$ et $0 \leq i, j \leq k$,

$$\begin{aligned} C_y^\infty(-e_i) &= \frac{1}{\lambda} y_i \left(\sum_{0 \leq h \leq k} y_h f_H(h) \left(1 - \sum_{h \leq l \leq k} e^{-a} \frac{a^{l-h}}{(l-h)!} \right) + 1 - \sum_{0 \leq h \leq k} y_h \right), \\ C_y^\infty(e_j) &= \frac{1}{\lambda} \left(1 - \sum_{0 \leq l \leq k} y_l \right) \sum_{0 \leq h \leq j} y_h f_H(h) e^{-a} \frac{a^{j-h}}{(j-h)!}, \\ C_y^\infty(-e_i + e_j) &= \frac{1}{\lambda} y_i \sum_{0 \leq h \leq j} y_h f_H(h) e^{-a} \frac{a^{j-h}}{(j-h)!}. \end{aligned}$$

Le processus Y peut donc être considéré comme "asymptotiquement markovien". Soit $G^\infty : \mathcal{D}^{k+1} \times \mathbb{R}^{k+1} \rightarrow \mathbb{R}$ la fonction définie par :

$$G^\infty(y; z) = \ln \left(1 + \sum_{v \in \mathcal{V}} C_y^\infty(v) (e^{vz} - 1) \right),$$

et soit $H^\infty : \mathcal{D}^{k+1} \times \mathbb{R}^{k+1} \rightarrow \mathbb{R}^+$ la transformée de Fenchel-Legendre de la fonction $G^\infty(y, \cdot)$:

$$H^\infty(y; u) = \sup_{z \in \mathbb{R}^{k+1}} (u \cdot z - G^\infty(y; z)),$$

où $u \cdot z$ représente le produit scalaire dans \mathbb{R}^{k+1} . Soit $T > 0$ et soit $\phi : [0, T] \rightarrow \mathcal{D}^{k+1}$ une fonction. Nous définissons l'action $I_T(\phi)$ de la trajectoire ϕ par :

$$I_T(\phi) = \int_0^T H^\infty(\phi(t), \phi'(t)) dt,$$

si la fonction ϕ est absolument continue et l'intégrale converge ; nous posons $I_T(\phi) = +\infty$ sinon. Soit $\rho > 0$, nous définissons l'ensemble ρ -réduit de \mathcal{D}^{k+1} par :

$$\mathcal{D}_\rho^{k+1} = \{ y \in \mathcal{D}^{k+1} : d(y, (\mathcal{D}^{k+1})^c) \geq \rho \}.$$

Pour un point $y^0 \in \mathcal{D}_\rho^{k+1}$ et $\eta \geq 0$, notons par $\Phi_{y^0;T}(\eta)$ l'ensemble des trajectoires ϕ dans \mathcal{D}_ρ^{k+1} issues du point y^0 et dont l'action est inférieure ou égale à η , i.e.,

$$\Phi_{y^0;T}(\eta) = \{ \phi : [0, T] \longrightarrow \mathcal{D}_\rho^{k+1} : \phi(0) = y^0 \text{ et } I_T(\phi) \leq \eta \}.$$

Nous notons par d la distance euclidienne dans \mathbb{R}^{k+1} , et nous notons par d_T la distance du supremum entre deux trajectoires sur l'intervalle $[0, T]$, i.e., pour $\phi, \psi : [0, T] \longrightarrow \mathbb{R}^{k+1}$,

$$d_T(\phi, \psi) = \sup_{0 \leq t \leq T} d(\phi(t), \psi(t)).$$

Soit T_ρ le temps de sortie du processus Y de l'ensemble \mathcal{D}_ρ^{k+1} , i.e.,

$$T_\rho = \inf \{ t \geq 0 : Y(t) \notin \mathcal{D}_\rho^{k+1} \}.$$

Nous considérons le processus $(Y^\rho(t), t \geq 0)$ arrêté au temps T_ρ :

$$\forall t \geq 0 \quad Y^\rho(t) = Y(t \wedge T_\rho).$$

Théorème 22.1.1 (PGD). *Nous avons le résultat suivant :*

Borne inférieure. *Pour tout $\delta, \gamma, \rho, \eta_0, T > 0$ avec $\rho > \delta$, asymptotiquement,*

$$P\left(d_T(Y, \phi) < \delta \mid Y(0) = y^0\right) \geq \exp(-m(I_T(\phi) + \gamma)),$$

uniformément sur tout $y^0 \in \mathcal{D}_\rho^{k+1}$, et sur tout $\phi \in \Phi_{y^0;T}(\eta_0)$.

Borne supérieure. *Pour tout $\delta, \gamma, \rho, \eta_0, T > 0$, asymptotiquement,*

$$P\left(d_T(Y^\rho, \Phi_{y^0;T}(\eta)) \geq \delta \mid Y^\rho(0) = y^0\right) \leq \exp(-m(\eta - \gamma)),$$

uniformément sur $y^0 \in \mathcal{D}_\rho^{k+1}$ et $\eta \leq \eta_0$.

Le reste du chapitre est dédié à la preuve de ce principe de grandes déviations. Nous introduisons d'abord les éléments nécessaires pour la preuve. Ensuite nous démontrons la borne inférieure et finalement la borne supérieure. La preuve est exactement la même pour tout $k \geq K$. Nous faisons donc la démonstration dans le cas $k = K$ seulement. La démonstration est une généralisation du théorème de Cramér dans \mathbb{R}^d , et repose sur des idées développées par Wentzell [82, 83, 84]. Nous ne pouvons pas appliquer directement

les résultats de Wentzell pour plusieurs raisons. D'une part, Wentzell considère des familles de processus de Markov qui dépendent d'un seul paramètre (équivalent à notre m), alors que nous considérons un régime asymptotique plus compliqué. D'autre part, les hypothèses de Wentzell ne sont pas vérifiées par notre famille de processus. Une autre différence importante est le fait que notre processus Y n'est pas markovien, mais plutôt "asymptotiquement markovien". Nous sommes donc obligés de faire toute la démonstration du théorème de grandes déviations. Nous commençons par introduire la transformée de Fenchel–Legendre d'une fonction ainsi que la log–Laplace et la transformée de Cramér d'une variable aléatoire, qui sont des objets centraux dans la preuve du théorème de Cramér dans \mathbb{R}^d . Nous introduisons ensuite les généralisations de ces objets au cas des processus : le cumulants conditionnel et la transformée de Cramér généralisée. Avant de passer à la partie finale de la preuve, nous donnons les estimations nécessaires sur le cumulants et sur sa transformée de Fenchel–Legendre. Nous faisons ensuite la démonstration de la borne inférieure du principe de grandes déviations, et nous finissons en démontrant la borne supérieure.

22.2 Outils

Nous introduisons ici les outils de base dont nous aurons besoin pour montrer le principe de grandes déviations 22.1.1.

22.2.1 Transformée de Fenchel–Legendre

Nous introduisons maintenant la transformée de Fenchel–Legendre d'une fonction convexe semicontinue inférieurement. Nous donnons sans preuve les propriétés les plus pertinentes pour la suite. Les résultats de cette section peuvent être trouvés dans la section 26 du livre de Rockafellar [64]. Soit $d \geq 1$ et soit $G : \mathbb{R}^d \rightarrow]-\infty, +\infty]$ une fonction convexe semicontinue inférieurement. Nous supposons que G n'est pas identiquement égale à $+\infty$. La *transformée de Fenchel–Legendre* $H : \mathbb{R}^d \rightarrow]-\infty, +\infty]$ de G est définie par la formule :

$$\forall u \in \mathbb{R}^d \quad H(u) = \sup_{z \in \mathbb{R}^d} (z \cdot u - G(z)),$$

où $z \cdot u$ représente le produit scalaire dans \mathbb{R}^d . La fonction H est à nouveau convexe et semicontinue inférieurement. De plus, la transformée de Fenchel–Legendre de H est G , i.e.,

$$\forall z \in \mathbb{R}^d \quad G(z) = \sup_{u \in \mathbb{R}^d} (z \cdot u - H(u)).$$

Proposition 22.2.1. *Nous avons les propriétés suivantes :*

- Si $G(z) < +\infty$ pour tout $z \in \mathbb{R}^d$, alors

$$\lim_{|u| \rightarrow \infty} \frac{H(u)}{|u|} = +\infty.$$

- Si H est différentiable en $u \in \mathbb{R}^d$, alors

$$H(u) = \nabla H(u) \cdot u - G(\nabla H(u)).$$

- Si de plus G est différentiable en $z = \nabla H(u)$, alors

$$\nabla G(\nabla H(u)) = u.$$

- Si la fonction G est $k \geq 2$ fois différentiable en $z \in \mathbb{R}^d$ et la matrice des dérivées secondes $(\partial^2 G / \partial z_i \partial z_j, 1 \leq i, j \leq d)$ est non dégénérée en z , alors la fonction H est k fois différentiable en $\nabla G(z)$ et

$$\frac{\partial^2 H}{\partial u_i \partial u_j}(\nabla G(z)) = \left(\frac{\partial^2 G}{\partial z_i \partial z_j}(z) \right)^{-1}, \quad 1 \leq i, j \leq d.$$

22.2.2 La log–Laplace

Soit $d \geq 1$. Soit μ une loi de probabilité à support fini sur \mathbb{R}^d vérifiant $\mu(\{0\}) > 0$. Il existe donc un ensemble fini V de vecteurs non nuls de \mathbb{R}^d , et des constantes $0 < C(v) < 1$, pour $v \in V$, tels que

$$\mu = \sum_{v \in V} C(v) \delta_v + \left(1 - \sum_{v \in V} C(v) \right) \delta_0.$$

Soit X une variable aléatoire de loi μ et soit $G : \mathbb{R}^d \rightarrow \mathbb{R}$ sa log–Laplace, i.e.,

$$\forall z \in \mathbb{R}^d \quad G(z) = \ln E(e^{z \cdot X}) = \ln \left(1 + \sum_{v \in V} C(v) (e^{z \cdot v} - 1) \right).$$

Proposition 22.2.2. *La fonction G a les propriétés suivantes :*

- *La fonction G est convexe et analytique.*
- *Si de plus l'ensemble de vecteurs V engendre \mathbb{R}^d , la fonction G est strictement convexe.*

Démonstration. La fonction G est évidemment analytique. La convexité découle de l'inégalité de Hölder. En effet, pour tout $y, z \in \mathbb{R}^d$ et pour tout $t \in [0, 1]$,

$$\begin{aligned} G(ty + (1-t)z) &= \ln \mathbb{E} \left((e^{y \cdot X})^t (e^{z \cdot X})^{1-t} \right) \\ &\leq \ln \left(\mathbb{E}(e^{y \cdot X})^t \mathbb{E}(e^{z \cdot X})^{1-t} \right) = tG(y) + (1-t)G(z). \end{aligned}$$

Dans l'inégalité de Hölder, nous avons égalité si et seulement si il existe des nombres strictement positifs α, β tels que $\alpha e^{y \cdot X} = \beta e^{z \cdot X}$ presque sûrement. Puisque $\mu(\{0\}) > 0$, nous en déduisons que $\alpha = \beta$. Or, $e^{y \cdot X} = e^{z \cdot X}$ presque sûrement si et seulement si

$$\forall v \in V \quad y \cdot v = z \cdot v,$$

et vu que V engendre \mathbb{R}^d , nous en déduisons que $y = z$, d'où la convexité stricte de la fonction G . \square

22.2.3 Entropie relative

Soit μ une loi de probabilité à support fini sur \mathbb{R}^d , et appelons V le support de μ . Pour toute loi de probabilité ν sur \mathbb{R}^d , nous définissons l'entropie relative de ν par rapport à μ par :

$$L(\nu | \mu) = \begin{cases} \sum_{v \in V} \nu(v) \ln \frac{\nu(v)}{\mu(v)} & \text{si } \nu \ll \mu, \\ +\infty & \text{sinon.} \end{cases}$$

Dans cette définition, nous supposons que $0 \ln 0 = 0$.

Lemme 22.2.3. *Pour toute loi de probabilité ν sur \mathbb{R}^d , l'entropie relative $L(\nu | \mu)$ est positive. De plus, $L(\nu | \mu) = 0$ si et seulement si $\nu = \mu$.*

Démonstration. Soit ν une loi de probabilité sur \mathbb{R}^d à support inclus dans V (autrement le résultat est trivial). Pour tout $x \geq 0$ nous avons $x \ln x \geq x - 1$, avec égalité si et seulement si $x = 1$. Nous en déduisons que :

$$\sum_{\nu \in V} \nu(\nu) \ln \frac{\nu(\nu)}{\mu(\nu)} = \sum_{\nu \in V} \mu(\nu) \frac{\nu(\nu)}{\mu(\nu)} \ln \frac{\nu(\nu)}{\mu(\nu)} \geq \sum_{\nu \in V} \mu(\nu) \left(\frac{\nu(\nu)}{\mu(\nu)} - 1 \right) = 0,$$

avec égalité si et seulement si $\nu(\nu)/\mu(\nu) = 1$ pour tout $\nu \in V$. \square

Lemme 22.2.4. *Pour toute loi de probabilité ν sur \mathbb{R}^d , nous avons :*

$$L(\nu | \mu) = \sup \left\{ \int_{\mathbb{R}^d} f(x) \nu(dx) - \ln \int_{\mathbb{R}^d} e^{f(x)} \mu(dx) \right\},$$

où le supremum est pris sur toutes les fonctions $f : \mathbb{R}^d \rightarrow \mathbb{R}$ bornées.

Démonstration. L'égalité est évidemment vraie pour des mesures ν n'étant pas absolument continues par rapport à μ . Supposons donc que $\nu \ll \mu$, et soit $f : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction bornée. Le logarithme étant une fonction concave, nous avons :

$$\begin{aligned} \int_{\mathbb{R}^d} f(x) \nu(dx) - L(\nu | \mu) &= \sum_{\nu \in V} f(\nu) \nu(\nu) - \sum_{\nu \in V} \nu(\nu) \ln \frac{\nu(\nu)}{\mu(\nu)} = \\ &= \sum_{\nu \in V} \nu(\nu) \ln \left(e^{f(\nu)} \frac{\mu(\nu)}{\nu(\nu)} \right) \leq \ln \left(\sum_{\nu \in V} \mu(\nu) e^{f(\nu)} \right) = \ln \int_{\mathbb{R}^d} e^{f(x)} \mu(dx). \end{aligned}$$

Nous en déduisons que

$$L(\nu | \mu) \geq \sup_f \left\{ \int_{\mathbb{R}^d} f(x) \nu(dx) - \ln \int_{\mathbb{R}^d} e^{f(x)} \mu(dx) \right\}.$$

Soit $f : \mathbb{R}^d \rightarrow [-\infty, +\infty[$ la fonction définie par :

$$\forall x \in \mathbb{R}^d \quad f(x) = \begin{cases} \ln \frac{\nu(x)}{\mu(x)} & \text{si } x \in V, \\ 0 & \text{sinon.} \end{cases}$$

La fonction f est bien bornée, et vérifie :

$$L(\nu | \mu) = \int_{\mathbb{R}^d} f(x) \nu(dx) - \ln \int_{\mathbb{R}^d} e^{f(x)} \mu(dx),$$

ce qui conclut la preuve. \square

Lemme 22.2.5. *Pour toute fonction bornée $f : \mathbb{R}^d \rightarrow \mathbb{R}$, nous avons :*

$$\ln \int_{\mathbb{R}^d} e^{f(x)} \mu(dx) = \sup \left\{ \int_{\mathbb{R}^d} f(x) \nu(dx) - L(\nu | \mu) \right\},$$

où le supremum est pris sur toutes les mesures de probabilité ν absolument continues par rapport à μ .

Démonstration. Soit $f : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction bornée et soit ν une mesure de probabilité absolument continue par rapport à μ . Comme dans la preuve du lemme précédent, nous avons :

$$\int_{\mathbb{R}^d} f(x) \nu(dx) - L(\nu | \mu) \leq \ln \int_{\mathbb{R}^d} e^{f(x)} \mu(dx).$$

Nous en déduisons que

$$\ln \int_{\mathbb{R}^d} e^{f(x)} \mu(dx) \geq \sup_{\nu} \left\{ \int_{\mathbb{R}^d} f(x) \nu(dx) - L(\nu | \mu) \right\}.$$

Pour réaliser l'égalité, nous définissons la mesure ν de support V par :

$$\forall v \in V \quad \nu(v) = \frac{\mu(v) e^{f(v)}}{\sum_{v \in V} \mu(v) e^{f(v)}},$$

ce qui conclut la démonstration du lemme. □

22.2.4 Transformée de Fenchel–Legendre de la log–Laplace

Comme dans la section 22.2.2, nous considérons une loi de probabilité μ de support fini sur \mathbb{R}^d vérifiant $\mu(\{0\}) > 0$ et X une variable aléatoire de loi μ . Soit $G(z)$ la log–Laplace de X :

$$G(z) = \ln \left(1 + \sum_{v \in V} C(v) (e^{v \cdot z} - 1) \right).$$

Nous considérons maintenant la transformée de Fenchel–Legendre de la fonction G :

$$\forall u \in \mathbb{R}^d \quad H(u) = \sup_{z \in \mathbb{R}^d} (u \cdot z - G(z)).$$

Théorème 22.2.6. *La fonction H vérifie :*

$$H(u) = \inf \left\{ L(\nu | \mu) : \int_{\mathbb{R}^d} x \nu(dx) = u \right\}.$$

Démonstration. Notons par $\bar{\nu}$ la moyenne d'une mesure ν , i.e., $\bar{\nu} = \int x \nu(dx)$. D'après le lemme 22.2.5, nous avons, pour tout $z \in \mathbb{R}^d$:

$$-G(z) = -\ln \int_{\mathbb{R}^d} e^{x \cdot z} \mu(dx) = \inf_{\nu} \left\{ L(\nu | \mu) - z \cdot \bar{\nu} \right\}.$$

Donc, pour tout $u \in \mathbb{R}^d$:

$$H(u) = \sup_{z \in \mathbb{R}^d} \left(u \cdot z + \inf_{\nu} \left\{ L(\nu | \mu) - z \cdot \bar{\nu} \right\} \right) = \sup_{z \in \mathbb{R}^d} \inf_{\nu} \left\{ L(\nu | \mu) - z \cdot (\bar{\nu} - u) \right\}.$$

Puisque pour tout $\nu \neq u$, nous avons $\sup_z \{ z \cdot (\nu - u) \} = +\infty$, nous en déduisons que :

$$\inf_{\nu} \sup_{z \in \mathbb{R}^d} \left\{ L(\nu | \mu) - z \cdot (\bar{\nu} - u) \right\} = \inf \left\{ L(\nu | \mu) : \bar{\nu} = u \right\}.$$

Il suffit donc de justifier l'échange entre sup et inf dans la formule précédente. Nous le faisons à l'aide du théorème minimax.

Théorème 22.2.7. *Soient X, Y deux espaces vectoriels. Soit A un ensemble compact de X et soit $f : A \times Y \rightarrow \mathbb{R}$ une fonction vérifiant :*

- *Pour tout $y \in Y$ l'application $x \mapsto f(x, y)$ est convexe et semi-continue inférieurement.*
- *Pour tout $x \in A$ l'application $y \mapsto f(x, y)$ est concave.*

Nous avons alors :

$$\inf_{x \in A} \sup_{y \in Y} f(x, y) = \sup_{y \in Y} \inf_{x \in A} f(x, y).$$

Ce théorème est dû à Ky Fan [34]. Pour toute mesure ν et tout $z \in \mathbb{R}^d$ nous posons $f(\nu, z) = L(\nu | \mu) - z \cdot (\bar{\nu} - u)$. Il s'agit de montrer que

$$\sup_{z \in \mathbb{R}^d} \inf_{\nu} f(\nu, z) = \inf_{\nu} \sup_{z \in \mathbb{R}^d} f(\nu, z).$$

L'entropie n'étant finie que pour des mesures de probabilité ν absolument continues par rapport à μ , nous voyons qu'il suffit de considérer dans les infimum précédents les mesures $\nu \ll \mu$. Puisque μ est à support fini $V \cup \{0\}$, l'ensemble de mesures ν peut être identifié avec le simplexe $S^{|V|+1}$, qui est un sous-ensemble compact de $\mathbb{R}^{|V|+1}$. Pour tout $z \in \mathbb{R}^d$, la fonction $\nu \mapsto f(\nu, z)$ est le

supremum de fonctions affines de v (lemme 22.2.4), elle est donc convexe et semi-continue inférieurement. Pour tout $v \ll \mu$, la fonction $z \mapsto f(v, z)$ est une fonction affine, et donc concave. Nous pouvons alors conclure la démonstration grâce au théorème minimax. \square

Notons par $\text{Conv}(V)$ l'enveloppe convexe de l'ensemble V :

$$\text{Conv}(V) = \left\{ \sum_{v \in V} \alpha(v)v : \alpha(v) \geq 0 \text{ et } \sum_{v \in V} \alpha(v) = 1 \right\}.$$

L'ensemble $\text{Conv}(V)$ est fermé.

Proposition 22.2.8. *La fonction H vérifie les propriétés suivantes :*

- *La fonction H est finie sur $\text{Conv}(V)$ et infinie ailleurs.*
- *La fonction H est analytique sur $\text{int}(\text{Conv}(V))$.*
- *La valeur $H(u)$ vaut zéro si et seulement si $u = E(X)$, où X est une variable aléatoire de loi μ .*

Démonstration. La première affirmation est une conséquence du théorème précédent. La deuxième affirmation découle du théorème 26.5 dans [64]. La troisième affirmation résulte du théorème précédent et du lemme 22.2.3. \square

22.2.5 Transformée de Cramér

Soit X une variable aléatoire discrète sur \mathbb{R}^d et soit G sa transformée de Laplace logarithmique. Pour $z \in \mathbb{R}^d$, définissons la variable aléatoire $\pi(z)$ par :

$$\pi(z) = \exp(z \cdot X - G(z)).$$

La variable aléatoire $\pi(z)$ a pour espérance 1. Nous pouvons donc définir une nouvelle mesure de probabilité Q par :

$$dQ = \pi(z) dP.$$

La mesure de probabilité Q est appelée la transformée de Cramér de la loi de X .

Proposition 22.2.9. *Sous la mesure de probabilité Q , nous avons, pour tout i, j dans $\{1, \dots, d\}$:*

$$E_Q(X) = \nabla G(z), \quad \text{cov}_Q(X_i, X_j) = \frac{\partial^2 G}{\partial z_i \partial z_j}(z).$$

Démonstration. Pour tout $i \in \{1, \dots, d\}$, nous avons :

$$\frac{\partial G}{\partial z_i}(z) = \frac{E_P(X_i e^{z \cdot X})}{E_P(e^{z \cdot X})} = E_P(X_i \pi(z)) = E_Q(X_i).$$

Également, pour tout $i, j \in \{1, \dots, d\}$:

$$\begin{aligned} \frac{\partial^2 G}{\partial z_i \partial z_j}(z) &= \frac{E_P(X_i X_j e^{z \cdot X})}{E_P(e^{z \cdot X})} - \frac{E_P(X_i e^{z \cdot X}) E_P(X_j e^{z \cdot X})}{E_P(e^{z \cdot X})^2} = \\ &= \frac{E_P(X_i X_j \pi(z)) - E_P(X_i \pi(z)) E_P(X_j \pi(z))}{E_P(e^{z \cdot X})} = \\ &= E_Q(X_i X_j) - E_Q(X_i) E_Q(X_j) = \text{Cov}_Q(X_i, X_j), \end{aligned}$$

comme voulu. □

22.2.6 Le cumulants conditionnel

Les outils que nous avons définis jusqu'à présent, nous les avons définis pour une loi à support fini dans \mathbb{R}^d quelconque. Nous reprenons maintenant notre processus $(Y(t))_{t \geq 0}$, et nous définissons quelques outils additionnels, cette fois-ci, spécifiques au processus. Définissons la fonction $G : \mathcal{S}^{\ell+1} \times \mathbb{R}^{K+1} \rightarrow \mathbb{R}$ par :

$$\forall x \in \mathcal{S}^{\ell+1} \quad \forall z \in \mathbb{R}^{K+1} \quad G(x; z) = \ln \left(1 + \sum_{v \in \mathcal{V}} C_x(v) (e^{v \cdot z} - 1) \right).$$

Conditionnellement à $X(0) = x$, l'incrément $Y(\tau) - Y(0)$ a pour log-Laplace la fonction $G(x; z/m)$, c'est-à-dire :

$$\forall x \in \mathcal{S}^{\ell+1} \quad \forall z \in \mathbb{R}^{K+1} \quad \ln E \left(e^{z \cdot (Y(\tau) - Y(0))} \mid X_0 = x \right) = G(x; z/m).$$

Nous appelons la fonction $G(x; \cdot/m)$ le *cumulants conditionnel* du processus Y . La proposition suivante résume les propriétés les plus importantes du cumulants :

Proposition 22.2.10. *Pour tout $x \in \mathcal{S}^{\ell+1}$, comme fonction de $z \in \mathbb{R}^{K+1}$, la fonction $G(x; z)$ vérifie les propriétés suivantes :*

- *La fonction $G(x; z)$ est convexe et continue.*
- *$G(x; 0) = 0$ et pour tout $z \in \mathbb{R}^{K+1}$, $G(x; z) < +\infty$.*
- *La fonction $G(x; z)$ est analytique et la matrice des dérivées secondes est définie positive.*

La proposition suivante résume les propriétés les plus importantes du cumulants en relation avec les incréments du processus Y :

Proposition 22.2.11. *Soit $x \in \Sigma^{\ell+1}$ et soient $i, j \in \{0, \dots, K\}$. Nous avons :*

$$\begin{aligned} \mathbb{E}\left(Y_i(\tau) - Y_i(0) \mid \mathcal{X}_0 = x\right) &= \frac{1}{m^2} \frac{\partial G}{\partial z_i}(x; 0), \\ \text{cov}\left(Y_i(\tau) - Y_i(0), Y_j(\tau) - Y_j(0) \mid \mathcal{X}_0 = x\right) &= \frac{1}{m} \frac{\partial^2 G}{\partial z_i \partial z_j}(x; 0). \end{aligned}$$

Les propositions 22.2.10 et 22.2.11 sont des conséquences directes des propriétés de la log-Laplace et de la transformée de Cramér (cf. les sections 22.2.2 et 22.2.5).

22.2.7 La transformée de Cramér généralisée

Les processus \mathcal{X}, Y sont des processus constants par morceaux qui sautent dans des instants multiples de τ . Soient $t_1 < t_2$ deux multiples positifs de τ , et notons par $[[t_1, t_2]]$ l'ensemble de multiples de τ dans l'intervalle $[t_1, t_2]$. Soit $(\mathcal{F}_t)_{t \geq 0}$ la filtration générée par le processus $(\mathcal{X}(t))_{t \geq 0}$ et soit $(Z(t))_{t \geq 0}$ un processus adapté à la filtration $(\mathcal{F}_t)_{t \geq 0}$. Posons :

$$\pi(t_1, t_2) = \exp \left(\sum_{t \in [[t_1, t_2 - \tau]]} Z(t) \cdot (Y(t + \tau) - Y(t)) - \sum_{t \in [[t_1, t_2 - \tau]]} G(\mathcal{X}(t); Z(t)/m) \right).$$

Lemme 22.2.12. *Le processus $\pi(0, t)_{t=0, \tau, \dots}$ est une martingale vis-à-vis de la filtration $(\mathcal{F}_t)_{t=0, \tau, \dots}$.*

Démonstration. Soit t un multiple de τ . Nous avons :

$$\begin{aligned} E(\pi(0, t + \tau) \mid \mathcal{F}_t) &= \pi(0, t) E(\pi(t, t + \tau) \mid \mathcal{F}_t) = \\ &= \pi(0, t) E\left(e^{Z(t) \cdot (Y(t+\tau) - Y(t))} \mid \mathcal{F}_t\right) e^{-G(\mathcal{X}(t); Z(t)/m)}. \end{aligned}$$

Or, $G(\cdot; \cdot/m)$ étant le cumulants conditionnel du processus Y , nous avons

$$\ln E\left(e^{Z(t) \cdot (Y(t+\tau) - Y(t))} \mid \mathcal{F}_t\right) = G(\mathcal{X}(t); Z(t)/m),$$

ce qui donne le résultat voulu. \square

Soit maintenant $T > 0$, soit $z : [0, T] \times \mathcal{D}^{K+1} \rightarrow \mathbb{R}^{K+1}$ une fonction mesurable, et introduisons la fonctionnelle π associée à $(z(t, Y(t)))_{0 \leq t \leq T}$, donnée par

$$\forall t \in [[0, T]] \quad \pi(0, t) = \exp \left(\sum_{t \in [[0, t-\tau]]} z(t, Y(t)) \cdot (Y(t + \tau) - Y(t)) - \sum_{t \in [[0, t-\tau]]} G(\mathcal{X}(t); z(t, Y(t))/m) \right).$$

Pour tout $x^0 \in \Sigma^{\ell+1}$, nous notons P_{x^0} la loi du processus Y sachant que $\mathcal{X}_0 = x^0$. La transformée de Cramér généralisée associée à la fonction $z(t, y)$ consiste à remplacer la mesure P_{x^0} par la mesure $P_{x^0}^z$ définie par : pour tout $A \in \mathcal{F}_T$,

$$P_{x^0}^z(A) = E_{x^0}(1_A \pi(0, T)).$$

D'après le lemme précédent, $P_{x^0}^z$ est bien une mesure de probabilité.

22.3 Bornes supplémentaires

Nous donnons ici quelques bornes supplémentaires que nous utiliserons dans la démonstration du PGD.

22.3.1 Les probabilités C_x

Nous montrons maintenant qu'asymptotiquement, les quantités C_x sont minorées par une constante strictement positive, uniformément sur $x \in \mathcal{S}^{\ell+1}$ tel que $\pi_K(x) \in \mathcal{D}_\rho^{K+1}$.

Lemme 22.3.1. Soit $\rho > 0$. Il existe $\underline{c} > 0$ tel que, pour $\ell, 1/q$ assez grands et ℓq assez proche de α , et pour tout $x \in \mathcal{S}^{\ell+1}$ vérifiant $\pi_K(x) \in \mathcal{D}_\rho^{K+1}$,

$$\forall v \in \mathcal{V} \quad C_x(v) \geq \underline{c}, \quad 1 - \sum_{v \in \mathcal{V}} C_x(v) \geq \underline{c}.$$

Démonstration. Soit $\rho > 0$ assez petit pour que

$$\forall k \in \{0, \dots, K+1\} \quad e^{-\alpha} \frac{\alpha^k}{k!} - \rho > 0.$$

Posons

$$\underline{c} = \min_{0 \leq k \leq K+1} \rho^2 f_H(0) \left(e^{-\alpha} \frac{\alpha^k}{k!} - \rho \right).$$

Pour $\ell, 1/q$ assez grands et ℓq assez proche de α ,

$$\forall j \in \{0, \dots, K+1\} \quad M_H(0, j) > e^{-\alpha} \frac{\alpha^j}{j!} - \rho.$$

Soit $x \in \mathcal{S}^{\ell+1}$ tel que $\pi(x) \in \mathcal{D}_\rho^{K+1}$. Pour tout $i \in \{0, \dots, K\}$, nous avons :

$$\begin{aligned} C_x(-e_i) &= x_i \sum_{0 \leq h \leq \ell} x_h f_H(h) \left(1 - \sum_{0 \leq t \leq K} M_H(h, t) \right) \\ &\geq \rho^2 f_H(0) M_H(0, K+1) > \rho^2 f_H(0) \left(e^{-\alpha} \frac{\alpha^{K+1}}{(K+1)!} - \rho \right) \geq \underline{c}. \end{aligned}$$

Également, pour tout $j \in \{0, \dots, K\}$:

$$C_x(e_j) = \left(1 - \sum_{0 \leq t \leq K} x_t \right) \sum_{0 \leq h \leq \ell} x_h f_H(h) M_H(h, j) \geq \rho^2 f_H(0) M_H(0, j) > \underline{c}.$$

De même, pour tout $i, j \in \{0, \dots, K\}$ tels que $i \neq j$:

$$C_x(-e_i + e_j) = x_i \sum_{0 \leq h \leq \ell} x_h f_H(h) M_H(h, j) \geq \rho^2 f_H(0) M_H(0, j) > \underline{c}.$$

Finalement :

$$\begin{aligned} 1 - \sum_{v \in \mathcal{V}} C_x(v) &\geq C_y(-e_0 + e_0) \\ &\geq x_i \sum_{0 \leq h \leq \ell} x_h f_H(h) M_H(h, 0) \geq \rho^2 f_H(0) M_H(0, 0) > \underline{c}, \end{aligned}$$

comme voulu. □

Les quantités C_x convergent vers les quantités $C_{\pi_K(x)}^\infty$. Nous montrons maintenant que cette convergence a lieu uniformément sur $x \in \mathcal{S}^{\ell+1}$. Dans l'énoncé du résultat qui suit, ainsi que dans sa démonstration, le symbole " \rightarrow " représente la limite quand $\ell, 1/q$ tendent vers l'infini et ℓq tend vers a .

Lemme 22.3.2. *Pour tout $v \in \mathcal{V}$,*

$$C_x(v) \rightarrow C_{\pi_K(x)}^\infty(v),$$

uniformément sur $x \in \mathcal{S}^{\ell+1}$.

Démonstration. D'après le lemme 4.4.2 de la partie I, pour tout $k \geq 0$, les probabilités $M_H(k+i, k)$ convergent vers 0, uniformément par rapport à $i \geq 1$. Il existe donc une suite $\varepsilon = \varepsilon(\ell, q)$ vérifiant $\varepsilon \rightarrow 0$ et

$$\begin{aligned} \forall k \in \{0, \dots, K\} \quad \forall i \geq 1 \quad M_H(k+i, k) < \varepsilon, \\ \forall i \in \{0, \dots, K\} \quad \forall j \in \{i, \dots, K\} \quad \left| M_H(i, j) - e^{-a} \frac{a^{j-i}}{(j-i)!} \right| < \varepsilon. \end{aligned}$$

Donc, pour $i \in \{0, \dots, K\}$,

$$\begin{aligned} |C_x(-e_i) - C_{\pi_K(x)}^\infty(-e_i)| = \\ x_i \sum_{0 \leq h \leq K} x_h f_H(h) \left(\sum_{0 \leq t \leq h-1} M_H(h, t) + \sum_{h \leq t \leq K} \left| M_H(h, t) - e^{-a} \frac{a^{t-h}}{(t-h)!} \right| \right) \\ + x_i \sum_{K+1 \leq h \leq \ell} x_h \sum_{0 \leq t \leq K} M_H(h, t) \leq f_H(0)(K+1)\varepsilon, \end{aligned}$$

qui tend vers 0 uniformément sur $x \in \mathcal{S}^{\ell+1}$. Également, pour tout $j \in \{0, \dots, K\}$,

$$\begin{aligned} |C_x(e_j) - C_{\pi_K(x)}^\infty(e_j)| = \left(1 - \sum_{0 \leq t \leq K} x_t \right) \sum_{0 \leq h \leq j} x_h f_H(h) \left| M_H(h, j) - e^{-a} \frac{a^{j-h}}{(j-h)!} \right| \\ + \left(1 - \sum_{0 \leq t \leq K} x_t \right) \sum_{j+1 \leq h \leq \ell} x_h f_H(h) M_H(h, j) < f_H(0)\varepsilon, \end{aligned}$$

qui tend vers 0 uniformément sur $x \in \mathcal{S}^{\ell+1}$. Finalement, pour tout i, j dans $\{0, \dots, K\}$,

$$\begin{aligned} |C_x(-e_i + e_j) - C_{\pi_K(x)}^\infty(-e_i + e_j)| = x_i \sum_{0 \leq h \leq j} x_h f_H(h) \left| M_H(h, j) - e^{-a} \frac{a^{j-h}}{(j-h)!} \right| \\ + x_i \sum_{j+1 \leq h \leq \ell} x_h f_H(h) M_H(h, j) < f_H(0)\varepsilon, \end{aligned}$$

qui tend vers 0 uniformément sur $x \in \mathcal{S}^{\ell+1}$. \square

22.3.2 Convergence uniforme de G et de ses dérivées

Comme dans la section précédente, le symbole " \rightarrow " représente la limite quand $\ell, 1/q$ tendent vers l'infini et ℓq tend vers a .

Proposition 22.3.3. *Soit $\mathcal{K} \subset \mathbb{R}^{K+1}$ un ensemble compact. Les convergences suivantes ont lieu uniformément sur $x \in \mathcal{S}^{\ell+1}$ et $z \in \mathcal{K}$:*

$$\begin{aligned} G(x; z) &\rightarrow G^\infty(\pi_K(x), z), \\ \forall i \in \{0, \dots, K\} \quad \frac{\partial G}{\partial z_i}(x; z) &\rightarrow \frac{\partial G^\infty}{\partial z_i}(\pi_K(x), z), \\ \forall i, j \in \{0, \dots, K\} \quad \frac{\partial^2 G}{\partial z_i \partial z_j}(x; z) &\rightarrow \frac{\partial^2 G^\infty}{\partial z_i \partial z_j}(\pi_K(x), z). \end{aligned}$$

Démonstration. Définissons d'abord les fonctions auxiliaires $\tilde{G}, \tilde{G}^\infty$ par :

$$\begin{aligned} \forall x \in \mathcal{S}^{\ell+1} \quad \forall z \in \mathbb{R}^{K+1} \quad \tilde{G}(x; z) &= \sum_{v \in \mathcal{V}} C_x(v)(e^{v \cdot z} - 1), \\ \forall y \in \mathcal{D}^{K+1} \quad \forall z \in \mathbb{R}^{K+1} \quad \tilde{G}^\infty(y; z) &= \sum_{v \in \mathcal{V}} C_y^\infty(v)(e^{v \cdot z} - 1). \end{aligned}$$

Lemme 22.3.4. *Les convergences suivantes ont lieu uniformément sur $x \in \mathcal{S}^{\ell+1}$ et $z \in \mathcal{K}$:*

$$\begin{aligned} \tilde{G}(x; z) &\rightarrow \tilde{G}^\infty(\pi_K(x); z), \\ \forall i \in \{0, \dots, K\} \quad \frac{\partial \tilde{G}}{\partial z_i}(x; z) &\rightarrow \frac{\partial \tilde{G}^\infty}{\partial z_i}(\pi_K(x); z), \\ \forall i, j \in \{0, \dots, K\} \quad \frac{\partial^2 \tilde{G}}{\partial z_i \partial z_j}(x; z) &\rightarrow \frac{\partial^2 \tilde{G}^\infty}{\partial z_i \partial z_j}(\pi_K(x); z). \end{aligned}$$

Démonstration. D'après le lemme 22.3.2, quand $\ell, 1/q$ tendent vers l'infini et ℓq tend vers a , les quantités $C_x(v)$ convergent vers les quantités $C_{\pi_K(x)}^\infty(v)$, uniformément sur $x \in \mathcal{S}^{\ell+1}$. Il existe donc une suite $\varepsilon = \varepsilon(\ell, q)$ qui tend vers 0 quand $\ell, 1/q$ tendent vers l'infini et ℓq tend vers a vérifiant

$$\forall v \in \mathcal{V} \quad |C_x(v) - C_{\pi_K(x)}^\infty(v)| < \varepsilon,$$

uniformément sur $x \in \mathcal{S}^{\ell+1}$. Nous avons donc, pour tout $x \in \mathcal{S}^{\ell+1}$ et $z \in \mathcal{K}$:

$$|\tilde{G}(x; z) - \tilde{G}^\infty(x; z)| \leq \sum_{v \in \mathcal{V}} |C_x(v) - C_{\pi_K(x)}^\infty(v)| |e^{v \cdot z} - 1| < |\mathcal{V}| \varepsilon \sup_{z \in \mathcal{K}, v \in \mathcal{V}} |e^{v \cdot z} - q|,$$

qui tend vers 0 uniformément sur $x \in \mathcal{S}^{\ell+1}$ et $z \in \mathcal{K}$. Les affirmations sur les dérivées de \tilde{G} se démontrent de façon similaire. \square

Remarquons maintenant que si X est une variable aléatoire à valeurs dans \mathbb{R}^{K+1} de loi

$$\sum_{v \in V} C_y^\infty(v) \delta_v + \left(1 - \sum_{v \in V} C_y^\infty(v)\right) \delta_0,$$

alors,

$$1 + \tilde{G}^\infty(x; z) = \mathbb{E}(e^{z \cdot X}) > 0.$$

Puisque $1 + \tilde{G}^\infty(x; z)$ est une fonction continue de x et z , elle atteint son infimum, \underline{c} , sur l'ensemble compact $\mathcal{S}^{\ell+1} \times \mathcal{K}$, et \underline{c} est strictement positif. Puisque \tilde{G} converge vers \tilde{G}^∞ , uniformément sur $x \in \mathcal{S}^{\ell+1}$ et $z \in \mathcal{K}$, il existe une suite $\varepsilon = \varepsilon(\ell, q)$ qui tend vers 0 avec ℓ, q et telle que pour tout $x \in \mathcal{S}^{\ell+1}$ et $z \in \mathcal{K}$,

$$\left| \frac{\tilde{G}(x; z) - \tilde{G}^\infty(\pi_K(x); z)}{1 + \tilde{G}^\infty(\pi_K(x); z)} \right| \leq \frac{|\tilde{G}(x; z) - \tilde{G}^\infty(\pi_K(x); z)|}{\underline{c}} < \varepsilon.$$

Le logarithme étant une fonction monotone, nous en déduisons que :

$$\begin{aligned} |G(x; z) - G^\infty(\pi_K(x); z)| &= \left| \ln \left(1 + \frac{\tilde{G}(x; z) - \tilde{G}^\infty(\pi_K(x); z)}{1 + \tilde{G}^\infty(\pi_K(x); z)} \right) \right| \\ &< \max(\ln(1 + \varepsilon), -\ln(1 - \varepsilon)). \end{aligned}$$

cette quantité tend vers 0 uniformément sur $x \in \mathcal{S}^{\ell+1}$ et $z \in \mathcal{K}$. Fixons maintenant $i \in \{0, \dots, K\}$ et dérivons G par rapport à z_i :

$$\frac{\partial G}{\partial z_i}(x; z) = (1 + \tilde{G}(x; z))^{-1} \frac{\partial \tilde{G}}{\partial z_i}(x; z).$$

L'expression pour la dérivée de G^∞ est similaire. Nous avons donc :

$$\begin{aligned} \left| \frac{\partial G}{\partial z_i}(x; z) - \frac{\partial G^\infty}{\partial z_i}(\pi_K(x); z) \right| &\leq (1 + \tilde{G}(x; z))^{-1} (1 + \tilde{G}^\infty(\pi_K(x); z))^{-1} \\ &\times \left| \frac{\partial \tilde{G}}{\partial z_i}(x; z) (1 + \tilde{G}^\infty(\pi_K(x); z)) - \frac{\partial \tilde{G}^\infty}{\partial z_i}(\pi_K(x); z) (1 + \tilde{G}(x; z)) \right|. \end{aligned}$$

Par convergence uniforme, il existe une constante strictement positive \underline{C} , minorant les fonctions $1 + \tilde{G}$ et $1 + \tilde{G}^\infty$ uniformément sur $\mathcal{S}^{\ell+1} \times \mathcal{K}$. Nous avons

donc,

$$\begin{aligned} \left| \frac{\partial G}{\partial z_i}(x; z) - \frac{\partial G^\infty}{\partial z_i}(\pi_K(x); z) \right| &\leq \frac{1}{\underline{C}^2} \left(\left| \frac{\partial \tilde{G}}{\partial z_i}(x; z) - \frac{\partial \tilde{G}^\infty}{\partial z_i}(\pi_K(x); z) \right| \right. \\ &\quad + \left| \frac{\partial \tilde{G}}{\partial z_i}(x; z) \tilde{G}^\infty(\pi_K(x); z) - \frac{\partial \tilde{G}^\infty}{\partial z_i}(\pi_K(x); z) \tilde{G}^\infty(\pi_K(x); z) \right| \\ &\quad \left. + \left| \frac{\partial \tilde{G}^\infty}{\partial z_i}(\pi_K(y); z) \tilde{G}^\infty(\pi_K(x); z) - \frac{\partial \tilde{G}^\infty}{\partial z_i}(\pi_K(x); z) \tilde{G}^\infty(x; z) \right| \right). \end{aligned}$$

Les fonctions \tilde{G}^∞ et $\partial \tilde{G}^\infty / \partial z_i$ sont continues, elles sont donc bornées par une constante \bar{C} sur le compact $\mathcal{D}^{K+1} \times \mathcal{K}$. Nous concluons grâce à la convergence uniforme de \tilde{G} et de ses dérivées. L'affirmation sur les dérivées secondes de G se démontre d'une façon similaire. \square

Corollaire 22.3.5. *Les fonctions G , G^∞ et ses dérivées premières et secondes sont uniformément bornées sur $x \in \mathcal{S}^{\ell+1}$ (ou $x \in \mathcal{D}^{K+1}$ dans le cas de G^∞) et $z \in \mathcal{K}$.*

22.3.3 La fonction G^∞ et sa transformée de Fenchel–Legendre

Rappelons que la fonction $G^\infty : \mathcal{D}^{K+1} \times \mathbb{R}^{K+1} \rightarrow \mathbb{R}$ est définie par :

$$\forall y \in \mathcal{D}^{K+1} \quad \forall z \in \mathbb{R}^{K+1} \quad G^\infty(y; z) = \ln \left(1 + \sum_{v \in \mathcal{V}} C_y^\infty(v) (e^{v \cdot z} - 1) \right).$$

Lemme 22.3.6. *La fonction $G^\infty(y; z)$ est uniformément continue en $y \in \mathcal{D}_\rho^{K+1}$, uniformément sur $z \in \mathbb{R}^{K+1}$.*

Démonstration. Nous avons, pour tout $y, y' \in \mathcal{D}^{K+1}$ et $z \in \mathbb{R}^{K+1}$:

$$\begin{aligned} G^\infty(y; z) - G^\infty(y'; z) &= \ln \left(\frac{1 + \sum_{v \in \mathcal{V}} C_y^\infty(v) (e^{v \cdot z} - 1)}{1 + \sum_{v \in \mathcal{V}} C_{y'}^\infty(v) (e^{v \cdot z} - 1)} \right) \\ &= \ln \left(1 + \frac{\sum_{v \in \mathcal{V}} (C_y^\infty(v) - C_{y'}^\infty(v)) (e^{v \cdot z} - 1)}{1 + \sum_{v \in \mathcal{V}} C_{y'}^\infty(v) (e^{v \cdot z} - 1)} \right). \end{aligned}$$

Pour démontrer le lemme, il suffit de voir que pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que pour tout $y, y' \in \mathcal{D}_\rho^{K+1}$ et $z \in \mathbb{R}^{K+1}$,

$$|y - y'| < \delta \quad \implies \quad \left| \frac{\sum_{v \in \mathcal{V}} (C_y^\infty(v) - C_{y'}^\infty(v)) (e^{v \cdot z} - 1)}{1 + \sum_{v \in \mathcal{V}} C_{y'}^\infty(v) (e^{v \cdot z} - 1)} \right| < \varepsilon.$$

Nous réécrivons z sous la forme $z = \lambda u$, avec $|u| = 1$. Soit $\mathcal{V}(u)$ le sous-ensemble des vecteurs $v \in \mathcal{V}$ maximisant le produit vu et prenons $v^* \in \mathcal{V}(u)$. Nous avons alors :

$$\left| \frac{\sum_{v \in \mathcal{V}} (C_y^\infty(v) - C_{y'}^\infty(v))(e^{v \cdot z} - 1)}{1 + \sum_{v \in \mathcal{V}} C_{y'}^\infty(v)(e^{v \cdot z} - 1)} \right| \leq \frac{\sum_{v \in \mathcal{V}} |C_y^\infty(v) - C_{y'}^\infty(v)| (e^{(v-v^*) \cdot u\lambda} + e^{-v^* \cdot u\lambda})}{(1 - \sum_{v \in \mathcal{V}} C_{y'}^\infty(v)) e^{-v^* \cdot u\lambda} + \sum_{v \in \mathcal{V}} C_y^\infty(v) e^{(v-v^*) \cdot u\lambda}}.$$

La quantité $e^{(v-v^*) \cdot u\lambda} + e^{-v^* \cdot u\lambda}$ est majorée par 2, et le numérateur est minoré par la somme $\sum_{v \in \mathcal{V}(u)} C_{y'}^\infty(v)$. Nous avons donc,

$$\left| \frac{\sum_{v \in \mathcal{V}} (C_y^\infty(v) - C_{y'}^\infty(v))(e^{v \cdot z} - 1)}{1 + \sum_{v \in \mathcal{V}} C_{y'}^\infty(v)(e^{v \cdot z} - 1)} \right| \leq 2 \frac{\sum_{v \in \mathcal{V}} |C_y^\infty(v) - C_{y'}^\infty(v)|}{\sum_{v \in \mathcal{V}(u)} C_{y'}^\infty(v)}.$$

Les coefficients $C_y^\infty(v)$ sont minorés par une constante strictement positive sur l'ensemble $y \in \mathcal{D}_\rho^{K+1}$. De plus, les coefficients $C_y^\infty(v)$ sont des fonctions absolument continues de $y \in \mathcal{D}_\rho^{K+1}$. Il suffit donc de prendre δ assez petit pour que

$$|y - y'| < \delta \implies \sup_{v' \subset \mathcal{V}} 2 \frac{\sum_{v \in \mathcal{V}} |C_y^\infty(v) - C_{y'}^\infty(v)|}{\sum_{v \in \mathcal{V}'} C_{y'}^\infty(v)} < \varepsilon,$$

et la démonstration est achevée. □

Considérons maintenant la fonction $H^\infty : \mathcal{D}^{K+1} \times \mathbb{R}^{K+1} \rightarrow \mathbb{R}^{K+1}$, transformée de Fenchel–Legendre de G^∞ :

$$\forall y \in \mathcal{D}^{K+1} \quad \forall u \in \mathbb{R}^{K+1} \quad H^\infty(y; u) = \sup_{z \in \mathbb{R}^{K+1}} (u \cdot z - G^\infty(y; z)).$$

Remarque 22.3.7. D'après le lemme 22.2.8, pour tout $y \in \text{int}(\mathcal{D}^{K+1})$, la fonction $H^\infty(y; u)$ est finie pour $u \in \text{Conv}(\mathcal{V})$ et infinie ailleurs. Si de plus, $u \in \text{int}(\text{Conv}(\mathcal{V}))$, la fonction $H^\infty(y; u)$ est analytique en u .

Le comportement de la fonction H^∞ pour y sur le bord de \mathcal{D}^{K+1} est plus délicat ; heureusement, nous ne serons pas concernés par ce cas.

Corollaire 22.3.8. Soit $\rho > 0$. La fonction $H^\infty(y; u)$ est uniformément continue en $y \in \mathcal{D}_\rho^{K+1}$, uniformément sur $u \in \mathbb{R}^{K+1}$.

Démonstration. Soit $\varepsilon > 0$. Par le lemme précédent, il existe $\delta > 0$ tel que pour tout $y, y' \in \mathcal{D}_\rho^{K+1}$ et pour tout $z \in \mathbb{R}^{K+1}$ nous avons :

$$|y - y'| < \delta \quad \implies \quad |G^\infty(y; z) - G^\infty(y'; z)| < \varepsilon.$$

Alors, pour tout $y, y' \in \mathcal{D}_\rho^{K+1}$ et pour tout $u \in \mathbb{R}^{K+1}$, si $|y - y'| < \delta$,

$$\begin{aligned} H^\infty(y; u) &= \sup_{z \in \mathbb{R}^{K+1}} (uz - G^\infty(y; z)) \\ &> \sup_{z \in \mathbb{R}^{K+1}} (uz - G^\infty(y'; z)) - \varepsilon = H^\infty(y'; u) - \varepsilon. \end{aligned}$$

Nous échangeons y et y' dans l'inégalité précédente et nous obtenons le résultat voulu. \square

22.3.4 La fonctionnelle d'action

Dans cette section, nous donnons quelques propriétés de la fonctionnelle d'action I_T . Définissons d'abord la fonction $\overline{G}^\infty : \mathbb{R}^{K+1} \rightarrow \mathbb{R}$ par :

$$\forall z \in \mathbb{R}^{K+1} \quad \overline{G}^\infty(z) = 0 \vee \sup_{y \in \mathcal{D}_\rho^{K+1}} G^\infty(y; z).$$

La fonction \overline{G}^∞ est construite de sorte que $\overline{G}^\infty \geq 0$ pour tout $z \in \mathbb{R}^{K+1}$, et

$$\forall y \in \mathcal{D}_\rho^{K+1} \quad \forall z \in \mathbb{R}^{K+1} \quad G^\infty(y; z) \leq \overline{G}^\infty(z).$$

De plus, la fonction \overline{G}^∞ est convexe et finie. Nous pouvons donc définir sa transformée de Fenchel–Legendre :

$$\forall u \in \mathbb{R}^{K+1} \quad \underline{H}^\infty(u) = \sup_{z \in \mathbb{R}^{K+1}} (u \cdot z - \overline{G}^\infty(z)).$$

La fonction \underline{H} est aussi convexe et positive, elle vérifie de plus :

$$\forall y \in \mathcal{D}_\rho^{K+1} \quad \forall u \in \mathbb{R}^{K+1} \quad \underline{H}^\infty(u) \leq H^\infty(y; u).$$

En particulier, la fonction \underline{H}^∞ est finie sur l'ensemble $\text{Conv}(\mathcal{V})$.

Lemme 22.3.9. Soit $T > 0$. La fonctionnelle d'action I_T vérifie les propriétés suivantes :

- Pour tout $\eta \geq 0$, l'ensemble des fonctions $\phi : [0, T] \rightarrow \mathcal{D}_\rho^{K+1}$ vérifiant $I_T(\phi) \leq \eta$ est absolument équicontinu.

• La fonctionnelle d'action est semicontinue inférieurement vis-à-vis de la convergence uniforme.

• Pour tout $\eta, \rho \geq 0$, l'ensemble

$$\Phi_T(\eta) = \bigcup_{y \in \mathcal{D}_\rho^{k+1}} \Phi_{y,T}(\eta) = \left\{ \phi : [0, T] \longrightarrow \mathcal{D}_\rho^{k+1} : I_T(\phi) \leq \eta \right\}$$

est compact vis-à-vis de la convergence uniforme.

Démonstration. Les fonctions $\phi : [0, T] \longrightarrow \mathcal{D}_\rho^{k+1}$ vérifiant $I_T(\phi) \leq \eta$ sont absolument continues par définition de I_T . Il suffit donc de montrer que les dérivées de ces fonctions sont uniformément intégrables. En effet, pour tout ensemble fini d'intervalles disjoints $[t_1, t'_1], \dots, [t_n, t'_n]$ dans $[0, T]$, grâce à la continuité absolue de ϕ , nous avons

$$\sum_{1 \leq i \leq n} |\phi(t'_i) - \phi(t_i)| \leq \sum_{1 \leq i \leq n} \int_{t_i}^{t'_i} |\phi'(s)| ds.$$

Or, par uniforme intégrabilité, pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que si $\sum (t'_i - t_i) < \delta$, alors cette dernière quantité est inférieure à ε , uniformément sur ϕ . Pour montrer l'uniforme intégrabilité des dérivées, nous remarquons que, d'après la proposition 22.2.1, la fonction \underline{H}^∞ vérifie

$$\lim_{|u| \rightarrow \infty} \frac{\underline{H}^\infty(u)}{|u|} = \infty.$$

Nous avons aussi, par définition de \underline{H}^∞ ,

$$\int_0^T \underline{H}^\infty(\phi'(s)) ds \leq \int_0^T \underline{H}^\infty(\phi(s); \phi'(s)) ds \leq \eta.$$

Soit $\varepsilon > 0$ et choisissons $R > 0$ assez grand pour que, pour tout $u \in \mathbb{R}^{k+1}$ avec $|u| > R$,

$$\frac{|u|}{\underline{H}^\infty(u)} < \frac{\varepsilon}{2\eta}.$$

Alors,

$$\begin{aligned} \sum_{1 \leq i \leq n} \int_{t_i}^{t'_i} |\phi'(s)| ds &\leq \sum_{1 \leq i \leq n} \int_{t_i}^{t'_i} |\phi'(s)| \mathbf{1}_{\{|\phi'(s)| \leq R\}} ds + \\ &\quad \sum_{1 \leq i \leq n} \int_{t_i}^{t'_i} \frac{|\phi'(s)|}{\underline{H}^\infty(\phi'(s))} \underline{H}^\infty(\phi'(s)) \mathbf{1}_{\{|\phi'(s)| > R\}} ds \leq R \sum_{1 \leq i \leq n} (t'_i - t_i) + \frac{\varepsilon}{2}. \end{aligned}$$

Nous concluons en prenant $\delta = \varepsilon/2R$.

Nous démontrons maintenant la semicontinuité inférieure de la fonctionnelle d'action. Soit $\phi_n : [0, T] \rightarrow \mathcal{D}_\rho^{K+1}$ une suite de fonctions convergeant uniformément vers $\phi : [0, T] \rightarrow \mathcal{D}_\rho^{K+1}$. Nous allons démontrer que :

$$\liminf_{n \rightarrow \infty} I_T(\phi_n) \geq I_T(\phi).$$

Quitte à extraire une sous-suite, il suffit de considérer le cas où $I_T(\phi_n)$ converge et

$$\lim_{n \rightarrow \infty} I_T(\phi_n) = \eta < \infty.$$

Cette convergence entraîne l'existence d'une constante positive C telle que

$$\forall n \geq 1 \quad I_T(\phi_n) \leq C.$$

La suite de fonctions (ϕ_n) est donc équicontinue. Nous en déduisons la continuité absolue de la fonction ϕ . En effet, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que si $[t_1, t'_1], \dots, [t_k, t'_k]$ sont des sous-intervalles disjoints de $[0, T]$ tels que $\sum_{1 \leq i \leq k} |t'_i - t_i| < \delta$, alors pour tout $n \geq 0$,

$$\sum_{1 \leq i \leq k} |\phi_n(t'_i) - \phi_n(t_i)| < \varepsilon.$$

Un passage à la limite dans cette inégalité permet de conclure la continuité absolue de la fonction ϕ . Soit maintenant $\varepsilon > 0$ quelconque, et choisissons $h, \delta > 0$ tels que :

$$\begin{aligned} \forall y, y' \in \mathcal{D}_\rho^{K+1} \quad |y - y'| < \delta &\implies |H^\infty(y; u) - H^\infty(y'; u)| < \varepsilon, \\ \forall s, t \in [0, T] \quad |s - t| < h &\implies |\phi_n(s) - \phi_n(t)| \vee |\phi(s) - \phi(t)| < \delta. \end{aligned}$$

Remarquons que le choix de δ est uniforme en u grâce au corollaire 22.3.8, et le choix de h est uniforme en n grâce à l'équicontinuité de la suite (ϕ_n) . Soit maintenant $0 = t_0 < t_1 < \dots < t_k = T$ une partition de l'intervalle $[0, T]$ de pas maximal plus petit que h . Alors, pour tout $n \geq 1$,

$$I_T(\phi_n) = \int_0^T H^\infty(\phi_n(t); \phi'_n(t)) dt \geq \sum_{i=0}^{k-1} \int_{t_i}^{t_{i+1}} H^\infty(\phi_n(t_i); \phi'_n(t)) dt - \varepsilon T.$$

Or, par l'inégalité de Jensen,

$$\sum_{i=0}^{k-1} \int_{t_i}^{t_{i+1}} H^\infty(\phi_n(t_i); \phi'_n(t)) dt \geq \sum_{i=0}^{k-1} (t_{i+1} - t_i) H^\infty\left(\phi_n(t_i); \frac{\phi_n(t_{i+1}) - \phi_n(t_i)}{t_{i+1} - t_i}\right).$$

Nous faisons tendre n vers l'infini, et nous obtenons, grâce à la semicontinuité inférieure de H^∞ :

$$\eta + \varepsilon T \geq \sum_{i=0}^{k-1} (t_{i+1} - t_i) H^\infty \left(\phi(t_i); \frac{\phi(t_{i+1}) - \phi(t_i)}{t_{i+1} - t_i} \right).$$

Nous utilisons encore la continuité uniforme de H^∞ dans le premier argument, pour déduire que cette dernière somme est minorée par :

$$\sum_{i=0}^{k-1} \int_{t_i}^{t_{i+1}} H^\infty \left(\phi(t); \frac{\phi(t_{i+1}) - \phi(t_i)}{t_{i+1} - t_i} \right) dt - \varepsilon T.$$

Soit $\psi : [0, T] \rightarrow \mathcal{D}_\rho^{K+1}$ la fonction constante par morceaux égale à $(\phi(t_{i+1}) - \phi(t_i))/(t_{i+1} - t_i)$ dans l'intervalle $[t_i, t_{i+1}[$. Nous avons obtenu la relation

$$\int_0^T H^\infty(\phi(t); \psi(t)) dt \leq \eta + 2\varepsilon T.$$

Quand h tend vers 0, $\psi(t)$ converge vers $\phi'(t)$ pour presque tout $t \in [0, T]$. Puisque la fonction H^∞ est semicontinue inférieurement, nous avons, presque pour tout t ,

$$H^\infty(\phi(t); \phi'(t)) \leq \liminf_{h \rightarrow 0} H^\infty(\phi(t); \psi(t)).$$

La fonction H^∞ étant positive, par le lemme de Fatou,

$$I_T(\phi) = \int_0^T H^\infty(\phi(t); \phi'(t)) dt \leq \liminf_{h \rightarrow 0} \int_0^T H^\infty(\phi(t); \psi(t)) dt \leq \eta + 2\varepsilon T.$$

Nous faisons tendre ε vers 0 et nous obtenons le résultat voulu. Finalement, nous montrons la compacité de l'ensemble

$$\Phi_T(\eta) = \bigcup_{y \in \mathcal{D}_\rho^{K+1}} \Phi_{y,T}(\eta) = \left\{ \phi : [0, T] \rightarrow \mathcal{D}_\rho^{K+1} : I_T(\phi) \leq \eta \right\}.$$

La précompacité est obtenue grâce au théorème d'Ascoli–Arzela, vu que l'ensemble est équicontinu. Puis la semicontinuité inférieure de la fonctionnelle d'action entraîne que l'ensemble est fermé. \square

Nous étudions ensuite les trajectoires qui minimisent l'action I_T . Nous définissons la fonction $b : \mathcal{D}^{K+1} \rightarrow \mathbb{R}^{K+1}$ ainsi :

$$\forall 0 \leq k \leq K \quad \forall x \in \mathcal{D}^{K+1}$$

$$b_k(x) = \frac{1}{\lambda} \sum_{0 \leq h \leq k} x_h f(h) e^{-a} \frac{a^{k-h}}{(k-h)!} - \frac{1}{\lambda} x_k \left(\sum_{0 \leq h \leq K} x_h (f(h) - 1) + 1 \right).$$

Nous montrons ensuite que la fonction I_T s'annule pour les trajectoires ϕ qui sont solutions du système d'équations différentielles $x'(t) = b(x(t))$.

Corollaire 22.3.10. *La fonctionnelle I_T satisfait les propriétés suivantes :*

- La fonctionnelle $I_T(\cdot)$ atteint son infimum sur tout ensemble fermé de trajectoires $\phi : [0, T] \rightarrow \mathcal{D}_\rho^{K+1}$.
- Nous avons $I_T(\phi) = 0$ si et seulement si ϕ est une solution du système $x'(t) = b(x(t))$.

Démonstration. La première affirmation découle du lemme précédent ; en effet considérons un ensemble fermé F de trajectoires $\phi : [0, T] \rightarrow \mathcal{D}_\rho^{K+1}$, et supposons que $I^* = \inf \{ I_T(\phi) : \phi \in F \} < +\infty$ (autrement l'affirmation est triviale). Il existe donc une suite $(\phi_n)_{n \geq 1}$ de trajectoires de F telle que $I^n = I_T(\phi_n)$ décroît vers I^* . Quitte à extraire une sous-suite, la suite $(\phi_n)_{n \geq 1}$ converge vers une trajectoire $\phi^* \in F$ telle que $I_T(\phi^*) = I^*$.

Démontrons maintenant la deuxième affirmation. Rappelons que la fonction $H(y; u)$ est la transformée de Fenchel–Legendre de la log–Laplace $G(y; z)$ de la loi

$$\mu = \sum_{v \in \mathcal{V}} C_y^\infty(v) \delta_v + \left(1 - \sum_{v \in \mathcal{V}} C_y^\infty(v) \right) \delta_0.$$

Par le lemme 22.2.3, nous avons que $H(y; u) = 0$ si et seulement si $u = E(X)$, où X est un vecteur aléatoire de loi μ . Calculons donc cette espérance. Les coefficients $C_y^\infty(v)$ sont donnés par : pour tout $y \in \mathcal{D}^{K+1}$ et $0 \leq i, j \leq K$,

$$\begin{aligned} C_y^\infty(-e_i) &= \frac{1}{\lambda} y_i \left(\sum_{0 \leq h \leq K} y_h f_H(h) \left(1 - \sum_{h \leq l \leq K} e^{-a} \frac{a^{l-h}}{(l-h)!} \right) + 1 - \sum_{0 \leq h \leq K} y_h \right), \\ C_y^\infty(e_j) &= \frac{1}{\lambda} \left(1 - \sum_{0 \leq l \leq K} y_l \right) \sum_{0 \leq h \leq j} y_h f_H(h) e^{-a} \frac{a^{j-h}}{(j-h)!}, \\ C_y^\infty(-e_i + e_j) &= \frac{1}{\lambda} y_i \sum_{0 \leq h \leq j} y_h f_H(h) e^{-a} \frac{a^{j-h}}{(j-h)!}. \end{aligned}$$

Nous avons donc, pour $k \in \{0, \dots, K\}$,

$$\begin{aligned} E(X_k) &= C_y^\infty(e_k) + \sum_{i:i \neq k} C_y^\infty(-e_i + e_k) - C_y^\infty(-e_k) - \sum_{j:j \neq k} C_y^\infty(-e_k + e_j) = \\ &= \frac{1}{\lambda}(1 - y_k) \sum_{0 \leq h \leq j} y_h f_H(h) e^{-a} \frac{a^{j-h}}{(j-h)!} - \frac{1}{\lambda} y_k \left(\sum_{0 \leq h \leq K} y_h f_H(h) + 1 \right) = b_k(y). \end{aligned}$$

Nous en déduisons que si ϕ est une solutions du système $x'(t) = b(x(t))$, alors $I_T(\phi) = 0$. Inversement, si $I_T(\phi) = 0$, alors $\phi'(t) = b(\phi(t))$ pour presque tout $t \in [0, T]$. Comme de plus ϕ est absolument continue, nous avons, pour tout $t \in [0, T]$,

$$\lim_{h \rightarrow 0} \frac{\phi(t+h) - \phi(t)}{h} = \lim_{h \rightarrow 0} \frac{1}{h} \int_t^{t+h} b(\phi(s)) ds = b(\phi(t)).$$

Nous voyons donc que ϕ est dérivable pour tout $t \in [0, T]$ et qu'elle est solution de l'équation $\phi'(t) = b(\phi(t))$. \square

22.4 Borne inférieure du PGD

Soient $\delta, \gamma, \rho, \eta, T > 0$ avec $\rho > \delta$. Notre but est de démontrer que, asymptotiquement,

$$P\left(d_T(Y, \phi) < \delta \mid Y(0) = y^0\right) \geq \exp\left(-m(I_{0,T}(\phi) + \gamma)\right),$$

uniformément sur $y^0 \in \mathcal{D}_\rho^{K+1}$ et $\phi \in \Phi_{y^0, T}(\eta)$. Fixons donc $\delta, \gamma, \rho, \eta, T > 0$. Nous définissons la distance d_T^τ entre deux trajectoires $[0, T] \rightarrow \mathcal{D}^{K+1}$ par :

$$d_T^\tau(\phi, \psi) = \sup_{t=0, \tau, \dots, \lfloor T/\tau \rfloor \tau} |\phi(t) - \psi(t)|.$$

Pour tout $x^0 \in \mathcal{S}^{\ell+1}$, nous notons par P_{x^0} la loi du processus $(X(t))_{t \geq 0}$ sachant $X(0) = x^0$. Également, nous notons par E_{x^0} l'espérance associée. Soient $y^0 \in \mathcal{D}_\rho^{K+1}$, $x^0 \in \mathcal{S}^{\ell+1}$ tels que $\pi_K(x^0) = y^0$ et $\phi : [0, T] \rightarrow \mathcal{D}_\rho^{K+1}$ une trajectoire vérifiant $\phi(0) = y^0$ et $I_{0,T}(\phi) \leq s_0$. Comme ϕ est absolument continue, Pour m assez grand, nous avons

$$P_{x^0}(d_T(Y, \phi) < \delta) \geq P_{x^0}(d_T^\tau(Y, \phi) < \delta/2).$$

De plus, d'après le lemme 22.3.9, l'inégalité est vraie uniformément sur $x^0 \in \pi_K^{-1}(\mathcal{D}_\rho^{K+1})$ et $\phi \in \Phi_{y^0;T}(\eta)$. Nous allons donc démontrer que, asymptotiquement,

$$P_{x^0}(d_T^\tau(Y, \phi) < \delta) \geq e^{-m(I(\phi)+\gamma)},$$

uniformément sur $x^0 \in \pi_K^{-1}(\mathcal{D}_\rho^{K+1})$ et $\phi \in \Phi_{y^0;T}(\eta)$.

Pour alléger les notations, nous enlevons désormais les indices τ et T et nous notons la distance d_T^τ tout simplement par d . Également, nous enlevons l'indice T dans la fonctionnelle d'action. Nous allons d'abord faire la preuve pour des trajectoires ϕ vérifiant l'hypothèse suivante :

Hypothèse (H). Il existe un compact $\mathcal{K}' \subset \text{int}(\text{Conv}(\mathcal{V}))$ tel que $\phi'(t) \in \mathcal{K}'$ pour tout $t \in [0, T]$.

Sous cette hypothèse, la fonction $H^\infty(y; u)$ est dérivable par rapport à u au point $\phi'(t)$ pour tout $y \in \mathcal{D}_\rho^{K+1}$ et $t \in [0, T]$ (remarque 22.3.7). Nous en déduisons l'existence d'un deuxième compact $\mathcal{K} \subset \mathbb{R}^{K+1}$ tel que

$$\forall y \in \mathcal{D}_\rho^{K+1} \quad \forall t \in [0, T] \quad \nabla_u H^\infty(y; \phi'(t)) \in \mathcal{K}.$$

Malheureusement, le processus Y ne reste pas forcément à l'intérieur de l'ensemble \mathcal{D}_ρ^{K+1} , et lorsqu'il touche le bord de \mathcal{D}^{K+1} , la fonction H^∞ peut être infinie. Pour surmonter cette difficulté, nous allons construire un processus auxiliaire, avec la même dynamique que le processus Y dans le triangle réduit \mathcal{D}_ρ^{K+1} , mais cette fois-ci, défini partout dans l'espace \mathbb{R}^{K+1} .

Processus auxiliaire. Nous construisons maintenant un processus auxiliaire à valeurs dans \mathbb{Z}^{K+1}/m , que nous notons par $(\tilde{Y}(t))_{t=0,\tau,\dots}$. Nous supposons, sans perte de généralité, que $\rho - \delta > 0$ et nous prenons ρ tel que $0 < \rho < (\rho - \delta)/2$. Sur l'ensemble $(\mathbb{Z}^{K+1}/m) \cap \mathcal{D}_\rho^{K+1}$, le processus \tilde{Y} a la même dynamique que le processus Y , et en dehors de cet ensemble, le processus \tilde{Y} évolue selon une marche aléatoire symétrique standard. Plus précisément, pour tout $t \geq 0$ et pour tout $v \in \mathcal{V}$

$$\begin{aligned} P\left(\tilde{Y}(t+\tau) = \tilde{Y} + \frac{v}{m} \mid \mathcal{X}(t)\right) &= C_{\mathcal{X}(t)}(v), & \text{si } \pi_K(\mathcal{X}(t)) \in \mathcal{D}_\rho^{K+1}, \\ P\left(\tilde{Y}(t+\tau) = \tilde{Y} + \frac{v}{m} \mid \mathcal{X}(t)\right) &= \frac{1}{|\mathcal{V}|+1}, & \text{sinon.} \end{aligned}$$

Nous définissons les fonctions $\tilde{G}, \tilde{H}, \tilde{G}^\infty, \tilde{H}^\infty$ associées au processus \tilde{Y} , qui sont les analogues des fonctions G, H, G^∞, H^∞ associées au processus Y . Remarquons

que la proposition 22.3.2 et le corollaire 22.3.5 restent vrais pour les fonctions $\tilde{G}, \tilde{G}^\infty$. Évidemment, pour tout $\phi \in \Phi_{y^0, T}(\eta)$:

$$P_{x^0}(d(Y, \phi) < \delta) = P_{x^0}(d(\tilde{Y}, \phi) < \delta),$$

reste vrai jusqu'au temps de sortie inclus. Nous pouvons donc travailler avec le processus \tilde{Y} au lieu du processus Y pour prouver notre estimée.

Nous posons, pour tout $t \in [0, T]$ et $y \in \mathbb{R}^{k+1}$:

$$z(t, y) = m \nabla_u \tilde{H}^\infty(y; \phi'(t)).$$

Puisque ϕ vérifie l'hypothèse (\mathcal{H}) , nous avons bien

$$\forall y \in \mathbb{R}^{k+1} \quad \forall t \in [0, T] \quad \frac{z(t, y)}{m} \in \mathcal{X}.$$

Introduisons la transformée de Cramér généralisée associée à z , comme dans la section 22.2.7.

$$\pi(0, T) = \exp \left(\sum_{s=0}^{T-\tau} z(s, \tilde{Y}(s)) (\tilde{Y}(s+\tau) - \tilde{Y}(s)) - \sum_{s=0}^{T-\tau} \tilde{G} \left(\mathcal{X}(s); \frac{z(s, \tilde{Y}(s))}{m} \right) \right).$$

La mesure de probabilité $P_{x^0}^z$ est absolument continue vis-à-vis de P_{x^0} , et a une densité strictement positive $\pi(0, T)$. Nous en déduisons que la mesure de probabilité P_{x^0} est absolument continue vis-à-vis de $P_{x^0}^z$ avec densité $\pi(0, T)^{-1}$. D'où

$$P_{x^0}(d(\tilde{Y}, \phi) < \delta) = E_{x^0}^z \left(\mathbf{1}_{\{d(\tilde{Y}, \phi) < \delta\}} \pi(0, T)^{-1} \right).$$

Nous allons d'abord minorer la probabilité de l'événement $\{d(\tilde{Y}, \phi) < \delta\}$, et nous allons ensuite majorer la densité $\pi(0, T)$. Nous avons

$$P_{x^0}^z(d(\tilde{Y}, \phi) < \delta) \geq 1 - P_{x^0}^z(d(\tilde{Y}, \phi) \geq \delta).$$

Définissons les processus Y^z et $Y^{\infty, z}$ ainsi : pour tout $t = 0, \tau, \dots$,

$$\begin{aligned} Y^z(t) &= \tilde{Y}(0) + \frac{1}{m} \sum_{s=0}^{t-\tau} \nabla_z \tilde{G} \left(\mathcal{X}(s); \frac{z(s, \tilde{Y}(s))}{m} \right), \\ Y^{\infty, z}(t) &= \tilde{Y}(0) + \frac{1}{m} \sum_{s=0}^{t-\tau} \nabla_z \tilde{G}^\infty \left(Y(s); \frac{z(s, \tilde{Y}(s))}{m} \right). \end{aligned}$$

Nous avons

$$\begin{aligned} P_{x^0}^z(d(\tilde{Y}, \phi) \geq \delta) &\leq P_{x^0}^z(d(\tilde{Y}, Y^z) \geq \delta/3) + \\ &\quad P_{x^0}^z(d(Y^z, Y^{\infty, z}) \geq \delta/3) + P_{x^0}^z(d(Y^{\infty, z}, \phi) \geq \delta/3). \end{aligned}$$

Il s'agit maintenant de contrôler les trois termes de droite de cette inégalité.

Lemme 22.4.1. *Le processus $(\tilde{Y}(t) - Y^z(t))_{t=0, \tau, \dots}$ est une martingale vis-à-vis de la tribu $(\mathcal{F}_t)_{t=0, \tau, \dots}$ et de la mesure de probabilité $P_{x^0}^z$.*

Démonstration. Il suffit de montrer que, pour tout $t = 0, \tau, \dots$,

$$E_{x^0}^z \left(\tilde{Y}(t + \tau) - \tilde{Y}(t) - \frac{1}{m} \nabla_z \tilde{G} \left(\mathcal{X}(t); \frac{z(t, \tilde{Y}(t))}{m} \right) \middle| \mathcal{F}_t \right) = 0.$$

Puisque $\pi(0, t)$ est \mathcal{F}_t -mesurable, il suffit de montrer que

$$E_{x^0}^z \left((\tilde{Y}(t + \tau) - \tilde{Y}(t)) \pi(t, t + 1) \middle| \mathcal{F}_t \right) = \frac{1}{m} \nabla_z \tilde{G} \left(\mathcal{X}(t); \frac{z(t, \tilde{Y}(t))}{m} \right).$$

Or, sachant \mathcal{F}_t , la log-Laplace de l'incrément $\tilde{Y}(t + \tau) - \tilde{Y}(t)$ est donnée par $\tilde{G}(\mathcal{X}(t); \cdot/m)$. Par le lemme 22.2.9, nous obtenons le résultat voulu. \square

Par l'inégalité de Doob,

$$\begin{aligned} P_{x^0}^z(d(\tilde{Y}, Y^z) \geq \delta/3) &\leq \left(\frac{3}{\delta}\right)^2 E_{x^0}^z(|\tilde{Y}(T) - Y^z(T)|^2) \\ &\leq \left(\frac{3}{\delta}\right)^2 \sum_{h=0}^K E_{x^0}^z((\tilde{Y}_h(T) - Y_h^z(T))^2). \end{aligned}$$

Pour chaque h dans $\{0, \dots, K\}$, nous décomposons le carré dans l'espérance comme suit :

$$\begin{aligned} |\tilde{Y}_h(T) - Y_h^z(T)|^2 &= \sum_{s=0}^{T-\tau} \left| \tilde{Y}_h(s + \tau) - \tilde{Y}_h(s) - \frac{1}{m} \frac{\partial \tilde{G}}{\partial z_h} \left(\mathcal{X}(s); \frac{z(s, \tilde{Y}(s))}{m} \right) \right|^2 \\ &\quad + 2 \sum_{0 \leq s < t \leq T-\tau} \left(\tilde{Y}_h(s + \tau) - \tilde{Y}_h(s) - \frac{1}{m} \frac{\partial \tilde{G}}{\partial z_h} \left(\mathcal{X}(s); \frac{z(s, \tilde{Y}(s))}{m} \right) \right) \\ &\quad \times \left(\tilde{Y}_h(t + \tau) - \tilde{Y}_h(t) - \frac{1}{m} \frac{\partial \tilde{G}}{\partial z_h} \left(\mathcal{X}(t); \frac{z(t, \tilde{Y}(t))}{m} \right) \right). \end{aligned}$$

Nous prenons les espérances conditionnelles par rapport à \mathcal{F}_s dans l'expression précédente. Par le lemme précédent, les termes croisés sont d'espérance nulle. Pour la somme des carrés, nous raisonnons comme dans le lemme précédent, et nous appliquons à nouveau le lemme 22.2.9, pour déduire que

$$\begin{aligned} E_{x_0}^z \left(\left(\tilde{Y}_h(s + \tau) - \tilde{Y}_h(s) - \frac{1}{m} \frac{\partial \tilde{G}}{\partial z_h} \left(\mathcal{X}(s); \frac{z(s, \tilde{Y}(s))}{m} \right) \tau \right)^2 \middle| \mathcal{F}_s \right) \\ = \frac{1}{m^2} \frac{\partial^2 \tilde{G}}{\partial z_h^2} \left(\mathcal{X}(s); \frac{z(s, \tilde{Y}(s))}{m} \right). \end{aligned}$$

Par l'hypothèse (\mathcal{H}) , $z(s, \tilde{Y}(s))/m$ reste dans un compact \mathcal{K} pour tout temps $s = 0, \tau, \dots, T - \tau$. Nous pouvons donc utiliser le corollaire 22.3.5 pour en déduire qu'il existe une constante $C > 0$ telle que

$$P_{x_0}^z \left(d(\tilde{Y}, Y^z) \geq \delta/3 \right) \leq \left(\frac{3}{\delta} \right)^2 (K + 1) T \frac{C}{m}.$$

Cette quantité tend vers 0 quand m tend vers l'infini. Nous majorons maintenant la probabilité de l'événement

$$d(Y^z, Y^{\infty, z}) \geq \delta'/3.$$

Nous avons, pour tout $t \geq 0$,

$$\begin{aligned} |Y^z(t) - Y^{\infty, z}(t)| \leq \\ \sum_{s=0}^{t-\tau} \left| \nabla_z \tilde{G} \left(\mathcal{X}(s); \frac{z(s, \tilde{Y}(s))}{m} \right) - \nabla_z \tilde{G}^\infty \left(Y(s); \frac{z(s, \tilde{Y}(s))}{m} \right) \right| \tau. \end{aligned}$$

Vu que $Y(s) = \pi_{\mathcal{K}}(\mathcal{X}(s))$ pour tout $s = 0, \tau, \dots$ et que $z(s, \tilde{Y}(s))/m \in \mathcal{K}$ pour tout $s = 0, \tau, \dots, T$, nous pouvons utiliser la proposition 22.3.3 pour en déduire que, asymptotiquement,

$$|Y^z(t) - Y^{\infty, z}(t)| \leq \delta/3.$$

Finalement, grâce au choix de z , et d'après la proposition 22.2.1, nous voyons que

$$Y^{\infty, z}(t) = \tilde{Y}(0) + \sum_{s=0}^{t-\tau} \nabla_z \tilde{G}^\infty \left(Y(s); \frac{z(s, \tilde{Y}(s))}{m} \right) \tau = \phi(0) + \sum_{s=0}^{t-\tau} \phi'(s) \tau.$$

La continuité absolue de ϕ entraîne que, pour m assez grand,

$$|Y^{\infty, z}(t) - \phi(t)| < \delta/3.$$

Nous concluons que, asymptotiquement,

$$P_{x_0}^z(d(\tilde{Y}, \phi) < \delta) \rightarrow 1.$$

Pour compléter la démonstration de la borne inférieure du PGD, nous allons majorer la densité $\pi(0, T)$. Il suffit de majorer la densité $\pi(0, T)$ sur l'événement $\{d(\tilde{Y}, \phi) < \delta\}$, et justement sur cet événement, $Y = \tilde{Y}$. Nous pouvons donc de nouveau travailler avec le processus original Y . Nous avons

$$\begin{aligned} \pi(0, T) = \\ \exp \left(\sum_{t=0}^{T-\tau} z(t, Y(t)) (Y(t+\tau) - Y(t)) - m \sum_{t=0}^{T-\tau} G \left(\mathcal{X}(t); \frac{z(t, Y(t))}{m} \right) \tau \right). \end{aligned}$$

L'argument dans l'exponentielle est égal à

$$\begin{aligned} & \sum_{t=0}^{T-\tau} z(t, Y(t)) \left((Y(t+\tau) - Y(t)) - \nabla_z G \left(\mathcal{X}(t); \frac{z(t, Y(t))}{m} \right) \tau \right) \\ & + \sum_{t=0}^{T-\tau} z(t, Y(t)) \left(\nabla_z G \left(\mathcal{X}(t); \frac{z(t, Y(t))}{m} \right) \tau - \nabla_z G^\infty \left(Y(t); \frac{z(t, Y(t))}{m} \right) \tau \right) \\ & + \sum_{t=0}^{T-\tau} \left(z(t, Y(t)) \nabla_z G^\infty \left(Y(t); \frac{z(t, Y(t))}{m} \right) \tau - m G^\infty \left(Y(t); \frac{z(t, Y(t))}{m} \right) \tau \right) \\ & + \sum_{t=0}^{T-\tau} \left(m G^\infty \left(Y(t); \frac{z(t, Y(t))}{m} \right) \tau - m G \left(\mathcal{X}(t); \frac{z(t, Y(t))}{m} \right) \tau \right). \end{aligned}$$

Nous allons majorer une par une chacune de ces quatre lignes.

Ligne 1. Pour tout $t = 0, \tau, \dots$, nous posons

$$\eta(t) = \sum_{t=0}^{T-\tau} z(t, Y(t)) \left((Y(t+\tau) - Y(t)) - \nabla_z G \left(\mathcal{X}(t); \frac{z(t, Y(t))}{m} \right) \tau \right).$$

Lemme 22.4.2. *Le processus $(\eta(t))_{t=0, \tau, \dots}$ est une martingale vis-à-vis de $(\mathcal{F}_t)_{t=0, \tau, \dots}$ et de la mesure de probabilité $P_{x_0}^z$.*

Démonstration. Il suffit de montrer que, pour tout $t = 0, \tau, \dots$, nous avons

$$\mathbb{E}_{x_0}^z \left(z(t, Y(t)) \left((Y(t + \tau) - Y(t)) - \nabla_z G \left(\mathcal{X}(t); \frac{z(t, Y(t))}{m} \right) \tau \right) \middle| \mathcal{F}_t \right) = 0.$$

Puisque $z(t, Y(t))$ est mesurable par rapport à \mathcal{F}_t , nous pouvons le sortir de l'espérance conditionnelle. Comme nous l'avons déjà vu dans la démonstration du lemme 22.4.1, l'espérance conditionnelle qui reste est nulle. \square

Par le théorème de Kolmogorov pour les martingales, pour tout $\varepsilon > 0$,

$$\mathbb{P}_{x_0}^z (|\eta(t)| \geq \varepsilon) \leq \varepsilon^{-2} \mathbb{E}_{x_0}^z (|\eta(T)|^2).$$

Nous décomposons le carré dans l'espérance comme suit :

$$\begin{aligned} |\eta(T)|^2 &= \sum_{s=0}^{T-\tau} z(s, Y(s))^2 \left(Y(s + \tau) - Y(s) - \nabla_z G \left(\mathcal{X}(s); \frac{z(s, Y(s))}{m} \right) \tau \right)^2 \\ &+ \sum_{0 \leq s < t \leq T-\tau} z(s, Y(s)) \left(Y(s + \tau) - Y(s) - \nabla_z G \left(\mathcal{X}(s); \frac{z(s, Y(s))}{m} \right) \tau \right) \\ &\quad \times z(t, Y(t)) \left(Y(t + \tau) - Y(t) - \nabla_z G \left(\mathcal{X}(t); \frac{z(t, Y(t))}{m} \right) \tau \right). \end{aligned}$$

Nous prenons les espérances conditionnelles par rapport à \mathcal{F}_s dans cette expression. Par le lemme précédent, les termes croisés sont d'espérance nulle. Chaque terme dans la somme des carrés s'écrit

$$\begin{aligned} &\sum_{i,j=0}^K \left(Y_i(s + \tau) - Y_i(s) - \frac{\partial G}{\partial z_i} \left(\mathcal{X}(s); \frac{z(s; Y(s))}{m} \right) \tau \right) \\ &\quad \times \left(Y_j(s + \tau) - Y_j(s) - \frac{\partial G}{\partial z_j} \left(\mathcal{X}(s); \frac{z(s; Y(s))}{m} \right) \tau \right) z_i(s, Y(s)) z_j(s, Y(s)). \end{aligned}$$

Or, $z(s, Y(s))$ est mesurable par rapport à \mathcal{F}_s , et sachant \mathcal{F}_s , la log-Laplace de l'incrément $Y(s + \tau) - Y(s)$ est $G(\mathcal{X}(s); \cdot/m)$. Donc, d'après le lemme 22.2.9, en prenant l'espérance conditionnelle par rapport à \mathcal{F}_s dans l'expression ci-dessus, nous obtenons

$$\sum_{i,j=0}^K \frac{\partial^2 G}{\partial z_i \partial z_j} \left(\mathcal{X}(s); \frac{z(s, Y(s))}{m} \right) \frac{z_i(s, Y(s))}{m} \frac{z_j(s, Y(s))}{m}.$$

Par l'hypothèse (\mathcal{H}) , $z(s, Y(s))/m$ reste dans le compact \mathcal{K} pour tout temps $s = 0, \tau, \dots, T$. Nous pouvons donc utiliser le corollaire 22.3.5 pour en déduire que chaque terme de la somme ci-dessus est borné par une constante positive C . Cela entraîne que

$$P_{x_0}^z \left(|\eta(T)| \geq \varepsilon \right) \leq \varepsilon^{-2} (K+1)^2 TC.$$

En prenant $\varepsilon = m\gamma/4$, nous concluons que avec probabilité qui tend vers 1,

$$|\eta(T)| < m\gamma/4.$$

Ligne 2. Vu que $z(s, Y(s))/m$ reste dans le compact \mathcal{K} pour tout $s = 0, \tau, \dots, T$, et qu'en plus $\nabla_z G(x; z) \rightarrow \nabla_z G^\infty(\pi_{\mathcal{K}}(x); z)$ uniformément sur $\mathcal{S}^{\ell+1} \times \mathcal{K}$, pour tout $\varepsilon > 0$, asymptotiquement,

$$\begin{aligned} \sum_{s=0}^{T-\tau} z(s, Y(s)) \left(\nabla_z G \left(\mathcal{X}(s); \frac{z(s, Y(s))}{m} \right) - \nabla_z G^\infty \left(Y(s); \frac{z(s, Y(s))}{m} \right) \right) \tau \\ \leq m \left(\sup_{z \in \mathcal{K}} |z| \right) T \varepsilon, \end{aligned}$$

avec probabilité 1. En prenant ε assez petit, nous concluons que, avec probabilité 1, asymptotiquement, la ligne 2 est majorée par $m\gamma/4$.

Ligne 3. Rappelons que notre choix de z a été : pour tout $t \in [0, T]$ et $y \in \mathbb{R}^{K+1}$,

$$z(t, y) = m \nabla_u \tilde{H}^\infty(y; \phi'(t)).$$

Or, pour tout $y \in \mathcal{D}_\rho^{K+1}$, nous avons $\tilde{H}^\infty(y; \cdot) = H^\infty(y; \cdot)$. Puisque sur l'événement $\{d(Y(s), \phi(s)) < \delta\}$, nous avons aussi $Y(s) \in \mathcal{D}_\rho^{K+1}$, nous concluons, d'après la proposition 22.2.1, que

$$\begin{aligned} \sum_{t=0}^{T-\tau} \left(z(t, Y(t)) \nabla_z G^\infty \left(Y(t); \frac{z(t, Y(t))}{m} \right) \tau - m G^\infty \left(Y(t); \frac{z(t, Y(t))}{m} \right) \tau \right) \\ = m \sum_{t=0}^{T-\tau} H^\infty(Y(s); \phi'(s)) \tau. \end{aligned}$$

Or, sur l'événement $d(Y(s), \phi(s)) < \delta$, en prenant δ assez petit, par la continuité uniforme de H^∞ dans le premier argument, nous avons

$$\sum_{t=0}^{T-\tau} H^\infty(Y(s); \phi'(s)) \tau \leq \sum_{t=0}^{T-\tau} H^\infty(\phi(s); \phi'(s)) \tau + \gamma/8.$$

Puisque ϕ est absolument continue, asymptotiquement, nous avons que l'expression précédente est majorée par

$$\int_0^T H^\infty(\phi(s); \phi'(s)) ds + \gamma/4.$$

Nous concluons que, asymptotiquement, avec probabilité qui tend vers 1, la ligne 3 est majorée par

$$m \left(\int_0^T H^\infty(\phi(s); \phi'(s)) ds + \gamma/4 \right).$$

Ligne 4. Vu que $z(s, Y(s))/m$ reste dans le compact \mathcal{K} pour tout $s = 0, \tau, \dots, T$, d'après la proposition 22.3.3, pour tout $\varepsilon > 0$, asymptotiquement,

$$\sum_{t=0}^{T-\tau} \left(G^\infty \left(Y(t); \frac{z(t, Y(t))}{m} \right) \tau - G \left(X(t); \frac{z(t, Y(t))}{m} \right) \tau \right) \leq T\varepsilon,$$

avec probabilité 1. Nous concluons que, pour ε assez petit, asymptotiquement, avec probabilité 1, la ligne 4 est majorée par $m\gamma/4$. Ceci achève la démonstration de la borne inférieure du PGD pour des trajectoires ϕ vérifiant l'hypothèse (\mathcal{H}) . Il s'agit maintenant d'étendre le résultat à toute trajectoire ϕ telle que $I(\phi) < \eta$.

Étant donnée une partition $0 = t_0 < t_1 < \dots < t_k = T$ de l'intervalle $[0, T]$, nous considérons la ligne polygonale $L : [0, T] \rightarrow \mathcal{D}_\rho^{K+1}$ passant par les points

$$(t_0, \phi(t_0)), (t_1, \phi(t_1)), \dots, (t_k, \phi(t_k)).$$

Le pas maximal de la partition est la longueur du plus grand intervalle $[t_i, t_{i+1}]$.

Lemme 22.4.3. *Pour tout $\varepsilon > 0$, il existe $h > 0$ tel que, pour tout partition de pas maximal inférieure à h , pour tout ϕ vérifiant $I(\phi) \leq \eta$,*

$$d_T(L, \phi) < \varepsilon, \quad I(L) \leq I(\phi) + \varepsilon.$$

Démonstration. Soit $\varepsilon > 0$. Par équicontinuité (lemma 22.3.9), il existe $h_1 > 0$ tel que, pour toute partition de pas maximal inférieur à h_1 , $d_T(L, \phi) < \varepsilon$. Soient $h_2, \delta' > 0$ tels que :

$$\begin{aligned} \forall y, y' \in \mathcal{D}_\rho^{K+1} \quad \forall u \in \mathbb{R}^{K+1} \\ |y - y'| < \delta' &\implies |H^\infty(y; u) - H^\infty(y'; u)| < \frac{\varepsilon}{2T}, \\ \forall s, t \in [0, T] \quad |s - t| < h_2 &\implies |\phi(s) - \phi(t)| < \delta'. \end{aligned}$$

Remarquons que le choix de δ' est uniforme en u grâce au corollaire 22.3.8, et le choix de h_2 est uniforme en ϕ tel que $I(\phi) < \eta$ par équicontinuité. Par l'inégalité de Jensen, nous avons :

$$\begin{aligned} \sum_{i=0}^{k-1} (t_{i+1} - t_i) H^\infty \left(\phi(t_i); \frac{\phi(t_{i+1}) - \phi(t_i)}{t_{i+1} - t_i} \right) &\leq \\ &\sum_{i=0}^{k-1} \int_{t_i}^{t_{i+1}} H^\infty(\phi(t_i); \phi'(t)) dt \leq \int_0^T H^\infty(\phi(t); \phi'(t)) dt + \frac{\varepsilon}{2}. \end{aligned}$$

Or,

$$\forall i \in \{0, \dots, k-1\} \quad \sup_{t_i \leq t \leq t_{i+1}} |L(t) - L(t_i)| < \delta'.$$

Donc,

$$\begin{aligned} I(L) = \int_0^T H^\infty(L(t); L'(t)) dt &= \sum_{i=0}^{k-1} \int_{t_i}^{t_{i+1}} H^\infty \left(L(t); \frac{\phi(t_{i+1}) - \phi(t_i)}{t_{i+1} - t_i} \right) dt \\ &\leq \sum_{i=0}^{k-1} (t_{i+1} - t_i) H^\infty \left(\phi(t_i); \frac{\phi(t_{i+1}) - \phi(t_i)}{t_{i+1} - t_i} \right) + \frac{\varepsilon}{2}. \end{aligned}$$

Nous combinons ces deux inégalités et nous obtenons le résultat voulu. \square

Soit $\varepsilon > 0$ et soit h associé à ε par le lemme précédent. Considérons les approximations polygonales L des fonctions ϕ telles que $I(\phi) < \eta$, construites à partir des partitions de $[0, T]$ de pas maximal inférieur à h . Puisque $I(L) < \eta + \varepsilon$, en particulier les trajectoires L vérifient $L'(t) \in \text{int}(\text{Conv}(\mathcal{V}))$ (autrement l'action ne serait pas finie). Cependant, les dérivés $L'(t)$ ne restent pas forcément dans un compact $\mathcal{K}' \in \text{int}(\text{Conv}(\mathcal{V}))$. Soit $\kappa > 0$ et considérons les lignes polygonales L_κ définies par

$$\forall t \in [0, T] \quad L_\kappa(t) = \kappa L(0) + (1 - \kappa)L(t).$$

L'ensemble des lignes polygonales L_κ vérifient l'hypothèse (\mathcal{H}) . De plus, pour κ assez petit, $d(L, L_\kappa) < \varepsilon$, uniformément sur l'ensemble des lignes polygonales L en question. En effet, pour tout $t \in [0, T]$,

$$|L(t) - L_\kappa(t)| \leq \kappa |L(0)| + \kappa |L(t)| \leq 2\kappa.$$

Nous estimons maintenant l'action d'une ligne polygonale L_κ . Au vu de la convexité de la fonction H^∞ , nous avons

$$\begin{aligned} I(L_\kappa) &= \int_0^T H^\infty(L_\kappa(t); L'_\kappa(t)) dt = \int_0^T H^\infty(L_\kappa(t); (1-\kappa)L'(t)) dt \\ &\leq \kappa \int_0^T H^\infty(L_\kappa(t); 0) dt + (1-\kappa) \int_0^T H^\infty(L_\kappa(t); L'(t)) dt. \end{aligned}$$

La première intégrale est bornée par

$$\kappa T \sup_{y \in \mathcal{D}_\rho^{k+1}} H^\infty(y; 0).$$

Pour κ assez petit, d'après le lemme précédent, la deuxième intégrale est bornée par

$$I(\phi) + 2\varepsilon.$$

Finalement, puisque L_κ vérifie l'hypothèse (\mathcal{H}) , nous pouvons utiliser la borne inférieure des grandes déviations pour conclure que, pour $\varepsilon < \delta/3 \wedge \gamma/3$,

$$P_{x^0}(d(Y, \phi) < \delta) \geq P_{x^0}(d(Y, L_\kappa) < \delta/3) \geq e^{-m(I(\phi)+\gamma)}.$$

22.5 Borne supérieure du PGD

Soit $\rho > 0$ et définissons

$$T_\rho = \inf \{ t \geq 0 : Y(t) \notin \mathcal{D}_\rho^{k+1} \}.$$

Nous considérons le processus $(Y^\rho(t), t \geq 0)$ défini par

$$\forall t \geq 0 \quad Y^\rho(t) = Y(t \wedge T_\rho).$$

Rappelons que notre objectif est de démontrer le résultat suivant : pour tout $\delta, \gamma, \eta_0, T > 0$, asymptotiquement,

$$P(d_T(Y^\rho, \Phi_{y^0}(\eta)) \geq \delta \mid Y(0) = y^0) \leq \exp(-m(\eta - \gamma)),$$

uniformément sur $y^0 \in \mathcal{D}_\rho^{k+1}$ et $\eta \leq \eta_0$. Nous construisons d'abord une approximation polygonale $(L(t))_{t \geq 0}$ du processus Y^ρ . Soit $N \geq 1$ et posons $\Delta t = T/N$.

Nous approximons le processus Y^ρ par la ligne polygonale $L(t)_{t \geq 0}$ passant par les points :

$$(0, Y^\rho(0)), (\Delta t, Y^\rho(\Delta t)), \dots, (T, Y^\rho(T)).$$

Soit $\delta' < \delta/2$ et introduisons les événements suivants :

$$A(i) = \left\{ \sup_{i\Delta t \leq t \leq (i+1)\Delta t} |Y(t) - Y(i\Delta t)| < \delta' \right\}, \quad 0 \leq i \leq N-1.$$

Si tous les événements $A(i)$ sont réalisés, nous avons $d(Y^\rho, L) < \delta$. C'est-à-dire, que si tous les événements $A(i)$ sont réalisés, et que nous avons en plus $d(Y^\rho, \Phi_{y^0}(\eta)) \geq \delta$, alors $L \notin \Phi_{y^0}(\eta)$. Il s'ensuit que, pour tout $x^0 \in \mathcal{S}^{\ell+1}$ vérifiant $\pi_K(x^0) \in \mathcal{D}_\rho^{K+1}$,

$$\begin{aligned} P_{x^0}(d(Y^\rho, \Phi_{\pi_K(x^0)}(\eta)) < \delta) &\leq \\ &P_{x^0}\left(\bigcup_{i=0}^{N-1} A(i)^c\right) + P_{x^0}\left(\bigcap_{i=0}^{N-1} A(i), L \notin \Phi_{\pi_K(x^0)}(\eta)\right). \end{aligned}$$

La première probabilité est majorée par

$$P_{x^0}\left(\bigcup_{i=0}^{N-1} A(i)^c\right) \leq N \sup_{x \in \mathcal{S}^{\ell+1}} P_x\left(\sup_{0 \leq t \leq \Delta t} |Y(t) - \pi_K(x)| > \delta'\right).$$

Lemme 22.5.1. *Pour tout $x^0 \in \mathcal{S}^{\ell+1}$, nous avons*

$$\begin{aligned} P_{x^0}\left(\sup_{0 \leq t \leq \Delta t} |Y(t) - \pi_K(x^0)| > \delta'\right) &\leq \\ &2 \sup_{x \in \mathcal{S}^{\ell+1}} \sup_{0 \leq t \leq \Delta t} P_x(|Y(t) - \pi_K(x)| > \delta'/2). \end{aligned}$$

Démonstration. Soit τ le temps d'arrêt défini par

$$\tau = \inf \{ t \geq 0 : |Y(t) - \pi_K(x^0)| > \delta' \} \wedge \Delta t.$$

Nous avons

$$\begin{aligned} P_{x^0}\left(\sup_{0 \leq t \leq \Delta t} |Y(t) - \pi_K(x^0)| > \delta'\right) &\leq P_{x^0}\left(|Y(\Delta t) - \pi_K(x^0)| > \delta'/2\right) \\ &\quad + P_{x^0}\left(|Y(\Delta t) - Y(\tau)| > \delta'/2\right). \end{aligned}$$

Or, par la propriété de Markov forte,

$$P_{x^0}(|Y(\Delta t) - Y(\tau)| > \delta'/2) \leq \sup_{x \in \mathcal{S}^{\ell+1}} \sup_{0 \leq t \leq \Delta t} P_x(|Y(t) - \pi_K(x)| > \delta'/2).$$

ce qui donne le résultat voulu. \square

D'après le lemme précédent, nous avons

$$P_{x^0} \left(\bigcup_{i=0}^{N-1} A(i)^c \right) \leq 2N \sup_{0 \leq t \leq \Delta t} \sup_{x \in \mathcal{S}^{\ell+1}} P_x(|Y(t) - \pi_K(x)| > \delta'/2).$$

Par l'inégalité de Chebyshev exponentielle, pour tout $\alpha > 0$,

$$P_x(|Y(t) - \pi_K(x)| > \delta'/2) \leq e^{-m\alpha\delta'/2} \sum_{\sigma \in \{-1, +1\}^{K+1}} E_x \left(e^{m\alpha\sigma(Y(t) - \pi_K(x))} \right).$$

Par le lemme 22.2.12, la variable aléatoire suivante a espérance égale à 1 :

$$\exp \left(\sum_{s=0}^{\lfloor t/\tau \rfloor \tau - \tau} m\alpha\sigma(Y(s+\tau) - Y(s)) - m \sum_{s=0}^{\lfloor t/\tau \rfloor \tau - \tau} G(X(s); \alpha\sigma)\tau \right).$$

Par le corollaire 22.3.5, il existe une constante $C(\alpha) > 0$ dépendante de α telle que pour $1/\ell, q, |\ell q - \alpha|$ assez petits,

$$\sup \{ G(x; \alpha\sigma) : x \in \mathcal{S}^{\ell+1}, \sigma \in \{-1, 1\}^{K+1} \} < C(\alpha).$$

Il s'ensuit que, pour tout $\sigma \in \{-1, 1\}^{K+1}$,

$$E_x \left(e^{m\alpha\sigma(Y(t) - \pi_K(x))} \right) = E_x \left(e^{m\alpha\sigma(Y(\lfloor t/\tau \rfloor \tau) - \pi_K(x))} \right) \leq e^{C(\alpha)mt}.$$

Nous concluons donc que

$$P_{x^0} \left(\bigcup_{i=0}^{N-1} A(i)^c \right) \leq 2^{K+2} N \exp(-m(\alpha\delta'/2 - C(\alpha)\Delta t)).$$

Nous choisissons maintenant $\alpha = 4\eta_0/\delta'$ et Δt tel que $C(\alpha)\Delta t \leq \eta_0$. Le terme de droite de cette inégalité est donc asymptotiquement négligeable vis-à-vis de $\exp(-m(\eta - \gamma))$. Nous procédons maintenant à la majoration de la probabilité

$$P_{x^0} \left(\bigcap_{i=0}^{N-1} A(i), L \notin \Phi_{\pi_K(x^0)}(\eta) \right).$$

Nous utilisons à nouveau l'inégalité de Chebyshev exponentielle. Nous avons

$$\begin{aligned} P_{x^0} \left(\bigcap_{i=0}^{N-1} A(i), L \notin \Phi_{\pi_K(x^0)}(\eta) \right) &= P_{x^0} \left(\bigcap_{i=0}^{N-1} A(i), I(L) > \eta \right) \\ &\leq e^{-m\eta} E_{x^0} \left(\mathbf{1}_{\bigcap_{i=0}^{N-1} A(i)} e^{mI(L)} \right). \end{aligned}$$

Il s'agit maintenant de borner l'action de la ligne polygonale L . Fixons $\varepsilon > 0$ et soit δ' assez petit pour que pour tout $y, y' \in \mathcal{D}_\rho^{K+1}$ et pour tout $u \in \mathbb{R}^{K+1}$,

$$|y - y'| < \delta' \quad \implies \quad |H^\infty(y; u) - H^\infty(y'; u)| < \varepsilon.$$

Si tous les événements $A(i)$ sont réalisés, nous pouvons borner l'action de la ligne polygonale L ainsi :

$$\begin{aligned} I(L) &= \sum_{i=1}^{N-1} \int_{i\Delta t}^{(i+1)\Delta t} H^\infty \left(L(t); \frac{Y^\rho((i+1)\Delta t) - Y^\rho(i\Delta t)}{\Delta t} \right) dt \\ &\leq \Delta t \sum_{i=1}^{N-1} H^\infty \left(Y^\rho(i\Delta t); \frac{Y^\rho((i+1)\Delta t) - Y^\rho(i\Delta t)}{\Delta t} \right) + \varepsilon T. \end{aligned}$$

Nous en déduisons que :

$$\begin{aligned} P_{x^0} \left(\bigcap_{i=0}^{N-1} A(i), L \notin \Phi_{\pi_K(x^0)}(\eta) \right) &\leq e^{-m(\eta - \varepsilon T)} E_{x^0} \left(\mathbf{1}_{\bigcap_{i=0}^{N-1} A(i)} \right. \\ &\quad \left. \times \prod_{i=0}^{N-1} \exp \left(m\Delta t H^\infty \left(Y^\rho(i\Delta t); \frac{Y^\rho((i+1)\Delta t) - Y^\rho(i\Delta t)}{\Delta t} \right) \right) \right). \end{aligned}$$

Par la propriété de Markov appliquée dans les instants $\Delta t, 2\Delta t, \dots, (N-1)\Delta t$, nous voyons que l'espérance de la formule est inférieure à

$$\left(\sup_{y \in \mathcal{D}_\rho^{K+1}} E_y \left(\mathbf{1}_{A(0)} \exp \left(m\Delta t H^\infty \left(y; \frac{Y^\rho(\Delta t) - y}{\Delta t} \right) \right) \right) \right)^N.$$

Nous remarquons que sur l'événement $A(0)$, $|Y^\rho(\Delta t) - y| < \delta'$. De plus,

$$Y^\rho(\tau) - y \in \frac{1}{m} \text{Conv}(\mathcal{V}) \quad \implies \quad Y^\rho(\Delta t) - y \in \Delta t \text{Conv}(\mathcal{V}).$$

Lemme 22.5.2. Soient $\varepsilon > 0$ et $y \in \mathcal{D}_\rho^{K+1}$. Il existe $n \geq 1$ et $z^0, \dots, z^n \in \mathbb{R}^{K+1}$ tels que :

$$\forall u \in \mathcal{U} = \left\{ u \in \mathbb{R}^{K+1} : |u| \leq \delta'/\Delta t \right\} \cap \text{Conv}(\mathcal{V})$$

$$H^\infty(y; u) \leq \max_{0 \leq i \leq n} (uz^i - G^\infty(y; z^i)) + \varepsilon,$$

où les vecteurs z^1, \dots, z^n dépendent de y , mais le nombre de vecteurs n dépend seulement de ε .

Avant de donner la démonstration, nous finissons la preuve de la borne supérieure des grandes déviations à l'aide du lemme. Soit $y \in \mathcal{D}_\rho^{K+1}$. Sur l'événement $A(0)$, nous avons bien

$$\frac{Y^\rho(\Delta t) - y}{\Delta t} \in \mathcal{U}.$$

Nous pouvons donc appliquer le lemme pour en déduire que

$$\begin{aligned} & \mathbb{E}_y \left(\mathbf{1}_{A(0)} \exp \left(m \Delta t H^\infty \left(y; \frac{Y^\rho(\Delta t) - y}{\Delta t} \right) \right) \right) \leq \\ & e^{m \Delta t \varepsilon} \mathbb{E}_y \left(\mathbf{1}_{A(0)} \exp \left(m \max_{0 \leq i \leq n} ((Y^\rho(\Delta t) - y)z^i - \Delta t G^\infty(y; z^i)) \right) \right) \leq \\ & e^{m \Delta t \varepsilon} \sum_{0 \leq i \leq n} \mathbb{E}_y \left(\mathbf{1}_{A(0)} \exp \left(m ((Y^\rho(\Delta t) - y)z^i - \Delta t G^\infty(y; z^i)) \right) \right). \end{aligned}$$

Observons ensuite que, pour tout $t \in [0, \Delta t]$, avec t multiple de τ , nous avons

$$\begin{aligned} (Y(t) - y)z^i - tG^\infty(y; z^i) = & \\ & \sum_{s=0}^{t-\tau} (Y(s+\tau) - Y(s))z^i - \sum_{s=0}^{t-\tau} G(\mathcal{X}(s); z^i)\tau \\ & + \sum_{s=0}^{t-\tau} \left(G(\mathcal{X}(s); z^i) - G^\infty(Y(s); z^i) \right) \tau \\ & + \sum_{s=0}^{t-\tau} \left(G^\infty(Y(s); z^i) - G^\infty(y; z) \right) \tau. \end{aligned}$$

Asymptotiquement, la deuxième ligne est plus petite que $t\varepsilon/2$ grâce à la proposition 22.3.3. Sur l'événement $A(0)$, il en est de même de la troisième ligne,

grâce à la continuité uniforme de G^∞ . Nous avons donc

$$\begin{aligned} \mathbb{E}_y \left(\mathbf{1}_{A(0)} e^{m \left((Y^\rho(\Delta t) - y) z^i - \Delta t G^\infty(y; z^i) \right)} \right) &\leq \\ e^{m \varepsilon \Delta t} \sup_{t \in [0, \Delta t]} \mathbb{E}_y \left(\exp \left(\sum_{s=0}^{t-\tau} (Y(s+\tau) - Y(s)) m z^i - \sum_{s=0}^{t-\tau} G(\mathcal{X}(s); z^i) \right) \right). \end{aligned}$$

Or, cette dernière espérance vaut 1 pour tout $t \in [0, \Delta t]$. Nous concluons que

$$P_{x^0} \left(\bigcap_{i=0}^{N-1} A(i), L \notin \Phi_{\pi_k(x^0)}(\eta) \right) \leq (n+1)^N e^{-m(\eta-3\varepsilon T)}.$$

Ceci achève la démonstration de la borne supérieure des grandes déviations. Nous procédons maintenant à la démonstration du lemme 22.5.2.

Démonstration. Soit $\varepsilon > 0$. Remarquons d'abord qu'il existe $h > 0$ tel que

$$\forall y \in \mathcal{D}_\rho^{K+1} \quad \forall u, u' \in U \quad |u - u'| < h \quad |H^\infty(y; u) - H^\infty(y; u')| < \varepsilon.$$

En effet, d'après le corollaire 22.3.8, la fonction H^∞ est uniformément continue en y , uniformément sur $u \in \mathbb{R}^{K+1}$. Il existe donc un recouvrement ouvert fini $V_1, \dots, V_{n'}$ de \mathcal{D}_ρ^{K+1} tel que, pour tout $1 \leq i \leq n'$ et pour tout $y, y' \in V_i$,

$$|H^\infty(y; u) - H^\infty(y'; u)| < \frac{\varepsilon}{2},$$

uniformément sur $u \in U$. Choisissons pour tout $1 \leq i \leq n'$ un représentant $y^i \in V_i$. La fonction $H^\infty(y; u)$ est aussi uniformément continue en u sur l'ensemble compact U . Pour $1 \leq i \leq n'$, il existe $h^i > 0$ tel que, pour tout $u, u' \in U$,

$$|u - u'| < h^i \implies |H^\infty(y^i; u) - H^\infty(y^i; u')| < \frac{\varepsilon}{2}.$$

Nous concluons en prenant $h \leq \min\{h^i : 1 \leq i \leq n'\}$. Considérons maintenant U_1, \dots, U_n un recouvrement ouvert de U dont chaque ensemble a un diamètre inférieur à $h/3$, et choisissons pour tout $1 \leq i \leq n$ un représentant $u^i \in U_i$. Remarquons ici que le choix de n est indépendant de y . Fixons $y \in \mathcal{D}_\rho^{K+1}$ et définissons, pour $k \geq 0$, l'ensemble $U(k)$ par

$$U(k) = \{ u \in U : H^\infty(y; u) \leq k\varepsilon \}.$$

Les ensembles emboîtés $U(k)$ sont convexes et vérifient

$$d(U(k), U \setminus U(k+1)) \geq h.$$

Nous allons nous servir des représentants u^i dans chaque bande $U(k+1) \setminus U(k)$ pour définir des z^i tels que l'inégalité

$$H^\infty(y; u) \leq \max_{1 \leq i \leq n} (uz^i - G^\infty(y; z^i)) + \varepsilon$$

soit vraie pour tout $u \in U(k+2) \setminus U(k+1)$. Fixons $k \geq 0$, et considérons le $h/3$ -voisinage de l'ensemble $U(k)$:

$$U_{h/3}(k) = \{ u \in U : d(u, U(k)) \leq h/3 \}.$$

Remarquons que tout comme $U(k)$, l'ensemble $U(k)_{h/3}$ est aussi convexe. Soit \mathcal{E}_k l'ensemble des vecteurs u^i présents dans la bande $U(k+1) \setminus U_{h/3}(k)$, et soit u^i un des représentants dans \mathcal{E}_k . Puisque l'ensemble $U(k)$ est convexe, il existe un hyperplan, $uz = d$, tangent à $U(k)$ et séparant le point u^i de l'ensemble $U(k)$. D'un côté de l'hyperplan nous avons le point ρ^* qui vérifie $H^\infty(y; \rho^*) < +\infty$, et de l'autre nous avons le point u^i , qui vérifie aussi $H^\infty(y; u^i) < +\infty$. Remarquons que nous avons de plus

$$\inf \{ H^\infty(y; u) : zu = d \} = k\varepsilon,$$

et que l'infimum est atteint au moins en un point \tilde{u}^i appartenant à l'intersection entre l'hyperplan et l'ensemble $U(k)$, en particulier $\tilde{u}^i \in \text{int}(U)$. Alors il existe une constante c telle que pour tout $u \in \mathbb{R}^{k+1}$,

$$H^\infty(y; u) \geq c(zu) - cd + k\varepsilon = (cz)u - G^\infty(y; cz).$$

En effet, d'après le théorème 37.3 de [64], nous avons

$$\begin{aligned} k\varepsilon &= \inf_{z \cdot u = d} H^\infty(y; u) = \inf_{z \cdot u = d} \sup_{z' \in \mathbb{R}^d} (u \cdot z' - G^\infty(y; z')) \\ &= \sup_{z' \in \mathbb{R}^d} \inf_{z \cdot u = d} (u \cdot z' - G^\infty(y; z')). \end{aligned}$$

Cette dernière quantité vaut $-\infty$ si z' n'est pas de la forme $z' = c'z$, avec $c' \in \mathbb{R}$. Il s'ensuit que

$$k\varepsilon = \sup_{c' \in \mathbb{R}} (c'd - G^\infty(y; c'z)).$$

Or, puisque $\tilde{u}^i \in \text{int}(U)$, le supremum précédent est atteint, en un point $c \in \mathbb{R}$. Nous prenons $z^i = cz$. Il s'ensuit que

$$G^\infty(y; z^i) = z^i \tilde{u}^i - k\varepsilon.$$

Remarquons que pour $u \in U(k)$:

$$k\varepsilon \geq H^\infty(y; u) \geq z^i u - G^\infty(y; z^i) = z^i(u - u^i) + k\varepsilon.$$

Donc, si $u \in U(K)$, alors $z^i u \leq d$. En appliquant ce même procédé à tous les vecteurs dans \mathcal{E}_k , nous voyons que l'ensemble des hyperplans obtenus séparent les ensembles $U(k)$ et $U \setminus U_{2h/3}(k)$, car tout point u dans la frontière de $U \setminus U_{2h/3}(k)$ est à distance plus petite que $h/3$ de l'un des vecteurs de \mathcal{E}_k . Nous en déduisons que tous les points dans $U \setminus U(k+1)$ sont séparés des $U(k)$, i.e., pour tout $u \in U \setminus U(k+1)$, il existe au moins un $u^i \in \mathcal{E}_k$ tel que

$$z^i u > z^i u^i = k\varepsilon + G^\infty(y; z).$$

En particulier, pour tout $u \in U(k+2) \setminus U(k+1)$ il existe au moins un $u^i \in \mathcal{E}_k$ tel que

$$H^\infty(y; u) - (z^i u - G^\infty(y; z^i)) < (k+2)\varepsilon - k\varepsilon = \varepsilon.$$

Nous prenons tous les z^i ainsi construits, et nous rajoutons le vecteur $z^0 = 0$ pour que le résultat soit vérifié par les vecteurs dans $U(1)$ aussi. L'ensemble de tous ces vecteurs z^i donne le résultat voulu. \square

Chapitre 23

Sortie d'un domaine

Soit $D \subset \mathcal{D}_\rho^{K+1}$ un domaine fermé dont la frontière ∂D est une courbe de classe C^∞ . L'objectif de ce chapitre est d'estimer le temps moyen de sortie du domaine D pour le processus $(Y(t))_{t \geq 0}$. Pour cela nous allons nous servir du principe de grandes déviations 22.1.1 obtenu dans le chapitre précédent, et nous allons suivre la stratégie proposée par Freidlin et Wentzell dans [38]. Tous les résultats de ce chapitre se généralisent à un nombre de coordonnées supérieure à $K + 1$, mais par simplicité, nous les formulons seulement pour les premières $K + 1$ coordonnées.

23.1 Perturbation aléatoire d'un système d'EDO

Le processus stochastique $(Y(t))_{t \geq 0}$ peut être vu comme une perturbation aléatoire du système d'équations différentielles $x'(t) = b(x(t))$ où

$$\forall 0 \leq k \leq K \quad \forall x \in \mathcal{D}^{K+1}$$

$$b_k(x) = \frac{1}{\lambda} \sum_{0 \leq h \leq k} x_h f(h) e^{-a} \frac{a^{k-h}}{(k-h)!} - \frac{1}{\lambda} x_k \left(\sum_{0 \leq h \leq K} x_h (f(h) - 1) + 1 \right).$$

Nous commençons par faire quelques remarques sur ce système d'équations. Pour $x \in \mathbb{R}^{K+1}$ et $\delta > 0$, nous notons par $B(x, \delta)$ la boule ouverte de centre x et de rayon δ . Soit $(x^x(t))_{t \geq 0}$ la solution du système d'équations avec condition initiale $x(0) = x$. Alors la trajectoire changée de temps $(x^x(\lambda t))_{t \geq 0}$ est une so-

lution du système d'équations (S_∞) considéré dans le chapitre 5. Il est facile de voir, par exemple à l'aide du lemme de Gronwall, que les trajectoires $(x^x(t))_{t \geq 0}$ dépendent de façon continue de la condition initiale x , i.e., pour tout $\varepsilon, t > 0$, il existe $\delta > 0$ tel que

$$|x - y| < \delta \quad \Rightarrow \quad |x^x(t) - x^y(t)| < \varepsilon.$$

Soit $\rho^* = (\rho_0^*, \dots, \rho_K^*)$ la projection sur les $K + 1$ premières coordonnées de la distribution de quasi-espèce autour de 0 associée au système (S_{f_H}) (cf. le chapitre 5). Par le théorème 5.5.1, la solution stationnaire ρ^* est stable dans l'ensemble compact \mathcal{D}_ρ^{K+1} . Soient $\varepsilon < \varepsilon'$ tels que pour toute condition initiale $x \in B(\rho^*, \varepsilon)$, la trajectoire $x^x(t)$ reste toujours dans la boule $B(\rho^*, \varepsilon')$. Considérons, pour tout $x \in \mathcal{D}_\rho^{K+1}$, le temps d'entrée dans la boule $B(x, \varepsilon)$:

$$T(x) = \inf \{ t \geq 0 : x^x(t) \in B(\rho^*, \varepsilon) \}.$$

Grâce au théorème 5.5.1, nous savons que $T(x) < +\infty$ pour tout $x \in \mathcal{D}_\rho^{K+1}$. Puisque la trajectoire $x^x(t)$ dépend de manière continue de x , la fonction $T(x)$ est sémicontinue supérieurement, et elle atteint donc son maximum dans l'ensemble compact \mathcal{D}_ρ^{K+1} , que nous appelons T_ε . De plus, la trajectoire $x^x(t)$ reste toujours à l'intérieur de l'ensemble \mathcal{D}^{K+1} . En effet, pour tout $k \geq 0$, nous avons,

$$b_k(x) \geq \frac{1}{\lambda} x_k (f_H(k) e^{-\alpha} - f_H(0)),$$

ce qui entraîne que pour tout $t \geq 0$,

$$x_k^x(t) \geq x_k \exp(t(f_H(k) e^{-\alpha} - f_H(0))) > 0.$$

Considérons pour chaque $x \in \mathcal{D}_\rho^{K+1}$ la fonction

$$\delta(x) = \inf \{ d(x^x(t), \partial \mathcal{D}^{K+1}) : t \geq 0 \}.$$

L'infimum précédent ne peut être réalisé que pour $t \in [0, T_\varepsilon]$. Nous en déduisons que δ est une fonction continue de x , et qu'elle atteint son minimum $\delta_\rho > 0$ sur l'ensemble compact \mathcal{D}_ρ^{K+1} .

23.2 Résultats supplémentaires

Nous donnons ici quelques lemmes utiles pour la suite. Ces résultats peuvent être trouvés dans [38], chapitre 4, pour des perturbations gaussiennes des systèmes d'équations. Nous formulons sans démonstration les équivalents dans

notre cas ; les démonstrations sont très similaires à celles de [38] et reposent sur le principe de grandes déviations 22.1.1, ainsi que sur les remarques de la section précédente. Dans la suite, nous utilisons la notation P_x, E_x pour dénoter les probabilités et espérances associées au processus $(Y(t))_{t \geq 0}$ partant du point x , et nous supposons implicitement que x est un point de l'ensemble Δ^{K+1} . Soit $D \subset \text{int}(\mathcal{D}_\rho^{K+1})$ un ensemble fermé dont la frontière est une courbe de classe C^∞ . D'après les remarques de la section précédente, nous pouvons supposer que ρ est assez petit pour que toute trajectoire $(x^x(t))_{t \geq 0}$ issue d'un point $x \in D$ reste toujours dans l'ensemble \mathcal{D}_ρ^{K+1} . À l'aide du principe de grandes déviations et du corollaire 22.3.10, nous déduisons la loi des grands nombres trajectorielle pour le processus $(Y(t))_{t \geq 0}$: pour tout $x \in D$, nous avons,

$$\lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha, \frac{m}{\ell} \rightarrow \alpha}} P_x \left(d_T(Y(t), x^x(t)) > \delta \right) = 0,$$

uniformément sur $x \in D$. Cela signifie que, si le point ρ^* n'appartient pas à l'ensemble D , le temps de sortie de D est fini avec très grande probabilité. Nous allons donc supposer que $\rho^* \in \text{int}(D)$.

Soit $\varepsilon > 0$ assez petit pour que $B(\rho^*, \varepsilon) \subset \text{int}(D)$. Définissons le temps τ_ε de sortie de l'ensemble $D \setminus B(\rho^*, \varepsilon)$ pour le processus $(Y(t))_{t \geq 0}$:

$$\tau_\varepsilon = \inf \{ t \geq 0 : Y(t) \notin D \setminus B(\rho^*, \varepsilon) \}.$$

Lemme 23.2.1. *Pour tout $\varepsilon > 0$, nous avons :*

- Il existe $\alpha, T_0 > 0$ tels que, pour toute trajectoire $\phi : [0, T] \rightarrow D \setminus B(\rho^*, \varepsilon)$,

$$\forall T > T_0 \quad I_T(\phi) > \alpha(T - T_0).$$

- Il existe $c, T_0 > 0$ tels que, pour m assez grand et pour tout $x \in D \setminus B(\rho^*, \varepsilon)$,

$$\forall T > T_0 \quad P_x(\tau_\varepsilon > T) \leq e^{-mc(T-T_0)}.$$

La fonctionnelle d'action I_T peut être généralisée pour mesurer l'action des courbes ϕ définies sur des intervalles de temps différents de $[0, T]$. En effet, soit $\phi : [T_1, T_2] \rightarrow \mathcal{D}_\rho^{K+1}$ une courbe avec $-\infty \leq T_1 < T_2 \leq +\infty$. Nous définissons l'action $I_{T_1 T_2}(\phi)$ de la courbe ϕ par :

$$I_{T_1 T_2}(\phi) = \int_{T_1}^{T_2} H^\infty(\phi(t); \phi'(t)) dt.$$

Nous définissons le quasi-potentiel associé à l'action I par : pour tout $x, y \in \mathcal{D}_\rho^{K+1}$,

$$V(x, y) = \inf \{ I_{T_1 T_2}(\phi) : \phi(T_1) = x, \phi(T_2) = y, -\infty \leq T_1 < T_2 \leq +\infty \}.$$

Le quasi-potentiel vérifie $V(x, y) \geq 0$. De plus, $V(\rho^*, \rho^*) = 0$ et $V(\rho^*, x)$ est une fonction continue en x . Prenons

$$V_0 = \min_{x \in \partial D} V(\rho^*, x).$$

Lemme 23.2.2. *Pour tout $x \in B(\rho^*, \varepsilon)$ et $d > 0$, il existe $h, T_0 > 0$ et une trajectoire $\phi^x : [0, T_0] \rightarrow \mathcal{D}_\rho^{K+1}$ satisfaisant : $\phi^x(0) = x$, il existe $T(x) < T_0$ tel que $d(\phi^x(T(x)), D) > h$, $\phi^x(t)$ ne rentre pas dans la boule $B(\rho^*, \varepsilon/2)$ une fois qu'elle est sortie de $B(\rho^*, \varepsilon)$, et $I_{T(x)}(\phi^x) < V_0 + d$.*

23.3 Temps de sortie

Nous énonçons maintenant le résultat principal de ce chapitre. Soient $D \subset D' \subset \text{int}(\mathcal{D}_\rho^{K+1})$ des ensembles fermés dont les frontières sont des courbes de classe C^∞ . Nous supposons que $\rho^* \in \text{int}(D)$, et que pour tout $x \in D$, la solution du système $x'(t) = b(x(t))$ avec condition initiale $x(0) = x$ reste dans D' pour tout $t \geq 0$. Nous définissons le temps de sortie de D' pour le processus $(Y(t))_{t \geq 0}$ par

$$\tau(D') = \inf \{ t \geq 0 : Y(t) \notin D' \}.$$

Théorème 23.3.1. *Pour tout $x \in D$, nous avons*

$$\lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha, \frac{m}{\ell} \rightarrow \alpha}} \frac{1}{m} \ln E_x(\tau(D')) = \min_{y \in \partial D} V(\rho^*, y) = V_0.$$

Démonstration. Tout le long de cette démonstration, par *asymptotiquement* nous voulons dire : pour m, ℓ assez grands, q assez petit, ℓq assez proche de α et m/ℓ assez proche de α . Nous commençons par montrer que, pour tout $d > 0$, asymptotiquement,

$$\frac{1}{m} \ln E_x(\tau(D')) < V_0 + d.$$

D'après les remarques dans la section 23.1, nous pouvons choisir $\varepsilon, T_1 > 0$ tels que toutes les trajectoires $(x^x(t))_{t \geq 0}$ avec condition initiale $x \in D$ rentrent

dans l'ensemble $B(\rho^*, \varepsilon/2)$ avant le temps T_1 , et ne sortent plus de l'ensemble $B(\rho^*, \varepsilon/2)$ après le temps T_1 . D'après le lemme précédent, nous pouvons aussi choisir $h, T_2 > 0$ tels que, pour tout point $x \in B(\rho^*, \varepsilon)$, il existe une trajectoire $(\phi^x(t))_{t \geq 0}$ vérifiant : $d(\phi^x(T(x)), D') > h$ avec $T(x) \leq T_2$, $\phi^x(t)$ ne rentre pas dans la boule $B(\rho^*, \varepsilon/2)$ après être sortie de $B(\rho^*, \varepsilon)$, et $I_{T(x)}(\phi^x) < V_0 + d/2$. Nous choisissons de plus ε assez petit pour que $B(\rho^*, 2\varepsilon) \subset \text{int}(D)$. Grâce au principe de grandes déviations 22.1.1, nous avons : pour tout $y \in B(\rho^*, \varepsilon)$,

$$P_y(d_{T(y)}(Y(t), \phi^y(t)) < h) \geq \exp(-m(I_{T(y)}(\phi^y) + d/2)) \geq e^{-m(V_0+d)}.$$

Puisque $d(\phi^y(T(y)), D') > 0$, nous avons

$$P_y(\tau(D') < T_2) \geq P_y(\tau(D') < T(y)) \geq e^{-m(V_0+d)}.$$

Nous notons par τ_ε le temps d'entrée dans la boule $B(\rho^*, \varepsilon)$:

$$\tau_\varepsilon = \inf \{ t \geq 0 : Y(t) \in B(\rho^*, \varepsilon) \}.$$

Soit $x \in D$, par la propriété de Markov forte, asymptotiquement,

$$\begin{aligned} P_x(\tau(D') < T_1 + T_2) &\geq E_x(1_{\tau_\varepsilon < T} P_{Y(\tau_\varepsilon)}(\tau(D') < T_2)) \\ &\geq P_x(\tau_\varepsilon < T_1) e^{-m(V_0+d)} \geq \frac{1}{2} e^{-m(V_0+d)}, \end{aligned}$$

où la dernière inégalité a été obtenue grâce à la loi des grands nombres trajectorielle pour le processus $(Y(t))_{t \geq 0}$ (section précédente). Nous décomposons l'espérance $E_x(\tau(D'))$ grâce à la propriété de Markov :

$$\begin{aligned} E_x(\tau(d)) &\leq \sum_{n \geq 0} (n+1)(T_1 + T_2) P_x(n(T_1 + T_2) < \tau(D') < (n+1)(T_1 + T_2)) \\ &= (T_1 + T_2) \sum_{n \geq 0} P_x(\tau(D') > n(T_1 + T_2)) \\ &\leq (T_1 + T_2) \sum_{n \geq 0} \left(1 - \min_{y \in D} P_y(\tau(D') \leq T_1 + T_\varepsilon)\right)^n \\ &= (T_1 + T_2) \left(\min_{y \in D} P_y(\tau(D') \leq T_1 + T_\varepsilon)\right)^{-1} \leq 2(T_1 + T_2) \varepsilon^{m(V_0+d)}. \end{aligned}$$

Donc, asymptotiquement,

$$\frac{1}{m} \ln E_x(\tau(D')) < V_0 + d.$$

Nous démontrons maintenant que, asymptotiquement,

$$\frac{1}{m} \ln E_x(\tau(D')) > V_0 - d.$$

Soit $U = B(\rho^*, \varepsilon/2) \cup (\mathcal{D}_\rho^{K+1} \setminus D')$. Nous définissons les suites de temps d'arrêts suivantes : soit $\tau_0^* = 0$ et prenons

$$\begin{aligned} \tau_1 &= \inf \{ t \geq \tau_0^* : Y(t) \notin B(\rho^*, \varepsilon) \}, & \tau_1^* &= \inf \{ t \geq \tau_1 : Y(t) \in U \}, \\ \vdots & & \vdots & \\ \tau_i &= \inf \{ t \geq \tau_{i-1}^* : Y(t) \notin B(\rho^*, \varepsilon) \}, & \tau_i^* &= \inf \{ t \geq \tau_i : Y(t) \in U \}, \\ \vdots & & \vdots & \end{aligned}$$

Considérons le processus $(Z_n)_{n \geq 0}$ à valeurs dans U défini par

$$\forall n \geq 0 \quad Z_n = Y(\tau_n^*).$$

Nous notons par Γ la bande $B(\rho^*, 2\varepsilon) \setminus B(\rho^*, \varepsilon)$. Remarquons que pour m assez grand, si $Y(\tau_n^*) \in B(\rho^*, \varepsilon/2)$, alors $Y(\tau_{n+1}^*) \in \Gamma$. Pour tout $n \geq 0$ et $x \in B(\rho^*, \varepsilon/2)$, nous avons

$$\begin{aligned} P(Z_{n+1} \in \mathcal{D}_\rho^{K+1} \setminus D' \mid Z_n = x) &\leq \max_{y \in \Gamma} P_y(\tau_1^* = \tau(D')) \\ &= \max_{y \in \Gamma} (P_y(\tau_1^* = \tau(D') < T) + P_y(\tau_1^* = \tau(D') \geq T)). \end{aligned}$$

D'après le lemme 23.2.1, pour T assez grand,

$$P_y(\tau_1^* = \tau(D') \geq T) \leq \frac{1}{2} e^{-m(V_0 - h)}.$$

Nous estimons ensuite la première des probabilités dans la somme. Il découle de la définition du quasi-potentiel que, pour ε suffisamment petit, les trajectoires $(Y(t))_{0 \leq t \leq T}$ vérifiant $\tau_1^* = \tau(D') < T$, sont à distance positive de l'ensemble

$$\{ \phi : [0, T] \rightarrow \mathcal{D}_\rho^{K+1} : \phi(0) = y, I_T(\phi) < V_0 - h/2 \}.$$

Nous déduisons du principe de grandes déviations que, asymptotiquement,

$$P_y(\tau_1^* = \tau(D') < T) < \frac{1}{2} e^{-m(V_0 - h)}.$$

Il s'ensuit que

$$P(Z_{n+1} \in \mathcal{D}_\rho^{K+1} \setminus D' \mid Z_n = x) \leq e^{-m(V_0 - h)}.$$

Soit N le premier temps de sortie de l'ensemble D' pour le processus $(Z_n)_{n \geq 0}$:

$$N = \inf \{ n \geq 0 : Z_n \notin D' \}.$$

Grâce à l'inégalité précédente, nous avons, pour tout $x \in B(\rho^*, \varepsilon)$ et $n \geq 1$,

$$P_x(N \geq n) \geq \left(1 - e^{-m(V_0-h)}\right)^{n-1}.$$

Puisque $\tau(D') = (\tau_1^* - \tau_0^*) + \dots + (\tau_N^* - \tau_{N-1}^*)$, nous pouvons réécrire l'espérance de $\tau(D')$ sous la forme

$$E_x(\tau(D')) = \sum_{n \geq 1} E_x(1_{N \geq n}(\tau_n^* - \tau_{n-1}^*)).$$

Nous utilisons la propriété de Markov forte pour minorer les espérances dans la somme :

$$E_x(1_{N \geq n}(\tau_n^* - \tau_{n-1}^*)) \geq E_x(1_{N \geq n}(\tau_n^* - \tau_n)) \geq P_x(N \geq n) \inf_{x \in \Gamma} E_x(\tau_1^*).$$

Ce dernier infimum est minoré par une constante positive t_1 , qui ne dépend pas de m , car les solutions du système d'équations $x'(t) = b(x(t))$ passent un temps positif pour aller de Γ à la boule $B(\rho^*, \varepsilon/2)$. Nous mettons ensemble toutes les estimations obtenues précédemment et nous obtenons

$$\begin{aligned} E_x(\tau(D')) &> t_1 \sum_{n \geq 1} \min_{z \in B(\rho^*, \varepsilon/2)} P_z(N \geq n) \\ &\geq t_1 \sum_{n \geq 1} \left(1 - e^{-m(V_0-h)}\right)^{n-1} = t_1 e^{m(V_0-h)}. \end{aligned}$$

Ceci entraîne le résultat voulu pour $x \in B(\rho^*, \varepsilon/2)$. Pour $x \in D$, nous utilisons le fait que $P_x(\tau(D') > \tau_1^*)$ tend vers 1 asymptotiquement. Donc,

$$\begin{aligned} E_x(\tau(D')) &= E_x(1_{\tau(D') > \tau_1^*} \tau(D')) \geq E_x(1_{\tau(D') > \tau_1^*} E_{Y(\tau_1^*)}(\tau(D'))) \\ &> t_1 e^{m(V_0-h)} P_x(\tau(D') > \tau_1^*) > \frac{t_1}{2} e^{m(V_0-h)} P_x(\tau(D') > \tau_1^*). \end{aligned}$$

Ceci entraîne le résultat voulu. □

Chapitre 24

Disparition de la quasi-espèce pour le paysage à un pic

Dans le cas du paysage à un pic, asymptotiquement, la dynamique du nombre de master sequences ne dépend pas des autres coordonnées. Nous pouvons donc travailler en une dimension, ce qui fait possible le calcul explicite de la fonctionnelle d'action et du quasipotential. Nous allons utiliser cet avantage pour décrire la disparition de la quasi-espèce. Tout le long de ce chapitre, nous supposons que $f_H : \mathbb{N} \rightarrow \mathbb{R}^+$ est la fonction de fitness du paysage à un pic :

$$\forall k \geq 0 \quad f_H(k) = \begin{cases} \sigma & \text{si } k = 0, \\ 1 & \text{si } k \geq 1, \end{cases}$$

où le paramètre $\sigma > 1$ représente l'avantage sélectif de la master sequence.

24.1 Calcul de la fonctionnelle d'action

Dans le cas du paysage à un pic, le principe de grandes déviations 22.1.1 est vrai avec $k = 0$. Pour $k = 0$, le processus $(Y(t))_{t \geq 0}$ est unidimensionnel, et ne peut sauter qu'à gauche ou à droite. Les probabilités asymptotiques de sauter à gauche ou à droite, $C_y^\infty(-e_0)$ et $C_y^\infty(e_0)$, que nous notons par la suite $g(y)$ et

$d(y)$, sont données par : pour tout $y \in [0, 1]$,

$$g(y) = \frac{1}{\lambda} y (y \sigma (1 - e^{-\alpha}) + 1 - y),$$

$$d(y) = \frac{1}{\lambda} (1 - y) y \sigma e^{-\alpha}.$$

Soit $G^\infty(y; z)$ la log-Laplace d'une variable aléatoire de loi

$$\mu = g(y)\delta_{-1} + d(y)\delta_1 + (1 - g(y) - d(y))\delta_0.$$

Nous avons, pour tout $y \in [0, 1]$ et pour tout $z \in \mathbb{R}$,

$$G^\infty(y; z) = \ln (1 - g(y) - d(y) + g(y)e^{-z} + d(y)e^z).$$

Nous calculons maintenant la transformée de Fenchel-Legendre $H^\infty(y; u)$ de la fonction $G^\infty(y; z)$, pour $y \in]0, 1[$. Par commodité, durant le calcul, nous omettons l'argument y dans $g(y)$ et $d(y)$. Nous savons que, pour tout u dans l'ensemble $] -\infty, -1[\cup] 1, +\infty[$, la fonction $h^\infty(y; u)$ vaut l'infini. Soit $u \in] -1, 1[$ et essayons de résoudre l'équation

$$\frac{d}{dz} G^\infty(y; z) = u.$$

Nous dérivons la fonction $G^\infty(y; z)$ par rapport à z et nous obtenons l'équation

$$\frac{-ge^{-z} + de^z}{1 - g - d + ge^{-z} + de^z} = u.$$

La seule solution de cette équation est donnée par

$$z(u) = \ln \frac{(1 - g - d)u + \sqrt{(1 - g - d)^2 u^2 + 4gd(1 - u^2)}}{2d(1 - u)}.$$

Nous avons alors

$$H^\infty(y; u) = uz(u) - G^\infty(y; z(u)).$$

Nous procédons maintenant au calcul de $G^\infty(y; z(u))$. D'une part, nous avons

$$ge^{-z(u)} = g \frac{2d(1 - u)}{(1 - g - d)u + \sqrt{(1 - g - d)^2 u^2 + 4gd(1 - u^2)}}.$$

Nous multiplions le numérateur et le dénominateur par le conjugué du dénominateur pour obtenir

$$ge^{-z(u)} = g \frac{2d(1 - u) \left((1 - g - d)u - \sqrt{(1 - g - d)^2 u^2 + 4gd(1 - u^2)} \right)}{-4gd(1 - u^2)}$$

$$= - \frac{(1 - g - d)u - \sqrt{(1 - g - d)^2 u^2 + 4gd(1 - u^2)}}{2(1 + u)}.$$

D'autre part, nous avons

$$\begin{aligned} de^{z(u)} &= d \frac{(1-g-d)u + \sqrt{(1-g-d)^2u^2 + 4gd(1-u^2)}}{2d(1-u)} \\ &= \frac{(1-g-d)u + \sqrt{(1-g-d)^2u^2 + 4gd(1-u^2)}}{2(1-u)}. \end{aligned}$$

Nous combinons ces deux expressions et nous obtenons

$$\begin{aligned} 1-g-d + ge^{-z(u)} + de^{z(u)} &= \\ 1-g-d + \frac{(1-g-d)u^2 + \sqrt{(1-g-d)^2u^2 + 4gd(1-u^2)}}{1-u^2} &= \\ = \frac{1-g-d + \sqrt{(1-g-d)^2u^2 + 4gd(1-u^2)}}{(1-u^2)}. \end{aligned}$$

Nous concluons que, pour tout $y \in]0, 1[$ et $u \in]-1, 1[$, la valeur de $H^\infty(y; u)$ est donnée par :

$$\begin{aligned} H^\infty(y; u) &= \\ u \ln \frac{(1-g(y)-d(y))u + \sqrt{(1-g(y)-d(y))^2u^2 + 4g(y)d(y)(1-u^2)}}{2d(y)(1-u)} & \\ - \ln \left(\frac{(1-g(y)-d(y)) + \sqrt{(1-g(y)-d(y))^2u^2 + 4g(y)d(y)(1-u^2)}}{(1-u^2)} \right). & \end{aligned}$$

En outre, nous avons, pour tout $y \in]0, 1[$,

$$H^\infty(y; -1) = -\ln g(y) \quad \text{et} \quad H^\infty(y; 1) = -\ln d(y).$$

Sous les conventions $0 \ln 0 = 0 \ln(0/0) = 0$, nous pouvons réécrire l'expression de $H^\infty(y; u)$ comme suit : pour tout $y \in]0, 1[$ et $u \in [-1, 1]$,

$$\begin{aligned} H^\infty(y; u) &= \\ \frac{1+u}{2} \ln \frac{(1-g(y)-d(y))u + \sqrt{(1-g(y)-d(y))^2u^2 + 4g(y)d(y)(1-u^2)}}{2d(y)(1+u)} &+ \\ \frac{1-u}{2} \ln \frac{-(1-g(y)-d(y))u + \sqrt{(1-g(y)-d(y))^2u^2 + 4g(y)d(y)(1-u^2)}}{2g(y)(1-u)} & \\ + (1+u) \ln(1+u) - (1-u) \ln(1-u) & \\ - \ln \left(1-g(y)-d(y) + \sqrt{(1-g(y)-d(y))^2u^2 + 4g(y)d(y)(1-u^2)} \right). & \end{aligned}$$

Nous pouvons bien vérifier que, pour tout $y \in]0, 1[$,

$$H^\infty(y; d(y) - g(y)) = 0.$$

Soit $\phi : [0, T] \rightarrow [\rho, 1 - \rho]$. L'action $I(\phi)$ de la trajectoire ϕ est donc donnée par

$$\begin{aligned} I(\phi) = & \int_0^T \left(\frac{1 + \phi'(t)}{2} \ln \frac{(1 - g(\phi(t)) - d(\phi(t)))\phi'(t)}{2d(\phi(t))(1 + \phi'(t))} \right. \\ & + \frac{\sqrt{(1 - g(\phi(t)) - d(\phi(t)))^2\phi'(t)^2 + 4g(\phi(t))d(\phi(t))(1 - \phi'(t)^2)}}{2d(\phi(t))(1 + \phi'(t))} \\ & + \frac{1 - \phi'(t)}{2} \ln \frac{-(1 - g(\phi(t)) - d(\phi(t)))\phi'(t)}{2g(\phi(t))(1 - \phi'(t))2g(\phi(t))(1 - \phi'(t))} \\ & + \frac{\sqrt{(1 - g(\phi(t)) - d(\phi(t)))^2\phi'(t)^2 + 4g(\phi(t))d(\phi(t))(1 - \phi'(t)^2)}}{2g(\phi(t))(1 - \phi'(t))} \\ & + (1 + \phi'(t)) \ln(1 + \phi'(t)) - (1 - \phi'(t)) \ln(1 - \phi'(t)) \\ & \left. - \ln(1 - g(\phi(t)) - d(\phi(t))) \right. \\ & \left. + \sqrt{(1 - g(\phi(t)) - d(\phi(t)))^2\phi'(t)^2 + 4g(\phi(t))d(\phi(t))(1 - \phi'(t)^2)} \right) dt. \end{aligned}$$

24.2 Calcul du quasi-potentiel

Nous donnons ici une expression explicite pour le quasi-potentiel $V(\rho^*, x)$. Soit $\rho > 0$. Nous rappelons que le quasi-potentiel $V(\rho^*, x)$ est défini pour tout $x \in [\rho, 1 - \rho]$ par

$$V(\rho^*, x) = \inf \{ I_{T_1 T_2}(\phi) : \phi(T_1) = \rho^*, \phi(T_2) = x, -\infty \leq T_1 < T_2 \leq +\infty \}.$$

Nous rappelons que $\rho^* = (\sigma e^{-\alpha} - 1)/(\sigma - 1)$. Définissons pour tout $x \in [\rho, 1 - \rho]$ la quantité $U(x)$ par

$$U(x) = \int_{\rho^*}^x \ln \left(\frac{g(y)}{d(y)} \right) dy = \int_{\rho^*}^x \ln \left(\frac{y\sigma(1 - e^{-\alpha}) + 1 - y}{(1 - y)\sigma e^{-\alpha}} \right) dy.$$

Nous remarquons que, pour tout $y \in [\rho, 1 - \rho]$,

$$\left(\frac{y\sigma(1 - e^{-\alpha}) + 1 - y}{(1 - y)\sigma e^{-\alpha}} \right)' = \frac{\sigma^2 e^{-\alpha}(1 - e^{-\alpha})}{((1 - y)\sigma e^{-\alpha})^2} > 0.$$

La fonction $g(y)/d(y)$ est donc strictement croissante sur $[\rho, 1 - \rho]$. De plus,

$$\frac{g(\rho^*)}{d(\rho^*)} = \frac{\frac{(\sigma e^{-\alpha} - 1)\sigma(1 - e^{-\alpha})}{\sigma - 1} + \frac{\sigma - \sigma e^{-\alpha}}{\sigma - 1}}{\frac{(\sigma - \sigma e^{-\alpha})\sigma e^{-\alpha}}{\sigma - 1}} = 1.$$

Nous concluons que $U(\rho^*) = 0$ et pour tout $x \in [\rho, \rho^*[\cap]\rho^*, 1 - \rho]$, $U(x) > 0$. De plus, pour tout $x \in [\rho, 1 - \rho]$,

$$G^\infty(x; U'(x)) = \ln \left(1 - g(x) - d(x) + g(x) \frac{d(x)}{g(x)} + d(x) \frac{g(x)}{d(x)} \right) = 0.$$

La fonction $U(x)$ est en effet le quasi-potentiel.

Théorème 24.2.1. *Pour tout $x \in [\rho, 1 - \rho]$, nous avons*

$$V(\rho^*, x) = U(x).$$

De plus, l'infimum dans la définition du quasi-potentiel $V(\rho^, x)$ n'est atteint que pour les trajectoires $\phi : [-\infty, T_2] \rightarrow [\rho, 1 - \rho]$, $T_2 < +\infty$, vérifiant*

$$\begin{cases} \phi'(t) = g(\phi(t)) - d(\phi(t)), \\ \phi(T_2) = x. \end{cases}$$

Démonstration. Soit $x \in [\rho, 1 - \rho]$. Soient $-\infty \leq T_1 < T_2 \leq +\infty$ et $\phi : [T_1, T_2] \rightarrow [\rho, 1 - \rho]$ une trajectoire vérifiant $\phi(T_1) = \rho^*$ et $\phi(T_2) = x$. D'après la définition de la transformée de Fenchel-Legendre, pour tout $y \in [0, 1]$, $u \in \mathbb{R}$ et $z \in \mathbb{R}$,

$$H^\infty(y; u) \geq uz - G^\infty(y; z).$$

Donc, pour tout $t \in [T_1, T_2]$,

$$H^\infty(\phi(t); \phi'(t)) \geq \phi'(t)U'(\phi(t)) - G^\infty(\phi(t); U'(\phi(t))) = \phi'(t)U'(\phi(t)).$$

Nous en déduisons que

$$I_{T_1 T_2}(\phi) = \int_{T_1}^{T_2} H^\infty(\phi(t); \phi'(t)) dt \geq \int_{T_1}^{T_2} \phi'(t)U'(\phi(t)) dt = U(x).$$

Nous concluons que

$$\forall x \in [\rho, 1 - \rho] \quad V(\rho^*, x) \geq U(x).$$

Nous montrons maintenant que l'infimum est atteint. Soit $T_2 < +\infty$ et soit $\phi : [-\infty; T_2] \rightarrow \mathbb{R}$ une solution du problème :

$$\begin{cases} \phi'(t) = g(\phi(t)) - d(\phi(t)), \\ \phi(T_2) = x. \end{cases}$$

Pour tout $y \in [\rho, 1 - \rho]$, nous avons

$$\begin{cases} g(y) - d(y) < 0 & \text{si } y < \rho^*, \\ g(y) - d(y) > 0 & \text{si } y > \rho^*. \end{cases}$$

Nous en déduisons que, pour tout $t \in [-\infty, T_2]$, $\phi(t) \in [\rho, 1 - \rho]$. De plus, puisque $\phi'(t) = g(\phi(t)) - d(\phi(t))$, la solution de l'équation

$$\frac{d}{dz} G^\infty(\phi(t); z) = \phi'(t)$$

est justement $U'(\phi(t))$, ce qui entraîne que

$$H^\infty(\phi(t); \phi'(t)) = \phi'(t)U'(\phi(t)) - G^\infty(\phi(t); U'(\phi(t))) = \phi'(t)U'(\phi(t)).$$

Nous en déduisons que, pour tout $t \in [-\infty, T_2]$, nous avons

$$\frac{d}{dt} U(\phi(t)) = U'(\phi(t))\phi'(t) = H^\infty(\phi(t); \phi'(t)).$$

Cette dernière quantité est strictement positive pour tout t vérifiant $\phi(t) \neq \rho^*$. En effet, $H^\infty(x; u) \geq 0$ et vaut 0 si et seulement si $u = d(x) - r(x)$. Or, si $\phi(t) \neq \rho^*$, alors

$$\phi'(t) = g(\phi(t)) - d(\phi(t)) \neq d(\phi(t)) - g(\phi(t)).$$

La trajectoire ϕ est monotone, croissante si $x > \rho^*$ et décroissante si $x < \rho^*$. En particulier, la limite $\lim_{t \rightarrow -\infty} U(\phi(t)) > 0$ existe et est finie. Supposons que $\phi(-\infty) \neq \rho^*$. Alors il existe $\varepsilon > 0$ tel que, pour tout $t \in]-\infty, T_2]$, $|\phi(t) - \rho^*| > \varepsilon$. Nous en déduisons qu'il existe $\delta > 0$ tel que, pour tout $t \in]-\infty, T_2]$,

$$\frac{d}{dt} U(\phi(t)) = H^\infty(\phi(t); \phi'(t)) > \delta,$$

ce qui est en contradiction avec le fait que la limite $\lim_{t \rightarrow -\infty} U(\phi(t)) > 0$ est finie. Nous calculons maintenant l'action de la trajectoire ϕ . Nous avons,

$$\begin{aligned} I_{T_1 T_2}(\phi) &= \int_{T_1}^{T_2} H^\infty(\phi(t); \phi'(t)) dt \\ &= \int_{T_1}^{T_2} \frac{d}{dt} U(\phi(t)) dt = U(\phi(T_2)) - U(\phi(T_1)) = U(x). \end{aligned}$$

Pour finir, nous montrons que l'infimum dans $V(\rho^*, \varepsilon)$ n'est réalisé que pour les trajectoires $\phi : [-\infty, T_2] \rightarrow [-\varepsilon, \varepsilon]$, $T_2 < \infty$ vérifiant

$$\begin{cases} \phi'(t) = g(\phi(t)) - d(\phi(t)), \\ \phi(T_2) = \varepsilon. \end{cases}$$

En effet, soit $\phi : [T_1, T_2] \rightarrow [\varepsilon, 1 - \varepsilon]$ une trajectoire telle que $\phi(T_1) = \rho^*$, $\phi(T_2) = \varepsilon$, et $I_{T_1 T_2}(\phi) = V(\rho^*, \varepsilon)$. Nous avons :

$$V(\rho^*, \varepsilon) = U(\varepsilon) = \int_{T_1}^{T_2} \frac{d}{dt} U(\phi(t)) dt.$$

Nous en déduisons que

$$I_{T_1 T_2}(\phi) = \int_{T_1}^{T_2} H^\infty(\phi(t); \phi'(t)) dt = \int_{T_1}^{T_2} \frac{d}{dt} U(\phi(t)) dt.$$

Or, comme nous avons vu dans la démonstration du théorème précédent, pour tout $t \in [T_1, T_2]$,

$$H^\infty(\phi(t); \phi'(t)) \geq U'(\phi(t))\phi'(t) = \frac{d}{dt} U(\phi(t)).$$

Puisque $G^\infty(\phi(t), U'(\phi(t))) = 0$ Nous en déduisons que, presque pour tout $t \in [T_1, T_2]$,

$$H^\infty(\phi(t); \phi'(t)) = U'(\phi(t))\phi'(t) - G^\infty(\phi(t), U'(\phi(t))).$$

Autrement dit, la trajectoire ϕ satisfait

$$\phi'(t) = \frac{d}{dz} G^\infty(\phi(t); U'(\phi(t))) = g(\phi(t)) - d(\phi(t)).$$

Ceci achève la démonstration du théorème. □

24.3 Trajectoire de disparition de la quasi-espèce

Considérons maintenant le processus $(Y(t))_{t \geq 0}$, et définissons, pour $\varepsilon > 0$, les temps

$$\begin{aligned} T_\varepsilon &= \inf \{ t \geq 0 : Y(t) < \varepsilon \}, \\ \theta_\varepsilon &= \sup \{ t \leq T_\varepsilon : Y(t) > \rho^* - \varepsilon \}. \end{aligned}$$

Pour tout $x \in [\varepsilon, 1 - \varepsilon]$, soit $(\phi^x(t))_{t \geq 0}$ la solution de l'équation différentielle

$$\begin{cases} \phi'(t) = g(\phi(t)) - d(\phi(t)), \\ \phi(0) = x. \end{cases}$$

Théorème 24.3.1. *Pour tout $\delta > 0$, et pour tout $x \geq \varepsilon$, nous avons*

$$\lim_{\substack{\ell, m \rightarrow \infty, q \rightarrow 0 \\ \ell q \rightarrow \alpha, \frac{m}{\ell} \rightarrow \alpha}} P_x \left(\sup_{\theta_\varepsilon \leq t \leq T_\varepsilon} |Y(t) - \phi^{\rho^* - \varepsilon}(t - \theta_\varepsilon)| > \delta \right) = 0.$$

Avant de procéder à la démonstration, nous formulons trois lemmes préparatoires.

Lemme 24.3.2. *Il existe une constante $L > 0$ telle que, pour tout $x, y \in \mathcal{D}_\rho^{K+1}$, il existe une fonction continue et dérivable $\phi : [0, T] \rightarrow \mathcal{D}_\rho^{K+1}$ vérifiant $\phi(0) = x$, $\phi(T) = y$, $T = |x - y|$ et*

$$I_T(\phi) \leq L|x - y|.$$

Démonstration. Considérons la fonction $\phi : [0, T] \rightarrow \mathcal{D}_\rho^{K+1}$ donnée par

$$\forall t \in [0, T] \quad \phi(t) = x + t \frac{y - x}{|y - x|}.$$

La fonction ϕ vérifie $\phi(0) = x$ et $\phi(T) = y$. De plus, pour tout $t \in [0, T]$,

$$\phi'(t) = \frac{y - x}{|y - x|} \in \text{Conv}(\mathcal{V}).$$

Nous prenons

$$L = \sup_{\substack{y \in \mathcal{D}_\rho^{K+1} \\ u \in \text{Conv}(\mathcal{V})}} H^\infty(y; u),$$

et nous obtenons le résultat voulu. □

Soit $C > 0$ assez grande telle que, pour tout $x \in [\rho^* - \varepsilon - \delta/C, \rho^* - \varepsilon + \delta/C]$,

$$\sup_{t \geq 0} |\phi^x(t) - \phi^{\rho^* - \varepsilon}(t)| < \delta/4.$$

Considérons l'ensemble Φ des trajectoires

$$\Phi = \{ \phi : [0, T] \rightarrow [\varepsilon, 1 - \varepsilon], \phi'(t) = g(\phi(t)) - d(\phi(t)), T > 0 \}.$$

Il existe $d > 0$ tel que, pour tout $T > 0$ et pour toute trajectoire $\phi : [0, T] \rightarrow [\varepsilon, 1 - \varepsilon]$ vérifiant $\phi(0) = \rho^*$, $\phi(T) = \varepsilon$ et $d_T(\phi, \Phi) > \delta/C$, nous avons

$$I_T(\phi) \geq V(\rho^*, \varepsilon) + d.$$

Soit $\gamma = d/(5L) \wedge \delta/C$. Nous définissons le temps d'arrêt τ par

$$\tau = \inf \{ t \geq 0 : Y(t) > \rho^* - \gamma/2 \text{ ou } Y(t) < \varepsilon \}.$$

Lemme 24.3.3. *Pour tout $x \in [\rho^* - \gamma, \rho^* - \gamma/2]$, nous avons, pour m assez grand,*

$$P_x(Y(\tau) < \varepsilon) \geq e^{-m(V(\rho^*, \varepsilon) + d/3)}.$$

Démonstration. Soit $x \in [\rho^* - \gamma, \rho^* - \gamma/2]$ et considérons la trajectoire $(\phi^x(t))_{t \geq 0}$ solution de l'équation différentielle :

$$\begin{cases} \phi'(t) = g(\phi(t)) - d(\phi(t)), \\ \phi(0) = x. \end{cases}$$

La trajectoire ϕ converge vers 0, il existe donc $T \geq 0$ tel que $\phi^x(T) = \varepsilon$. De plus, la trajectoire $(\phi(t))_{0 \leq t \leq T}$ vérifie $I_T(\phi) \leq V(\rho^*, \varepsilon)$. Par le lemme précédent, il existe une trajectoire $(\phi^1(t))_{0 \leq t \leq T_1}$, $T_1 > 0$, vérifiant $\phi^1(0) = \varepsilon$, $\phi^1(T_1) = \varepsilon'$ et $I_{T_1}(\phi^1) < d/6$, pour $\varepsilon' < \varepsilon$ assez proche de ε . Soit $(\tilde{\phi}(t))_{0 \leq t \leq T+T_1}$ la trajectoire obtenue par concaténation des trajectoires ϕ^x et ϕ^1 . La trajectoire $\tilde{\phi}$ vérifie $\tilde{\phi}(0) = x$, $\tilde{\phi}(T+T_1) = \varepsilon'$ et $I_{T+T_1}(\tilde{\phi}) \leq V(\rho^*, \varepsilon) + d/6$. Nous avons, d'après le principe de grandes déviations 22.1.1, pour m assez grand,

$$P_x(Y(\tau) < \varepsilon) \geq P_x(d_{T+T_1}(Y, \tilde{\phi}) < (\varepsilon - \varepsilon')/2) \geq e^{-m(V(\rho^*, \varepsilon) + d/6 + d/6)},$$

comme voulu. □

Lemme 24.3.4. *Pour tout $x \in [\rho^* - \gamma, \rho^* - \gamma/2]$, nous avons*

$$P_x\left(Y(\tau) < \varepsilon, \sup_{\theta_\varepsilon \leq t \leq T_\varepsilon} |Y(t) - \phi^{\rho^* - \varepsilon}(t - \theta_\varepsilon)| > \delta\right) \leq e^{-m(V(\rho^*, \varepsilon) + 2d/3)}.$$

Démonstration. Grâce au lemme 23.2.1, il existe $T > 0$ tel que

$$P_x(\tau > T) \leq e^{-m(V(\rho^*, \varepsilon) + d)}.$$

Il suffit donc de majorer la probabilité

$$P_x\left(\tau \leq T, Y(\tau) < \varepsilon, \sup_{\theta_\varepsilon \leq t \leq T_\varepsilon} |Y(t) - \phi^{\rho^* - \varepsilon}(t - \theta_\varepsilon)| > \delta\right).$$

Soit $\phi : [0, T] \rightarrow [\varepsilon, 1 - \varepsilon]$ une trajectoire et définissons

$$t_\gamma = \inf \{ t \geq 0 : \phi(t) = \varepsilon + \gamma/2 \} \wedge T.$$

Nous montrons maintenant que toute trajectoire $\phi : [0, T] \rightarrow [\varepsilon, 1 - \varepsilon]$ telle que $\phi(0) \in [\rho^* - \gamma, \rho^* - \gamma/2]$, $d_{t_\gamma}((\phi)_{0 \leq t \leq t_\gamma}, \Phi) > \delta/C$, $I_T(\phi) \leq V(\rho^*, \varepsilon) + 0.65d$, vérifie nécessairement $t_\gamma = T$. En effet, soit $\phi : [0, T] \rightarrow [\varepsilon, 1 - \varepsilon]$ une telle trajectoire et supposons par l'absurde que $t_\gamma < T$. Alors la trajectoire ϕ restreinte à l'intervalle $[0, t_\gamma]$ vérifie $I_{0, t_\gamma}(\phi) < V(\rho^*, \varepsilon) + 0.65d$. D'après le lemme 24.3.2, il existe une trajectoire $\phi^1 : [0, T_1] \rightarrow [\varepsilon, 1 - \varepsilon]$, $T_1 \geq 0$, vérifiant $\phi^1(0) = \rho^*$, $\phi^1(T_1) = \phi(0)$ et $I_{0, T_1}(\phi^1) \leq 0.2d$. De même, il existe une trajectoire $\phi^2 : [0, T_2] \rightarrow [\varepsilon, 1 - \varepsilon]$, $T_2 \geq 0$, vérifiant $\phi^2(0) = \phi(t_\gamma)$, $\phi^2(T_2) = \varepsilon$ et $I_{0, T_2}(\phi^2) \leq 0.1d$. Nous construisons une trajectoire $\tilde{\phi} : [0, T_1 + t_\gamma + T_2]$ par concaténation des trajectoires ϕ^1 , ϕ restreinte à $[0, t_\gamma]$, et ϕ^2 . La trajectoire $\tilde{\phi}$ vérifie $\tilde{\phi}(0) = \rho^*$, $\tilde{\phi}(T_1 + t_\gamma + T_2) = \varepsilon$,

$$d_{T_1 + t_\gamma + T_2}(\tilde{\phi}, \Phi_{T_1 + t_\gamma + T_2}) > \delta/C \quad \text{et} \quad I_{T_1 + t_\gamma + T_2}(\tilde{\phi}) < V(\rho^*, \varepsilon) + d,$$

ce qui est absurde vu le choix de d . Pour $x \in [\varepsilon, 1 - \varepsilon]$, $T > 0$ et $\eta > 0$, nous appelons $\Phi_{x, T}(\eta)$ l'ensemble des trajectoires $\phi : [0, T] \rightarrow [-\varepsilon, \varepsilon]$ issues de x et ayant une action inférieure à η . Supposons maintenant que

$$d_T(Y, \Phi_{x, T}(V(\rho^*, \varepsilon) + 0.65d)) < \gamma/2.$$

Alors il existe $\phi \in \Phi_{x, T}(V(\rho^*, \varepsilon) + 0.65d)$ à distance inférieure à $\gamma/2$ de Y . Soit le temps t_γ associé à la trajectoire ϕ est égal à T , et dans ce cas, les événements $\tau \leq T, Y(\tau) < \varepsilon$ ne sont pas réalisés tous les deux. Soit les événements $\tau \leq T, Y(\tau) < \varepsilon$ sont réalisés, et dans ce cas, $t_\gamma < T$ et $d_{t_\gamma}((\phi)_{0 \leq t \leq t_\gamma}, \Phi) \leq \delta/C$. À l'aide de l'inégalité triangulaire, nous en déduisons qu'il existe $x \in [\rho^* - \varepsilon - \delta/C, \rho^* + \varepsilon + \delta/C]$ tel que

$$\sup_{\theta_\varepsilon \leq t \leq T_\varepsilon} |Y(t) - \phi^x(t - \theta_\varepsilon)| \leq \delta/2.$$

Nous avons alors

$$\sup_{\theta_\varepsilon \leq t \leq T_\varepsilon} |Y(t) - \phi^{\rho^* - \varepsilon}(t - \theta_\varepsilon)| \leq \delta.$$

Nous concluons que

$$\begin{aligned} P_x(\tau \leq T, Y(\tau) < \varepsilon, \sup_{\theta_\varepsilon \leq t \leq T_\varepsilon} |Y(t) - \phi^{\rho^* - \varepsilon}(t)| > \delta) \\ \leq P_x(d_T(Y, \Phi_{x, T}(V(\rho^*, \varepsilon) + 0.65d)) \geq \gamma/2) \\ \leq \exp(-m(V(\rho^*, \varepsilon) + 2d/3)), \end{aligned}$$

comme voulu. \square

Démonstration du théorème. D'après les deux lemmes précédents, pour tout $x \in [\rho^* - \gamma, \rho^* - \gamma/2]$, nous avons,

$$P_x \left(Y(\tau) < \varepsilon, \sup_{\theta_\varepsilon \leq t \leq T_\varepsilon} |Y(t) - \phi^{\rho^* - \varepsilon}(t)| > \delta \right) \leq P_x(Y(\tau) < \varepsilon) e^{-md/3}.$$

Nous introduisons les suites de temps d'arrêts suivantes : soit $\tau_0^* = 0$ et prenons

$$\begin{aligned} \tau_1 &= \inf \{ t \geq \tau_0^* : Y(t) < \rho^* - \gamma \}, \\ \tau_1^* &= \inf \{ t \geq \tau_1 : Y(t) < \varepsilon \text{ ou } Y(t) > \rho^* - \gamma/2 \}, \\ &\vdots \\ \tau_i &= \inf \{ t \geq \tau_{i-1}^* : Y(t) < \rho^* - \gamma \}, \\ \tau_i^* &= \inf \{ t \geq \tau_i : Y(t) < \varepsilon \text{ ou } Y(t) > \rho^* - \gamma/2 \}, \\ &\vdots \end{aligned}$$

et soit N le temps :

$$N = \inf \{ n \geq 0 : Y(\tau_n^*) < \varepsilon \}.$$

Nous avons donc,

$$\begin{aligned} &P_x \left(\sup_{\theta_\varepsilon \leq t \leq T_\varepsilon} |Y(t) - \phi^{\rho^* - \varepsilon}(t)| > \delta \right) \\ &= \sum_{n \geq 1} P_x \left(N = n, \sup_{\theta_\varepsilon \leq t \leq T_\varepsilon} |Y(t) - \phi^{\rho^* - \varepsilon}(t)| > \delta \right) \\ &= \sum_{n \geq 1} E_x \left(\mathbf{1}_{Y(\tau_1^*) > \rho^* - \gamma/2}, \dots, \mathbf{1}_{Y(\tau_{n-1}^*) > \rho^* - \gamma/2}, \right. \\ &\quad \left. P_{Y(\tau_n)} \left(Y(\tau) < \varepsilon, \sup_{\theta_\varepsilon \leq t \leq T_\varepsilon} |Y(t) - \phi^{\rho^* - \varepsilon}(t)| > \delta \right) \right) \\ &\leq \sum_{n \geq 1} E_x \left(\mathbf{1}_{Y(\tau_1^*) > \rho^* - \gamma/2}, \dots, \mathbf{1}_{Y(\tau_{n-1}^*) > \rho^* - \gamma/2}, \mathbf{1}_{Y(\tau_n^*) < \varepsilon} \right) e^{-md/3} = e^{-md/3}. \end{aligned}$$

Soit maintenant $x \in [\varepsilon, 1]$. Nous avons

$$\begin{aligned} &P_x \left(\sup_{\theta_\varepsilon \leq t \leq T_\varepsilon} |Y(t) - \phi^{\rho^* - \varepsilon}(t)| > \delta \right) \leq \\ &\quad P_x(Y(\tau) < \varepsilon) + P_x \left(Y(\tau) > \rho^* - \gamma/2, \sup_{\theta_\varepsilon \leq t \leq T_\varepsilon} |Y(t) - \phi^{\rho^* - \varepsilon}(t)| > \delta \right). \end{aligned}$$

Pour tout $x > \rho^* - \gamma/2$, la première probabilité vaut 0, et pour $x \in [\varepsilon, \rho^* - \gamma/2]$, d'après la loi des grands nombres trajectorielle que nous avons déduit du

corollaire 22.3.10, la première probabilité tend vers 0. Grâce à la propriété de Markov forte appliquée au temps

$$\tau' = \inf \{ t \geq 0 : Y(t) \in [\rho^* - \gamma, \rho^* - \gamma/2] \},$$

nous concluons que

$$\begin{aligned} P_x \left(Y(\tau) > \rho^* - \gamma/2, \sup_{\theta_\varepsilon \leq t \leq T_\varepsilon} |Y(t) - \phi^{\rho^* - \varepsilon}(t)| > \delta \right) = \\ E_x \left(\mathbf{1}_{Y(\tau) > \rho^* - \gamma/2} P_{Y(\tau')} \left(\sup_{\theta_\varepsilon \leq t \leq T_\varepsilon} |Y(t) - \phi^{\rho^* - \varepsilon}(t)| > \delta \right) \right). \end{aligned}$$

Nous avons montré précédemment que cette dernière quantité tend vers 0. \square

Chapitre 25

Conclusion

Nous avons montré que dans le cas du paysage à un pic, le modèle de Moran présente un phénomène de seuil d'erreur et une distribution de quasi-espèce. De plus, la distribution de la quasi-espèce pour le modèle de Moran est la même que pour le modèle d'Eigen, $\mathcal{Q}(\sigma, \alpha)$.

Comme pour le modèle de Wright-Fisher, l'étude du processus de Moran met en évidence l'impact de la taille de la population dans la quasi-espèce. Afin de pouvoir créer une quasi-espèce, la taille de la population doit être assez grande. La courbe critique exprimant cette condition n'a pas la même expression que la courbe critique pour le modèle de Wright-Fisher. Ceci montre que le seuil d'erreur dans les modèles à population finie, est sensible au mécanisme de production.

Nous avons en outre développé un principe de grandes déviations afin de traiter le cas des fonctions de fitness dépendants de la classe de Hamming. Comme application du principe de grandes déviations nous avons montré que le temps moyen que le modèle de Moran reste près de la distribution de la quasi-espèce est d'ordre exponentiel en la taille de la population, non seulement dans le cas du paysage à un pic, mais aussi pour des paysages de fitness dépendants de la classe de Hamming et constants à partir d'un certain rang. Nous avons aussi exploité le principe de grandes déviations pour décrire la trajectoire de disparition de la quasi-espèce dans le cas du paysage à un pic. Ces applications nous donnent une aperçu de l'utilité du principe de grandes déviations. Nous espérons pouvoir nous en servir de ce principe de grandes déviations pour obtenir

nombreux d'autres résultats intéressants.

Annexe A

Lemmes sur les équations différentielles linéaires

Nous donnons ici quelques lemmes sur les équations différentielles linéaires, et sur leur comportement asymptotique en temps long.

Lemme A.1. Soient $\alpha : [0, +\infty[\rightarrow [0, +\infty[$ et $\beta : [0, +\infty[\rightarrow \mathbb{R}$ des fonctions Lipschitz et soit $(z(t), t \geq 0)$ la solution de l'équation différentielle

$$z'(t) = \alpha(t) + \beta(t)z(t).$$

Si $z(0) \geq 0$ alors $z(t) \geq 0$ pour tout $t \geq 0$.

Démonstration. La trajectoire $(z(t), t \geq 0)$ est continue. S'il existe $t^* \geq 0$ tel que $z(t^*) = 0$, alors

$$z'(t^*) = \alpha(t^*) \geq 0,$$

et donc $z(t) \geq 0$ pour tout $t \geq 0$. □

Lemme A.2. Soient $\alpha, \tilde{\alpha} : [0, +\infty[\rightarrow [0, +\infty[$ et $\beta, \tilde{\beta} : [0, +\infty[\rightarrow \mathbb{R}$ des fonctions Lipschitz vérifiant

$$\forall t \geq 0, \quad \alpha(t) \leq \tilde{\alpha}(t), \quad \beta(t) \leq \tilde{\beta}(t).$$

Soient $(y(t), t \geq 0)$ et $(z(t), t \geq 0)$ les solutions des équations différentielles

$$y'(t) = \alpha(t) + \beta(t)y(t), \quad z'(t) = \tilde{\alpha}(t) + \tilde{\beta}(t)z(t).$$

Si $z(0) \geq y(0) \geq 0$ alors $z(t) \geq y(t)$ pour tout $t \geq 0$.

Démonstration. Nous avons

$$z'(t) - y'(t) = \tilde{\alpha}(t) - \alpha(t) + (\tilde{\beta}(t) - \beta(t))z(t) + \tilde{\beta}(t)(z(t) - y(t)).$$

D'après le lemme précédent, $z(t) \geq 0$ pour tout $t \geq 0$. Et donc, en appliquant encore le lemme précédent, $z(t) - y(t) \geq 0$ pour tout $t \geq 0$. \square

Lemme A.3. Soient $\alpha, \beta : [0, +\infty[\rightarrow [0, +\infty[$ des fonctions Lipschitz, et supposons qu'il existe $\alpha^*, \beta^* \in]0, +\infty[$ tels que

$$\lim_{t \rightarrow \infty} \alpha(t) = \alpha^*, \quad \lim_{t \rightarrow \infty} \beta(t) = \beta^*.$$

Soit $(y(t), t \geq 0)$ la solution de l'équation différentielle

$$y'(t) = \alpha(t) - \beta(t)y(t).$$

Alors, pour toute condition initiale $y(0) \in \mathbb{R}$,

$$\lim_{t \rightarrow \infty} y(t) = \frac{\alpha^*}{\beta^*}.$$

Démonstration. Soit $\varepsilon > 0$ tel que $\alpha^* - \varepsilon, \beta^* - \varepsilon > 0$. Soit $T \geq 0$ assez grand pour que

$$\forall t \geq T, \quad |\alpha(t) - \alpha^*| < \varepsilon, \quad |\beta(t) - \beta^*| < \varepsilon.$$

Soient $(\underline{y}(t), t \geq 0)$ et $(\bar{y}(t), t \geq 0)$ les solutions des équations différentielles

$$\underline{y}'(t) = (\alpha^* - \varepsilon) - (\beta^* + \varepsilon)\underline{y}(t), \quad \bar{y}'(t) = (\alpha^* + \varepsilon) - (\beta^* - \varepsilon)\bar{y}(t),$$

avec $\underline{y}(0) = \bar{y}(0) = y(T)$. D'après le lemme précédent, pour tout $t \geq 0$,

$$\underline{y}(t) \leq y(T+t) \leq \bar{y}(t).$$

Or, $\underline{y}(t)$ et $\bar{y}(t)$ convergent :

$$\lim_{t \rightarrow \infty} \underline{y}(t) = \frac{\alpha^* - \varepsilon}{\beta^* + \varepsilon}, \quad \lim_{t \rightarrow \infty} \bar{y}(t) = \frac{\alpha^* + \varepsilon}{\beta^* - \varepsilon}.$$

Nous en déduisons que

$$\frac{\alpha^* - \varepsilon}{\beta^* + \varepsilon} \leq \liminf_{t \rightarrow \infty} y(t) \leq \limsup_{t \rightarrow \infty} y(t) \leq \frac{\alpha^* + \varepsilon}{\beta^* - \varepsilon}.$$

Nous faisons tendre ε vers 0 et nous obtenons le résultat voulu. \square

Lemme A.4 (Gronwall). Soit $T > 0$, et soit $f(t)$, $t \in [0, T]$, une fonction positive qui vérifie l'inégalité suivante : il existe $C, \alpha > 0$ tels que

$$\forall t \in [0, T], \quad f(t) \leq C + \alpha \int_0^t f(s) ds.$$

Alors pour tout $t \in [0, T]$,

$$f(t) \leq Ce^{\alpha t}.$$

Démonstration. Nous avons par hypothèse

$$\alpha f(t) \left(C + \alpha \int_0^t f(s) ds \right)^{-1} \leq \alpha.$$

Nous intégrons des deux côtés pour obtenir

$$\ln \left(C + \alpha \int_0^t f(s) ds \right) - \ln C \leq \alpha t.$$

D'où

$$f(t) \leq C + \alpha \int_0^t f(s) ds \leq Ce^{\alpha t},$$

comme voulu. □

Annexe B

Lumping

Soit \mathcal{E} un ensemble fini ou dénombrable, et soit $(X_n)_{n \geq 0}$ une chaîne de Markov à valeurs dans \mathcal{E} de matrice de transition p . Le résultat de "lumping" classique permet de réduire la taille de l'espace d'états de la chaîne de Markov $(X_n)_{n \geq 0}$. Ce résultat peut être trouvé dans la section 6.3 du livre de Kemeny et Snell [50], formulé et démontré pour des chaînes de Markov à espace d'états fini. La démonstration pour les chaînes de Markov à espace d'états dénombrable est littéralement la même. Soit $(\mathcal{E}_1, \mathcal{E}_2, \dots)$ une partition finie ou dénombrable de \mathcal{E} . Définissons la fonction $f : \mathcal{E} \rightarrow \{1, 2, \dots\}$ par

$$\forall x \in \mathcal{E}, \quad f(x) = \sum_{k \geq 1} k 1_{x \in \mathbb{D}^{k+1}}.$$

Nous disons que la chaîne de Markov est réductible vis-à-vis de la partition $(\mathcal{E}_1, \mathcal{E}_2, \dots)$, si pour toute distribution initiale π de X_0 , le processus $(f(X_n))_{n \geq 0}$ est une chaîne de Markov à valeurs dans $\{1, 2, \dots\}$, dont les probabilités de transition ne dépendent pas de π .

Théorème B.1 (Lumping). *La chaîne de Markov est réductible vis-à-vis de la partition $(\mathcal{E}_1, \mathcal{E}_2, \dots)$ si et seulement si la condition suivante est satisfaite :*

$$\forall i, j \in \{1, 2, \dots\} \quad \forall x, y \in \mathcal{E}_i \quad \sum_{z \in \mathcal{E}_j} p(x, z) = \sum_{z \in \mathcal{E}_j} p(y, z).$$

Si la condition du théorème est vérifiée, nous notons $\tilde{p}(i, j)$ la valeur commune des sommes ci-dessus. Le processus $(f(X_n))_{n \geq 0}$ est donc une chaîne de Markov à valeurs dans $\{1, 2, \dots\}$ de matrice de transition \tilde{p} .

Bibliographie

- [1] Domingos Alves and Jose Fernando Fontanari. Error threshold in finite populations. *Phys. Rev. E*, 57:7008–7013, 1998.
- [2] Jon P. Anderson, Richard Daifuku, and Lawrence A. Loeb. Viral error catastrophe by mutagenic nucleosides. *Annual Review of Microbiology*, 58(1):183–205, 2004.
- [3] Krishna B. Athreya and Peter E. Ney. *Branching processes*. Dover Publications, Inc., Mineola, NY, 2004. Reprint of the 1972 original [Springer, New York; MR0373040].
- [4] Ellen Baake and Wilfried Gabriel. Biological evolution through mutation, selection, and drift: An introductory review. *Ann. Rev. Comp. Phys. VII*, pages 203–264, 2000.
- [5] Chikafusa Bessho and Naoki Kuroda. A note on a more general solution of Eigen’s rate equation for selection. *Bull. Math. Biol.*, 45(1):143–149, 1983.
- [6] Christof K. Biebricher. *Darwinian Selection of Self-Replicating RNA Molecules*, pages 1–52. Springer US, Boston, MA, 1983.
- [7] Alexander S. Bratus, Artem S. Novozhilov, and Yuri S. Semenov. Linear algebra of the permutation invariant Crow-Kimura model of prebiotic evolution. *Math. Biosci.*, 256:42–57, 2014.
- [8] James J. Bull and Claus O. Wilke. Lethal Mutagenesis of Bacteria. *Genetics*, 180(2):1061–1070, 2008.
- [9] Leonard Carlitz. Permutations with prescribed pattern. *Math. Nachr.*, 58:31–53, 1973.

- [10] Raphaël Cerf. *Critical population and error threshold on the sharp peak landscape for a Moran model*. *Memoirs of the American Mathematical Society*. 2015.
- [11] Raphaël Cerf. Critical population and error threshold on the sharp peak landscape for the Wright-Fisher model. *Ann. Appl. Probab.*, 25(4):1936–1992, 2015.
- [12] Raphaël Cerf and Joseba Dalmau. A probabilistic representation of the quasispecies distribution. *arXiv preprint*, 2016.
- [13] Raphaël Cerf and Joseba Dalmau. Quasispecies on class-dependent fitness landscapes. *Bulletin of Mathematical Biology*, 78(6):1238–1258, 2016.
- [14] Raphaël Cerf and Joseba Dalmau. The distribution of the quasispecies for a Moran model on the sharp peak landscape. *Stochastic Processes and their Applications*, 126(6):1681 – 1709, 2016.
- [15] Shane Crotty, Craig E. Cameron, and Raul Andino. RNA virus error catastrophe: Direct molecular test by using ribavirin. *Proceedings of the National Academy of Sciences*, 98(12):6895–6900, 2001.
- [16] Joseba Dalmau. Convergence of a Moran model to Eigen’s quasispecies model. *arXiv preprint*, 2014.
- [17] Joseba Dalmau. The distribution of the quasispecies for the Wright–Fisher model on the sharp peak landscape. *Stochastic Processes and their Applications*, 125(1):272 – 293, 2015.
- [18] Joseba Dalmau. Distribution of the quasispecies for a Galton-Watson process on the sharp peak landscape. *J. Applied Probability*, 53:606–613, 2016.
- [19] Lloyd Demetrius, Peter Schuster, and Karl Sigmund. Polynucleotide evolution and branching processes. *Bulletin of Mathematical Biology*, 47(2):239 – 262, 1985.
- [20] Narendra M. Dixit, Piyush Srivastava, and Nisheeth K. Vishnoi. A finite population model of molecular evolution: theory and computation. *J. Comput. Biol.*, 19(10):1176–1202, 2012.
- [21] Esteban Domingo. Viruses at the Edge of Adaptation. *Virology*, 270(2):251–253, 2000.

- [22] Esteban Domingo. Quasispecies theory in virology. *Journal of Virology*, 76(1):463–465, 2002.
- [23] Esteban Domingo, Christof Biebricher, Manfred Eigen, and John J. Holland. *Quasispecies and RNA virus evolution: principles and consequences*. Landes Bioscience, Austin, Tex., 2001.
- [24] Esteban Domingo, Mercedes Dávila, and Juan Ortín. Nucleotide sequence heterogeneity of the {RNA} from a natural population of foot-and-mouth-disease virus. *Gene*, 11(3–4):333–346, 1980.
- [25] Esteban Domingo, Encarnación Martínez-Salas, Francisco Sobrino, Juan Carlos de la Torre, Agustín Portela, Juan Ortín, Cecilio López-Galindez, Pilar Pérez-Breña, Nieves Villanueva, Rafael Nájera, Scott Van-dePol, David Steinhauer, Nicholas DePolo, and John Holland. The quasispecies (extremely heterogeneous) nature of viral {RNA} genome populations: biological relevance — a review. *Gene*, 40(1):1–8, 1985.
- [26] Esteban Domingo, Donna Sabo, Tadatsugu Taniguchi, and Charles Weissmann. Nucleotide sequence heterogeneity of an RNA phage population. *Cell*, 13(4):735–744, 1978.
- [27] John W. Drake and John J. Holland. Mutation rates among rna viruses. *Proceedings of the National Academy of Sciences*, 96(24):13910–13913, 1999.
- [28] Manfred Eigen. Self-organization of matter and the evolution of biological macromolecules. *Naturwissenschaften*, 58(10):465–523, 1971.
- [29] Manfred Eigen. Error catastrophe and antiviral strategy. *Proceedings of the National Academy of Sciences*, 99(21):13374–13376, 2002.
- [30] Manfred Eigen, John McCaskill, and Peter Schuster. The molecular quasispecies. *Advances in Chemical Physics*, 75:149–263, 1989.
- [31] Manfred Eigen and Peter Schuster. The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle. *Naturwissenschaften*, 64(11):541 – 565, 1977.
- [32] Manfred Eigen and Peter Schuster. The hypercycle. A principle of natural self-organization. Part B: The abstract hypercycle. *Naturwissenschaften*, 65(1):7 – 41, 1978.

- [33] Manfred Eigen and Peter Schuster. The hypercycle. A principle of natural self-organization. Part C: The realistic hypercycle. *Naturwissenschaften*, 65(7):341 – 369, 1978.
- [34] Ky Fan. Minimax theorems. *Proc. Nat. Acad. Sci. U. S. A.*, 39:42–47, 1953.
- [35] Ronald A. Fisher. On the dominance ratio. *Proceedings of the Royal Society of Edinburgh*, 42:321–341, 1922.
- [36] Ronald A. Fisher. *The Genetical Theory of Natural Selection*. Oxford University Press, Oxford, 1930.
- [37] Edward J. Fox and Lawrence A. Loeb. Lethal Mutagenesis: Targeting the Mutator Phenotype in Cancer. *Seminars in Cancer Biology*, 20(5):353–359, 2010. Mutator Phenotype in Cancer: Origins and Consequences.
- [38] 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, Heidelberg, third edition, 2012. Translated from the 1979 Russian original by Joseph Szücs.
- [39] Francis Galton and Henry W. Watson. On the probability of the extinction of families. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 4:138–144, 1875.
- [40] John B. S. Haldane. A mathematical theory of natural and artificial selection, part v: Selection and mutation. *Mathematical Proceedings of the Cambridge Philosophical Society*, 23:838–844, 7 1927.
- [41] John B. S. Haldane. The origin of life. *Rationalist Annual*, (148):3–10, 1929.
- [42] Theodore E. Harris. *The theory of branching processes*. Die Grundlehren der Mathematischen Wissenschaften, Bd. 119. Springer-Verlag, Berlin; Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.
- [43] Gregory R. Hart and Andrew L. Ferguson. Error catastrophe and phase transition in the empirical fitness landscape of hiv. *Phys. Rev. E*, 91:032705, Mar 2015.
- [44] John Holland, Katherine Spindler, Frank Horodyski, Elizabeth Grabau, Stuart Nichol, and Scott VandePol. Rapid evolution of RNA genomes. *Science*, 215(4540):1577–1585, 1982.

- [45] John J. Holland, Elizabeth A. Grabau, Charlotte L. Jones, and Bert L. Semler. Evolution of multiple genome mutations during long-term persistent infection by vesicular stomatitis virus. *Cell*, 16(3):495–504, 1979.
- [46] Edward C. Holmes and Andrés Moya. Is the quasispecies concept relevant to rna viruses? *Journal of Virology*, 76(1):460–462, 2002.
- [47] Peter Jagers. *Branching processes with biological applications*. Wiley-Interscience [John Wiley & Sons], London-New York-Sydney, 1975. Wiley Series in Probability and Mathematical Statistics—Applied Probability and Statistics.
- [48] Billy L. Jones. Analysis of Eigen’s equations for selection of biological molecules with fluctuating mutation rates. *Bull. Math. Biology*, 39(3):311–316, 1977.
- [49] Billy L. Jones, Richard H. Enns, and Sadanand S. Rangnekar. On the theory of selection of coupled macromolecular systems. *Bulletin of Mathematical Biology*, 38(1):15–28, 1976.
- [50] John G. Kemeny and J. Laurie Snell. *Finite Markov chains*. Springer-Verlag, New York-Heidelberg, 1976. Reprinting of the 1960 original, Undergraduate Texts in Mathematics.
- [51] Thomas G. Kurtz. Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Probability*, 7:49–58, 1970.
- [52] Adam S. Luring and Andino Raul. Quasispecies Theory and the Behavior of RNA Viruses. *PLoS Pathog*, 6(7):1–8, 07 2010.
- [53] Ira Leuthäusser. An exact correspondence between Eigen’s evolution model and a two-dimensional Ising system. *J. Chem. Phys.*, 84(3):1884–1885, 1986.
- [54] Ira Leuthäusser. Statistical mechanics of Eigen’s evolution model. *J. Statist. Phys.*, 48(1-2):343–360, 1987.
- [55] Lawrence A. Loeb, John M. Essigmann, Farhad Kazazi, Jue Zhang, Karl D. Rose, and James I. Mullins. Lethal mutagenesis of hiv with mutagenic nucleoside analogs. *Proceedings of the National Academy of Sciences*, 96(4):1492–1497, 1999.

- [56] John McCaskill. A stochastic theory of macromolecular evolution. *Biological Cybernetics*, 50:63–73, 1984.
- [57] Patrick A. P. Moran. Random processes in genetics. *Proceedings of the Cambridge Philosophical Society*, 54:60, 1958.
- [58] Patrick A. P. Moran. Global stability of genetic systems governed by mutation and selection. *Math. Proc. Cambridge Philos. Soc.*, 80(2):331–336, 1976.
- [59] Patrick A. P. Moran. Global stability of genetic systems governed by mutation and selection. II. *Math. Proc. Cambridge Philos. Soc.*, 81(3):435–441, 1977.
- [60] Fabio Musso. A stochastic version of the Eigen model. *Bull. Math. Biol.*, 73(1):151–180, 2011.
- [61] Martin Nowak and Peter Schuster. Error thresholds of replication in finite populations mutation frequencies and the onset of muller’s ratchet. *Journal of Theoretical Biology*, 137(4):375 – 395, 1989.
- [62] Aleksandr Ivanovich Oparin. *The origin of life on the earth / Translated from the Russian by Ann Synge*. New York: Academic Press, 1957. <http://www.biodiversitylibrary.org/bibliography/4528>.
- [63] Jeong-Man Park, Enrique Muñoz, and Michael W. Deem. Quasispecies theory for finite populations. *Phys. Rev. E*, 81:011902, 2010.
- [64] R. Tyrrell Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.
- [65] David B. Saakian. A new method for the solution of models of biological evolution: derivation of exact steady-state distributions. *J. Stat. Phys.*, 128(3):781–798, 2007.
- [66] David B. Saakian, Christof K. Biebricher, and Chin-Kun Hu. Lethal mutants and truncated selection together solve a paradox of the origin of life. *PLoS ONE*, 6(7):1–12, 07 2011.
- [67] David B. Saakian, Michael W. Deem, and Chin-Kun Hu. Finite population size effects in quasispecies models with single-peak fitness landscape. *Europhysics Letters*, 98(1):18001, 2012.

- [68] David B. Saakian and Chin-Kun Hu. Exact solution of the eigen model with general fitness functions and degradation rates. *Proceedings of the National Academy of Sciences of the United States of America*, 103(13):4935–4939, 2006.
- [69] David Seifert, Francesca Di Giallonardo, Karin J. Metzner, Huldrych F. Günthard, , and Niko Beerenwinkel. A framework for inferring fitness landscapes of patient-derived viruses using quasispecies theory. *Genetics*, 199(1):192–203, 2015.
- [70] Yuri S. Semenov, Alexander S. Bratus, and Artem S. Novozhilov. On the behavior of the leading eigenvalue of Eigen’s evolutionary matrices. *Math. Biosci.*, 258:134–147, 2014.
- [71] Yuri S. Semenov and Artem S. Novozhilov. Exact solutions for the selection–mutation equilibrium in the crow–kimura evolutionary model. *Mathematical Biosciences*, 266:1 – 9, 2015.
- [72] Yuri S. Semenov and Artem S. Novozhilov. On eigen’s quasispecies model, two-valued fitness landscapes, and isometry groups acting on finite metric spaces. *Bulletin of Mathematical Biology*, 78(5):991–1038, 2016.
- [73] Eugene Seneta. *Nonnegative matrices and Markov chains*. Springer Series in Statistics. Springer-Verlag, New York, second edition, 1981.
- [74] Vladimir Shevelev. Number of permutations with prescribed up-down structure as a function of two variables. *Integers*, 12(4):529–569, 2012.
- [75] Ricard V. Solé and Thomas S. Deisboeck. An error catastrophe in cancer? *Journal of Theoretical Biology*, 228(1):47–54, 2004.
- [76] Jörg Swetina and Peter Schuster. Self–replication with errors. a model for polynucleotide replication. *Biophys. Chem.*, 16(4):329–45, 1982.
- [77] Pedro Tarazona. Error thresholds for molecular quasispecies as phase transitions: From simple landscapes to spin-glass models. *Phys. Rev. A*, 45:6038–6050, Apr 1992.
- [78] Colin J. Thompson and John L. McBride. On Eigen’s theory of the self-organization of matter and the evolution of biological macromolecules. *Math. Biosci.*, 21:127–142, 1974.

- [79] Kushal Tripathi, Rajesh Balagam, Nisheeth K. Vishnoi, and Narendra M. Dixit. Stochastic simulations suggest that HIV-1 survives close to its error threshold. *preprint*, 2012.
- [80] Marco Vignuzzi, Jeffrey K. Stone, and Raul Andino. Ribavirin and lethal mutagenesis of poliovirus: molecular mechanisms, resistance and biological implications. *Virus Research*, 107(2):173–181, 2005. Virus Entry into Error Catastrophe as a New Antiviral Strategy.
- [81] Edward D. Weinberger. *A Stochastic Generalization of Eigen's Theory of Natural Selection*. Ph.D. Dissertation. The Courant Institute of Mathematical Sciences, New York University, 1987.
- [82] Alexander D. Wentzell. Rough Limit Theorems on Large Deviations for Markov Stochastic Processes. I. *Theory of Probability & Its Applications*, 21(2):227–242, 1977.
- [83] Alexander D. Wentzell. Rough Limit Theorems on Large Deviations for Markov Stochastic Processes, II. *Theory of Probability & Its Applications*, 21(3):499–512, 1977.
- [84] Alexander D. Wentzell. *Limit theorems on large deviations for Markov stochastic processes*, volume 38 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1990. Translated from the Russian.
- [85] Claus Wilke. Quasispecies theory in the context of population genetics. *BMC Evolutionary Biology*, 5:1–8, 2005.
- [86] Sewall Wright. Evolution in mendelian populations. *Genetics*, 16(2):97–159, 1931.

Index

- (\mathcal{R}) , 64
 (\mathcal{S}) , 42, 88
 $(\mathcal{S}_\infty^{\text{sp}})$, 44, 89
 (\mathcal{S}_∞) , 43, 89
 (\mathcal{S}_f) , 61
 (\mathcal{J}_1) , 189
 (\mathcal{J}_2) , 207
 C_x , 260
 C_y^∞ , 261
 D_n , 168
 D_t , 178
 $E(k, i)$, 49
 $G(x, z)$, 270
 G^∞ , 261
 H^∞ , 261
 $I(p, t)$, 108
 $I_K(p, t)$, 131
 $I_K(r, \xi, \beta, t)$, 133
 I_T , 261
 I_ℓ , 132
 $\text{Li}_s(z)$, 49
 M_H , 13, 35
 M_H^ℓ , 210
 M_H^{K+1} , 210
 M_∞ , 43
 $N^k(x)$, 108, 163
 O_n^ℓ , 119, 181, 190
 O_n^{K+1} , 181, 194
 O_n , 169
 $P_{x^0}^z$, 272
 R_n , 113, 176
 S_n , 176
 T_ρ , 262
 $U(x)$, 312
 $U_{n,l}$, 176
 V , 139
 $V(x, y)$, 305
 V_1 , 133
 V_L , 136
 $Y(t)$, 259
 Y^ρ , 262
 Z_n^ℓ , 193
 Z_n^θ , 185
 Z_n^{K+1} , 202
 \mathcal{A}^ℓ , 32
 $(\mathcal{A}^\ell)^m$, 107
 Δ^{k+1} , 260
 $\Lambda^*(t)$, 241
 Φ'_O , 198
 Φ_H , 178
 Φ_O , 178
 Φ_O^{K+1} , 182, 200
 Φ_O^ℓ , 182, 191
 $\Phi_T(\cdot)$, 280
 $\Phi_{y^0, T}(\cdot)$, 261
 Ψ_H , 114
 Ψ_O , 114
 Ψ_O^{K+1} , 120
 Ψ_O^ℓ , 120
 $\Sigma^{\ell+1}$, 260

- $(\Theta_n^\ell)_{n \geq 0}$, 129
 $(Z_n^\theta)_{n \geq 0}$, 122
 \mathcal{A} , 32
 $\mathcal{B}(t)$, 133
 \mathcal{D}_ρ^{k+1} , 261
 \mathcal{D}_K , 131
 \mathcal{M}'_H , 196
 \mathcal{M}_H , 53, 177
 \mathcal{M}_H^{K+1} , 196
 \mathcal{M}_H , 113
 \mathcal{N} , 120, 191, 200
 \mathcal{N}_K , 182
 \mathcal{O} , 112, 169
 \mathcal{O}_H , 112, 169
 $\mathcal{Q}(\sigma, \alpha)$, 47
 \mathcal{R} , 113, 177
 \mathcal{R}' , 197
 \mathcal{S}^N , 37
 \mathcal{S}_H , 113, 177
 \mathcal{S}_O , 114, 178
 \mathcal{T}^ℓ , 192
 \mathcal{T}^θ , 121
 \mathcal{T}^{K+1} , 202
 \mathcal{T}_K^θ , 184
 \mathcal{V} , 260
 \mathcal{W}^* , 120, 191, 200
 \mathcal{W}_K^* , 182
 \mathbb{D}^{K+1} , 122, 185
 \mathbb{H} , 111, 168
 δ_i , 213
 $\delta_i(x)$, 222
 ℓ , 32
 γ_i , 213
 $\gamma_i(x)$, 222
 κ , 32
 \mathcal{D}^{k+1} , 260
 $\mathcal{S}^{\ell+1}$, 260
 $\mathcal{X}(t)$, 259
 μ_O^ℓ , 124, 186, 204
 μ^θ , 125
 μ_O^{K+1} , 124, 186, 204
 ν_O , 124, 171, 186, 204
 ν_K , 124
 ν_k , 170
 $\omega_{\text{exit}}^\theta$, 185
 ω_θ^i , 184
 $\overline{\Phi}_O$, 198
 $\phi(\beta, \varepsilon, \chi, \rho)$, 229
 $\phi(\beta, \varepsilon, \rho)$, 213
 $\pi(i)$, 211
 $\pi(t_1, t_2)$, 271
 π_ℓ , 119, 182, 190
 π_k , 260
 π_{K+1} , 119, 182, 200
 $\mathcal{P}_{\ell+1}^m$, 112, 169
 \preceq , 115, 180
 $\psi(\alpha)$, 108, 163
 $\rho(\beta, \varepsilon)$, 213
 $\rho(\beta, \varepsilon, \chi)$, 229
 $\rho^*(\alpha)$, 108
 σ , 33
 τ , 128, 205, 259
 τ^* , 128, 205
 τ_K , 187
 τ_K^* , 187
 τ_0 , 128, 187, 206
 $\overline{\Phi}_O$, 190
 d_T^τ , 284
 d_H , 33
 d_T , 262
 o_{enter}^ℓ , 120, 191
 o_{exit}^ℓ , 120, 191
 o_{enter}^{K+1} , 120, 200
 o_{exit}^{K+1} , 120, 200
 q , 33
 w^* , 32
 $x(j \leftarrow u)$, 161
 $x_\nu(t)$, 33

z^ℓ , 128, 193, 206

z^θ , 128

z^{K+1} , 128, 202, 206

Résumé. Le concept de quasi-espèce, introduit par Manfred Eigen dans les années 70, décrit l'état d'équilibre d'une population subissant des forces de mutation et sélection. La plupart de modèles classiques présentant un phénomène de quasi-espèce sont déterministes et considèrent une population de taille infinie. L'objectif de cette thèse est d'étudier plusieurs modèles stochastiques, dont la taille de la population est finie, afin de montrer que le phénomène de la quasi-espèce est aussi présent dans ces modèles comme dans les modèles déterministes. Nous étudions en détail les modèles de Galton–Watson, Wright–Fisher et Moran. Nous confirmons que les trois modèles présentent un phénomène de quasi-espèce, et que la distribution de la quasi-espèce est la même pour les trois modèles, ainsi que pour le modèle original d'Eigen. De plus, nous décrivons explicitement la distribution de cette quasi-espèce pour le paysage à un pic, ainsi que pour des fonctions de fitness qui ne dépendent que de la distance de Hamming au génotype le plus adapté.

Abstract. Manfred Eigen introduced the concept of quasispecies in the early 70s, in order to describe the steady-state distribution of a population subject to mutation and selection forces. Most classical models showing a quasispecies phenomenon are deterministic and deal with a population of an infinite size. The aim of this thesis is to study several stochastic and finite-population models, and to show that a quasispecies phenomenon arises in these models too. We study in detail the Galton–Watson, Wright–Fisher and Moran models. We confirm that a quasispecies phenomenon is present in the three models, and that the distribution of this quasispecies is common to all three models as well as to Eigen's original model. Moreover, we describe explicitly the distribution of the quasispecies for the sharp peak landscape, as well as for fitness functions which depend only on the Hamming distance to the fittest genotype.