

Mémoire de Magistère

Bastien Mallein

18 octobre 2011

Table des matières

1	Curriculum Vitae	3
I	Introduction au Domaine de la recherche	5
2	Les marches aléatoires branchantes	6
1	Introduction aux marches aléatoires branchantes	8
2	Méthode de décomposition en épine dorsale et Lemme de regroupement	11
3	La position de l'individu le plus à droite	15
4	Marche aléatoire branchante avec sélection	18
II	Mémoire de M2	20
3	Divergence globale des processus coalescents spatiaux	21
1	À propos des processus coalescents	22
2	Le cas des coalescents non-spatiaux	26
3	Processus coalescent spatial	33
4	Résultats utiles sur les coalescents spatiaux	35
5	Divergence du coalescent de Kingman spatial	37
6	Divergence générale d'un processus Λ -coalescent	45
7	Le cas du Beta-coalescent	49
8	Conclusion	56
9	Annexe 1 : Preuve du théorème de Schweinsberg	57
4	Vulgarisation pour le site Culturemath : Le coalescent de Kingman	63
1	Introduction du coalescent de Kingman	65
2	Construction du coalescent de Kingman	70
3	Descente de l'infini du coalescent de Kingman	73
4	Taille des groupes de cousins	77
5	Mutation neutre et classification phylogénique	81
6	Conclusion	83
7	Preuve du Théorème 2.1	83

III	Stage de M1	85
5	Convergence de certain processus de particules vers des processus de Dawson-Watanabe	86
1	Introduction	88
2	Dérivée de Radon-Nikodým des processus de particules	92
3	Quelques calculs relatif à la marche aléatoire branchante	95
4	Convergence de la marche aléatoire branchante avec dérive	97
5	Convergence du processus de contact	102
6	Convergence pour le modèle du voteur	107
7	Le modèle de Lotka-Volterra, une modification du modèle du voteur	108
8	Conclusion	111
9	Annexe	111
IV	Mémoire de M1	123
6	Théorèmes limites pour certaines classes d'urnes aléatoires	124
1	Introduction	126
2	Généralités sur les urnes	126
3	Solution générale pour les urnes sacrificielles	135
4	Urnas semi-sacrificielles	138
5	Urnas triangulaires	144
6	Conclusion	152
7	Vulgarisation pour le site CultureMath : L'Urne de Pólya	153
1	Introduction	155
2	Série génératrice	156
3	Urne aléatoire	160
4	Conclusion	167

Chapitre 1

Curriculum Vitae

Année 2010-2011, troisième année de magistère

Cours d'École Doctorale : Arbres aléatoires et Coalescences (J.-F. Le Gall, J. Bertoin et É. Pardoux) :

- Processus de coalescence échangeables (J. Bertoin) ;
- Limites continues d'arbres et de graphes aléatoires (J.-F. Le Gall).

Inscription administrative au M2 (Probabilités et statistiques) à Paris 11 (Orsay). Cours de M2 suivis à Paris 6 (Jussieu) :

- Théorèmes limites et grandes déviations (Z. Shi) ;
- Processus de Markov (I. Kourkova) ;
- Introduction aux processus de Lévy (J. Bertoin) ;
- Fonctionnelles exponentielles de processus de Lévy (M. Yor) ;
- Temps locaux du mouvement brownien et théorie des excursions (M. Yor) ;
- Méthodes probabilistes pour les EDP (D. Talay et M. Bossy) ;
- Dynamique des populations structurées (S. Méléard, V. Bansaye et V.C. Tran) ;
- Processus de branchement et équation FKPP (J. Berestycki) ;
- Stochastic partial differential equations (L. Zambotti) ;
- Mémoire de M2 sur les processus coalescents spatiaux, sous la direction de Jean Bertoin.

Cours de M1 suivis à l'ENS :

- EDP (N. Burq) ;
- Groupe de travail Estimation de densité (G. Biau, O. Catoni et G. Stoltz).

Année 2009-2010, deuxième année de magistère

Cours de M1 suivis à l'ENS :

- Statistiques (G. Biau) ;
- Systèmes dynamiques (V. Baladi) ;
- Groupe de travail de probabilités (T. Bodinau et P. Marchal).

Inscription pédagogique au M2 (Probabilités et statistiques) à Paris 11. Cours de M2 suivis à Paris 11 :

- Calcul stochastique et processus de Markov (J.-F. Le Gall) ;

- Concentration de la mesure et sélection de modèles (P. Massart);
- Grandes déviations (R. Rossignol);
- Quelques modèles de la physique statistique (W. Werner).

Stage à University of British Columbia (Vancouver, Canada), sous la direction de Ed Perkins, du 1^{er} mars au 14 juillet, sur les super-processus comme limites de processus de particules.

École d'été "PIMS Summer school in Probability 2010" à Seattle :

- Exchangeable coalescents (J. Bertoin);
- Random surfaces and quantum gravity (S. Sheffield);
- Scaling Limits and SLE (G. Lawler);
- Mixing Times of Markov Chains (E. Lubetzky, Y. Peres et D. Wilson);
- Dirichlet Form Theory and Invariance Principle (Zhenqing Chen).

Année 2008-2009, première année de magistère

L3 de mathématiques, mention TB. Cours de L3 suivis à l'ENS :

- Topologie et calcul différentiel (F. Paulin);
- Logique (F. Loeser);
- Intégration et probabilités (J. Bertoin);
- Algèbre 1 (E. Ullmo);
- Analyse complexe et harmonique (L. Saint-Raymond).

M1 de mathématiques, mention TB. Cours de M1 suivis à l'ENS :

- Processus aléatoires (W. Werner);
- Géométrie différentielle (P. Pansu);
- Analyse fonctionnelle et EDP (G. Carlier);
- Exposé de maîtrise sur les urnes aléatoires, sous la direction de Philippe Marchal.

Première partie

Introduction au Domaine de la
recherche

Chapitre 2

Les marches aléatoires branchantes

Bastien MALLEIN

Sous la direction de Zhan SHI

20 juin 2011

Résumé

Nous allons détailler ici quelques-unes des propriétés les mieux connues des marches aléatoires branchantes, notamment sur la position de l'individu le plus à droite dans ce processus. Nous introduirons également la méthode de décomposition en épine dorsale, un résultat d'importance dans l'étude de ces processus. Pour finir, nous nous intéresserons à une évolution récente de ces modèles, les marches aléatoires branchantes avec sélection.

Sommaire

1	Introduction aux marches aléatoires branchantes	8
1.1	Définition du modèle	9
1.2	Notations et quantités associées aux arbres marqués	9
1.3	Quelques quantités associées au processus de points \mathcal{L}	11
2	Méthode de décomposition en épine dorsale et Lemme de regroupement	11
2.1	Martingales additives exponentielles et propriété de branchement	11
2.2	Changement de probabilité	13
2.3	Le Lemme de regroupement	14
3	La position de l'individu le plus à droite	15
3.1	Vitesse de dispersion vers la droite	15
3.2	Un premier raffinement	16
3.3	Dernières avancées	17
4	Marche aléatoire branchante avec sélection	18

1 Introduction aux marches aléatoires branchantes

De très nombreux modèles de la théorie des probabilités décrivent l'évolution de populations au cours du temps. L'un des plus connus est le processus de Galton-Watson, dont le comportement détaillé est connu de longue date [1]. Dans un arbre de Galton-Watson, chaque individu engendre à la génération suivante un nombre aléatoire A d'enfants, indépendamment du reste du processus. Nous pouvons en particulier citer les résultats très classiques suivants :

- la probabilité d'extinction d'un arbre partant d'un unique individu à l'instant initial est égal à la plus petite solution sur $[0, 1]$ de l'équation $q = \mathbb{E}(q^A)$; en particulier $q = 1$ si et seulement si $m = \mathbb{E}(A) \leq 1$;
- si $m > 1$, alors le nombre d'individus présents à l'instant n divisé par m^n tend vers une constante (aléatoire) W , de plus on a :

$$\mathbb{P}(W = 0) = q \iff \mathbb{E}(A \log^+ A) < +\infty \text{ et } \mathbb{P}(W = 0) = 1 \iff \mathbb{E}(A \log^+ A) = +\infty,$$

où $\log^+ A = \max\{\log A, 0\}$.

Cette seconde propriété donne un ordre de grandeur du nombre d'individus présents dans l'arbre en cas de survie de la population. Remarquons en particulier que dans ce processus, on a soit extinction presque sûre, soit divergence presque sûre vers $+\infty$, à vitesse exponentielle dans la plupart des cas.

Ce processus peut par la suite être enrichi de la façon suivante : à chaque individu x de l'arbre, on associe une position $V(x)$. Cette position s'obtient de la façon suivante : tous les enfants d'un individu donné sont positionnés par rapport à leur parent selon un processus de points \mathcal{L} , indépendamment du reste du processus. Cette construction autorise en particulier que la nouvelle position de deux frères soit corrélée. Ce modèle décrit ainsi une population qui se reproduit tout en se déplaçant sur la droite réelle, colonisant ainsi tout l'espace disponible. Ce processus a été introduit par J.-F. C. Kingman au début des années 1970.

L'ajout de cette composante spatiale augmente la richesse de ce processus, car peut être interprétée de nombreuses façons différentes : comme la position de l'individu au moment de sa naissance bien sûr, mais aussi comme sa date de naissance, ou son degré d'adaptation au milieu. Ces interprétations permettent de relier ce modèle à de nombreux autres, dans des domaines différents : en informatique par exemple, où de nombreux algorithmes utilisent des arbres de recherche assez semblables (L. Devoye, S. Janson, P. Flajolet); en physique théorique, selon les travaux de B. Derrida *et al.*; et bien sûr en biologie, à laquelle ce modèle emprunte son vocabulaire.

Cette composante spatiale peut même être interprétée comme le logarithme de la probabilité d'être sélectionné dans un certain ensemble compact. Ceci forme un lien avec les fractales aléatoires, auto-similaires en loi, introduites par Mandelbrot, au début des années 70 également. Dans ce modèle, on part d'un domaine compact K_0 donné divisé en un certain nombre de sous-domaines de même forme. Pour chacun d'entre eux, on tire une variable aléatoire de Bernoulli indépendante, de paramètre p_i dépendant du sous-domaine. Le compact K_1 est défini comme l'ensemble de ces sous-domaines pour lesquels la variable aléatoire vaut 1. On réalise alors le même découpage pour chacun de ces sous-domaines, et on définit par récurrence une suite de sous-domaines. Le compact $K = \bigcap K_n$ est appelé fractale aléatoire auto-similaire en loi. Ce

modèle, qui permet par exemple d'étudier les lois limites de certains modèles de la physique statistique, a donné naissance à la théorie très féconde des cascades multiplicatives. Les liens entre ces deux domaines sont nombreux, et chaque résultat obtenu dans l'un des deux se traduit automatiquement dans le second.

Nous allons maintenant passer à une définition plus précise d'une marche aléatoire branchante, avec les notations qui y sont liées.

1.1 Définition du modèle

Soit $\mathcal{L} = (l_1, \dots, l_n)$ un processus de points, c'est-à-dire une variable aléatoire donnant un certain nombre n de valeurs réelles (ce nombre n peut également être aléatoire et représente le nombre total d'enfants d'un individu). La marche aléatoire branchante associée à \mathcal{L} est le processus défini comme suit. On commence par un unique individu situé en 0 à l'instant initial. A l'instant 1, cet individu meurt et laisse place à des enfants situés sur les points d'une copie de \mathcal{L} . De même, à chaque instant $n \in \mathbb{N}$, tous les individus de la génération $n - 1$ meurent, en donnant naissance à des individus situés sur les points de copies indépendantes de \mathcal{L} , translatées de la position de leur parent.

Plusieurs variantes de ce modèle existent, comme par exemple le processus CMJ (étudié en particulier par Crump, Mode et Jagers) : dans ce cas, le processus de points \mathcal{L} est à valeurs dans \mathbb{R}^+ , et représente non plus la position de l'individu, mais l'instant auquel celui-ci naît. En d'autres termes le processus \mathcal{L} associé à un individu donné représente l'âge auquel celui-ci aura des enfants. On s'intéressera dans ce cas non plus aux individus présents à la génération n , mais aux individus présents à l'instant t . Ce processus est un analogue à temps continu du processus de Galton-Watson, avec des complications dues au fait que les générations se superposent, un « oncle » peut être plus jeune que son « neveu ».

Une autre variante est le processus de branchement à plusieurs types, dans lequel chaque individu possède un « type » σ . Un individu donne alors naissance à des enfants qui peuvent être, ou non, du même type que lui. Les individus d'un même type se reproduisent suivant la même loi, mais on peut imaginer des différences de reproduction entre les différents types. L'intérêt de ce modèle est d'étudier les interactions entre individus qui peuvent accélérer la colonisation du milieu.

Nous allons maintenant définir un certain nombre de notations relatives aux arbres qui nous permettront par la suite de décrire en détail le comportement de la marche aléatoire branchante.

1.2 Notations et quantités associées aux arbres marqués

Dans toute la suite nous noterons $\mathbb{N} = \{1, 2, \dots\}$ l'ensemble des entiers strictement positifs. La marche aléatoire branchante peut être vue comme une variable aléatoire à valeur dans l'ensemble des arbres marqués. Nous allons donc préciser un certain nombre de notations relatives aux arbres. Nous posons pour commencer $\mathcal{U} = \bigcup_{n \in \mathbb{N}} \mathbb{N}^n \cup \{\emptyset\}$ l'ensemble des suites finies d'entiers, où la suite de longueur nulle est notée \emptyset .

Dans la définition d'un arbre, une suite finie indique le chemin à parcourir depuis la racine \emptyset de l'arbre pour arriver au sommet auquel elle correspond. Le sommet (u_1, \dots, u_n) représente le

$u_n^{\text{ième}}$ fils du $u_{n-1}^{\text{ième}}$ fils ... du $u_1^{\text{ième}}$ fils de l'individu origine. On définit pour commencer un certain nombre de notations sur l'élément $u = (u_1, \dots, u_n) \in \mathcal{U}$:

- on note $|u| = n$ la longueur de la suite, avec la convention $|\emptyset| = 0$ (c'est la génération de l'individu u) ;
- pour $k \leq n$, on note $u|_k = (u_1, \dots, u_k)$ la restriction de u aux k premiers éléments, avec la convention $u|_0 = \emptyset$ ($u|_k$ est l'ancêtre de la génération k de u) ;
- on note $\Pi u = (u_1, \dots, u_{n-1})$ le prédécesseur de la suite u (Πu est le parent de l'individu u) ;
- pour $v = (v_1, \dots, v_m)$ on note $u.v = (u_1, \dots, u_n, v_1, \dots, v_m)$.

Définition 1.1. Un arbre est un sous-ensemble de \mathcal{U} vérifiant les propriétés suivantes :

- $\emptyset \in T$ (l'individu origine est bien dans l'arbre généalogique) ;
- si $u \in T$ alors $\Pi u \in T$ (le parent d'un individu est dans l'arbre généalogique) ;
- si $u \in T$ alors il existe $A_u \in \mathbb{N} \cup \{0, +\infty\}$ tel que pour tout $v \leq A_u$, on a $u.v \in T$ (on numérote les enfants à partir de 1, sans sauter de numéro).

L'ensemble des arbres est noté \mathbb{T} .

Remarquons au passage que nous autorisons un individu donné à avoir une infinité d'enfants, ou à n'en avoir aucun. Il pourra arriver que des restrictions soient nécessaires. Nous allons maintenant définir la notion d'arbre marqué, qui permet de tenir compte de la position de chaque individu.

Après avoir défini la structure d'arbre généalogique de notre population, nous allons maintenant ajouter la composante spatiale de la marche aléatoire branchante. On définit ainsi un arbre marqué : à chaque sommet x d'un arbre T on associe un réel $V(x)$. L'ensemble $\{(x, V(x)), x \in T\}$, noté également (X, V) est appelé arbre marqué. On note \mathcal{T} l'ensemble des arbres marqués.

Il pourra être utile par la suite de regarder les sous-arbres d'un arbre marqué donné. Pour cela, soit (T, V) un arbre marqué et x un sommet de T , on note (T_x, V_x) le sous-arbre marqué de (T, V) issu de x défini de la façon suivante :

- u est un sommet de T_x si $x.u$ est un sommet de T ;
- $V_x(u) = V(u) - V(x)$.

Nous pouvons alors définir la marche aléatoire branchante de façon formelle de la manière suivante : soit $\mathcal{L} = (l_1, \dots, l_n)$ un processus de points et $(\mathcal{L}^{(u)}, u \in \mathcal{U})$ des copies i.i.d. de \mathcal{L} , on pose A_u le nombre de points de $\mathcal{L}^{(u)}$; et T l'arbre défini par ces $(A_u, u \in \mathcal{U})$, c'est-à-dire :

$$x \in T \iff x = \emptyset \text{ ou } \forall 1 < k \leq |x|, x_k \leq A_{x|_{k-1}}.$$

On pose ensuite $V(\emptyset) = 0$, et, pour tout $x \in T$, tout $u \leq A_x$ correspondant à un enfant de x , on définit $V(x.u)$ comme la somme de $V(\Pi x.u)$ et de la $u^{\text{ième}}$ valeur de $\mathcal{L}^{(x.u)}$. La marche aléatoire branchante est alors définie par (T, V) . Notons que tout sous-arbre de la marche aléatoire branchante issue d'un sommet donné est une copie de la marche aléatoire branchante, indépendante du reste du processus.

Par la suite nous aurons également besoin de distinguer un chemin infini particulier dans un arbre, ce chemin sera appelé une épine. Un arbre marqué avec épine est la donnée d'un arbre marqué (T, V) et d'un élément $u \in \mathbb{N}^{\mathbb{N}}$ tel que pour tout $k \in \mathbb{N}$, $u|_k \in T$. L'ensemble des arbres marqués avec épine est alors noté $\tilde{\mathcal{T}}$.

Nous allons définir des filtrations, correspondant à différentes gammes d'informations, sur les ensembles \mathcal{T} et $\tilde{\mathcal{T}}$:

- on note $\mathcal{F}_n = \sigma((x, V(x)), ; x \in T : |x| \leq n)$ la filtration des individus présents jusqu'en l'instant n ;
- $\tilde{\mathcal{F}}_n = \mathcal{F}_n \vee \sigma(u_{|n})$ la filtration des individus présents et de l'épine jusqu'en l'instant n ;
- $\mathcal{G}_n = \sigma(V(u_{|k}); k \leq n)$ la filtration du mouvement de l'épine ;
- $\tilde{\mathcal{G}}_n = \sigma(V(u_{|k}), A_{u_{|k}}; k \leq n)$ la filtration du mouvement de l'épine et du nombre de ses enfants à chaque instant.

1.3 Quelques quantités associées au processus de points \mathcal{L}

Après avoir défini quelques quantités relatives aux arbres, nous nous intéressons au processus de points \mathcal{L} . La mesure d'intensité de ce processus est l'unique mesure μ vérifiant pour toute fonction ϕ mesurable positive :

$$\mathbb{E} \left[\sum_{|x|=1} \phi(V(x)) \right] = \int \phi d\mu.$$

La transformée de Laplace de la mesure d'intensité de \mathcal{L} joue un rôle important dans l'étude de la marche aléatoire branchante. Pour tout $\theta \in \mathbb{R}$, on pose :

$$\Lambda(\theta) = \mathbb{E} \left[\sum_{|x|=1} \exp(\theta V(x)) \right] = \int_{\mathbb{R}} \mu(dx) e^{\theta x}$$

et $\kappa(\theta) = \log(\Lambda(\theta))$.

De plus on notera $\Lambda'(\theta) = \mathbb{E} \left[\sum_{|x|=1} V(x) \exp(\theta V(x)) \right]$, même si Λ n'est définie qu'au seul point θ et que cette notation n'est donc pas bien définie. Bien entendu, $\kappa'(\theta)$ désignera alors $\frac{\Lambda'(\theta)}{\Lambda(\theta)}$.

Pour finir, on notera $q \in [0, 1]$ la probabilité que la marche aléatoire branchante issue d'un unique individu à l'instant origine s'éteigne. Étant donné que l'arbre généalogique de ce processus est un arbre de Galton-Watson, c'est la plus petite solution de l'équation $\mathbb{E}(q^n) = q$, où n suit la loi du nombre de points de \mathcal{L} .

Nous allons maintenant indiquer quelques méthodes utilisées dans l'étude des marches aléatoires branchantes, à l'origine de nombreux résultats obtenus sur ce modèle.

2 Méthode de décomposition en épine dorsale et Lemme de regroupement

2.1 Martingales additives exponentielles et propriété de branchement

L'étude de la marche aléatoire branchante passe souvent par l'étude fine de certaines martingales qui lui sont associées, notamment les martingales additives exponentielles, définies comme

suit. Pour tout $\theta \in \mathbb{R}$ tel que $\kappa(\theta) < +\infty$, on pose :

$$W_n^{(\theta)} = \sum_{|x|=n} \exp(\theta V(x) - n\kappa(\theta)),$$

la propriété de martingale se vérifiant simplement grâce à l'égalité suivante :

$$W_{n+p}^{(\theta)} = \sum_{|x|=n} \exp(\theta V(x) - n\kappa(\theta)) \sum_{y>x, |y|=n+p} \exp(\theta(V(y) - V(x)) - p\kappa(\theta)), \quad (2.1)$$

où $y > x$ indique que y est un descendant de x . C'est une application directe de la propriété de branchement : les arbres issus de chacun des individus x à la génération n forment des marches aléatoires branchantes indépendantes et de même loi.

Les martingales ainsi définies sont positives, elles convergent donc presque sûrement vers des limites $W^{(\theta)}$. Biggins a déterminé en 1977 quand ces limites étaient triviales ou non, une preuve simple a par la suite été apportée par Lyons [9], utilisant principalement la décomposition en épine dorsale.

Théorème 2.1. *Soit $\theta \in \mathbb{R}$ tel que $\kappa(\theta) < +\infty$ et $\kappa'(\theta)$ existe et est fini. On pose alors*

$$X^{(\theta)} = \sum_{|x|=1} \exp(\theta V(x)).$$

Les assertions suivantes sont équivalentes :

1. $\mathbb{P}(W^{(\theta)} = 0) < 1$;
2. $\mathbb{P}(W^{(\theta)} = 0) = q$;
3. $\mathbb{E}(W^{(\theta)}) = 1$;
4. $\mathbb{E}[X^{(\theta)} \log^+(X^{(\theta)})] < +\infty$ et $\theta\kappa'(\theta) < \kappa(\theta)$.

L'équivalence entre les deux premières assertions est une conséquence directe de la propriété de branchement et des résultats standards sur les arbres de Galton-Watson. On pose N_n le nombre d'individus présents à l'instant n , en faisant tendre p vers $+\infty$ dans 2.1, on obtient :

$$W^{(\theta)} = \sum_{|x|=n} \exp(\theta V(x) - \kappa(\theta)) W_x^{(\theta)},$$

où $(W_x^{(\theta)})$ sont des variables aléatoires i.i.d. de même loi que $W^{(\theta)}$.

Par conséquent, on a :

$$\mathbb{P}(W^{(\theta)} = 0) = \mathbb{E} \left[\mathbb{P}(W^{(\theta)} = 0)^{N_n} \right],$$

et $\mathbb{P}(W^{(\theta)} = 0) \leq \mathbb{P}(N_\infty = 0)$, donc $\mathbb{P}(W^{(\theta)} = 0) = q$ ou 0.

Le détail de ce résultat est obtenu en étudiant la marche aléatoire changée de probabilité par rapport à celle que nous avons définie, avec pour dérivée de Radon-Nikodým $M^{(\theta)}$. Nous allons maintenant donner une description de la marche aléatoire branchante sous la loi

$$\mathbb{Q}_{|\mathcal{F}_n}^{(\theta)} = M_n^{(\theta)} \cdot \mathbb{P}_{|\mathcal{F}_n}.$$

2.2 Changement de probabilité

Dans le but de décrire $\mathbb{Q}^{(\theta)}$, nous allons décrire un processus comme un arbre marqué avec épine. En effet, cette loi s'écrit de façon simple en distinguant une lignée particulière, qui se reproduira différemment de tous les autres individus du processus, de la façon suivante.

Théorème 2.2. *On suppose $\kappa(\theta) < +\infty$.*

On pose $\widehat{\mathcal{L}}^\theta$ un processus de points dont la loi admet la densité $\sum_{x \in \mathcal{L}} e^{\theta l_x - \kappa(\theta)}$ par rapport à la loi de \mathcal{L} .

Sous $\mathbb{Q}^{(\theta)}$, la marche aléatoire branchante se comporte ainsi :

- *à l'instant 0, il y a un unique individu situé en 0 noté w_0 ;*
- *à chaque instant n , les individus de la génération n meurent en laissant place à un certain nombre d'enfants, les enfants des individus différents de w_n sont distribués selon des copies indépendantes de \mathcal{L} , et les enfants de w_n sont distribués selon une copie de $\widehat{\mathcal{L}}^\theta$;*
- *l'individu w_{n+1} est l'enfant x de w_n avec probabilité $\frac{e^{\theta V(w_n, x)}}{\sum_{u=1}^A w_n e^{\theta V(w_n, u)}}$.*

En résumé, sous $\mathbb{Q}^{(\theta)}$, la marche aléatoire branchante est donnée par un individu spécial, l'épine, qui donne un certain nombre d'enfants normaux, plus un enfant spécial, les autres individus se reproduisant normalement.

Ce théorème est la décomposition en épine dorsale de la marche aléatoire branchante. Il permet en particulier de démontrer les résultats les plus anciens de convergence de la position de l'individu le plus à droite. Rappelons également des résultats classiques sur les changements de probabilités, qui permettent d'obtenir entre autres les résultats de Biggins.

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_n))$ un espace de probabilité filtré muni de deux probabilités \mathbb{P} et \mathbb{Q} telles que $\mathbb{Q}|_{\mathcal{F}_n} \ll \mathbb{P}|_{\mathcal{F}_n}$. On pose pour $n \in \mathbb{N}$:

$$\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_n} = X_n \text{ et } X = \limsup_{n \rightarrow +\infty} X_n.$$

La proposition suivante permet de préciser le comportement de (X_n) .

Proposition 2.1. *(X_n) est une \mathbb{P} -martingale positive, donc*

$$X_n \xrightarrow[n \rightarrow +\infty]{} X \text{ } \mathbb{P}\text{-p.s. et } X < +\infty \text{ } \mathbb{P}\text{-p.s.}$$

De plus pour tout $A \in \mathcal{F}$, on a :

$$\mathbb{Q}(A) = \mathbb{E}_{\mathbb{P}}(X \mathbf{1}_A) + \mathbb{Q}(A \cap \{X = +\infty\}).$$

Ce résultat peut également être utilisé pour obtenir des résultats sur la martingale, car on a de plus la dichotomie suivante.

Proposition 2.2.

$$\begin{aligned} \mathbb{Q} \ll \mathbb{P} &\iff X < +\infty \text{ } \mathbb{Q}\text{-p.s.} \iff \mathbb{E}_{\mathbb{P}}(X) = 1 ; \\ \mathbb{Q} \perp \mathbb{P} &\iff X = +\infty \text{ } \mathbb{Q}\text{-p.s.} \iff \mathbb{E}_{\mathbb{P}}(X) = 0. \end{aligned}$$

2.3 Le Lemme de regroupement

Nous allons maintenant nous intéresser à une conséquence intéressante de la décomposition en épine dorsale de la marche aléatoire branchante, le Lemme de regroupement, qui permet de remplacer l'étude de nombreuses fonctionnelles additives sur les lignées de la marche aléatoire branchante en des fonctionnelles sur une marche aléatoire simple. De nombreux résultats obtenus dans le cadre des marches aléatoires simples peuvent alors être étendus au cas des marches aléatoires branchantes. De façon heuristique, il suffit de remarquer que lorsqu'on suit une branche au hasard de notre marche, le processus des positions est une marche aléatoire.

Lemme 2.1 (Lemme de regroupement). *Soit $\theta \in \mathbb{R}$ tel que $\kappa(\theta) < +\infty$. Il existe une marche aléatoire (S_n) telle que pour tout $n \in \mathbb{N}$, pour toute fonction mesurable $g : \mathbb{R}^n \rightarrow \mathbb{R}^+$, on a*

$$\mathbb{E} \left[\sum_{|x|=n} g(V(x_1), \dots, V(x_n)) \right] = \mathbb{E} [\exp(\theta S_n - n\kappa(\theta)) g(S_1, \dots, S_n)].$$

Preuve. Ce lemme peut se prouver par récurrence, mais observons qu'il peut être vu comme une conséquence de la décomposition en épine dorsale :

$$\begin{aligned} & \mathbb{E}_{\mathbb{Q}(\theta)} \left[e^{n\kappa(\theta) - \theta V(w_n)} g(V(w_0), \dots, V(w_n)) \right] \\ &= \mathbb{E}_{\mathbb{Q}(\theta)} \left[\sum_{|x|=n} e^{n\kappa(\theta) - \theta V(x)} g(V(x_1), \dots, V(x_n)) \mathbf{1}_{\{x=w_n\}} \right], \end{aligned}$$

or on sait que conditionnellement à \mathcal{F}_n , la probabilité qu'un individu donné soit dans l'épine est donnée par :

$$\mathbb{Q}^{(\theta)}(x = w_n | \mathcal{F}_n) = \frac{e^{\theta V(x) - n\kappa(\theta)}}{W_n^{(\theta)}},$$

par récurrence.

Par conséquent, on obtient :

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}(\theta)} \left[e^{n\kappa(\theta) - \theta V(w_n)} g(V(w_0), \dots, V(w_n)) \right] &= \mathbb{E}_{\mathbb{Q}(\theta)} \left[\frac{1}{W_n^{(\theta)}} \sum_{|x|=n} g(V(x_1), \dots, V(x_n)) \right] \\ &= \mathbb{E} \left[\sum_{|x|=n} g(V(x_1), \dots, V(x_n)) \right], \end{aligned}$$

or sous $\mathbb{Q}^{(\theta)}$, $(V(w_n), n \geq 0)$ est une marche aléatoire. On peut donc conclure. \square

Ce Lemme est utilisé la plupart du temps avec $\theta = 1$ et un processus de points \mathcal{L} renormalisé vérifiant $\kappa(\theta) = 0$, ce qui en simplifie l'expression.

Nous allons maintenant voir quelques résultats qui ont été pour la plupart obtenus en utilisant des variantes de ces méthodes.

3 La position de l'individu le plus à droite

Afin de s'intéresser à la manière dont la population colonise l'espace, on peut s'intéresser à la position de l'individu le plus à droite dans une marche aléatoire branchante (ou bien sûr la particule la plus à gauche, cela revient au même). Les premiers résultats sur ce problème datent des années 75, mais les derniers raffinements n'ont été obtenus que récemment.

On notera dans toute la suite B_n l'individu le plus à droite dans la marche aléatoire branchante, c'est-à-dire :

$$B_n = \max_{|x|=n} V(x), \text{ avec la convention } \max \emptyset = -\infty.$$

3.1 Vitesse de dispersion vers la droite

Le premier résultat que nous citons ici permet de donner la vitesse de dispersion vers la droite v de la marche aléatoire branchante, plus précisément, nous verrons que $\frac{B_n}{n}$ converge vers une constante v , qui peut être calculée.

Théorème 3.1. *Si*

$$\text{il existe } \phi > 0 \text{ tel que } \kappa(\phi) < +\infty, \tag{2.2}$$

alors il existe une constante $v \geq 0$ *telle que*

$$\frac{B_n}{n} \xrightarrow[n \rightarrow +\infty]{} v \text{ p.s.,}$$

sur l'ensemble de non-extinction. De plus $v = \inf_{\theta \geq 0} \left\{ \frac{\kappa(\theta)}{\theta} \right\}$.

Ce théorème peut être étendu pour donner l'asymptotique du nombre d'individus situés à droite de na , pour $a < v$, dans l'échelle exponentielle, en fonction de la transformée de Legendre κ^* de κ , définie par :

$$\kappa^*(a) = \sup_{\theta \geq 0} \{ \theta a - \kappa(\theta) \}.$$

Observons en particulier que $v = \sup\{a \in \mathbb{R} | \kappa^*(a) < 0\}$.

Remarque 3.1. Sous l'hypothèse d'intégrabilité (2.2), κ^* est une fonction convexe semi-continue inférieurement croissante dont la valeur minimale est $-\kappa(0) = -\log \mathbb{E}(N)$, qui est négative, car on suppose la marche aléatoire branchante sur-critique.

Nous pouvons alors formuler le théorème suivant.

Théorème 3.2. *Si l'hypothèse 2.2 est vérifiée, on a, pour tout* $a < v$:

$$\frac{1}{n} \log \sum_{|x|=n} \mathbf{1}_{\{V(x) > na\}} \xrightarrow[n \rightarrow +\infty]{} -\kappa^*(a) \text{ p.s.}$$

sur l'ensemble de non-extinction.

De plus, pour tout $a > v$:

$$\frac{1}{n} \log \sum_{|x|=n} \mathbf{1}_{\{V(x) > na\}} \xrightarrow[n \rightarrow +\infty]{} -\infty \text{ p.s.}$$

Remarque 3.2. Lorsque le processus de points choisi est la donnée de N points i.i.d. de loi ν , on a alors :

$$\kappa(\theta) = \log \left(\mathbb{E} \left(\sum_{i=1}^N e^{\theta X_i} \right) \right) = \log N - L^*(\theta),$$

où L^* est la transformée de Cramer de ν .

Le résultat précédent peut alors être vu ainsi : lorsqu'on considère une marche aléatoire de loi ν , la probabilité qu'elle soit au dessus de na à l'étape n est de l'ordre de $e^{-n\Lambda^*(a)}$. Nous nous intéressons alors à la marche aléatoire branchante, à l'étape n on a $N^n = e^{n \log N}$ individus, qui ont suivi des marches aléatoires au moins en partie indépendantes donc des évènements dont la probabilité était exponentiellement faible peuvent devenir des évènements possibles, et ce tant que $\log N - \Lambda^*(\theta) < 0$. C'est d'ailleurs exactement le résultat que nous obtenons.

Le théorème que nous citons est néanmoins bien plus général puisqu'on ne suppose pas a priori l'indépendance des déplacements à l'intérieur d'une même fratrie.

Les théorèmes obtenus ici peuvent s'étendre aux processus CMJ, des marches aléatoires branchantes à valeurs dans $\mathbb{R}^+ \times \mathbb{R}$, où la première coordonnée du déplacement d'un individu représente l'âge de son parent à sa naissance. En d'autres termes, la première coordonnée de $V(x)$ indique la date à laquelle l'individu x naît. Dans ce cas, on peut obtenir le même genre de résultats asymptotiques en comptant les individus non en fonction de leur génération mais de leur date de naissance.

3.2 Un premier raffinement

Il est naturel de s'intéresser à l'écart entre la prédiction de la position de l'individu le plus à droite nv et sa position réelle. Il faudra s'attendre à des hypothèses d'intégrabilité supplémentaires, car si par exemple il y a toujours au moins un enfant placé exactement à distance v de son parent, alors l'approximation obtenue précédemment est exacte.

Théorème 3.3. *S'il existe $0 < \theta < +\infty$ tel que $\theta v - \kappa(\theta) = 0$, alors*

$$B_n - nv \xrightarrow[n \rightarrow +\infty]{} -\infty \text{ p.s.}$$

sur l'ensemble de non-extinction.

Il est maintenant naturel de s'interroger sur la vitesse de cette divergence, de préciser le second ordre de grandeur de la position de l'individu le plus à droite. Ces résultats nécessitent de plus fortes hypothèses d'intégrabilité pour être obtenues. On peut montrer, sous réserve de bonnes conditions d'intégrabilité, que la différence entre B_n et $n\theta$ est d'ordre logarithmique, et qu'à cet ordre, il existe des fluctuations presque sûres, ainsi qu'une limite en probabilité. On obtient le résultat suivant :

Théorème 3.4. *Supposons qu'il existe $\varepsilon > 0$ tel que $\mathbb{E}(N^{1+\varepsilon}) < +\infty$, $\kappa(\theta + \varepsilon) < +\infty$ et $\kappa(-\varepsilon) < +\infty$, alors on a*

$$-\frac{3}{2\theta} = \liminf \frac{B_n - n\Gamma}{\log n} < \limsup \frac{B_n - n\Gamma}{\log n} = -\frac{1}{2\theta} \text{ p.s.}$$

et

$$\frac{B_n - nv}{\log n} \xrightarrow[n \rightarrow +\infty]{(P)} -\frac{3}{2\theta} \text{ p.s.}$$

sur l'ensemble de non-extinction.

Ce résultat peut être obtenu grâce au Lemme de regroupement, en utilisant des estimations finies sur les marches aléatoires, comme dans l'article d'Aïdekon et Shi [3].

3.3 Dernières avancées

Pour finir, les derniers résultats sur l'individu le plus à droite donne également une limite en loi pour la position de l'individu le plus à droite à laquelle on soustrait le présent équivalent. Ce résultat est prouvé dans [2]. Ce théorème est obtenu avec des hypothèses d'intégrabilité moins fortes que le précédent, mais en supposant que la marche aléatoire branchante considérée est critique, c'est-à-dire qui vérifie les conditions suivantes :

$$\kappa(0) > 0, \quad \kappa(-1) = 0 \text{ et } \kappa'(-1) = 0.$$

Il est alors plus naturel de regarder le symétrique du processus, ainsi que la position de l'individu le plus à gauche, que l'on note :

$$M_n = \min_{|x|=n} V(x),$$

qui tend vers $+\infty$ presque sûrement, grâce au Théorème 3.3. Ce théorème reste assez général. En effet, sous des hypothèses suffisantes, une marche aléatoire branchante peut être transformée en marche aléatoire branchante critique, par une transformation affine du processus de points \mathcal{L} .

Nous introduisons alors la martingale dérivée, définie pour tout $n \geq 0$ par

$$\partial W_n = \sum_{|x|=n} V(x) e^{-V(x)},$$

cette martingale converge presque sûrement vers ∂W_∞ , qui est strictement positive sur l'ensemble de non-extinction de la marche aléatoire branchante, sous des hypothèses suffisantes.

Théorème 3.5. *On suppose que la distribution de \mathcal{L} n'est pas à valeurs dans un réseau, on introduit les variables aléatoires suivantes :*

$$X = \sum_{|x|=1} e^{-V(x)}, \quad \tilde{X} = \sum_{|x|=1} V(x)_+ e^{-V(x)} \text{ et } \hat{X} = \sum_{x=1} V(x)^2 e^{-V(x)},$$

et on suppose :

$$\mathbb{E}(X(\ln_+ X)^2) < +\infty, \quad \mathbb{E}(\tilde{X} \ln_+(\tilde{X})) < +\infty, \text{ et } \mathbb{E}(\hat{X}) < +\infty.$$

Il existe une constante $c > 0$ telle que pour tout $x \in \mathbb{R}$:

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left(M_n \geq \frac{3}{2} \log n + x \right) = \mathbb{E}(e^{-ce^x \partial W_\infty}).$$

Après ces résultats bien connus, nous allons maintenant introduire une variante importante des marches aléatoires branchantes, qui permet de tenir compte de la sélection naturelle dans les processus de population.

4 Marche aléatoire branchante avec sélection

Dans la plupart des modèles probabilistes de dynamique des populations, on a une dichotomie assez gênante extinction-divergence à vitesse exponentielle vers $+\infty$. Autrement dit, une population ne peut atteindre un équilibre simple, comme ceci est observé dans la nature. Ces observations nous permettent donc de supposer qu'une composante essentielle manque au modèle. Celle-ci pourrait être la sélection naturelle, qui peut être représentée de manière assez simple dans le formalisme de la marche aléatoire branchante. La position V de l'individu, représente alors son degré d'adaptation au milieu, et selon la logique darwinienne, les plus adaptés survivent.

On fixe un entier N représentant la capacité d'accueil du milieu, le nombre d'individus qui peuvent se nourrir. La marche aléatoire branchante avec sélection évolue alors comme une marche aléatoire branchante classique, avec la règle supplémentaire suivante : si à n'importe quel instant donné il y a plus de N individus dans la population, seuls les N plus à droite survivent, les autres sont immédiatement tués sans descendance.

Afin d'éviter l'extinction presque sûre du processus, on supposera $\mathbb{P}(\mathcal{L} = \emptyset) = 0$. Le but devient alors d'étudier la position de l'individu le plus à droite dans ce processus modifié, en fonction du nombre N d'individu conservés. Un couplage trivial montre que $v^N \geq v^{N+1} \geq v$, l'individu le plus à droite d'un processus très sélectif est toujours moins à droite que si on avait gardé d'avantage d'individus, puisqu'on a ainsi éliminé un certain nombre de lignées qui pourraient contenir le « vrai » élément le plus à droite.

Une conjecture de Derrida et Brunet donne la vitesse de convergence de v^N vers v . Cette conjecture prétend que :

$$v^N - v = \frac{c_1}{(\log n)^2} + \frac{c_2 + o(1)}{\log \log n},$$

Bérard et Gouéré (c.f. [7]) ont récemment prouvé que l'équivalent de cette suite était bien de cette forme, avec une constante c_1 qui dépend du modèle. On conjecture que la constante c_2 est en revanche universelle.

Ces résultats ont de forts intérêts pratiques : en informatique notamment, ils mesurent l'erreur faite lorsqu'on remplace un arbre de recherche (à croissance exponentielle) par les N meilleures branches à chaque étape, soit une exploration bien plus rapide.

L'intérêt est également évident pour le modèle biologique, bien que l'hypothèse de capacité d'accueil du milieu constante soit assez restrictive. Des modifications du modèle utilisant une capacité d'accueil $\phi(n)$ dépendant du temps, voire de la population présente à l'instant précédent, pourraient être intéressantes à étudier également. Un certain nombre de questions classiques en biologie peuvent alors se poser encore : comment se comporte l'arbre généalogique lorsqu'on remonte dans le temps ? Quel est l'âge du plus récent ancêtre commun ? Existe-t-il des modèles limites d'intérêt dans cette étude ?

Bibliographie

- [1] K.B. Athreya and P.E. Ney. *Branching Processes*. Grundlehren der mathematischen Wissenschaften. Springer, 1972.

- [2] E. Aïdékon. Convergence in law of the minimum of a branching random walk. *arXiv :1101.1810*, 2011.
- [3] Elie Aïdékon and Zhan Shi. Weak convergence for the minimal position in a branching random walk : a simple proof. *Periodica Mathematica Hungarica*, 61(1-2) :43–54, 2010.
- [4] J.D. Biggins. Lindley-type equations in the branching random walk. *Stochastic Processes and Applications*, 75(1) :105–133, 1998.
- [5] J.D. Biggins. Branching out. *arXiv :1003.4715*, 2010.
- [6] J.D. Biggins and A.E. Kyprianou. Fixed points of the smoothing transform : the boudary case. *Electronic Journal of Probability*, 10(17) :609–631, 2005.
- [7] Jean Bérard and Jean-Baptiste Gouéré. Brunet-derrida behavior of branching-selection particle systems on the line. *arXiv :0811.2782*, 2010.
- [8] Yueyun Hu and Zhan Shi. Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. *The Annals of Probability*, 37(2) :742–789, 2009.
- [9] Russel Lyons. A simple path to biggins’ martingale convergence for branching random walk. *Classical and Modern Branching Processes*, pages 217–221, 1997.

Deuxième partie

Mémoire de M2

Chapitre 3

Divergence globale des processus coalescents spatiaux

Bastien MALLEIN

Sous la direction de Jean BERTOIN

10 février 2011

Résumé

Lorsqu'on cherche à étudier le patrimoine génétique d'une population, on a souvent besoin de s'intéresser à l'arbre généalogique des individus formant cette population. C'est dans ce but que Kingman a introduit, en 1982, un modèle décrivant la généalogie d'une population [5, 6]. Ce modèle a par la suite été généralisé indépendamment par Pitman [8] et Sagitov [9] à des processus appelés Λ -coalescents. Dans ces modèles, on peut en particulier s'intéresser à la hauteur de cet arbre généalogique, c'est-à-dire à l'âge du plus récent ancêtre commun à une famille de n individus.

Il apparaît que dans de nombreux cas, cet âge est borné indépendamment de n . Ce résultat a été prouvé par Schweinsberg [10] qui a donné une condition nécessaire et suffisante pour la finitude de la hauteur de l'arbre généalogique d'un Λ -coalescent. Le but de ce mémoire est de prouver, comme dans [1], que lorsqu'on autorise les individus à se déplacer sur un graphe infini, alors cette borne n'existe jamais. Plus précisément nous montrerons qu'aucun processus Λ -coalescent spatial sur un graphe infini, partant d'une population infinie en un point ne devient fini en temps fini, et dans certains cas, nous serons même capables de donner un ordre de grandeur du nombre d'ancêtres qui engendrent, en un temps t , une famille de n individus.

Je tiens à remercier M. Bertoin qui m'a proposé ce sujet, et m'a apporté son aide pendant la rédaction de ce mémoire, ainsi que Yasmine pour sa relecture attentive.

1 À propos des processus coalescents

Un processus coalescent est un modèle de la généalogie de n individus. C'est un processus stochastique qui à chaque instant t associe la partition de la population selon le critère suivant : deux individus sont dans le même ensemble s'ils ont un ancêtre commun âgé de moins de t , i.e. les lignées ancestrales de l'arbre généalogique se sont regroupées avant la génération $-t$. On appelle ces processus « coalescents » car au cours du temps, les lignées ancestrales se rejoignent et fusionnent en une seule. On peut alors voir ce processus comme un ensemble de particules qui se regroupent en sous-ensembles plus gros à certains instants. Lorsqu'on étudie des modèles mathématiques de populations, comme le modèle de Wright-Fisher, l'un de ces processus coalescents revient très souvent pour en donner la généalogie, c'est le coalescent de Kingman.

Ce modèle suppose que durant l'évolution, les coagulations de lignées ancestrales se font de manière microscopique, c'est-à-dire que seuls deux individus sont concernés par un épisode de coagulation. Ce processus a par la suite été généralisé en autorisant des évolutions plus brusques dans les lignées ancestrales. Ces processus sont appelés des Λ -coalescents. Nous allons commencer par citer les principaux résultats que nous démontrerons ici, puis après quelques notations qui nous permettront d'écrire de manière plus précise ce qu'est un processus à valeurs dans les espaces de partitions, nous introduirons tous ces processus coalescents, ainsi qu'une de leurs propriétés que nous nous proposons d'étudier ici : l'existence d'un ancêtre commun à toute la population. Par la suite nous étendrons ce genre de notions en permettant aux individus de se déplacer sur un graphe.

Remarque 1.1. Dans tout ce mémoire, nous aurons besoin de nombreuses constantes, que nous appellerons c, c_1, c_2, \dots et C, C_1, C_2, \dots , souvent respectivement choisies assez petites et assez grandes, qui pourront varier d'une ligne à l'autre de nos équations, par soucis de simplicité.

1.1 Principaux résultats

Nous présenterons ici quelques résultats de l'étude des processus coalescents obtenus par Angel, Berestycki et Limic dans leur article [1], portant sur l'étude de l'existence d'un ancêtre commun dans les processus coalescents spatiaux. Le premier théorème auquel nous nous intéresserons concerne le coalescent de Kingman spatial, pour lequel on peut donner un ordre de grandeur du nombre d'ancêtres de la famille en remontant de t dans le temps, ou autrement dit du nombre de particules présentes à l'instant t .

Commençons par définir la fonction $\log^* n$ ou fonction logarithme itérée. Elle est définie de la manière suivante :

$$\log^* n = \min\{k \in \mathbb{N} : \log^{(k)} n < 1\},$$

où on a posé $\log^{(k)} n = \underbrace{\log \circ \dots \circ \log}_k n$.

Cette fonction diverge vers $+\infty$, mais de manière extrêmement lente. On sait par exemple que pour tout $n < 10^{1656520}$, $\log^* n$ est inférieur ou égale à 4.

Théorème 1.1. *Considérons le coalescent de Kingman spatial sur un graphe infini G de degré maximal fini D partant de n individus tous situés en $u \in G$. Posons N_t^n le nombre d'individus*

présents à l'instant $t > 0$, il existe deux constantes $c, C > 0$ dépendant uniquement de D et t telles que pour tout $\varepsilon > 0$ on ait :

$$\mathbb{P}(c \text{Vol}[B(u, (1 - \varepsilon) \log^* n)] \leq N_t^n \leq C \text{Vol}[B(u, (1 + \varepsilon) \log^* n)]) \xrightarrow{n \rightarrow +\infty} 1,$$

On obtient donc la divergence vers l'infini de la taille de la population au temps t lorsqu'on part de $n \rightarrow +\infty$ individus à l'origine, ce qui est bien différent du comportement qu'adopte le coalescent non-spatial. Un résultat similaire peut aussi être démontré dans le cas des processus Beta-coalescents spatiaux, pour lesquels on a encore une estimation du nombre d'individus restants à un instant donné.

Théorème 1.2. *Soit $1 < \alpha < 2$, et considérons un Beta($\alpha, 2 - \alpha$)-coalescent spatial sur un graphe infini G de degré maximal fini D partant de n individus tous situés en $u \in G$. Posons N_t^n le nombre d'individus présents à l'instant $t > 0$, il existe deux constantes $c, C > 0$ dépendant uniquement de α, D et t telles que :*

$$\mathbb{P}(c \text{Vol}[B(u, c \log \log n)] \leq N_t^n \leq C \text{Vol}[B(u, C \log \log n)]) \xrightarrow{n \rightarrow +\infty} 1,$$

Plus généralement, nous pouvons prouver que pour n'importe quel processus Λ -coalescent spatial (Λ sans masse de Dirac en 1) sur un graphe infini, la taille de la famille au temps t diverge.

Théorème 1.3. *Soit Λ une mesure finie sur $[0, 1]$ et considérons le processus Λ -coalescent spatial sur un graphe infini G partant de n individus tous situés en $u \in G$. Soit N_t^n le nombre d'individus restants à l'instant $t > 0$, on a :*

$$N_t^n \xrightarrow{n \rightarrow +\infty} +\infty \text{ p.s.}$$

Finalement, nous pouvons également nous intéresser au nombre de lignées généalogiques persistant sur le long terme, particulièrement dans le cas où le graphe G est de la forme \mathbb{Z}^d , pour lequel, grâce aux propriétés asymptotiques sur la taille des boules, les résultats précédents s'écrivent bien plus facilement. Nous avons alors l'estimation suivante :

Théorème 1.4. *Supposons que le coalescent que nous considérons est le coalescent de Kingman sur $G = \mathbb{Z}^d$, posons $m_n = \log^* n$ et $\delta > 0$ fixé. Il existe deux constantes $c, C > 0$ dépendant uniquement de d, δ telles que si $d > 2$:*

$$\mathbb{P}\left(cm_n^{d-2} \leq N_{\delta m_n^2}^n \leq Cm_n^{d-2}\right) \xrightarrow{n \rightarrow +\infty} 1,$$

et si $d = 2$:

$$\mathbb{P}(c \log m_n \leq N_{\delta m_n^2}^n \leq C \log m_n) \xrightarrow{n \rightarrow +\infty} 1.$$

Si le coalescent considéré est un Beta($\alpha, 2 - \alpha$) coalescent, nous avons le même résultat en prenant $m_n = \log \log n$.

En réalité dans ce dernier théorème, le fait que le coalescent considéré soit de Kingman influe uniquement sur l'étalement initial des individus. Ensuite sur un temps long, la probabilité de trouver simultanément deux individus en un même sommet du graphe devient faible, et ceux-ci fusionnent avec une certaine probabilité, ne dépendant que de $\Lambda([0, 1])$. Voir plus de deux individus sur un même sommet est encore plus rare, donc on ne peut pas réellement différencier un coalescent de l'autre sur un temps long, et le comportement de tout coalescent devient donc similaire à celui du coalescent de Kingman.

Pour démontrer ces résultats, les processus que nous aurons réellement besoin d'étudier sont ceux donnant le nombre de blocs d'un processus coalescent, c'est-à-dire le nombre d'ancêtres communs à la génération $-t$ du processus. Nous présenterons toutefois les processus à valeurs dans les espaces de partitions, car les objets sont alors plus intuitifs, en particulier la notion de limite projective nous permettra de définir un processus recouvrant tous les autres. Commençons par quelques résultats sur les espaces de partitions.

1.2 Quelques notions sur les espaces de partitions

Commençons par énumérer les individus que nous considérons à l'instant initial. Soit \mathbb{N} l'ensemble des entiers naturels strictement positifs. On notera $[n]$ l'ensemble $\{1, \dots, n\}$, qui correspondra donc à une population de n individus que nous souhaitons étudier.

Une partition de $B \subset \mathbb{N}$ est une collection $\pi = (\pi_i, i \in \mathbb{N})$ de sous-ensembles disjoints de B tels que $\bigcup_{i \in \mathbb{N}} \pi_i = B$, indexés selon l'ordre croissant de leur plus petit élément, i.e. :

$$\forall i \leq j, \min \pi_i \leq \min \pi_j,$$

avec la convention $\min \emptyset = +\infty$. Le bloc π_i représentera l'individu qui est l'ancêtre commun de chacun des individus dont le numéro est dans π_i , et aucun des autres.

On dira que $B' \subset B$ est un bloc de π si c'est un élément non-vide de π . Une partition π est dite plus fine qu'une autre π' si chaque bloc de π est inclus dans un bloc de π' . Dans ce cas on dit également que π' est plus grossière que π .

L'ensemble des partitions de B est noté \mathcal{P}_B , et on écrira également \mathcal{P}_n pour $\mathcal{P}_{[n]}$ et \mathcal{P}_∞ pour $\mathcal{P}_{\mathbb{N}^*}$. L'un des éléments remarquables de cet ensemble est la partition en singletons $(\{1\}, \{2\}, \dots)$, notée $\mathbf{0}_n \in \mathcal{P}_n$ ou $\mathbf{0}_\infty \in \mathcal{P}_\infty$. Le nombre de blocs d'une partition π est le cardinal de l'ensemble des blocs (non-vides par définition) de π , noté $\#\pi$:

$$\#\pi = \#\{i \in \mathbb{N}^* | \pi_i \neq \emptyset\} = \max\{i \in \mathbb{N}^* | \pi_i \neq \emptyset\}.$$

Si B' est un sous-ensemble de B , toute partition $\pi \in \mathcal{P}_B$ induit naturellement une partition sur B' notée $\pi|_{B'}$ et définie par :

$$\pi|_{B'} = (\pi_i \cap B', i \in \mathbb{N}) \text{ convenablement réordonnée.}$$

On définit alors la notion de compatibilité de partitions. Si $B' \subset B$, $\pi' \in \mathcal{P}_{B'}$ et $\pi \in \mathcal{P}_B$ sont dites compatibles si et seulement si on a $\pi' = \pi|_{B'}$. On peut alors définir la limite projective d'une suite de partitions compatibles au sens suivant.

Si π^1, π^2, \dots sont des partitions de $[1], [2], \dots$, elles sont dites compatibles si pour toute paire d'entiers $p \leq n$, $\pi^n|_{[p]} = \pi^p$. On peut alors prouver sans difficulté le lemme suivant, qui nous servira par la suite à construire nos processus coalescents.

Lemme 1.1. Soient π^1, π^2, \dots des partitions de $[1], [2], \dots$. On a :

$$(\pi^n)_{n \in \mathbb{N}^*} \text{ compatibles} \iff \exists \pi \in \mathcal{P}_\infty | \forall n \in \mathbb{N}^*, \pi^n = \pi|_{[n]}$$

Cette propriété est évidente lorsqu'on représente \mathcal{P}_∞ sous forme des feuilles d'un arbre. Les nœuds au niveau n sont les partitions de \mathcal{P}_n , et les liens entre nœuds sont définis par la relation de compatibilité. La suite (π^n) représente alors un chemin sur cet arbre, et π est la feuille au bout du chemin.

Nous voyons ainsi qu'il est équivalent de construire un processus à valeurs dans \mathcal{P}_∞ , et de construire une suite de processus sur les partitions de $[n]$ pour tout $n \in \mathbb{N}^*$ de telle manière que les trajectoires soient toutes compatibles. Nous construirons souvent ainsi nos processus, sur un nombre fini d'individus, puis nous passerons à la limite projective.

$$\text{On munit maintenant } \mathcal{P}_\infty \text{ de la distance } d(\pi, \pi') = \frac{1}{\max\{p \in \mathbb{N}^* | \pi|_{[p]} = \pi'|_{[p]}\}}.$$

Propriété 1.1. (\mathcal{P}_∞, d) est un espace métrique compact.

Preuve. On part de π^n une suite de partitions de \mathbb{N} . Nous allons extraire par récurrence une suite de partitions compatible. Posons $n_1 = 1$.

Soit $k \in \mathbb{N}$, on suppose qu'il existe une infinité d'entiers p tels que $\pi_{[k]}^p = \pi_{[k]}^{n_k}$. Parmi cette infinité d'entiers, une infinité d'entre eux ont des restrictions à $k+1$ égales, par principe des tiroirs. On choisit alors n_{k+1} comme le plus petit entier supérieur à n_k tel qu'il existe une infinité d'entiers q tels que $\pi_{[k+1]}^q = \pi_{[k+1]}^{n_{k+1}}$ et $\pi_{[k]}^q = \pi_{[k]}^{n_k}$.

La suite $(\pi_{[k]}^{n_k})_{k \in \mathbb{N}}$ est une suite de partitions compatible, il existe donc une partition π^∞ de \mathcal{P}_∞ . De plus on a $\pi^{n_k} \xrightarrow[k \rightarrow +\infty]{} \pi^\infty$. \square

Munissons maintenant l'ensemble des partitions d'une loi de composition interne Coag, qui permettra de réécrire des résultats de coagulation, par :

$$\text{Coag}(\pi, \pi') = \left(\bigcup_{j \in \pi'_i} \pi_j \right)_{i \in \mathbb{N}},$$

le résultat de la coagulation de π par π' . Autrement dit un bloc de la coagulation de π par π' est l'union des blocs de π dont l'indice est dans un même bloc de π' .

L'application Coag est 1-lipschitzienne et est compatible avec l'application de restriction. Pour $n \in \mathbb{N}$, on a :

$$\text{Coag}(\pi, \pi')|_{[n]} = \text{Coag}(\pi|_{[n]}, \pi') = \text{Coag}(\pi|_{[n]}, \pi'|_{[n]}).$$

On appelle partition simple une partition dans laquelle il n'existe qu'un seul bloc qui ne soit ni vide, ni un singleton. Soit $\pi \in \mathcal{P}_n$ possédant k blocs et $\sigma \in \mathcal{P}_k$ une partition simple avec un bloc à p éléments, alors la partition $\text{Coag}(\pi, \sigma)$ est appelé une coagulation de p blocs de π . Autrement dit, on choisit p blocs (non-vides) de π et on les réunit en un seul. Cette opération est l'évolution élémentaire d'un Λ -coalescent, chaque évènement de coagulation pourra s'écrire sous cette forme.

Dans toute la suite, si B est un bloc de \mathbb{N} qui n'est pas un singleton, nous noterons σ_B la partition de \mathcal{P}_∞ simple dont le seul bloc non-trivial est B . Par exemple $\sigma_{\{1,2\}} = (\{1, 2\}, \{3\}, \{4\}, \dots)$.

2 Le cas des coalescents non-spatiaux

Nous allons maintenant définir un processus coalescent non-spatial, et étudier quelques-unes de ses propriétés de base avant de nous tourner vers les coalescents spatiaux et les résultats que nous souhaitons démontrer. Certaines des propriétés évoquées ici pourront être utiles par la suite car nous pouvons assez aisément coupler un coalescent spatial avec son pendant non-spatial. Définissons pour commencer le premier de ces coalescents, introduit par Kingman.

2.1 Le coalescent de Kingman

De façon heuristique, le coalescent de Kingman correspond à la généalogie de modèles de populations pour lesquels chaque individu n'engendre à un instant donné qu'une portion infime de la population totale. On dit qu'une partition π' peut être obtenue par coagulation d'une paire de blocs de π si il existe une paire d'entiers $i < j$ tels que $\pi_i, \pi_j \neq \emptyset$ et $\pi' = \text{Coag}(\pi, \sigma_{\{i,j\}})$.

Nous allons maintenant définir le coalescent de Kingman pour un nombre fini d'individus.

Définition 2.1. Soit $n \in \mathbb{N}^*$, un n -coalescent de Kingman est un processus de Markov $(\Pi_t^n)_{t \in \mathbb{R}^+}$ à valeurs dans \mathcal{P}_n partant d'une partition $\pi_0 \in \mathcal{P}_n$ et possédant la dynamique suivante.

Le taux de transition d'une partition π à une partition pouvant être obtenue comme la coagulation de deux blocs de π est égal à 1, et tous les autres taux sont égaux à 0.

Remarquons que la partition $([n], \emptyset, \emptyset \dots)$ est un état absorbant de la chaîne de Markov. De plus la dynamique peut se résumer à ce qui suit : si à un instant donné nous avons k « lignées » (i.e. $\#\Pi_t^n = k$), à taux $\frac{k(k-1)}{2}$, deux d'entre elles choisies uniformément au hasard se réunissent.

Remarque 2.1. Le processus décrivant le nombre de blocs d'un n -coalescent de Kingman est également un processus de Markov, c'est un processus de mort avec taux de mort au niveau k égal à $\frac{k(k-1)}{2}$ (c'est le nombre de paires de blocs pouvant coaguler). On note ce processus D_t^n dans la suite.

Nous allons maintenant vérifier que les n -coalescents forment une suite de processus de Markov compatibles, c'est-à-dire que la restriction d'un n -coalescent à $[k]$ est un k -coalescent. C'est intuitivement assez évident puisque c'est la généalogie d'une famille de k individu, ayant oublié les $n - k$ autres.

Lemme 2.1. *Pour tout $n \geq 2$, le processus $\Pi_{|[n-1]}^n$ est un $(n - 1)$ -coalescent de Kingman.*

Preuve. Considérons Π^n un n -coalescent de Kingman partant de la partition π . Lorsqu'on s'intéresse à la restriction de ce processus deux cas se présentent.

Commençons par le cas où le bloc de π contenant n contient également un autre entier $i < n$. Alors de la restriction de Π^n à \mathcal{P}_{n-1} on peut remonter au processus original en ajoutant n à la partition contenant i . Par conséquent, $\Pi_{|[n-1]}^n$ est un processus de Markov avec les mêmes transitions qu'un $n - 1$ -coalescent partant de $\pi_{|[n-1]}$.

Supposons maintenant que le bloc $\{n\} \in \pi$, et que π possède k blocs. Dans ce cas deux types de sauts peuvent se produire, et l'un d'eux implique un individu qu'on ne peut voir dans le processus restreint. Soit τ le premier instant de saut, et A l'évènement $\{\{n\} \in \Pi_\tau^n\}$ sur lequel l'individu n n'est pas impliqué. Soit τ' le deuxième instant de saut.

Remarquons alors que le premier temps de saut de $\Pi_{[n-1]}^n$ est $\tilde{\tau} = \tau + \mathbf{1}_{\{A^c\}}\tau'$. De plus, τ est une variable aléatoire exponentielle de paramètre $\frac{k(k-1)}{2}$, A est un évènement indépendant de τ de probabilité $1 - \frac{2}{k}$, et τ' est une variable aléatoire exponentielle indépendante de paramètre $\frac{(k-1)(k-2)}{2}$. Nous avons alors :

$$\begin{aligned} \mathbb{P}(\tau + \mathbf{1}_{\{A^c\}}\tau' > s) &= \frac{2}{k}\mathbb{P}(\tau + \tau' > s) + \left(1 - \frac{2}{k}\right)\mathbb{P}(\tau > s) \\ &= \frac{2}{k}\left(\frac{k}{2}e^{-\frac{(k-1)(k-2)}{2}s} - \frac{k-2}{2}e^{-\frac{k(k-1)}{2}s}\right) + \left(1 - \frac{2}{k}\right)e^{-\frac{k(k-1)}{2}s} \\ &= e^{-\frac{(k-1)(k-2)}{2}s}. \end{aligned}$$

Le premier instant de saut pour le processus restreint est donc de loi exponentielle de paramètre $\frac{(k-1)(k-2)}{2}$. De plus $\Pi_{[n-1]}^n$ est uniforme sur les partitions de $[n-1]$ et indépendant de $\tilde{\tau}$. C'est donc bien le premier saut d'un $n-1$ -coalescent de Kingman.

En utilisant la propriété de Markov forte, grâce aux deux cas déjà traités, on obtient bien que $\Pi_{[n-1]}^n$ est un $n-1$ -coalescent de Kingman. \square

Remarquons en passant que nous pouvons réduire l'étude des coalescents de Kingman aux processus partant de la partition en singletons $\mathbf{0}_n$.

Propriété 2.1. Soit $\pi_0 \in \mathcal{P}_n$ possédant k blocs, Π^n un n -coalescent de Kingman partant de la partition en singletons et $\Pi^k = \Pi_{[k]}^n$.

Le processus $\Pi = \text{Coag}(\pi_0, \Pi^n) = \text{Coag}(\pi_0, \Pi^k)$ est un coalescent de Kingman partant de la partition π_0

Preuve. On a clairement $\Pi_0 = \pi_0$, et les taux de transition correspondent, ce qui prouve notre point. \square

En utilisant la propriété de Markov, nous voyons qu'un processus n -coalescent de Kingman arrivé dans un état comportant k blocs se comporte comme un k -coalescent de Kingman, pour lequel les individus sont les blocs du processus arrêté. C'est encore une fois intuitif, nous étudions la généalogie des k ancêtres de notre population de n individus, qui fait bien partie de la généalogie de ces n individus.

Nous allons maintenant utiliser la propriété de compatibilité des n -coalescents, nous pouvons appliquer le théorème d'extension de Kolmogorov pour définir simultanément des n -coalescents compatibles pour tout entier n . Imposons de plus que tous ces processus aient des trajectoires càdlàg. Le Lemme 1.1 nous permet alors de définir le processus obtenu par limite projective vérifiant $\Pi_t|_{[n]} = \Pi_t^n$. C'est ce processus qui sera appelé coalescent de Kingman.

Définition 2.2. Il existe un processus, unique en loi, noté Π^K à valeurs dans \mathbb{P}_∞ , partant de π , tel que, pour tout $n \in \mathbb{N}$, $\Pi_{[n]}^K$ est un n -coalescent de Kingman partant de $\pi_{[n]}$. Ce processus Π^K est appelé coalescent de Kingman.

Comme dans le cas fini, la coagulation d'une partition π avec un coalescent de Kingman partant de la partition en singletons $\mathbf{0}_\infty$ est un coalescent de Kingman partant de π , encore une fois par limite projective. Posons $D_t = \#\Pi_t^K$ le processus du nombre de blocs du coalescent de Kingman.

Nous allons maintenant nous intéresser à l'existence de l'ancêtre commun du coalescent de Kingman. Si celui-ci existe, alors son âge est une borne supérieure pour celui de l'ancêtre commun de toute sous-famille finie d'individus.

Remarquons que si à un instant donné le processus arrive dans un état π n'ayant qu'un nombre fini de blocs non-vides k , alors il existe un ancêtre commun, car à partir de cet instant, le processus se comporte comme un k -coalescent de Kingman partant de la partition en singletons, pour lequel on remplace chaque individu par l'un des blocs non-vides de π . Comme l'espace des états de Π^k est fini et que $([k], \emptyset, \emptyset, \dots)$ est absorbant, Π^K arrive presque sûrement dans l'état $(\mathbb{N}, \emptyset, \emptyset, \dots)$ en un certain temps.

Théorème 2.1. *Le coalescent de Kingman descend de l'infini presque sûrement, i.e.*

$$\forall t > 0, D_t < +\infty \text{ p.s.}$$

Preuve. Notons $D_t^n = \#\Pi_{[n]_t}^K$, c'est un processus du nombre de blocs d'un n -coalescent de Kingman, et nous avons de plus :

$$D_t = \lim_{n \rightarrow +\infty} D_t^n \text{ p.s.}$$

Soit $(e_i)_{i \in \mathbb{N}}$ une suite de variables aléatoires i.i.d. de loi exponentielle de paramètre 1. Connaissant la loi des processus de mort D_t^n , on a :

$$\mathbb{P}(D_t^n \geq k) = \mathbb{P}\left(\sum_{j=k+1}^n \frac{2}{j(j-1)} e_j > t\right) \leq \mathbb{P}\left(\sum_{j=k+1}^{+\infty} \frac{2}{j(j-1)} e_j > t\right).$$

Or remarquons que, en utilisant Fubini-Tonelli,

$$\mathbb{E}\left(\sum_{j=2}^{+\infty} \frac{2}{j(j-1)} e_j\right) = \sum_{j=2}^{+\infty} \frac{2}{j(j-1)} = 2.$$

Par conséquent la série $\sum \frac{e_j}{j(j-1)}$ converge p.s., ce qui nous donne bien :

$$\mathbb{P}(D_t \geq k) = \lim_{n \rightarrow +\infty} \mathbb{P}(D_t^n \geq k) \leq \mathbb{P}\left(\sum_{j=k+1}^{+\infty} \frac{2}{j(j-1)} e_j > t\right).$$

Comme la borne ci-dessus tend vers 0 quand k tend vers l'infini, on obtient bien le résultat escompté. \square

Nous obtenons en particulier la loi de l'âge du dernier ancêtre commun du coalescent de Kingman, si on note $T = \inf\{t > 0 | D_t = 1\}$, on a :

$$T \stackrel{(d)}{=} \sum_{j=2}^{+\infty} \frac{2}{j(j-1)} e_j.$$

De plus pour tout n -coalescent de Kingman, le temps d'atteinte de 1 pour D^n est dominé par T .

Lorsque nous étudierons le coalescent de Kingman spatial, il sera utile de connaître son comportement pour de petites valeurs de t . Nous allons donc nous intéresser à la « vitesse de descente de l'infini » du coalescent de Kingman, c'est-à-dire à un équivalent de D_t au voisinage de 0.

Proposition 2.2.

$$\lim_{t \rightarrow 0} t D_t = 2 \text{ p.s.}$$

Preuve. On note à nouveau

$$T^n = \inf\{t > 0 | \#\Pi_t^K = n\} \stackrel{(d)}{=} \sum_{j=n+1}^{+\infty} \frac{2}{j(j-1)} e_j.$$

Dans ce cas nous avons :

$$\mathbb{E}(T^n) = \frac{2}{n} \text{ et } \text{Var}(T^n) = \sum_{j=n+1}^{+\infty} \left(\frac{2}{j(j-1)} \right)^2 = O(n^{-3}).$$

Dès lors, nous obtenons rapidement que :

$$\mathbb{P} \left(\left| T^n - \frac{2}{n} \right| > \frac{1}{n \log n} \right) = O \left(\frac{(\log n)^2}{n} \right).$$

Grâce au lemme de Borel-Cantelli, on obtient facilement que :

$$n^2 T^{n^2} \xrightarrow[n \rightarrow +\infty]{} 2 \text{ p.s.}$$

d'où on tire, par un argument de monotonie,

$$n T^n \xrightarrow[n \rightarrow +\infty]{} 2 \text{ p.s.}$$

Ceci est bien équivalent par monotonie à $\lim_{t \rightarrow 0} t \#\Pi_t^K = 2$ p.s. □

Dernier fait à propos du coalescent de Kingman : le fait qu'il descende de l'infini, et sa dynamique sur les partitions possédant un nombre fini de blocs le caractérisent de manière unique. Autrement dit, il existe une unique loi d'entrée de l'infini pour le coalescent de Kingman.

Théorème 2.2. *Il existe un unique (en loi) processus de Markov Π à valeurs dans \mathcal{P}_∞ qui vérifie les propriétés suivantes :*

- $\lim_{t \rightarrow 0} \Pi_t = \mathbf{0}_\infty$;
- Π descend de l'infini p.s. ;
- Le taux de transition d'une partition π comportant un nombre fini de blocs à une partition π' obtenue par coagulation de deux blocs de π est 1 et tous les autres taux sont nuls.

Preuve. Soit Π un tel processus de Markov. On pose, pour $n \in \mathbb{N}$, $T_n = \inf\{t > 0 \mid \#\Pi_t = n\}$. C'est une suite décroissante (par la troisième propriété) de temps d'arrêt, finis presque sûrement (par la deuxième propriété), qui tend vers 0 presque sûrement (par la première propriété).

On voit de plus que $(\Pi_{t+T_n})_{t \in \mathbb{R}^+}$ est la coagulation de la partition Π_{T_n} avec un processus n -coalescent de Kingman Γ^n (on remplace les n ancêtres de la population par n individus « neufs »).

On choisit maintenant $k \in \mathbb{N}$. Pour n assez grand, on a $\Pi_{|[k]_{T_n}} = \mathbf{0}_k$, car $T_n \xrightarrow{n \rightarrow +\infty} 0$ et $\Pi_t \xrightarrow{t \rightarrow 0} \mathbf{0}_\infty$. Par conséquent pour n assez grand on a

$$\Pi_{|[k]_{t+T_n}} = \Gamma^n_{|[k]_t}.$$

Par propriété de compatibilité, la restriction de Γ^n à $[k]$ est un k -coalescent de Kingman. Dès lors en prenant la limite pour $n \rightarrow +\infty$, on a $\Pi_{|[k]}$ est un k -coalescent de Kingman.

Dès lors Π est un coalescent de Kingman partant de la partition en singletons. \square

Passons maintenant à un cas plus général de processus coalescent, pour lequel on autorise un individu à engendrer une grande partie de la population d'un seul coup.

2.2 Cas général de coalescence échangeable et Λ -coalescent

Le coalescent de Kingman est un processus généalogique d'une population de grande taille pour laquelle un individu ne donne naissance qu'à une portion infime de la population totale. Mais pour certaines populations, cette hypothèse est fautive, et une très faible partie de la population parentale peut engendrer une descendance très grande tandis que la plupart des autres individus n'engendrent pas ou très peu d'enfants. C'est le cas par exemple d'une population de virus quand la sélection est très forte, ou bien de certaines espèces marines. Pour modéliser la généalogie de ce type de population, on autorise le regroupement en un seul bloc de plusieurs éléments d'un seul coup. On peut aussi autoriser plusieurs coagulations simultanées.

Dans ce travail, nous nous intéresserons au cas particulier suivant : lorsqu'un évènement de coagulation se présente, tous les blocs impliqués coagulent en un seul bloc (autrement dit, un seul individu a engendré une descendance très nombreuse, mais nous n'autorisons pas plusieurs évènements de coagulation simultanés). Comme précédemment pour le coalescent de Kingman, nous commençons par donner les taux de transition pour une famille finie d'individus.

L'évènement que nous venons de décrire peut être modélisé de la manière suivante : on choisit $x \in]0, 1]$ qui sera la portion des blocs coagulant en un seul. Chaque bloc aura une probabilité x d'appartenir au bloc coagulé après l'évènement, et une probabilité $1 - x$ de rester disjoint. Cet évènement arrivera à un certain taux donné par une mesure sur $]0, 1]$ ν , pour laquelle $\nu([a, b])$ représente le taux avec lequel cet évènement arrive, pour $x \in]a, b[$. Afin de considérer des processus généraux, nous ajouterons une part de coagulation de Kingman, pour laquelle deux blocs coagulent à taux constant pour toute paire de blocs. Nous pouvons donc écrire les taux de la manière suivante.

Soit ν une mesure σ -finie sur $[0, 1]$ donnant le taux auquel se produit la coagulation d'une portion x des individus présents, et $c \in \mathbb{R}^+$ le taux de coagulation bipartite. Supposons que nous partions d'une partition π contenant n blocs. Le taux de transition vers une partition π' résultant de la coagulation de k blocs de π est donné par :

$$\lambda_{n,k} = \mathbf{1}_{\{k=2\}}c + \int_{]0,1]} x^k(1-x)^{n-k}\nu(dx).$$

Remarquons que, souhaitant que tous ces taux soient finis, on imposera la condition d'intégrabilité suivante : $\int_{]0,1]} x^2\nu(dx) < +\infty$. On peut alors réécrire ces taux de manière plus simple en définissant la mesure finie sur $[0; 1]$ suivante :

$$\Lambda(dx) = c\delta_0(dx) + x^2\nu(dx).$$

Dans ce cas, le taux s'écrira (avec la convention $0^0 = 1$) :

$$\lambda_{n,k} = \int_{]0,1]} x^{k-2}(1-x)^{n-k}\Lambda(dx).$$

C'est de cette mesure finie Λ que vient le nom Λ -coalescent donné à ce processus.

Ces taux de transition nous permettent sans difficulté de définir ce qu'est un processus Λ -coalescent sur \mathcal{P}_n .

Définition 2.3. Un Λ, n -coalescent est un processus de Markov Π^n à valeurs dans \mathcal{P}_n possédant les taux de transition suivants : on passe d'une partition π contenant k blocs à une partition π' résultant de la coagulation de p blocs de π à taux $\lambda_{k,p}$.

On remarque que le premier saut (partant d'une partition de n blocs) arrive à taux $\lambda_n = \sum_{p=2}^n \binom{n}{p} \lambda_{n,p}$. Remarquons que la propriété de compatibilité est respectée par les processus Λ, n -coalescents.

Encore une fois, le processus du nombre de blocs D_t^n est un processus de mort, de taux de passage de k à $k-p+1$ étant égal à $\binom{k}{p} \lambda_{k,p}$.

Propriété 2.3. La restriction de Π^n à \mathcal{P}_{n-1} est un $\Lambda, n-1$ -coalescent, et le processus du nombre de bloc $\#\Pi^n$ est un processus de mort de taux de passage de n à $n-k$ égal à $\binom{n}{k} \lambda_{n,k}$.

Preuve. Il suffit de réaliser le même genre de calculs que dans le cas du coalescent de Kingman. Considérons Π^n un Λ, n -coalescent partant de la partition π . Nous considérons encore les deux cas suivants.

Lorsque le bloc de π contenant n contient également un autre entier $i < n$, $\Pi^n_{|[n-1]}$ est un processus de Markov avec les mêmes transitions qu'un $n-1$ -coalescent partant de $\pi_{|[n-1]}$, car on peut remonter à Π^n .

Si nous supposons $\{n\} \in \pi$, et que π possède k blocs. Soit τ le premier instant de saut (de loi exponentielle de paramètre λ_k), A l'évènement $\{\Pi_\tau^n \neq \pi\}$ sur lequel l'individu n n'est pas le seul impliqué (de probabilité $1 - \frac{(k-1)\lambda_{k,2}}{\lambda_k}$) et τ' le deuxième instant de saut (de loi conditionnelle exponentielle de paramètre $\lambda_{\#\Pi_\tau^n}$). En particulier, conditionnellement à A^c , τ' est une variable aléatoire de loi exponentielle λ_{k-1} .

Encore une fois, le premier temps de saut de $\Pi_{[n-1]}^n$ est $\tilde{\tau} = \tau + \mathbf{1}_{\{A^c\}}\tau'$. Le même calcul que précédemment nous montre sans difficultés que le premier instant de saut pour le processus restreint est de loi exponentielle de paramètre λ_{k-1} . De plus $\Pi_{[n-1]}^n$ est indépendant de $\tilde{\tau}$, et a bien la loi voulue.

On a dès lors $\Pi_{[n-1]}^n$ est un $\Lambda, n-1$ -coalescent. \square

On peut alors définir un processus Λ -coalescent à valeurs dans \mathcal{P}_∞ grâce au théorème d'extension de Kolmogorov encore une fois. Là encore on pourra faire en sorte que chacun des processus Π^n soient càdlàg, ce qui permettra que le processus défini par limite projective le soit également.

Définition 2.4. Un processus de Markov Π^Λ à valeurs dans \mathcal{P}_∞ est un Λ -coalescent si chacune de ses projections sur \mathcal{P}_n est un Λ, n -coalescent.

Comme précédemment avec le coalescent de Kingman, on va s'intéresser à l'existence d'un ancêtre commun pour cette famille infinie. Le résultat principal est dû à Schweinsberg [10], qui a donné une condition nécessaire et suffisante à la descente de l'infini pour de tels processus. Dans le cas où le Λ -coalescent descend de l'infini, nous avons aussi une borne sur l'âge de l'ancêtre commun le plus récent de toute population, quelle que soit sa taille.

Théorème 2.3 (Schweinsberg). *Soit Λ une mesure finie sur $[0, 1]$ sans masse de Dirac en 1. Pour $n \in \mathbb{N}$ on définit :*

$$\phi(n) = \sum_{k=2}^n (k-1) \binom{n}{k} \lambda_{n,k}.$$

Alors le processus Π^Λ descend de l'infini si et seulement si la série $\sum \frac{1}{\phi(n)}$ converge.

De plus l'espérance du temps d'atteinte de la partition $(\mathbb{N}, \emptyset, \dots)$ est majorée par $\sum_{n=2}^{+\infty} \frac{1}{\phi(n)}$.

La preuve de ce théorème est donnée en annexe. Remarquons que nous avons encore l'alternative suivante : soit le processus descend de l'infini presque sûrement, soit il reste infini presque sûrement. Le résultat a ensuite été précisé dans [2], qui donne la vitesse de descente de l'infini du processus, c'est-à-dire qu'il existe une fonction déterministe $v(t)$ tel que $D_t \sim_0 v(t)$.

Remarque 2.2. On peut simplement s'intéresser aux processus Λ -coalescents pour Λ une mesure de probabilité sur $[0, 1]$, quitte à changer le temps du processus coalescent. En d'autres termes on a :

$$(\Pi_{\alpha t}^\Lambda)_{t \in \mathbb{R}^+} \stackrel{(d)}{=} (\Pi_t^{\alpha \Lambda})_{t \in \mathbb{R}^+}.$$

Nous pouvons construire un processus Λ -coalescent, si Λ n'a pas de mesure de Dirac en 0, en utilisant une mesure de Poisson M sur $\mathbb{R}^+ \times [0, 1]$ d'intensité $dt \otimes \frac{1}{x^2} \Lambda(dx)$ de la manière suivante. On pose Π_t le processus défini de la manière suivante : on pose $\Pi_0 = \mathbf{0}_\infty$, et si $\delta_{t,r}$ est un atome de M , on réalise la coagulation d'une proportion r de blocs de Π_{t-} .

Pour cela on pose $(B_i)_{i \in \mathbb{N}}$ une suite de variables aléatoires de Bernoulli i.i.d. de paramètre r , et on pose π' la partition simple aléatoire pour laquelle un entier i appartient au bloc non-trivial si et seulement si $B_i = 1$. Posons enfin $\Pi_t = \text{Coag}(\Pi_{t-}, \pi')$, qui réalise bien la coagulation d'une proportion r des blocs de Π_{t-} . Le processus ainsi construit est bien un processus Λ -coalescent.

Dans tout ce que nous avons fait jusqu'ici, nous n'avons pas permis aux individus de se déplacer. Nous allons donc maintenant nous intéresser à ce cas, pour lequel les résultats que

nous avons trouvés restent localement les mêmes, mais le comportement global est différent. En particulier, il n'existe pas de borne sur l'âge du dernier ancêtre commune indépendante de la taille de la famille considérée.

3 Processus coalescent spatial

Nous allons ajouter l'existence d'un processus de migration au sein de notre population. Pour cela nous associerons à chaque individu une marche aléatoire indépendante à temps continu de taux de transition ρ sur un graphe connexe G infini. Ce processus a été introduit pour la première fois par Limic et Sturm [7].

3.1 Quelques notions sur les graphes

Soit $G = (S, V)$ un graphe infini, avec S un ensemble de sommets, et V un ensemble d'arêtes entre deux sommets. Pour $u, v \in S$, on note $u \sim v$ si u et v sont liés par une arête. On définit $d(u, v)$ la distance entre u et v par le nombre minimal d'arêtes qu'il faut parcourir pour aller de u à v .

Le degré de $u \in S$, d_u , est le nombre de voisins de u dans G . Un graphe G est dit de degré maximal fini D si tout sommet u a un degré inférieur ou égal à D .

On note $B(u, k)$ la boule de centre u et de rayon k . On note $\text{Vol } B(u, k)$ le nombre de sommets contenus dans cette boule. Si G est un graphe infini de degré maximal D , connexe, nous pouvons borner le volume de la manière suivante :

$$k \leq \text{Vol } B(u, k) \leq D^k.$$

Une marche aléatoire à temps continu sur D de taux de transition ρ est définie de la manière suivante : au bout d'un temps aléatoire exponentiel de paramètre ρ , le processus saute d'une position u à une position $v \sim u$, choisi uniformément au hasard parmi les voisins de u . Nous avons donc les taux de transition suivants :

$$\begin{cases} p_{u,v} = \frac{\rho}{d_u} & \text{si } u \sim v \\ p_{u,v} = 0 & \text{sinon.} \end{cases}$$

3.2 Définition d'un processus coalescent spatial

On définit encore de la même manière ces processus coalescents pour un nombre fini d'individus, avant de passer à la limite projective en utilisant une propriété de compatibilité. Soit $n \in \mathbb{N}$, les processus que nous étudierons prendront leurs valeurs dans $\mathcal{P}_n \times G^n = \mathcal{P}_n^G$, où l'élément (π, l) représente une famille d'individus dans laquelle l'ancêtre des individus π_i est situé à la position l_i . Nous noterons :

$$\#_v(\pi, l) = \#\{i \in [n] \mid \pi_i \neq \emptyset \text{ et } l_i = v\}.$$

Définition 3.1. Soit $n \in \mathbb{N}$ et Λ une mesure finie sur $[0, 1]$. Un processus Λ, n -coalescent spatial est un processus de Markov à valeurs dans \mathcal{P}_n^G noté $(\Pi_t^{\Lambda, n}, B_t)_{t \in \mathbb{R}^+}$ évoluant de la manière suivante :

- les blocs situés en un même sommet se comportent comme dans un Λ -coalescent non-spatial ;
- chaque bloc se déplace sur le graphe selon une marche aléatoire de taux de transition ρ indépendamment des évènements de coalescence.

On vérifie alors que la restriction d'un Λ, n -coalescent spatial est un $\Lambda, (n - 1)$ -coalescent spatial. Pour cela on sépare encore les cas entre $\{n\} \in \pi$ auquel cas on peut avoir à considérer d'autres sauts de Π^n , et $\{n\} \notin \Pi^n$ pour lequel le premier saut suffit, exactement comme dans le cas du Λ -coalescent non-spatial.

Ensuite par limite projective, on définit un Λ -coalescent comme suit :

Définition 3.2. Un Λ -coalescent spatial est un processus de Markov $(\Pi_t^\Lambda, B_t)_{t \in \mathbb{R}^+}$ à valeurs dans $\mathbb{P}_\infty \times G^{\mathbb{N}}$ tel que la restriction de Π^Λ à $[n]$ et la projection de B sur les n premières variables est un processus Λ, n -coalescent.

Il existe un tel processus, et on a unicité en loi de celui-ci pour une loi initiale donnée. Comme nous nous intéresserons maintenant uniquement à la descente de l'infini de ces processus, nous allons définir, comme dans le cas non-spatial, le processus de Markov de mort correspondant au nombre de blocs du processus coalescent.

Soit Λ une mesure finie sur $[0, 1]$. On note $(X_t^n)_{t \in \mathbb{R}^+}$ le processus de Markov à valeurs dans l'ensemble des fonctions $G \rightarrow \mathbb{N}$ qui à un sommet v associe le nombre d'individus considérés dans la projection sur $[n]$ d'un Λ -coalescent spatial correspondant. On note $(X_t)_{t \in \mathbb{R}^+}$ la limite croissante des X^n . On peut préciser les taux de transition de (X^n) :

- le taux de transition de x à $x + \delta_v - \delta_u$ est $x(u)\rho p(u, v)$;
- le taux de transition de x à $x - (k - 1)\delta_u$ est $\binom{x(u)}{k} \lambda_{x(u), k}$;

où on a posé δ_u la fonction valant 1 en u et 0 ailleurs.

On notera N_t^n le nombre total d'individus présents dans le processus coalescent à l'instant t .

3.3 Descente de l'infini : vitesse et heuristique

Dans de nombreux cas que nous étudierons, il a été prouvé que le coalescent non-spatial se concentre autour d'une fonction déterministe $v(t)$ au voisinage de 0, quand la taille de la population devient grande. L'idée est alors de partir d'une grande population en un seul point u du graphe, et de considérer que pour des instants assez petits τ_n , le nombre d'individus ayant migré est négligeable devant le nombre d'individus qui restent en place.

Nous pourrions alors estimer le nombre d'individus sédentaires par $v(\tau_n)$, et le nombre d'individus ayant migré par $f(n) = \rho \int_0^{\tau_n} v(s) ds$, uniformément distribués sur les voisins. Nous prendrions enfin un temps suffisamment petit pour qu'aucun individu ne soit à plus d'une unité de distance de u , et que les individus ayant migré n'aient pas commencé à coalescer.

Cette approximation peut être appliquée à nouveau à la configuration que nous venons d'obtenir dans laquelle il reste de l'ordre de $f(n)$ individus en chaque sommet voisin de u ainsi qu'en u . On suppose qu'il n'y a que peu d'interactions entre les individus en u et les individus sur les sites voisins de u . Jusqu'à l'instant $\tau_{f(n)}$, cette hypothèse tient. Nous pouvons appliquer encore et encore la même approximation jusqu'à ce que $f \circ \dots \circ f(n)$ soit de l'ordre de 1.

L'idée sera alors de nous arrêter légèrement plus tôt dans notre approximation afin d'obtenir que l'évolution décrite soit d'autant plus probable lorsque la taille de la population deviendra grande. Dans le processus, les individus se seront alors suffisamment étalés sur le graphe pour obtenir de bonnes estimations jusqu'au temps constant.

L'idée de la preuve du comportement aux temps longs du coalescent spatial est de partir de la situation étalée que nous venons de décrire. Dans ce cas les évènements de coalescence multiple deviennent rares, donc le processus se comporte d'avantage comme une marche aléatoire coalescente. En utilisant des résultats connus sur ces processus, il est possible de borner l'ordre de grandeur du nombre d'individus restants en un temps long.

Toutefois, avant de nous intéresser aux preuves des Théorèmes 1.1 à 1.3, nous aurons besoin de plusieurs résultats techniques, que nous avons placés ici.

4 Résultats utiles sur les coalescents spatiaux

Commençons cette série de lemmes utiles par une estimation de grandes déviations sur les sommes de variables aléatoires exponentielles indépendantes, auxquelles nous aurons souvent affaire par la suite :

Lemme 4.1. *Soit I un ensemble d'indices, $(e_i)_{i \in I}$ une suite de variables aléatoires i.i.d de loi exponentielle de paramètre 1 et $(\mu_i)_{i \in I}$ une suite de paramètres positifs vérifiant $\sum_{i \in I} \mu_i < +\infty$. Posons $S = \sum_{i \in I} \mu_i e_i$, on a pour tout $0 < \varepsilon < 1$,*

$$\mathbb{P}(S < (1 - \varepsilon)\mathbb{E}(S)) \leq \exp\left(-\frac{\varepsilon^2 \mathbb{E}(S)^2}{4\text{Var}(S)}\right) \quad (3.1)$$

et pour tout $0 < \varepsilon < \frac{2\text{Var}(S)}{\mathbb{E}(S) \sup_{i \in I} \mu_i}$,

$$\mathbb{P}(S > (1 + \varepsilon)\mathbb{E}(S)) \leq \exp\left(-\frac{\varepsilon^2 \mathbb{E}(S)^2}{4\text{Var}(S)}\right). \quad (3.2)$$

Preuve. Soit $0 < \lambda < \frac{1}{2} \inf_{i \in I} \frac{1}{\mu_i}$, on a en utilisant l'inégalité de Markov :

$$\begin{aligned} \mathbb{P}(S > (1 + \varepsilon)\mathbb{E}(S)) &\leq e^{-\lambda(1+\varepsilon)\mathbb{E}(S)} \prod_{i \in I} \mathbb{E}(e^{\lambda \mu_i e_i}) \\ &\leq e^{-\lambda(1+\varepsilon)\mathbb{E}(S)} \prod_{i \in I} \frac{1}{1 - \lambda \mu_i} \\ &< e^{-\lambda(1+\varepsilon)\mathbb{E}(S)} \exp\left(\sum_{i \in I} \lambda \mu_i + \lambda^2 \mu_i^2\right) \\ &\leq e^{-\lambda \varepsilon \mathbb{E}(S) + \lambda^2 \text{Var}(S)} \end{aligned}$$

où nous avons utilisé que pour tout $x < \frac{1}{2}$, $-\log(1 - x) < x + x^2$. Nous pouvons maintenant choisir $\lambda = \frac{\varepsilon \mathbb{E}(S)}{2\text{Var}(S)}$, ce qui n'est possible que si $\varepsilon < \frac{2\text{Var}(S)}{\mathbb{E}(S) \sup_{i \in I} \mu_i}$. Et dans ce cas nous obtenons le résultat escompté.

L'autre inégalité s'obtient de la même manière : on a pour tout $\lambda < 0$

$$\begin{aligned}
\mathbb{P}(S < (1 - \varepsilon)\mathbb{E}(S)) &\leq e^{-\lambda(1-\varepsilon)\mathbb{E}(S)} \prod_{i \in I} \mathbb{E}(e^{\lambda\mu_i e_i}) \\
&\leq e^{-\lambda(1-\varepsilon)\mathbb{E}(S)} \prod_{i \in I} \frac{1}{1 - \lambda\mu_i} \\
&< e^{-\lambda(1-\varepsilon)\mathbb{E}(S)} \exp\left(\sum_{i \in I} \lambda\mu_i + \lambda^2\mu_i^2\right) \\
&\leq e^{-\lambda\varepsilon\mathbb{E}(S) + \lambda^2\text{Var}(S)}.
\end{aligned}$$

En posant $\lambda = -\frac{\varepsilon\mathbb{E}(S)}{2\text{Var}(S)}$, nous obtenons la borne voulue. \square

Afin d'étudier les coalescents spatiaux, nous aurons souvent besoin d'étudier le comportement de sous-ensembles de la population totale, pour cela il sera utile de nous intéresser aux liens entre le nombre d'individus dans le coalescent et le nombre d'individus dans chacune des sous-populations considérées.

Soit Π un processus coalescent partant de n individus à l'instant 0. Nous séparons ces n individus initiaux en sous-ensembles B_1, \dots, B_r . On notera X^{B_k} le processus $\#\Pi_{|B_k}^n$ du nombre d'ancêtres commun à la sous-population B_k en chaque point dans le coalescent spatial et N^{B_k} le nombre total d'individus dans ce processus.

Lemme 4.2. *Pour tout $t \in \mathbb{R}^+$, nous avons :*

$$\forall v \in G, \forall k \leq r, X_t^{B_k}(v) \leq X_t(v) \leq \sum_{i=1}^r X_t^{B_i}(v)$$

et

$$\forall k \leq r, N_t^{B_k} \leq N_t \leq \sum_{i=1}^r N_t^{B_i}$$

Preuve. On compte dans X^{B_i} tous les individus dont au moins un descendant appartient à B_i . Dans ce cas la première inégalité est clairement vérifiée, on a bien $X_t^{B_k}(v) \leq X_t^n(v)$ et de plus chaque individu de Π étant compté au moins une fois dans la somme, $X_t(v) \leq \sum_{i=1}^r X_t^{B_i}(v)$.

Notons enfin que la seconde inégalité se déduit de la première en sommant sur tous les sommets v . \square

Nous allons maintenant démontrer un résultat nous permettant de lier des coalescents spatiaux et non-spatiaux. Pour cela nous allons coupler un coalescent non-spatial avec son pendant spatial partant de n individus au même endroit.

X^n est le processus du nombre de blocs d'un Λ, n -coalescent spatial partant de n individus en un sommet u . On note Z_t^n le nombre d'individus dans le coalescent qui ont quitté le sommet u avant le temps t et M_t^n le nombre d'individus encore présents au temps t qui n'ont jamais quitté ce point. Enfin notons Y^n un Λ, n -coalescent non-spatial.

Lemme 4.3. *Il existe un couplage de X^n et Y^n qui vérifie pour tout t :*

$$M_t^n \leq Y_t^n \leq M_t^n + Z_t^n \quad (3.3)$$

et

$$Y_t^n - Z_t^n \leq X_t^n(u) \leq Y_t^n + Z_t^n. \quad (3.4)$$

Preuve. Nous pouvons tout d'abord remarquer que la loi de M^n est celle d'un Λ -coalescent non spatial pour lequel les individus meurent à taux ρ . De plus Z_t^n compte alors le nombre de morts de ce processus jusqu'à l'instant t . Nous pouvons alors réaliser un couplage de Y^n et (M^n, Z^n) de telle sorte que l'inégalité (3.3) soit réalisée. En effet, dans Y^n , les « morts » continuent à coalescer avec les autres individus, et dans $M^n + Z^n$, les « morts » sont toujours comptés dans le processus, mais gelés (ils ne prennent plus part aux événements de coalescence). On obtient bien :

$$M_t^n \leq Y_t^n \leq M_t^n + Z_t^n.$$

Pour la deuxième partie, remarquons tout d'abord que M_t^n est clairement inférieur à $X_t^n(u)$, car dans X^n certains individus « morts » peuvent revenir en u , dès lors $Y_t^n - Z_t^n \leq X_t^n(u)$ grâce à l'inégalité précédente. De plus le nombre total d'individus présents dans le coalescent spatial ne peut excéder $M_t^n + Z_t^n$ puisque dans cette approximation, les individus qui ont à un moment quitté u sont considérés comme gelés. Donc en particulier, $X_t^n(u) \leq Y_t^n + Z_t^n$, nous avons l'inégalité (3.4). \square

Une étude plus fine de ce couplage est possible. Pour cela notons \mathcal{G} la tribu engendrée par Y^n , $\mathcal{F}_t = \sigma(\Pi_s^n, s \leq t)$ la filtration canonique de Π^n et soit $\mathcal{F}_t^* = \sigma(\mathcal{F}_t, \mathcal{G})$ la filtration de X connaissant Y . Les processus suivants :

$$S_t^n = Z_t^n - \int_0^t \rho M_s^n ds \text{ et } V_t^n = S_t^{n2} - \int_0^t \rho M_s ds$$

sont clairement des martingales par rapport à (\mathcal{F}_t) , mais nous pouvons aussi prouver que ce sont des martingales par rapport à (\mathcal{F}_t^*) .

Lemme 4.4. *Pour tout $a < b$, nous avons la domination stochastique suivante :*

$$\mathbb{P}(Z_b^n - Z_a^n > x | \mathcal{G}) \leq \mathbb{P}\left(\text{Poisson}\left(\rho \int_a^b Y_s^n ds\right) > x \mid \mathcal{G}\right).$$

Preuve. Connaissant \mathcal{G} , Z^n est un processus de saut pur arrivant à taux $\rho M_t^n \leq \rho Y_t^n$, d'où on tire le caractère de martingale et la domination précédente. \square

5 Divergence du coalescent de Kingman spatial

Nous allons maintenant tenter de borner le nombre total d'individus présents dans un coalescent de Kingman spatial. Soit Π^n un tel coalescent partant de n individus en un site $u \in G$, et X^n le processus du nombre de blocs de Π^n . Nous allons estimer le comportement de X au cours du temps par approximations successives.

5.1 Étude fine du coalescent de Kingman non-spatial

Puisque nous avons liés les coalescents spatiaux et non-spatiaux, nous commencerons par démontrer quelques résultats sur le coalescent de Kingman non-spatial, notamment sur les quantités Y^n et Z^n du couplage de X^n avec son pendant non-spatial du Lemme 4.3, qui nous permettront de déduire des bornes utiles pour X^n . Pour commencer, nous allons affiner l'approximation dont nous disposons sur la vitesse de descente de l'infini du coalescent de Kingman.

Lemme 5.1. *Soit t_n une suite de réels positifs décroissant vers 0 tel que $nt_n \xrightarrow{n \rightarrow +\infty} +\infty$, et Y^n un n -coalescent de Kingman non-spatial, on a pour $0 < \varepsilon < \frac{1}{2}$ et n assez grand,*

$$\mathbb{P}\left(\left|\frac{t_n}{2}Y_{t_n}^n - 1\right| > \varepsilon\right) \leq \exp\left(-\frac{\varepsilon^2}{t_n}\right). \quad (3.5)$$

Preuve. Soit $\varepsilon > 0$, nous allons calculer le temps S_n nécessaire pour passer de n individus à $m_n = \left\lfloor \frac{2(1+\varepsilon)}{t_n} \right\rfloor$. C'est une somme de temps exponentiels de paramètre $\binom{k}{2}$. On a en particulier :

$$\mathbb{E}(S_n) = \sum_{k=m_n+1}^n \frac{2}{k(k-1)} = \frac{2}{m_n} - \frac{2}{n} \sim \frac{2}{m_n}$$

et

$$\text{Var}(S_n) = \sum_{k=m_n+1}^n \left(\frac{2}{k(k-1)}\right)^2 \sim \frac{2}{3m_n^3}.$$

Nous pouvons alors utiliser le Lemme 4.1 : pour tout $\eta < \frac{m_n(m_n+1)\text{Var}(S_n)}{\mathbb{E}(S_n)} (\sim \frac{2}{3})$,

$$\mathbb{P}(S_n > \mathbb{E}(S_n)(1 + \eta)) \leq \exp\left(-\frac{\eta^2 \mathbb{E}(S_n)^2}{4\text{Var}(S_n)}\right).$$

Nous obtenons alors, pour n assez grand et $\varepsilon < \frac{1}{2}$:

$$\begin{aligned} \mathbb{P}\left(\frac{t_n}{2}Y_{t_n}^n > 1 + \varepsilon\right) &= \mathbb{P}(S_n > t_n) \\ &= \mathbb{P}\left(S_n > \left(1 + \left(\frac{t_n}{\mathbb{E}(S_n)} - 1\right)\right) \mathbb{E}(S_n)\right) \\ &\leq \exp\left(-\frac{\left(\frac{t_n}{\mathbb{E}(S_n)} - 1\right)^2 \mathbb{E}(S_n)^2}{4\text{Var}(S_n)}\right) \\ &\leq \exp\left(-\frac{3\varepsilon^2}{2t_n}(1 + o(1))\right). \end{aligned}$$

L'autre sens de l'inégalité s'obtient de la même manière en estimant la probabilité que

$$\tilde{S}_n = \sum_{k=\left\lfloor \frac{2(1-\varepsilon)}{t_n} \right\rfloor}^n \frac{2}{k(k-1)} e_i$$

soit inférieure à t . On conclut une fois encore en utilisant l'inégalité (3.1) du Lemme 4.1. \square

Nous aurons également besoin d'une estimation du nombre total d'« émigrants » qui quittent u au cours du temps, autrement dit de la quantité Z_∞^n dans notre précédente écriture. Pour ce qui est de l'heuristique, il est assez simple de voir que cette quantité est de l'ordre de $2\rho \log n$. En effet, entre t et $t + dt$, Z_t^n croit d'environ $\rho X_t(u)dt$ individus.

Étant donné qu'il y a au début bien plus d'individus dans notre coalescent spatial, nous obtenons $Z_\infty^n \sim \int_0^M \rho X_s(u)ds$, où M est une constante arbitraire. Or on sait qu'il y a au voisinage de $t = 0$ de l'ordre de $\frac{2}{t} \wedge n$ individus à l'origine, ce qui nous donne notre approximation. Plus précisément, nous allons démontrer le résultat de grandes déviations suivant.

Lemme 5.2. *Soit $\varepsilon > 0$, il existe $c, C > 0$ tel que :*

$$\mathbb{P}\left(\left|\frac{Z_\infty^n}{\log n} - 2\rho\right| > \varepsilon\right) < Cn^{-c}. \quad (3.6)$$

Et de plus pour tout $M > 0$,

$$\mathbb{P}(Z_\infty^n > M) < Cn^C e^{-M}. \quad (3.7)$$

Preuve. Soit Π^n un n -coalescent de Kingman spatial, on considère la suite $(T_k)_{0 \leq k \leq n-1}$ des temps auxquels il arrive quelque chose aux individus n'ayant jamais quitté u . C'est une suite de temps d'arrêt. De plus, l'évènement est soit un évènement de coagulation, soit un évènement de migration, donc à chaque instant T_k , le nombre d'individus n'ayant jamais quitté u diminue de 1, car soit deux lignées ont fusionné, soit un individu a quitté u . A l'instant T_k , il reste donc $n - k$ individus qui n'ont jamais quitté le sommet u .

Le $k + 1$ ^{ième} temps d'arrêt T_k suit donc une loi exponentielle de paramètre

$$\frac{(n-k)(n-k-1)}{2} + \rho(n-k) = \frac{(n-k)(n-k-1+2\rho)}{2}.$$

De plus l'évènement correspond à une mutation avec probabilité $\frac{2\rho}{n-k-1+2\rho}$ et à une fusion avec probabilité $\frac{n-k-1}{n-k-1+2\rho}$. Posons alors ξ_k une variable aléatoire de Bernouilli de paramètre $\frac{2\rho}{k-1+2\rho}$, nous avons alors :

$$Z_\infty^n \stackrel{(d)}{=} \sum_{i=1}^n \xi_i.$$

Le reste est une simple affaire de grandes déviations : pour tout $\lambda > 0$, nous avons :

$$\begin{aligned} \mathbb{E}(e^{-\lambda Z_\infty^n}) &= \prod \left(1 - \frac{2\rho}{i+2\rho}(1 - e^{-\lambda})\right) \\ &\leq \exp\left(-\sum_{i=1}^n \frac{2\rho}{i+2\rho}(1 - e^{-\lambda})\right) \\ &< \exp\left(-2\rho(1 - e^{-\lambda})(C + \log n)\right). \end{aligned}$$

Donc en utilisant l'inégalité de Markov, nous avons :

$$\begin{aligned} \mathbb{P}(Z_\infty^n < 2(1 - \varepsilon)\rho \log n) &\leq \exp(2\lambda((1 - \varepsilon)\rho \log n)\mathbb{E}(e^{-\lambda Z_\infty^n})) \\ &< \exp\left(2\lambda((1 - \varepsilon)\rho \log n - 2\rho(1 - e^{-\lambda})(C + \log n))\right) \\ &\leq \exp\left((-2\lambda\rho\varepsilon + O(\lambda^2)) \log n + C\right). \end{aligned}$$

Par conséquent pour λ assez petit, le coefficient devant $\log n$ est négatif, donc on peut majorer cette probabilité par une puissance négative de n , on obtient le premier sens de l'inégalité 3.6.

Nous appliquons maintenant l'inégalité de Markov à $e^{\lambda Z_\infty^n}$. Nous avons de la même manière :

$$\begin{aligned} \mathbb{P}(Z_\infty^n > M) &\leq \exp(-\lambda M)\mathbb{E}(e^{\lambda Z_\infty^n}) \\ &< \exp\left(-\lambda M - 2\rho(e^\lambda - 1)(C + \log n)\right). \end{aligned}$$

Dès lors, pour $M = 2\rho(1 + \varepsilon) \log n$ et λ assez petit, nous avons l'autre sens de l'inégalité 3.6, et pour $\lambda = 1$, nous avons également l'inégalité 3.7. \square

5.2 Estimations du comportement du coalescent de Kingman spatial

Rappelons que nous avons démontré précédemment que pour des temps suffisamment petits, $X_t(u)$ se comporte comme $\frac{2}{t}$, et que le nombre d'individus quittant u est de l'ordre de $2\rho \log n$. Nous allons montrer que cette approximation tient jusqu'à un temps τ_n , pour lequel avec grande probabilité, il restera $\frac{2}{\tau_n}$ individus en u , et $2\rho \log n$ individus auront migré, et se seront uniformément distribués sur les voisins de u .

Commençons par un premier lemme sur le comportement du coalescent de Kingman spatial immédiatement après son départ de l'origine.

Lemme 5.3. *Soit $a_0, a_1, \varepsilon > 0$, on pose $\tau_n = a(\log n)^{-3}$ pour $a \in [a_0, a_1]$ et on définit l'évènement*

$$A = \{\forall v \in G, X_{\tau_n}^n(v) \in [(1 - \varepsilon)Q(v), (1 + \varepsilon)Q(v)]\}$$

où

$$Q(v) = \begin{cases} \frac{2}{\tau_n} & v = u \\ \frac{2\rho}{d_u} \log n & v \sim u \\ 0 & \text{sinon.} \end{cases}$$

Alors il existe C dépendant uniquement de ε, a_0, a_1 et d_u tel que $\mathbb{P}(A^c) \leq \frac{C}{\log n}$.

Preuve. Nous posons Z_t^n le nombre d'individus distincts ayant quitté u sur l'intervalle $[0, t]$ (chaque individu n'est compté qu'au plus une fois), et Y_t^n le nombre total d'individus dans le couplage de notre processus avec le coalescent non-spatial défini dans le Lemme 4.3. Nous avons :

$$Y_t^n - Z_t^n \leq X_t^n(u) \leq Y_t^n + Z_t^n \text{ p.s.}$$

Nous avons maintenant besoin d'estimer $Y_{\tau_n}^n$ et $Z_{\tau_n}^n$. Or le comportement de $Y_{\tau_n}^n$ est connu, grâce au Lemme 5.1, nous savons que :

$$\mathbb{P}\left(\left|\frac{\tau_n}{2}Y_{\tau}^n - 1\right| > \varepsilon\right) \leq Ce^{-\frac{\varepsilon}{\tau_n}} \leq \frac{C}{n}.$$

De plus, grâce à l'inégalité (3.6) du Lemme 5.2, nous avons :

$$\mathbb{P}\left(\left|\frac{Z_{\infty}^n}{\log n} - 2\rho\right| > \frac{\varepsilon}{2}\right) \leq Cn^{-c}.$$

Il nous reste à borner $Z_{\infty}^n - Z_{\tau_n}^n$, le nombre d'individus quittant le sommet u après l'instant τ_n . Rappelons que M_t^n est le nombre d'individus qui sont restés en u tout au long de l'intervalle $[0, t]$. Nous pouvons alors utiliser les équations (3.5) et (3.7) et le Lemme 4.3 pour obtenir

$$\begin{aligned} \mathbb{P}(Z_{\infty}^n - Z_{\tau_n}^n > \frac{\varepsilon}{2} \log n) &\leq \mathbb{P}\left(Y_{\tau_n}^n > \frac{3}{\tau_n}\right) + \mathbb{P}\left(Z_{\infty}^n - Z_{\tau_n}^n > \frac{\varepsilon}{2} \log n \text{ et } Y_{\tau_n}^n < \frac{3}{\tau_n}\right) \\ &\leq \mathbb{P}\left(Y_{\tau_n}^n > \frac{3}{\tau_n}\right) + \mathbb{P}\left(Z_{\infty}^n - Z_{\tau_n}^n > \frac{\varepsilon}{2} \log n \mid M_{\tau_n}^n < \frac{3}{\tau_n}\right) \\ &\leq \frac{C}{n} + C\left(\frac{3}{\tau_n}\right)^C e^{-\frac{\varepsilon}{2} \log n} \\ &\leq \frac{C}{n} + \frac{C(\log n)^C}{n^{\frac{\varepsilon}{2}}} \\ &\leq \frac{C}{\log n}. \end{aligned}$$

Dès lors nous avons également, avec probabilité au moins $1 - \frac{C}{\log n}$, $|Z_{\tau_n}^n - 2\rho \log n| < \varepsilon \log n$ et $\left|Y_{\tau_n}^n - \frac{2}{\tau_n}\right| < \frac{2}{\tau_n}\varepsilon$. En particulier, le nombre d'émigrants est négligeable devant le nombre d'individus restants sur place. Grâce à l'inégalité vérifiée par $X_t^n(u)$, nous obtenons la bonne borne pour cette quantité, pour n assez grand, donc pour tout n quitte à changer la constante C .

De plus, conditionnellement à $Z_{\tau_n}^n$, la probabilité qu'il y ait plus d'un seul saut avant le temps τ_n est au plus de l'ordre de $2\rho\tau_n Z_{\tau_n}^n$, soit avec grande probabilité au plus de l'ordre de $C(\log n)^{-2}$. Avec probabilité $1 - \frac{C}{\log n}$, chaque individu fait donc au plus un saut. De même la probabilité qu'un évènement de coagulation impliquant certains des individus ayant migré avant l'instant τ_n arrive avant ce même instant est également dominé par $C(\log n)^{-1}$, donc aucun n'arrive avant le temps τ_n avec grande probabilité.

Enfin, $Z_{\tau_n}^n$ est concentré au voisinage de $2\rho \log n$, et conditionnellement au nombre total d'émigrants, les destinations des individus sont indépendantes et identiquement distribuées selon la loi $p(u, \cdot)$. Par conséquent, comme $p(u, v) = \frac{1}{d_u}$ pour $v \sim u$, nous avons, grâce au théorème central limite, que la probabilité avec laquelle il existe un décalage d'au moins $\varepsilon \log n$ par rapport aux valeurs moyennes est dominée par $\frac{C}{n^c}$.

Par conséquent, nous pouvons majorer la probabilité de A^c par $\frac{C}{\log n}$. \square

Passons maintenant à la seconde partie de notre raisonnement. Grâce à ce lemme, nous pouvons contrôler avec grande probabilité le comportement du coalescent spatial, ce qui nous permet de l'amener en un temps τ_n d'une population concentrée en un point avec n individus, à une population concentrée sur ce point et ses voisins immédiats avec de l'ordre de $(\log n)$ individus en chaque point, et $(\log n)^3$ en u . L'idée est alors de ré-appliquer ce lemme entre les instants τ_n et $\tau_n + \tau_{\log n}$, et continuer les itérations jusqu'à ce que cette approximation ne soit plus valide. C'est ce que nous allons faire maintenant.

Rappelons que $\log^{(k)} n = \underbrace{\log \circ \dots \circ \log}_k n$. Dans ce cas, soit $n \in \mathbb{N}$, pour $k \in \mathbb{N}$ on pose $t_0 = 0$

et $t_{k+1} = t_k + (\log^{(k+1)} n)^{-3}$. Posons ensuite $m_n = \log^* n$ le pas jusqu'auquel nous aimerions poursuivre notre approximation pour obtenir le théorème. Pour des raisons de calcul, nous ne les pousserons que jusqu'à l'étape m'_n , où on a posé :

$$m'_n = \min \left\{ k \in \mathbb{N} : \log^{(k)} n < \text{Vol } B(u, m_n)^2 \right\}.$$

Observons tout d'abord que pour n assez grand, $\text{Vol } B(u, m_n)^2 > 10$, donc à partir de ce rang, $m'_n \leq m_n$. De plus si D est le degré maximal du graphe, alors en majorant le volume de la boule par D^{m_n} , on obtient :

$$\begin{aligned} m'_n &> \min \left\{ k \in \mathbb{N} : \log^{(k+1)} n < D^{2m_n} \right\} \\ &\geq \min \left\{ k \in \mathbb{N} : \log^{(k+2)} n < 2(\log D)m_n \right\} \\ &\geq \min \left\{ k \in \mathbb{N} : \log^{(k+\log^*(Cm_n))} n < 1 \right\} \\ &\geq m_n - \log^*(Cm_n + 1). \end{aligned}$$

Soit $m_n - m'_n < \log^*(C(\log^* n)) = o(\log^* n)$, donc m_n et m'_n sont équivalents.

Dernier fait, nous pouvons minorer $\log^{(m'_n)} n$ grâce à la définition de m'_n , en utilisant le fait que, dans un graphe connexe, $k \leq \text{Vol } B(u, k)$:

$$m_n^2 \leq \log^{m'_n} n,$$

donc en particulier cette quantité tend vers l'infini quand n tend vers l'infini. Nous pouvons donc appliquer le Lemme 5.3 m'_n fois à la suite et espérer garder une bonne approximation, jusqu'à l'instant $t_{m'_n}$, il y a toujours un très grand nombre d'individus en chaque point considéré.

Lemme 5.4. Soit B_k l'évènement défini par :

$$B_k = \left\{ \forall v \in G, X_{t_k}^n(v) \in \left[\left(\frac{1}{D} - \varepsilon \right) Q_k(v), (D + \varepsilon) Q_k(v) \right] \right\},$$

où D est le degré maximal de G , et Q_k est défini par :

$$Q_k(v) = \begin{cases} (\log^{(k)} n)^3 & d(u, v) < k \\ 2\rho \log^{(k)} n & d(u, v) = k \\ 0 & \text{sinon.} \end{cases}$$

On a alors :

$$\mathbb{P} \left(\bigcap_{k=1}^{m'_n} B_k \right) \xrightarrow{n \rightarrow +\infty} 1.$$

Preuve. Nous allons utiliser l'inégalité suivante, valable en toute généralité :

$$\mathbb{P} \left(\bigcap_{k=1}^{m'_n} B_k \right) \leq \mathbb{P}(B_1^c) + \sum_{k=2}^{m'_n} \mathbb{P}(B_k^c \cap B_{k-1}) \leq \mathbb{P}(B_1^c) + \sum_{k=2}^{m'_n} \mathbb{P}(B_k^c | B_{k-1}).$$

Le Lemme 5.3 nous donne déjà une borne sur $\mathbb{P}(B_1^c)$, l'idée est ici de majorer $\mathbb{P}(B_{k+1}^c | B_k)$ en utilisant successivement le Lemme 4.2 pour nous ramener d'un système dans lequel les individus sont dispersés sur une boule de rayon k à $\text{Vol } B(u, k)$ systèmes, chacun partant d'un certain nombre d'individus en un point, sur lequel on a des bornes grâce à B_k , puis le Lemme 5.3 pour caractériser l'évolution de chacun de ces sommets.

Supposons que B_k soit vérifié, dans ce cas, à l'instant t_k nous avons à chaque site de $B(u, k)$ un nombre d'individus compris entre $\frac{\log^{(k)} n}{2D}$ et $2D(\log^{(k)} n)^3$ (en choisissant $\varepsilon > 0$ assez petit pour que ces bornes soient vérifiées).

Nous séparons alors les individus en sous-groupes selon leur localisation géographique à l'instant t_k , et considérons l'évolution entre les instants t_k et t_{k+1} de chacun de ces sous-ensembles indépendamment des autres. Le Lemme 5.3 peut être appliqué, car pour chacun de ces sommets, le temps $\tau = (\log^{k+1} n)^{-3}$ est bien compris entre $a_0(\log q)^{-3}$, et $a_1(\log q)^{-3}$ (où q le nombre d'individus est compris entre $\frac{\log^{(k)} n}{2D}$ et $2D(\log^{(k)} n)^3$). En choisissant n assez grand, il existe a_0 et a_1 tel que cette estimation tienne uniformément pour tout $k \leq m'_n$ (car $\log m'_n \xrightarrow{n \rightarrow +\infty} +\infty$).

Dans ces conditions, nous appliquons donc le Lemme 5.3, et pour chacune des évolutions que nous étudions, le nombre d'individus qui restent en place est d'environ $(1 \pm \varepsilon)(\log^{(k+1)} n)^3$, et environ $\frac{1 \pm \varepsilon}{d_v} \log q$ individus se déplacent vers les sites voisins, avec probabilité $1 - C(\log q)^{-1}$, où q est le nombre d'individus présents en v . Dès lors, étant donnée la borne sur q induite par B_k , la probabilité que, sachant que B_k est réalisé, que cet évènement ne se réalise pas pour un sommet v fixé est dominée par $\frac{C}{\log^{(k+1)} n - \log 2D} \leq \frac{C'}{\log^{(k+1)} n}$.

Remarquons alors que si B_k est vérifié, et que ces égalités sont vérifiées pour tout $v \in B(u, k)$, alors grâce au Lemme 4.2, B_{k+1} est aussi réalisé (dès lors que n est choisi suffisamment grand, indépendamment de k). Dès lors la probabilité que sachant B_k , B_{k+1} ne soit pas réalisé est majorée par :

$$\frac{C \text{ Vol } B(u, k)}{\log^{(k+1)} n} \leq \frac{C \text{ Vol } B(u, m_n)}{\log^{(k+1)} n}.$$

Ensuite, en utilisant l'inégalité citée plus haut, la probabilité qu'il existe un k tel que B_k ne soit pas vérifié est dominée par :

$$C \text{ Vol } B(u, m_n) \sum_{k=1}^{m'_n} \frac{1}{\log^{(k+1)} n} \leq \frac{C}{\log^{(m'_n+1)} n} \sum_{k=0}^{m'_n} \frac{1}{2^k} \leq \frac{2C}{\log^{(m'_n+1)} n}.$$

Grâce à notre choix de m'_n , cette quantité est au plus $\frac{2C}{\text{Vol } B(u, m_n)}$, qui tend vers 0 quand n tend vers $+\infty$, ce qui conclut notre preuve. \square

5.3 Extension de l'estimation à temps fini

On a obtenu une bonne description du comportement du coalescent de Kingman spatial jusqu'au temps $t_{m'_n}$, qui est $o(1)$. Nous allons étendre cette estimation jusqu'à un temps fini. La minoration s'obtient sans trop de difficultés, mais pour la majoration, nous aurons besoin des outils suivants. Montrons tout d'abord que les individus, après l'instant $t_{m'_n}$ ne s'éloignent plus beaucoup du point u .

Lemme 5.5. *Soit $\varepsilon, t > 0$, avec grande probabilité il n'y a pas d'individus en dehors de la boule $B(u, (1 + \varepsilon)m_n)$ à ou avant l'instant t .*

Preuve. Le Lemme 5.4 nous indique qu'avec grande probabilité, à l'instant $t_{m'_n}$, il y a au plus $3 \text{Vol } B(u, m'_n)(\log^{(m'_n)} n)^3$ individus tous situés dans la boule de centre u et de rayon m'_n . Rappelons que $\log^{(m'_n)} n \leq D^{2m_n}$, et $m'_n \leq m_n$ donc le nombre total d'individus est dominé par CD^{7m_n} .

Nous allons majorer le nombre d'individus présents à l'instant t en ignorant simplement les événements de coagulation. Chaque individu fait alors avant l'instant t un nombre de sauts suivant une loi de Poisson de paramètre $\rho(t - t_{m'_n})$. La probabilité qu'il existe un individu qui fasse plus de εm_n sauts avant l'instant t est alors dominée par $CD^{7m_n} \frac{(\rho t)^{-\varepsilon m_n}}{|\varepsilon m_n|!}$.

Cette dernière quantité tend vers 0, nous obtenons bien le résultat escompté. \square

Nous allons maintenant étudier comment se comportent les individus conditionnés à rester dans un espace clos (ici la boule de centre u et de rayon $(1 + \varepsilon)m_n$, dans ce cas, les épisodes de coagulation sont assez fréquents pour obtenir une décroissance similaire à celle du coalescent de Kingman non-spatial.

Lemme 5.6. *Soit $t < 2$, et un ensemble S de sommets, supposons qu'à l'instant 0 tous les individus sont dans S . On note Z_t le nombre d'individus qui sont sortis de S au moins une fois entre les instants 0 et t , et N_t^S le nombre d'individus dans S à l'instant t . Dès lors pour tout $\varepsilon > 0$, il existe $c > 0$ tel que :*

$$\mathbb{P} \left(N_t^S > Z_t + \frac{4 + \varepsilon}{t} |S| \right) \leq e^{-\frac{c|S|}{t}}.$$

Preuve. On pose Q_t le nombre d'individus restés dans S entre les instants 0 et t , on a clairement $N_t^S \leq Z_t + Q_t$. Par conséquent,

$$\mathbb{P} \left(N_t^S > Z_t + \frac{4 + \varepsilon}{t} |S| \right) \leq \mathbb{P} \left(Q_t > \frac{4 + \varepsilon}{t} |S| \right).$$

Il nous suffit donc de majorer Q_t pour conclure. Pour cela remarquons que le taux de coagulation total à l'instant s pour les individus dénombrés dans Q_t est :

$$\sum_{v \in S} \binom{\tilde{X}_s(v)}{2} \geq |S| \binom{Q_s}{2},$$

en utilisant la convexité de $x \mapsto \binom{x}{2}$ et en posant $\tilde{X}_s(v)$ le nombre d'individus présents en v n'ayant jamais quitté l'ensemble S .

Sur l'ensemble $Q_t < 2|S|$, comme $t < 2$, il n'y a rien à démontrer. Sur l'ensemble $Q_s > 2|S|$ pour tout $s < t$, nous remarquons que le taux de coagulation est plus grand que $\frac{1}{2|S|} \binom{Q_s}{2}$, donc Q_t est dominé par un coalescent de Kingman non spatial ralenti par un facteur $2|S|$. Le Lemme 5.1 nous permet de conclure, car Q_t est concentré au voisinage de $\frac{4|S|}{t}$ pour n assez grand. \square

Tous ces résultats nous permettent finalement de démontrer le Théorème 1.1

Preuve du Théorème 1.1. La borne supérieure s'obtient simplement grâce aux deux lemmes précédents. On voit grâce au Lemme 5.6 qu'aucun individu ne quitte $B(u, (1 + \varepsilon)m_n)$ et grâce au Lemme 5.5 que le nombre d'individus dans cette boule est dominé par une constante multipliée par le volume de cette boule.

Pour la borne inférieure, notons qu'à l'instant $t_{m'_n}$ il y a au moins un individu en chaque point de la boule $B(u, m'_n)$, choisissons-en une pour chaque sommet. La probabilité qu'un de ces individus ne bouge avant l'instant t est indépendante du mouvement des autres individus et majorée par $e^{-\rho t}$. Par conséquent, le nombre total d'individus présents à l'instant t étant supérieur au nombre d'individus que nous avons choisis et qui n'ont pas bougés. Ce nombre est minoré avec une grande probabilité par une constante multipliée par le volume de la boule. Enfin, nous avons, pour n assez grand, $m_n - m'_n \leq \varepsilon m_n$.

Nous en tirons qu'avec grande probabilité, le nombre d'individus présents à l'instant t est compris entre $\text{Vol } B(u, (1 - \varepsilon)m_n)$ et $\text{Vol } B(u, (1 + \varepsilon)m_n)$. \square

Notons le fait intéressant suivant, utile pour la preuve du Théorème 1.4 : un coalescent de Kingman spatial partant de n individus en u domine de manière stochastique à l'instant t une collection de variables aléatoires de Bernoulli i.i.d. distribuées sur $B(u, m_n)$.

Remarque 5.1. Si on fait quelques hypothèses sur la forme du graphe, par exemple dans le cas où ce graphe est \mathbb{Z}^d , nous pouvons préciser notre propos. En effet dans ce cas, le volume d'une boule de centre u et de rayon k est de l'ordre de $c_0 k^d$. Dans ce cas nous pouvons donner un ordre de grandeur du nombre d'individus présents à l'instant t , entre $\text{clog}^* n^d$ et $C \text{clog}^* n^d$.

6 Divergence générale d'un processus Λ -coalescent

Nous nous intéressons maintenant à un Λ -coalescent spatial, où Λ est une mesure finie sur $[0, 1]$. Nous allons nous appliquer à montrer la divergence de ce processus vers $+\infty$. Pour cela, intéressons-nous tout d'abord à la taille de l'arbre généalogique d'un tel processus coalescent, stoppé au dernier ancêtre commun.

6.1 Longueur totale de l'arbre généalogique du Λ -coalescent non-spatial

Pour cela, soit Λ une mesure finie sur $[0, 1]$, on pose (Y_t) le processus du nombre de blocs d'un Λ -coalescent non-spatial, et (Y_t^n) le processus du nombre de blocs de sa restriction à n .

Nous allons tout d'abord nous intéresser à :

$$K_t^n = \int_0^t (Y_s^n - 1) ds,$$

qui est une quantité qui approxime bien la taille de l'arbre généalogique de la population. Nous nous intéressons à cette quantité, car si pour t petit, Y^n est une bonne estimation du nombre de blocs présents à l'origine du graphe, alors ρK^n est une bonne estimation du nombre d'émigrants dans le processus.

Lemme 6.1. *Pour tout $t > 0$ fixé, $K_t^n \xrightarrow[n \rightarrow +\infty]{} +\infty$ p.s.*

Preuve. Soit Π le Λ -coalescent spatial, on note \sim^t la relation d'équivalence induite par la partition Π_t sur \mathbb{N} . Soit $n \in \mathbb{N}$, on pose :

$$\tau_n = \inf\{t > 0 \mid \exists j < n : n \sim^t j\}$$

l'instant auquel l'individu n fusionne pour la première fois dans le Λ, n -coalescent.

On a alors $Y_s^n = Y_s^{n-1} + \mathbf{1}_{\{s < \tau_n\}}$, donc par conséquent :

$$K^n = K^{n-1} + \tau_n \wedge t.$$

On pose alors $\mathcal{F}_n = \sigma(Y^k, k \leq n)$, dans ce cas, conditionnellement à \mathcal{F}_{n-1} , le taux de coagulation de l'individu n avec un individu d'indice plus petit est à l'instant s égal à :

$$\int_{[0,1]} \frac{1}{x^2} x(1 - (1-x)^{Y_s^{n-1}}) d\Lambda(x).$$

On utilise alors que $(1-x)^p \geq 1 - px$, on obtient que le taux de coagulation est au plus égal à Y_s^{n-1} . Dans ce cas nous obtenons :

$$\begin{aligned} \mathbb{E}(t \wedge \tau_n | \mathcal{F}_{n-1}) &= \int_0^t \mathbb{P}(\tau_n > s | \mathcal{F}_{n-1}) ds \\ &\geq \int_0^t \exp\left(-\int_0^s Y_u^{n-1} du\right) ds \\ &\geq \int_0^t \exp\left(-s - \int_0^s Y_u^{n-1} - 1 du\right) ds \\ &= e^{-K^{n-1}} \int_0^t e^{-s} ds = e^{-K_t^{n-1}} (1 - e^{-t}). \end{aligned}$$

Notons maintenant que K^n est croissant en n (la limite existe donc presque sûrement), il suffit donc de montrer que pour tout $A > 0$, le temps d'arrêt par rapport à (\mathcal{F}_n) :

$$T^A = \inf\{n \geq 1 \mid K^n > A\}$$

est fini presque sûrement. Soit M_n la martingale définie par :

$$M_n = K^n - \sum_{k=2}^n \mathbb{E}(\tau_k \wedge t | \mathcal{F}_{k-1}).$$

Par théorème d'arrêt appliqué à M_n , on a :

$$\mathbb{E}(M_{n \wedge T^A}) = 0.$$

De plus sur l'ensemble $K^n < A$, $\mathbb{E}(t \wedge \tau_n | \mathcal{F}_{n-1})$ est minorée par $e^{-A}(1 - e^{-t})$. On obtient donc :

$$\begin{aligned} \mathbb{E}(K^{n \wedge T^A}) &= \mathbb{E} \left(\sum_{k=1}^{n \wedge T^A} \mathbb{E}(t \wedge \tau_k | \mathcal{F}_{k-1}) \right) \\ &\geq c \mathbb{E}(n \wedge T^A) - 1. \end{aligned}$$

On en tire $c \mathbb{E}(n \wedge T^A) \leq 1 + \mathbb{E}(K^{n \wedge T^A}) \leq 1 + A + t$, car $A \leq K^{n \wedge T^A} \leq A + t$. Par conséquent nous obtenons en passant à la limite $\mathbb{E}(T^A) < +\infty$, d'où $T^A < +\infty$ p.s.

On conclut que $K^n \xrightarrow[n \rightarrow +\infty]{} +\infty$ p.s. \square

6.2 Divergence du coalescent spatial

Rappelons les notations employées dans le Lemme 4.3, (X^n, Y^n) les processus couplés des nombres de blocs dans le Λ, n -coalescent spatial et non-spatial. Z^n est le processus du nombre d'immigrants hors du sommet u , et M_t le processus du nombre d'individus restés en u tout au long de $[0, t]$.

Lemme 6.2. *Nous avons $Z_t^n \xrightarrow[n \rightarrow +\infty]{} +\infty$ p.s.*

Preuve. Comme Z_t^n est croissant en n , pour montrer $Z_t^n \xrightarrow[n \rightarrow +\infty]{} +\infty$ p.s. il nous suffit de montrer que $\forall A > 0, \mathbb{P}(Z_t^n \leq A) \xrightarrow[n \rightarrow +\infty]{} 0$.

Rappelons pour commencer que $S_t^n = Z_t^n - \int_0^t \rho M_s^n ds$ est une \mathcal{F}^* -martingale. Sur l'évènement $\{Z_t^n \leq A\}$, nous avons pour tout $s \leq t$:

$$M_s^n \geq Y_s^n - Z_s^n \geq Y_s^n - A.$$

Par conséquent, $S_t^n \leq A + \rho A t - \int_0^t \rho Y_s^n ds$. Grâce au Lemme 6.1, pour n assez grand, sur l'évènement $\{Z_t^n \leq A\}$, nous avons :

$$S_t^n \leq -\frac{1}{2} \int_0^t \rho Y_s^n ds.$$

On a alors en utilisant l'inégalité maximale de Doob, sur les variables aléatoires conditionnées par l'évolution de Y^n pour obtenir :

$$\mathbb{P}(Z_t^n \leq A | \mathcal{G}) \leq \mathbb{P} \left(\sup_{s \leq t} |S_s^n| \geq \frac{1}{2} \int_0^t \rho Y_s^n ds \middle| \mathcal{G} \right) \leq \frac{16}{\int_0^t \rho Y_s^n ds}.$$

On utilise encore le fait que $\int_0^t Y_s^n ds \xrightarrow[n \rightarrow +\infty]{} +\infty$ p.s. pour conclure, on a bien $\mathbb{P}(Z_t^n \leq A) \xrightarrow[n \rightarrow +\infty]{} 0$. \square

Nous allons maintenant nous intéresser à un évènement local : la probabilité qu'en un point v , au moins m individus soient réunis avant l'instant t . Notons un tel évènement :

$$E_{m,t,v}^n = \left\{ \sup_{[0,t]} X_s^n(v) \geq m \right\}.$$

Lemme 6.3. *Pour tout $v \in G$, pour tout $t > 0$ et tout $m \in \mathbb{N}$, on a :*

$$\mathbb{P}(E_{m,t,v}^n) \xrightarrow{n \rightarrow +\infty} 1.$$

Preuve. Soit $v \in G$, $t > 0$, $m \in \mathbb{N}$. Remarquons pour commencer que si $v = u$, le résultat est évident. Nous allons démontrer ensuite le résultat pour v voisin de u . Soit $\varepsilon > 0$, on pose :

$$\tau = \varepsilon \min\{(\rho m)^{-1}, \lambda_m^{-1}, t\}.$$

On sait que $Z_\tau^n \xrightarrow{n \rightarrow +\infty} +\infty$, donc de même, par loi des grands nombres, il existe N assez grand tel qu'avec probabilité au moins $1 - \varepsilon$, v reçoive au moins m individus de u si on commence avec au moins N individus.

Concentrons-nous sur ces m individus, ils quitteront v à taux ρ , et un évènement de coalescence interviendra entre plusieurs d'entre eux à taux λ_m . Par définition du temps τ , aucuns de ces individus ne coalesceront entre eux ou ne quitteront v avec probabilité au moins $1 - 2\varepsilon$.

Dès lors, pour tout $n \geq N$, on a $\mathbb{P}(X_\tau^n(v) \geq m) \geq 1 - 3\varepsilon$. ε étant choisi arbitrairement petit, on obtient bien :

$$\mathbb{P}(E_{m,t,v}^n) \xrightarrow{n \rightarrow +\infty} 1.$$

Pour v situé à distance k de u , on choisit w voisin de v tel que $d(u, w) = k - 1$. On suppose par récurrence que pour tout $m' \in \mathbb{N}$, $\mathbb{P}(E_{m', \frac{t}{2}, w}^n) \xrightarrow{n \rightarrow +\infty} 1$. Dans ce cas, on utilise la propriété de Markov forte, et on raisonne comme précédemment en faisant partir un nombre m' assez grand d'individus de w . \square

Ce dernier lemme nous permet de démontrer la non-descente de l'infini de tout Λ -coalescent spatial. En effet on sait contrôler la probabilité qu'au moins m individus passent en un point donné pour tout point. L'idée sera maintenant de contrôler la probabilité qu'au moins m individus soient simultanément présents à l'instant t , en des endroits suffisamment éloignés les uns des autres pour qu'avec une très grande probabilité, ces derniers n'interagissent pas.

Théorème 6.1. $N_t^n \xrightarrow{n \rightarrow +\infty} +\infty$ *p.s.*

Preuve. En utilisant encore une fois la monotonie en n et t de N_t^n , il nous suffit de montrer que pour tout $t > 0$ et $m \in \mathbb{N}$, $\mathbb{P}(N_t^n \geq m) \xrightarrow{n \rightarrow +\infty} 1$.

Soit $\varepsilon, \eta > 0$ et $m \in \mathbb{N}$ fixé. On pose S un sous-ensemble de m sommets de G tel que la distance entre deux sommets soit d'au moins $\frac{1}{\eta}$. Grâce au Lemme 6.3 nous savons que :

$$\forall v \in S, \mathbb{P}(E_{1,\varepsilon,v}^n) \xrightarrow{n \rightarrow +\infty} 1.$$

Notons maintenant $A_{v,\varepsilon}$ la probabilité que le premier (s'il y en a un) des individus qui rentrent au point v avant l'instant ε reste en v jusqu'à l'instant ε . Cette probabilité est au moins égale à $e^{-\rho\varepsilon}$ si un tel individu existe.

Par conséquent, en choisissant ε assez petit et n assez grand, nous voyons qu'avec probabilité au moins $1 - \eta$, à l'instant ε chacun des m sommets de S possède au moins un individu. De plus conditionnellement à cet évènement de forte probabilité, la probabilité que deux de ces mêmes individus, situés à l'instant ε à une distance d'au moins $\frac{1}{\eta}$ les uns des autres fusionnent avant l'instant t tend vers 0 quand $\eta \rightarrow 0$.

On en tire bien $\mathbb{P}(N_t^n \geq m) \xrightarrow[n \rightarrow +\infty]{} 1$. □

Remarque 6.1. Notons pour conclure que même si nous sommes partis de la situation initiale comptant n individus en un point pour prouver la divergence du Λ -coalescent spatial partant d'un nombre infini d'individus, ce résultat peut être généralisé à tout type de condition initiale donnée. En effet de deux choses l'une : soit il existe un sommet avec un nombre infini d'individus à l'instant initial et nous pouvons directement appliquer notre théorème (par le Lemme 4.2) ; soit il existe une infinité de sommets avec au moins un individu, et dans ce cas la preuve du Théorème 1.3 peut être facilement adaptée.

7 Le cas du Beta-coalescent

Nous allons maintenant parler brièvement d'un autre type de Λ -coalescents, qui est presque aussi bien compris que le coalescent de Kingman : c'est la famille des Beta-coalescents, appelés ainsi de part leurs liens avec la loi Beta. Dans ce cas comme dans celui de Kingman, nous disposons d'une estimation de la vitesse de divergence du nombre d'individus présents.

En réalité, nous allons nous intéresser à un cas un peu plus général. Nous supposons simplement disposer d'un Λ -coalescent dont la mesure caractéristique Λ est supposée posséder une densité suffisamment régulière au voisinage de 0, au sens où $\Lambda(dx) = g(x)dx$, pour laquelle il existe $\alpha \in (1, 2)$ et $B > 0$ tel que $g(x) \sim_0 Bx^{1-\alpha}$. Cela inclut en particulier le cas où Λ est la loi Beta($\alpha, 2 - \alpha$), cas dans lequel le processus est appelé Beta-coalescent.

Commençons par estimer le taux total de coalescence au rang n .

Lemme 7.1. λ_n est croissant en n , de plus il existe $c > 0$ tel que $\lambda_n \sim cn^\alpha$.

Preuve. La croissance en n est une conséquence directe de la consistance du processus. Pour ce qui est de l'équivalent, on a juste à vérifier que :

$$\begin{aligned} \lambda_n &= \sum_{k=2}^n \binom{n}{k} \int_0^1 x^{k-2} (1-x)^{n-k} d\Lambda(x) \\ &= \int_0^1 \frac{1}{x^2} (1 - (1-x)^n - nx(1-x)^{n-1}) g(x) dx. \end{aligned}$$

Par conséquent, nous obtenons bien l'équivalent souhaité pour λ_n . □

7.1 Borne inférieure pour le processus

Nous aurons besoin par la suite des paramètres suivantes :

$$\beta = \frac{\alpha - 1}{2}; \tau_n = an^{-\beta} \text{ pour un certain } a \in [a_0, a_1]; \gamma = \min \left\{ 1 - \frac{\alpha}{2}, \frac{\beta}{2}, \frac{1}{8} \right\},$$

et observons qu'en particulier, γ vérifie :

$$\alpha + \gamma - 2 \leq \frac{\alpha}{2} - 1 \leq -\gamma < 0.$$

Nous notons X^n le processus du nombre de blocs d'un Λ, n -coalescent spatial, $N^n = X^n(1)$ et Y^n le coalescent non-spatial couplé dans le Lemme 4.3, nous allons commencer par étudier la quantité

$$K_n = \int_0^{\tau_n} Y_s^n ds,$$

de la même façon que dans le Lemme 6.1, mais nous aurons ici besoin d'un résultat plus précis. Cette étape est grossièrement la même que le Lemme 5.2 dans le cas du coalescent de Kingman, on étudie l'ordre de grandeur du nombre d'individus qui quittent u avant l'instant τ_n .

Lemme 7.2. *Il existe c, C dépendant uniquement de Λ tels que :*

$$\mathbb{P}(K_n \geq n^{2-\alpha+\gamma}) \leq Cn^{-\gamma},$$

$$\mathbb{P}(K_n \leq cn^\gamma) \leq Cn^{-\gamma}.$$

Preuve. La borne supérieure découle de manière assez immédiate de l'inégalité de Markov. En effet il suffit de constater que si $Y_t^n = k$, alors Y^n reste dans k pendant un temps exponentiel de paramètre λ_k puis décroît strictement. Par conséquent nous avons :

$$\mathbb{E}(K_n) \leq \tau_n + \sum_{k=2}^n \frac{k}{\lambda_k} \leq Cn^{2-\alpha},$$

d'où nous tirons notre estimation.

Pour la borne inférieure, nous montrerons qu'avec grande probabilité, les premiers $M = n^{1-\alpha+\gamma}$ sauts arrivent tous avant l'instant τ_n , et que Y^n reste plus grand que $\frac{n}{2}$ après tous ces sauts. Ne considérer que ces premiers sauts dans K^n nous donnera la borne inférieure que nous cherchons.

Commençons par estimer la quantité B_m , le nombre d'individus perdus lors d'un événement de coalescence sur une population de m individus. On peut supposer que $g \leq Cx^{1-\alpha}$, par conséquent nous avons :

$$\binom{n}{k} \lambda_{n,k} \leq C \binom{n}{k} \int_0^1 x^{k-1-\alpha} (1-x)^{n-k} dx \leq C \frac{n! \Gamma(k-\alpha)}{k! \Gamma(n-\alpha+1)} \leq Cn^\alpha k^{-\alpha-1}.$$

Or cette borne nous permet d'obtenir :

$$\mathbb{P}(B_m > k) \leq \frac{1}{\lambda_m} \sum_{i=k+1}^m \binom{m}{i} \lambda_{m,i} \leq Ck^{-\alpha}. \quad (3.8)$$

En particulier, $\mathbb{E}(B_m) < C$, où C ne dépend que de Λ . Dès lors le nombre total d'individus ayant coalescé dans les M premiers sauts a une espérance d'au plus CM . On note t_k l'instant du $k^{\text{ième}}$ saut, nous avons alors :

$$\mathbb{P}\left(Y_{t_M}^n \leq \frac{n}{2}\right) \leq \frac{CM}{n - \frac{n}{2}} \leq Cn^{\alpha-2+\gamma} \leq Cn^{-\gamma}.$$

De plus, sur l'évènement de forte probabilité $\{Y_{t_M}^n > \frac{n}{2}\}$, le taux de chacun des M premiers sauts est d'au moins $\lambda_{\frac{n}{2}}$. En appliquant une nouvelle fois l'inégalité de Markov, nous obtenons :

$$\mathbb{P}\left(t_M > \tau_n, Y_{t_M}^n \geq \frac{n}{2}\right) \leq \frac{M}{\lambda_{\frac{n}{2}}\tau_n} \leq Cn^{-1+\beta+\gamma} \leq Cn^{-\gamma}.$$

Nous obtenons alors que la probabilité qu'on ait fait M sauts avant τ_n et que au cours de ces M sauts on soit resté au-dessus de $\frac{n}{2}$ est d'au moins $1 - Cn^{-\gamma}$. Notons que sur cet évènement de forte probabilité, nous avons :

$$K_n = \int_0^{\tau_n} Y_s^n ds \geq \int_0^{t_M} \frac{n}{2} ds \geq \frac{nt_M}{2}.$$

Il suffit donc pour conclure de démontrer que $\mathbb{P}(t_M \leq cn^{\gamma-1}) \leq Cn^{-\gamma}$. Mais chaque saut arrivant à taux au plus λ_n , t_M est minoré par la somme de M variables aléatoires i.i.d. exponentielles de paramètre λ_n . Or l'espérance de cette somme est $\frac{M}{\lambda_n} \sim cn^{\gamma-1}$, il nous suffit alors d'utiliser le Lemme 4.1 pour conclure :

$$\mathbb{P}\left(\sum_{i=1}^M E_i \leq \frac{c}{2}n^{\gamma-1}\right) < \exp\left(-\frac{1}{16}n^{\alpha-1+\gamma}\right) \leq Cn^{-\gamma}.$$

□

Nous allons maintenant trouver une borne inférieure au nombre d'individus quittant l'origine. Ce lemme est une version plus précise du Lemme 6.2, pour lequel on prouve simplement la divergence vers l'infini de ce nombre d'immigrants.

Lemme 7.3. *Il existe $C > 0$ tel que :*

$$\mathbb{P}(Z_{\tau_n}^n < n^\gamma) \leq Cn^{-\gamma}$$

Preuve. Soit $a \in (0, 1)$, on définit un temps aléatoire $T_a = \inf\{t > 0 : Z_t^n \geq aY_t^n\}$. Définissons alors les évènements suivants :

$$A = \{Z_{\tau_n}^n < n^\gamma\}, A_1 = \{Z_{\tau_n \wedge T_a}^n < n^\gamma\} \text{ et } A_2 = A \cap \{\tau_n > T^a\}.$$

On a en particulier $A \subset A_1 \cup A_2$. Il nous suffit donc de majorer les probabilités des évènements A_1 et A_2 .

Nous commençons par borner A_1 . Rappelons les notations du Lemme 4.3 : \mathcal{G} est la tribu engendrée par Y^n et (\mathcal{F}_t^*) la tribu engendrée par (X_t^n) et \mathcal{G} .

T_a est un temps d'arrêt relativement à la filtration (\mathcal{F}_t^*) . Comme Y^n est décroissant de limite 1, et que Z_t^n devient supérieur à 1 presque sûrement, ce temps d'arrêt est fini presque sûrement.

Considérons la (F_t^*) -martingale S_t^n stoppée à l'instant T_a , nous utilisons l'inégalité maximale de Doob pour obtenir l'inégalité suivante :

$$\begin{aligned} & \mathbb{P} \left(\sup_{s < T_a \wedge \tau_n} |S_s^n| \geq \delta \int_0^{\tau_n} Y_s^n ds \middle| \mathcal{G} \right) \\ & \leq \frac{4\rho \mathbb{E} \left(\int_0^{T_a \wedge \tau_n} M_s^n ds \middle| \mathcal{G} \right)}{\delta^2 \left(\int_0^{\tau_n} Y_s^n ds \right)^2} \wedge 1 \\ & \leq \frac{4\rho}{\delta^2 \int_0^{\tau_n} Y_s^n ds} \wedge 1, \end{aligned}$$

où la dernière inégalité s'obtient grâce à l'inégalité $M_s^n \leq Y_s^n$, qui implique que

$$\mathbb{E} \left(\int_0^{T_a \wedge \tau_n} M_s^n ds \middle| \mathcal{G} \right) \leq \int_0^{\tau_n} Y_s^n ds.$$

Posons maintenant A_s l'évènement :

$$A_s = \left\{ 1 - \delta - a < \frac{Z_s^n}{\rho \int_0^s Y_u^n du} < 1 + \delta \right\}.$$

En utilisant que jusqu'à l'instant T_a , $M_s^n \geq (1 - a)Y_s^n$ et $Y_s^n - M_s^n \leq Z_s^n \leq aY_s^n$, l'inégalité précédente nous permet d'obtenir :

$$\mathbb{P}(A_{s \wedge T_a}^c \middle| \mathcal{G}) \leq \frac{4\rho}{\delta^2 \int_0^{\tau_n} Y_s^n ds}.$$

En effet, on a :

$$\mathbb{P} \left(Z_s^n \geq (1 + \delta) \int_0^s Y_s^n ds \middle| \mathcal{G} \right) \leq \mathbb{P} \left(S_s^n \geq \delta \int_0^s Y_s^n ds \right)$$

et de même :

$$\mathbb{P} \left(Z_s^n \leq (1 - \delta - a) \int_0^s Y_s^n ds \middle| \mathcal{G} \right) \leq \mathbb{P} \left(|S_{s \wedge T_a}^n| \geq \delta \int_0^s Y_s^n ds \right).$$

Dès lors, en prenant l'espérance, et choisissant a et δ tel que $1 - a - \delta > \frac{1}{2}$, nous obtenons grâce au Lemme 7.2 l'inégalité suivante :

$$\mathbb{P}(A_1) \leq \frac{C}{n^\gamma} + n^{\alpha-2-\gamma} \leq Cn^{-\gamma}.$$

Intéressons-nous maintenant à l'évènement A_2 , qui est clairement inclus dans $\{Y_{\tau_n}^n \leq an^\gamma\}$, il nous suffit donc de majorer $\mathbb{P}(Y_{\tau_n}^n \leq n^\gamma)$. Pour cela, remarquons que grâce à l'équation 3.8, la probabilité qu'un épisode de coalescence concerne plus de k individus est dominée par $Ck^{-\alpha}$.

Par conséquent, la probabilité qu'à un instant donné Y^n tombe dans l'intervalle $[n^\gamma + 1, 2n^\gamma]$ est d'au moins $1 - Cn^{-\alpha\gamma}$. De plus Y^n restera alors un temps exponentiel de paramètre inférieur à λ_{2n^γ} une fois dans cet intervalle.

Nous pouvons donc majorer la probabilité que $Y_{\tau_n}^n$ soit plus petit que n^γ par la probabilité qu'une variable aléatoire exponentielle de paramètre λ_{2n^γ} soit plus petite que τ_n ajoutée à celle qu'un saut de taille au moins n^γ se produise au bon moment, d'où :

$$\mathbb{P}(Y_{\tau_n}^n \leq n^\gamma) \leq Cn^{-\alpha\gamma} + (1 - e^{-\lambda_{2n^\gamma}\tau_n}) \leq Cn^{-\gamma}.$$

Ce qui conclut la preuve de ce lemme. □

Nous pouvons maintenant obtenir une borne inférieure sur le comportement du coalescent. Commençons par un lemme semblable à la borne inférieure du Lemme 5.3.

Lemme 7.4. *Soit $1 < a_0 < a_1$, et considérons le Λ -coalescent spatial partant de n individus au point u et aucun ailleurs. On pose $\tau_n = an^{-\beta}$ pour $a \in [a_0, a_1]$, et définissons l'évènement :*

$$A = \left\{ \forall v \sim u \text{ ou } v = u, X_{\tau_n}(v) \geq \frac{n^\gamma}{4d_u} \right\}.$$

Il existe c et C dépendant uniquement de a_0 , a_1 et d_u tels que $\mathbb{P}(A^c) \leq Cn^{-c}$.

Preuve. La partie sur la minoration $X_{\tau_n}^n(u)$ a déjà été vue précédemment, car $\mathbb{P}(Y_{\tau_n}^n \leq n^\gamma) \leq Cn^{-\gamma}$.

Pour le cas $v \sim u$, le Lemme 7.3 montre qu'avec forte probabilité (au moins $1 - Cn^{-\gamma}$, au moins n^γ individus immigreront. Il est alors évident qu'avec forte probabilité, chaque sommet v reçoit une part égale de ces individus.

Il nous reste à estimer la probabilité qu'un évènement de coalescence se produise avant l'instant τ_n pour les n^γ premiers individus à sauter. En effet un évènement de coalescence arrive à taux minoré par $\lambda_{n^\gamma} \geq cn^{\alpha\gamma}$. Dès lors nous obtenons :

$$P(X_{\tau_n}^n(v) \leq \frac{n^\gamma}{4d_u}) \leq Cn^{-\gamma} + \mathbb{P}\left(\frac{e}{\lambda_{n^\gamma}} \leq \tau_n\right) \leq Cn^{-\gamma},$$

où e est une variable aléatoire exponentielle de paramètre 1.

Finalement, nous devons estimer la probabilité que l'un des n^γ premiers individus à sauter fasse plus de deux sauts dans les τ_n premiers instants. Or le taux de saut de ces n^γ individus est ρn^γ , qui est inférieur au taux de coalescence de ces n^γ individus. Donc la probabilité d'un évènement de migration qui empêcherait d'obtenir la borne inférieure requise est encore une fois majorée par $Cn^{-\gamma}$. □

Preuve du Théorème 1.2 (borne inférieure). On pose $f^{(k)}(n) = \underbrace{f \circ \dots \circ f}_{k \text{ termes}}$, où $f(n) = \frac{n^\gamma}{4d_u}$. On définit alors la séquence de temps (t_k) par $t_0 = 0$ et $t_{k+1} = t_k + f^{(k+1)}(n)^{-\beta}$. Nous allons itérer la minoration précédente une quantité $m_n = \frac{\log \log n}{-2 \log \gamma}$, on peut vérifier que pour tout $k \leq m_n$, $f^{(k)}(n) \geq c \exp(\sqrt{\log n})$.

On pose alors B_k l'évènement pour lequel il y a au moins $f^{(k)}(n)$ individus en chaque site de la boule de centre u et de rayon k et aucune en dehors. On a $\mathbb{P}(B_1^c) \leq Cn^{-c}$, et pour tout $k \geq 1$, si on suppose B_k , en utilisant le Lemme 4.2 on peut considérer, pour minorer le nombre d'individus en $v \in B(u, k)$ ou le nombre d'individus en $w \sim v \in B(u, k)$, l'évolution d'une portion $f^{(k)}(n)$ de l'ensemble des individus présents en v à l'instant t_k pendant un temps $\tau_{f^{(k)}(n)}$. Le Lemme 7.2 nous donne alors une borne sur la probabilité, sachant que B_k est réalisé, que B_{k+1} le soit.

Nous obtenons en résumé :

$$\mathbb{P}\left(\left(\bigcap_{k=1}^{m_n} B_k\right)^c\right) \leq \sum_{k=1}^{m_n} C \text{Vol } B(u, k) f^{(k)}(n)^{-c} \leq C m_n \text{Vol } B(u, m_n) f^{(m_n)}(n)^{-c}.$$

Cette borne est encore majorée, en utilisant $\text{Vol } B(u, m_n) \leq D^{m_n}$ par :

$$C m_n D^{m_n} \exp(-c\sqrt{\log n}) \xrightarrow{n \rightarrow +\infty} 0.$$

□

7.2 Borne supérieure

Ayant la borne inférieure de notre théorème, nous nous tournons maintenant vers la borne supérieure. Pour cela nous aurons besoin d'un équivalent du Lemme 5.6 dans le cas des Λ -coalescents considérés.

Lemme 7.5. *Considérons un Λ -coalescent sur un graphe G partant d'une situation initiale quelconque, A un ensemble de sommets de G , et soit Q_t le nombre d'individus qui restent dans A tout au long de l'intervalle $[0, t]$. Il existe c et C dépendant uniquement de Λ tels que :*

$$\mathbb{P}(Q_t > C|A|t^{-\frac{1}{\alpha-1}}) \leq \exp(-c|A|).$$

Preuve. Si des individus quittent A , nous les ignorerons par la suite, donc tout se passera comme si elles avaient été tuées. Mais même en ignorant la migration, ce qui rend Q_t petit est l'action d'épisodes de coalescence. Le taux de coalescence en un site v contenant q individus est de l'ordre de Cq^α , et à chaque épisode de coalescence, au moins un individu disparaît, ce qui nous permet de minorer le taux total de décroissance de Q_t par :

$$\sum_{v \in A} c(X_t(v))^\alpha \geq c|A|^{1-\alpha} Q_t^\alpha$$

en utilisant l'inégalité de Jensen (on a $\alpha > 1$).

Nous pouvons donc dominer Q_t par une chaîne de mort pour laquelle le taux de décroissance de i à $i-1$ est donné par $c|A|^{1-\alpha} i^\alpha$. Il est alors facile de conclure en utilisant le Lemme 4.1. Si on note e_k une suite de variables aléatoires exponentielles indépendantes de paramètre $\mu_k = c|A|^{1-\alpha} k^\alpha$ et $S_K = \sum_{k=K+1}^{+\infty} e_k$, on a :

$$\mathbb{E}(S_K) = \sum_{k=K+1}^{+\infty} \mu_k^{-1} \sim c_1 |A|^{\alpha-1} K^{1-\alpha},$$

$$\text{Var}(S_K) = \sum_{k=K+1}^{+\infty} \mu_k^{-2} \sim c_2 |A|^{2\alpha-2} K^{1-2\alpha}.$$

En particulier, $\frac{\text{Var}(S_K)}{\mathbb{E}(S_K)\mu_K}$ est asymptotiquement constant, donc en utilisant l'inégalité (3.2) pour $\varepsilon > 0$ assez petit. On en tire :

$$\mathbb{P}(S_K > 2\mathbb{E}(S_K)) \leq \exp\left(-c \frac{\mathbb{E}(S_K)^2}{\text{Var}(S_K)}\right) \leq e^{-cK}$$

Prenons maintenant K tel que $\mathbb{E}(S_K) \leq \frac{t}{2}$, nous pouvons en conclure :

$$\mathbb{P}(Q_t > K) < e^{-cK}.$$

Nous remarquons finalement que pour C assez grand, $K = Ct^{-\frac{1}{\alpha-1}}|A|$ fonctionne. \square

Nous allons maintenant majorer X_{τ_n} , il ne nous restera qu'à étendre cette approximation encore une fois pour obtenir la majoration espérée.

Lemme 7.6. *Soit $a_0, a_1 > 0$, considérons X^n le coalescent spatial partant de n individus au sommet u et aucun ailleurs. Soit $\tau_n = an^{-\beta}$ pour $a \in [a_0, a_1]$, et notons l'évènement $A = \{X_{\tau_n}^n \leq C_1 Q\}$, où on a posé :*

$$Q(v) = \begin{cases} n^{\frac{3}{4}} & v = u \\ n^{2-\alpha+\gamma} & d(u, v) \leq r = \left\lceil \frac{4}{\alpha-1} \right\rceil \\ 0 & \text{sinon.} \end{cases}$$

Il existe deux constantes C et C_1 dépendant uniquement de a_0, a_1 et Λ telles que

$$\mathbb{P}(A^c) \leq Cn^{-\gamma}.$$

Preuve. On sait qu'avec une probabilité assez grande au plus $n^{2-\alpha+\gamma}$ individus quittent u avant l'instant τ_n , la borne est donc obtenue pour $0 < d(u, v) \leq r$.

Quelques-unes des n^γ individus peuvent coalescer, mais nous prétendons qu'il est impossible que l'un d'entre eux fasse plus de r sauts avant l'instant τ_n . En effet la probabilité qu'un individu fasse r sauts avant l'instant τ_n est majorée par $C(\rho n^{-\beta})^r$, et au plus n^γ individus quittent l'origine.

Dès lors si r est tel que $1 - \beta r < -\gamma$, la probabilité qu'un individu parmi n^γ fasse plus de r sauts est majorée par $1 - (1 - \rho^r n^{-r\beta})^{n^\gamma} \leq Cn^{-\gamma}$, ce qui traite le cas $d(u, v) > r$.

Pour le cas $v = u$, il suffit de prendre S un ensemble de sommets contenant u de taille $C_2 \log n$. Pour C_2 assez grand, grâce au Lemme 7.5, on sait qu'avec probabilité au plus $Cn^{-\gamma}$, on a $Q_{\tau_n} \leq C|A|n^{\frac{1}{2}} < n^{\frac{3}{4}}$. De plus on sait que $X_{\tau_n}(u) < Q_{\tau_n} + Z_{\tau_n}$, donc grâce au Lemme 4.4, on conclut sans difficultés que $X_{\tau_n}(u)$ est également petit. \square

Preuve du Théorème 1.2 (borne supérieure). Notons que pour tout α , nous avons $\gamma \leq \frac{\beta}{2} < \frac{1}{2}$. Posons C_1 la constante du Lemme 7.6, et $C_2 = C_1 \text{Vol } B(u, r)$.

On pose $f(n) = C_2 n^{\frac{3}{4}}$, et comme précédemment, $f^{(k)}(n) = f \circ \dots \circ f(n)$. On pose de plus $t_0 = 0$ et $t_{k+1} = t_k + f^{(k+1)}(n)^{-\beta}$.

Soit A_i l'évènement tel que pour tout $v \in B(u, ir)$, $X_{t_i}(v) \leq f^{(i)}(n)$ et aucun individu n'est hors de la boule. Posons $m_n = \min\{k \in \mathbb{N} : f^{(k+1)}(n) \leq \log n\}$, le point maximal auquel on itérera le Lemme 7.6. Il est clair que $\log n < f^{(m_n)}(n) < \log n^2$ pour n assez grand. On peut aussi vérifier que $m_n \sim c \log \log n$ et que $t_{m_n} = o(1)$.

Nous obtenons alors comme précédemment :

$$\mathbb{P} \left(\left(\bigcap_{k=1}^{m_n} A_i \right)^c \right)_{n \rightarrow +\infty} \rightarrow 0.$$

Par conséquent, avec forte probabilité, à l'instant t_{m_n} , le nombre d'individus restants est dominé par $C f^{(m_n)}(n) \text{Vol } B(u, m_n r)$, et tous les individus sont dans la boule de centre u et de rayon $m_n r$. Pour étendre notre approximation jusqu'à un instant t constant, nous posons $A = B(u, M \log \log n)$, pour M précisé ultérieurement.

Pour qu'un individu quitte l'ensemble A , il faut qu'il survive jusqu'à l'instant τ_k et fasse au moins $M \log \log n - m_n r$ sauts avant l'instant t . Dès lors, l'espérance du nombre d'individus quittant A est dominé par :

$$\begin{aligned} & C f^{(m_n)}(n) \text{Vol } B(u, m_n r) \exp(-c(M \log \log n - m_n r)) \\ & \leq C (\log n)^2 \text{Vol } B(u, C \log \log n) \exp(-c \log \log n (cM - C)). \end{aligned}$$

En prenant M assez grand, cette quantité tend vers 0 quand $n \rightarrow +\infty$. Dès lors tous les individus restent dans A jusqu'à l'instant t , utiliser le Lemme 7.5 nous permet de conclure, on a :

$$P(N_t^n \leq C \text{Vol } B(u, C \log \log n)) \leq o(1) + \exp(-c \log \log n) \xrightarrow[n \rightarrow +\infty]{} 0.$$

□

8 Conclusion

On a vu tout au long de ce mémoire comment il était possible d'estimer le comportement d'un Λ -coalescent sur un graphe infini général. On remarque pour commencer que pour un temps assez petit, le Λ -coalescent spatial est équivalent au Λ -coalescent non-spatial auquel on ajoute un processus de migration sur les voisins du site considéré. Ensuite, en itérant l'approximation précédente sur chacun de ces sites occupés un certain nombre de fois, nous pouvons estimer quel sera l'ordre de grandeur de la diffusion de la population sur le graphe, c'est-à-dire à trouver l'ordre de grandeur du rayon m_n de la boule qui sera presque sûrement couverte par les n premiers individus.

La minoration du nombre total d'individus présents dans le processus s'en déduit assez facilement, puisqu'on peut alors réaliser une minoration stochastique du nombre d'individus présents à temps constant par une somme de variables aléatoires de Bernoulli i.i.d. de paramètre $e^{-\rho t}$.

Pour la majoration, nous avons besoin d'une meilleure estimation, ainsi, en partant de la même situation que dans le cas de la minoration, c'est à dire des individus raisonnablement dispersés sur une boule, il nous faut d'abord affirmer qu'avec une bonne probabilité aucun des individus restants ne fera un nombre de pas grand devant l'ordre de grandeur de m_n . Il faut enfin trouver un moyen de majorer le nombre d'individus présents à l'instant t dans un Λ -coalescent spatial à valeurs dans un graphe de taille importante, mais finie, par une constante multipliée par la taille du graphe. Ceci est possible dans le cas du coalescent de Kingman spatial ou du Beta-coalescent spatial, mais nous n'avons obtenu, en toute généralité qu'une minoration nous permettant de prouver la divergence de tout Λ -coalescent spatial sur un graphe de taille infinie.

Il est également possible d'étudier le comportement sur des temps longs de ces Λ -coalescent spatiaux, pour obtenir la manière de réaliser une limite d'échelle et espérer obtenir un processus non-trivial. Il est intéressant de noter que si le détail de la loi du Λ -coalescent a de fortes incidences sur le comportement des individus au voisinage des points de forte intensité, cette incidence est bien plus faible lorsque les individus sont initialement dispersés sur le graphe, la probabilité d'une coagulation de plus de deux individus en un même instant diminuant alors beaucoup.

9 Annexe 1 : Preuve du théorème de Schweinsberg

Nous allons prouver ici la condition nécessaire et suffisante à la descente de l'infini d'un processus Λ -coalescent non-spatial de Schweinsberg. Pour cela nous utiliserons une démonstration tirée de [4].

Rappelons que si Λ est une mesure finie sur $[0, 1]$, on note

$$\lambda_{n,k} = \int_0^1 x^{k-2}(1-x)^{n-k} \Lambda(dx) \text{ et } \lambda_n = \sum_{k=2}^n \binom{n}{k} \lambda_{n,k},$$

respectivement le taux de coagulation de k individus choisis parmi n et le taux auquel un événement de coalescence apparaît parmi n individus. Rappelons alors ce théorème que nous souhaitons démontrer.

Théorème 9.1 (Schweinsberg). *Soit Λ une mesure finie sur $[0, 1]$, on pose*

$$\phi(n) = \sum_{k=2}^n k \binom{n}{k} \lambda_{n,k},$$

On a alors :

$$\text{Le } \Lambda\text{-coalescent descend de l'infini} \iff \sum_{n=2}^{+\infty} \frac{1}{\phi(n)} < +\infty.$$

Commençons par prouver que ce théorème donne une condition suffisante à la descente de l'infini du Λ -coalescent. On pose (Π_t) un processus Λ -coalescent.

Lemme 9.1. Soit ξ le temps d'atteinte de $\{\mathbb{N}, \emptyset, \emptyset, \dots\}$ du Λ -coalescent Π , défini par :

$$\xi = \inf\{t > 0 \mid \Pi_t = \{\mathbb{N}, \emptyset, \emptyset, \dots\}\},$$

on a alors :

$$\mathbb{E}(\xi) \leq \sum_{n=2}^{+\infty} \frac{1}{\phi(n)},$$

donc en particulier, si la série est finie, le coalescent descend de l'infini presque sûrement.

Preuve. Nous pouvons supposer que $\sum_{n=2}^{+\infty} \frac{1}{\phi(n)} < +\infty$, sinon il n'y a rien à démontrer.

Intéressons-nous au processus du nombre de blocs de $(\Pi_t|_{[n]}) : D_t^n$. Nous pouvons sans difficulté calculer les taux de transitions de ce processus de Markov. Le taux de transition de p à $p - k + 1$ est $\binom{p}{k} \lambda_{p,k}$. Le générateur infinitésimal de D_t^n est donc :

$$G^n(f) : l \mapsto \sum_{k=2}^l \binom{l}{k} \lambda_{l,k} (f(l - k + 1) - f(l)).$$

Remarquons alors que $\phi(n) = \sum_{k=2}^n k \binom{n}{k} \lambda_{n,k}$ est une fonction croissante de n . Nous avons supposé que la série $\sum \frac{1}{\phi(n)}$ converge, nous pouvons donc poser :

$$f(l) = \sum_{l+1}^{+\infty} \frac{1}{\phi(n)}.$$

Comme ϕ est croissante, nous avons $f(l - k + 1) - f(l) \geq \frac{k}{\phi(l)}$, donc on a :

$$G^n(f)(l) \geq \sum_{k=2}^l \binom{l}{k} \lambda_{l,k} \frac{k}{\phi(l)} \geq 1.$$

Dès lors le processus $f(D_t^n) - \int_0^t G^n f(D_s^n) ds$ est une martingale, de plus $\xi_n = \inf\{t > 0 \mid D_t^n = 1\}$ est un temps d'arrêt fini presque sûrement. Soit $k \geq 1$, on utilise le Théorème d'arrêt au temps d'arrêt borné $\xi_n \wedge k$, on a :

$$\mathbb{E}(f(D_{\xi_n \wedge k}^n)) - \mathbb{E}\left(\int_0^{\xi_n \wedge k} G^n(f)(D_s^n) ds\right) = f(n).$$

En utilisant la minoration de $G^n(f)$ par 1, il vient :

$$\mathbb{E}(\xi_n \wedge k) \leq \mathbb{E}(f(D_{\xi_n \wedge k}^n) - f(n)).$$

On fait alors tendre k vers $+\infty$, par convergence monotone (f et D^n sont toutes deux des fonctions décroissantes), on obtient :

$$\mathbb{E}(\xi_n) \leq f(1) - f(n).$$

Nous pouvons finalement passer à la limite quand $n \rightarrow +\infty$, par hypothèse $f(n) \xrightarrow{n \rightarrow +\infty} 0$, et $(\xi_n)_{n \in \mathbb{N}}$ croît vers $\xi_\infty = \inf\{t > 0 | D_t = 1\} = \xi$. On en tire bien :

$$\mathbb{E}(\xi) \leq \sum_{n=2}^{+\infty} \frac{1}{\phi(n)}.$$

□

Pour prouver que cette condition est également nécessaire, nous allons maintenant supposer que la série des $\frac{1}{\phi(n)}$ diverge, et que le coalescent descend de l'infini. Nous allons essayer ensuite de trouver une contradiction.

Commençons par un résultat standard de théorie des grandes déviations :

Lemme 9.2. *Soit (X_n) une suite i.i.d de variables aléatoires de paramètre $x < \frac{1}{4}$, on pose $S_n^{(x)} = \sum_{k=1}^n X_k$, pour tout $n_0 \in \mathbb{N}$ on a :*

$$\mathbb{P}\left(\exists n \geq n_0 : S_n^{(x)} \geq \frac{n}{2}\right) \leq \frac{e^{-n_0 f(x)}}{1 - e^{-f(x)}}$$

avec f une fonction équivalente en 0 à $-\frac{1}{2} \log x$.

Preuve. Par inégalité de Markov, nous avons :

$$\mathbb{P}\left(S_n^{(x)} \geq \frac{n}{2}\right) \leq e^{-\frac{nt}{2}} \mathbb{E}(e^{t S_n^{(x)}}) \leq \exp\left(-n \left[\frac{t}{2} - \log(xe^t + 1 - x)\right]\right).$$

On applique alors cette inégalité en $t = \log \frac{1}{x}$, nous obtenons $\mathbb{P}(S_n^{(x)} \geq \frac{n}{2}) \leq e^{-nf(x)}$, où :

$$f(x) = \frac{1}{2} \log\left(\frac{1}{x}\right) - \log(2 - x).$$

La fonction f est strictement positive sur $]0, \frac{1}{4}[$, donc en utilisant la convergence d'une série géométrique, nous obtenons bien :

$$\mathbb{P}\left(\exists n \geq n_0 : S_n^{(x)} \geq \frac{n}{2}\right) \leq \frac{e^{-n_0 f(x)}}{1 - e^{-f(x)}}.$$

De plus f a le bon équivalent en 0. □

Nous nous intéressons maintenant au premier instant de coagulation impliquant plus de la moitié de la population.

Lemme 9.3. *Supposons que le Λ -coalescent descende de l'infini. Alors avec probabilité 1, le temps d'arrêt*

$$\tau = \inf\left\{t > 0 | D_t < \frac{D_t}{2}\right\} > 0. \text{ p.s.}$$

De plus si nous posons $\tau_n = \inf\{t > 0 | D_t^n < \frac{D_t^n}{2}\}$, alors $\tau_n \xrightarrow{n \rightarrow +\infty} \tau$ p.s.

Preuve. Il est clair que le taux de coagulation relatif au coalescent de Kingman ne joue aucun rôle dans ce théorème, donc nous pouvons supposer $\Lambda(\{0\}) = 0$. Soit N une mesure de Poisson sur $\mathbb{R}^+ \times [0, 1]$ d'intensité $dt \otimes \frac{1}{x^2} \Lambda(dx)$. En utilisant les notations du lemme précédent, nous allons montrer que :

$$N \left(\left\{ (t, x) : t \leq 1, \exists n \geq 4 | S_n^{(x)} \geq \frac{n}{2} \right\} \right) < +\infty,$$

en calculant son espérance, ce qui nous permettra de conclure en utilisant la construction d'un Λ -coalescent à l'aide d'une mesure de Poisson.

Nous avons supposé que le processus descend presque sûrement de l'infini, donc pour tout $\varepsilon > 0$, $D_\varepsilon < +\infty$. Nous déduisons de ce résultat qu'il n'y a qu'un nombre fini de sauts sur $[0, 1]$ impliquant plus de la moitié des particules. De plus par continuité à droite, $D_{0+} = D_0$, par conséquent 0 n'est pas un instant de saut et on a $\tau > 0$. Enfin, nous savons que $D_{\tau-} < +\infty$, dès lors nous avons pour tout $n \geq D_{\tau-}$, $\tau_n = \tau$. Nous en déduisons que $\tau_n \xrightarrow[n \rightarrow +\infty]{} \tau$ p.s. Nous obtenons également :

$$\mathbb{E} \left(N \left(\left\{ (t, x) : t \leq 1, \exists n \geq 4 | S_n^{(x)} \geq \frac{n}{2} \right\} \right) \right) = \int_0^1 \frac{1}{x^2} \mathbb{P} \left(\exists n \geq 4 : S_n^{(x)} \geq \frac{n}{2} \right) \Lambda(dx).$$

Nous pouvons maintenant utiliser le Lemme 9.2 pour obtenir :

$$\int_0^1 \frac{1}{x^2} \mathbb{P} \left(\exists n \geq 4 : S_n^{(x)} \geq \frac{n}{2} \right) \Lambda(dx) \leq \int_0^{\frac{1}{4}} \frac{e^{-4f(x)}}{(x^2(1 - e^{-f(x)}))} \Lambda(dx) + \int_{\frac{1}{4}}^1 \frac{1}{x^2} \Lambda(dx).$$

Ces deux termes sont bornés car l'intégrande est majorée par 8 dans le premier terme, et par 16 dans le second. Nous obtenons donc que le nombre d'événements de coagulation impliquant plus de la moitié des individus avant l'instant 1 est fini, ce qui conclut la preuve. \square

Nous supposons que le coalescent descend de l'infini et que $\sum \frac{1}{\phi(n)}$ diverge. Nous allons ensuite définir une surmartingale et utiliser un théorème d'arrêt. Posons f la fonction décroissante définie par :

$$f(n) = \exp \left(- \sum_{k=2}^{n+1} \frac{1}{\phi(k)} \right).$$

Lemme 9.4. *Il existe une constante C tel que pour tout $n \geq 1$, $(e^{-Ct} f(D_t^n))_{t \leq \tau_n}$ est une surmartingale positive.*

Preuve. Rappelons que le générateur infinitésimal de D^n est :

$$G^n(f) : l \mapsto \sum_{k=2}^l \binom{l}{k} \lambda_{l,k} (f(l-k+1) - f(l)).$$

Il suffit d'étudier le processus avant l'instant τ_n pour savoir si celui-ci descend de l'infini, puisque ceci arrive dans les premiers instants, et $\tau > 0$. Jusqu'à cet instant, D_t^n est égal au

processus \tilde{D}^n dans lequel les sauts de plus de la moitié de la taille courante sont ignorés, dont le générateur infinitésimal est :

$$A^n(f) : l \mapsto \sum_{k=2}^{\lfloor \frac{l}{2} \rfloor} \binom{l}{k} \lambda_{l,k} (f(l-k+1) - f(l)).$$

Nous utiliserons désormais le générateur A^n . Remarquons que $\phi(n) = \sum_{k=2}^n k \binom{n}{k} \lambda_{n,k} = \int_0^1 \frac{1}{x^2} ((1-x)^n - 1 + nx) \Lambda(dx)$. Or il existe $c > 0$ tel que pour tout n ,

$$c(e^{-nx} - 1 + nx) \leq (1-x)^n - 1 + nx \leq e^{-nx} - 1 + nx.$$

Posons alors $\psi(n) = \int_0^1 \frac{1}{x^2} (e^{-nx} - 1 + nx) \Lambda(dx)$, on a alors $c\psi \leq \phi \leq \psi$. Nous avons également :

$$\frac{\psi(n)}{n} = \int_0^1 (1 - e^{-nx}) \int_x^1 \frac{1}{u^2} \Lambda(du) dx \xrightarrow{n \rightarrow +\infty} \int_0^1 \frac{1}{x} \Lambda(dx) > 0.$$

$h(q) = \frac{\psi(q)}{q}$ est une fonction concave, donc $h(\frac{q}{2}) \geq \frac{h(q)}{2}$, d'où $\psi(\frac{q}{2}) \geq \frac{\psi(q)}{4}$. On a pour l assez grand :

$$A^n(f)(l) = \sum_{k=2}^{\frac{l}{2}} \binom{l}{k} \lambda_{l,k} f(l) \left(\exp\left(\sum_{i=l-k+2}^l \frac{1}{\phi(i)} \right) - 1 \right) \leq C f(l) \sum_{k=2}^{\frac{l}{2}} \binom{l}{k} \lambda_{l,k} \frac{k}{\phi(\frac{l}{2})},$$

où on a utilisé que pour x assez petit, $e^x - 1 \leq Cx$.

On en tire $A^n(f)(l) \leq C f(l) \frac{\phi(l)}{\phi(\frac{l}{2})} \leq C f(l)$, dès lors $(e^{-Ct} f(\tilde{D}_t^n))$ est une surmartingale. \square

Lemme 9.5. *Si $\sum \frac{1}{\phi(n)}$ diverge, alors Π ne descend pas de l'infini.*

Preuve. On suppose que le coalescent descend de l'infini, on pose $T_j^{(n)} = \inf\{t > 0 \mid \tilde{D}_t^n \leq j\}$. On applique le théorème d'arrêt à la surmartingale précédente pour le temps d'arrêt $T_j^{(n)}$, on en tire :

$$f(n) \geq \mathbb{E}(e^{-CT_j^{(n)}} f(\tilde{D}_{T_j^{(n)}}^n) \mathbf{1}_{\{T_j^{(n)} < \tau_n\}}).$$

On passe à la limite quand $n \rightarrow +\infty$, on a $f(n) \xrightarrow{n \rightarrow +\infty} 0$ par hypothèse. De plus quand $t < \tau_n$, on a $\tilde{D}_t^n = D_t^n$. Ensuite, $T_j^{(n)} \xrightarrow{n \rightarrow +\infty} T_j = \inf\{t > 0 \mid D_t \leq j\} > 0$, donc $f(D_{T_j}) > 0$, car $D_{T_j} < +\infty$. Enfin $\tau_n \xrightarrow{n \rightarrow +\infty} \tau > 0$ et $T_j^j \xrightarrow{j \rightarrow +\infty} 0$, donc pour j assez grand, avec grande probabilité, $T_j < \tau$.

Ceci entraîne $\mathbb{E}(e^{-CT_j} \mathbf{1}_{\{T_j < \tau\}}) = 0$, donc $T_j = +\infty$ sur l'évènement $T_j < \tau$, ce qui ne se peut puisque le processus descend de l'infini presque sûrement et que pour j assez grand, l'évènement précédent est non-négligeable. On a bien une contradiction, donc Π ne descend pas de l'infini. \square

Bibliographie

- [1] O. Angel, N. Berestycki, and V. Limic. Global divergence of spatial coalescents. *Preprint*, 2009.
- [2] J. Berestycki, N. Berestycki, and V. Limic. The Λ -coalescent speed of coming down from infinity. *Ann. Probab.*, 2009. To appear.
- [3] J. Bertoin and J.-F. Le Gall. Stochastic flows associated to coalescent processes. *Prob. Theory Related Fields*, pages 261–288, 2003.
- [4] C. Foucart. Distinguished exchangeable coalescents and generalized fleming-viot processes with immigration. 2010. To appear.
- [5] J.F.C. Kingman. The coalescent. *Stochastic Process. App.*, pages 235–248, 1982.
- [6] J.F.C. Kingman. On the genealogy of large populations. *J. App. Probab.*, pages 29–43, 1982.
- [7] V. Limic and A. Sturm. The spatial Λ -coalescent. *Electron J. Probab.*, pages 363–393, 2006.
- [8] J. Pitman. Coalescents with multiple collisions. *Ann Probab.*, pages 1870–1902, 1999.
- [9] S. Sagitov. The general coalescent with asynchronous mergers of ancestor lines. *J. Appl. Prob.*, pages 1116–1125, 1999.
- [10] J. Schweinsberg. A necessary and sufficient condition for the Λ coalescent to come down from infinity. *Electr. Comm. Probab.*, pages 1–11, 2000.

Chapitre 4

Vulgarisation pour le site Culturemath : Le coalescent de Kingman

Bastien MALLEIN,

29 novembre 2010

Résumé

Lorsque l'on s'intéresse au patrimoine génétique d'une population, il est souvent intéressant d'étudier également les liens de parenté entre les individus de cette population. Le processus coalescent de Kingman est un modèle probabiliste qui associe à un petit nombre d'individus pris au hasard dans une population leur arbre généalogique. En utilisant ce modèle, on peut tester des hypothèses sur la dissémination de mutations. Nous étudierons ici quelques propriétés du processus coalescent de Kingman.

Je tiens à remercier Grégory Ginot et Éric Vandendriessche qui m'ont permis de publier cet article de vulgarisation mathématique.

Sommaire

1	Introduction du coalescent de Kingman	65
1.1	Le modèle de Wright-Fisher	65
1.2	Étude fine de la généalogie	66
1.3	Le coalescent de Kingman	70
2	Construction du coalescent de Kingman	70
2.1	Quelques notions sur les espaces de partitions	71
2.2	Le coalescent de Kingman, construction détaillée	72
3	Descente de l'infini du coalescent de Kingman	73
3.1	Borne sur l'âge de l'ancêtre commun le plus récent	73
3.2	La vitesse de descente de l'infini du coalescent de Kingman	75
4	Taille des groupes de cousins	77
4.1	Une autre construction du coalescent de Kingman	78
4.2	Taille des familles d'individus	79
5	Mutation neutre et classification phylogénique	81
6	Conclusion	83
7	Preuve du Théorème 2.1	83

1 Introduction du coalescent de Kingman

Une part entière de la théorie des probabilités est consacrée à l'étude des modèles de populations utilisés en biologie. A l'aide de ces modèles, on essaie d'estimer des quantités dont l'observation directe est difficile, de prévoir l'avenir, ou de tester une hypothèse théorique. Mais un modèle mathématique est construit sous de nombreuses hypothèses simplificatrices, il faut donc toujours garder du recul sur les résultats obtenus. Il y a une grande différence entre la réalité et la construction mathématique censée s'en approcher. Dans le modèle de nombreux effets sont négligés ce qui rend les calculs plus simples, mais les résultats faux. Cela peut toutefois suffire pour obtenir des ordres de grandeurs. De nombreux modèles différents sont bien entendu nécessaires pour reproduire toute la diversité du monde vivant. Nous nous intéresserons uniquement à l'un d'entre eux, un modèle aléatoire assez simple : le modèle de Wright-Fisher.

Le but sera ici d'étudier l'arbre généalogique d'une population qui évolue en suivant le modèle de Wright-Fisher. Cet arbre généalogique est bien entendu aléatoire également. Nous étudierons alors comment il évolue lorsque la population considérée devient très grande. Nous en tirerons un modèle limite, le coalescent de Kingman, qui se dégage naturellement de cette étude. Ce modèle décrit l'arbre généalogique d'un nombre fini de lignées choisies au hasard dans une population ambiante infinie. On choisit n individus dans une population de N individus et on étudie l'arbre généalogique de leurs lignées, et on fait tendre N vers $+\infty$ en gardant n fixé. Le modèle obtenu est suffisamment simple pour autoriser des calculs nous permettant, par exemple, de déterminer une borne sur l'âge de l'ancêtre commun de ces lignées.

L'étude de la généalogie des populations est nécessaire en biologie lorsqu'on s'intéresse au patrimoine génétique d'une population. On peut par exemple se demander quand un gène est-il apparu au cours du temps, ou si un certain type de mutations a un impact sur les individus ou si la mutation considérée du gène est neutre. J.F.C. Kingman apporta pour la première fois une réponse à cette question en 1981 dans ses articles *On the genealogy of large populations* [5] et *The coalescent* [4]. Dans son étude, Kingman s'est intéressé à la généalogie de populations évoluant selon le modèle de Wright-Fisher. Nous introduirons tout d'abord ces deux modèles avant de calculer plusieurs quantités relatives au coalescent de Kingman : l'âge moyen du dernier ancêtre commun à une population, le nombre d'individus partageant la même mutation génétique, le nombre de descendants après un temps t de l'ancêtre à la descendance la plus abondante, ou de celui ayant eu la descendance la plus faible, etc. Les résultats obtenus ont une application en biologie, par exemple grâce à la formule d'Ewens on peut tester l'hypothèse qu'une mutation est neutre ou non.

1.1 Le modèle de Wright-Fisher

Dans tout ce qui suit, nous considérerons que chaque individu de la population que nous étudions descend d'un seul parent (haploïde), et possède le même patrimoine génétique que ce parent. Ainsi nous pouvons facilement définir des lignées d'ancêtres de nos individus. L'arbre généalogique est alors bien plus simple à décrire. Remarquons pour commencer que lorsque deux lignées ancestrales se rejoignent, elles restent communes pour toutes les générations précédentes. Cette remarque évidente est la raison pour laquelle le processus que nous étudierons sera appelé processus coalescent.

FIGURE 4.1 – Évolution au cours du temps d’une population selon le modèle de Wright-Fisher, avec l’indication des descendants de la première génération.

Dans le cas où un individu descend de deux parents, comme l’espèce humaine, nous pouvons par exemple appliquer ce modèle en nous intéressant uniquement aux relations mères-filles, afin de trouver la mère de toutes les femmes, l’Ève mitochondriale. De même on peut essayer de retrouver le père de tous les hommes, l’Adam-chromosome Y, dont le chromosome Y est l’ancêtre de celui de tous les hommes.

Introduisons maintenant le modèle de Wright-Fisher. Dans ce modèle, la taille de la population est supposée constante au cours du temps, égale à un entier N . La population évolue de génération en génération, à chaque étape tous les individus meurent, et donnent naissance à la génération suivante. De plus nous supposons qu’aucun individu particulier dans la population n’est favorisé par un quelconque avantage. Dans ce modèle, tout se passe comme si chaque individu choisissait parmi la population de la génération précédente son père uniformément au hasard, indépendamment du choix des autres membres de la population et des liens de parenté de ses ancêtres ou descendants.

Regardé dans le sens conventionnel d’écoulement du temps, un individu a un nombre de descendants qui suit une loi dite binomiale de paramètres N et $\frac{1}{N}$. Autrement dit chaque enfant a une chance sur N d’être le fils d’un parent donné. Mais on n’a alors pas d’indépendance du nombre d’enfants de chaque parent à une génération donnée. La connaissance du nombre d’enfants d’un individu particulier modifie la loi du nombre d’enfant de chacun des autres de sa génération. Par exemple savoir qu’un individu donné a eu N enfants nous permet de déduire qu’aucun des autres n’en a eu, puisque la population est supposée constante au cours du temps. A contrario, le fait de savoir quel parent a choisi un individu n’influence aucunement le choix des autres : lorsqu’on remonte le temps, le processus s’exprime en termes d’événements indépendants. Il est par conséquent plus simple de réaliser des calculs en remontant dans le temps. Ce modèle est donc parfait pour étudier la généalogie d’une population.

1.2 Étude fine de la généalogie

Nous allons maintenant étudier l’arbre généalogique d’un échantillon d’une population se reproduisant selon le modèle de Wright-Fisher. Rappelons pour commencer une remarque simple : lorsque deux lignées généalogiques admettent un ancêtre commun à une génération donnée, les lignées restent confondues pour toutes les générations précédentes.

Considérons pour commencer deux individus dans une population ambiante de N personnes. La probabilité pour que ces deux individus soient des frères est égale à $\frac{1}{N}$. En effet, il y a N^2 choix possibles de parents pour chacun des deux individus, et pour N de ces choix, le parent choisi est le même. Dès lors, la probabilité pour que deux individus ne soient pas frères est de $1 - \frac{1}{N}$. On s’intéresse alors à la probabilité que leur grand-père soit différent. Il faut que leurs pères soient différents, et que les pères de ceux-ci soient différents. Or connaître les enfants d’une génération n’influence pas les liens de parenté de la génération précédente. Par conséquent la probabilité que les grand-pères des deux enfants soient différents est égale à $(1 - \frac{1}{N})^2$. Ce résultat

FIGURE 4.2 – Les lignées du modèle de Wright-Fisher, remontées au cours du temps.

peut se généraliser sans difficulté. On appelle la lignée d'un individu le chemin reliant un individu à ses ancêtres successifs. La probabilité pour qu'après k générations, deux lignées choisies soient toujours distinctes est égale à $(1 - \frac{1}{N})^k$.

Nous pouvons de la même façon calculer quelle est la probabilité que l lignées restent toutes distinctes pendant k générations. Commençons par nous intéresser à une seule génération. Nous considérons alors les liaisons parent-enfant permettant de laisser les l lignées distinctes. Le premier individu choisit son parent, puis le second parmi les $N - 1$ parents restants, le troisième parmi les $N - 2$ restants, etc. En définitive nous obtenons que la probabilité que l lignées choisies restent distinctes après une génération est égale à :

$$\frac{N(N-1)\cdots(N-l+1)}{N^l} = \left(1 - \frac{1}{N}\right)\cdots\left(1 - \frac{l-1}{N}\right).$$

De même après k générations, la probabilité que nos l lignées soient restées distinctes est alors égale à $\left[\left(1 - \frac{1}{N}\right)\cdots\left(1 - \frac{l-1}{N}\right)\right]^k$.

Nous pouvons maintenant introduire le coalescent de Kingman. C'est un processus construit comme une limite des arbres généalogiques d'échantillons de taille finie dans une population ambiante de plus en plus grande, tendant vers $+\infty$. Bien sûr, après une génération il n'y a plus aucune chance pour que deux individus pris au hasard soient frères, car on a $\frac{1}{N} \xrightarrow{N \rightarrow +\infty} 0$. Pour éviter cette situation triviale, nous allons prendre une limite d'échelle : en même temps que la population grandira, nous accélérerons le passage du temps, ce qui nous permettra d'obtenir une limite non-triviale. Ainsi, plutôt que de regarder le processus à N individus à la génération k fixée, nous regarderons pour tout $t \in \mathbb{R}^+$ le processus à N individus à la génération $[Nt]$ ($[\cdot]$ désigne la partie entière, le plus grand entier inférieur ou égal à ce nombre).

Ce processus n'a pas forcément d'interprétation biologique simple. Mais lorsque nous nous intéresserons à une population de taille N donnée, on pourra approximer son comportement par celui d'un coalescent de Kingman, et exporter les résultats de la manière suivante : si l'on trouve par exemple un ancêtre commun à l'instant t , dans le modèle « réel », cela correspond à un ancêtre commun à la génération Nt . Nous allons alors avoir besoin d'un résultat classique sur les limites.

Propriété 1.1. Soit $x \in \mathbb{R}$, on a :

$$\lim_{n \rightarrow +\infty} \left(1 + \frac{x}{n}\right)^n = e^x.$$

Preuve. Soit $x \in \mathbb{R}$ et $n \in \mathbb{N}$, on pose $u_n = \left(1 + \frac{x}{n}\right)^n$ et $v_n = \ln u_n$. On a :

$$v_n = n \ln \left(1 + \frac{x}{n}\right) = x \frac{\ln\left(1 + \frac{x}{n}\right) - \ln(1)}{\frac{x}{n}}.$$

Or $x \mapsto \ln x$ est dérivable en 1 de dérivée égale à 1, par conséquent on a :

$$\lim_{n \rightarrow +\infty} v_n = x,$$

d'où, par continuité de $x \mapsto e^x$ on obtient :

$$\lim_{n \rightarrow +\infty} u_n = e^x.$$

□

En utilisant cette propriété nous remarquons alors que la probabilité que, dans le processus limite, deux individus aient des lignées distinctes jusqu'au temps t est égale à $\lim_{N \rightarrow +\infty} (1 - \frac{1}{N})^{\lfloor Nt \rfloor} = e^{-t}$. En particulier, la probabilité que deux individus n'aient aucun ancêtre commun est égale à 0, car dans ce cas les lignées sont distinctes pour tout $t \in \mathbb{R}^+$, donc en particulier quand $t \rightarrow +\infty$. En d'autre terme, si on note T l'âge (aléatoire) de l'ancêtre commun de ces deux individus choisis dans la population, on a :

$$\mathbb{P}(T > t) = e^{-t}.$$

Définition 1.1. Une variable aléatoire e est dite de loi exponentielle de paramètre $\lambda > 0$ si c'est une variable aléatoire à valeurs dans \mathbb{R}^+ vérifiant pour tout $t \geq 0$:

$$\mathbb{P}(e > t) = e^{-\lambda t}.$$

Autrement dit T est une variable aléatoire exponentielle de paramètre 1. Rappelons quelques propriétés très classiques des variables aléatoires exponentielles :

Propriété 1.2. Si e est une variable aléatoire exponentielle de paramètre 1, alors $\frac{1}{\lambda}e$ est une variable aléatoire exponentielle de paramètre λ .

Si e est une variable aléatoire exponentielle de paramètre λ , elle vérifie la propriété d'absence de mémoire, autrement dit :

$$\mathbb{P}(e > t + s | e > t) = \mathbb{P}(e > s).$$

Nous aurons également besoin tout au long de notre étude du lemme suivant.

Lemme 1.1 (Lemme des réveils). Soit T_1, \dots, T_k des variables aléatoires indépendantes de loi exponentielles de paramètres t_1, \dots, t_k . On note T le minimum de T_1, \dots, T_k et α l'indice de la variable aléatoire réalisant ce minimum, c'est-à-dire que l'on a $T = T_\alpha$.

T est une variable aléatoire exponentielle de paramètre $t_1 + \dots + t_k$, et α est une variable aléatoire indépendante de T , telle que $\mathbb{P}(\alpha = i) = \frac{t_i}{t_1 + \dots + t_k}$.

Preuve. On calcule simplement la probabilité que $\min T_1, \dots, T_k$ soit plus grande que t et que $\alpha = i$:

$$\mathbb{P}(\alpha = i, T > t) = \mathbb{P}(T_i > t, \forall j \neq i T_j > T_i).$$

Cette dernière probabilité est une fonction de T_i , que l'on peut écrire de la manière suivante :

$$\begin{aligned} \mathbb{P}(\alpha = i, T > t) &= \int_t^{+\infty} t_i e^{-t_i s} \prod_{j \neq i} \mathbb{P}(T_j > s) ds \\ &= \int_t^{+\infty} t_i e^{-t_1 s - \dots - t_k s} ds \\ &= e^{-(t_1 + \dots + t_k)t} \frac{t_i}{t_1 + \dots + t_k}. \end{aligned}$$

FIGURE 4.3 – Le coalescent de Kingman comme limite d'échelle d'arbre généalogiques.

Dès lors en prenant en particulier $t = 0$ ou en sommant sur tous les indices allant de 1 à k , on voit que T et α ont bien les lois que nous avons annoncé, de plus ces deux variables aléatoires sont bien indépendantes. Ce lemme est parfois appelé Lemme des réveils, car si on considère ces variables aléatoires T_1, \dots, T_k comme k réveils distincts, nous voyons que le fait de savoir que le premier des réveils ait sonné à un instant donné ne nous permet pas de dire de manière plus précise lequel des réveils a sonné. Par exemple, même si un réveil a tendance à sonner beaucoup plus tôt que les autres, le fait d'avoir attendu longtemps n'augmente ni ne diminue la probabilité que ce soit celui-ci qui ait sonné. \square

Ce lemme sera très important dans toute la suite, car nous travaillerons beaucoup avec des variables aléatoires exponentielles, et nous utiliserons sans arrêt ce résultat, ainsi que l'absence de mémoire des variables aléatoires de loi exponentielle.

Choisissons maintenant k individus, on note T_k le premier instant (aléatoire) auquel deux au moins des lignées généalogiques que nous étudions se rejoignent. Nous obtenons alors de la même manière que précédemment :

$$\mathbb{P}(T > t) = \exp[-(1 + \dots + (k-1))t] = \exp\left(-\frac{k(k-1)}{2}t\right),$$

car on vérifie immédiatement que $1 + \dots + (n-1) = \frac{n(n-1)}{2}$ par récurrence.

T_k est alors une variable aléatoire exponentielle de paramètre $\frac{k(k-1)}{2}$. Notons que lorsqu'on regarde k lignées généalogiques distinctes, on peut trouver $\frac{k(k-1)}{2}$ paires de lignées généalogiques distinctes. Chacune d'entre elles peut se regrouper à des instants égaux à des variables aléatoires exponentielles de paramètre 1. Nous pouvons déterminer le premier instant auquel l'un de ces regroupements deux à deux se produit. Pour cela nous utilisons le Lemme 1.1.

Nous observons donc que si les instants auxquels les lignées se regroupent deux à deux sont indépendants, le premier instant auquel les deux premières de ces k lignées se regroupent est aussi une variable aléatoire exponentielle de paramètre $\frac{k(k-1)}{2}$. On retrouve la quantité précédente, avec de plus l'information suivante : quand un changement arrive dans k lignées, c'est forcément deux, et seulement deux de ces lignées choisies au hasard qui se regroupent.

Ce résultat a donc encouragé Kingman à introduire son modèle de coagulation dans lequel les seuls événements de regroupement de lignées sont des regroupement deux à deux de lignées, indépendantes les unes des autres. La figure 4.3 illustre ainsi comment on sélectionne quelques individus dans une population de plus en plus grande. Le troisième graphe montre le processus obtenu lorsqu'on fait tendre la taille de cette population vers $+\infty$, en raccourcissant les longueurs des branches rouges de façon proportionnelle. L'arbre ainsi obtenu est appelé le processus coalescent de Kingman.

FIGURE 4.4 – Une représentation du coalescent de Kingman à 7 individus.

FIGURE 4.5 – Quelques partitions associées au coalescent de Kingman.

1.3 Le coalescent de Kingman

Notre processus d'intérêt, le coalescent de Kingman, est l'arbre généalogique d'une population de n individus. En particulier, il décrit qui est plus proche parent avec qui, et quel âge a l'ancêtre commun de chaque paire d'individus. Dans le modèle de Kingman, si il reste k ancêtres à l'instant t , alors au bout d'un temps exponentiel de paramètre $\frac{k(k-1)}{2}$, deux lignées choisies au hasard parmi toutes se regroupent en une seule. Ce processus s'appelle processus coalescent, car lorsque deux lignées se rassemblent, elles restent ensemble pour toujours, comme collées.

Dans ce modèle, chaque paire de lignées se regroupe après un temps exponentiel de paramètre 1, indépendamment des autres individus. La figure 4.4 est une représentation du coalescent de Kingman, l'axe indique le temps qui remonte dans le passé, et la longueur avant que deux lignes se rejoignent représente le degré de parenté entre les individus situés à l'extrémité. Afin de définir plus précisément les paramètres du modèle et de pouvoir réaliser des calculs, nous allons avoir besoin de quelques notions sur les espaces de partitions, avec lesquelles on code cet arbre généalogique.

2 Construction du coalescent de Kingman

Pour construire le coalescent de Kingman, nous avons besoin d'une manière de coder un arbre généalogique. Il faut en particulier écrire lesquelles de ces lignées se rejoignent à un instant donné, et à quel instant ces lignées se rejoignent. Pour cela nous allons, à chaque instant t associer les sous-ensembles de la population initiale constitués des descendants des lignées existantes à la génération t . Plus précisément, l'ensemble $\{\{1, 2\}, \{3\}, \{4\}\}$ décrit une situation où les individus 1 et 2 descendent de la même lignée lorsqu'on remonte à l'instant t , mais d'une lignée distincte de celles des individus 3 et 4. Si la partition suivante est $\{\{1, 2, 4\}, \{3\}\}$, alors cela signifie que la lignée commune à 1 et 2 et celle de 4 se rejoignent avant qu'on ait trouvé un lien de parenté avec 3. De plus, l'instant auquel deux lignées se rejoignent, qui est aussi l'âge de l'ancêtre commun le plus récent de ces deux lignées, est donné par l'instant auquel on a un changement dans la répartition des ensembles. En quelque sorte, on représente un ancêtre à la génération t par l'ensemble de ses descendants. La figure 4.5 correspond aux relations entre l'arbre généalogique associé au coalescent de Kingman et les partitions présentes dans le coalescent.

Nous allons commencer par préciser quelques propriétés relatives aux partitions. Nous les utiliserons par la suite pour construire un coalescent de Kingman partant d'une infinité d'individus à l'instant initial.

FIGURE 4.6 – Le sous-arbre des quatre premiers individus de notre population.

2.1 Quelques notions sur les espaces de partitions

Nous souhaitons construire l'arbre généalogique de n individus. Pour cela, à chaque instant t , nous associons la partition de cette population constitué des descendants des lignées existantes t générations dans le passé. Une partition des entiers entre 1 et n est un certain nombre de sous-ensembles deux à deux distincts (aucun entier n'est dans deux ensembles à la fois, aucun individu ne descend de deux lignées à la fois) et dont l'union donne tous les entiers (tout le monde possède un ancêtre). L'ensemble de ces partitions est noté \mathcal{P}_n .

Soit $\pi \in \mathcal{P}_n$ une partition des entiers entre 1 et n , on note $\#\pi$ le nombre d'ensembles présents dans cette partition, correspondant au nombre de lignées restantes. Nous avons par exemple, dans la partition suivante $\#\{\{1, 3, 4, 6\}, \{2, 5, 7\}\} = 2$. Notons que la partition correspondant à la population initiale est la partition en singletons : une partition dans laquelle chaque individu correspond à une lignée. Nous allons maintenant préciser comment passer d'une partition à une partition plus ancienne. Lorsqu'un événement de coagulation arrive, deux lignées choisies uniformément au hasard se rejoignent. Par conséquent, on passe d'une partition à une partition plus ancienne en réalisant la réunion de deux des ensembles présents pris au hasard.

Nous allons maintenant détailler une autre propriété des arbres généalogiques, qui se traduit sur les partitions qui sont associées. On peut passer de l'arbre généalogique à n individus à celui à $n - 1$ individus simplement en oubliant le $n^{\text{ième}}$ individu. On dira que deux arbres de n et $k \leq n$ individus sont compatibles si l'arbre généalogique des k premiers individus du grand arbre est égal au petit arbre. Cela se traduit de la manière suivante sur les partitions : à tout instant t , la partition des entiers entre 1 et k obtenue en prenant les intersections des ensembles de la partition associée au grand arbre avec l'ensemble $\{1, \dots, k\}$ est égale à la partition associée au petit arbre.

Cette remarque est importante, car elle permet non seulement de descendre (à partir d'un arbre généalogique, on peut retrouver l'arbre généalogique d'une partie de la population concernée) mais également de remonter vers un arbre de taille infinie. En effet, si on est capable de définir une suite d'arbres généalogiques de plus en plus grands et compatibles les uns avec les autres (comme le sont par exemple les arbres de la figure 4.6, le petit arbre en gras et le plus grand contenant également les pointillés) ; alors on est capable de définir les liens de parenté d'une famille infinie d'individus. Plus précisément, deux arbres sont dits compatibles lorsque la propriété suivante est vérifiée : le sous-arbre du plus grand arbre constitué des individus présents dans l'arbre le plus petit (celui des individus 1 à 4 par exemple) correspond exactement avec le plus petit arbre. En fait, si on cherche un renseignement dans n'importe lequel des deux arbres, soit celui-ci n'existe pas dans l'un des deux arbres, soit il est exactement le même pour les deux. Dans ce cas, si on peut fournir toute l'information permettant de réaliser l'arbre de n'importe quelle famille finie, on peut alors également fournir l'arbre généalogique infini, en ajoutant une nouvelle branche à chaque nouvel individu découvert. C'est ce que nous allons formaliser par la suite.

Pour réaliser cette remontée, on raisonne de la manière suivante : partons d'une suite $(\Pi^n)_{n \in \mathbb{N}}$

d'arbres généalogiques compatibles tel que Π^n parte de n individus à l'instant initial. A chaque instant t , on regarde la suite de partitions associée $(\Pi_t^n)_{n \in \mathbb{N}}$. On construit alors la partition de \mathbb{N} qui soit compatible avec toutes les partitions finies de la manière suivante. Deux entiers p et q sont situés dans le même ensemble si et seulement si pour tout entier m plus grand que p et q , ces deux entiers sont situés dans le même ensemble de Π_t^m . Cette partition infinie permettra de construire un arbre généalogique sur une population infinie. En effet nous obtenons pour chaque instant t une partition infinie Π_t , et le passage d'une partition à une partition plus ancienne se fait forcément par la réunion d'un certain nombre d'ensembles. Nous avons donc bien un arbre généalogique.

Nous allons maintenant construire le processus coalescent de Kingman à proprement parlé.

2.2 Le coalescent de Kingman, construction détaillée

Commençons donc la construction de l'arbre généalogique d'une population de n individus. Cet arbre généalogique sera appelé le *n -coalescent de Kingman*. À chaque instant t , on associe la partition Π_t de n correspondant aux lignées restantes lorsqu'on remonte de t dans le temps.

On a clairement $\Pi_0 = \{\{1\}, \dots, \{n\}\}$, la partition associée à l'instant 0 est constitué des n ensembles disjoints. Nous attendons alors un temps exponentiel de paramètre $\frac{n(n-1)}{2}$, puis deux individus choisis uniformément au hasard sont regroupés dans un même ensemble. De manière plus générale, si à un instant donné il reste k ensembles (on a $\#\Pi_t = k$), on attend un temps exponentiel de paramètre $\frac{k(k-1)}{2}$, puis on choisit deux des ensembles uniformément au hasard que l'on fusionne ensemble pour obtenir la partition suivante. Cela correspond à deux des lignées de la population qui, choisis uniformément au hasard se rejoignent.

Ce processus est bien défini pour tout entier n , notons $(\Pi_t^{(n)})_{t \in \mathbb{R}^+}$ ce n -coalescent de Kingman. Nous nous intéressons maintenant à $\Pi_k^{(n)}$ le processus restreint à k individus. Il est naturel d'espérer obtenir un k -coalescent de Kingman, puisque c'est un arbre généalogique, celui de k individus pris dans une population infinie suivant l'évolution du modèle de Wright-Fisher. Par calcul sur les variables aléatoires exponentielles, on peut prouver le théorème suivant.

Théorème 2.1. *Pour tout $n \geq 2$ et $k \leq n$, le processus $\Pi_k^{(n)}$ est un k -coalescent de Kingman, un coalescent partant de k individus à l'instant initial.*

La preuve de ce théorème est donnée en annexe. Il est donc possible de construire pour tout $n \in \mathbb{N}$ un n -coalescent de Kingman de telle manière que pour tout instant t , la suite des arbres généalogiques n -coalescents de Kingman $(\Pi_t^n)_{n \in \mathbb{N}}$ soit compatible. Il existe alors pour tout instant t une partition infinie dont la restriction à n vaut $\Pi_t^{(n)}$. On note cette partition Π_t . Le processus $(\Pi_t)_{t \in \mathbb{R}^+}$ est appelé processus coalescent de Kingman. Il est caractérisé par la propriété suivante :

$$\forall n \in \mathbb{N}, \Pi_t|_n \text{ est un } n\text{-coalescent de Kingman.}$$

Il existe une unique loi pour ce processus, c'est pourquoi on l'appelle le coalescent de Kingman infini. Il est difficile de travailler directement avec ce processus, c'est pourquoi dans la plupart des preuves que nous verrons, nous nous intéresserons à des coalescents de Kingman de taille n finie, puis nous passerons à la limite quand $n \rightarrow +\infty$ pour obtenir des résultats sur ce coalescent infini, ou bien nous raisonnerons par récurrence.

Le coalescent de Kingman est intéressant à étudier, car il « chapeaute » tous les arbres généalogiques de taille finie à la fois. Un résultat sur ce processus a donc une certaine importance, car on en déduit simplement un bon résultat sur tout les arbres de Kingman de taille finie.

3 Descente de l'infini du coalescent de Kingman

Considérons Π un coalescent de Kingman partant d'un nombre infini d'individus. Ce processus descend de l'infini, c'est-à-dire que pour tout instant positif, le nombre d'ancêtres de la population est fini. C'est un résultat intéressant, car cela signifie que quelle que soit la taille de la population de laquelle nous partons, nous savons non seulement qu'ils ont un ancêtre commun, mais aussi que l'âge du dernier ancêtre commun est borné par une valeur indépendante de la taille de la population. En d'autres termes, que l'on prenne deux ou deux mille individus, l'âge du dernier ancêtre commun est inférieur à la même borne finie.

Nous nous intéresserons également à la vitesse de descente de l'infini de ce processus, car il se trouve que pour des instants assez petits, le nombre d'individus est assez bien approximé par une fonction déterministe.

3.1 Borne sur l'âge de l'ancêtre commun le plus récent

Afin de calculer l'âge de l'ancêtre commun le plus récent d'un coalescent de Kingman, la seule information dont nous ayons besoin est le nombre d'ancêtres de la population à chaque instant t . En effet, nous savons que nous avons atteint l'ancêtre commun le plus récent lorsqu'on atteint l'instant où il ne reste plus qu'une seule lignée dans le processus.

Notons D_t^n le nombre de lignées distinctes à l'instant t dans le n -coalescent de Kingman. Ce processus $(D_t^n)_{t \in \mathbb{R}^+}$ peut être décrit de la manière suivante. Nous avons $D_0^n = n$. Nous attendons alors un temps exponentiel de paramètre $\frac{n(n-1)}{2}$, à cet instant nous avons un évènement de coagulation, donc D_t^n diminue de 1. Puis nous attendons un temps exponentiel de paramètre $\frac{(n-1)(n-2)}{2}$, et on diminue encore de 1, etc.

Posons e_1, e_2, \dots des variables aléatoires indépendantes exponentielles de paramètre 1. On pose, pour tout $k \leq n$:

$$S_{n,k} = \sum_{i=k+1}^n \frac{2}{i(i-1)} e_i,$$

l'instant auquel le processus D_t^n arrive dans l'état k . Grâce à la définition de D_t^n , $S_{n,k}$, le premier instant auquel D_t^n vaut k , est bien la somme de variables aléatoires de paramètres $\frac{n(n-1)}{2}, \dots, \frac{k(k-1)}{2}$ indépendantes, ce qui est la variable que nous avons décrite plus haut. Or $D_t^n \leq k$ indique que le premier instant auquel D_t^n atteint k est avant l'instant t , d'où nous tirons l'égalité suivante :

$$\mathbb{P}(D_t^n \leq k) = \mathbb{P}(S_{n,k} \leq t).$$

Nous allons maintenant nous intéresser au coalescent de Kingman infini. Le nombre de lignées de la population infinie présentes à la génération t est clairement égal à la limite du nombre de lignées présentes dans les coalescents restreints à n , pour $n \rightarrow +\infty$. On pose $D_t = \lim_{n \rightarrow +\infty} D_t^n$

le nombre total d'individus à l'instant t dans le coalescent de Kingman. Posons également

$$S_k = \sum_{i=k+1}^{+\infty} \frac{2}{i(i-1)} e_i,$$

nous pouvons donc maintenant passer à la limite quand $n \rightarrow +\infty$ d'un côté et de l'autre de l'inégalité. Des résultats de la théorie de l'intégration nous donnent :

$$\mathbb{P}(D_t \leq k) = \lim_{n \rightarrow +\infty} \mathbb{P}(D_t^n \leq k) = \lim_{n \rightarrow +\infty} \mathbb{P}(S_{n,k} \leq t) = \mathbb{P}(S_k \leq t).$$

Remarquons pour commencer que S_k est une variable aléatoire d'espérance finie, donc qui est finie également. En effet :

$$\mathbb{E}(S_k) = \sum_{i=k+1}^{+\infty} \frac{2}{i(i-1)} \mathbb{E}(e_i) = \sum_{i=k+1}^{+\infty} \frac{2}{i-1} - \frac{2}{i} = \frac{2}{k}.$$

De plus $\lim_{k \rightarrow +\infty} \mathbb{E}(S_k) = 0$, donc on sait que la probabilité pour S_k d'être plus grande que t devient très faible pour k très grand. Ceci nous indique que la variable aléatoire D_t est, pour tout $t > 0$ une variable aléatoire finie presque sûrement :

$$\mathbb{P}(D_t < +\infty) = \lim_{k \rightarrow +\infty} \mathbb{P}(D_t \leq k) = 1 - \lim_{k \rightarrow +\infty} \mathbb{P}(S_k > t) = 1.$$

C'est un résultat très important appelé la propriété de descente de l'infini du coalescent de Kingman. Il signifie qu'en tout instant strictement positif, il ne reste qu'un nombre fini de lignées dans le coalescent. Et ce bien que l'on soit parti d'un nombre infini de lignées à l'instant 0. Autrement dit un nombre fini de lignées engendrent une population infinie. De plus, l'échantillon de population que nous regardons est grand, moins le nombre de lignées qui lui est associé croit, et on finit par atteindre un nombre maximal. Dans la prochaine section, nous verrons comment se produit ce phénomène de goulot d'étranglement, qui fait passer le nombre d'individus étudiés de l'infini à une quantité finie. Nous mettrons ainsi en évidence une « vitesse de descente de l'infini » du coalescent de Kingman, la vitesse à laquelle le nombre de lignées diminue lorsqu'on s'écarte de l'instant d'origine. On retiendra donc :

$$\forall t > 0, D_t < +\infty \text{ presque sûrement.}$$

Le deuxième résultat intéressant est que nous obtenons la valeur de l'âge du plus récent ancêtre commun du coalescent de Kingman. En effet, cet âge est le temps d'atteinte de 1 pour le processus du nombre de blocs D_t . Posons T l'âge de cet ancêtre commun, nous avons :

$$\mathbb{P}(T < t) = \mathbb{P}(D_t = 1) = \mathbb{P}(S_1 < t).$$

Nous obtenons par conséquent le résultat suivant :

$$T = \sum_{k=2}^{+\infty} \frac{2}{k(k-1)} e_k.$$

Nous pouvons calculer l'espérance de l'âge du plus récent ancêtre commun, nous avons $\mathbb{E}(T) = 2$. L'espérance étant fini, l'âge aléatoire est lui-même fini presque sûrement. Mais rappelons que l'arbre généalogique de tout échantillon de taille fini peut être considéré comme un sous-arbre du coalescent de Kingman, par conséquent l'âge de leur plus récent ancêtre commun est majoré par T , qui est une constante indépendante de la taille de l'échantillon. Autrement dit, il existe un âge au-delà duquel quelque soit l'échantillon choisi dans la population, on ne peut trouver deux lignées distinctes. Nous interprétons ceci comme l'âge de l'ancêtre commun de toute la population considérée.

Nous allons maintenant tenter d'appliquer les résultats obtenus sur notre modèle à une espèce réelle. Nous avons obtenu que l'âge de l'ancêtre commun d'une population infinie est en moyenne égal à 2. Cependant, dans notre modèle, nous avons accéléré le passage du temps d'un facteur proportionnel à la taille de la population considérée. Rappelons également qu'une unité du temps (discret) est égal à une génération. Par conséquent, si l'approximation de l'évolution de la vraie population par le coalescent de Kingman est assez précise, nous obtenons que l'âge du récent ancêtre commun d'un groupe de N individus est en moyenne égal à environ $2 \times N$ générations dans le passé de ce groupe.

Lorsqu'on applique ce modèle à la population humaine, on obtient un nombre bien trop grand. Il y a six milliards d'êtres humains, et une génération humaine représente environ 25 ans. Mais nous avons également supposé dans notre modèle que la population ambiante restait constante au cours du temps, or la population humaine a cru à vitesse exponentielle durant plusieurs dizaines de siècles. Si on remonte d'une dizaine de milliers d'années, il y avait alors quelques milliers d'êtres humains seulement, et la taille de cette population évolue peu lorsqu'on remonte encore dans le passé, jusqu'à une centaine de milliers d'années. L'estimation que nous obtenons est de l'ordre de 50-60000 ans, bien plus raisonnable et proche des estimations biologiques de plusieurs généticiens, comme Spencer Wells. Mais les nombreuses approximations du modèle nous empêchent d'avoir une trop grande confiance dans cette estimation. Remarquons que cet âge correspond à celui de la mère de nos les mères, pas au premier ancêtre commun que l'on peut trouver à toute l'humanité, c'est à dire qui engendre toute l'humanité, mais à la fois par ses fils et ses filles. En effet cet individu, d'après la plupart des estimations aurait vécu aux environs de l'an 1000.

3.2 La vitesse de descente de l'infini du coalescent de Kingman

Nous allons maintenant étudier le nombre d'individus dans le coalescent de Kingman aux premiers instants. Ceci nous donnera le comportement du coalescent avec un très grand nombre d'individus. En fait, pour t proche de 0, on a environ $\frac{2}{t}$ individus. Pour cela, nous reprenons simplement l'évaluation précédente :

$$\mathbb{P}(D_t \leq k) = \mathbb{P}(S_k \leq t),$$

où on rappelle que $S_k = \sum_{i=k+1}^{+\infty} \frac{2}{i(i-1)} e_i$. Le temps d'atteinte du niveau k , que nous noterons T_k est donc égal en loi à S_k .

Nous calculons alors :

$$\mathbb{E}(T_k) = \sum_{i=k+1}^{+\infty} \frac{2}{i(i-1)} = \sum_{i=k+1}^{+\infty} \frac{2}{i-1} - \frac{2}{i} = \frac{2}{k},$$

$$\mathbb{V}\text{ar}(T_k) = \sum_{i=k+1}^{+\infty} \frac{4}{(i(i-1))^2} \sim \frac{4}{k^3}.$$

Connaissant l'espérance et la variance de T_k , nous pouvons calculer la probabilité que T_k soit loin de $\frac{2}{k}$. Nous avons en effet :

$$\begin{aligned} \mathbb{P}\left(\left|T_k - \frac{2}{k}\right| \geq \frac{\ln k}{k}\right) &= \mathbb{P}\left(\left(T_k - \frac{2}{k}\right)^2 \geq \left(\frac{\ln k}{k}\right)^2\right) \\ &\leq \frac{k^2}{\ln k^2} \mathbb{E}\left(T_k - \frac{2}{k}\right)^2 \\ &\leq \frac{C \ln k^2}{k}. \end{aligned}$$

Nous allons maintenant utiliser un résultat très important et très utile de la théorie des probabilités, le lemme de Borel-Cantelli. Celui-ci postule que si les probabilités associées d'une suite d'évènements décroissent assez vite, alors presque sûrement, il y a seulement un nombre fini de ces évènements qui sont vérifiés. Plus précisément :

Lemme 3.1. *Soit A_k une suite d'évènements. Si $\sum_{k=0}^{+\infty} \mathbb{P}(A_k) < +\infty$ alors on a :*

$$\mathbb{P}(A_k \text{ n'est plus vérifié à partir d'un certain rang } k_0) = 1.$$

Preuve. On considère, pour chaque entier k la fonction f_k qui vaut 1 si A_k est vérifié et 0 sinon. Dans ce cas, nous voyons que $\sum_{k=0}^{+\infty} \mathbb{P}(A_k)$ est l'espérance de la fonction $\sum_{k=0}^{+\infty} f_k$. Cette espérance est finie, donc la fonction l'est aussi (si la fonction était infinie, son espérance le serait aussi). Cela signifie en particulier que l'on est sûr qu'au plus un nombre fini d'évènements sont vrais. Par conséquent, pour k_0 assez grand, A_k n'est plus vérifié pour tout $k \geq k_0$. \square

Nous allons maintenant appliquer ce lemme à la suite des évènements :

$$B_p = \left\{ \left| T_{p^2} - \frac{2}{p^2} \right| \geq \frac{\ln p^2}{p^2} \right\}.$$

On a vu précédemment que la probabilité d'un tel évènement est de l'ordre de $\frac{\ln p}{p^2}$, grâce au théorème de Bienaymé-Tchebychev la série des probabilités est donc convergente. Par conséquent, nous savons que presque sûrement, il existe un rang au delà duquel on a :

$$\left| T_{p^2} - \frac{2}{p^2} \right| \leq \frac{\ln p^2}{p^2}.$$

Par conséquent, nous avons :

$$\lim_{p \rightarrow +\infty} p^2 T_{p^2} = 2 \text{ presque sûrement,}$$

On obtient donc pour tout entier n compris entre p^2 et $(p+1)^2$:

$$p^2 T_{(p+1)^2} \leq n T_n \leq (p+1)^2 T_{p^2},$$

donc par théorème des gendarmes, nous obtenons $\lim_{n \rightarrow +\infty} n T_n = 2$ presque sûrement. Ce résultat se traduit de la même manière sur le processus du nombre de blocs D_t . On obtient :

$$\lim_{t \rightarrow 0} t D_t = 2 \text{ p.s.}$$

ce qui est bien ce que nous souhaitons démontrer, au voisinage de 0, D_t se comporte comme $\frac{2}{t}$, il y a environ N individus à l'instant $\frac{2}{N}$.

Ce résultat peut être précisé. Grâce à un résultat de la théorie des martingales, on peut prouver qu'en réalité, les n -coalescents de Kingman se comportent dans leurs premiers instants comme la fonction $t \mapsto \frac{2}{t + \frac{2}{n}}$. Ces résultats ont été développés par Julien Berestycki, Nathanaël Berestycki, et Vlada Limic, pour des classes plus générales de processus coalescents [2].

Nous allons maintenant nous intéresser à la taille des groupes de descendants d'un individu donné. Pour ce faire, il sera utile de donner une autre construction du coalescent de Kingman, qui donne directement accès à ces quantités, une fois leur définition précisée. Il nous faudra bien entendu démontrer que les deux constructions donnent le même processus coalescent, mais par la suite, en réalisant des comparaisons, nous pourrons obtenir des résultats d'autant plus forts sur le coalescent de Kingman.

4 Taille des groupes de cousins

Nous nous sommes jusqu'ici intéressés au nombre de lignées distinctes présentes à l'instant t dans l'arbre généalogique. Nous allons maintenant étudier le nombre de descendants de chacune de ces lignées. Autrement dit, en étudiant un ancêtre à la $t^{\text{ième}}$ génération, quel est la taille de la famille qu'il a engendré. Nous allons obtenir des estimations intéressantes sur nombre de descendants de la lignée la plus prolifique, et de celle ayant le moins de descendants.

Pour les obtenir, nous aurons besoin d'une nouvelle construction du coalescent de Kingman. En effet, dans notre construction actuelle, nous connaissons les « probabilités de transition », c'est-à-dire la façon de passer d'une situation donnée à une autre. Mais celle-ci ne donne pas la loi de la répartition des tailles des familles. Notre plan est de construire un autre processus d'une manière différente, en donnant cette répartition à chaque instant. Nous vérifierons ensuite que ce processus est bien un coalescent de Kingman. Grâce à l'unicité de celui-ci, nous obtiendrons l'égalité entre les deux processus, en particulier la distribution des tailles de familles, qui nous intéresse. Nous allons commencer par une manière de mesurer la taille des partitions de \mathbb{N} . Nous nous intéresserons ensuite à la manière de répartir les descendants de n familles. Nous finirons en construisant la suite des partitions apparaissant au cours du temps.

FIGURE 4.7 – Construction sur $[0, 1]$ de $K(5)_{|10}$, les dix premiers éléments de la partition de \mathbb{N} en 5 ensembles.

4.1 Une autre construction du coalescent de Kingman

Nous nous intéressons à une partition de \mathbb{N} en n ensembles. Remarquons pour commencer que l'un d'entre eux au moins est infini. Afin de connaître leur taille, nous allons essayer de donner un sens à la proportion d'individus dans chacun de ces blocs. Nous souhaitons par exemple que l'ensemble des entiers pairs soit considéré comme ayant une proportion $\frac{1}{2}$, celui contenant les multiples de 3 comme ayant une proportion $\frac{1}{3}$, etc.

Nous définissons ainsi la proportion d'un bloc B de \mathbb{N} :

$$|B| = \lim_{n \rightarrow +\infty} \frac{1}{n} |B \cap \{1, \dots, n\}| \text{ si cette limite existe.}$$

Cette limite n'existe pas toujours, et dans ce cas-là, nous noterons $|B| = \partial$, ce qui signifiera que cette quantité n'est pas définie. Par exemple, le bloc D constitué des entiers compris entre 2^{2p} et 2^{2p+1} pour tout entier p n'a pas de proportion définie. En effet, on a :

$$\lim_{p \rightarrow +\infty} \frac{1}{2^{2p} - 1} |D \cap \{1, \dots, 2^{2p} - 1\}| = \frac{1}{3}$$

$$\lim_{p \rightarrow +\infty} \frac{1}{2^{2p+1} - 1} |D \cap \{1, \dots, 2^{2p+1} - 1\}| = \frac{2}{3},$$

donc la suite n'a pas de limite.

Nous allons maintenant donner la procédure inverse : un moyen d'obtenir une partition de \mathbb{N} possédant des proportions fixées. Soient $\{x_1, \dots, x_n\}$ des valeurs comprises entre 0 et 1 telles que $x_1 + \dots + x_n = 1$. Nous construisons une partition aléatoire Π possédant n blocs de telle manière que le $i^{\text{ième}}$ bloc possède la proportion x_i . Pour cela on choisit $(U_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires indépendantes de loi uniforme sur $[0, 1[$. L'individu n est placé dans la partition i si la valeur U_n appartient à l'intervalle $[\sum_{j=1}^{i-1} x_j, \sum_{j=1}^i x_j[$. Dans ce cas, la loi des grands nombres assure que la partition Π possède les bonnes proportions.

Nous allons maintenant montrer qu'un coalescent de Kingman peut être construit en utilisant cette méthode. Nous prenons deux ensembles de lois uniformes sur $[0, 1]$ indépendantes et uniformément distribuées $(U_n)_{n \in \mathbb{N}}$ et $(V_n)_{n \in \mathbb{N}}$. Les variables aléatoires $(U_n)_{n \in \mathbb{N}}$ permettent de former les partitions de \mathbb{N} selon la méthode précédente, et les variables aléatoires $(V_n)_{n \in \mathbb{N}}$ permettent de définir les proportions de la partition aléatoire que nous allons définir.

Pour tout $n \in \mathbb{N}$, on regarde V_1, \dots, V_{n-1} qui, distribuées sur le segment $[0, 1]$ comme formant une subdivision de ce segment en n intervalles. On note $K(n)$ la partition aléatoire de \mathbb{N} construit sur cette subdivision, qui comporte donc n individus.

Nous allons maintenant essayer de définir un coalescent de Kingman avec ces quantités. Pour espérer un tel résultat, il faut qu'après un temps exponentiel de paramètre $\frac{n(n-1)}{2}$, on passe du bloc $K(n)$ au bloc $K(n-1)$. Posons donc D_t un processus à valeurs entières défini de la manière suivante : quand t tend vers 0, D_t tend vers l'infini, et en tout instant positif t , si D_t est égal à k , après un temps exponentiel de paramètre $\frac{k(k-1)}{2}$ D_t diminue d'une unité.

Posons $\Pi'_t = K(D_t)$ pour tout instant $t > 0$ et Π'_0 la partition de \mathbb{N} en singletons. Nous pouvons alors prouver que le processus $(\Pi'_t)_{t \geq 0}$ est aussi un processus coalescent de Kingman. En effet, lors d'un saut de D_t , on passe d'un processus $K(n)$ à $K(n-1)$, construit par le retrait au hasard d'une partition. Il se trouve que c'est simplement la coagulation au hasard de deux blocs. Par conséquent, lorsqu'on arrive dans la partition $K(n)$, le processus se comporte comme un coalescent de Kingman. Ceci étant vrai pour tout entier n , on obtient en passant à la limite quand n tend vers l'infini que Π' est un coalescent de Kingman.

Mais le coalescent de Kingman est unique en loi, par conséquent, tout résultat que nous obtenons sur l'une des deux versions que nous avons construites s'étend aussitôt à l'autre. C'est le cas en particulier de la distribution de la partition $K(n)$, la partition des familles obtenues lorsqu'il ne nous reste que n individus dans le coalescent de Kingman.

4.2 Taille des familles d'individus

Avec cette nouvelle représentation, il devient aisé de calculer la taille des groupes de cousins. En effet, lorsqu'on arrive à n lignées, les proportions de chacune des fratries sont distribuées comme la taille de la subdivision de $[0, 1]$ par une collection de n variables aléatoires uniformes indépendantes. Le problème est que les calculs sur les longueurs des intervalles créés par cette partition sont malaisés. Nous allons donc donner le résultat d'un lemme technique, dont la preuve sera donnée en annexe. Ce lemme donne une représentation permettant de manipuler aisément ces longueurs d'intervalles.

Lemme 4.1. *Soit e_1, \dots, e_n n variables aléatoires indépendantes de loi exponentielle de paramètre 1 et $\gamma_n = e_1 + \dots + e_n$. La loi de la taille des subdivisions de $[0, 1]$ par $n-1$ variables aléatoires indépendantes uniformes est égale à celle de $(\frac{e_1}{\gamma_n}, \dots, \frac{e_n}{\gamma_n})$. En d'autres termes, ce jeu de variables aléatoires réalise la même subdivision de $[0, 1]$ que les variables aléatoires uniformes.*

Preuve. Ce lemme se prouve par un simple changement de variables, il faut démontrer que pour toute fonction continue bornée f , l'espérance de f est égale pour les deux jeux de variables aléatoires. Pour cela, notons tout d'abord que si on pose x_1, \dots, x_n les variables aléatoires donnant les tailles de la subdivision, on a :

$$\mathbb{E}(f(x_1, \dots, x_n)) = n! \int_{0 \leq s_1 \leq \dots \leq s_{n-1} \leq 1} f(s_1, s_2 - s_1, \dots, 1 - s_{n-1}) ds_1, \dots, ds_n.$$

En effet, nous réalisons un réarrangement dans l'ordre croissant de V_1, \dots, V_{n-1} , il y a $n!$ manières de le faire, ce qui donne bien ce que nous obtenons.

Nous pouvons également écrire :

$$\mathbb{E}(f(\frac{e_1}{\gamma_n}, \dots, \frac{e_n}{\gamma_n})) = \int_{\mathbb{R}^n} f(\frac{s_1}{s_1 + \dots + s_n}, \dots, \frac{s_n}{s_1 + \dots + s_n}) e^{-s_1 - s_2 - \dots - s_n} ds_1 \dots ds_n,$$

ce qui en réalisant le changement de variables $u = s_1 + \dots + s_n$ se réécrit :

$$\mathbb{E}(\frac{e_1}{\gamma_n}, \dots, \frac{e_n}{\gamma_n}) = \int_{\mathbb{R}} \int_{s_1 + \dots + s_{n-1} \leq u} f(\frac{s_1}{u}, \dots, \frac{s_{n-1}}{u}, \frac{u - s_1 - \dots - s_{n-1}}{u}) e^{-u} ds_1 \dots ds_{n-1} du.$$

On réalise ensuite le changement de variables $v_i = \frac{s_i}{u}$, on obtient :

$$\mathbb{E}\left(\frac{e_1}{\gamma_n}, \dots, \frac{e_n}{\gamma_n}\right) = \int_{\mathbb{R}} \int_{v_1 + \dots + v_{n-1} \leq 1} f(v_1 \dots v_{n-1}, 1 - v_1 + \dots + v_{n-1}) u^{n-1} e^{-u} dv_1 \dots dv_{n-1} du.$$

On réalise alors que $\int_0^{+\infty} u^{n-1} e^{-u} du = n!$ par intégrations par parties successives, et on réalise un dernier changement de variables $s_i - s_{i-1} = v_i$ et on retrouve bien la première espérance. On obtient par conséquent qu'il revient au même de considérer les variables aléatoires (x_1, \dots, x_n) et $(\frac{e_1}{\gamma_n}, \dots, \frac{e_n}{\gamma_n})$. \square

Ce lemme nous permet alors de revoir d'un œil nouveau les proportions de chacun des groupes de parentés dans le coalescent de Kingman lorsque nous avons n lignées restantes. Cette représentation est bien plus agréable à manipuler pour ce qui est du minimum et du maximum. Nous en déduisons en particulier les résultats suivants. On note A_n et a_n respectivement la plus grande et la plus petite proportion d'individus lorsqu'on remonte à n lignées.

On a alors $nA_n - \ln n \xrightarrow[n \rightarrow +\infty]{} G$ et $n^2 a_n \xrightarrow[n \rightarrow +\infty]{} e$, où on a posé G une variable aléatoire vérifiant $\mathbb{P}(G \leq x) = \exp(-e^{-x})$ et e une variable aléatoire de loi exponentielle. En d'autres termes, quand n devient grand, la plus grande famille a une taille de l'ordre de $\frac{\ln n + G}{n}$ et la plus petite de l'ordre de $\frac{e}{n^2}$.

En effet, grâce à la représentation en loi exponentielle, nous avons $\gamma_n A_n$ la loi de la plus grande de n variables aléatoires indépendantes et de même loi exponentielle de paramètre 1. Nous avons donc :

$$\begin{aligned} \mathbb{P}(\gamma_n A_n - \ln n \leq x) &= \mathbb{P}(\max e_1, \dots, e_n \leq x + \ln n) \\ &= \mathbb{P}(e_1 \leq x + \ln n, \dots, e_n \leq x + \ln n) \\ &= \mathbb{P}(e_1 \leq x + \ln n)^n = \left(1 - \frac{e^{-x}}{n}\right)^n. \end{aligned}$$

Grâce à un lemme précédent, on obtient :

$$\gamma_n A_n - \ln n \xrightarrow[n \rightarrow +\infty]{} G,$$

avec $\mathbb{P}(G \leq x) = \exp(-e^{-x})$. Or la loi des grands nombres, indiquant que la moyenne d'un grand nombre d'observations indépendantes converge vers sa moyenne nous permet de dire que

$$\frac{\gamma_n}{n} \xrightarrow[n \rightarrow +\infty]{} 1,$$

on obtient finalement que $nA_n - \ln n \xrightarrow[n \rightarrow +\infty]{} G$. On dit que G suit la loi de Gumble.

Pour ce qui est de a_n , on a de la même manière :

$$\begin{aligned} \mathbb{P}(n\gamma_n a_n \geq x) &= \mathbb{P}(\min e_1, \dots, e_n \geq \frac{x}{n}) \\ &= \mathbb{P}(e_1 \geq \frac{x}{n})^n \\ &= (e^{-\frac{x}{n}})^n = e^{-x}. \end{aligned}$$

On obtient donc que $n^2 a_n \xrightarrow{n \rightarrow +\infty} e$, comme annoncé.

Grâce à l'équivalent en 0 du coalescent de Kingman, nous pouvons également réécrire ce résultat en terme de temps passé et non de nombre de lignées restantes. On note maintenant, si Π est un coalescent de Kingman, $M(\Pi_t)$ la taille de la plus grande famille à l'instant t et $m(\Pi_t)$ la taille de la plus petite famille. Nous avons alors :

$$\frac{2}{t} M(\Pi_t) + \ln \frac{t}{2} \xrightarrow{t \rightarrow 0} G \text{ et}$$

$$\frac{m(\Pi_t)}{t^2} \xrightarrow{t \rightarrow 0} \frac{e}{4}.$$

Ces estimations peuvent être étendues à d'autres quantités, la taille de la deuxième plus grande famille ou d'autres calculs similaires. L'idée force est ici que nous avons réussi à calculer les tailles des familles, qui n'étaient pas accessibles à première vue, en construisant un autre processus, qui a la même loi que le premier. Utiliser ce double regard nous a permis de répondre à de nouvelles questions. Nous allons maintenant nous intéresser au patrimoine génétique des individus, en fonction des mutations qui se répartissent le long de l'arbre généalogique.

5 Mutation neutre et classification phylogénique

Nous allons maintenant enrichir l'arbre généalogique des individus auxquels nous nous intéressons par un certain nombre d'évènements de mutation. Plus précisément, nous considérons que chaque individu, à chaque génération, a une probabilité (faible) de subir une mutation neutre. Une mutation neutre est sans impact sur le nombre de descendants que l'individu concerné peut avoir, mais se conserve comme une signature chez tous ses descendants. Nous allons maintenant nous intéresser aux ensembles d'individus qui possèdent les mêmes mutations, plutôt qu'à ceux qui descendent de la même lignée. C'est donc un changement radical dans notre étude du processus. Jusqu'ici nous nous intéressions uniquement au comportement du processus à un instant t fixé. Maintenant, nous regardons le processus comme un arbre, achevé, qui remonte jusqu'à un ancêtre commun. Nous plaçons ensuite un certain nombre de marques sur les branches de cet arbre, et nous nous intéresserons au nombres de sous-groupes créés quand nous coupons les branches de l'arbre en ces marques.

Nous créons ces marques de la manière suivante : à chaque épisode de coagulation t , nous lançons une variable aléatoire exponentielle de paramètre $\frac{\theta}{2}$ pour chaque individu présent dans le coalescent. Si certaines de ces variables aléatoires sont plus petites que l'instant de coagulation suivant, nous marquons la branche associée aux individus concernés à une distance de l'instant t donnée par la variable aléatoire exponentielle. Nous relançons ensuite une autre variable aléatoire exponentielle que l'on compare à la distance entre cet instant de mutation et l'instant de coagulation suivant, puis on continue. Un tel arbre est dit marqué par un processus de Poisson d'intensité $\frac{\theta}{2}$ par unité de longueur des branches.

Nous allons maintenant étudier comment n individus choisis se répartissent en fonction des mutations qu'ils possèdent. Regardons précisément la dynamique de ces n individus. Chaque individu va subir une mutation selon une loi exponentielle de paramètre $\frac{\theta}{2}$ et un évènement

FIGURE 4.8 – Un coalescent de Kingman marqué, avec les différents génotypes indiqués par des couleurs.

de coalescence peut arriver selon une loi exponentielle de paramètre $\frac{n(n-1)}{2}$. Grâce au Lemme des réveils 1.1, le premier de ces événements arrive suivant une loi exponentielle de paramètre $\frac{n(n-1)+\theta n}{2}$, et c'est un événement de coagulation avec une probabilité de $\frac{n(n-1)}{n(n-1)+\theta n} = \frac{n-1}{n-1+\theta}$, ou un événement de mutation avec pour probabilité $\frac{\theta}{n-1+\theta}$. Dans ce cas, on considère la lignée ayant muté comme étant perdu, on l'oublie dans notre calcul suivant.

Plus généralement, si à un instant donné il reste k individus, l'évènement suivant sera un évènement de mutation avec la probabilité $\frac{\theta}{k-1+\theta}$ et un évènement de coagulation avec une probabilité de $\frac{k-1}{k-1+\theta}$. Cet évènement arrivera selon une variable aléatoire exponentielle de paramètre $\frac{k(k-1)+\theta k}{2}$.

Nous allons maintenant réétudier notre parcours dans le sens opposé du temps. En fait, nous suivons alors le sens conventionnel du temps, du passé vers le futur, car dans le coalescent de Kingman nous remontions le temps. Nous partons de la mutation la plus récente de l'ancêtre commun de la population de n individus. Le premier évènement qui arrive ensuite est une mutation avec la probabilité $\frac{\theta}{1+\theta}$ et une coagulation avec la probabilité $\frac{1}{1+\theta}$. En d'autres termes, le deuxième individu observé possède un génome différent avec une probabilité $\frac{\theta}{1+\theta}$, et possèdera les mêmes mutations que l'ancêtre commun avec la probabilité $\frac{1}{1+\theta}$. Le $k^{\text{ième}}$ individu qui arrive sera causé par une mutation avec la probabilité $\frac{\theta}{k-1+\theta}$, et possèdera donc une nouvelle mutation, qui le mettra à part de tous les autres, ou avec la probabilité $\frac{k-1}{k-1+\theta}$ sera causé par une coagulation. Dans ce cas, le nouvel individu sera le frère d'un individu existant choisi uniformément au hasard, et partagera donc exactement les mêmes mutations que lui.

Nous pouvons alors généraliser ceci pour tout entier n , donc pour la population infinie, qui se distribuera selon cette dynamique. Celle-ci s'appelle le processus du restaurant chinois, car il rappelle le processus suivant. Un client arrive dans un restaurant chinois et s'installe à la première table disponible. Un deuxième arrive. Avec la probabilité $\frac{1}{1+\theta}$ il connaît le premier client et s'installe à sa table, sinon il s'installe à la table suivante. Un troisième client entre, il connaît l'un des deux clients déjà attablés avec la probabilité $\frac{1}{2+\theta}$ et s'installe à la table de ce dernier dans ce cas-là, sinon il s'installe à la table suivante, etc.

Ce processus est exactement le même que celui selon lequel se créent les partitions correspondant aux mutations, et les clients assis à une même table représentent les individus partageant les mêmes gènes. Le numéro de la table représente l'ordre dans lequel ces diverses mutations ont apparu sur l'arbre du coalescent de Kingman. Ce processus est bien connu, nous connaissons en particulier la probabilité que les n premiers individus soient divisés en k ensembles de taille n_1, \dots, n_k dans l'ordre indiqué, grâce à la formule d'échantillonnage d'Ewens [3], qui se prouve par récurrence :

$$p_{n_1, \dots, n_k} = \frac{\theta^k}{\theta(\theta-1) \dots (\theta-n+1)} \prod_{i=1}^k (n_i - 1)!$$

La distribution des proportions de chacune de ces classes est plus compliquée à décrire, et utilise

l'algèbre beta-gamma.

Nous pouvons néanmoins conclure en signalant que le nombre total de génotypes distincts observés dans un coalescent de Kingman à n individus est de l'ordre de $\theta \log n$. En effet, à chaque ajout d'un nouvel individu, celui-ci possède un génotype différent du précédent avec la probabilité $\frac{\theta}{n+\theta}$. On peut alors conclure grâce à des résultats de la théorie des grandes déviations.

6 Conclusion

Nous avons, tout au long de cet article, exposé un moyen de modéliser la généalogie d'une population à l'aide d'un modèle, le coalescent de Kingman. Nous avons également étudié plusieurs quantités liées à ce modèle, comme l'âge du dernier ancêtre commun, ou le nombre de familles restant lorsque nous remontons assez loin en arrière. Ces quantités ne peuvent pas s'appliquer aux modèles biologiques sans tenir compte des nombreuses hypothèses que nous faisons, comme la constance de la taille de la population ambiante au cours du temps. Mais elles permettent néanmoins de donner des ordres de grandeurs de quantités apparaissant dans l'étude de l'évolution de certaines populations réelles.

La formule d'échantillonnage d'Ewens est en elle-même un résultat très intéressant, car elle représente la manière dont les individus se répartissent les mutations survenues au cours du temps, lorsque celles-ci sont neutres. Ainsi, en étudiant la proportion d'individus possédant telle ou telle mutation d'un gène, si nous observons une répartition différente de celle exprimée dans la formule d'Ewens, nous saurons que ces mutations ne sont pas neutres, et ont un impact sur la descendance des individus. Cette formule permet donc de confirmer ou d'infirmer l'hypothèse selon laquelle une certaine mutation est neutre.

De nombreux raffinements du coalescent de Kingman existent. On peut par exemple autoriser la coagulation simultanée d'un grand nombre d'individus, c'est-à-dire que l'on peut autoriser que certains individus exceptionnels engendrent d'un seul coup une portion non-négligeable de la population totale. On peut également autoriser que plusieurs de ces coagulations simultanées se produisent en même temps. Ces généralisations s'appellent Λ -coalescents ou processus de coalescence échangeable.

Nous pouvons également autoriser les individus à se déplacer sur une carte de taille finie ou infinie. Dans ce cas, on suppose que les individus sont localisés en certains endroits précis appelés colonies, et qu'un individu passe d'une colonie à l'autre de manière aléatoire. Cette nouvelle composante du problème change de manière radicale le problème posé. En effet, si dans le cas du coalescent de Kingman non-spatial, il existe presque sûrement un ancêtre commun à cette population, lorsqu'on permet aux individus de se déplacer sur un graphe infini, cet ancêtre commun n'existe pas presque sûrement.

7 Preuve du Théorème 2.1

Considérons Π^n un n -coalescent de Kingman qui à l'instant t est dans la partition π . Lorsqu'on s'intéresse à la restriction de ce processus à $n - 1$ individus, deux cas se présentent.

Commençons par le cas où le bloc de π contenant n contient également un autre entier $i < n$. Alors de la restriction de Π^n à \mathcal{P}_{n-1} on peut remonter au processus original en ajoutant à tout

instant $t > 0$ l'entier n à la partition contenant i . Par conséquent, $\Pi_{|n-1}^{(n)}$ saute au mêmes instants que le processus non restreint, et arrive dans les mêmes partitions, restreintes à $n-1$. C'est donc un $n-1$ -coalescent de Kingman étant à l'instant t dans la partition $\Pi_t^n|_{n-1}$.

Supposons maintenant que le bloc $\{n\} \in \pi$, et que π possède k blocs, autrement dit l'individu n n'a pas encore d'ancêtre commun dans la population. Dans ce cas deux types de sauts peuvent se produire, et l'un d'eux implique l'individu n , qu'on ne peut voir dans le processus restreint. Soit τ le premier instant de saut, et A l'évènement $\{\{n\} \in \Pi_\tau^n\}$ sur lequel l'individu n n'est pas impliqué. Soit $\tau + \tau'$ le deuxième instant de saut.

Remarquons alors que le premier temps de saut de $\Pi_{|n-1}^n$ est $\tilde{\tau} = \tau + \mathbf{1}_{\{A^c\}}\tau'$, c'est-à-dire que le premier instant de saut du processus restreint est égal au premier instant de saut du processus non-restreint si cet évènement ne concerne pas l'individu n , au deuxième instant de saut si le premier évènement est une fusion de n avec un autre individu. De plus, τ est une variable aléatoire exponentielle de paramètre $\frac{k(k-1)}{2}$, A est un évènement indépendant de τ de probabilité $1 - \frac{2}{k}$, et τ' est une variable aléatoire exponentielle indépendante de paramètre $\frac{(k-1)(k-2)}{2}$. Nous avons alors :

$$\begin{aligned} \mathbb{P}(\tau + \mathbf{1}_{\{A^c\}}\tau' > s) &= \frac{2}{k}\mathbb{P}(\tau + \tau' > s) + \left(1 - \frac{2}{k}\right)\mathbb{P}(\tau > s) \\ &= \frac{2}{k}\left(\frac{k}{2}e^{-\frac{(k-1)(k-2)}{2}s} - \frac{k-2}{2}e^{-\frac{k(k-1)}{2}s}\right) + \left(1 - \frac{2}{k}\right)e^{-\frac{k(k-1)}{2}s} \\ &= e^{-\frac{(k-1)(k-2)}{2}s}. \end{aligned}$$

Le premier instant de saut pour le processus restreint est donc de loi exponentielle de paramètre $\frac{(k-1)(k-2)}{2}$. De plus $\Pi_{|n-1}^n$ est uniforme sur les partitions de $\{1, \dots, n-1\}$ et indépendant de $\tilde{\tau}$. C'est donc bien le premier saut d'un $n-1$ -coalescent de Kingman.

On passe donc d'une partition à $k-1$ blocs à une partition à $k-2$ blocs comme un $n-1$ coalescent de Kingman, et par récurrence descendante, grâce aux deux cas déjà traités, on obtient bien que Π_{n-1}^n est un $n-1$ -coalescent de Kingman, car au bout d'un certain temps, l'individu n admet un ancêtre commun avec l'un au moins des autres individus.

Bibliographie

- [1] O. Angel, N. Berestycki, and V. Limic. Global divergence of spatial coalescents. *Preprint*, 2009.
- [2] J. Berestycki, N. Berestycki, and V. Limic. The Λ -coalescent speed of coming down from infinity. *Annals of Probabilities*, 2009. To appear.
- [3] W.J. Ewens. The sampling theory of selectively neutral alleles. *Theoretical Population Biology*, 3.
- [4] J.F.C. Kingman. The coalescent. *Stochastic Processes and Applications*, pages 235–248, 1982.
- [5] J.F.C. Kingman. On the genealogy of large populations. *Journal of Applied Probabilities*, pages 29–43, 1982.

Troisième partie

Stage de M1

Chapitre 5

Convergence de certain processus de particules vers des processus de Dawson-Watanabe

Bastien MALLEIN

Sous la direction d'Edwin A. PERKINS

8 septembre 2010

Traduit le 7 juin 2011

Résumé

Nous étudions ici la convergence de certains processus de particules vers des super-processus de Dawson-Watanabe, en utilisant la convergence de leurs dérivées de Radon-Nikodým par rapport à la loi de la marche aléatoire branchante, comme l'ont fait Lalley et Zheng dans leur article [8] pour les processus SIR en dimension 2 et 3. Les processus que nous étudierons de cette manière sont les marches aléatoires branchantes avec dérive, les processus de contact, le modèle du voteur et les modèles de Lotka-Volterra.

Je tiens à remercier toutes les personnes qui m'ont aidé à réaliser ce rapport, et en particulier Ed Perkins, qui m'a présenté ce sujet, pour la patience et le temps dont il a usé avec moi. Je remercie aussi Max Fathi pour sa relecture attentive, ainsi que Jean-François Le Gall, qui m'a offert cette opportunité de partir en stage à UBC.

Sommaire

1	Introduction	88
2	Dérivée de Radon-Nikodým des processus de particules	92
2.1	Sur les processus de particules à temps continu	92
2.2	Dérivée de Radon-Nikodým pour les processus de particules à temps discret	94
3	Quelques calculs relatif à la marche aléatoire branchante	95
3.1	Représentation de la marche aléatoire branchante à temps continu	95
3.2	Calcul des moments	96
3.3	Calculs pour la marche aléatoire branchante à temps discret	97
4	Convergence de la marche aléatoire branchante avec dérive	97
4.1	Marche aléatoire branchante à temps continu avec dérive symétrique θ	97
4.2	Marche aléatoire branchante avec dérive, modèle à temps discret	99
4.3	Une marche aléatoire branchante avec dérive asymétrique (θ, θ')	101
5	Convergence du processus de contact	102
5.1	Processus de contact à temps continu	103
5.2	Processus de contact à temps discret	105
6	Convergence pour le modèle du voteur	107
6.1	Modèle du voteur à long rayon d'interaction	107
6.2	Autour du modèle du voteur à court rayon d'interaction	108
7	Le modèle de Lotka-Volterra, une modification du modèle du voteur	108
7.1	Modèle de Lotka-Volterra à temps continu	109
7.2	Modèle de Lotka-Volterra à temps discret	110
8	Conclusion	111
9	Annexe	111
9.1	Autour de la marche aléatoire	112
9.2	Nombre moyen de voisins d'une particule	112

1 Introduction

Les processus de particules que nous étudierons par la suite sont des processus de naissance et de mort de particules. Ce sont des processus qui peuvent évoluer de la façon suivante : chaque particule peut donner naissance à un certain nombre d'enfants en un site voisin du sien, ou mourir. Ces processus évoluent sur le réseau \mathbb{Z}^d . Le taux auquel une particule située sur le site x donne naissance ou meurt ne dépend que du nombre de particules situées en x et du nombre de particules dans un voisinage de x .

Nous allons maintenant introduire quelques notations et du vocabulaire que nous utiliserons tout au long de ce rapport. Pour commencer, précisons l'heuristique des processus de particules à travers deux définitions. Nous nous intéresserons en effet à la fois à des processus à temps continu et à temps discret, les problèmes se posant dans chacun des cas sont voisins, mais différents. Il est donc intéressant de dresser un parallèle entre ces deux types de processus.

Nous notons p un noyau de marche aléatoire symétrique irréductible sur \mathbb{Z}^d , vérifiant :

$$p(0) = 0 \text{ et } \sum_x x^i x^j p(x) = \delta_{i,j} \sigma^2 < +\infty,$$

qui représente la manière selon laquelle les particules se déplacent sur le réseau. Pour toute fonction bornée $\phi : \mathbb{Z}^d \rightarrow \mathbb{R}$, nous notons $P\phi(x) = \sum_e p(e)\phi(x+e)$ la moyenne locale de ϕ par rapport à p .

Définition 1.1. Un système de particules à temps continu, de noyau p , taux de naissance b_t et taux de mort k_t est un processus de Markov $\xi_t : \mathbb{Z}^d \rightarrow \mathbb{N}$, tel que $\xi_t(x)$ représente le nombre de particules présentes en x à l'instant t . Cette population croît si un voisin produit un enfant en x , et décroît si une particule située en x meurt. Les deux taux $b_t(x)$ et $k_t(x)$ auxquels ces événements se produisent sont des fonctions de $x, t, \xi_{t-}(x)$ et $V_{t-}(x) = \sum_e p(e)\xi_{t-}(x+e)$. Nous pouvons donc écrire :

$$\begin{cases} \xi_t(x) \rightarrow \xi_t(x) + 1 \text{ à taux } V_t(x)b_t(x) \\ \xi_t(x) \rightarrow \xi_t(x) - 1 \text{ à taux } \xi_t(x)k_t(x). \end{cases}$$

Définition 1.2. Un système de particules à temps discret, de noyau p , et distribution du nombre d'enfants Π est une chaîne de Markov $\xi_n : \mathbb{Z}^d \rightarrow \mathbb{N}$, tel que $\xi_n(x)$ représente le nombre de particules présentes en x à l'instant n . La probabilité $\Pi_{n,x}(\cdot)$ sur \mathbb{N} est une fonction de $n, x, \xi_n(x)$ et $V_n(x) = \sum_e p(e)\xi_n(x+e)$, et représente le nombre d'enfants produits à la génération n par les individus en x . Nous pouvons réécrire ceci de la manière suivante :

$$\xi_{n+1}(x) = k \text{ avec probabilité } \Pi_{n,x}(k).$$

Remarque 1.1. Il existe quelques différences dans nos définitions des modèles à temps continu ou à temps discret. Dans le premier cas, nous donnons deux taux différents qui jouent tous deux un rôle à la fois sur la vitesse à laquelle le processus évolue (doubler b_t et k_t revient à considérer $(\xi_{2t}, t \geq 0)$) ; et sur le nombre d'enfants d'un individu donné avant sa mort (grossièrement, si b_t et k_t restent constants durant toute la vie de la particule, le nombre d'enfants de celle-ci suit une loi géométrique de paramètre $\frac{b_t(x)}{b_t(x)+k_t(x)}$, donc doubler b_t et k_t ne change rien).

Dans le modèle à temps discret, nous ne tenons compte que du nombre d'enfants produit en chaque site, sous-entendu par les individus des sites voisins. Les variations de vitesse du processus peuvent être introduites par la suite, par mise à l'échelle, en considérant qu'une étape en position x à l'instant t correspond à un saut dans le temps de $\gamma_n \sim b_n + k_n$, à condition que cette quantité soit sensiblement la même à proximité de toutes les particules en vie à l'instant n .

Nous allons maintenant définir le type de processus limites que nous espérons pouvoir trouver, les super-processus de Dawson-Watanabe, comme processus à valeurs dans l'espace $\mathcal{M}_f(\mathbb{R}^d)$ des mesures finies sur \mathbb{R}^d .

Définition 1.3. Un super-processus de Dawson-Watanabe de taux de branchement $\gamma > 0$, de dérive $\theta_t : \mathbb{R}^d \rightarrow \mathbb{R}$, de coefficient de diffusion σ^2 , issu de $X_0 \in \mathcal{M}_f(\mathbb{R}^d)$ est un processus $(X_t, t \geq 0)$ à valeurs dans $\mathcal{M}_f(\mathbb{R}^d)$, adapté et presque sûrement continu sur un espace de probabilité muni d'une filtration complète $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, qui résout le problème de martingales suivant :

$$\forall \phi \in C_b^\infty(\mathbb{R}^d), M_t(\phi) = X_t(\phi) - X_0(\phi) - \int_0^t X_s(\gamma \frac{\sigma^2 \Delta \phi}{2}) ds - \int_0^t X_s(\theta_s \phi) ds \quad (5.1)$$

est une (\mathcal{F}_t) -martingale continue issue de 0 et de variation quadratique :

$$\langle M(\phi) \rangle_t = \gamma \int_0^t X_s(\phi^2) ds.$$

L'existence et unicité en loi de la solution de ce problème de martingales est bien connue, posons $\mathbb{P}_{X_0}^{\gamma, \theta, \sigma^2}$ la loi de cette solution sur $\Omega_{X,D} = \mathcal{D}([0, +\infty[, \mathcal{M}_f(\mathbb{R}^d))$ l'ensemble des fonctions càdlàg à valeurs dans $\mathcal{M}_f(\mathbb{R}^d)$.

Remarque 1.2. Un processus de Dawson-Watanabe sans dérive peut aussi être appelé supermouvement Brownien.

Dans ce rapport, nous essaierons de prouver la convergence de certaines suites de processus de particules, convenablement renormalisées, vers des processus de Dawson-Watanabe de la manière suivante. Nous allons calculer la dérivée de Radon-Nikodým du processus de particules que nous étudions contre un autre processus de particules pour lequel la convergence vers un super-processus est connue. Nous nous intéresserons ensuite à la convergence de la suite de dérivées de Radon-Nikodým, et essaierons de prouver que la limite est la dérivée de Radon-Nikodým d'un super-processus de Dawson-Watanabe par rapport à un autre. Le lemme suivant nous permettra de finir les preuves.

Lemme 1.1. Soit X_n, X des variables aléatoires à valeurs dans un espace métrique E toutes définies sur le même espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, et soit L_n, L des variables aléatoires réelles positives sur ce même espace, de moyenne 1.

On pose \mathbb{Q}_n, \mathbb{Q} les probabilités sur (Ω, \mathcal{F}) de dérivée de Radon-Nikodým respective L_n et L par rapport à X_n et X sous la loi \mathbb{P} , si on a :

$$(X_n, L_n) \xrightarrow[n \rightarrow +\infty]{} (X, L),$$

alors les \mathbb{Q}_n -distributions de X_n convergent faiblement vers la \mathbb{Q} -distribution de X .

Preuve. Soit $\phi \in \mathcal{C}_b(E)$, par définition on a :

$$\mathbb{E}_{\mathbb{Q}_n}(\phi(X_n)) = \mathbb{E}_{\mathbb{P}}(L_n \phi(X_n)).$$

Comme $(X_n, L_n) \xrightarrow[n \rightarrow +\infty]{\Longrightarrow} (X, L)$, nous avons, pour tout $A > 0$,

$$\mathbb{E}((L_n \wedge A)\phi(X_n)) \xrightarrow[n \rightarrow +\infty]{} \mathbb{E}((L \wedge A)\phi(X)).$$

De plus, comme $\mathbb{E}(L_n) = 1$, et $L_n \geq 0$, nous obtenons :

$$\mathbb{E}(|L_n \wedge A - L_n|) = \mathbb{E}((L_n - A)\mathbf{1}_{\{L_n > A\}}) = 1 - A - \mathbb{E}((L_n - A)\mathbf{1}_{\{L_n \leq A\}}),$$

converge vers $\mathbb{E}(|L \wedge A - L|)$ quand $n \rightarrow +\infty$. Nous pouvons donc calculer :

$$\begin{aligned} & \limsup_{n \rightarrow +\infty} |\mathbb{E}(L_n \phi(X_n) - L \phi(X))| \\ & \leq \limsup_{n \rightarrow +\infty} \|\phi\|_{\infty} \mathbb{E}(|(L_n \wedge A) - L_n| + |(L \wedge A) - L|) \\ & \leq 2\|\phi\|_{\infty} \mathbb{E}(|(L \wedge A) - L|). \end{aligned}$$

On conclut en faisant tendre A vers $+\infty$. □

Nous avons maintenant besoin d'un processus de particules dont la limite est bien connue, que nous pourrions utiliser comme loi de référence par rapport à tous les autres processus. Le processus que nous choisirons sera la marche aléatoire branchante. Une marche aléatoire branchante est le plus simple des processus que nous avons défini plus haut. Pour le modèle continu, cela revient à prendre les taux de naissance et de mort tous deux égaux à une même constante γ .

Définition 1.4. Une marche aléatoire à temps continu et taux γ et de noyau p est un processus de Markov $\xi_t : \mathbb{Z}^d \rightarrow \mathbb{N}$ qui évolue de la manière suivante :

$$\begin{cases} \xi_t(x) \rightarrow \xi_t(x) + 1 \text{ à taux } \gamma V_t(x) \\ \xi_t(x) \rightarrow \xi_t(x) - 1 \text{ à taux } \gamma \xi_t(x). \end{cases}$$

Afin de prouver la convergence des processus de particules que nous étudierons vers des super-processus de Dawson-Watanabe, nous les renormaliserons en temps, espace et masse, et nous ne tiendrons compte que de la densité de particules au voisinage d'un point. Plus précisément, si ξ^N est une suite de marches aléatoires branchantes de taux γN , nous définissons la suite de processus à valeurs dans $\mathcal{M}_f(\mathbb{R}^d)$ suivante :

$$X_t^N = \frac{1}{N} \sum_x \xi_t^N(x) \delta_{\frac{x}{\sqrt{N}}}.$$

Le support de X_t^N est inclus dans le réseau renormalisé $Z_N^d = \frac{\mathbb{Z}^d}{\sqrt{N}}$. La convergence en loi de cette suite de processus à valeurs dans les espaces de mesure est étudiée dans le théorème suivant :

Théorème 1.1 (Dawson-Watanabe). Si $X_0^N \xrightarrow[N \rightarrow +\infty]{} \mu \in \mathcal{M}_f(\mathbb{R}^d)$ alors :

$$X^N \xrightarrow[N \rightarrow +\infty]{} X,$$

où X est un super-mouvement Brownien de distribution initiale μ de taux 2γ et de diffusion σ^2 . La convergence dans $\mathcal{M}_f(\mathbb{R}^d)$ est la convergence faible des mesures.

On peut prouver le même type de résultat avec des marches aléatoires branchantes à taux discret.

Définition 1.5. Une marche aléatoire à temps discret de noyau p est une chaîne de Markov $\xi_n : \mathbb{Z}^d \rightarrow \mathbb{N}$ qui évolue de la manière suivante :

$$\xi_{n+1}(x) = k \text{ avec probabilité } e^{-V_n(x)} \frac{(V_n(x))^k}{k!}.$$

En d'autres termes, c'est un processus de particules à temps discret pour lequel à chaque étape, toutes les particules meurent en donnant naissance à un nombre d'enfants défini par la distribution suivante :

$$\Pi_{n,x}(k) = e^{-V_n(x)} \frac{(V_n(x))^k}{k!}.$$

Remarque 1.3. Nous avons ici choisi d'utiliser un modèle pour lequel la distribution des enfants suit une loi de Poisson pour simplifier les futurs calculs. D'autres lois peuvent néanmoins être utilisées, comme une distribution géométrique, plus proche du modèle à temps continu, mais les calculs deviennent alors plus complexe.

Soit ξ^N une suite de marches aléatoires à temps discret à taux 1, nous considérons le processus renormalisé accéléré d'un facteur γN suivant :

$$X_t^N = \frac{1}{N} \sum_x \xi_{[\gamma N t]}^N(x) \delta_{\frac{x}{\sqrt{N}}}.$$

On peut encore étudier la convergence de (X^N) .

Théorème 1.2 (Dawson-Watanabe). Si $X_0^N \xrightarrow[N \rightarrow +\infty]{} \mu \in \mathcal{M}_f(\mathbb{R}^d)$, alors on a :

$$X^N \xrightarrow[N \rightarrow +\infty]{} X,$$

où X est un super-mouvement Brownien de taux de branchement γ et de diffusion σ^2 .

Remarque 1.4. La disparition du facteur 2 dans le taux de branchement du processus limite est naturel, si nous nous souvenons des règles de correspondance entre les processus à temps continu et discret. Une multiplication par γ des deux taux de naissance et de mort dans la marche aléatoire branchante à temps continu correspond à une accélération de taux 2γ dans son pendant à temps discret.

2 Dérivée de Radon-Nikodým des processus de particules

2.1 Sur les processus de particules à temps continu

Quelques notations utiles

Nous allons définir quelques quantités, valables pour tous ces systèmes de particules, qui nous seront utiles par la suite, quand nous voudrions calculer les dérivées de Radon-Nikodým.

Pour commencer, nous allons caractériser un processus de particules par un ensemble dénombrable de variables aléatoires dont la loi jointe sera facile à calculer.

Proposition 2.1. *Soit X_t un système de particules de taux de naissance b_t et de taux de mort k_t , nous notons :*

- $T_0 = 0$ et $T_{n+1} = \inf \{t > T_n | X_t \neq X_{T_n}\}$;
- $t_n = T_{n+1} - T_n$;
- $x_n \in \mathbb{Z}^d$ tel que $X_{T_{n+1}}(x_n) \neq X_{T_n}(x_n)$ qui est unique presque sûrement ;
- $\delta_n = X_{T_{n+1}}(x_n) - X_{T_n}(x_n) \in \{-1, 1\}$.

Dans ce cas, connaissant \mathcal{F}_{T_n} , la loi jointe de (t_n, x_n, δ_n) est donnée par

$$\mathbb{P}(t_n \in [t, t + dt[, x_n = x, \delta_n = 1 | \mathcal{F}_{T_n}) = \exp\left(-\int_{T_n}^t X_{T_n}(Pb_s + k_s) ds\right) P(X_{T_n})(x) b_{T_n+t-}(x) dt,$$

$$\mathbb{P}(t_n \in [t, t + dt[, x_n = x, \delta_n = -1 | \mathcal{F}_{T_n}) = \exp\left(-\int_{T_n}^t X_{T_n}(Pb_s + k_s) ds\right) X_{T_n}(x) k_{T_n+t-}(x) dt.$$

De plus, il existe une bijection continue entre $(X_{t \wedge T_n}, t \geq 0)$ and $(t_k, x_k, \delta_k, 0 \leq k < n)$.

Preuve. Nous allons utiliser la propriété de Markov, on peut écrire :

$$\mathbb{P}(t_n \in [t, t + dt[, x_n = x, \delta_n = \delta | \mathcal{F}_{T_n}) = \mathbb{P}^{X_{T_n}}(t_0 \in [t + T_n, t + T_n + dt[, x_0 = x, \delta_0 = \delta).$$

Ensuite, pour déterminer quand, où et de quel type est le premier évènement, il suffit d'utiliser le "Lemme des réveils".

La dernière affirmation est évidente. □

Nous connaissons maintenant une façon de construire un processus de particules de naissance et de mort. Il nous suffit de construire cet ensemble continu de variables aléatoires, les unes après les autres.

Dérivée de Radon-Nikodým des processus de particules à temps continu

Proposition 2.2. *Soit \mathbb{P} la loi d'un système de particules de taux de naissance b_t , de taux de mort k_t , et \mathbb{Q} la loi d'une modification de ce processus de particules avec taux $b_t(1 + \alpha_t)$ et $k_t(1 + \beta_t)$, où α, β sont des fonctions continues bornées sur $\mathbb{R}^+ \times \mathbb{Z}^d$. La dérivée de Radon-Nikodým du processus jusqu'à l'instant t peut s'écrire :*

$$\frac{d\mathbb{Q}}{d\mathbb{P}} \Big|_{\mathcal{F}_t} (X) = \exp\left\{-\int_0^t X_s(P(b_s\alpha_s) + k_s\beta_s) ds\right\} \prod_{n|T_n < t} (1 + \mathbf{1}_{\{\delta_n=1\}}\alpha_{T_{n+1}-}(x_n) + \mathbf{1}_{\{\delta_n=-1\}}\beta_{T_{n+1}-}(x_n)).$$

Preuve. Nous utilisons pour commencer la Proposition 2.1, pour calculer la dérivée de Radon-Nikodým L_n de $(X_{t \wedge T_n}, t \geq 0)$ sous les lois \mathbb{P} et \mathbb{Q} :

$$\begin{aligned} L_n &= \exp \left\{ \int_0^{T_n} X_s (P(b_s \alpha_s) + k_s \beta_s) ds \right\} \\ &\quad \times \prod_{k=0}^{n-1} (1 + \mathbf{1}_{\{\delta_k=1\}} \alpha_{T_{k+1}-}(x_k) + \mathbf{1}_{\{\delta_k=-1\}} \beta_{T_{k+1}-}(x_k)). \end{aligned}$$

Nous pouvons étendre cette égalité de la manière suivante, soit $n_t = \inf\{n > 0 | T_n > t\}$, nous pouvons calculer la dérivée de Radon-Nikodým jusqu'à l'instant t , de la manière suivante :

$$\begin{aligned} \mathbb{E} \left(\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_t} \middle| \mathcal{F}_{T_n} \right) &= \mathbb{E} \left(\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_t} \mathbf{1}_{\{n \leq n_t\}} \middle| \mathcal{F}_{T_n} \right) \\ &\quad + \mathbb{E} \left(\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_t} \mathbf{1}_{\{n > n_t\}} \middle| \mathcal{F}_{T_n} \right) \\ &= L_n \mathbf{1}_{\{n \leq n_t\}} + L_{n_t-1} \frac{\mathbb{Q}(T_{n_t} > t)}{\mathbb{P}(T_{n_t} > t)} \mathbf{1}_{\{n > n_t\}}. \end{aligned}$$

En effet, l'information que nous obtenons sur le processus entre les instants T_{n_t-1} et t est uniquement que $T_{n_t} > t$. Il suffit maintenant de faire tendre $n \rightarrow +\infty$ pour conclure.

$$\begin{aligned} \left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_t} (X) &= \exp \left\{ - \int_0^t X_s (P(b_s \alpha_s) + k_s \beta_s) ds \right\} \\ &\quad \prod_{n | T_n < t} (1 + \mathbf{1}_{\{\delta_n=1\}} \alpha_{T_{n+1}-}(x_n) + \mathbf{1}_{\{\delta_n=-1\}} \beta_{T_{n+1}-}(x_n)). \end{aligned}$$

□

Nous allons maintenant écrire cette dérivée de façon plus compacte. Dans ce but, nous allons rappeler l'expression de l'exponentielle d'une martingale càdlàg.

Définition 2.1. Soit M_t une martingale càdlàg (continue à droite avec des limites à gauche), l'exponentielle de M est la martingale notée $\mathcal{E}(M)$, définie par :

$$\mathcal{E}(M)_t = \exp \left(M_t - \frac{1}{2} [M, M]_t^c \right) \prod_{s \leq t} (1 + \Delta M_s) \exp(-\Delta M_s).$$

Si M est une martingale purement discontinue, alors on a :

$$\mathcal{E}(M)_t = \exp \left(M_t - \sum_{s \leq t} \Delta M_s \right) \prod_{s \leq t} (1 + \Delta M_s).$$

Remarque 2.1. La martingale $\mathcal{E}(M)$ est l'unique solution de l'équation différentielle stochastique :

$$Z_t = 1 + \int_0^t Z_s dM_s.$$

Soit (X_t) un processus de particules de loi \mathbb{P} , nous pouvons définir les processus de sauts suivants :

- Le processus de naissance $B_t(x) = \sum_{0 < s \leq t} \Delta X_s(x)^+$;
- Le processus de mort $K_t(x) = \sum_{0 < s \leq t} \Delta X_s(x)^-$.

Bien entendu, nous avons $X_t(x) - X_0(x) = B_t(x) - K_t(x)$. De plus, la propriété de Markov montre que le processus :

$$\widehat{B}_t(x) = B_t(x) - \int_0^t b_s(x) P(X_s)(x) ds \text{ et}$$

$$\widehat{K}_t(x) = K_t(x) - \int_0^t k_s(x) X_s(x) ds$$

sont des \mathbb{P} -martingales.

Dès lors, on a :

$$\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_t} = \mathcal{E}(M)_t,$$

où on pose :

$$M_t = \sum_x (\alpha(x) \cdot \widehat{B}(x))_t + (\beta(x) \cdot \widehat{K}(x))_t.$$

2.2 Dérivée de Radon-Nikodým pour les processus de particules à temps discret

On note \mathbb{P} la loi d'un processus de particules à temps discret de distribution du nombre d'enfants Π , et \mathbb{Q} une modification de ce processus de particules avec distribution Π' .

Ces dérivées de Radon-Nikodým sont plus faciles à calculer, en utilisant l'indépendance du nombre d'enfants à l'instant $n+1$ sur chaque site, conditionnellement à \mathcal{F}_n . La dérivée de Radon-Nikodým peut alors s'exprimer comme un produit sur le temps et l'espace :

$$\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_N} (X) = \prod_{n=0}^N \prod_x L_{n,x},$$

où on a noté :

$$L_{n,x} = \frac{\mathbb{Q}(\xi_{n+1}(x) = X_{n+1}(x))}{\mathbb{P}(\xi_{n+1}(x) = X_{n+1}(x))}.$$

En utilisant les définitions que nous avons donné des processus de particules à temps discret, nous pouvons écrire :

$$L_{n,x}(X) = \frac{\Pi'_{n,x}(X_{n+1})}{\Pi_{n,x}(X_{n+1})}.$$

3 Quelques calculs relatif à la marche aléatoire branchante

Comme nous l'avons dit dans l'introduction, nous allons utiliser une marche aléatoire branchante comme loi de référence pour presque tous nos systèmes de particules. En d'autres termes, pour étudier la convergence des dérivées de Radon-Nikodým, nous étudierons des fonctionnelles de marches aléatoires branchantes (à l'exception du modèle de Lotka-Volterra, pour lequel il sera plus simple d'utiliser le modèle du voteur comme modèle de référence). Par conséquent, quelques résultats sur les méthodes de calcul des moments d'une marche aléatoire branchantes pourront être utiles par la suite.

3.1 Représentation de la marche aléatoire branchante à temps continu

Soient $\Lambda^n(x, y)$ et $\Lambda^n(x)$ des processus de Poisson continus d'intensité $\gamma p(x, y)$ et γ respectivement. Une marche aléatoire branchante ξ_t à taux γ est la seule solution forte du problème suivant :

$$\xi_t(x) = \xi_0(x) + \sum_{y,n} \int_0^t \mathbf{1}_{\{\xi_{s-}(y) > n\}} d\Lambda_s^n(x, y) - \sum_n \int_0^t \mathbf{1}_{\{\xi_{s-}(x) > n\}} d\Lambda_s^n(x).$$

Si Λ est un processus de Poisson d'intensité λ , nous noterons son processus compensé $\widehat{\Lambda}_t = \Lambda_t - \lambda t$, et pour toutes fonctions ϕ , on pose :

$$\xi_t(\phi) = \sum_x \phi(x) \xi_t(x) \text{ et}$$

$$\mathcal{L}_\gamma \phi(x) = \gamma(\mathbb{P}\phi(x) - \phi(x)).$$

Nous pouvons donc réécrire ξ de la façon suivante :

$$\xi_t(\phi) = \xi_0(\phi) + M_t(\phi) + A_t(\phi) \tag{5.2}$$

où nous avons posé

$$M_t(\phi) = \left[\sum_{x,y,n} \int_0^t \phi(x) \mathbf{1}_{\{\xi_{s-}(y) > n\}} d\widehat{\Lambda}_s^n(x, y) - \sum_{x,n} \int_0^t \phi(x) \mathbf{1}_{\{\xi_{s-}(x) > n\}} d\widehat{\Lambda}_s^n(x, y) \right] \text{ et}$$

$$A_t(\phi) = \int_0^t \xi_s(\mathcal{L}_\gamma \phi) ds.$$

$M_t(\phi)$ est une martingale, de variation quadratique :

$$\langle M(\phi) \rangle_t = \gamma \int_0^t \xi_s(\phi^2 + P(\phi^2)) ds.$$

3.2 Calcul des moments

Nous voulons étendre l'équation (5.2) à des fonctions $\phi(t, x) = \phi_t(x) \in \mathcal{C}^{1,3}([0, +\infty[\times \mathbb{R}^d)$. D'après l'égalité de Riemann-Stieltjes, on a :

$$\xi_t(x)\phi_t(x) = \xi_0(x)\phi_0(x) + \int_0^t \xi_s \dot{\phi}_s(x) ds + \int_0^t \phi_s(x) d\xi_s(x).$$

Nous pouvons alors sommer sur tout $x \in \mathbb{Z}^d$, pour obtenir :

$$\xi_t(\phi_t) = \xi_0(\phi_0) + M_t(\phi) + A_t(\phi),$$

où on a noté :

$$M_t(\phi) = \left[\sum_{x,y,n} \int_0^t \phi_s(x) \mathbf{1}_{\{\xi_{s-}(y) > n\}} d\widehat{\Lambda}_s^n(x, y) - \sum_{x,n} \int_0^t \phi_s(x) \mathbf{1}_{\{\xi_{s-}(y) > n\}} d\widehat{\Lambda}_s^n(x, y) \right]$$

et

$$A_t(\phi) = \int_0^t \sum_x \xi_s(\mathcal{L}_\gamma \phi_s + \dot{\phi}_s) ds.$$

$M_t(\phi)$ est encore une martingale de variation quadratique :

$$\langle M(\phi) \rangle_t = \gamma \int_0^t \xi_s(\phi_s^2 + P(\phi_s^2)) ds.$$

Nous appelons M la mesure-martingale orthogonale de la marche aléatoire branchante.

Définition 3.1. Soit ξ_t une marche aléatoire à taux γ , la martingale-mesure orthogonale de ξ est le processus M_t à valeurs dans $\mathcal{M}_f(\mathbb{R}^d)$ vérifiant pour tout $\phi \in \mathcal{C}^{1,3}([0, +\infty[\times \mathbb{R}^d)$:

$$M(\phi)_t = \xi_t(\phi) - \xi_0(\phi) - \int_0^t \sum_x \xi_s(\mathcal{L}_\gamma \phi_s + \dot{\phi}_s) ds.$$

Notons P_t le semigroupe de la marche aléatoire continue B^γ sur \mathbb{Z}^d qui saute à taux γ vers l'un de ses voisins selon la distribution p . En d'autres termes,

$$P_t(\phi)(x) = \mathbb{E}(\phi(x + B_t^\gamma)).$$

Si nous appliquons l'équation précédente à la fonction $\phi_s(x) = P_{t-s}\phi(x)$, on a $A_t(\phi) = 0$, dès lors :

$$\mathbb{E}(\xi_t(\phi)) = \xi_0(P_t(\phi)),$$

De plus, nous pouvons ainsi calculer les autres moments par récurrence, en utilisant la formule d'Itô. Ainsi, pour le deuxième moment, nous avons :

$$\mathbb{E}(\xi_t(\phi)\xi_t(\psi)) = \xi_0(P_t(\phi))\xi_0(P_t(\psi)) + \int_0^t \mathbb{E}(\xi_s(P_{t-s}\phi P_{t-s}\psi)) ds.$$

On peut « lire » cette formule de la manière suivante, pour $\phi = \mathbf{1}_{\{x\}}$ et $\psi = \mathbf{1}_{\{y\}}$: la probabilité qu'il y ait une particule en x et une autre en y est égale à la probabilité que deux particules issues de deux particules initiales différentes se retrouvent aux bons endroits, plus celle que la particule en position x et celle en position y descendent d'un même ancêtre commun à l'instant s , puis se soient déplacées de manière indépendante pendant un temps $t - s$ pour se retrouver dans la bonne position.

La formule étendue s'en déduit par linéarité.

3.3 Calculs pour la marche aléatoire branchante à temps discret

Le même genre de calculs peuvent s'étendre aux marches aléatoires branchantes à temps discret. En fait, nous pouvons aisément prouver par récurrence que :

$$\mathbb{E}(\xi_n(\phi)) = \xi_0(P_n(\phi)),$$

$$\mathbb{E}(\xi_n(\phi)\xi_n(\psi)) = \xi_0(P_n(\phi))\xi_0(P_n(\psi)) + \sum_{k=0}^{n-1} \mathbb{E}(\xi_k(P_{n-k}(\phi)P_{n-k}(\psi))),$$

où on a posé $P_n(\phi)(x) = \mathbb{E}(\phi(x + B_n))$, avec B_n une marche aléatoire sur \mathbb{Z}^d de noyau p .

4 Convergence de la marche aléatoire branchante avec dérive

Dans cette section, nous mettons en valeur l'une des difficultés qui peut apparaître lorsqu'on s'intéresse aux marches aléatoires à temps continu, difficulté qui disparaît lorsqu'on s'intéresse aux marches aléatoires branchantes à temps discret. Dans ce but, nous étudierons l'une des modifications les plus simples de la marche aléatoire branchante. Nous ajoutons uniquement une dérive, une différence entre la probabilité de naissance et de mort pour chaque particule.

4.1 Marche aléatoire branchante à temps continu avec dérive symétrique θ

Pour commencer, nous étudions un modèle dans lequel la dérive est distribuée de façon symétrique sur les taux de naissance et de mort. Avec ce modèle, on peut prouver la convergence de la dérivée de Radon-Nikodým, sans difficultés.

Définition 4.1. Une marche aléatoire branchante à temps continu avec dérive symétrique $\theta \in \mathcal{C}_b([0, +\infty[\times \mathbb{Z}^d)$ est un processus de particules dans lequel chaque particule a un taux de naissance accéléré et un taux de mort ralenti du même facteur. Ce système de particules évolue de la façon suivante :

$$\begin{cases} \xi_t(x) \rightarrow \xi_t(x) + 1 \text{ à taux } V_t(x)(1 + \theta(x, t)) \\ \xi_t(x) \rightarrow \xi_t(x) - 1 \text{ à taux } \xi_t(x)(1 - \theta(x, t)) \end{cases}$$

Une suite de tels processus, convenablement renormalisée, convergera vers un super-processus de Dawson-Watanabe de la manière suivante. On note θ^N une suite de fonctions continues bornées sur $\mathbb{R}^+ \times \frac{\mathbb{Z}^d}{\sqrt{N}}$, que nous étendons par interpolation à des fonctions $\mathcal{C}_b(\mathbb{R}^+ \times \mathbb{R}^d)$. Nous supposons que θ^N converge uniformément vers θ .

Théorème 4.1. Soit ξ_t^N une suite de marches aléatoires branchantes à taux N avec dérivée symétrique $\frac{\theta^N}{N}$ sur le réseau renormalisé d'un facteur \sqrt{N} , on pose :

$$X_t^N = \frac{1}{N} \sum_{x \in \mathbb{Z}^d} \xi_t^N(x) \delta_{\frac{x}{\sqrt{N}}}.$$

Si $X_0^N \xrightarrow[N \rightarrow +\infty]{} \mu$, où $\mu \in \mathcal{M}_f(\mathbb{R}^d)$, alors nous avons :

$$X^N \xrightarrow[N \rightarrow +\infty]{} X,$$

où X est un super-processus de Dawson-Watanabe avec taux de branchement 2, diffusion σ^2 et dérivée θ .

Preuve. Soit P_θ la loi du processus X^N , et P^N la loi de la marche aléatoire branchante à taux N , renormalisée de la même manière, que nous noterons aussi X^N . La dérivée de Radon-Nikodým pour ces processus renormalisés est égal à celle des systèmes de particules dont ils sont tirés :

$$\left. \frac{dP_\theta^N}{dP^N} \right|_{\mathcal{F}_t} = \mathcal{E} \left(\frac{1}{N} M^N(\theta^N) \right)_t,$$

où M^N est la mesure-martingale orthogonale de la marche aléatoire branchante ξ^N .

Nous utilisons des approximations $\phi_\varepsilon^N \in \mathcal{C}_b^{1,3}(\mathbb{R}^+ \times \mathbb{R}^d)$ de θ^N vérifiant $\|\phi_\varepsilon^N - \theta^N\|_\infty \leq \varepsilon$. De la même façon, on note ϕ_ε une approximation de θ .

De plus, nous savons que la séquence de marches aléatoires branchantes à taux N , renormalisées en espace par \sqrt{N} et en masse par N , converge en loi vers le super-mouvement Brownien. Dès lors, nous avons simplement à prouver la convergence en loi jointe avec la dérivée de Radon-Nikodým. On a, pour toute fonction $\phi \in \mathcal{C}_b^{1,3}(\mathbb{R}^+ \times \mathbb{R}^d)$:

$$M_t^N(\phi) = \xi_t^N(\phi) - \xi_0^N(\phi) - \int_0^t \xi_s^N(\mathcal{L}_N \phi) ds - \int_0^t \xi_s^N(\dot{\phi}_s) ds.$$

De la même manière, on a :

$$[M^N(\phi)]_t = \sum_{x,y,n} \int_0^t \phi_{s-}(x)^2 \mathbf{1}_{\{\xi_{s-}^N(y) > n\}} d\widehat{\Lambda}_s^n(x,y) + \sum_{x,n} \phi_{s-}(x)^2 \mathbf{1}_{\{\xi_{s-}^N(x) > n\}} d\widehat{\Lambda}_s^n(x),$$

on voit donc que $[\frac{1}{N} M^N(\phi)]_t - \int_0^t X_s^N(\phi_s^2 + P^N(\phi_s^2)) ds$ est une martingale de variation quadratique tendant vers 0, donc

$$[\frac{1}{N} M^N(\phi)]_t \xrightarrow[n \rightarrow +\infty]{L^2} 2 \int_0^t X_s(\phi_s^2) ds.$$

De plus, la fonction :

$$X \in D([0, +\infty[, \mathcal{M}_f(\mathbb{R}^d)) \mapsto \int_0^t X_s(\phi) ds$$

est continue, donc en posant ϕ^n un ensemble dénombrable de fonctions dans $\mathcal{C}_b^{1,3}(\mathbb{R}^+ \times \mathbb{R}^d)$, on a :

$$\left(X^N, \left(\frac{1}{N} M_t^N(\phi^n) \right), \left(\frac{1}{2N^2} [M^N(\phi^n)]_t \right) \right) \xrightarrow{n \rightarrow +\infty} \left(X, (M_t(\phi^n)), \int_0^t X_s(\phi_s^2) ds \right).$$

Nous utilisons ensuite ce résultat pour l'ensemble dénombrable $(\phi_\varepsilon^N)_{N \in \mathbb{N}, \varepsilon \in \mathbb{Q}}$ des approximations de θ^N , ce qui nous permet de prouver la convergence jointe :

$$\left(X^N, \frac{1}{N} M_t^N(\theta^N), \frac{1}{2N^2} [M^N(\theta^N)]_t \right) \xrightarrow{n \rightarrow +\infty} \left(X, M_t(\theta), \int_0^t X_s(\theta_s^2) ds \right).$$

Cette convergence jointe nous donne en particulier la convergence jointe de X^N et de la dérivée de Radon-Nikodým, on en conclut la convergence de P_θ^N vers P^{2,θ,σ^2} . \square

Étudions maintenant le même genre de convergence pour des processus à temps discret.

4.2 Marche aléatoire branchante avec dérive, modèle à temps discret

Définition 4.2. Une marche aléatoire branchante à temps discret avec dérive $\theta \in \mathcal{C}_b(\mathbb{R}^+ \times \mathbb{R}^d)$ est un système de particules dans lequel chaque particule a un nombre d'enfants augmenté en moyenne d'un facteur θ . Ce processus de particule évolue de la manière suivante :

$$\xi_{n+1}(x) = k \text{ avec probabilité } e^{-\lambda} \frac{\lambda^k}{k!}$$

où on a posé $\lambda = \lambda_n(x) = (1 + \theta_n(x))V_n(x)$.

Comme dans le cas précédent, on considère une suite θ^N de fonctions continues bornées sur $\frac{\mathbb{N}}{N} \times \frac{\mathbb{Z}^d}{\sqrt{N}}$, que nous étendons par interpolation à des fonctions de $\mathcal{C}_b(\mathbb{R}^+ \times \mathbb{R}^d)$. On suppose que θ^N converge uniformément vers θ . On peut alors prouver le même type de résultats :

Théorème 4.2. Soit ξ^N une suite de marches aléatoires branchantes à temps discret sur la matrice renormalisée, avec dérive $\frac{\theta_n^N}{N}$. On pose :

$$X_t^N = \frac{1}{N} \sum_{x \in \mathbb{Z}^d} \xi_{[Nt]}^N(x) \delta_{\frac{x}{\sqrt{N}}}.$$

Si $X_0^N \xrightarrow{n \rightarrow +\infty} \mu$ où $\mu \in \mathcal{M}_f(\mathbb{R}^d)$, alors on a :

$$X^N \xrightarrow{n \rightarrow +\infty} X,$$

où X est un super-processus de Dawson-Watanabe avec taux de branchement 1, diffusion σ^2 et dérive θ .

Preuve. Nous supposons dans un premier temps que les dérivées θ^N et θ sont dans $\mathcal{C}_b^{1,3}(\mathbb{R}^+ \times \mathbb{R})$, et que la convergence est uniforme en espace.

Comme précédemment, nous étudions la dérivée de Radon-Nikodým de nos processus. On pose \mathbb{P}_θ^N la loi de $(X_t^N)_{t \geq 0}$ et \mathbb{P}^N la loi de la marche aléatoire branchante sans dérive, renormalisée de la même manière. Dans le modèle à temps discret, on a :

$$\frac{dP_\theta^N}{dP^N} \Big|_{\mathcal{F}_t} = \prod_{k=0}^{\lfloor Nt \rfloor} \prod_x L_{n,x},$$

où on a posé :

$$L_{n,x} = e^{-\frac{1}{N} V_n^N(x) \theta_{\frac{n}{N}}^N(x)} \left(1 + \frac{\theta_{\frac{n}{N}}^N(x)}{N} \right)^{\xi_{n+1}^N(x)}.$$

On peut maintenant réécrire

$$\frac{dP_\theta^N}{dP^N} \Big|_{\mathcal{F}_t} = (1 + o_P(1)) \exp \left[\frac{1}{N} \sum_{k=0}^{\lfloor Nt \rfloor} \xi_{n+1}^N \left(\theta_{\frac{n}{N}}^N \right) - \xi_n^N(P\theta_{\frac{n}{N}}^N) + \frac{1}{N^2} \sum_{k=0}^{\lfloor Nt \rfloor} \xi_{n+1}^N(\theta_{\frac{n}{N}}^{N^2}) \right].$$

On étudie la convergence de chaque partie de la dérivée de Radon-Nikodým, on commence par

$$\frac{1}{N^2} \sum_{k=0}^{\lfloor Nt \rfloor} \xi_{n+1}^N \left(\theta_{\frac{n}{N}}^{N^2} \right) = \int_0^t X_s^N \left(\theta_s^{N^2} \right) ds + o(1),$$

et de la même manière que précédemment, nous obtenons la convergence jointe avec celle du processus.

Nous nous intéressons ensuite à :

$$\begin{aligned} & \frac{1}{N} \sum_{k=0}^{\lfloor Nt \rfloor} \xi_{n+1}^N \left(\theta_{\frac{n}{N}}^N \right) - \xi_n^N(P\theta_{\frac{n}{N}}^N) \\ &= \frac{1}{N} \sum_{k=0}^{\lfloor Nt \rfloor} \left[\xi_{n+1}^N \left(\theta_{\frac{n+1}{N}}^N \right) - \xi_n^N \left(\theta_{\frac{n}{N}}^N \right) \right] - \left[\xi_{n+1}^N \left(\theta_{\frac{n+1}{N}}^N \right) - \xi_{n+1}^N \left(\theta_{\frac{n}{N}}^N \right) \right] - \left[\xi_n^N(P\theta_{\frac{n}{N}}^N) - \xi_n^N \left(\theta_{\frac{n}{N}}^N \right) \right] \\ &= X_{\lfloor Nt \rfloor}^N \left(\theta_{\frac{\lfloor Nt \rfloor}{N}}^N \right) - X_0^N \left(\theta_0^N \right) - \frac{1}{N^2} \sum_{k=0}^{\lfloor Nt \rfloor} \xi_{n+1}^N \left(\dot{\theta}_{\frac{n}{N}}^N \right) + \xi_n^N(\mathcal{L}_N \theta_{\frac{n}{N}}^N) + o(1) \\ &= X_{\lfloor Nt \rfloor}^N \left(\theta_{\frac{\lfloor Nt \rfloor}{N}}^N \right) - X_0^N \left(\theta_0^N \right) - \int_0^t X_s^N \left(\dot{\theta}_s^N + \mathcal{L}_N \theta_s^N \right) ds + o(1). \end{aligned}$$

Ainsi on obtient la loi jointe du processus et de sa dérivée de Radon-Nikodým.

Nous utilisons ensuite une suite d'approximations de θ par des fonctions régulières pour compléter la preuve. \square

Remarque 4.1. Cette preuve de la convergence du processus à temps discret est très proche de celle de la convergence de la marche aléatoire à dérive symétrique. Ici l'espérance du nombre d'enfants de chaque particule est égal à $\frac{1+\theta_t(x)}{2}$, ce qui nous amène à considérer une dérive $\frac{\theta}{2}$ pour le processus à temps discret. De plus, le taux de branchement du processus limite vaut 2 dans le cas de la marche aléatoire branchante à dérive symétrique, au lieu de 1 dans le modèle à temps discret.

Mais s'intéresser à $\tilde{X}_t^N = \frac{1}{N} \sum_{x \in \mathbb{Z}^d} \xi_{[2Nt]}^N(x) \delta_{\frac{x}{\sqrt{N}}}$ nous aurait conduit à un super-mouvement Brownien, avec taux de branchement 2 et dérive θ , ainsi qu'on l'attendait.

Nous allons maintenant nous intéresser à une autre marche aléatoire branchante à temps continu, pour laquelle la dérive n'est pas distribuée de façon égale sur les taux de naissance et de mort. Prouver la convergence grâce à la dérivée de Radon-Nikodým du processus devient alors un exercice bien plus ardu.

4.3 Une marche aléatoire branchante avec dérive asymétrique (θ, θ')

Définition 4.3. Une marche aléatoire avec dérive asymétrique $(\theta, \theta') \in \mathcal{C}_b^{1,3}(\mathbb{R}^+ \times \mathbb{R}^d)^2$ est une marche aléatoire branchante dont on a accéléré le taux de naissance par un certain facteur, et le taux de décès par un autre. Le système de particules évolue de la façon suivante :

$$\begin{cases} \xi_t(x) \rightarrow \xi_t(x) + 1 \text{ à taux } V_t(x)(1 + \theta_t(x)) \\ \xi_t(x) \rightarrow \xi_t(x) - 1 \text{ à taux } \xi_t(x)(1 + \theta'_t(x)) \end{cases}$$

Le théorème que nous aimerions prouver est le suivant : soit (θ^N, θ'^N) des suites fonctions continues bornées sur $\mathbb{R}^+ \times \frac{\mathbb{Z}^d}{\sqrt{N}}$, que nous étendons par interpolation à des fonctions dans $\mathcal{C}_b(\mathbb{R}^+ \times \mathbb{R}^d)$. Nous supposons que θ^N et θ'^N convergent uniformément vers θ et θ' .

Théorème 4.3. Soit ξ_t^N une suite de marches aléatoires branchantes à taux N avec dérive asymétrique $\left(\frac{\theta^N}{N}, \frac{\theta'^N}{N}\right)$, on pose :

$$X_t^N = \frac{1}{N} \sum_{x \in \mathbb{Z}^d} \xi_t^N(x) \delta_{\frac{x}{\sqrt{N}}}.$$

Si $X_0^N \xrightarrow[N \rightarrow +\infty]{} \mu$ où $\mu \in \mathcal{M}_f(\mathbb{R}^d)$, alors on a :

$$X^N \xrightarrow[N \rightarrow +\infty]{} X,$$

où X est un super-processus de Dawson-Watanabe avec taux de branchement 2, diffusion σ^2 et dérive $\frac{\theta - \theta'}{2}$.

Nous voudrions prouver ce théorème de la même façon que précédemment.

Comme nous l'avions fait précédemment, nous notons $P_{\theta, \theta'}^N$ la loi du processus ξ^N , et rappelons que \mathbb{P}^N est la loi de la marche aléatoire branchante à taux N . La dérivée de Radon-Nikodým de ce processus est

$$\left. \frac{dP_{\theta}^N}{d\mathbb{P}^N} \right|_{\mathcal{F}_t} = \mathcal{E}(B^N(\theta^N) - K^N(\theta'^N))_t,$$

où $B^N(\theta^N)_t = \sum_{0 < s \leq t} \sum_x \theta_s^N(x) \Delta \xi_s^N(x)^+$, est défini comme le processus de naissance, qui saute de $\theta_s(x)$ s'il y a une naissance sur le site x à l'instant s , et de la même façon, on note $K^N(\theta'^N)_t = \sum_{0 < s \leq t} \sum_x \theta_s^N(x) \Delta \xi_s^N(x)^-$.

On s'intéresse maintenant à la convergence en loi de cette dérivée de Radon-Nikodým.

Commençons par rappeler que

$$B_t^N(\theta^N) = \sum_{x, y, n} \int_0^t \theta_{s-}^N(x) \mathbf{1}_{\{\xi_{s-}^N(y) > n\}} d\widehat{\Lambda}_s^n(x, y).$$

De plus, nous avons :

$$\begin{aligned} [B^N(\theta^N)]_t &= \sum_{x, y, n} \int_0^t \theta_{s-}^N(x)^2 \mathbf{1}_{\{\xi_{s-}^N(y) > n\}} d\Lambda_s^n(x, y) \\ &= \frac{1}{2} [M^N(\theta^N)]_t + \frac{1}{2} M_t^N(\theta^{N^2}). \end{aligned}$$

Nous voyons donc que la variation quadratique est assez simple à étudier, en utilisant les mêmes arguments que précédemment. Il nous faut maintenant prouver la convergence en loi de $\frac{1}{N} B_t^N(\theta)$ vers $\frac{1}{2} M_t(\theta)$. Mais cette partie semble assez problématique, en raison des différences importantes existant entre B^N et K^N les processus de vie et de mort.

Remarque 4.2. Afin de donner un équivalent du processus à dérive asymétrique, nous devons considérer la suite de processus modifiée à temps discrets suivante $Y_t^N = X_{t + \frac{\theta_t(x)}{N}}^N$, car le processus à temps continu évolue à un taux $2 + \frac{\theta_t(x)}{N}$, et conserver le même type de dérive que dans le processus X^N . Il apparait ainsi que les difficultés de calculs proviennent d'un problème de temps, au lieu de conserver la même échelle de temps, dans le cas du processus asymétrique le temps est accéléré d'un facteur $(1 + \frac{\theta - \theta'}{2})$ à l'ordre 0, et cette modification nous amène à des effets difficiles à contrôler pour la dérivée de Radon-Nikodým.

5 Convergence du processus de contact

Un processus de contact est un système de particules de vie et de mort, qui représente l'évolution d'une épidémie dans une population. Les « particules » représentent les individus infectés. En chaque site, il y a uniquement un nombre fini de personnes qui sont susceptibles de tomber malade. Chaque personne infectée peut contaminer une autre voisine à un certain taux, et guérir à un autre taux. Ces dynamiques ont déjà été beaucoup étudié, et la convergence a été prouvée dans [6], mais nous cherchons ici à étudier la possibilité d'utiliser une dérivée de Radon-Nikodým.

5.1 Processus de contact à temps continu

Le processus de contact à temps continu peut être défini comme suit : On suppose qu'en chaque site du réseau \mathbb{Z}^d se situe un village de M individus. Chaque individu infecté peut contaminer un individu d'un village voisin à taux $\frac{1}{M}p(x, y)$. La cible est effectivement contaminée si elle n'était pas déjà malade. Si nous comparons ceci à une marche aléatoire branchante, nous remarquons que le taux de naissance est moins élevé, donc nous avons également besoin de modifier le taux de mort, pour éviter l'écueil de la dérive asymétrique. Chaque particule guérira à un taux $1 + \theta$. Formellement, nous posons la définition suivante.

Définition 5.1. Un processus de contact avec dérive θ et taille de village M est un processus de particules à temps continu avec taux de naissance $(1 - V_{t-}(x))$ et taux de mort $(1 + \theta)$, c'est-à-dire un processus de Markov évoluant de la manière suivante :

$$\begin{cases} \xi_t(x) \rightarrow \xi_t(x) + 1 & \text{at rate } V_t(x)(1 - \frac{X_t(x)}{M}) \\ \xi_t(x) \rightarrow \xi_t(x) - 1 & \text{at rate } \xi_t(x)(1 + \theta). \end{cases}$$

Nous pouvons obtenir le théorème suivant sur la convergence d'une suite de processus de contact.

Théorème 5.1. Soit ξ^N une suite de processus de contact à taux N , dérive $\frac{\theta}{N}$ et taille de village N , on pose :

$$X_t^N = \frac{1}{N} \sum_{x \in \mathbb{Z}^d} \xi_t^N(x) \delta_{\frac{x}{\sqrt{N}}}.$$

Si $X_0^N \xrightarrow[N \rightarrow +\infty]{} \mu$, où μ est une mesure finie sans atomes sur \mathbb{R}^d , nous avons :

$$X^N \xrightarrow[N \rightarrow +\infty]{} X,$$

où X est un super-processus de Dawson-Watanabe avec taux de branchement 2, diffusion σ^2 et dérive $\theta - b$, où on a posé :

$$b = \sum_{n=0}^{+\infty} \mathbb{E}(p(B_n)) = \sum_{n=1}^{+\infty} \mathbb{P}(B_n = 0),$$

avec B_n une marche aléatoire sur \mathbb{Z}^d de noyau p .

Ce théorème a déjà été prouvé dans [6], mais d'une façon différente. On prouve pour commencer l'existence de valeurs d'adhérence par un argument de tension, puis on montre que toute valeur d'adhérence est solution du problème de martingales (5.1). Le but ici serait de donner une autre preuve plus directe, en utilisant la dérivée de Radon-Nikodým.

Si nous notons Q_θ^N la loi du processus X^N , nous avons encore, grâce à la Proposition 2.2 :

$$\left. \frac{dQ_\theta^N}{dP^N} \right|_{\mathcal{F}_t} = \mathcal{E} \left(\frac{1}{N} \tilde{M}^N \right)_t,$$

où :

$$\tilde{M}_t^N = \sum_{x,y,n} \int_0^t \xi_s^N(x) \mathbf{1}_{\{\xi_s^N(y) > n\}} d\hat{\Lambda}_s^n(x,y) - \sum_{x,n} \int_0^t \theta \mathbf{1}_{\{\xi_s^N(x) > n\}} d\hat{\Lambda}_s^n(x).$$

Nous allons maintenant traiter le cas $\theta = b$, pour lequel nous souhaiterions prouver que la dérivée de Radon-Nikodým converge en loi vers 1.

Mais si nous avons cette convergence, nous pourrions calculer l'espérance du logarithme de cette variable aléatoire :

$$\begin{aligned} \frac{1}{N^2} \mathbb{E}([\tilde{M}^N]_t) &= \mathbb{E} \left(\frac{1}{N^2} \sum_{x,y,n} \int_0^t \xi_s^N(x)^2 \mathbf{1}_{\{\xi_s^N(y) > n\}} d\Lambda_s^n(x,y) - \frac{1}{N^2} \sum_{x,n} \int_0^t b^2 \mathbf{1}_{\{\xi_s^N(x) > n\}} d\Lambda_s^n(x) \right) \\ &= \int_0^t \frac{1}{N} \sum_{x,y} \mathbb{E}((\xi_s^N(x)^2 - b^2) \xi_s^N(y)) p(x,y) ds \xrightarrow{n \rightarrow +\infty} 0. \end{aligned}$$

Mais nous pouvons prouver que (calculs portés en annexe) :

$$\mathbb{E} \left(\int_0^t \frac{1}{N} \sum_{x,y} (\xi_s^N(x) - b) \xi_s^N(y) p(x,y) ds \right) \xrightarrow{n \rightarrow +\infty} 0. \quad (5.3)$$

Ce phénomène est appelé simplification en champ moyen. Il apparait qu'au voisinage de chaque particule, le nombre moyen de voisin tend vers la constante b .

Ces calculs nous mènent aux résultat suivant : si la dérivée de Radon-Nikodým converge vers 1 en loi, alors on a :

$$\mathbb{E} \left(\int_0^t \frac{1}{N} \sum_{x,y} (\xi_t^N(x) - b)^2 \xi_t^N(y) p(x,y) ds \right) \xrightarrow{n \rightarrow +\infty} 0.$$

Mais comme b n'est pas entier, nous pouvons choisir $\varepsilon > 0$ tel que $]b - \varepsilon, b + \varepsilon[\cap \mathbb{N} = \emptyset$, dans ce cas, nous avons également :

$$\mathbb{E} \left(\int_0^t \frac{1}{N} \sum_{x,y} (\xi_t^N(x) - b)^2 \xi_t^N(y) p(x,y) ds \right) \geq \mathbb{E} \left(\int_0^t \frac{1}{N} \sum_{x,y} \varepsilon^2 \xi_t^N(y) p(x,y) ds \right) = t X_0^N(\varepsilon^2).$$

Il apparait donc que la dérivée de Radon-Nikodým ne peut converger dans ce cas précis. La simplification en champ moyen n'apparait pas dans ce type de calculs.

Remarque 5.1. Même si nous avons nous ayons "symétrisé" le processus de contact en accélérant le taux de guérison des particules côtoyant de nombreux malades, il semble difficile de terminer une telle preuve de la convergence du processus de contact, car la « fonction » X_t^N que nous intégrons par rapport à la mesure-martingale M n'a pas de limite en dimension $d > 2$, en particulier cette limite ne peut pas être b .

Il est impossible d'obtenir une convergence séparée (la mesure martingale d'une part, et la fonction intégrée d'autre part, comme réalisé dans [8]). De plus, nous ne pouvons plus utiliser la simplification en champ moyen de notre processus, car la dérivée de Radon-Nikodým conserve les carrés de variations microscopiques dans la dérive.

5.2 Processus de contact à temps discret

Nous souhaitons maintenant voir quelles seront les difficultés lorsque nous traiterons du processus de contact à temps discret. Commençons par définir la distribution du nombre d'enfant. Notons qu'il y a deux types possibles d'enfants d'une marche aléatoire branchante qui disparaissent dans le processus de contact : ceux qui représentent l'infection d'enfants déjà infectés, ce qui se produit avec probabilité $\frac{\xi_n(x)}{M}$; et ceux qui visent plusieurs fois le même individu non-infecté. Il y aura dans ce cas un seul nouvel individu dans le processus de contact, là où il y en avait plusieurs dans le cadre de la marche aléatoire branchante. Pour le moment, nous ignorerons ce second terme.

Un calcul assez simple nous montre que la nouvelle distribution d'enfants doit être :

$$\Pi_{n,x}(k) = e^{-(1-\frac{\xi_n(x)}{M})V_n(x)} \frac{\left((1 - \frac{\xi_n(x)}{M})V_n(x) \right)^k}{k!}.$$

Nous pouvons maintenant définir ce processus de contact modifié à temps discret.

Définition 5.2. Un processus de contact à temps discret modifié avec dérive θ et taille de village M est un processus de particules à temps discret dont le nombre d'enfants suit la distribution suivante :

$$\Pi_{n,x}(k) = e^{-\lambda} \frac{\lambda^k}{k!},$$

où $\lambda = \lambda_n(x) = (1 - \frac{\xi_n(x)}{M})V_n(x)(1 + \theta)$.

Remarque 5.2. Ce processus de contact modifié est une bonne approximation du modèle initial. En effet le nombre de particules par site dans le processus de contact est dominé par le nombre de particules par site de la marche aléatoire branchante avec dérive θ associée. De plus, l'espérance conditionnelle du nombre de particules infectées par deux voisines ou plus à l'instant $n + 1$ au site x est clairement borné par une constante multipliée par le carré du nombre de voisins de cette particule. Finalement, un calcul assez simple montre que le nombre de ces erreurs jusqu'à l'instant t est un $o_P(N)$, quand $N \rightarrow +\infty$, donc ces erreurs n'apparaissent pas dans le processus limite.

Nous étudions la possible convergence d'une suite ξ^N de processus de contact modifiés avec dérive $\frac{\theta}{N}$ et taille de village N . Nous notons également, de la même manière que pour la marche aléatoire branchante avec dérive :

$$X_t^N = \frac{1}{N} \sum_{x \in \mathbb{Z}^d} \xi_{[Nt]}^N(x) \delta_{\frac{x}{\sqrt{N}}}.$$

Nous pouvons calculer la dérivée de Radon-Nikodým de ces processus par rapport aux marches aléatoires branchantes. Nous commençons en calculant

$$L_{n,x} = \exp \left[-V_n(x) \left((1 - \frac{\xi_n^N(x)}{N}) \left(1 + \frac{\theta}{N} \right) - 1 \right) \right] \left((1 - \frac{\xi_n^N(x)}{N}) \left(1 + \frac{\theta}{N} \right) \right)^{\xi_{n+1}(x)}.$$

Nous pouvons aisément calculer le logarithme de la dérivée de Radon-Nikodým de ce processus de contact modifié contre la marche aléatoire branchante :

$$\begin{aligned} \ln(L_t^N) &= \frac{1}{N} \sum_{n=0}^{Nt} \sum_x (\xi_{n+1}^N(x) - V_n(x))(\theta - \xi_n^N(x)) \\ &\quad - \frac{1}{2N^2} \sum_{n=0}^{Nt} \sum_x \xi_{n+1}^N(x)(\xi_n^N(x)^2 + \theta^2) - 2V_n(x)\xi_n^N(x)\theta + o_P(1). \end{aligned}$$

La convergence de la part de cette dérive qui implique le terme en θ est déjà bien connue. Il nous reste juste à tenir compte de la part de dérive reliée à $\xi_n^N(x)$. Nous pouvons établir, de la même manière que pour la marche aléatoire branchante à temps continu que :

$$\mathbb{E} \left(\frac{1}{N} \sum_{n=0}^{\lfloor Nt \rfloor} \frac{(\xi_n^N(x) - \frac{b}{2})V_n(x)}{N} \right) \xrightarrow{n \rightarrow +\infty} 0.$$

Mais la convergence de la dérivée de Radon-Nikodým suppose que :

$$\mathbb{E} \left(\frac{1}{N^2} \sum_{n=0}^{Nt} \sum_x \xi_{n+1}^N(x)(\xi_n^N(x)^2 - \frac{b^2}{4}) \right) = \mathbb{E} \left(\frac{1}{2N} \sum_{n=0}^{Nt} \sum_x V_n(x)(\xi_n^N(x)^2 - \frac{b^2}{4}) \right) \xrightarrow{n \rightarrow +\infty} 0,$$

et de la même manière que précédemment, on peut montrer qu'avoir ces deux convergences de manière simultanée est impossible.

Remarque 5.3. Une fois encore, nous observons que le problème est ici d'obtenir une limite pour l'intégrale de $\xi_n(x)$ contre la martingale orthogonale de la marche aléatoire branchante. Or cette suite de fonctions n'admet pas de limite, ne serait-ce que càdlàg.

En dimension 1, dans [7], comme le super-mouvement Brownien admet une densité continue par rapport à la mesure de Lebesgue, les calculs peuvent être terminés, et on trouve une limite pour la dérivée de Radon-Nikodým conjointe à la limite des suites de marches aléatoires branchantes. Dans [8], ces méthodes sont utilisées sur un modèle de contagion SIR, pour lequel les individus qui ont été infectés sont par la suite immunisés contre la maladie. Comme le super-mouvement Brownien possède une mesure d'occupation à densité continue en dimension 2 et 3, la preuve de la convergence utilisant une dérivée de Radon-Nikodým peut être utilisée, et permet de conclure.

La difficulté dans la preuve de la convergence de la dérivée de Radon-Nikodým est différente de la difficulté rencontrée pour la marche aléatoire avec dérive. Ce n'est pas seulement des différences de vitesse locales qui sont ici en jeu, c'est l'existence d'une simplification en champ moyen, qui n'apparaît pas dans la dérivée de Radon-Nikodým, pour laquelle les processus sont observés avec tous les détails microscopiques. La possibilité de conclure se ramène alors à l'existence d'une limite càdlàg pour la dérive utilisée $\xi_n(x)$, ce qui est faux pour des dimensions plus grandes que 2.

Nous allons maintenant donner quelques autres exemples de processus de particules dont la limite s'obtient par simplification en champ moyen, pour lesquels le calcul de la limite de la dérivée de Radon-Nikodým se révèle inopérant.

6 Convergence pour le modèle du voteur

Nous allons maintenant introduire deux types de modèles de voteur, avec court ou long rayon d'interaction. Ces processus ont déjà été étudiés à de nombreuses reprises, et la convergence d'une suite de ces processus convenablement renormalisés vers des super-mouvements Browniens a été prouvé dans [3].

Un modèle du voteur est un système de particules tel qu'à chaque site, nous avons un certain nombre d'individus. Chacun d'entre eux peut avoir une opinion, 0 ou 1. Avec taux 1, l'un d'entre eux choisi l'un de ses voisins uniformément au hasard, et adopte son opinion. Le processus est donc défini de la manière suivante.

Définition 6.1. Un modèle du voteur à temps continu avec taille de village M est un système de particules $\xi_t : \mathbb{Z}^d \rightarrow \{0, \dots, M\}$ avec taux de naissance $(1 - \frac{\xi_t(x)}{M})$ et taux de décès $(1 - \frac{V_t(x)}{M})$. Par conséquent, ce modèle évolue de la façon suivante :

$$\begin{cases} \xi_t(x) \rightarrow \xi_t(x) + 1 \text{ à taux } V_t(x) \left(1 - \frac{X_t(x)}{M}\right) \\ \xi_t(x) \rightarrow \xi_t(x) - 1 \text{ à taux } \xi_t(x) \left(1 - \frac{V_t(x)}{M}\right). \end{cases}$$

Définition 6.2. Un modèle du voteur à temps discret avec taille de village M est un système de particules $\xi_n : \mathbb{Z}^d \rightarrow \{0, \dots, M\}$ avec une distribution du nombre d'enfants Binomiale $\left(M, \frac{V_n(x)}{M}\right)$. Par conséquent, ce modèle évolue comme suit :

$$\xi_{n+1}(x) = k \text{ avec probabilité } \binom{M}{k} \left(\frac{V_n(x)}{M}\right)^k \left(1 - \frac{V_n(x)}{M}\right)^{M-k}.$$

6.1 Modèle du voteur à long rayon d'interaction

L'étude du modèle du voteur à long rayon d'interaction consiste à l'étude d'une suite de modèles du voteur dans lequel la taille de chaque village tend vers $+\infty$. En d'autres termes, chaque individu possède un nombre de voisins qui tend vers $+\infty$. Soit ξ_t^N une suite de modèles du voteur à temps continu de vitesse N et de taille de village M_N . On pose :

$$X_t^N = \frac{1}{N} \sum_{x \in \mathbb{Z}^d} \xi_t^N(x) \delta_{\frac{x}{\sqrt{N}}}.$$

Le théorème suivant a été prouvé dans [3].

Théorème 6.1. Si $X_0^N \xrightarrow[N \rightarrow +\infty]{} \mu \in \mathcal{M}_f(\mathbb{R}^d)$, alors :

$$X^N \xrightarrow[N \rightarrow +\infty]{} X$$

où X est un super-mouvement Brownien avec taux de branchement 2 et diffusion σ^2 .

Une fois encore, nous calculons la dérivée de Radon-Nikodým de ce processus, de loi R^N :

$$\frac{dR^N}{dP^N} \Big|_{\mathcal{F}_t} = \mathcal{E} \left(\frac{1}{M_N} V_t^N \right),$$

où :

$$V_t^N = \sum_{x,y,n} \int_0^t \xi_s^N(x) \mathbf{1}_{\{\xi_{s-}^N(y) > n\}} d\widehat{\Lambda}_s^n(x,y) + \sum_{x,n} \int_0^t P^N X_s^N(x) \mathbf{1}_{\{\xi_{s-}^N(X) > n\}} d\widehat{\Lambda}_s^n(x,y).$$

Les difficultés que nous avons déjà rencontré dans le processus de contact se retrouvent de la même manière pour le modèle du voteur.

6.2 Autour du modèle du voteur à court rayon d'interaction

Le modèle du voteur à court rayon d'interaction est défini comme un modèle du voteur pour lequel il n'y a qu'un unique individu en chaque site. Dans ce cas, nous avons également convergence vers le super-mouvement Brownien, mais cette fois avec un taux de branchement multiplié par $\gamma = \mathbb{P}(\forall n > 0, B_n \neq 0)$ (voir [3]). Nous pouvons encore calculer la dérivée de Radon-Nikodým, de la même façon que pour le modèle du voteur à long rayon d'interaction, mais cette fois-ci contre une marche aléatoire branchante avec taux γN .

Dans le modèle à temps continu, on a :

$$\frac{dR^N}{dP^N} \Big|_{\mathcal{F}_t} = \mathcal{E}(\tilde{V}_t^N),$$

où on a noté :

$$\begin{aligned} \tilde{V}_t^N &= \gamma \left(\sum_{x,y,n} \int_0^t (\xi_s^N(x) - b) \mathbf{1}_{\{\xi_{s-}^N(y) > n\}} d\widehat{\Lambda}_s^n(x,y) \right. \\ &\quad \left. + \sum_{x,n} \int_0^t (P^N X_s^N(x) - b) \mathbf{1}_{\{\xi_{s-}^N(X) > n\}} d\widehat{\Lambda}_s^n(x,y) \right), \end{aligned}$$

en utilisant la relation $b = \frac{\gamma - 1}{\gamma}$. Ces résultats sont semblables à ceux relatifs aux processus de contact, et conduisent aux mêmes genres de difficultés. Notons que résoudre ce problème permettrait également de finir la preuve pour le modèle du voteur à long rayon d'interaction.

7 Le modèle de Lotka-Volterra, une modification du modèle du voteur

Nous étudions ici le modèle de Lotka-Volterra, qui est une modification du modèle du voteur. Par conséquent, la dérivée de Radon-Nikodým sera calculée par rapport au modèle du voteur, pas par rapport à la marche aléatoire branchante. De plus, pour les calculs relatifs au modèle

du voteur, nous pourrions utiliser la dualité de celui-ci avec la marche aléatoire coalescente (B_t^x) . Un modèle de Lotka-Volterra est un modèle de compétition entre deux espèces, 0 et 1. Quand l'un des individus meurt, il est immédiatement remplacé par un nouvel individu, dont l'espèce dépend des concentrations de 0 et de 1 au voisinage du site. Nous utiliserons ici les notations suivantes, plus agréables à manipuler, pour les nombres de voisins de chaque espèce d'un site :

$$f_n^i(x) = \sum_e p(e) \mathbf{1}_{\{\xi_n(x+e)=i\}}, \quad i \in \{0, 1\}.$$

La convergence de l'un de ces processus vers un super-processus de Dawson-Watanabe a déjà été prouvé dans [4], en utilisant un argument de tension, et la caractérisation des valeurs d'adhérence par le problème de martingales (5.1). Regardons maintenant la dérivée de Radon-Nikodým du processus à temps continu.

7.1 Modèle de Lotka-Volterra à temps continu

Définition 7.1. Un modèle de Lotka-Volterra à temps continu avec paramètres d'interaction α_0 et α_1 est un système de particules $\xi_t : \mathbb{Z}^d \rightarrow \{0, 1\}$ avec taux de naissance $(1 - \xi_t(x))(f_t^0(x) + \alpha_0 f_t^1(x))$ et taux de mort $f_t^0(x)(f_t^1(x) + \alpha_1 f_t^0(x))$. Ce processus évolue de la manière suivante :

$$\begin{cases} \xi_t(x) \rightarrow \xi_t(x) + 1 \text{ à taux } (1 - \xi_t(x))(f_t^1(x) + (\alpha_0 - 1)f_t^1(x)^2) \\ \xi_t(x) \rightarrow \xi_t(x) - 1 \text{ à taux } \xi_t(x)(f_t^0(x) + (\alpha_1 - 1)f_t^0(x)^2) \end{cases}$$

Soit ξ_t^N une suite de modèles de Lotka-Volterra avec vitesse N et paramètres d'interaction $1 + \frac{\theta_0}{N}, 1 + \frac{\theta_1}{N}$. On note comme d'habitude le processus normalisé de loi \tilde{R}^N :

$$X_t^N = \frac{1}{N} \sum_{x \in \mathbb{Z}^d} \xi_t^N(x) \delta_{\frac{x}{\sqrt{N}}}.$$

Commençons par quelques notations. $(B^x)_{x \in \mathbb{Z}^d}$ est une marche aléatoire coalescente, où B^x est une marche aléatoire commençant en x . Si B^x et B^y se rencontrent à n'importe quel moment, alors ils coalescent, i.e. si $B_t^x = B_t^y$ alors pour tout $s \geq t$, $B_s^x = B_s^y$. Nous posons également :

$$\tau(x, y) = \inf\{t > 0 \mid B_t^x = B_t^y\}$$

$$\beta = \sum_{e, e'} p(e)p(e') \mathbb{P}(\tau(0, e) = \tau(0, e') = +\infty, \tau(e, e') < +\infty)$$

$$\delta = \sum_{e, e'} p(e)p(e') \mathbb{P}(\tau(0, e) = \tau(0, e') = +\infty)$$

Notre but serait alors de prouver le théorème suivant :

Théorème 7.1. Si $X_0^N \xrightarrow[N \rightarrow +\infty]{} \mu \in \mathcal{M}_f(\mathbb{R}^d)$, alors :

$$X^N \xrightarrow[N \rightarrow +\infty]{} X,$$

où X est un super-processus de Dawson-Watanabe avec taux de branchement 2γ et dérive $\beta\theta_0 - \delta\theta_1$.

Pour calculer la dérivée de Radon-Nikodým de ce processus par rapport au modèle du voteur renormalisé, nous devons donner une représentation du modèle du voteur, de la même manière que nous avons calculé les moments de la marche aléatoire branchante. Or on sait que le modèle du voteur à court rayon d'interaction (ξ_t) est la seule solution forte du problème suivant :

$$\xi_t(x) = \xi_0(x) + \sum_{x,y} \int_0^t (\xi_{s-}(y) - \xi_{s-}(x)) d\Lambda_s(x, y).$$

Grâce à cette formule, on peut écrire la dérivée de Radon-Nikodým de la manière suivante :

$$\left. \frac{d\tilde{R}^N}{dR^N} \right|_{\mathcal{F}_t} = \mathcal{E}(M_t^{1,N} - M_t^{0,N}),$$

où on a posé :

$$M_t^{1,N} = \sum_{x,y} \int_0^t \xi_{s-}^N(y) (1 - \xi_{s-}^N(x)) f_s^1(x) d\hat{\Lambda}_s(x, y) \text{ et}$$

$$M_t^{0,N} = \sum_{x,y} \int_0^t \xi_{s-}^N(x) (1 - \xi_{s-}^N(y)) f_s^0(x) d\hat{\Lambda}_s(x, y).$$

La convergence de la variation quadratique divisée par N est établi de la même manière que pour le processus de contact, et nous pouvons également prouver de la même manière qu'on ne peut pas avoir la convergence de la dérivée de Radon-Nikodým .

7.2 Modèle de Lotka-Volterra à temps discret

Définition 7.2. Un modèle de Lotka-Volterra à temps discret avec paramètres d'interaction α_0 et α_1 est un système de particules $\xi_n : \mathbb{Z}^d \rightarrow \{0, 1\}$ avec distribution du nombre d'enfants évoluant de la façon suivante :

$$\xi_n(x) \rightarrow \begin{cases} \xi_{n+1}(x) = 0 & \text{avec probabilité } f_n^0(x) - \varepsilon_n(x) \\ \xi_{n+1}(x) = 1 & \text{avec probabilité } f_n^1(x) + \varepsilon_n(x) \end{cases}$$

où nous notons $\varepsilon_n(x) = (\alpha_0 - 1)f_n^0(x)^2(1 - \xi_n(x)) - (\alpha_1 - 1)f_n^1(x)^2\xi_n(x)$.

Soit ξ_t^N une suite de modèles de Lotka-Volterra avec paramètres d'interaction $1 + \frac{\theta_0}{N}, 1 + \frac{\theta_1}{N}$. Nous notons également la loi du processus renormalisé de la façon usuelle \tilde{R}^N :

$$X_t^N = \frac{1}{N} \sum_{x \in \mathbb{Z}^d} \xi_{[Nt]}^N(x) \delta_{\frac{x}{\sqrt{N}}}.$$

Théorème 7.2. Soit $X_0^N \xrightarrow{N \rightarrow +\infty} \mu \in \mathcal{M}_f(\mathbb{R}^d)$, alors :

$$X^N \xrightarrow{N \rightarrow +\infty} X,$$

où X est un super-processus de Dawson-Watanabe avec taux de branchement γ et dérivée $\beta\theta_0 - \delta\theta_1$, où :

$$\beta = \sum_{e,e'} p(e)p(e') \mathbb{P}(\tau(0, e) = \tau(0, e') = +\infty, \tau(e, e') < +\infty)$$

$$\delta = \sum_{e,e'} p(e)p(e') \mathbb{P}(\tau(0, e) = \tau(0, e') = +\infty)$$

On calcule maintenant la dérivée de Radon-Nikodým de ce processus contre le modèle du voteur, on a :

$$L_t^N = \prod_{n=0}^{\lfloor Nt \rfloor} \prod_x \left(1 + \xi_{n+1}(x) \frac{\varepsilon_n(x)}{f_n^1(x)} + (1 - \xi_{n+1}(x)) \frac{\varepsilon_n(x)}{f_n^0(x)} \right),$$

qui peut être réécrit immédiatement, en utilisant le fait que la mesure martingale orthogonale du modèle du voteur à temps discret est la mesure purement atomique M telle que :

$$\forall k \in \mathbb{N}, \forall x \in \mathbb{Z}^d, \quad M(k, x) = \xi_k(x) - f_{k-1}^1(x) = M_{k,x};$$

comme le produit sur l'espace et le temps de :

$$L_{n,x} = \left(1 + M_{n,x} \frac{\varepsilon_n(x)}{f_n^0(x)f_n^1(x)} \right).$$

Une fois encore, nous voudrions montrer la convergence jointe du processus et de sa variation quadratique, avec les mêmes problèmes que l'on a toujours eu : ε_n ne converge pas vers une constante.

8 Conclusion

Nous avons vu à travers tous ces exemples que les deux plus grands problèmes rencontrés lorsqu'on tente de prouver la convergence de la dérivée de Radon-Nikodým de notre processus vers la limite espérée. Le premier d'entre eux ne concerne que les processus à temps continus, car la dérivée de Radon-Nikodým prend en compte le décalage de temps en $O_P(\frac{1}{N})$ lorsque les taux ne sont pas symétriques.

Le second problème est que notre processus, quand il n'y a pas limite càdlàg de la dérive, il devient alors plus difficile, et parfois même impossible.

Par conséquent, prouver la convergence de la plupart des processus que nous avons étudié, il reste nécessaire de prouver la tension de cette suite de processus, puis caractériser la limite en utilisant le problème de martingales (5.1).

9 Annexe

Dans cette section, nous allons prouver l'équation (5.3). Ici, nous notons ξ_t^N une suite de marches aléatoires branchantes avec taux N et noyau symétrique $p_N = p(\frac{\cdot}{\sqrt{N}})$ sur $\frac{\mathbb{Z}^d}{\sqrt{N}}$, vérifiant :

$$\frac{1}{N} \xi_0^N \xrightarrow[N \rightarrow +\infty]{} \mu.$$

On note $X_t^N = \frac{1}{N} \xi_t^N$, qui est compris comme une mesure sur \mathbb{R}^d .

Pour les calculs qui suivront, nous écrirons η une marche aléatoire branchante avec vitesse N et noyau p sur \mathbb{Z}^d , commençant avec une unique particule à l'instant $t = 0$ en position O . Soit $(\eta^{x,i})_{x \in \mathbb{Z}^d, i \in \mathbb{N}}$ une suite de marches aléatoires branchantes indépendantes de vitesse N , partant d'une particule unique en x à l'instant $t = 0$.

Nous notons également :

- $(B_n)_{n \in \mathbb{N}}$ une marche aléatoire sur \mathbb{Z}^d de noyau p ,
- $\Pi(t)$ un processus de Poisson d'intensité 1,
- $V_t = B_{\Pi(t)}$ une marche aléatoire (à temps continu) sur \mathbb{Z}^d de noyau p et vitesse 1,
- V_t', V_t'', \dots des copies indépendantes de V_t .

9.1 Autour de la marche aléatoire

Ici nous nous intéresserons à V_t la trajectoire d'une unique particule dans une marche aléatoire branchante. La borne supérieure que nous donnerons ici sera très utile plus tard.

Lemme 9.1. *Il existe $C > 0$ tel que pour tout $s > 0$ et $x \in \mathbb{Z}^d$:*

$$\mathbb{E}(p(x + V_s)) < C(1 + s)^{-d/2}$$

Preuve. Nous avons $E(p(x + V_s)) = \sum_{k=0}^{+\infty} P(\Pi(s) = k) \mathbb{E}(p(x + V_k))$, qui peut être réécrit en utilisant le fait que :

$$\mathbb{E}(p(V_k + x)) = \mathbb{P}(V_{k+1} = -x) < C(1 + k)^{-d/2}.$$

Ensuite nous appliquons un résultat de grands déviations pour le processus de Poisson, nous avons :

$$\mathbb{P}\left(\Pi(s) < \frac{s}{2}\right) \leq e^{-cs},$$

pour un certain $c > 0$. Nous avons également :

$$\mathbb{E}(p(x + V_s)) \leq \mathbb{P}(\Pi_s < \frac{s}{2}) + \left(1 + \frac{s}{2}\right)^{-d/2}.$$

□

9.2 Nombre moyen de voisins d'une particule

Nous étudions maintenant le comportement du nombre moyen de voisins d'une particule dans la marche aléatoire branchante. Comme dans le processus limite, seule les densités locales ont un sens, nous allons étudier la quantité suivante :

$$Z_t^N(\phi) = \frac{1}{N} \sum_{x,y} \phi(x) \xi_t^N(x) \xi_t^N(y) p_N(x, y).$$

Dans ce but, nous nous intéressons pour commencer à cette valeur quand la marche aléatoire branchante commence avec une unique particule en $t = 0$, on note :

$$Z_t^N(\phi) = \sum_{x,y} \phi(x) \eta_t(x) \eta_t(y) p(x, y).$$

Lemme 9.2. *Il existe $b > 0$ tel que pour toute suite $\tau_N \rightarrow 0$ telle que $N\tau_N \rightarrow +\infty$, nous avons :*

$$\mathbb{E}(Z_{\tau_N}^{1N}(1)) \xrightarrow{N \rightarrow +\infty} b.$$

De plus, nous avons :

$$\begin{aligned} b &= \int_0^{+\infty} \mathbb{E}(p(V_s)) ds = \int_0^{+\infty} \mathbb{P}(V_s + W = 0) ds \\ &= \sum_{n=0}^{+\infty} \mathbb{E}(p(B_n)) = \sum_{n=1}^{+\infty} \mathbb{E}(\mathbf{1}_{\{B_n=0\}}). \end{aligned}$$

Preuve. Nous savons que pour tout $t > 0$,

$$\begin{aligned} \mathbb{E}\left(\sum_{x,y} \eta_t(x)\eta_t(y)p(x,y)\right) &= \sum_{x,y} \eta_0(P_t(x))\eta_0(P_t(y))p(x,y) \\ &\quad + N \sum_{x,y} \int_0^t \mathbb{E}(\eta_u(P_{t-u}(x))\eta_u(P_{t-u}(y))) \\ &\quad + \eta_u(P^N(P_{t-u}(x))P_{t-u}(y)))p(x,y) du, \end{aligned}$$

qui peut être réécrit, en échangeant la somme et l'espérance, et en utilisant la condition initiale :

$$\mathbb{E}(Z_t^{1N}(1)) = \mathbb{E}(p(V_{2Nt})) + 2N \int_0^t \mathbb{E}(p(V_{2Ns})) ds.$$

Nous utilisons maintenant le Lemme 9.1. pour observer que :

$$\mathbb{E}(p(V_{2N\tau_N})) \leq C(1 + 2N\tau_N)^{-d/2} \xrightarrow{N \rightarrow +\infty} 0.$$

De plus nous pouvons calculer l'autre terme :

$$\begin{aligned} 2N \int_0^{\tau_N} \mathbb{E}(p(V_{2Ns})) ds &= \int_0^{2N\tau_N} \mathbb{E}(p(V_s)) ds \\ &\xrightarrow{N \rightarrow +\infty} \int_0^{+\infty} \mathbb{E}(p(V_s)) ds, \end{aligned}$$

qui est fini en utilisant encore la borne donnée par le Lemme 9.1.

Nous avons alors :

$$\begin{aligned}
b &= \int_0^{+\infty} \mathbb{E}(p(V_s)) ds \\
&= \sum_{n=0}^{+\infty} \mathbb{E}(p(B_n)) \int_0^{+\infty} P(\Pi(s) = n) ds \\
&= \sum_{n=0}^{+\infty} \mathbb{E}(p(B_n)) \\
&= \sum_{n=0}^{+\infty} \sum_e p(e) \mathbb{P}(B_n = e) \\
&= \sum_{n=1}^{+\infty} \mathbb{P}(B_n = 0) \\
&= \sum_{n=0}^{+\infty} \mathbb{P}(B_{n+1} = 0) \int_0^{+\infty} P(\Pi(s) = n) ds \\
&= \int_0^{+\infty} \mathbb{P}(V_s + W = 0) ds
\end{aligned}$$

□

Nous voyons ainsi qu'en commençant d'une unique particule, le nombre moyen de voisins devient très rapidement égal à b . Nous nous intéressons maintenant à la variance de cette quantité, et tentons de trouver une borne supérieure.

Lemme 9.3. *Il existe une constante positive $C > 0$ telle que pour tout $t > 0$ et $N \in \mathbb{N}$:*

$$\mathbb{E}(Z_t'^N(1)^2) \leq CNt.$$

Preuve. Tout au long de cette preuve, nous noterons C une constante assez grande, qui pourra être modifiée au fur et à mesure des calculs, mais reste indépendante de N et t . Nous calculons le second moment de $Z_t'^N(1)$,

$$\mathbb{E}(Z_t'^N(1)^2) = \sum_{a,b,c,d} \mathbb{E}(\eta_t(a)\eta_t(b)\eta_t(c)\eta_t(d))p(a,b)p(c,d).$$

Nous avons à calculer une borne supérieure pour le quatrième moment de la marche aléatoire branchante. Nous utiliserons des symétries pour réduire autant que possible les prochains calculs. Nous utiliserons souvent la borne du Lemme 9.1. Pour commencer, nous allons réduire ces calculs

à des calculs de troisième moment d'une marche aléatoire branchante :

$$\begin{aligned}
& \mathbb{E}(Z_t'^N(1)^2) \\
= & \mathbb{E}(p(V_{2Nt}))^2 \\
& + 4N \sum_{a,b} \int_0^t \mathbb{E}(\eta_s(a)\eta_s(b)\eta_s(1)) \\
& \mathbb{E}(p(b-a+V_{2N(t-s)}))\mathbb{E}(p(V_{2N(t-s)}))ds \\
& + 8N \sum_{a,b,c} \int_0^t \mathbb{E}(\eta_s(a)\eta_s(b)\eta_s(c)) \\
& \mathbb{E}(p(b-a+V_{2N(t-s)}))\mathbb{E}(p(c-a+V_{2N(t-s)}))ds.
\end{aligned}$$

Nous pouvons alors donner les inégalités suivantes :

$$\begin{aligned}
\mathbb{E}(Z_t'^N(1)^2) & \leq 1 + CN \int_0^t \frac{1}{(1+N(t-s))^{d/2}} \\
& \sum_{a,b} \mathbb{E}(\eta_s(a)\eta_s(b)\eta_s(1))\mathbb{E}(p(b-a+V_{2N(t-s)}))ds.
\end{aligned}$$

Il nous suffit juste de donner une bonne borne supérieure pour l'espérance sous l'intégrale :

$$\begin{aligned}
& \sum_{a,b} \mathbb{E}(\eta_s(a)\eta_s(b)\eta_s(1))\mathbb{E}(p(b-a+V_{2N(t-s)})) \\
= & \mathbb{E}(p(V_{2Nt})) \\
& + 2N \int_0^s \mathbb{E}(\eta_u(1)^2)\mathbb{E}(p(V_{2N(t-u)}))du \\
& + 4N \sum_{a,b} \int_0^s \mathbb{E}(\eta_u(a)\eta_u(b))\mathbb{E}(p(b-a+V_{2N(t-u)}))du.
\end{aligned}$$

Nous pouvons maintenant utiliser le fait que $\mathbb{E}(\eta_u(1)^2) = 1 + 2Nu$, pour borner ce terme par :

$$C(1 + \int_{N(t-s)}^{Nt} \frac{1+Nt}{(1+u)^{d/2}} du) \leq C(1+Nt).$$

En utilisant ceci, nous avons finalement :

$$\mathbb{E}(Z_t'^N(1)^2) \leq C(1+Nt)(1 + \int_0^t \frac{1}{(1+N(t-s))^{d/2}}) \leq C(1+Nt).$$

Ce qui termine la preuve de notre lemme. □

Remarque 9.1. Avoir le même genre de bornes dans le cas $d = 2$ n'augmente pas trop la difficulté, nous avons juste à tenir compte de quelques modifications logarithmiques.

Nous avons maintenant assez d'outils pour prouver le résultat le plus important de cette section : que les quantités X_t^N et Z_t^N sont proches l'une de l'autre, ce qui nous permet d'obtenir la convergence du processus, au moins pour $t \geq \varepsilon$ pour le nombre moyen de voisins d'un individu vers b . Nous allons démontrer le théorème suivant.

Théorème 9.1. *Pour toute fonction ϕ continue Lipschitzienne, pour tous $t > 0$, nous avons :*

$$\mathbb{E} \left(\sup_{s \in [0, t]} \left| \int_0^s Z_u^N(\phi) - bX_u^N(\phi) du \right| \right) \xrightarrow{n \rightarrow +\infty} 0.$$

Preuve. La preuve de ce théorème sera fait en plusieurs étapes. Pour commencer, nous remplacerons les quantités $X_t^N(\phi)$ et $Z_t^N(\phi)$ par des approximations : au lieu de compter une particule dans la position qu'elle occupe, on la compte là où était son ancêtre peu de temps auparavant, et en les multipliant soit par le nombre de descendants de cet ancêtre, soit par le nombre moyen de voisins parmi les descendants. Nous avons déjà vu plus tôt que ces deux quantités sont proches.

Nous remplacerons également b par une approximation, et nous prouverons alors le théorème pour les quantités modifiées.

Pour finir, nous prouverons que les approximations que nous avons choisi sont assez bonnes pour résoudre ce problème.

Nous commençons par fixer une suite $\tau_N \xrightarrow{N \rightarrow +\infty} 0$ tel que $N\tau_N \xrightarrow{N \rightarrow +\infty} +\infty$. Nous remonterons de τ_N dans le passé pour trouver la position de l'ancêtre d'une particule en x , cette quantité est choisie de façon à ce qu'il y ait beaucoup de sauts durant l'intervalle de temps τ_N , mais que ce temps tende vers 0.

En étudiant la propriété de branchement, nous observons qu'une marche aléatoire branchante commençant avec k particules est exactement la somme de k marches aléatoires branchantes indépendantes, chacune commençant d'une unique particule.

Pour $t > 0$, en utilisant la propriété de Markov, on a une famille $(\eta^{z,i})_{z \in \frac{z^d}{\sqrt{N}}, i \in \mathbb{N}}$ de marches aléatoires branchantes indépendantes issues d'une unique particule située en $\sqrt{N}z$, tel que :

$$\xi_t^N(x) = \sum_z \sum_{i=1}^{\xi_{t-\tau_N}^N(z)} \eta_{\tau_N \wedge t}^{z,i}(\sqrt{N}x),$$

où nous utilisons la convention suivante : pour $s < 0$, $\xi_s^N(x) = \xi_0^N(x)$.

Nous pouvons alors poser les approximations suivantes :

$$\tilde{X}_t^N(\phi) = \frac{1}{N} \sum_z \phi(z) \sum_{i=1}^{\xi_{t-\tau_N}^N(z)} \eta_{\tau_N \wedge t}^{z,i}(1)$$

pour X^N , et pour Z^N :

$$\tilde{Z}_t^N(\phi) = \frac{1}{N} \sum_z \phi(z) \sum_{i=0}^{\xi_{t-\tau_N}^N(z)} \sum_{x,y} \eta_{\tau_N \wedge t}^{z,i}(x) \eta_{\tau_N \wedge t}^{z,i}(y) p(x,y)$$

Pour finir, nous noterons $\tilde{b}_N = \mathbb{E}(Z_{\tau_N \wedge t}^N(1))$ une approximation de b , prouvons maintenant le théorème pour ces quantités.

Lemme 9.4. *Pour toute fonction ϕ continue Lipschitzienne, pour tout $t > 0$ on a :*

$$\mathbb{E} \left(\sup_{s \in [0, t]} \left| \int_0^s \tilde{Z}_u^N(\phi) - \tilde{b}_N \tilde{X}_u^N(\phi) du \right| \right) \xrightarrow{N \rightarrow +\infty} 0.$$

Preuve. Nous commençons par une notation :

$$\begin{aligned} X_s^{z,i} &= \eta_{s \wedge t}^{z,i}(1) \quad \text{and} \\ Z_s^{z,i} &= \sum_{x,y} \eta_{s \wedge t}^{z,i}(x) \eta_{s \wedge t}^{z,i}(y) p(x,y). \end{aligned}$$

Nous pouvons réécrire la formule de la différence que nous calculons de la manière suivante :

$$\begin{aligned} & \tilde{Z}_t^N(\phi) - \tilde{b}_N \tilde{X}_t^N(\phi) \\ &= \frac{1}{N} \sum_z \phi(z) \sum_{i=0}^{\xi_{t-\tau_N}^N(z)} Z_{\tau_N}^{z,i} - \tilde{b}_N X_{\tau_N}^{z,i}. \end{aligned}$$

Par conséquent $\tilde{Z}_t^N(\phi) - \tilde{b}_N \tilde{X}_t^N(\phi)$ est la somme de variables aléatoires indépendantes d'espérance nulle, donc l'espérance de la somme est nulle, et la variance vaut :

$$\begin{aligned} & \mathbb{E}((\tilde{Z}_t^N(\phi) - \tilde{b}_N \tilde{X}_t^N(\phi))^2) \\ &= \frac{1}{N^2} \sum_z \phi(z)^2 \mathbb{E}(\xi_{t-\tau_N}^N(z)) \mathbb{E}((Z_{\tau_N}^{z,i} - \tilde{b}_N X_{\tau_N}^{z,i})^2) \\ &= \frac{X_0^N(P_{t-\tau_N}(\phi^2))}{N} \mathbb{E}((Z'_{\tau_N \wedge t}(1) - \tilde{b}_N \eta_{\tau_N \wedge t}(1))^2) \\ &\leq C \|\phi\|_\infty^2 X_0^N(1) \tau_N. \end{aligned}$$

Dans la dernière inégalité, nous avons utilisé le Lemme 9.2., qui donne $\tilde{b}_N \xrightarrow{n \rightarrow +\infty} b$, le Lemme 9.3., le fait que $N\tau_N \xrightarrow{n \rightarrow +\infty} +\infty$ et l'inégalité de Cauchy-Schwarz pour obtenir la borne supérieure suivante :

$$\begin{aligned} & \mathbb{E}((Z'_{\tau_N \wedge t}(1) - \tilde{b}_N \eta_{\tau_N \wedge t}(1))^2) \\ &= \mathbb{E}((Z'_{\tau_N \wedge t}(1))^2) - 2\tilde{b}_N \mathbb{E}(Z'_{\tau_N \wedge t}(1) \eta_{\tau_N \wedge t}(1)) + \mathbb{E}((\eta_{\tau_N \wedge t}(1))^2) \\ &\leq C(1 + (\tau_N \wedge t)N) \leq CN\tau_N \end{aligned}$$

Nous pouvons donc dominer l'intégrale pour $t \geq \tau_N$:

$$\begin{aligned} & \int_{\tau_N}^t \mathbb{E}(|\tilde{Z}_s^N(\phi) - \tilde{b}_N \tilde{X}_s^N(\phi)|) ds \\ &\leq \int_{\tau_N}^t \mathbb{E}((\tilde{Z}_s^N(\phi) - \tilde{b}_N \tilde{X}_s^N(\phi))^2)^{1/2} ds \\ &\leq C\tau_N^{1/2} \|\phi\|_\infty X_0^N(1)^{1/2} t \xrightarrow{n \rightarrow +\infty} 0. \end{aligned}$$

Pour dominer l'intégrale entre 0 et τ_N , il nous suffira d'utiliser les inégalités suivantes :

$$\begin{aligned}
& \int_0^{\tau_N} \mathbb{E}(|\tilde{Z}_s^N(\phi) - \tilde{b}_N \tilde{X}_s^N(\phi)|) ds \\
& \leq \|\phi\|_\infty \int_0^{\tau_N} \tilde{b}_N \mathbb{E}(\tilde{X}_s^N(1)) + \mathbb{E}(\tilde{Z}_s^N(1)) ds \\
& \leq \|\phi\|_\infty (X_0^N(1) \tau_N b_{\tau_N} + X_0^N(1) \mathbb{E}(Z_{\tau_N}^N(1))) \\
& \leq C \|\phi\|_\infty X_0^N(1) \tau_N \xrightarrow{n \rightarrow +\infty} 0.
\end{aligned}$$

Les deux bornes obtenues nous permettent de finir la preuve. En effet nous avons :

$$\begin{aligned}
\mathbb{E}(\sup_{s \in [0, t]} |\tilde{Z}_s^N(\phi) - \tilde{b}_N \tilde{X}_s^N(\phi)|) & \leq \mathbb{E}(\int_0^t |\tilde{Z}_s^N(\phi) - \tilde{b}_N \tilde{X}_s^N(\phi)| du) \\
& \leq \int_0^t \mathbb{E}(|\tilde{Z}_s^N(\phi) - \tilde{b}_N \tilde{X}_s^N(\phi)|) du \\
& \leq C \|\phi\|_\infty X_0^N(1) (\tau_N)^{1/2} t \xrightarrow{n \rightarrow +\infty} 0.
\end{aligned}$$

□

Nous allons maintenant prouver que les approximations des quantités que nous avons choisies sont proches de celles attendues.

Lemme 9.5. *Pour toute fonction ϕ continue Lipschitzienne, pour tout $t > 0$, on a :*

$$\mathbb{E}(\int_0^t |X_s^N(\phi) - X_s^{N, \tau_N}(\phi)| ds) \xrightarrow{N \rightarrow +\infty} 0.$$

Preuve. Commençons par donner une représentation de la différence entre les deux termes en fonction de marches aléatoires branchantes indépendantes :

$$\begin{aligned}
|X_t^N(\phi) - \tilde{X}_t^N(\phi)| & = \left| \frac{1}{N} \sum_z \sum_{i=1}^{\xi_{t-\tau_N}^N(z)} \sum_x (\phi(x) - \phi(z)) \eta_{\tau_N \wedge t}^{z,i}(x\sqrt{N}) \right| \\
& \leq \frac{1}{N} \sum_z \sum_{i=1}^{\xi_{t-\tau_N}^N(z)} \sum_x |\phi(x) - \phi(z)| \eta_{\tau_N \wedge t}^{z,i}(x\sqrt{N}) \\
& \leq \frac{C}{N} \sum_z \sum_{i=1}^{\xi_{t-\tau_N}^N(z)} \sum_x \|z - x\| \eta_{\tau_N \wedge t}^{z,i}(x\sqrt{N}).
\end{aligned}$$

Cette dernière majoration est la somme de variables aléatoires i.i.d., donc d'espérance au plus

$$\frac{C}{N} \mathbb{E}(\xi_{t-\tau_N}^N(1)) \mathbb{E}(\sum_x \frac{\|x\|}{\sqrt{N}} \eta_{\tau_N \wedge t}(x)),$$

et de plus, nous avons :

$$\mathbb{E}\left(\sum_x \|x\| \eta_{\tau_N \wedge t}(x)\right) = \mathbb{E}(|V_{N\tau_N}|) \leq C\sqrt{N\tau_N}.$$

Ce qui nous permet de finir la preuve, car :

$$\mathbb{E}\left(\int_0^t |X_s^N(\phi) - \tilde{X}_s^N(\phi)| ds\right) \leq CtX_0^N(1)\sqrt{\tau_N} \xrightarrow{n \rightarrow +\infty} 0.$$

□

Pour finir, il nous faut prouver que \tilde{Z}^N est une bonne approximation de Z^N . En écrivant Z^N comme somme des termes $\eta^{z,i}$, nous pouvons séparer la preuve en deux parties disjointes :

$$\begin{aligned} Z_t^N(\phi) &= \sum_{x,y} \phi(x)\xi_t^N(x)\xi_t^N(y)p(x,y) \\ &= \sum_{z,z'} \sum_{i=1}^{\xi_{t-\tau_N}^N(z)} \sum_{j=1}^{\xi_{t-\tau_N}^N(z')} \sum_{x,y} \phi(x)\eta_{\tau_N \wedge t}^{z,i}(x)\eta_{\tau_N \wedge t}^{z',j}(y)p(x,y) \\ &= \sum_z \sum_{i=1}^{\xi_{t-\tau}^N(z)} \sum_{x,y} \phi(x)\eta_{\tau_N \wedge t}^{z,i}(x)\eta_{\tau_N \wedge t}^{z,i}(y)p(x,y) \\ &+ \sum_{z,z'} \sum_{i=1}^{\xi_{t-\tau_N}^N(z)} \sum_{j=1, (z,i) \neq (z',j)}^{\xi_{t-\tau_N}^N(z')} \sum_{x,y} \phi(x)\eta_{\tau_N \wedge t}^{z,i}(x)\eta_{\tau_N \wedge t}^{z',j}(y)p(x,y). \end{aligned}$$

Commençons par montrer que le premier de ces deux termes (qui ne compte comme voisins que les individus qui sont proches parents les uns des autres – et voisins) est à lui seul proche de $\tilde{Z}_t^N(\phi)$.

Lemme 9.6. *Nous noterons par la suite :*

$$Z_t^{N,1}(\phi) = \sum_z \sum_{i=1}^{\xi_{t-\tau}^N(z)} \sum_{x,y} \phi(x)\eta_{\tau_N \wedge t}^{z,i}(x)\eta_{\tau_N \wedge t}^{z,i}(y)p(x,y)$$

Pour toute fonction ϕ continue Lipschitzienne, pour tout $t > 0$, on a :

$$\mathbb{E}\left(\int_0^t |Z_s^{N,1}(\phi) - \tilde{Z}_s^N(\phi)| ds\right) \xrightarrow{N \rightarrow +\infty} 0.$$

Preuve. De la même manière que précédemment, nous commencerons par estimer la différence

entre les deux quantités évaluées de façon probabiliste :

$$\begin{aligned}
|Z_t^{N,1}(\phi) - \tilde{Z}_t^N(\phi)| &\leq \frac{1}{N} \sum_z \sum_{i=1}^{\xi_{t-\tau_N}^N(z)} \sum_{x,y} \left| \phi\left(\frac{x}{\sqrt{N}}\right) - \phi(z) \right| \eta_{\tau_N \wedge t}^{z,i}(x) \eta_{\tau_N \wedge t}^{z,i}(y) p(x,y) \\
&\leq \frac{C}{N} \sum_z \sum_{i=1}^{\xi_{t-\tau_N}^N(z)} \sum_{x,y} \left\| z - \frac{x}{\sqrt{N}} \right\| \eta_{\tau_N \wedge t}^{z,i}(x) \eta_{\tau_N \wedge t}^{z,i}(y) p(x,y).
\end{aligned}$$

Une fois encore, cette dernière majoration est la somme de variables aléatoires i.i.d., donc son espérance est au plus égale à :

$$\frac{C}{N} \mathbb{E}(\xi_{t-\tau_N}^N(1)) \mathbb{E}\left(\sum_{x,y} \frac{\|x\|}{\sqrt{N}} \eta_{\tau_N \wedge t}(x) \eta_{\tau_N \wedge t}(y) p(x,y)\right),$$

de plus, on a :

$$\begin{aligned}
&\mathbb{E}\left(\sum_{x,y} \|x\| \eta_s(x) \eta_s(y) p(x,y)\right) \\
&= \mathbb{E}(\|V_{Ns}\| p(V_{2Ns})) \\
&+ N \int_0^s \mathbb{E}(\|V_{N(s-u)+V'_{Nu}}\| + \|V_{N(s-u)} + V'_{Nu} + W\|) p(V_{2N(s-u)}) du \\
&\leq C\sqrt{Ns} \left(1 + \int_0^s \frac{N du}{(1+Nu)^{d/2}}\right) \\
&\leq C\sqrt{Ns},
\end{aligned}$$

où pour la première majoration, nous avons utilisé l'inégalité suivante :

$$\begin{aligned}
\mathbb{E}(\|V_{t-s} + V'_s\| p(V_{2(t-s)})) &= \mathbb{E}(\|V_{t-s} + V'_s\| \mathbb{E}(p(V_{2(t-s)}) | V_{t-s})) \\
&\leq \frac{C}{(1+t-s)^{d/2}} \mathbb{E}(\|V_t\|).
\end{aligned}$$

Nous obtenons maintenant que :

$$\mathbb{E}\left(\int_0^t |Z_s^{N,1}(\phi) - \tilde{Z}_s^N(\phi)| ds\right) \leq Ct X_0^N(1) \sqrt{\tau_N} \xrightarrow{N \rightarrow +\infty} 0.$$

Ce qui termine la preuve, très similaire à la précédente. \square

Nous avons maintenant besoin d'un dernier résultat pour borner le terme d'interférences entre les particules qui sont voisines, mais qui ne sont pas de proches parents (i.e. que leur dernier ancêtre commun, s'il existe, soit âgé de plus de τ_N).

Lemme 9.7. *Nous noterons par la suite :*

$$Z_t^{N,2}(\phi) = \frac{1}{N} \sum_{z,z'} \sum_{i=1}^{\xi_{t-\tau_N}^N(z)} \sum_{j=1, (z,i) \neq (z',j)}^{\xi_{t-\tau_N}^N(z')} \sum_{x,y} \phi(x) \eta_{\tau_N \wedge t}^{z,i}(x) \eta_{\tau_N \wedge t}^{z',j}(y) p(x, y).$$

Pour toute fonction ϕ continue Lipschitzienne, pour tout $t > 0$, nous avons :

$$\mathbb{E} \left(\sup_{s \in [0,t]} \left| \int_0^s Z_u^{N,2}(\phi) du \right| \right) \xrightarrow{n \rightarrow +\infty} 0.$$

Preuve. Commençons par borner $\mathbb{E}(|Z_t^{N,2}(\phi)|)$ pour $t > \tau_N$:

$$\begin{aligned} \mathbb{E}(|Z_t^{N,2}(\phi)|) &\leq \frac{\|\phi\|_\infty}{N} \sum_{z,z'} \mathbb{E}(\xi_{t-\tau_N}^N(z) \xi_{t-\tau_N}^N(z')) \\ &\quad \sum_{x,y} \mathbb{E}(\eta_{\tau_N}(x)) \mathbb{E}(\eta_{\tau_N}(y)) p(x+z, y+z') \\ &\leq \frac{\|\phi\|_\infty}{N} \sum_{z,z'} \mathbb{E}(\xi_{t-\tau_N}^N(z) \xi_{t-\tau_N}^N(z')) \mathbb{E}(p(z' - z + V_{2N\tau_N})) \\ &\leq \frac{C}{N} \sum_{z,z'} \xi_0^N(z) \xi_0^N(z') \mathbb{E}(p(z' - z + V_{2Nt})) \\ &\quad + C \int_{\tau_N}^t \xi_0^N(1) \mathbb{E}(p(V_{2Ns})) ds \\ &\leq CN X_0^N(1)^2 \frac{1}{(1+Nt)^{d/2}} + CX_0^N(1) \int_{2N\tau_N}^{2Nt} \frac{ds}{(1+s)^{d/2}}. \end{aligned}$$

Par conséquent, on a :

$$\begin{aligned} \int_{\tau_N}^t \mathbb{E}(|Z_s^{N,2}(\phi)|) ds &\leq CX_0^N(1)^2 \int_{N\tau_N}^{Nt} \frac{ds}{(1+s)^{d/2}} \\ &\quad + \frac{C}{N} X_0^N(1) \int_{2N\tau_N}^{2Nt} \int_{2N\tau_N}^s \frac{du}{(1+u^{d/2})} \\ &\leq C(X_0^N(1)^2 + X_0^N(1)t) \int_{N\tau_N}^{+\infty} \frac{ds}{(1+s)^{d/2}} \\ &\xrightarrow{N \rightarrow +\infty} 0. \end{aligned}$$

Nous nous intéressons maintenant à l'autre intégrale, en utilisant le fait que μ est sans atome

pour donner la conclusion :

$$\begin{aligned}
& \left| \int_0^{\tau_N} \mathbb{E}(Z_s^{N,2}(\phi)) ds \right| \\
& \leq \frac{C}{N^2} \sum_{z,z'} \xi_0^N(z) \xi_0^N(z') \int_0^{2N\tau_N} \mathbb{E}(p(z' - z + V_s)) ds \\
& \leq \frac{C}{N^2} \sum_{z,z'} \xi_0^N(z) \xi_0^N(z') \sum_{n=0}^{+\infty} \mathbb{E}(p(z' - z + B_n)) \int_0^{2N\tau_N} \mathbb{P}(\Pi(s) = n) ds \\
& \leq C X_0^N(1)^2 \frac{1}{(1+N)^{d/2}} \int_0^{2N\tau_N} \mathbb{P}(\Pi(s) > N) ds \\
& \quad + C X_0^N \times X_0^N(\{\|y-x\| < \varepsilon\}) + C X_0^N(1)^2 \sum_{n=\varepsilon\sqrt{N}}^N (1+n)^{-d/2} ds \\
& \xrightarrow{N \rightarrow +\infty} 0.
\end{aligned}$$

□

Ce dernier lemme nous permet de finir la preuve du Théorème 9.1, et en particulier la preuve de l'équation (5.3). □

Bibliographie

- [1] M. Bramson, R. Durrett, and G. Swindle. Statistical mechanics of crabgrass. *The Annals of Probability*, 17(2), 1989.
- [2] J.T. Cox and R. Durrett. Hybrid zones and voter model interfaces. *Bernoulli*, 1(4), 1995.
- [3] J.T. Cox, R. Durrett, and E.A. Perkins. Rescaled voter models converge to super-Brownian motion. *The Annals of Probability*, 28(1), 2000.
- [4] J.T. Cox and E.A. Perkins. Rescaled lotka-volterra models converge to super-Brownian motion. *The Annals of Probability*, 33(3), 2005.
- [5] J.T. Cox and E.A. Perkins. Survival and coexistence in stochastic spatial lotka-volterra models. *Probability Theory and Related Fields*, 139(1-2), 2006.
- [6] R. Durrett and E.A. Perkins. Rescaled contact processes converge to super-Brownian motion in two or more dimensions. *Probability Theory and Related Fields*, 114, 1999.
- [7] S. P. Lalley. Spacial epidemics : critical behavior in one dimension. *Probability Theory and Related Fields*, 2007.
- [8] S. P. Lalley and X. Zheng. Spacial epidemics and local times for critical branching random walks in dimensions 2 and 3. *Probability Theory and Related Fields*, 2009.
- [9] J.-F. Le Gall and M. Merle. On the occupation measure of super-brownian motion.
- [10] S. Sugitani. Some properties for the measure-valued branching diffusion processes. *J. Math. Soc. Japan*, 41(3), 1989.

Quatrième partie

Mémoire de M1

Chapitre 6

Théorèmes limites pour certaines classes d'urnes aléatoires

Vincent BRAULT et Bastien MALLEIN

Sous la direction de Philippe MARCHAL

22 juin 2009

Revu le 20 mai 2011

Résumé

Les urnes aléatoires ont été étudiées de nombreuses fois depuis l'émergence de la théorie des probabilités, par exemple par Jacob Bernoulli et Laplace [7]. Le but est ici de déterminer comment évoluent les populations de boules de chaque couleur. Plus tard, Pólya, Ehrenfest [2], ou Friedman [6], par exemple, ont étudié de telles urnes, certaines d'entre elles servant à modéliser certains systèmes dynamiques : contagion d'une population pour l'urne de Pólya, mélange moléculaire d'un gaz pour Ehrenfest...

Dans le présent mémoire, nous nous attacherons à développer des outils généraux d'analyse permettant la résolution de certaines classes d'urnes, pour lesquelles les méthodes probabilistes montrent des limites. Les deux classes d'urnes que nous étudierons ici sont les urnes semi-sacrificielles et les urnes triangulaires. Les résultats exposés par la suite illustreront la variété des phénomènes de convergence, de type théorème de la limite centrale, pouvant se produire dans les urnes aléatoires. Nous commencerons par quelques notations sur les urnes aléatoires, avant de nous intéresser à plusieurs cas particuliers.

Nous remercions Philippe Marchal pour son soutien tout au long de la rédaction de ce mémoire, ainsi que Philippe Flajolet qui a aimablement répondu à nos questions.

Sommaire

1	Introduction	126
2	Généralités sur les urnes	126
2.1	Définition du modèle	126
2.2	Outils mathématiques	127
2.3	Exemples d'application à des urnes aléatoires simples	131
3	Solution générale pour les urnes sacrificielles	135
3.1	Intégration du système différentiel	135
3.2	Quantités homogènes	136
4	Urnés semi-sacrificielles	138
4.1	Propriétés analytiques des fonctions de base	138
4.2	Propriétés probabilistes des urnes semi-sacrificielles	141
5	Urnés triangulaires	144
5.1	Résultats de la théorie des martingales	144
5.2	Fonctions génératrices	145
5.3	Loi de la variable aléatoire limite	148
6	Conclusion	152

1 Introduction

Une urne aléatoire est un objet défini de la manière suivante : on considère une urne contenant des boules de deux couleurs : noires et blanches. On étudie l'évolution de sa composition lorsqu'on lui applique la règle suivante fixée un grand nombre de fois :

- on tire une boule au hasard dans l'urne ;
- on regarde sa couleur, puis on la replace dans l'urne ;
- si la boule est blanche, on ajoute α boules blanches et β boules noires dans l'urne ;
- si la boule est noire, on ajoute γ boules blanches et δ boules noires.

Les entiers α, β, γ et δ définissent l'urne de manière unique. On pourra autoriser α ou δ à être négatif, dans ce cas les boules ne seront pas ajoutées mais retirées.

De nombreuses urnes aléatoires peuvent être étudiées simplement grâce à des outils de théorie des probabilités, comme les martingales. Mais dans certains cas, notamment lorsqu'on autorise à enlever des boules, ces outils deviennent plus difficiles à utiliser. Il est alors utile de passer par de l'analyse et le dénombrement pour obtenir des résultats. Dans d'autres cas, alors que les probabilités donnent un ordre de grandeur, l'utilisation de résultats d'analyse permet d'obtenir un résultat très fort de convergence.

Les résultats présentés ici sont issus de l'article [4].

2 Généralités sur les urnes

2.1 Définition du modèle

Définition 2.1. Nous considérons l'urne aléatoire évoluant de la manière suivante : à chaque tour si on a pioché une boule blanche, on ajoute α boules blanches et β boules noires, sinon, on en ajoute γ blanches et δ noires. Cette urne sera notée des deux manières équivalentes suivantes :

$$[\alpha, \beta, \gamma, \delta] \text{ ou } \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \quad (\alpha, \delta) \in \mathbb{Z}^2, (\beta, \gamma) \in \mathbb{N}^2.$$

Une telle urne est dite équilibrée si elle satisfait $\alpha + \beta = \gamma + \delta = \sigma$, σ est appelé le jeu de l'urne.

Nous n'étudierons ici que des urnes équilibrées. En particulier, le nombre total de boules présentes dans l'urne à chaque instant est déterministe. À chaque nouvelle étape, ce nombre augmente de σ .

Notation 2.1. On notera a_0 et b_0 le nombre de boules blanches et noires présentes dans l'urne à l'instant initial, a_n et b_n les variables aléatoires correspondant au nombre de boules après n itérations. On pose enfin $s_n = a_n + b_n = s_0 + n\sigma$ le nombre total de boules présentes dans l'urne à l'instant n .

Exemple 2.1. L'urne d'Ehrenfest, représentant le mixage moléculaire d'un gaz dans un récipient contenant deux chambres, séparées par une paroi, est symbolisée par l'urne $\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$. À l'instant n , on a a_n molécules de gaz dans la chambre de droite et b_n dans la chambre de gauche. À

FIGURE 6.1 – Nombre de boules noires au cours du temps, dans une urne d’Ehrenfest contenant 100 boules blanches à l’instant initial.

chaque instant, on choisit une molécule, qui traverse la paroi. On considérera la situation initiale d’un gaz entièrement contenu dans la première chambre, qui correspond à $a_0 = N$, $b_0 = 0$. On a $\forall n \in \mathbb{N}$, $s_n = N$.

On ne considérera par la suite que des urnes dont les conditions initiales et les règles sont agencées de telle manière que l’action définie soit toujours possible, par exemple qu’on ne puisse pas demander de retirer deux boules blanches, s’il n’y en a plus qu’une présente dans l’urne.

On peut vérifier que ces urnes sont celles qui vérifient les deux conditions suivantes :

- Si $\alpha < 0$, alors α divise γ et a_0
- Si $\delta < 0$, alors δ divise β et b_0

2.2 Outils mathématiques

Série génératrice

Définition 2.2. Une histoire d’une urne aléatoire est la donnée des boules successivement tirées dans l’urne. Chaque boule est considérée comme unique.

On note respectivement h_n et $h_n(a, b)$ le nombre d’histoires possibles de longueur n et le nombre de ces évolutions conduisant à $a_n = a$ et $b_n = b$. Ces quantités dépendent bien entendu de a_0 et b_0 , mais on ne notera pas cette dépendance par soucis de simplicité.

On définit alors les séries génératrices associées à l’urne comme :

$$H(z) = \sum_n h_n \frac{z^n}{n!} \text{ et } H(x, y, z) = \sum_{n,a,b} h_n(a, b) x^a y^b \frac{z^n}{n!}.$$

Exemple 2.2. Pour l’urne d’Ehrenfest partant de 4 boules blanches, une histoire de longueur 6 peut être (en soulignant à chaque étape la boule tirée, qui est remplacée par une boule noire) :

$$\underline{b}bb\underline{b} \rightarrow \underline{b}bbn \rightarrow n\underline{b}bn \rightarrow \underline{n}bnn \rightarrow \underline{b}bn\underline{n} \rightarrow \underline{b}bnb \rightarrow n\underline{b}nb.$$

Partons maintenant de $a_0 = N$ boules blanches et $b_0 = 0$ boules noires. On a $h_n = N^n$, car à chaque étape, on a N boules dans l’urne, donc N choix possibles d’une nouvelle boule. Déterminer $h_n(a, b)$ est plus délicat, mais on a par exemple :

- on a N choix possibles lors de la première étape : on tire l’une des N boules blanches au hasard, qui conduisent toutes à une composition $(N - 1, 1)$ de l’urne. Par conséquent on a $h_1(N - 1, 1) = N$;
- on a ici deux possibilités : soit on tire une autre boule blanche ($N - 1$ possibilités), soit on tire la boule noire. Ce qui donne au total $N(N - 1)$ possibilités d’arriver à la composition $(N - 2, N)$, on écrit donc $h_2(N, 0) = N$ et $h_2(N - 2, 2) = N(N - 1)$;
- On a $h_3(N - 1, 1) = N^2 + 2N(N - 1)$ et $h_3(N - 3, 3) = N(N - 1)(N - 2)$.

Pour donner une expression générale, pour tout $n \in \mathbb{N}$, on essaiera de déterminer la série génératrice associée.

Notation 2.2. On notera $[z^n]f$ le coefficient en z^n de la fonction analytique f . Par exemple, on aura $[z^n]H(z) = h_n$ et $[z^n]H(x, y, z) = \sum_{a,b} h_n(a, b)x^a y^b$.

Notons que la probabilité d'avoir a boules blanches et b boules noires à l'instant n est donnée par la formule :

$$\mathbb{P}(a_n = a, b_n = b) = \frac{h_n(a, b)}{h_n} = \frac{[x^a y^b z^n]H_n(x, y, z)}{[z^n]H(z)}.$$

La détermination des séries génératrices associées à l'urne permet donc d'accéder à la loi du nombre de boules à tout instant donné, par développement des fonctions en séries entières. Nous introduirons par la suite des outils nous permettant de calculer ces séries génératrices.

Proposition 2.1. *Les fonctions H définies précédemment admettent un rayon de convergence non-nul, et sont donc définies au voisinage de 0. De plus, nous avons :*

- si $\sigma > 0$, on a $H(z) = \frac{1}{(1 - \sigma z)^{\frac{s_0}{\sigma}}}$ et $H(x, y, z)$ est analytique, pour $R > 1$ sur des domaines du type $|x| < R$, $|y| < R$ et $z < \frac{1}{\sigma R^\sigma}$;
- si $\sigma = 0$, on a $H(z) = e^{s_0 z}$ et $H(x, y, z)$ est analytique sur \mathbb{C}^3 ;
- si $\sigma < 0$, on a $H(z) = (1 - \sigma z)^{-\frac{s_0}{\sigma}}$, et $H(x, y, z)$ est un polynôme.

Preuve. On notera que la série génératrice $H(z) = H(1, 1, z)$ est assez simple à déterminer pour une urne équilibrée. En effet, on sait que $s_n = s_0 + n\sigma$, par conséquent $h_n = \prod_{k=0}^{n-1} (s_0 + k\sigma)$. On distingue ensuite les cas en fonction de la valeur du jeu σ .

- Si $\sigma > 0$, on a :

$$h_n = \sigma^n \prod_{k=0}^{n-1} \left(\frac{s_0}{\sigma} + k \right) = \sigma^n \frac{\Gamma\left(n + \frac{s_0}{\sigma}\right)}{\Gamma\left(\frac{s_0}{\sigma}\right)},$$

d'où on tire :

$$H(z) = \frac{1}{(1 - \sigma z)^{\frac{s_0}{\sigma}}}.$$

De plus, comme $h_n(a, b) \leq h_n$, pour $R > 1$, $|x| < R$ et $|y| < R$ on a également :

$$[z^n]H(x, y, z) \leq (s_0 + n\sigma)R^{s_0+n\sigma}\sigma^n \frac{\Gamma\left(\frac{s_0}{\sigma} + n\right)}{\Gamma\left(\frac{s_0}{\sigma}\right)\Gamma(n+1)},$$

par conséquent, pour $|z| < \frac{1}{\sigma R^\sigma}$, la série converge.

- Si $\sigma = 0$, on a $h_n = s_0^n$, et par conséquent :

$$H(z) = \exp(s_0 z).$$

De plus, $h_n(a, b) \leq h_n$, donc pour $R > 1$, $|x| < R$ et $|y| < R$ on a

$$[z^n]H(x, y, z) \leq s_0 R^{s_0} \frac{s_0^n}{n!},$$

donc la série converge pour tout $z \in \mathbb{C}$.

- Si $\sigma < 0$, la longueur totale des histoires possibles est limitée par $-\frac{s_0}{\sigma}$. C'est en effet le nombre d'étapes nécessaires pour complètement vider l'urne. Par conséquent :

$$H(z) = (1 - \sigma z)^{-\frac{s_0}{\sigma}},$$

et $H(x, y, z)$ est un polynôme. □

Remarque 2.1. 1. Les urnes considérées sont équilibrées, on sait donc que $a_n + b_n = s_0 + n\sigma$, ce qui se traduit par l'égalité suivante :

$$H(x, y, z) = \sum_n \sum_a h_n(a, s_0 + n\sigma - a) x^a y^{s_0 + n\sigma - a} \frac{z^n}{n!} = y^{s_0} H\left(\frac{x}{y}, 1, zy^\sigma\right).$$

On pourra donc simplifier les études ultérieures en considérant la fonction de deux variables $H(x, z) = H(x, 1, z)$.

2. Notons également que les dérivées partielles de H donnent des informations sur les moments de a_n et b_n . En effet, on a :

$$\partial_x H(x, 1, z) = \sum_{a, b, n} h_n(a, b) a x^{a-1} \frac{z^n}{n!},$$

et en prenant cette égalité en $x = 1$, on a :

$$\partial_x H(x, 1, z)|_{x=1} = \sum_n h_n \frac{z^n}{n!} \sum_{a, b} a \mathbb{P}(a_n = a, b_n = b) = \sum_n h_n \frac{z^n}{n!} \mathbb{E}(a_n).$$

Plus généralement, on obtient, pour tout couple d'entiers k, h , la formule :

$$\frac{[z^n](\partial_x^k \partial_y^h H(x, y, z))_{x, y=1, 1}}{[z^n]H(z)} = \mathbb{E}(a_n(a_n - 1)\dots(a_n - k + 1)b_n\dots(b_n - h + 1)),$$

qui sera utilisée plusieurs fois par la suite.

Opérateur différentiel

On remarque que dans la série génératrice à trois variables, le monôme $x^a y^b$ « représente » la situation dans laquelle l'urne contient a boules blanches et b boules noires. Notons alors que $x\partial_x x^a y^b = ax^a y^b$, c'est-à-dire que cette transformation multiplie la situation considérée par le nombre d'histoires partant de cette situation et sélectionnant une boule blanche.

Définition 2.3. À toute l'urne $[\alpha, \beta, \gamma, \delta]$, on associe l'opérateur différentiel D , défini par :

$$D = x^{\alpha+1} y^\beta \partial_x + x^\gamma y^{\delta+1} \partial_y.$$

L'opérateur différentiel ainsi défini associe à l'« état » (a, b) de l'urne les états suivants : a fois l'« état » $(a + \alpha, b + \beta)$ plus b fois l'« état » $(a + \gamma, b + \delta)$, tout comme l'ensemble des possibilités envisageables dans l'urne aléatoires à l'étape suivante.

Exemple 2.3. L'opérateur associé à l'urne d'Ehrenfest est $D = y\partial_x + x\partial_y$.

On peut lier l'opérateur et les séries génératrices associées à une même urne aléatoire. En effet, on a par récurrence $D^n(x^{a_0}y^{b_0}) = \sum_{a,b} h_n(a,b)x^a y^b$, donc en sommant par rapport à n , on obtient :

$$H(x, y, z) = \sum_n D^n(x^{a_0}y^{b_0}) \frac{z^n}{n!} = e^{zD}(x^{a_0}y^{b_0}).$$

Système différentiel

Nous allons maintenant utiliser un système différentiel associé à notre urne aléatoire afin de déterminer la série génératrice associée à l'urne.

Définition 2.4. À l'urne $[\alpha, \beta, \gamma, \delta]$ on associe le système différentiel autonome suivant

$$\begin{cases} \dot{X} = X^{\alpha+1}Y^\beta \\ \dot{Y} = X^\gamma Y^{\delta+1}. \end{cases}$$

Notons au passage que le système différentiel associé à l'urne est indépendant de la situation initiale de l'urne.

Théorème 2.1. *On considère une urne possédant à l'instant initial a_0 boules blanches et b_0 noires. Si $(X(t), Y(t))$ est la solution du système différentiel associé à l'urne, valant (x, y) à $t = 0$ alors*

$$H(x, y, z) = X(z)^{a_0} Y(z)^{b_0},$$

dès lors que les quantités sont définies.

Preuve. On sait que :

$$H(x, y, z) = \sum_n D^n(x^{a_0}y^{b_0}) \frac{z^n}{n!},$$

or, on peut noter que :

$$\partial_t(X(t)^{a_0}Y(t)^{b_0}) = D(x^{a_0}y^{b_0})(X(t), Y(t)).$$

Par conséquent, on peut réécrire de la manière suivante :

$$H(X(t), Y(t), z) = \sum_n \partial_t^n(X(t)^{a_0}Y(t)^{b_0}) \frac{z^n}{n!}.$$

Grâce au théorème de Cauchy-Kovalevskaya, on sait que X et Y sont analytiques au voisinage de l'origine, on peut donc écrire, en utilisant la formule de Taylor,

$$H(X(t), Y(t), z) = X(z+t)^{a_0} Y(z+t)^{b_0},$$

d'où le résultat annoncé pour $t = 0$. □

Exemple de l'urne d'Ehrenfest [-1,1,1,-1]

L'urne d'Ehrenfest a été introduite en 1907 pour illustrer certains paradoxes de la mécanique statistique naissante. Le système différentiel associé à cette urne est :

$$\begin{cases} \dot{X} = Y \\ \dot{Y} = X, \end{cases}$$

qui s'intègre de manière immédiate en

$$\begin{cases} X(t) = xcht + ysht \\ Y(t) = xsht + ycht. \end{cases}$$

La série génératrice associée à notre urne est donc :

$$H(x, y, z) = (xchz + yshz)^N,$$

par conséquent $h_n(N, 0) = \frac{1}{2^N} \sum_{k=0}^N \binom{N}{k} (N-2k)^n$, par développement en séries entières. Sachant que $h_n = N^n$, on en déduit que la probabilité pour que tout le gaz retourne dans la première moitié de la chambre au temps n vaut

$$\frac{1}{N^n 2^N} \sum_{k=0}^N \binom{N}{k} (N-2k)^n.$$

2.3 Exemples d'application à des urnes aléatoires simples

L'urne de Pólya

On considère une population d'individus chez lesquels il existe deux versions d'un gène, A et B. Chaque individu possède l'un de ces allèles. On se demande comment va évoluer la proportion de personnes possédant l'allèle A. Ceci est l'interprétation la plus courante de l'urne que nous allons regarder maintenant : l'urne de Pólya.

Le principe de cette urne est très simple : quand on tire une boule, on en ajoute une de la même couleur. On la note donc $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

L'urne est équilibrée car $\alpha + \beta = \gamma + \delta = \sigma = 1$, on a donc en particulier :

$$h_n = \frac{\Gamma(s_0 + n)}{\Gamma(s_0)} = n! \binom{s_0 + n - 1}{s_0 - 1}.$$

On résout ensuite le système différentiel associé à l'urne :

$$\begin{cases} \dot{X} = X^2 \\ \dot{Y} = Y^2 \\ X(0) = x \text{ et } Y(0) = y. \end{cases}$$

Ce système s'intègre de la manière suivante :

$$\frac{\dot{X}(t)}{X(t)^2} = 1 \text{ et } \frac{\dot{Y}(t)}{Y(t)^2} = 1,$$

donc on obtient :

$$\begin{cases} X(t) = \frac{x}{1-xt} \\ Y(t) = \frac{y}{1-yt}, \end{cases}$$

d'où on déduit l'expression de la fonction génératrice associée à l'urne :

$$H(x, y, z) = \left(\frac{x}{1-zx} \right)^{a_0} \left(\frac{y}{1-zy} \right)^{b_0}.$$

Calculer le développement en série entière n'est pas forcément aisé alors nous allons nous aider de l'opérateur différentiel. Ici, il vaut : $D = x^2\partial_x + y^2\partial_y$, on a donc :

$$\begin{aligned} D(x^a y^b) &= ax^{a+1}y^b + bx^a y^{b+1} \\ &= x^a y^b (ax + by) \\ &= 1! \binom{a}{a-1} \binom{b-1}{b-1} x^{a+1} y^b + \binom{a-1}{a-1} \binom{b}{b-1} x^a y^{b+1} \end{aligned}$$

De la même manière, on peut calculer la seconde itération de D :

$$\begin{aligned} D^2(x^a y^b) &= x^a y^b (ax + by)(ax + by) + x^a y^b (ax^2 + by^2) \\ &= x^a y^b (a(a+1)x^2 + b(b+1)y^2 + 2abxy) \\ &= 2! x^a y^b \left(\frac{a(a+1)}{2} x^2 + \frac{b(b+1)}{2} y^2 + abxy \right) \\ &= 2! \sum_{i=0}^2 \binom{a-1+i}{a-1} \binom{b-1+(2-i)}{b-1} x^{a+i} y^{b+2-i} \end{aligned}$$

On obtient aisément, par récurrence sur n ,

$$D^n(x^{a_0} y^{b_0}) = n! \sum_{i=0}^n \binom{a_0-1+i}{a_0-1} \binom{b_0-1+(n-i)}{b_0-1} x^{a_0+i} y^{b_0+n-i},$$

ce qui permet de calculer les $h_n(a, b)$, sachant que $D^n(x^{a_0} y^{b_0}) = \sum_{a,b} h_n(a, b) x^a y^b$:

$$h_n(a, b) = \begin{cases} n! \binom{a-1}{a_0-1} \binom{b-1}{b_0-1} & \text{si } a_0 \leq a \leq a_0 + n \text{ et si } b = n - a + s_0 \\ 0 & \text{sinon.} \end{cases}$$

Ainsi, nous pouvons dire que

$$\mathbb{P}(A_n = a, B_n = b) = \begin{cases} \frac{\binom{a-1}{a_0-1} \binom{b-1}{b_0-1}}{\binom{n+s_0-1}{s_0-1}} & \text{si } a_0 \leq a \leq a_0 + n \text{ et si } b = n - a + s_0 \\ 0 & \text{sinon.} \end{cases}$$

Ce résultat peut également être obtenu simplement par des méthodes directes, mais nous utilisons ici un raisonnement qui peut être reproduit de façon mécanique pour une grande classe d'urnes.

FIGURE 6.2 – Évolution du nombre de boules noires au cours de 30 simulations d'une urne de Pólya contenant une boule de chaque couleur au départ.

L'urne des records

L'exemple par les éléments que nous avons mis en place Nous allons maintenant nous intéresser à l'urne des records. Cette urne est définie de la manière suivante : on ajoute toujours une boule blanche quelle que soit la boule tirée. On la note par conséquent $\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$. On peut remarquer que cette urne est équilibrée : $\alpha + \beta = \gamma + \delta = \sigma = 1$. On résout pour commencer le système différentiel associé :

$$\begin{cases} \dot{X} = X^2 \\ \dot{Y} = XY \\ X(0) = x \text{ et } Y(0) = y, \end{cases}$$

de la manière suivante :

$$\begin{cases} X(t) = \frac{x}{1-xt} \\ Y(t) = \frac{y}{1-xt}. \end{cases}$$

Nous pouvons alors calculer la fonction génératrice associée à l'urne :

$$H(x, y, z) = \left(\frac{x}{1-zx} \right)^{a_0} \left(\frac{y}{1-zx} \right)^{b_0} = \frac{x^{a_0} y^{b_0}}{(1-zx)^{s_0}},$$

de plus, comme $\sigma = 1 > 0$, on a $H(z) = \frac{1}{(1-z)^{s_0}}$.

Nous supposons désormais que nous commençons avec une seule boule noire, donc $a_0 = 0$ et $s_0 = b_0 = 1$. Ainsi, on a $h_n = n!$, de plus :

$$H(x, y, z) = \frac{y}{1-zx} = y \frac{1}{1-zx} = y \sum_n (zx)^n = y \sum_n x^n n! \frac{z^n}{n!} = \sum_n x^n y n! \frac{z^n}{n!}.$$

On en déduit :

$$h_n(a, b) = \begin{cases} n! & \text{si } a = n \text{ et } b = 1 \\ 0 & \text{sinon.} \end{cases}$$

Par conséquent

$$\mathbb{P}(a_n = a, b_n = b) = \begin{cases} 1 & \text{si } a = n \text{ et } b = 1 \\ 0 & \text{sinon,} \end{cases}$$

comme attendu (l'urne est déterministe).

Performance du système L'utilisation de tous ces outils pour obtenir qu'à l'instant n , on a n boules blanches et une boule noire semble disproportionnée. On peut toutefois modifier notre système pour donner un résultat intéressant de combinatoire sur le nombre de records dans une permutation.

Soit τ une permutation de n éléments. Nous appelons un record de τ tout entier $j \in \{1, \dots, n\}$ qui vérifie $\forall i < j, \tau(i) < \tau(j)$, c'est-à-dire que $\tau(j)$ est plus grand que tous ses prédécesseurs. Nous noterons $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$ le nombre de permutation n ayant k record, ce nombre est également appelé le nombre de Stirling de première espèce.

On peut vérifier que les nombres de Stirling sont définies par la relation de récurrence suivante : pour $1 \leq k \leq n-1$

$$\left[\begin{smallmatrix} n+1 \\ k \end{smallmatrix} \right] = \left[\begin{smallmatrix} n \\ k-1 \end{smallmatrix} \right] - n \left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right],$$

avec $\left[\begin{smallmatrix} n \\ 0 \end{smallmatrix} \right] = 0$ et $\left[\begin{smallmatrix} n \\ n \end{smallmatrix} \right] = 1$.

La série génératrice associée aux nombres de Stirling est

$$\sum_{n=0}^{\infty} \left(\sum_{k=0}^n \left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right] w^k \right) \frac{z^n}{n!} = (1+z)^w.$$

Nous allons modifier légèrement le système différentiel associé à notre urne, en y ajoutant une variable w qui « indique qu'on a tiré une boule noire », de la manière suivante :

$$\begin{cases} \dot{X} = X^2 \\ \dot{Y} = wXYX(0) = x \text{ et } Y(0) = y. \end{cases}$$

On interprète l'histoire de l'urne ainsi. On suppose les boules blanches que nous rajouterons sont numérotées de n à 1. On place les boules blanches dans l'urne d'une manière particulière :

- si on tire une boule blanche, on place la nouvelle boule (blanche) juste à droite de celle-ci. Comme la suite des nombres portés par les boules blanches décroît, cette nouvelle boule n'est pas un record ;
- si on tire la boule noire, on place la boule blanche en première position, elle sera un nouveau record puisque toutes les autres boules ajoutées auront un nombre plus petit.

On obtient ainsi une la permutation au temps n , en regardant la suite des nombres de gauche à droite.

Lorsque l'on résout le système, on obtient :

$$\begin{cases} X(t) = \frac{x}{1-xt} \\ Y(t) = \frac{y}{(1-xt)^w}. \end{cases}$$

Ainsi, la fonction génératrice avec une boule noire et aucune boule blanche au début devient :

$$H_w(x, y, z) = \frac{y}{(1-zx)^w}.$$

Le coefficient en w^k marque que l'on a réalisé k choix de la boule noire, soit k records pour la permutation obtenue. On a bien :

$$[w^k x^n y z^n] H_w = \frac{1}{n!} \left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right].$$

On pourra également noter que le nombre C_n de choix de la boule noires à l'instant n est donnée par :

$$C_n = \sum_{i=1}^n B_i,$$

où les B_i sont des variables aléatoires indépendantes et suivent des lois de Bernoulli de paramètre $\frac{1}{i}$. L'espérance du nombre de records varie donc comme $\log n$ quand n tend vers $+\infty$.

Pour plus d'information, nous invitons le lecteur à lire [3].

3 Solution générale pour les urnes sacrificielles

3.1 Intégration du système différentiel

On peut observer que les solutions au système différentiel associé à une urne aléatoires se déplacent le long de certaines courbes intégrales dont la forme dépend du signe de la différence entre nombre de boules blanches ajoutées dans l'urne en fonction de la couleur de l'urne tirée.

Définition 3.1. Pour les urnes équilibrées, on pose $p = \gamma - \alpha = \beta - \delta$, on dit alors que

- si $p > 0$, l'urne est altruiste, la couleur choisie n'est pas favorisée, au sens où lorsqu'on tire une boule d'une couleur, on ajoute d'avantage de boules de l'autre couleur ;
- si $p = 0$, l'urne est neutre, cas peu intéressant, car alors la composition de l'urne est déterminée au cours du temps, comme dans le cas de l'urne des records ;
- si $p < 0$, l'urne est égoïste, car la couleur choisie est favorisée.

Proposition 3.1. Toute solution (X, Y) du système différentiel associé à l'urne $[\alpha, \beta, \gamma, \delta]$

$$\begin{cases} \dot{X} = X^{\alpha+1}Y^{\beta} \\ \dot{Y} = X^{\gamma}Y^{\delta+1} \end{cases}$$

vérifie $X(t)^p - Y(t)^p = X(0)^p - Y(0)^p$.

Preuve. On calcule simplement :

$$\begin{aligned} \partial_t(X(t)^p - Y(t)^p) &= p(X^{p-1}\dot{X} - Y^{p-1}\dot{Y}) \\ &= p(X^{\alpha+p}Y^{\beta} - X^{\gamma}Y^{\delta+p}) \\ &= 0. \end{aligned}$$

□

Exemple 3.1. Cette égalité est bien vérifiée dans le cas de l'urne d'Ehrenfest, on a

$$(xcht + ysht)^2 - (ycht + xsht)^2 = x^2 - y^2.$$

3.2 Quantités homogènes

En remarquant que $H(x, y, z) = y^{\sigma_0} H\left(\frac{x}{y}, 1, zy^{\sigma}\right)$, on constate que la dépendance en y peut être ignorée. On peut donc supposer $y = 1$, on cherche alors une classe de fonctions permettant d'exprimer les fonctions génératrices associées aux urnes aléatoires concernées dans ce cas particulier. Lorsque l'urne est dite sacrificielle, nous allons exhiber une telle classe de fonctions.

Définition 3.2. Une urne sacrificielle est une urne vérifiant $\alpha \leq -1$, les boules blanches sont sacrifiées au sens où lorsqu'on en pioche une, leur population diminue.

Nous fixons maintenant une urne sacrificielle $[\alpha, \beta, \gamma, \delta]$, rappelons que $\sigma = \alpha + \beta$ et $p = \gamma - \alpha > 0$.

Notation 3.1. Pour $x \in \mathbb{C}$ voisin de 0, on pose :

$$\Delta(x) = (1 - x^p)^{\frac{1}{p}}$$

en utilisant une détermination holomorphe de la racine p ème sur ce voisinage.

On pose également, pour (X, Y) solution du système différentiel associé à l'urne considérée, avec pour conditions initiales $(x_0, 1)$ les fonctions ξ et η définies comme suit :

$$\begin{cases} \xi(t) = \frac{X\left(\frac{t}{\Delta(x_0)^{\sigma}}\right)}{\Delta(x_0)} \\ \eta(t) = \frac{Y\left(\frac{t}{\Delta(x_0)^{\sigma}}\right)}{\Delta(x_0)}. \end{cases}$$

Proposition 3.2. Les fonctions ξ et η définies ci-dessus vérifient les équations suivantes :

$$\begin{cases} \dot{\xi} = \xi^{\alpha+1}(\xi^p + 1)^{\frac{\beta}{p}} \\ \dot{\eta} = \eta^{\delta+1}(\eta^p - 1)^{\frac{\gamma}{p}}, \end{cases}$$

avec pour conditions initiales $\xi(0) = \frac{x_0}{\Delta(x_0)}$ et $\eta(0) = \frac{1}{\Delta(x_0)}$.

Preuve. On a par définition de (X, Y) :

$$\dot{X} = X^{\alpha+1}Y^{\beta}.$$

Or on sait que $Y^p - X^p = 1 - x_0^p = \Delta(x_0)^p$, par conséquent, on peut exprimer \dot{X} en fonction de X :

$$\dot{X} = X^{\alpha+1}(X^p + \Delta(x_0)^p)^{\frac{\beta}{p}}.$$

Dès lors, en utilisant que $\sigma = \alpha + \beta$, on a bien :

$$\dot{\xi} = \xi^{\alpha+1}(\xi^p + 1)^{\frac{\beta}{p}}.$$

L'égalité pour η s'obtient de la même manière. □

On peut alors écrire, en notant $\dot{\xi} = \frac{d\xi}{dt}$ et en intégrant :

$$t = \int_{\frac{x_0}{\Delta(x_0)}}^{\xi(t)} \frac{w^{-\alpha-1}}{(1+w^p)^{\frac{\beta}{p}}} dw,$$

puis par le changement de variables $\zeta = w^{-\alpha}$ ($\alpha < -1$) :

$$t = \frac{1}{\alpha} \int_{\left(\frac{x_0}{\Delta(x_0)}\right)^{-\alpha}}^{\xi(t)^{-\alpha}} \frac{d\zeta}{(1+\zeta^{\frac{-p}{\alpha}})^{\frac{\beta}{p}}}.$$

Définition 3.3. On définit, pour $\lambda \in \mathbb{Q}^+$, $r \in \mathbb{N}^*$ les fonctions suivantes :

- la fonction pseudo-arcsinus par la formule intégrale :

$$J_{r,\lambda}(u) = \int_0^u \frac{d\xi}{(1+\xi^r)^\lambda},$$

- la fonction pseudo-sinus comme inverse local de $J_{r,\lambda}$ dans un voisinage complexe de 0 :

$$S_{r,\lambda}(J_{r,\lambda}(u)) = J_{r,\lambda}(S_{r,\lambda}(u)) = u,$$

- la fonction pseudo-cosinus est définie pour $s \in \mathbb{Q}^*$ dans un voisinage complexe de 0 par :

$$C_{r,s,\lambda}(u) = (1 + S_{r,\lambda}(u)^r)^{\frac{1}{s}}.$$

Ces fonctions sont bien définies, en effet J est analytique dans $B(0, \frac{1}{2})$ dans lequel on a une détermination analytique de $\ln(1+z^r)$ donc de $\frac{1}{(1+z^r)^\lambda}$. J est donc définie dans ce voisinage simplement connexe de 0 comme primitive d'une fonction analytique, et est de l'ordre de $u + O(u^{r+1})$ au voisinage de 0.

Par conséquent, J est inversible dans un voisinage de 0, et S est bien définie, et de l'ordre de $u + O(u^{r+1})$. C est alors définie dans un voisinage de 0 dans lequel $1 + S^r$ ne s'annule pas, est analytique et est de l'ordre de $1 + O(z^r)$.

Théorème 3.1. Soit une urne équilibrée $[\alpha, \beta, \gamma, \delta]$ avec $\alpha < 0$ et $\delta \neq 0$, on pose :

$$\lambda = \frac{\beta}{p}, \quad r = \frac{-p}{\alpha}, \quad s = \frac{-p}{\delta}.$$

On note également les fonctions

$$J = J_{r,\lambda}, \quad S = S_{r,\lambda}, \quad C = C_{r,s,\lambda},$$

on peut alors exprimer H par la formule suivante :

$$H(x, 1, z) = \Delta(x)^{s_0} S(-\alpha z \Delta(x)^\sigma + J(x^{-\alpha} \Delta(x)^\alpha)) \frac{-a_0}{\alpha} C(-\alpha z \Delta(x)^\sigma + J(x^{-\alpha} \Delta(x)^\alpha)) \frac{-b_0}{\delta}.$$

Preuve. On a déjà vu que :

$$t = \frac{1}{\alpha} \int_{\left(\frac{x}{\Delta(x)}\right)^{-\alpha}}^{\xi(t)^{-\alpha}} \frac{d\zeta}{(1 + \zeta^r)^\lambda},$$

soit, avec les notations précédentes,

$$-\alpha t = J(\xi(t)^{-\alpha}) - J(x^{-\alpha} \Delta(x)^\alpha).$$

Par conséquent, on a

$$\xi^{-\alpha} = S(-\alpha t + J(x^{-\alpha} \Delta(x)^\alpha)).$$

En utilisant le fait que $\xi^p - \eta^p = 1$, on obtient pour $\eta^{-\delta}$ l'expression :

$$C(-\alpha t + J(x^{-\alpha} \Delta(x)^\alpha)).$$

On a alors l'expression voulue pour H . □

Exemple 3.2. L'utilisation de ces fonctions pour l'urne d'Erhenferst donne :

$$J(u) = J_{2, \frac{1}{2}}(u) = \operatorname{argsh}(u).$$

Par conséquent, on a :

$$S(u) = \operatorname{sh}(u) \text{ et } C(u) = \operatorname{ch}(u),$$

ce qui nous donne :

$$H(x, 1, z) = (1 - x^2)^{\frac{N}{2}} \operatorname{sh}^N \left(z + \operatorname{argsh} \left(\frac{x}{\sqrt{1 - x^2}} \right) \right),$$

qui est bien égal au résultat obtenu précédemment :

$$H(x, 1, z) = (x \operatorname{ch}(z) + \operatorname{sh}(z))^N.$$

4 Urnes semi-sacrificielles

Définition 4.1. Une urne semi-sacrificielle est une urne $[\alpha, \beta, \gamma, \delta]$ telle que lorsqu'on tire une boule blanche, on enlève des boules blanches, mais lorsqu'on tire une boule noire, on en ajoute, i.e. $\alpha < 0, \delta > 0, \beta\gamma \neq 0$ et donc $\sigma > 0$ (car β et γ sont positifs).

4.1 Propriétés analytiques des fonctions de base

La forme particulière des urnes semi-sacrificielles permet d'obtenir certains résultats analytiques sur les fonctions pseudo-sinus et pseudo-cosinus associées, en particulier sur leurs rayons de convergence, qui permettront par la suite d'obtenir certains résultats probabilistes pour les urnes.

Lemme 4.1. *Pour une urne semi-sacrificielle, le rayon de convergence commun ρ des fonctions C et S définies précédemment est donné par :*

$$\rho = \frac{1}{r} \frac{\Gamma\left(\frac{1}{r}\right) \Gamma\left(1 - \frac{1}{r} - \frac{1}{s}\right)}{\Gamma\left(\frac{1}{s}\right)}.$$

Preuve. On remarque que grâce au théorème 2 appliqué à l'urne, avec pour conditions initiales $a_0 = -\alpha$ et $b_0 = 0$, on a

$$H(0, 1, z) = S(-\alpha z).$$

Par conséquent, les termes du développement de Taylor de S en 0 sont strictement positifs, il suffit donc, pour déterminer son rayon de convergence d'étudier S sur \mathbb{R}^+ .

Or S est définie comme l'inverse de J au voisinage de 0, et J est analytique en tout point de \mathbb{R}^+ . Il suffit de trouver dans quel domaine J est injectif pour en déduire le domaine dans lequel S est analytique. Or $J'(u) = (1 + u^r)^{-\lambda} > 0$. Par conséquent, on en déduit $\rho = J(+\infty) = \int_0^{+\infty} \frac{du}{(1 + u^r)^\lambda}$, d'où l'expression recherchée, grâce aux résultats sur la fonction bêta d'Euler.

Montrons ensuite que $S(z)^r \neq -1$, pour $|z| < \rho$. En effet, par formule de la dérivée de l'inverse, S est solution de l'équation différentielle :

$$f' = (1 + f^r)^\lambda.$$

Mais si f solution de l'équation précédente était analytique au voisinage de z_0 vérifiant $f(z_0)^r = -1$, on aurait :

$$f \sim a + ba(z - z_0)^n, \quad f' \sim nba(z - z_0)^{n-1} \text{ et } (1 + f^r)^\lambda \sim (-rb)^\lambda (z - z_0)^{n\lambda},$$

avec $n > 1$, $a^r = 1$. L'égalité est alors impossible, car on aurait $\lambda n = n - 1$ avec $\lambda > 1$.

Par conséquent, $S^r \neq -1$ sur le disque de convergence, donc C est analytique dans ce domaine. \square

Nous pouvons maintenant étudier ensuite le comportement de S et C au voisinage de ρ .

Lemme 4.2. *Posons $r_0 = r\lambda - 1$ et $r_1 = r(\lambda + 1) - 1$, on a, au voisinage de ρ :*

$$S(z) = (r_0(\rho - z))^{-\frac{1}{r_0}} \left[1 - \frac{\lambda}{r_1} (r_0(\rho - z))^{\frac{r}{r_0}} + O\left((\rho - z)^{\frac{2r}{r_0}}\right) \right],$$

$$C(z) = (r_0(\rho - z))^{-\frac{r}{s r_0}} \left[1 + \frac{1}{s} \left(1 - \frac{r\lambda}{r_1} \right) (r_0(\rho - z))^{\frac{r}{r_0}} + O\left((\rho - z)^{\frac{2r}{r_0}}\right) \right].$$

Preuve. On commence par donner un développement asymptotique de $J(u)$ pour $u \rightarrow +\infty$, grâce à son expression comme intégrale de $(1 + \xi^r)^{-\lambda}$:

$$J(u) - \rho = - \int_u^\infty \frac{d\xi}{(1 + \xi^r)^\lambda}.$$

Or on sait que :

$$(1 + \xi^r)^{-\lambda} = \xi^{-\lambda r} \sum_{k=0}^{+\infty} \binom{-\lambda}{k} \xi^{-kr},$$

d'où, en échangeant série et intégrale, et en étudiant le développement à l'ordre 2 :

$$J(u) = \rho - \frac{1}{r_0 u^{r_0}} + \frac{1}{r_1 u^{r_1}} + 0 \left(\frac{1}{u^{r(\lambda+2)-1}} \right),$$

or S étant défini comme l'inverse de J , on peut en déduire un équivalent de S pour $z \rightarrow \rho$. De plus, $C(z) = (1 + S(z)^r)^{\frac{1}{\sigma}}$, on obtient son développement par composition. \square

Lemme 4.3. *On note $\rho(x)$ le rayon de convergence de la série $H(x, 1, z)$ vu comme une fonction de z . Il existe un voisinage complexe de 1 pour x dans lequel ce rayon de convergence est analytique et vérifie :*

$$\rho(x) = x^{-\sigma} \sum_{k=0}^{+\infty} \binom{-\lambda}{k} \frac{1}{\sigma + kp} \left(\frac{1 - x^p}{x^p} \right)^k.$$

Preuve. Connaissant le rayon de convergence ρ de S et C , et l'expression de H , le rayon de convergence $\rho(x)$ est déterminé par la relation :

$$-\alpha \rho(x) \Delta(x)^\sigma + J(x^{-\alpha} \Delta(x)^\alpha) = \rho,$$

or on a :

$$\rho - J(u) = \int_u^\infty \frac{d\xi}{(1 + \xi^r)^\lambda}.$$

Par conséquent, on obtient, en posant $e(x) = \frac{\Delta(x)}{x}$:

$$\rho(x) = x^{-\sigma} \frac{-1}{\alpha e(x)^\sigma} \int_{e(x)^\alpha}^\infty \frac{d\xi}{(1 + \xi^r)^\lambda},$$

donc, par changement de variable $\xi = v^\alpha$ ($\alpha \leq -1$) on obtient :

$$\begin{aligned} \rho(x) &= x^{-\sigma} e(x)^{-\sigma} \int_0^{e(x)} \frac{v^{\alpha-1} dv}{(1 + v^{-p})^\lambda} \\ &= x^{-\sigma} e(x)^{-\sigma} \int_0^{e(x)} \frac{v^{\alpha-1+\lambda p} dv}{(1 + v^p)^\lambda}. \end{aligned}$$

Dès lors, par développement en séries entières, on a :

$$\rho(x) = x^{-\sigma} e(x)^{-\sigma} \sum_{k=0}^{+\infty} \binom{-\lambda}{k} \int_0^{e(x)} v^{\alpha+\beta+kp-1},$$

et donc en intégrant, on obtient :

$$\rho(x) = x^{-\sigma} \sum_{k=0}^{+\infty} \frac{e(x)^{kp}}{\sigma + kp}.$$

Ce qui donne l'égalité désirée. \square

4.2 Propriétés probabilistes des urnes semi-sacrificielles

Grâce à tous les outils développés précédemment, nous sommes maintenant prêts à démontrer un théorème de type théorème de la limite centrale pour les urnes semi-sacrificielles.

Théorème 4.1. *Soit une urne semi-sacrificielle, et a_n le nombre de boules blanches présentes à l'instant n dans l'urne. On a alors*

$$\frac{a_n}{n} \xrightarrow[n \rightarrow +\infty]{} \mu \text{ p.s. et}$$

$$\frac{a_n - n\mu}{\sqrt{n}} \xrightarrow[n \rightarrow +\infty]{} N(0, \xi^2),$$

où on a posé $\mu = \frac{\sigma\gamma}{\beta + \gamma}$ et $\xi^2 = \frac{\beta\gamma\sigma}{-\alpha + \beta + 2\gamma} \left(\frac{\alpha - \gamma}{\beta + \gamma} \right)^2$.

Pour démontrer ce théorème, nous aurons besoin d'un lemme d'analyse fonctionnelle sur les séries génératrices nous donnant le comportement des coefficients d'une série entière en fonction de son pôle en 1 :

Lemme 4.4. *Soit f une fonction holomorphe au voisinage du domaine*

$$\Delta = \Delta(\eta, \phi) = \{z \in \mathbb{C} \mid |z| \leq 1 + \eta, |\text{Arg}(z - 1)| \geq \phi\},$$

sauf en 1 au voisinage duquel il existe $a \in \mathbb{R}$ tel que $f = O(|z - 1|^a)$.

On a alors $[z^n]f(z) = O(n^{-a-1})$.

Preuve. Comme $(1 - z)^a$ est de module minoré par une quantité strictement positive sur tout compact de Δ ne contenant pas 1, la condition d'existence d'un voisinage de 1 dans Δ et $K > 0$ tel que $f(z) < K(1 - z)^{-a-1}$ est équivalente à l'existence d'une constante K tel que $f(z) < K(1 - z)^{-a-1}$ sur tout $\Delta \setminus \{1\}$.

Soit $n > 2|a| + 4$, on pose $f_n = [z^n]f$, on a alors par la formule de Cauchy :

$$f_n = \frac{1}{2i\pi} \int_o f(z) \frac{dz}{z^{n+1}},$$

où o est un contour $\partial(\Delta \setminus D(1, \varepsilon))^+$.

On calcule alors l'intégrale sur chacun des contours :

- sur le petit cercle, pour $n \geq 4$ on obtient un majorant de l'intégrale en

$$\frac{1}{2\pi} K \left(\frac{1}{n} \right)^a \left(1 - \frac{1}{n} \right)^{-n-1} \left(\frac{2\pi i}{n} \right),$$

soit une majoration en $5Kn^{-a-1}$;

- sur la partie rectiligne, on pose $\omega = e^{i\phi}$ et on réalise le changement de variable $z \rightarrow 1 + \frac{\omega t}{n}$, on obtient une majoration du terme en :

$$\int_1^{En} K \left(\frac{t}{n} \right)^a \left| 1 + \frac{\omega t}{n} \right|^{-n-1} \frac{dt}{n},$$

soit, après majoration, on en tire encore une majoration en cKn^{-a-1} ;

- sur le grand cercle, on obtient une majoration exponentielle décroissante en :

$$K \frac{\eta^a}{(1 + \eta)^n}.$$

En sommant ces termes on obtient bien le résultat demandé. \square

Le lemme suivant permet d'obtenir un résultat similaire.

Lemme 4.5. *Soit f une fonction holomorphe au voisinage du domaine*

$$\Delta = \Delta(\eta, \phi) = \{z \in \mathbb{C} \mid |z| \leq 1 + \eta, |\text{Arg}(z - 1)| \geq \phi\},$$

sauf en 1 au voisinage duquel il existe $a \in \mathbb{R}$ tel que $f = o(|z - 1|^a)$.

Dans ce cas, on a $[z^n]f(z) = o(n^{-a-1})$.

Preuve. On raisonne de la même manière que précédemment, la preuve est donnée en détail dans [5]. \square

Corollaire 4.2. *Si f est analytique au voisinage de $\Delta \setminus \{1\}$, avec pour équivalent en 1 $K(1 - z)^a$, on a alors*

$$[z^n]f(z) \sim \frac{K}{\Gamma(-a)} n^{-a-1}.$$

Preuve. On applique la remarque précédente à $f - K(1 - z)^a$, le reste provient du développement en séries entières de $K(1 - z)^a$. \square

du Théorème 4.1. On a, grâce à l'expression de H à partir des fonctions S et C , et à leurs propriétés :

$$H(x, 1, z) \underset{z \rightarrow \rho(x)}{\sim} \sigma^{-\frac{s_0}{\sigma}} (\rho(x) - z)^{-\frac{s_0}{\sigma}}.$$

C'est également la seule singularité jouant vraiment un rôle pour x au voisinage de 1, car au voisinage des autres singularités, la fonction tend vers l'infini. Or, on sait que :

$$[z^n](\rho(x) - z)^{-\frac{s_0}{\sigma}} = \rho(x)^{-n - \frac{s_0}{\sigma}} \frac{\Gamma(\frac{s_0}{\sigma} + n)}{\Gamma(\frac{s_0}{\sigma})n!} \sim \frac{n^{\frac{s_0}{\sigma} - 1}}{\rho(x)^{n + \frac{s_0}{\sigma}} \Gamma(\frac{s_0}{\sigma})},$$

en particulier, pour x proche de 1, certains résultats pour les séries génératrices permettent alors d'affirmer que :

$$[z^n]H(x, 1, z) \sim \frac{\sigma^{-\frac{s_0}{\sigma}}}{\Gamma(\frac{s_0}{\sigma})} \rho(x)^{-n - \frac{s_0}{\sigma}} n^{\frac{s_0}{\sigma} - 1} \left[1 + O\left(n^{-\frac{p}{\sigma}}\right)\right].$$

Or on sait que l'on a, par définition :

$$\begin{aligned} H(x, 1, z) &= \sum_{n,a,b} h_n(a, b) x^a \frac{z^n}{n!} \\ &= \sum_n h_n \frac{z^n}{n!} \sum_a \mathbb{P}(a_n = a) x^a \\ &= \sum_n h_n \frac{z^n}{n!} \mathbb{E}(x^{a_n}). \end{aligned}$$

De plus, on sait que :

$$h_n = \sigma^n \frac{\Gamma(\frac{s_0}{\sigma} + n)}{\Gamma(\frac{s_0}{\sigma})}.$$

On peut donc calculer un équivalent de $\chi_n(x) = \mathbb{E}(x^{a_n})$, qui caractérise la loi de a_n :

$$\chi_n(x) \sim \left(\frac{\rho(1)}{\rho(x)} \right)^{n + \frac{s_0}{\sigma}} \left[1 + O\left(n^{-\frac{p}{\sigma}}\right) \right].$$

On peut alors reprendre la démonstration du théorème central limite. On étudie la fonction caractéristique de $\frac{a_n - n\mu}{\sqrt{n}}$:

$$\Phi_n(\lambda) = \mathbb{E}\left(e^{i\lambda \frac{a_n - n\mu}{\sqrt{n}}}\right) = \chi_n\left(e^{i\frac{\lambda}{\sqrt{n}}}\right) e^{-i\sqrt{n}\mu\lambda},$$

et on cherche un équivalent de Φ_n quand n tend vers $+\infty$:

$$\Phi_n(\lambda) \sim \left[e^{-i\frac{\mu\lambda}{\sqrt{n}}} \left(\frac{\rho(1)}{\rho\left(e^{i\frac{\lambda}{\sqrt{n}}}\right)} \right) \right]^n \left(\frac{\rho(1)}{\rho\left(e^{i\frac{\lambda}{\sqrt{n}}}\right)} \right)^{\frac{s_0}{n\sigma}} \left[1 + O\left(n^{-\frac{p}{\sigma}}\right) \right]^{\frac{1}{n}}.$$

Cherchons tout d'abord le développement asymptotique de :

$$\frac{1 - e^{ip\frac{\lambda}{\sqrt{n}}}}{e^{ip\frac{\lambda}{\sqrt{n}}}} = -ip\frac{\lambda}{\sqrt{n}} - p^2\frac{\lambda^2}{2n} + O\left(\frac{1}{n\sqrt{n}}\right),$$

puis celui de $\rho\left(e^{i\frac{\lambda}{\sqrt{n}}}\right)$:

$$\frac{1}{\sigma} \left[1 - \frac{i\sigma\lambda}{\sqrt{n}} \frac{\gamma}{\beta + \gamma} + \frac{\lambda^2}{2n} \left(\sigma^2 \frac{\beta - \gamma}{\beta + \gamma} + p^2 \left(\frac{\beta}{(\sigma + p)(\sigma + 2p)} - \frac{\lambda^2}{\sigma + 2p} \right) \right) \right] + O\left(\frac{1}{n\sqrt{n}}\right).$$

On pose alors $\mu = \frac{\sigma\gamma}{\beta + \gamma}$, on obtient :

$$\frac{\rho(1)}{\rho\left(e^{i\frac{\lambda}{\sqrt{n}}}\right)} = 1 + i\mu\frac{\lambda}{\sqrt{n}} - \frac{\lambda^2}{2n} \left[\frac{\sigma^2(\beta^2 + \gamma^2)}{(\beta + \gamma)^2} + p^2 \left(\frac{\beta}{(\sigma + p)(\sigma + 2p)} - \frac{\lambda^2}{\sigma + 2p} \right) \right] + O\left(\frac{1}{n\sqrt{n}}\right),$$

puis en développant

$$e^{-i\frac{\mu\lambda}{\sqrt{n}}} \left(\frac{\rho(1)}{\rho\left(e^{i\frac{\lambda}{\sqrt{n}}}\right)} \right) = 1 - \frac{\lambda^2 \xi^2}{2n} + O\left(\frac{1}{n\sqrt{n}}\right),$$

où on a posé :

$$\xi^2 = \frac{\beta\gamma}{-\alpha + \beta + 2\gamma} \left(\frac{\alpha - \gamma}{\beta + \gamma} \right)^2 \sigma.$$

FIGURE 6.3 – Proportion de boules blanches dans l’urne $[-1, 4, 2, 1]$ partant de 100 boules blanches au cours de 250 étapes.

Par conséquent, on en tire :

$$\lim_{n \rightarrow +\infty} \Phi_n(\lambda) = e^{-\frac{\lambda^2 \xi^2}{2}}$$

qui est la fonction caractéristique de la loi gaussienne centrée de paramètre ξ^2 , le théorème de Lévy permet d’en déduire la convergence en loi de la suite. \square

Remarque 4.1. On notera que pour toute urne semi-sacrificielle, quelle que soit la composition de départ de l’urne, la proportion de boules blanches $\frac{a_n}{s_0+n\sigma}$ converge presque sûrement vers la quantité $\frac{\gamma}{\beta+\gamma}$. On obtient un résultat de convergence classique pour les écarts à la moyenne.

On notera également que μ est un rationnel strictement inférieur à $\frac{1}{2}$.

Exemple 4.1. La proportion de boules blanches dans l’urne $\begin{pmatrix} -1 & 4 \\ 2 & 1 \end{pmatrix}$ tend presque sûrement vers $\frac{1}{3}$.

5 Urnes triangulaires

Définition 5.1. Une urne $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ est dite triangulaire si elle vérifie $\gamma = 0$.

Elle est alors de la forme $\begin{pmatrix} \alpha & \sigma - \alpha \\ 0 & \sigma \end{pmatrix}$ puisqu’on la suppose toujours équilibrée. On suppose de plus $0 < \alpha < \sigma$.

Remarque 5.1. On impose la condition $0 < \alpha < \sigma$. En effet,

- le cas $\alpha = 0$ nous donne une urne $[0, \sigma, 0, \sigma]$, et en supposant que σ divise a_0 et b_0 , on se ramène au cas de l’urne des records, déjà traitée,
- le cas $\alpha = \sigma$ se ramène de la même manière au cas de l’urne de Pólya.

5.1 Résultats de la théorie des martingales

On peut ici utiliser la théorie des martingales pour avoir une idée de l’ordre de grandeur de a_n .

Proposition 5.1. La variable aléatoire $\frac{a_n}{n^{\frac{\alpha}{\sigma}}}$ converge presque sûrement vers une variable aléatoire dans L^1 .

Preuve. On définit la filtration $F_n = \sigma(a_0, \dots, a_n)$, et on calcule :

$$\begin{aligned}\mathbb{E}(a_{n+1}|F_n) &= \frac{a_n}{s_0 + n\sigma}(a_n + \alpha) + \left(1 - \frac{a_n}{s_0 + n\sigma}\right)a_n \\ &= a_n \left(1 + \frac{\alpha}{s_0 + n\sigma}\right).\end{aligned}$$

On pose alors :

$$X_n = \frac{a_n}{\prod_{k=0}^n \left(1 + \frac{\alpha}{s_0 + k\sigma}\right)}.$$

C'est une martingale positive finie presque sûrement, qui converge donc p.s vers $X \in L^1$, or

$$\prod_{k=0}^n \left(1 + \frac{\alpha}{s_0 + k\sigma}\right) \sim Cn^{\frac{\alpha}{\sigma}}.$$

Par conséquent, on en déduit que $\frac{a_n}{n^{\frac{\alpha}{\sigma}}}$ converge p.s. vers $d \in L^1$. On en déduit que presque sûrement, a_n est de l'ordre de $n^{\frac{\alpha}{\sigma}}$. \square

On essaiera ici de déterminer grâce à nos outils analytiques quelle est la loi de la variable aléatoire d vers laquelle converge $\frac{a_n}{n^{\frac{\alpha}{\sigma}}}$.

5.2 Fonctions génératrices

Commençons par déterminer quelle est la série génératrice associée à l'urne triangulaire $\begin{pmatrix} \alpha & \sigma - \alpha \\ 0 & \sigma \end{pmatrix}$.

Proposition 5.2. *On a*

$$H(x, 1, z) = x^{a_0}(1 - \sigma z)^{-\frac{b_0}{\sigma}} \left(1 - x^\alpha \left(1 - (1 - \sigma z)^{\frac{\alpha}{\sigma}}\right)\right)^{-\frac{a_0}{\alpha}}.$$

Par conséquent, on a, pour $k \in \mathbb{N}$,

$$\mathbb{P}(a_n = a_0 + k\alpha) = \frac{\Gamma(n+1)\Gamma(\frac{s_0}{\sigma})}{\Gamma(\frac{s_0}{\sigma} + n)} \binom{k + \frac{a_0}{\alpha} - 1}{k} \sum_{i=0}^k (-1)^i \binom{k}{i} \binom{n + \frac{b_0 - \alpha i}{\sigma} - 1}{n}.$$

Preuve. Le système différentiel associé à une urne triangulaire est :

$$\begin{cases} \dot{X} = X^{\alpha+1}Y^{\sigma-\alpha} \\ \dot{Y} = Y^{\sigma+1}. \end{cases}$$

Une solution (X, Y) vérifiant $(X(0), Y(0)) = (x, y)$ vérifie :

$$Y(t) = y(1 - \sigma y^\sigma t)^{-\frac{1}{\sigma}},$$

dès lors on a :

$$\dot{X}X^{-\alpha-1} = Y^{\sigma-\alpha},$$

qui s'intègre en

$$X(t) = x \left(1 - x^\alpha \left(1 - (1 - \sigma y^\sigma t)^{\frac{\alpha}{\sigma}} \right) \right)^{-\frac{1}{\alpha}}.$$

On en déduit l'expression de H :

$$H(x, y, z) = x^{a_0} y^{b_0} \left(1 - x^\alpha \left(1 - (1 - \sigma y^\sigma z)^{\frac{\alpha}{\sigma}} \right) \right)^{-\frac{a_0}{\alpha}} (1 - \sigma y^\sigma z)^{-\frac{b_0}{\sigma}}.$$

On sait que a_n est de la forme $a_0 + k\alpha$, de par la forme triangulaire de l'urne. Par conséquent :

$$[x^{a_0+k\alpha}]H(x, 1, z) = \binom{k + \frac{a_0}{\alpha} - 1}{k} (1 - \sigma z)^{-\frac{b_0}{\sigma}} \left(1 - (1 - \sigma z)^{\frac{\alpha}{\sigma}} \right)^k.$$

Le résultat s'obtient alors par développement en séries entières de z de cette fonction. \square

Exemple 5.1. On étudie l'urne triangulaire $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$. Les résultats généraux sur les urnes triangulaires donnent, pour $a_0 = 1$ et $b_0 = 0$:

$$H(x, 1, z) = \frac{x}{(1 - x(1 - \sqrt{1 - 2z}))},$$

que l'on développe en séries entières pour obtenir le coefficient suivant devant x^k :

$$[x^k]H(x, 1, z) = (1 - \sqrt{1 - 2z})^{k-1}.$$

Or si on pose $\Phi(z) = 1 - \sqrt{1 - 2z}$, cette fonction vérifie :

$$\Phi = \frac{z}{2 - \Phi},$$

on peut donc calculer $[z^n]\Phi^{k-1}$ grâce à la formule d'inversion de Lagrange, qui donne :

$$[z^n]\Phi^{k-1} = \frac{k-1}{n!} \partial_t^{n-1} \left(\frac{t^{k-2}}{(2-t)^n} \right) \Big|_{t=0}.$$

En développant on obtient, pour $n \geq k-1$:

$$[z^n]\Phi^{k-1} = \frac{k-1}{n} \binom{2n-k}{n-1} 2^{-2n+k-1}.$$

De même on calcule le coefficient de $H(1, 1, z)$:

$$[z^n]H(1, 1, z) = \frac{1}{2^n} \binom{2n}{n},$$

FIGURE 6.4 – Nombre de boules blanches au cours de 10 simulations de l'urne $[1, 1, 0, 2]$ partant d'une boule blanche au départ. On voit nettement que les courbes sont de la forme $c\sqrt{x}$.

on en tire alors :

$$\mathbb{P}(a_n = k) = \frac{k-1}{n} 2^{k-1} \frac{\binom{2n-k}{n-1}}{\binom{2n}{n}}.$$

On en déduit que $\mathbb{P}(a_n = \xi\sqrt{n}) \sim \frac{\xi}{2\sqrt{n}} e^{-\frac{\xi^2}{4}}$, par conséquent $\frac{a_n}{\sqrt{n}}$ converge presque sûrement vers une variable aléatoire suivant une loi dite de Rayleigh, proportionnelle à $\sqrt{\mathbf{e}}$, avec \mathbf{e} variable aléatoire exponentielle.

Grâce à l'expression simple de H , on peut sans difficultés calculer des équivalents des différents moments de a_n , comme nous l'avions remarqué plus tôt.

Proposition 5.3. *La variable aléatoire a_n vérifie*

$$\mathbb{E}(a_n) = a_0 \frac{\Gamma(\frac{s_0}{\sigma})}{\Gamma(\frac{s_0+2\alpha}{\sigma})} n^{\frac{\alpha}{\sigma}} + O\left(n^{\frac{\alpha}{\sigma}-1}\right) \text{ et,}$$

$$\mathbb{V}\text{ar}(a_n) = a_0 \left[(a_0 + \alpha) \frac{\Gamma(\frac{s_0}{\sigma})}{\Gamma(\frac{s_0+\alpha}{\sigma})} - a_0 \left(\frac{\Gamma(\frac{s_0}{\sigma})}{\Gamma(\frac{s_0+2\alpha}{\sigma})} \right)^2 \right] n^{\frac{2\alpha}{\sigma}} + O\left(n^{\frac{\alpha}{\sigma}}\right).$$

Preuve. On sait que :

$$H(x, 1, z) = \sum_{n=0}^{+\infty} \left(\sum_a h_n(a, \sigma n - a) x^a \right) \frac{z^n}{n!}.$$

Par conséquent, on a :

$$\begin{aligned} [\partial_x H(x, 1, z)]_{x=1} &= \sum_{n=0}^{+\infty} \left(\sum_a a h_n(a, \sigma n - a) \right) \frac{z^n}{n!} \\ &= \sum_{n=0}^{+\infty} h_n \mathbb{E}(a_n) \frac{z^n}{n!}. \end{aligned}$$

Or, on connaît l'expression de $H(x, 1, z)$ et par le calcul, on obtient :

$$[\partial_x H(x, 1, z)]_{x=1} = a_0 (1 - \sigma z)^{-\frac{s_0+\alpha}{\sigma}},$$

et connaissant h_n , on développe sans difficulté en séries entières la fonction précédente pour obtenir :

$$\mathbb{E}(a_n) = a_0 \frac{\Gamma(\frac{s_0+\alpha}{\sigma} + n) \Gamma(\frac{s_0}{\sigma})}{\Gamma(\frac{s_0}{\sigma} + n) \Gamma(\frac{s_0+\alpha}{\sigma})}.$$

On a un équivalent du terme ainsi obtenu par la formule de Stirling :

$$\mathbb{E}(a_n) = a_0 \frac{\Gamma(\frac{s_0}{\sigma})}{\Gamma(\frac{s_0+\alpha}{\sigma})} n^{\frac{\alpha}{\sigma}} + O\left(n^{\frac{\alpha}{\sigma}-1}\right).$$

On obtient de la même manière la formule pour la variance en dérivant deux fois la fonction H . □

5.3 Loi de la variable aléatoire limite

Théorème 5.1. *On pose $\xi = \frac{k}{n^{\frac{\alpha}{\sigma}}}$, la variable aléatoire a_n vérifie :*

$$\mathbb{P}(a_n = a_0 + \alpha \xi n^{\frac{\alpha}{\sigma}}) = \frac{1}{n^{\frac{\alpha}{\sigma}}} \frac{\Gamma(\frac{s_0}{\sigma})}{\Gamma(\frac{a_0}{\alpha})} \xi^{\frac{a_0}{\alpha}-1} \sum_{k=0}^{+\infty} \frac{(-1)^k \xi^k}{\Gamma(\frac{b_0-k\alpha}{\sigma}) k!} + O\left(\frac{1}{n^{\frac{\alpha}{\sigma}}}\right).$$

En particulier, $\frac{a_n}{n^{\frac{\alpha}{\sigma}}}$ converge en loi vers une mesure absolument continue par rapport à la mesure de Lebesgue. On connaît également la loi limite, de densité par rapport à la mesure de Lebesgue donnée par :

$$\frac{\Gamma(\frac{s_0}{\sigma})}{\Gamma(\frac{a_0}{\alpha})} \xi^{\frac{a_0}{\alpha}-1} \sum_{k=0}^{+\infty} \frac{(-1)^k \xi^k}{\Gamma(\frac{b_0-k\alpha}{\sigma}) k!}.$$

Pour démontrer ce théorème, nous aurons besoin du lemme suivant sur la fonction Γ d'Euler :

Lemme 5.1 (Formule des compléments). *La fonction Γ d'Euler vérifie l'égalité suivante :*

$$\frac{1}{\Gamma(1-z)\Gamma(z)} = \frac{\sin(\pi z)}{\pi}.$$

Preuve. On utilise l'écriture de Γ suivante :

$$\frac{1}{\Gamma(z)} = \lim_{n \rightarrow +\infty} \frac{z(z+1)\dots(z+n)}{n^z n!}.$$

On peut alors écrire

$$\begin{aligned} \frac{1}{\Gamma(1-z)\Gamma(z)} &= \lim_{n \rightarrow +\infty} \frac{z(z+1)\dots(z+n)(1-z)(2-z)\dots(n+1-z)}{n^{z+1-z} (n!)^2} \\ &= \lim_{n \rightarrow +\infty} \frac{z(n+1-z)}{n(n!)^2} (1-z^2)(2^2-z^2)\dots(n^2-z^2) \\ &= \lim_{n \rightarrow +\infty} \frac{z(n+1-z)}{n} (1-z^2)\left(1-\frac{z^2}{2^2}\right)\dots\left(1-\frac{z^2}{n^2}\right). \end{aligned}$$

Or le produit infini $h(z) = \pi z \prod_{n=1}^{+\infty} \left(1 - \frac{z^2}{n^2}\right)$ est normalement convergent sur tout disque de centre 0, car la série $\sum_{n=1}^{+\infty} \frac{z^2}{n^2}$ l'est.

On calcule alors sa dérivée logarithmique, qui est, par théorème de Weierstrass :

$$\frac{h'(z)}{h(z)} = \frac{1}{z} + \sum_{n=1}^{+\infty} \frac{2z}{z^2 - n^2}.$$

Il se trouve que l'on peut calculer cette limite par développement en séries de Fourier de $f(t) = \cos(ut)$, $t \in [-\pi, \pi]$, pour $u \in \mathbb{R}$. La série est normalement convergente, et on a convergence ponctuelle :

$$\cos(ut) = \frac{\sin(u\pi)}{u\pi} + 2u \sin(u\pi) \sum_{n=1}^{+\infty} (-1)^n \frac{\cos(nt)}{\pi(u^2 - n^2)},$$

pour $t = \pi$, on a en particulier :

$$\pi \frac{\cos(\pi u)}{\sin(\pi u)} = \frac{1}{u} + \sum_{n=1}^{+\infty} \frac{2u}{u^2 - n^2}.$$

Par identification de deux fonctions entières égales sur la droite réelle, on a :

$$\frac{h'(z)}{h(z)} = \pi \cotan(\pi z) = \frac{\sin'(\pi z)}{\sin(\pi z)}.$$

De plus, $\lim_{z \rightarrow 0} \frac{h(z)}{z} = \lim_{z \rightarrow 0} \frac{\sin(\pi z)}{z} = \pi$, dès lors, on en déduit :

$$\frac{\sin(\pi z)}{\pi} = z \prod_{n=1}^{+\infty} \left(1 - \frac{z^2}{\pi^2 n^2}\right).$$

On a bien en définitive :

$$\frac{1}{\Gamma(1-z)\Gamma(z)} = \frac{\sin(\pi z)}{\pi}.$$

□

On utilise ce lemme pour calculer l'intégrale suivante :

Lemme 5.2. Soit $z \in \mathbb{C}$, on a

$$\frac{1}{\Gamma(z)} = \frac{1}{2i\pi} \int_h u^{-z} e^u du,$$

où h est un contour de Hankel, un chemin entourant l'axe réel négatif contenu dans l'union d'un cône strict issu de 0 et d'un disque centré en 0.

Preuve. Posons $h_{a,r}$ le contour parcouru dans le sens direct réunion des deux demi-droites $\{|z| > r, |\text{Arg}(z)| = \pi - a\}$ et du cercle de centre 0 et de rayon r . On suppose que h est inclus dans le domaine $\{\text{Arg}(z) \in]\pi - d, \pi + d[\cup D(0, r')\}$.

On applique le théorème de Cauchy dans l'ouvert simplement connexe $\mathbb{C} \setminus \mathbb{R}^-$ au contour constitué d'une partie de $h_{a,r}$, de h et du cercle de centre 0 et de rayon R , on a :

$$|(Re^{i\theta})^{-z} e^{Re^{i\theta}}| = O(R^b e^{-R \sin(d)}),$$

FIGURE 6.5 – Dessin du contour utilisé.

FIGURE 6.6 – Représentation du contour γ_R .

indépendamment de u , par conséquent en faisant tendre R vers $+\infty$, on obtient :

$$\frac{1}{2i\pi} \int_h u^{-z} e^u du = \frac{1}{2i\pi} \int_{h_{a,r}} u^{-z} e^u du.$$

Faisons maintenant tendre a vers 0, on note c_r le chemin constitué du cercle de centre 0 et de rayon r parcouru dans le sens trigonométrique en partant de $-r$, on a :

$$\frac{1}{2i\pi} \int_h u^{-z} e^u du = \frac{1}{2i\pi} \left[\int_{c_r} u^{-z} e^u du - e^{-i\pi z} \int_r^{+\infty} u^{-z} e^u du + e^{i\pi z} \int_r^{+\infty} u^{-z} e^u du \right].$$

On suppose alors $Re(1-z) > 0$, dans ce cas l'intégrale sur c_r tend vers 0 quand $r \rightarrow 0$, on obtient alors

$$\frac{1}{2i\pi} \int_h u^{-z} e^u du = \sin(\pi z) \Gamma(1-z) = \frac{1}{\Gamma(z)}$$

Par identification des deux fonctions holomorphes sur l'ensemble $\{Re(1-z) > 0\}$, on étend cette égalité à tout $z \in \mathbb{C}$. \square

du Théorème 5.1. On commence par développer H en séries entières de x , on obtient :

$$H(x, 1, z) = x^{a_0} (1 - \sigma z)^{\frac{b_0}{\sigma}} \sum_{k=0}^{+\infty} \binom{k + \frac{a_0}{\alpha} - 1}{k} x^{k\alpha} \left(1 - (1 - \sigma z)^{\frac{\alpha}{\sigma}}\right)^k.$$

On obtient alors, connaissant l'expression de $h_n = \sigma^n \frac{\Gamma(\frac{s_0}{\sigma} + n)}{\Gamma(\frac{s_0}{\sigma})}$:

$$\mathbb{P}(a_n = a_0 + k\alpha) = \frac{\Gamma(n+1) \Gamma(\frac{s_0}{\sigma}) \binom{k + \frac{a_0}{\alpha} - 1}{k} [z^n]}{\sigma^n \Gamma(\frac{s_0}{\sigma} + n)} \left((1 - \sigma z)^{\frac{b_0}{\sigma}} \left(1 - (1 - \sigma z)^{\frac{\alpha}{\sigma}}\right)^k \right).$$

On pose

$$f_k(z) = (1-z)^{-\frac{b_0}{\sigma}} \left(1 - (1-z)^{\frac{\alpha}{\sigma}}\right)^k,$$

on cherche à calculer $[z^n] f_k(\sigma z) = \sigma^n [z^n] f_k(z)$, grâce à la formule de Cauchy.

On pose γ_R un chemin qui fait le tour de 0 et qui évite la valeur 1. On a alors

$$[z^n] f_k(z) = \frac{1}{2i\pi} \int_{\gamma_R} f(u) \frac{du}{u^{n+1}}.$$

On fait tendre le grand rayon R vers $+\infty$, la longueur du cercle varie comme R , et sur le cercle de centre 0 et de rayon R , $|f|$ est dominée par $R^{\frac{k\alpha}{\sigma}}$, donc pour n assez grand, la contribution due au grand cercle devient négligeable devant celle au voisinage de la singularité. On réalise

FIGURE 6.7 – Représentation du contour g .

le changement de variables $n(1-u) \rightarrow t$. On note g un contour qui entoure l'axe réel négatif, l'intégrale calculée précédemment est donc équivalente à :

$$\frac{1}{2in\pi} \int_g \left(\frac{t}{n}\right)^{-\frac{b_0}{\sigma}} \left(1 - \left(\frac{t}{n}\right)^{\frac{\alpha}{\sigma}}\right)^k \frac{dz}{\left(1 - \frac{t}{n}\right)^{n+1}},$$

on pose $\xi = \frac{k}{n^{\frac{\alpha}{\sigma}}}$, on a donc :

$$[z^n]f_k(z) = \frac{1}{2i\pi} \frac{1}{n^{1-\frac{b_0}{\sigma}}} \int_g t^{-\frac{b_0}{\sigma}} \left(1 - \left(\frac{t}{n}\right)^{\frac{\alpha}{\sigma}}\right)^{\xi n^{\frac{\alpha}{\sigma}}} \frac{dt}{\left(1 - \frac{t}{n}\right)^{n+1}}.$$

Ceci nous conduit à l'approximation suivante :

$$[z^n]f_k(z) \sim \frac{1}{2i\pi} \frac{1}{n^{1-\frac{b_0}{\sigma}}} \int_g t^{-\frac{b_0}{\sigma}} e^{-\xi t^{\frac{\alpha}{\sigma}} + t} dt.$$

Nous pouvons calculer cette intégrale en développant $e^{-\xi t^{\frac{\alpha}{\sigma}}}$ en séries entières :

$$\begin{aligned} \int_g t^{-\frac{b_0}{\sigma}} e^{-\xi t^{\frac{\alpha}{\sigma}} + t} dt &= \sum_{k=0}^{+\infty} \int_g \frac{(-\xi)^k}{k!} t^{\frac{k\alpha-b_0}{\sigma}} e^t dt \\ &= 2i\pi \sum_{k=0}^{+\infty} \frac{(-\xi)^k}{\Gamma\left(\frac{b_0-k\alpha}{\sigma}\right)k!}, \end{aligned}$$

grâce au lemme précédent. On en tire :

$$[z^n]f_k(z) = \frac{1}{n^{1-\frac{b_0}{\sigma}}} \sum_{k=0}^{+\infty} \frac{(-\xi)^k}{\Gamma\left(\frac{b_0-k\alpha}{\sigma}\right)k!}.$$

De plus, grâce à la formule de Stirling, on a les équivalents suivants :

$$\begin{aligned} \binom{k + \frac{a_0}{\alpha} - 1}{k} &\sim \frac{k^{\frac{a_0}{\alpha} - 1}}{\Gamma\left(\frac{a_0}{\alpha}\right)} \text{ et} \\ \frac{n!}{\Gamma\left(\frac{s_0}{\sigma} + n\right)} &\sim \frac{1}{n^{\frac{s_0}{\sigma} - 1}}. \end{aligned}$$

On obtient alors, en définitive (rappelons que $h_n = \sigma^n \frac{\Gamma\left(\frac{s_0}{\sigma} + n\right)}{\Gamma\left(\frac{s_0}{\sigma}\right)}$) :

$$\begin{aligned} \mathbb{P}(a_n = a_0 + k\alpha) &= \frac{\Gamma\left(\frac{s_0}{\sigma}\right)n!}{\Gamma\left(\frac{s_0}{\sigma} + n\right)\sigma^n} \binom{k + \frac{a_0}{\alpha} - 1}{k} \sigma^n [z^n]f_k(z) \\ &\sim \frac{\Gamma\left(\frac{s_0}{\sigma}\right)}{\Gamma\left(\frac{a_0}{\alpha}\right)} \xi^{\frac{a_0}{\alpha} - 1} \frac{1}{n^{\frac{\alpha}{\sigma}}} \sum_{k=0}^{+\infty} \frac{(-\xi)^k}{k! \Gamma\left(\frac{b_0 - k\alpha}{\sigma}\right)}. \end{aligned}$$

□

Remarque 5.2. On notera que les urnes triangulaires sont nettement plus sensibles aux changements de conditions initiales que les urnes semi-sacrificielles. De plus, les loi limites possibles de $\frac{a_n}{n^\alpha}$ semblent être peu connues.

6 Conclusion

Dans ce document, nous avons tout d'abord développé des outils généraux pour l'analyse des urnes aléatoires. Dans un second temps, nous nous sommes concentrés sur deux types d'urnes particulières : les urnes semi-sacrificielles et les urnes triangulaires, et grâce à des résultats d'analyse complexe, nous avons obtenu des résultats de convergence très différents.

Les urnes semi-sacrificielles suivent un régime de convergence semblable au théorème de la limite centrale pour des variables aléatoires indépendantes et identiquement distribuées. Pour les urnes triangulaires, nous avons obtenus un résultat de type limite locale, et la convergence vers une distribution peu connue.

Bibliographie

- [1] Jean Dieudonné. *Calcul infinitésimal*. Hermann, 1997.
- [2] Paul Ehrenfest and Tatiana Ehrenfest. Über zwei bekannte Einwände gegen das Boltzmannsche H-Theorem. *Physikalische Zeitschrift*, 8(9) :311–314, 1907.
- [3] Philippe Flajolet. *Analytic combinatorics*. Cambridge University Press.
- [4] Philippe Flajolet, Philippe Dumas, and Vincent Puyhaubert. Some exactly solvable models of urn process theory. *Discrete Mathematics and Computer Science*, AG :59–118, 2006.
- [5] Philippe Flajolet and Andrew M. Odlyzko. Singularity analysis of generating functions. *SIAM Journal of Algebraic and Discrete Methods*, (2) :216–240, 1990.
- [6] Bernard Friedman. A simple urn model. *Communications in Pure and Applied Mathematics*, 2 :59–70, 1949.
- [7] Bernard Friedman. *Théorie analytique des probabilités*. Édition Jacques Gabay, 1995. Réimpression des éditions de 1819 et 1820.
- [8] Jim Pitman. Combinatorial stochastic processes. *Lecture Notes of Saint-Flour*, Rapport technique 621 :231, 2006.

Chapitre 7

Vulgarisation pour le site CultureMath : L'Urne de Pólya

Bastien MALLEIN

29 novembre 2010

Résumé

Comment évolue au cours du temps une population dans laquelle deux versions d'un même gène co-existent, et quelle sera la proportion, dans la population finale, des sous populations caractérisées par cette différence génétique? Ce problème et bien d'autres peuvent être modélisés par un outil probabiliste : une urne aléatoire. Nous présenterons ici cet objet, ainsi qu'une manière de calculer les probabilités d'évolution de celui-ci grâce à des séries entières.

Je remercie Grégory Ginot pour sa relecture attentive de toutes les versions préparatoires à ce texte, ainsi qu'Éric Vandendriessche qui a accepté de publier cette version, et pour ses nombreux conseils éclairés.

Sommaire

1	Introduction	155
2	Série génératrice	156
2.1	Propriétés des séries génératrices	156
2.2	Calcul des coefficients de fonctions « utiles »	157
2.3	La suite de Fibonacci	159
3	Urne aléatoire	160
3.1	Série génératrice associée à l'urne aléatoire	161
3.2	L'urne de Pólya	164
4	Conclusion	167

1 Introduction

Le but de cet article est de présenter un objet étudié en théorie des probabilités : les urnes aléatoires. Une urne aléatoire est une boîte opaque, contenant des boules de deux couleurs, disons rouge et bleu. Sur la boîte on a inscrit une règle d'évolution, sous la forme suivante :

- Tirer une boule au hasard.
- Regarder sa couleur et la remettre dans l'urne.
- Si la boule est rouge, ajouter α boules rouges (α pouvant être négatif, dans ce cas on retirera des boules rouges) et β boules bleues.
- Si la boule est bleue, ajouter γ boules rouges et δ boules bleues (là encore, δ peut être positif ou négatif).
- Recommencer à la première étape.

α, β, γ et δ sont des entiers fixés, et on suppose qu'avant le premier tirage, le nombre de boules de chaque couleur présentes dans la boîte est connu. On cherche alors comment la composition de cette urne évoluera au fil des tirages. Le nombre de boules présentes dans l'urne après n tirages forme deux suites de variables aléatoires que l'on appelle une chaîne de Markov, c'est-à-dire que seul la valeur de la suite à l'instant n influe sur la loi de la $(n+1)^{\text{ième}}$ variable aléatoire. Insistons sur le fait que dans notre règle d'évolution, on a supposé que la boule tirée est remise dans l'urne.

Un des intérêts des urnes aléatoires est qu'elles donnent des modèles, simples à comprendre et à utiliser, pour simuler des phénomènes divers, en physique, en biologie... On peut par exemple s'intéresser à l'évolution d'un gaz dans deux chambres reliées l'une à l'autre par un trou, pour savoir s'il est possible que tout le gaz rentre dans la première chambre. C'est dans ce but qu'Ehrenfest a introduit l'urne qui porte son nom, afin de démontrer certains paradoxes de la thermodynamique qui annonçait qu'une telle évolution était impossible, le gaz devant forcément s'équilibrer entre les deux chambres. L'urne évolue ainsi : lorsqu'on tire une boule d'une couleur, on la remplace par une boule de l'autre couleur. Le lien avec le modèle est le suivant. On suppose que l'écoulement dans les chambres se fait une molécule après l'autre. Les boules rouges représentent les molécules présentes dans la chambre de gauche, et les bleues celles de la chambre de droite. A chaque étape, une molécule au hasard parmi toutes les molécules présentes change de chambre. On retrouve bien le problème d'Ehrenfest.

Nous pouvons également citer la modélisation d'évolutions de populations. Des modèles déterministes, comme la suite de Fibonacci, que nous expliciterons en Section 2.3. sont connus depuis bien longtemps. Les urnes aléatoires permettent d'obtenir d'autres types de modèles, aléatoires, comme l'urne de Pólya que nous présenterons en Section 3.2, qui permet de modéliser l'évolution d'une population dans laquelle deux versions d'un même gène coexistent. Pour cette urne, les boules rouges représentent alors des individus possédant l'une des versions du gène, et les boules bleues ceux possédant l'autre version. La boule que l'on tire est alors considérée comme un descendant possédant l'une des deux versions du gène. On ajoute donc dans l'urne une boule de la même couleur que celle que nous avons piochée, et remise.

Afin de calculer la probabilité pour qu'après n étapes, il y ait un certain nombre de boules rouges dans l'urne, nous aurons besoin d'une notion très utile en combinatoire, celle de série génératrice. L'idée est d'associer une fonction à une suite, de telle sorte que des égalités sur les quantités de la suite (comme des formules de récurrence) se transforment en équations fonc-

tionnelles (équation différentielle, équation du second degré, etc.). On utilise alors des résultats d'analyse afin de déterminer cette fonction, pour ainsi remonter à la suite inconnue.

D'une certaine manière, la série génératrice (la fonction associée) "retient" toute l'information contenue dans la suite, y compris pour les très grandes valeurs, alors qu'en combinatoire, la plupart des raisonnements se font à entier n fixé. Cette "connaissance" des grandes valeurs nous permet alors d'appliquer les relations connues à toute la suite à la fois, et donc d'obtenir une information supplémentaire. De plus les outils d'analyse permettent une écriture très simple de certaines relations compliquées.

Nous commencerons par présenter les séries génératrices, que nous appliquerons tout d'abord au calcul de la fameuse suite de Fibonacci. Ensuite nous reviendrons aux urnes aléatoires en général, avant d'appliquer les résultats présentés au cas de l'urne de Pólya.

2 Série génératrice

Présentons maintenant ce qu'est une série génératrice. On a déjà expliqué qu'il s'agissait d'une fonction associée à une suite permettant notamment deux opérations : connaissant la fonction, on doit pouvoir retrouver la suite, et les relations connues sur les suites doivent pouvoir se traduire sur la fonction. Nous définissons la série génératrice associée à la suite $(a_n)_{n \in \mathbb{N}}$ comme :

$$f : z \in \mathbb{C} \mapsto a_0 + a_1z + a_2z^2 + \dots$$

Cette fonction est bien définie pour z assez petit, pourvu que $(a_n)_{n \in \mathbb{N}}$ ne croisse pas trop vite. Par exemple, si $a_n \leq A^n$ pour tout n , la limite de $(\sum_{k=0}^n a_k z^k)_{n \in \mathbb{N}}$ existe pour tout $|z| < \frac{1}{A}$. Donc dans ce cas, la série génératrice est bien définie, et on peut calculer avec.

2.1 Propriétés des séries génératrices

Nous devons pour commencer vérifier que les séries génératrices caractérisent les suites auxquelles elles sont associées de manière non équivoque, c'est-à-dire que si deux séries génératrices sont égales, alors les suites associées le sont aussi.

Prenons deux suites $(a_n)_{n \in \mathbb{N}}$ et $(b_n)_{n \in \mathbb{N}}$. Si nous avons :

$$a_0 + a_1z + a_2z^2 + \dots = b_0 + b_1z + b_2z^2 + \dots$$

alors pour $z = 0$ nous avons $a_0 = b_0$. Soustrayons $a_0 = b_0$, et divisons par z nous avons :

$$a_1 + a_2z + a_3z^2 + \dots = b_1 + b_2z + b_3z^2 + \dots$$

Ce qui nous permet de voir que $a_1 = b_1$, en faisant tendre z vers 0. En réitérant le processus nous obtenons bien que les suites $(a_n)_{n \in \mathbb{N}}$ et $(b_n)_{n \in \mathbb{N}}$ sont identiques.

Par conséquent, à chaque fonction est associée au plus une série génératrice, mais cette preuve nous donne davantage : connaissant une fonction qui est une série génératrice, nous

pouvons calculer la suite de termes associée en utilisant la méthode précédente. Par exemple, pour $f(z) = \frac{1}{1+z^2}$, nous avons :

$$\begin{aligned} a_0 &= f(0) = 1 \\ a_1 &= \lim_{z \rightarrow 0} \frac{1}{z} (f(z) - a_0) = \lim_{z \rightarrow 0} \frac{-z}{1+z^2} = 0 \\ a_2 &= \lim_{z \rightarrow 0} \frac{1}{z} \left(\frac{-z}{1+z^2} - 0 \right) = \lim_{z \rightarrow 0} \frac{-1}{1+z^2} = -1 \\ a_3 &= \lim_{z \rightarrow 0} \frac{1}{z} \left(\frac{-1}{1+z^2} + 1 \right) = \lim_{z \rightarrow 0} \frac{z}{1+z^2} = 0 \\ &\dots \end{aligned}$$

Nous pourrions bien sûr itérer ce processus pour obtenir pour tout entier n :

$$a_n = \lim_{z \rightarrow 0} \frac{1}{z^n} [f(z) - (a_0 + a_1z + a_2z^2 + \dots + a_{n-1}z^{n-1})],$$

mais cette formule nous demande de calculer tous les termes jusqu'au rang $n-1$ pour connaître le $n^{\text{ième}}$ terme, alors que nous souhaiterions connaître une formule valable quel que soit l'entier n . Notons que $(a_n)_{n \in \mathbb{N}}$ ne dépend que des valeurs de f pour z proche de 0. On pourra par conséquent toujours supposer que z est assez petit, et restreindre f à une fonction définie sur un voisinage de 0.

Une autre propriété intéressante de ces séries est leur comportement intuitif vis-à-vis des opérations usuelles. Notamment, si l'on pose :

$$f(z) = a_0 + a_1z + a_2z^2 + \dots \text{ et } g(z) = b_0 + b_1z + b_2z^2 + \dots,$$

on a alors (pour des z assez petits) :

$$\begin{aligned} f'(z) &= a_1 + 2a_2z + 3a_3z^2 + \dots \\ f(z) + g(z) &= (a_0 + b_0) + (a_1 + b_1)z + (a_2 + b_2)z^2 \\ f(z)g(z) &= (a_0 + a_1z + a_2z^2 + \dots)(b_0 + b_1z + b_2z^2 + \dots) \\ &= a_0b_0 + (a_1b_0 + a_0b_1)z + (a_2b_0 + a_1b_1 + a_0b_2)z^2 + \dots \end{aligned}$$

Par conséquent, f' est la série génératrice associée à la suite $((n+1)a_{n+1})_{n \in \mathbb{N}}$, $f+g$ est celle associée à $(a_n + b_n)_{n \in \mathbb{N}}$, et fg à $(\sum_{k=0}^n a_k b_{n-k})_{n \in \mathbb{N}}$.

Intéressons-nous maintenant à la détermination de séries génératrices remarquables, qui nous permettront de calculer par la suite les développements des fonctions que nous obtiendrons.

2.2 Calcul des coefficients de fonctions « utiles »

Les séries génératrices que nous aurons à développer par la suite pourront être exprimées comme somme de fonctions plus simples de la forme $z \rightarrow \frac{1}{(1-z)^i}$. Par conséquent nous allons

déterminer quelles sont les suites associées à ces fonctions. Nous commençons par calculer le développement de $z \rightarrow \frac{1}{1-z}$.

Soit g_n la fonction définie (pour $|z| < 1$) comme suit :

$$g_n(z) = 1 + z + z^2 + \dots + z^n.$$

On remarque que

$$\begin{aligned} z g_n(z) - g_n(z) &= z + z^2 + \dots + z^{n+1} - (1 + z + \dots + z^n) \\ &= z^{n+1} - 1. \end{aligned}$$

Par conséquent $g_n(z) = \frac{z^{n+1} - 1}{z - 1}$. En faisant tendre n vers $+\infty$, nous obtenons alors :

$$g_n(z) \xrightarrow{n \rightarrow +\infty} \frac{1}{1-z}$$

Ce qui nous donne en définitive l'égalité suivante valable pour $|z| < 1$:

$$1 + z + z^2 + \dots = \frac{1}{1-z},$$

autrement dit, la suite associée à la fonction $z \mapsto \frac{1}{1-z}$ est la suite constante égale à 1.

Dès lors nous pouvons écrire, pour $a, b, c \in \mathbb{R}$:

$$\begin{aligned} \frac{a}{bz + c} &= \frac{a}{c} \frac{1}{1 - \left(-\frac{bz}{c}\right)} \\ &= \frac{a}{c} \left(1 - \frac{b}{c}z + \frac{b^2}{c^2}z^2 + \dots\right), \end{aligned}$$

donc par conséquent le $n^{\text{ième}}$ coefficient associé à cette fonction s'écrira $a \frac{(-b)^n}{c^{n+1}}$.

Nous allons généraliser un peu plus notre propos, et développer $f_p : z \rightarrow \frac{1}{(1-z)^p}$.

Pour cela procédons par récurrence. On sait que pour $p = 1$, tous les coefficients sont égaux à 1. Pour $p = 2$, remarquons que $f_2 = f_1'$. Par conséquent, il suffit de dériver l'écriture de f_1 en série génératrice pour obtenir :

$$f_2(z) = 1 + 2z + 3z^2 + \dots$$

De manière générale, on remarque que $f_{p+1} = \frac{1}{p} f_p'$. On obtient ainsi par récurrence que, pour tout $n \in \mathbb{N}$:

$$f_p(z) = 1 + \binom{p}{1}z + \binom{p+1}{2}z^2 + \binom{p+2}{3}z^3 + \dots$$

Soit par conséquent, le $n^{\text{ième}}$ coefficient associé à f_p est égal à $\binom{n+p-1}{n} = \binom{n+p-1}{p-1}$.

Nous allons maintenant appliquer ces résultats une première fois à la suite de Fibonacci.

2.3 La suite de Fibonacci

Historique

La suite de Fibonacci a été introduite au XIII^e siècle par Leonardo Fibonacci. Cette suite peut représenter, par exemple, l'évolution d'une population de lapins qui se reproduisent de la manière suivante. Un jeune couple de lapins met un mois à devenir adulte, et tout couple de lapins adultes donne naissance à un nouveau couple chaque mois. On note f_n le nombre de couples de lapins présents au début du $n^{\text{ième}}$ mois, et on suppose qu'au premier mois, on ne possède qu'un unique couple de jeunes lapins.

Nous commençons par déterminer la formule de récurrence. Le nombre de couples de lapins naissant le $(n+2)^{\text{ième}}$ mois est égal au nombre de couples âgés de plus d'un mois présents. Or il y a $f_{n+1} - f_n$ couples qui sont nés au $(n+1)^{\text{ième}}$ mois qui sont donc des jeunes qui ne se reproduisent pas. Chacun des f_n couples âgés d'au moins un mois donne alors naissance à un nouveau couple. On a en définitive :

$$f_{n+2} = 2f_n + (f_{n+1} - f_n) = f_{n+1} + f_n.$$

Les conditions initiales sont données par $f_0 = 0$ et $f_1 = 1$.

Cette suite est très connue en raison de son lien avec le nombre d'or $\phi = \frac{1 + \sqrt{5}}{2}$. En particulier on remarque que le rapport $\frac{f_{n+1}}{f_n}$ tend vers ϕ quand n tend vers l'infini, comme nous le montrerons ci-dessous.

Construction et calcul de la série génératrice

Soit ψ la série génératrice associée à la suite de Fibonacci, qui est définie par :

$$\psi(z) = f_0 + f_1z + f_2z^2 + \dots$$

Nous utilisons alors la relation de récurrence pour calculer explicitement ψ . En effet, en écrivant $f_{n+2} = f_{n+1} + f_n$ dans la formule précédente, nous obtenons :

$$\begin{aligned} \psi(z) &= f_0 + f_1z + f_2z^2 + \dots \\ &= f_0 + f_1z + (f_2z^2 + f_3z^3 + f_4z^4 + \dots) \\ &= z + ((f_0 + f_1)z^2 + (f_1 + f_2)z^3 + (f_2 + f_3)z^4 + \dots) \\ &= z + z^2(f_0 + f_1z + f_2z^2 + \dots) + z(f_0 + f_1z + f_2z^2 + \dots) \\ &= z + z\psi(z) + z^2\psi(z). \end{aligned}$$

On vient d'appliquer à la série génératrice la relation de récurrence que nous connaissions pour la suite. Cela se traduit par une égalité fonctionnelle, particulièrement simple dans ce cas particulier, qui nous permet de déterminer ψ :

$$\psi(z) = \frac{z}{1 - z - z^2}.$$

Ce n'est malheureusement pas une forme que l'on sait facilement développer en série entière. Nous allons donc la décomposer en éléments simples. Commençons par déterminer les solutions de $z^2 + z - 1 = 0$. On obtient $\omega_1 = \frac{\sqrt{5}-1}{2} = \frac{1}{\phi}$ et $\omega_2 = -\frac{1+\sqrt{5}}{2} = -\phi$, où on a posé $\phi = \frac{1+\sqrt{5}}{2}$ le nombre d'or. On écrit alors :

$$\psi(z) = \frac{a}{\omega_1 - z} + \frac{b}{\omega_2 - z}.$$

En réduisant au même dénominateur il vient :

$$\psi(z) = \frac{-(a+b)z + \left(-a\phi + \frac{b}{\phi}\right)}{z^2 + z - 1} = \frac{-z}{z^2 + z - 1},$$

d'où on tire le système :

$$\begin{cases} a + b = 1 \\ \frac{b}{\phi} - a\phi = 0 \end{cases},$$

soit $a = \frac{1}{1+\phi^2}$ et $b = \frac{\phi^2}{1+\phi^2}$

En utilisant les calculs du paragraphe précédent et en déterminant la série génératrice associée à ψ , nous obtenons alors pour f_n :

$$\begin{aligned} f_n &= a \frac{1}{\omega_1^{n+1}} + b \frac{1}{\omega_2^{n+1}} \\ &= \frac{1}{1+\phi^2} \left(\phi^{n+1} + \frac{(-1)^{n+1}}{\phi^{n-1}} \right) \end{aligned}$$

On en déduit une expression explicite pour la suite de Fibonacci ne nécessitant pas le calcul des n premiers termes pour donner f_{n+1} . En particulier, on obtient bien, grâce à un petit calcul, que $\frac{f_{n+1}}{f_n} \xrightarrow{n \rightarrow +\infty} \phi$.

3 Urne aléatoire

Les urnes aléatoires ont été introduites il y a plus de trois siècles, et les premiers résultats ont été démontrés par Jacob Bernoulli et Laplace [3]. Par la suite de nombreuses urnes particulières ont été introduites pour modéliser certains problèmes, parmi lesquelles les urnes d'Ehrenfest, Friedman ou encore Pólya sont certainement les plus connues.

Nous allons maintenant utiliser les outils que nous venons de développer pour étudier l'urne de Pólya. Comme l'approche que nous avons choisi s'étend sans grandes difficultés à de très nombreux types d'urnes, nous nous placerons pour commencer dans un cadre général. Nous laissons à la curiosité du lecteur l'application des résultats obtenus ici à d'autres urnes aléatoires, comme l'urne d'Ehrenfest, ou bien d'autres¹.

1. Par exemple l'urne de Friedman qui suit la règle suivante : lorsqu'on tire une boule d'une couleur, on ajoute

FIGURE 7.1 – Arbre des histoires d'une urne de Pólya

3.1 Série génératrice associée à l'urne aléatoire

Commençons par nous intéresser à une urne aléatoire générale. Rappelons que la règle est la suivante : si on a tiré une boule rouge, on ajoute α boules rouges et β bleues à l'urne, et si on a tiré une boule bleue, on en ajoute γ rouges et δ bleues. On note a_0 et b_0 le nombre (déterministe) de boules respectivement rouges et bleues présentes à l'instant 0, et a_n et b_n le nombre (aléatoire) de boules de chaque couleur présentes dans l'urne après n étapes.

Nous cherchons quelle est la probabilité pour une urne donnée d'arriver après n étapes à une configuration avec a boules rouges et b boules bleues. Pour cela nous avons besoin de déterminer le nombre d'« histoires » possibles pour l'urne conduisant à une configuration avec a boules rouges et b boules bleues en n étapes. Une « histoire » est donnée par la suite des boules tirées, chacune étant supposées distinctes. Par exemple, pour l'urne de Pólya partant d'une boule rouge et d'une boule bleue, il existe deux histoires menant à deux boules rouges et deux boules bleues après deux étapes : celle où on a d'abord tiré la boule rouge et celle où on a d'abord tiré la boule bleue. Il existe également deux histoires menant à trois boules rouges et une bleue selon la boule rouge tirée à la deuxième étape, l'« ancienne » ou la « nouvelle ».

La figure 7.1 montre que pour l'urne de Pólya, après 2 étapes, chacune des compositions possibles est donnée par deux histoires. On remarquera que l'on distingue bien dans la deuxième étape le choix de la première boule rouge et celui de la deuxième dans la partie gauche du tableau, et de même dans la partie droite pour ce qui est des boules bleues.

Toutes les histoires possibles de longueur n sont équiprobables. Par conséquent, si on note $h_n(a, b)$ le nombre d'histoires de longueur n menant à une urne composée de a boules rouges et b boules bleues et h_n le nombre total d'histoires de longueur n , alors on a :

$$\mathbb{P}(a_n = a, b_n = b) = \frac{h_n(a, b)}{h_n}.$$

Nous allons calculer $h_n(a, b)$ grâce à la notion de série génératrice. Comme la quantité $h_n(a, b)$ croît très vite quand n croît, nous allons utiliser une série génératrice exponentielle, associée à $\frac{h_n(a, b)}{n!}$, où $n! = 1 \times \dots \times n$, qui croira donc à un rythme bien plus acceptable.

Il y a néanmoins un autre problème, nous avons trois quantités à retenir simultanément : le nombre de boules rouges, le nombre de boules bleues, et le nombre d'étapes que nous avons effectué. Néanmoins, après n étapes, l'urne ne peut être que dans un nombre fini d'états différents, par conséquent, la somme suivante :

$$\sum_{a,b} h_n(a, b) x^a y^b$$

une boule de l'autre couleur. Cette urne modélise par exemple une campagne électorale où les deux candidats sont tellement mauvais que toute personne écoutant un électeur convaincu de voter pour l'un parler choisit aussitôt de voter pour l'autre (Toute ressemblance avec une élection existante ou ayant existé est purement fortuite).

est finie quel que soit l'entier n . C'est un polynôme en deux variables x et y . Si x et y sont maintenant pensés comme des paramètres, la connaissance de cette quantité (dépendant de n) pour tout couple (x, y) nous permettra de remonter au polynôme puis aux coefficients. Étudions donc la série génératrice exponentielle associée à ces quantités :

$$H_{x,y}(z) = \left(\sum_{a,b} h_0(a,b)x^a y^b \right) + \left(\sum_{a,b} h_1(a,b)x^a y^b \right) \frac{z}{1} + \left(\sum_{a,b} h_2(a,b)x^a y^b \right) \frac{z^2}{2} + \dots$$

Ceci nous permet donc d'introduire la série génératrice exponentielle associée à l'urne comme la fonction de trois variables x, y, z suivante :

$$H : (x, y, z) \mapsto H_{x,y}(z).$$

Comme nous l'avons vu précédemment, connaître cette fonction nous permet de remonter aux coefficients, en considérant tout d'abord (x, y) comme des paramètres de notre fonction de la variable z .

Pour $x = y = 1$, la fonction $z \mapsto H(1, 1, z)$ est intéressante en elle-même. En effet, h_n est la somme des $h_n(a, b)$ pour tous les couples (a, b) possibles (le nombre d'histoires de longueur n est bien égal à la somme sur tous les états possibles du nombre d'histoires de longueur n menant à un état donné). Par conséquent, comme $H(1, 1, z)$ est la somme de tous les termes $h_n(a, b) \frac{z^n}{n!}$, pour a, b, n entiers, en commençant par faire les sommes par rapport à a et b à entier n fixé, on obtient que $H(1, 1, z)$ s'écrit comme la somme des $h_n \frac{z^n}{n!}$. Autrement dit, nous venons de prouver que la fonction $z \mapsto H(1, 1, z)$ est la série génératrice exponentielle associée à la suite (h_n) .

Afin d'établir une égalité fonctionnelle pour H , il va nous falloir utiliser une représentation astucieuse de notre urne aléatoire, utilisant des équations aux dérivées partielles.

Représentation d'une urne aléatoire par une dérivation

Nous pouvons remarquer que dans la série génératrice précédente, une urne avec a boules rouges et b boules bleues est représentée par le monôme $x^a y^b$. Or si on définit les fonctions suivantes :

$$u : x, y \mapsto x$$

$$v : x, y \mapsto y,$$

on a $x^a y^b = u^a(x, y)v^b(x, y)$. On dit alors que l'urne contenant a boules rouges et b boules bleues est codée par la fonction $u^a v^b$.

Lorsqu'on réalise un tirage de cette urne, on crée a histoires pour lesquelles on a tiré une boule rouge et b pour lesquelles c'est une boule bleue qui est tirée. Par conséquent, à une fonction $u^a v^b$, on associe a fonctions $u^{a+\alpha} v^{b+\beta}$ et b fonctions $u^{a+\gamma} v^{b+\delta}$. Cela revient à appliquer l'opérateur aux dérivées partielles suivant à notre fonction :

$$\mathcal{D} = u^{\alpha+1} v^\beta \partial_x + u^\gamma v^{\delta+1} \partial_y.$$

En effet on a $\mathcal{D}(u^a v^b) = a u^{a+\alpha} v^{b+\beta} + b u^{a+\gamma} v^{b+\delta}$. Par conséquent l'évolution de l'urne est gouvernée par cet opérateur, que l'on associe à l'urne. On peut se servir de celui-ci pour réécrire la série génératrice associée à l'urne.

Système différentiel associé à l'urne

La série génératrice que nous cherchons est obtenue en itérant la dérivation associée à l'urne. Supposons que nous commençons avec a_0 boules rouges et b_0 boules bleues. L'état est représenté par $u^{a_0}v^{b_0}$. Les états accessibles après une étape, (comptés avec le nombre de manières possibles pour arriver à chacun d'entre eux) sont représentés par $\mathcal{D}(u^{a_0}v^{b_0})$. Après deux étapes, c'est $\mathcal{D}(\mathcal{D}(u^{a_0}v^{b_0})) = \mathcal{D}^2(u^{a_0}v^{b_0})$ qui code la somme des états accessibles, puis $\mathcal{D}^3(u^{a_0}v^{b_0})$ après trois étapes, etc...

On peut alors réécrire la série génératrice en regroupant tous les termes associés à z^n . Grâce à l'opérateur \mathcal{D} , nous pouvons alors écrire :

$$H(x, y, z) = (u^{a_0}v^{b_0})(x, y) + \mathcal{D}(u^{a_0}v^{b_0})(x, y) + \frac{1}{2!}\mathcal{D}^2(u^{a_0}v^{b_0})(x, y) + \frac{1}{3!}\mathcal{D}^3(u^{a_0}v^{b_0})(x, y) + \dots$$

Ceci nous donne un moyen d'obtenir une équation fonctionnelle satisfaite par H . Pour cela, appliquons l'opérateur \mathcal{D} à l'égalité ci-dessus.

$$\begin{aligned} D(H)(u, v, z) &= \sum_{n=0}^{+\infty} D^{n+1}(u^a v^b) \frac{z^n}{n!} \\ &= \sum_{n=0}^{+\infty} D^{n+1}(u^a v^b) \partial_z \frac{z^{n+1}}{(n+1)!} \\ &= \partial_z \sum_{n=1}^{+\infty} D^n(u^a v^b) \frac{z^n}{n!} \\ &= \partial_z H. \end{aligned}$$

H vérifie par conséquent l'équation aux dérivées partielles suivante :

$$\partial_z H = x^{\alpha+1} y^\beta \partial_x H + x^\gamma y^{\delta+1} \partial_y H.$$

Pour résoudre cette équation différentielle on va chercher un couple de fonctions $(X(z), Y(z))$ qui se comportent, lorsqu'on les dérive par rapport à z , de la même manière que u et v lorsqu'on leur applique \mathcal{D} . Autrement dit on veut que, pour tout choix de $(a, b) \in \mathbb{N}$, on ait :

$$(X^a Y^b)'(z) = aX(z)^{a+\alpha} Y(z)^{b+\beta} + bX(z)^{a+\gamma} Y(z)^{b+\delta}.$$

Or on sait que :

$$(X^a Y^b)' = aX^{a-1} X' Y^b + bX^a Y^{b-1} Y'.$$

Donc en choisissant successivement $a = 1, b = 0$ et $a = 0, b = 1$, il vient que X et Y vérifient en particulier le couple d'équations différentielles suivant :

$$\begin{cases} \dot{X} = X^{\alpha+1} Y^\beta \\ \dot{Y} = X^\gamma Y^{\delta+1} \end{cases}$$

Maintenant, si les équations différentielles suivantes sont vérifiées par un couple de fonctions (X_x, Y_y) satisfaisant la condition initiale $X_x(0) = x, Y_y(0) = y$, alors on a :

$$\partial_z(X_x^a Y_y^b)(z) = D(u^a v^b)(X(z), Y(z)) = aX(z)^{a+\alpha} Y(z)^{b+\beta} + bX(z)^{a+\gamma} Y(z)^{b+\delta}.$$

Or $H(x, y, z) = \sum_{n=0}^{+\infty} D^n(u^{a_0} v^{b_0})(x, y) \frac{z^n}{n!}$, donc on a :

$$\begin{aligned} H(x, y, z) &= \sum_{n=0}^{+\infty} D^n(u^{a_0} v^{b_0})(X_x(0), Y_y(0)) \frac{z^n}{n!} \\ &= \sum_{n=0}^{+\infty} \partial_z^n(X_x^{a_0} Y_y^{b_0})(0) \frac{z^n}{n!} \\ &= (X_x(z)^{a_0} Y_y(z)^{b_0}) \end{aligned}$$

Nous venons d'obtenir une méthode générale permettant de calculer les valeurs de $h_n(a, b)$. Grâce à cette formule, si on est capable de résoudre le système d'équations différentielles précédent pour une urne aléatoires, alors on est capable de calculer la probabilité d'arriver dans un état donné en un certain nombre d'étapes. Nous allons maintenant appliquer les résultats obtenus au cas de l'urne de Pólya.

3.2 L'urne de Pólya

Rappelons que l'urne de Pólya permet d'étudier l'évolution d'une population de deux espèces. Elle a été introduite par Pólya [4] et été étudiée maintes fois par la suite ([5] pour un bref aperçu historique). L'urne est définie avec les quantités suivantes : $\alpha = 1, \beta = 0, \gamma = 0$ et $\delta = 1$, autrement dit, quand on tire une boule d'une couleur, on en ajoute une de la même couleur. Le système associé à l'urne est alors le suivant :

$$\begin{cases} \dot{X} = X^2 \\ \dot{Y} = Y^2 \end{cases}$$

qui se résout immédiatement en $X_x(t) = \frac{x}{1-xt}, Y_y(t) = \frac{y}{1-yt}$. On a alors :

$$H(x, y, z) = \left(\frac{x}{1-xz}\right)^{a_0} \left(\frac{y}{1-yz}\right)^{b_0}.$$

On peut alors développer H en séries entières pour calculer $h_n(a, b)$. Pour cela nous commencerons par rappeler que :

$$\left(\frac{x}{1-xz}\right)^{a_0} = x + \binom{a_0}{a_0-1} x^2 z + \binom{a_0+1}{a_0-1} x^3 z^2 + \dots$$

On doit maintenant réaliser le produit des deux développements suivants pour obtenir le développement en séries génératrices de notre urne. Rappelons que :

$$(a_0 + a_1 z + a_2 z^2 + \dots)(b_0 + b_1 z + b_2 z^2 + \dots) = a_0 b_0 + (a_1 b_0 + a_0 b_1) z + (a_2 b_0 + a_1 b_1 + a_0 b_2) z^2 + \dots$$

En d'autres termes, au produit de deux séries génératrices de coefficients (a_n) et (b_n) est associé la suite $(a_0b_n + a_1b_{n-1} + \dots + a_nb_0)$. Dans notre cas, la série entière de z $H_{x,y}(z) = H(x, y, z)$ (nous considérons de nouveau x et y comme des paramètres) possède les coefficients suivants :

$$\binom{b_0 + n - 3}{b_0 - 1} xy^{n+1} + \binom{a_0}{a_0 - 1} \binom{b_0 + n - 4}{b_0 - 1} x^2 y^n + \dots + \binom{a_0 + n - 3}{a_0 - 1} x^{n+1} y.$$

Or rappelons que ces coefficients s'écrivent également (c.f. Section 3.1, la définition de $H(x, y, z)$) :

$$\frac{1}{n!} (h_n(0, n+2)y^{n+2} + h_n(1, n+1)xy^{n+1} + \dots + h_n(n+1, 1)x^{n+1}y).$$

Ceci nous permet par conséquent de calculer le nombre d'histoires de longueur n menant à un état particulier de l'urne :

$$\frac{h_n(a, b)}{n!} = \begin{cases} \binom{a-1}{a_0-1} \binom{b-1}{b_0-1} & \text{si } a_0 \leq a \leq a_0 + n \text{ et si } b = n + a_0 + b_0 - a \\ 0 & \text{sinon.} \end{cases}$$

De plus, en se souvenant que la série génératrice exponentielle associée à cette suite est $H(1, 1, z) = \frac{1}{(1-z)^{a_0+b_0}}$, on obtient l'égalité :

$$\frac{h_n}{n!} = \binom{n + a_0 + b_0 - 1}{a_0 + b_0 - 1}.$$

Finalement, toutes les histoires étant équiprobables, si on note a_n le nombre de boules rouges présentes dans l'urne à l'instant n et b_n le nombre de boules bleues présentes à cet instant, nous obtenons les probabilités suivantes, qui nous permettent de déterminer l'évolution de toute urne (on utilisera le fait que le nombre total de boules à l'instant n est connu, $a_n + b_n = a_0 + b_0 + n$, par conséquent nous pouvons nous intéresser au seul nombre de boules rouges, celui de boules bleues s'en déduit) :

$$\mathbb{P}(a_n = a, b_n = n + a_0 + b_0 - a) = \mathbb{P}(a_n = a) = \frac{\binom{a-1}{a_0-1} \binom{n+a_0+b_0-a-1}{b_0-1}}{\binom{n+a_0+b_0-1}{a_0+b_0-1}} \text{ si } a_0 \leq a \leq a_0 + n$$

Nous pouvons maintenant interpréter les résultats obtenus. Commençons par une observation. On remarque que si $a_0 = b_0 = 1$, c'est-à-dire que l'on commence avec une boule rouge et une boule bleue dans l'urne, alors toutes les urnes possibles au temps n ont exactement la même probabilité d'apparaître.

Mais cette propriété n'est pas vraie quel que soit le nombre de boules avec lequel nous débutons : par exemple, si nous étudions une urne dans laquelle il y a au départ deux boules de chaque couleur, soit $a_0 = b_0 = 2$, nous obtenons grâce à un calcul rapide :

$$\mathbb{P}(a_n = a) = 6 \frac{(a-1)(n+3-a)}{(n+3)(n+2)(n+1)}.$$

FIGURE 7.2 – Distribution des probabilités d'apparition d'une urne après 30 tours en fonction des conditions initiales.

Ainsi, dans ce cas, on voit clairement que tous les états possibles au temps n ne sont pas équiprobables. En particulier, une urne qui possède une moitié de boules rouges apparaît avec la probabilité $\frac{3(n+4)^2}{2(n+3)(n+2)(n+1)}$ à l'étape n alors qu'une urne ne possédant que deux boules rouges à l'étape n n'apparaît qu'avec la probabilité $\frac{6}{(n+3)(n+2)}$.

Sur la figure 7.2, nous avons représenté la probabilité de posséder a boules rouges pour une urne de Pólya ayant subi 30 étapes d'évolution, à partir de toutes les situations initiales possibles avec 10 boules, de 1 rouges et 9 bleues à 9 rouges et 1 bleues.

Chacune de ces distributions de probabilités est « piquée » au voisinage de la proportion de boules qu'il contient. Autrement dit, une urne de Pólya favorise dans le futur les configurations qui possèdent un nombre moyen de boules voisin du sien.

Ainsi une urne de Pólya commençant avec une proportion r de boules rouges parmi ses boules rend plus probable les compositions d'urnes pour lesquelles cette proportion est conservée. En particulier, si on commence avec $\frac{1}{3}$ de boules rouges dans l'urne, on a de fortes chances que ce rapport soit conservé pour toujours. Ce phénomène s'accroît quand le nombre total de boules augmente, autrement dit, plus le nombre initial de boules est grand, plus il est difficile de modifier le ratio de boules rouges présentes dans l'urne.

Or une urne aléatoire vérifie une propriété remarquable : si à un moment donné nous regardons dans l'urne le nombre de boules de chaque couleur et que nous replaçons ensuite le couvercle, et recommençons à faire évoluer cette urne, alors l'urne se comporte comme si on avait dès le début commencé avec la composition que nous venons de découvrir. Cette propriété s'appelle la propriété de Markov. L'urne évolue à partir du temps n exactement comme si on avait à l'instant initial commencé avec a_n boules rouges et b_n boules bleues.

Revenons à l'urne de Pólya. On vient de voir que si la composition initiale est de une boule rouge et une boule bleue, alors à un instant donné toutes les compositions de l'urne sont équiprobables. Mais alors le ratio de boules rouges dans l'urne est une variable aléatoire choisie uniformément au hasard. Or si on repart ensuite d'un instant donné, ce ratio aura tendance à se conserver ! Par conséquent, le ratio de boules rouges tend vers une constante, mais cette constante est aléatoire de loi uniforme.

Autrement dit, si on trace le graphe représentant l'évolution du nombre de boules rouges présentes dans l'urne au cours du temps, on obtient pour chaque réalisation de l'expérience une courbe proche d'une droite linéaire, mais dont la pente est aléatoire, et varie de telle façon qu'à un temps donné, sa distribution sur l'axe des ordonnées varie selon une loi uniforme (à entier n fixé, les probabilités d'atteindre chacun des points (n, k) pour $1 \leq k \leq n + 1$ sont égales).

On a représenté sur la figure 7.3 l'évolution du nombre de boules rouges de 11 réalisations d'urnes de Pólya, entre les instants 0 et 500. On remarquera l'allure de droite de ces courbes, ainsi que leur caractère uniforme.

L'urne de Pólya modélise donc une population autostable, pour laquelle tout équilibre à un

FIGURE 7.3 – 11 réalisations d’urnes de Pólya au cours du temps, partant d’une boule rouge et d’une boule bleue

instant donné tend à être conservé à l’infini, avec une probabilité d’autant plus grande que le nombre d’individus est grand, cet équilibre restant toutefois aléatoire. Nous pourrions préciser quelle est la probabilité, partant d’une situation donnée, que ces populations s’éloignent de manière significative de leur équilibre, c’est ce dont s’occupe la théorie des grandes déviations. De nombreuses autres questions peuvent encore se poser sur le comportement à l’infini de telles urnes.

4 Conclusion

On a vu dans cet article comment nous pouvions utiliser des séries génératrices pour calculer certaines quantités (suite récurrente d’ordre 2, comme la suite de Fibonacci, probabilité pour une variable aléatoire d’avoir une valeur donnée, nombres d’objets de taille n dans un ensemble...) . Le nombre d’applications possibles ne s’arrête certainement pas ici. Il y a encore de nombreux cas pour lesquels cette méthode se révèle l’une des plus aisées, comme pour le calcul des nombres de Catalan, qui comptent le nombre d’arbres planaires à n branches (ou bien le nombre de parenthésages de longueur n). Dans ce cas la résolution de la relation de récurrence se transforme en la résolution d’une équation du second degré. Un autre exemple est le nombre t_n de triplets d’entiers naturels de somme égale à n (pour celui-ci nous avons déjà fait tous les calculs dans la Section 2.2, il suffit de considérer le développement de $z \mapsto \frac{1}{(1-z)^3}$ de deux manières distinctes), et encore bien d’autres.

En ce qui concerne les urnes aléatoires en général, il est possible de s’intéresser de plus près au système différentiel que nous leur avons associé, comme dans l’article [1]. Ainsi il devient possible d’étudier le comportement à long terme de certaines classes d’urnes aléatoires. Dans ce cas, il apparaît que les proportions de boules rouges et bleues peuvent se comporter de manière bien plus étrange que présenté ici.

Bibliographie

- [1] Philippe Flajolet, Philippe Dumas, and Vincent Puyhaubert. Some exactly solvable models of urn process theory. *Proceedings of Fourth Colloquium on Mathematics and Computer Science*, AG :59–118, 2006.
- [2] Norman L. Johnson and Samuel Kotz. Urns models and their application. 1977.
- [3] Pierre-Simon Laplace. *Théorie analytique des probabilités*. 1819. Oeuvres complètes, Tome 7, Réédition 1886.
- [4] Georges Pólya. Sur quelques points de la théorie des probabilités. *Annales de l’institut Henri Poincaré*, pages 117–161, 1930.

- [5] P. Tautu. *Stochastic spatial processes in biology : A concise historical survey*, volume 1212. Springer-Verlag, Berlin and New York, lecture notes in math. edition, 1986.