Université de Montréal

Convergence de martingales sur promenades aléatoires avec branchement: preuve conceptuelle

par

Éric Nguyen

Département de mathématiques et de statistique
Faculté des arts et sciences

Mémoire présenté à la Faculté des arts et sciences
en vue de l'obtention du grade de M.Sc.
en mathématiques

Juin 2009

Université de Montréal
Faculté des études supérieures et postdoctorales

Ce mémoire intitulé:

Convergence de martingales sur promenades aléatoires
avec branchement: preuve conceptuelle

présenté par:

Éric Nguyen

a été évalué par un jury composé des personnes suivantes:

Richard Duncan
président-rapporteur

Anatole Joffe
directeur de recherche

Louigi Addario-Berry
membre du jury

Résumé

Étant donné une promenade aléatoire avec branchement, la transformée de Laplace du processus ponctuel formé par les positions des individus, normalisée par son espérance, est une martingale. L'espérance de la limite de cette martingale est dictée par le théorème de Biggins qui fût initialement démontrée de façon analytique.

Une démonstration conceptuelle et probabilistique du théorème est détaillée. L'idée de la preuve fût initialement introduite par Lyons, Pemantle et Peres en 1995 pour le théorème de Kesten-Stigum sur les processus de branchement. Par la mise en relation de nouvelles mesures sur des espaces de probabilité, la convergence d'une martingale se trouve lié à son comportement asymptotique par rapport à une autre mesure spécifique.

Mots-clés. promenade aléatoire avec branchement, théorème de Biggins, processus de branchement, théorème de Kesten-Stigum, preuve conceptuelle.

Abstract

Given a branching random walk, the Laplace transform of the point process given by the positions of the individuals, normalized by its expectation, is a martingale. The expection of the limit of this martingale is given by Biggins' theorem, which was proven analytically.

A conceptual proof, using probabilistic methods, of the theorem is detailed. The initial idea was introduced by Lyons, Pemantle and Peres in 1995 inside their proof of Kesten-Stigum's theorem on branching processes. By a clever choice of measures on related probability spaces, the convergence of a martingale is tightly bound to its asymptotic behaviour with respect to another measure.

Keywords. branching random walk, Biggins' theorem, branching process, Kesten-Stigum's theorem

Table des matières

1. Introduction

Les processus de branchement et les promenades aléatoires avec branchement ont été longuement étudiés. Les problèmes liés à la convergence des martingales associés à ces processus ont été résolus de façon analytique.

En 1995, Lyons, Pemantle et Peres [6] publièrent une démonstration probabiliste du théorème de Kesten-Stigum sur la convergence d'une de ces martingales associées aux processus de branchement. Peu de temps après, Lyons [5] généralisa cette démonstration pour établir le théorème de Biggins dans les promenades aléatoires avec branchement. Le but du présent travail est d'exposer d'une façon didactique l'article de Lyons.

Le lemme 2 de la section 4.1, présenté en annexe, est un point central de la démarche probabiliste. De façon générale, soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé, muni d'une filtration $\mathcal{F}_n$. Sur cet espace, soit μ et ν deux mesures telles que la restriction ν_n de ν sur $\mathcal{F}_n$ est absolument continue par rapport à la restriction μ_n de μ sur $\mathcal{F}_n$. Soit $X_n = \frac{d\nu_n}{d\mu_n}$ la densité de ν_n par rapport à μ_n. Il se trouve que X_n est une martingale, dont la convergence sous la mesure ν amène un scénario dichotomique : si X_n diverge ν-p.p., alors μ et ν sont mutuellement singuliers, alors que si X_n converge ν-p.p., alors ν est absolument continue par rapport à μ.

Dans le cas des processus de branchement, où l'on modélise des populations, le théorème de Kesten-Stigum permet de connaître les conditions sous lesquelles les populations d'individus croissent de façon attendue. Le théorème de Biggins, quant à lui, se rapporte à la distribution des positions des individus de la population dans un espace. La possibilité de démontrer ces deux théorèmes passe par un judicieux choix d'une nouvelle mesure ν telle que la densité sujette à la filtration $\mathcal{F}_n$ est la martingale étudiée.

Dans la section 3, le théorème de Biggins sera démontrée en toute intégralité, alors que pour la section 2, les idées et concepts permettant de reconstruire la démonstration de Lyons, Pemantle et Peres [6] pour le théorème de Kesten-Stigum seront abordés.

2. Processus de branchement et Théorème de Kesten-Stigum

Les processus de Galton-Watson ont été introduits pour modéliser la survivance des noms de famille. En 1873, Francis Galton se posa le problème de déterminer s'il y avait une différence entre la survie des noms de famille des nobles et des paysans. Un an après, le révérend Henry William Watson trouva la solution. Ils copublièrent 'On the probability of extinction of families' en 1874. Ce fut la naissance des processus de Galton-Watson.

L'article initial de Galton et Watson contenait une erreur qui ne fût corrigée que plus tard. Bien avant, Bienaymé avait entamé des études sur les processus de branchement, sans commettre la même erreur. C'est ainsi que Heyde et Seneta[2] suggèrent l'appellation de processus de Bienaymé-Galton-Watson.

Une *loi de reproduction* est une loi L dictant le nombre d'enfants d'un individu. Pour $k \in \mathbb{N}$, p_k est la probabilité selon L d'avoir k enfants. Un processus de branchement est décrit comme ceci : soit une particule au temps 0, elle a une progéniture de taille L_1, une variable i.i.d. à L, et pour chaque enfant σ, nous assignons une progéniture de taille L_σ, i.i.d. à L. Le processus continue ainsi à l'infini ou jusqu'à extinction de la population.

Définition. Soit L une loi de reproduction à valeurs dans les entiers naturels. Un *processus de Galton-Watson* est un processus stochastique $\{Z_n\}$ tel que $Z_0 = 1$ et

$$Z_{n+1} = \sum_{i=1}^{Z_n} L_i^{(n+1)}$$

où $L_i^{(n+1)}$ est i.i.d. à L.

Ainsi, Z_n représente le nombre d'individus de la génération n. Les processus de branchement tels que l'espérance du nombre d'enfants $m := E[L]$ est supérieure à 1 sont appelés *supercritiques*. Les processus discutés dans ce travail le seront tous.

L'*extinction* du processus est l'événement $\{\exists n,\ Z_n = 0\}$. La probabilité d'extinction est désignée par q.

En conditionant sur la non-extinction du processus, l'objectif est d'étudier le comportement asymptotique de Z_n pour les grandes valeurs de n. Pour ce faire, nous étudions la martingale $W_n := \frac{Z_n}{m^n}$. Rappelons la définition d'une martingale :

Définition. Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé muni d'une filtration $\{\mathcal{F}_n\}$. Soit $\{X_n\}$ un processus stochastique (à temps discret) adapté à $\mathcal{F}_n$ tel que pour tout n, $E[|X_n|] < \infty$, alors si

$$E[X_{n+1}|\mathcal{F}_n] = X_n \text{ p.s.}$$

$\{X_n\}$ est une $\mathcal{F}_n$*−martingale* (à temps discret).

Soient $\mathcal{F}_n$ les sous-tribus engendrées par les Z_k, $k \leq n$. La vérification que W_n est une $\mathcal{F}_n$−martingale est immédiate :

$$\begin{aligned}
E[W_{n+1}|\mathcal{F}_n] &= E[\frac{Z_{n+1}}{m^{n+1}}|\mathcal{F}_n] \\
&= E[\frac{\sum_{i=1}^{Z_n} L_i^{(n+1)}}{m^{n+1}}|\mathcal{F}_n] \\
&= \frac{\sum_{i=1}^{Z_n} E[L_i^{(n+1)}|\mathcal{F}_n]}{m^{n+1}} \\
&= \frac{\sum_{i=1}^{Z_n} m}{m^{n+1}} \\
&= \frac{mZ_n}{m^{n+1}} \\
&= \frac{Z_n}{m^n} = W_n.
\end{aligned}$$

La martingale W_n est non-négative. Par le théorème de convergence des martingales, cité en annexe en 4.2, il en découle que sa limite $W := \lim_{n\to\infty} W_n$ existe et est finie p.s. Par le lemme de Fatou, on observe que

$$\begin{aligned}
E[W] &= E[\lim\ \inf_{n\to\infty} W_n] \\
&\leq \lim\ \inf_{n\to\infty} E[W_n] \\
&= \lim\ \inf_{n\to\infty} E[\frac{Z_n}{m^n}] \\
&= 1.
\end{aligned}$$

Lorsque $E[W] = 1$, les Z_n croissent de façon prévisible, soit de l'ordre de grandeur de leur moyenne m^n. Le théorème de Kesten-Stigum donne les conditions nécessaires et suffisantes sous lesquelles.

Théorème. *(Kesten-Stigum, 1966) Soit* $1 < m < \infty$. *Les énoncés suivants sont équivalents :*

(1) $P[W = 0] = q$

(2) $E[W] = 1$

(3) $E[L \log L] < \infty$

où q est la probabilité d'extinction.

2.1. **L'arbre de Galton-Watson.** Pour aborder l'étude des processus de Galton-Watson, il est utile d'introduire la notion d'*arbre*. Pour les détails, le lecteur est invité à consulter l'article de Neveu [8].

L'idée de l'arbre est intuitive : soit un processus de Galton-Watson, avec ancêtre initial σ. L'arbre t est construit avec une racine marqué par σ. Ensuite, de σ sont issues Z_1 branches vers des noeuds représentant les Z_1 enfants de σ. À chaque enfant τ de σ, L_τ branches partent de τ vers des noeuds représentant les membres de la progéniture de τ, et ainsi de suite. L'espace des arbres peut être probabilisé de plusieurs façons. Par exemple, si les copies L_τ sont i.i.d. à L, on obtient les processus de Galton-Watson. À chaque génération n, l'arbre t contient alors Z_n noeuds. L'arbre est fini si le processus s'éteint, et est infini si le processus se répète sans fin.

Nous donnons une orientation de l'arbre vers le bas : la racine est la première particule en haut de l'arbre, et les générations suivantes se placent en dessous, une à la suite de l'autre.

Pour un arbre t, avec $\sigma \in t$ un noeud, le sous-arbre de t ayant σ comme racine est noté par t_σ. La restriction de t à ses n premières générations est désigné par $t|_n$: il est l'arbre fini contenant les mêmes n premiers niveaux que t.

2.1.1. *L'espace topologique.* Soit Ω_T l'espace des arbres. Pour t, $t' \in \Omega_T$, soit $d(t,t') := (\frac{1}{2})^k$ où $k = sup\{n : t|_n = t'|_n\}$. La valeur de $d(t,t')$ dépend donc de la similitude entre t et t' : pour t et t' identique, $d(t,t') = 0$, et $d(t,t') = (\frac{1}{2})^n$ pour le plus petit n tel que $t|_{n+1} \neq t'|_{n+1}$. Notons que $d(t,t') = \infty$ si t et t' ont des racines différentes. Il est clair que $d(t,t')$ satisfait l'inégalité du triangle ; il définit alors une distance sur Ω_T, qui devient un espace métrique. Pour $r \in \mathbb{R}$, soient $B_{t,r} := \{t' \in \Omega_T : d(t,t') < r\}$ les boules ouvertes. Les ensembles ouverts sont alors définis naturellement comme les ensembles U tels que pour tout $x \in U$, il existe une boule ouverte de centre x incluse dans U. Ils génèrent ainsi une topologie sur Ω_T. La tribu borélienne Σ est ainsi obtenue à partir des ensembles ouverts et fermés.

Soit $t \in \Omega_T$. L'ensemble $[t]_n := \{t' \in \Omega_T : t|_n = t'|_n\}$ est l'ensemble des arbres t' tel que les n premiers niveaux de t et de t' sont identiques. Il est clair que $[t]_n \in \Sigma$ par la définition de $d(t,t')$.

2.1.2. *Une mesure induite.* Soit L une loi de reproduction, et l'espace des arbres Ω_T. Soit μ_L la mesure canonique sur Ω_T induite par le processus de branchement. Pour un sous-espace $U \in \Sigma$, $\mu_L(U)$ donne la probabilité que des arbres ayant été construits selon la loi de reproduction L appartiennent à U. Par exemple, si $P[L = 1] = 0$, et si U est le sous-ensemble de Ω_T formé par tous les arbres dont la racine possède un enfant, alors $\mu_L(U) = 0$.

2.2. **L'arbre biaisé par à la taille et une autre mesure induite.**

2.2.1. *Motivation.* Nous allons construire une deuxième mesure $\hat{\mu}_L$, sur Ω_T, telle que la restriction à chacune des générations donne lieu à une mesure absolument continue par rapport à celle de μ_L, et dont la dérivée de Radon-Nikodym est W_n. Autrement dit, si Σ_n est la sous-tribu engendrée par les n premières générations des arbres de Ω_T, et si $\mu_{L,n}$ et $\hat{\mu}_{L,n}$ sont les restrictions respectives de μ_L et $\hat{\mu}_L$ à Σ_n, l'égalité

$$\hat{\mu}_{L,n}([t]_n) = W_n(t) \cdot \mu_{L,n}([t]_n)$$

est recherchée pour tout $t \in \Omega_T$ et pour tout n.

En appliquant le lemme 2 de la section 4.1 directement aux mesures μ_L et $\hat{\mu}_L$, $E[W]$ dépend du comportement de W sous la mesure $\hat{\mu}_L$, ce qui permettra de conclure le théorème de Kesten-Stigum de façon probabiliste.

2.2.2. *Le biais par à la taille.* L'obtention de la nouvelle mesure $\hat{\mu}_L$ passe par une nouvelle façon de construire des arbres. L'*arbre biaisé par à la taille* est une traduction de l'anglais *size-biased tree.*

Soit L la loi de reproduction d'un processus de Galton-Watson supercritique. Sa loi biaisée par la taille, $\hat{L}$, est définie par

$$P[\hat{L} = k] = \frac{k \cdot P[L = k]}{m}.$$

Notons que $P[\hat{L} = 0] = 0$. Les familles engendrées par $\hat{L}$ ont tendance à avoir plus d'enfants que celles engendrées par L puisque $E[\hat{L}] \geq E[L]$ par l'inégalité de Jensen.

De façon intuitive, le biais par à la taille peut être observé dans divers problèmes [?]. Par exemple, pour un arbre t avec au moins n générations, la probabilité de choisir un parent à la génération $n-1$ en choisissant de manière uniforme un individu de la génération n est une situation où la taille crée un biais : plus sa progéniture directe est grande, plus le parent a de chances d'être choisi.

2.2.3. *La nouvelle construction d'arbres.* Doté de la nouvelle loi $\hat{L}$, l'arbre biaisé se construit ainsi : la particule initiale, v_0, donne naissance à un nombre $\hat{L}_1$ d'individus, où $\hat{L}_1$ est i.i.d. à $\hat{L}$. Parmi les enfants de v_0, un individu v_1 est choisi de façon aléatoire, avec probabilité uniforme. L'individu v_1 donne naissance à $\hat{L}_2$ enfants, où $\hat{L}_2$ est i.i.d. à $\hat{L}$, alors que la reproduction se continue avec un arbre de Galton-Watson ordinaire avec loi L pour les autres enfants de v_0. Parmi les $\hat{L}_2$ enfants de v_1, de nouveau, un individu v_2 est choisi de façon aléatoire. Il aura $\hat{L}_3$ enfants. Les autres enfants de v_1 engendrent des arbres de Galton-Watson ordinaires avec la loi L. Le processus se poursuit ainsi, de façon indéfinie puisque $\hat{L} \geq 1$ p.s.

2.2.4. *La densité recherchée.* Soit $t \in \Omega_T$ un arbre. Un *rayon* est une ligne commençant à la racine de t et descendant à l'infini vers les noeuds plus profonds sans jamais remonter vers la racine; il noté par ξ. Pour $\xi \subseteq t$, la paire (t, ξ) est appelé *arbre avec rayon* et l'espace des arbres avec rayon est désigné par Ω_T^*. Soit v un noeud de t situé à la génération n. L'ensemble $[t, v]_n \subseteq \Omega_T^*$ est l'ensemble de tous les arbres avec rayon (t', ξ) tel que $t' \in [t]_n$ et ξ passe par v.

Pour le même arbre t avec v à la n-ième génération, supposons que la racine v_0 de t possède k enfants $\sigma_1, \sigma_2, ..., \sigma_k$, donnant lieu à k sous-arbres $t_{\sigma_1}, t_{\sigma_2}, ..., t_{\sigma_k}$. Le noeud v appartient à l'un de ses sous-arbres, t_{σ_i}. La mesure μ_L satisfait

$$
\begin{aligned}
\mu_L([t]_n) &= P[L=k]\cdot\prod_{j=1}^{k}\mu([t_{\sigma_j}]_{n-1}) \\
&= p_k\cdot\mu([t_{\sigma_i}]_{n-1})\cdot\prod_{j\neq i}\mu([t_{\sigma_j}]_{n-1}) \qquad (2.1)
\end{aligned}
$$

où $p_k := P[L=k]$. Soit $\hat{\mu}_L^*$ la mesure canonique sur Ω_T^* induite par la construction de l'arbre biaisé. Nous avons que σ_i est choisi avec probabilité $\frac{1}{k}$, et donc, après substitution avec l'équation 2.1,

$$
\begin{aligned}
\hat{\mu}_L^*([t,v]_n) &= P[\hat{L}=k]\cdot\frac{1}{k}\cdot\hat{\mu}_L^*([t_{\sigma_i},v]_{n-1})\cdot\prod_{j\neq i}\mu([t_{\sigma_j}]_{n-1}) \\
&= \frac{k\cdot p_k}{m\cdot k}\cdot\hat{\mu}_L^*([t_{\sigma_i},v]_{n-1})\cdot\prod_{j\neq i}\mu([t_{\sigma_j}]_{n-1}) \\
&= \frac{1}{m}\cdot\frac{\hat{\mu}_L^*([t_{\sigma_i},v]_{n-1})}{\mu([t_{\sigma_i}]_{n-1})}\cdot\mu_L([t]_n).
\end{aligned}
$$

En réitérant sur les $n-i$, l'équation suivante est obtenue :

$$\hat{\mu}_L^*([t,v]_n)=\frac{1}{m^n}\mu_L([t]_n).$$

Soit $\hat{\mu}_L$ la projection de $\hat{\mu}_L^*$ sur l'espace des arbres Ω_T. La martingale W_n apparaît donc puisque

$$
\begin{aligned}
\hat{\mu}_L([t]_n) &= \frac{Z_n(t)}{m^n}\mu_L([t]_n) \\
&= W_n(t)\mu_L([t]_n).
\end{aligned}
$$

La démonstration de Lyons, Pemantle et Peres se termine par l'utilisation d'un lemme de Seneta sur les processus de branchement avec immigration. Pour les détails de la démonstration originale, le lecteur est invité à les consulter dans [6]. Nous déduirons ce théorème de celui de Biggins.

3. Promenade aléatoire avec branchement

L'arbre de Galton-Watson modélise de nombreux problèmes, notamment en génétique, en épidémiologie et en physique nucléaire. Par contre, il peut devenir intéressant d'attacher à chacun des individus de l'arbre une position dans $\mathbb{R}^n$. Prenons l'exemple d'une graine semée, d'où pousse une plante. Cette plante laisse tomber au sol des graines que le vent déplace de manière aléatoire. Un certain nombre d'entre elles donnent naissance à de nouvelles plantes, dont le déplacement autour du plant initial est aléatoire et dictée par la même loi. Ces derniers reprennent le cycle, laissent tomber au sol des graines qui germeront en d'autres plantes autour de leur parent respectif. La promenade aléatoire avec branchement modélise la distribution des plantes.

La promenade aléatoire avec branchement est définie de la façon suivante : prenons un processus de Galton-Watson avec une loi de reproduction L et ayant comme particule initiale τ. située à l'origine de $\mathbb{R}^n$. Soit L_τ le nombre (aléatoire) d'enfants de τ, où L_τ est identiquement distribué à L. Chacun des enfants effectue un déplacement $X_i \in \mathbb{R}^n$. Pour l'individu τ, l'élément aléatoire $\mathcal{L}_\tau := \{X_i\}_{i=1}^{L_\tau}$ sera identiquement distribué et indépendant à un élément aléatoire commun $\mathcal{L} := \{X_i\}_{i=1}^{L}$, la *loi de déplacement de la progéniture*. Pour chaque individu de la nouvelle génération, le même exercice de reproduction et de déplacement se répète, et le processus continue infiniment, ou jusqu'à extinction de l'arbre. L'ensemble des positions de chacun des individus de l'arbre, à chacune des générations, est appelé *promenade aléatoire avec branchement.*

Pour simplifier, nous supposerons que les déplacements se font sur $\mathbb{R}$.

Soit σ une particule du processus. La génération de σ est désignée par $|\sigma|$. $X(\sigma)$ est le déplacement de σ par rapport à son ancêtre direct, et $S(\sigma)$ est sa position dans $\mathbb{R}$. Aussi, $p(\sigma)$ dénote l'ancêtre direct de σ.

3.1. **Motivation.** Étant donné une promenade aléatoire avec branchement, une question à considérer est le rythme de déplacement auquel l'ensemble des particules s'étendent dans l'espace. Plusieurs façon d'étudier cette dispersion dans l'espace sont possibles. Entre autres, la transformée de Laplace du processus ponctuel formé par les particules a été longuement étudiée. Sans entrer dans les détails formels d'un processus ponctuel, la notion est introduite ici.

Considérons une promenade aléatoire avec branchement. Pour la n-ième génération, soit un sous-ensemble $S \subseteq \mathbb{R}$, et soit $\xi_n(S)$ le nombre aléatoire d'individus de la génération n dont la position se retrouve dans S. ξ_n est donc le *processus ponctuel* généré par la promenade à la génération n, et peut être représenté par $\xi_n = \sum_{i=1}^{Z_n} \delta_{S_i}$, où S_i sont les positions des individus de la n-ième génération, et δ est la mesure de Dirac.

Pour $\alpha \in \mathbb{R}$, la transformée de Laplace de ξ_n est $\int_0^\infty e^{-\alpha t} d\xi_n(t) = \sum_{|\sigma|=n} e^{-\alpha S(\sigma)}$. La transformée de Laplace de ξ_n normalisé par son espérance est une martingale non-négative.

3.2. **Une martingale.** Considérons une promenade aléatoire avec branchement, avec loi de déplacement $\mathcal{L} = \{X_i\}_{i=1}^{L}$, où les X_i sont i.i.d. Pour $\alpha \in \mathbb{R}$, la variable aléatoire

$$\langle \alpha, \mathcal{L} \rangle := \sum_{i=1}^{L} e^{-\alpha X_i}$$

représente la transformée de Laplace du processus ponctuel des déplacements dictés par $\mathcal{L}$. Soit $m(\alpha) := E[\langle \alpha, \mathcal{L} \rangle]$, et

$$m'(\alpha) := -E[\sum_{i=1}^{L} X_i e^{-\alpha X_i}].$$

sa dérivée. Pour obtenir un processus de branchement supercritique, nous supposerons que $1 < m(0) < \infty$. Pour $\alpha \in \mathbb{R}$, et $m(\alpha) < \infty$, définissons

$$W_n(\alpha) := \frac{\sum_{|\sigma|=n} e^{-\alpha S(\sigma)}}{m(\alpha)^n}.$$

Soit $\mathcal{F}_n$ les sous-tribus engendrés par $W_1(\alpha), ..., W_n(\alpha)$. $W_n(\alpha)$ est effectivement une martingale puisque

$$\begin{aligned}
E[W_{n+1}(\alpha)|\mathcal{F}_n] &= E[\frac{\sum_{|\sigma|=n+1} e^{-\alpha S(\sigma)}}{m(\alpha)^{n+1}}|\mathcal{F}_n] \\
&= E[\frac{\sum_{|\sigma|=n+1} e^{-\alpha X(\sigma)} e^{-\alpha S(p(\sigma))}}{m(\alpha)^{n+1}}|\mathcal{F}_n] \\
&= E[\frac{\sum_{|\sigma|=n} (e^{-\alpha S(\sigma)} \langle \alpha, \mathcal{L}_\sigma \rangle)}{m(\alpha)^{n+1}}|\mathcal{F}_n] \\
&= \frac{\sum_{|\sigma|=n} (e^{-\alpha S(\sigma)} E[\langle \alpha, \mathcal{L}_\sigma \rangle])}{m(\alpha)^{n+1}} \\
&= \frac{m(\alpha) \sum_{|\sigma|=n} e^{-\alpha S(\sigma)}}{m(\alpha)^{n+1}} \\
&= \frac{\sum_{|\sigma|=n} e^{-\alpha S(\sigma)}}{m(\alpha)^{n}} \\
&= W_n(\alpha).
\end{aligned}$$

$W_n(\alpha)$ est non-négatif, donc la martingale converge vers une limite finie p.s. Soit $W(\alpha) := \limsup_{n\to\infty} W_n(\alpha)$.

3.3. Théorème de Biggins.

Théorème. *(Biggins, 1977) Soit $\alpha \in \mathbb{R}$, tel que $m(\alpha) < \infty$ et $m'(\alpha) < \infty$. Les énoncés suivants sont équivalents :*

(1) $P[W(\alpha) = 0] = q$

(2) $P[W(\alpha) = 0] < 1$

(3) $E[W(\alpha)] = 1$

(4) $E[\langle \alpha, \mathcal{L} \rangle \log \langle \alpha, \mathcal{L} \rangle] < \infty$ et $\alpha m'(\alpha)/m(\alpha) < \log m(\alpha)$.

3.4. **Démonstration du théorème de Biggins.** La démonstration du théorème de Biggins sera divisée en plusieurs sous-sections. L'équivalence de 1) et 2), ainsi que l'implication 3) $\Rightarrow$ 2) seront abordés. Ensuite, grâce à l'élaboration de mesures additionnelles, de la section 3.4.3 à 3.4.6, et d'un lemme, en 3.4.7, les implications 2) $\Rightarrow$ 4) et 4) $\Rightarrow$ 3) seront ensuite explicitement démontrées.

Dans ce qui suit, $\alpha \in \mathbb{R}$ est fixé.

3.4.1. *L'équivalence de 1) et 2).* Les énoncés 1) et 2) sont équivalents dû à la propriété "*zéro-un*" du processus de branchement.

Pour un processus de branchement, on appelle une propriété *héritée* une caractéristique de l'arbre telle que, si l'arbre possède cette propriété, alors les sous-arbres descendant de la racine possèdent cette même caractéristique, et telle que tous les arbres finis possèdent cette caractéristique. Le lemme suivant est démontré en 4.1 :

Lemme. *Étant donné la non-extinction de l'arbre, toute propriété héritée a une probabilité de 0 ou de 1.*

Soit une promenade aléatoire et soit t le processus de Galton-Watson inscrit dans cette promenade. Il se trouve que la propriété $W(\alpha) = 0$ est une propriété héritée : pour un arbre t fini, $W(\alpha) = 0$ nécessairement. Pour un arbre t infini, avec sous-arbres t_i, il est clair que si $W(\alpha) = 0$ pour t, alors $W(\alpha) = 0$ pour tout sous-arbre t_i, autrement il y aurait contradiction avec la définition de la propriété héritée. Donc $P[W(\alpha) = 0]$ est soit q (la probabilité d'avoir un arbre fini) soit 1, d'où vient l'équivalence des deux propositions 1) et 2).

3.4.2. *L'implication 3) ⇒ 2).* Cette implication est triviale : si $P[W(\alpha) = 0] = 1$, alors $E[W(\alpha)] = 0 \neq 1$.

3.4.3. *Espace des arbres marqués Ω et $\mathcal{F}_n$.* Alors que la promenade aléatoire avec branchement a été définie précédemment, l'espace topologique dont elle fait partie n'a pas été formalisé. L'espace des arbres marqués sera l'espace utilisé.

Soit t un arbre ayant r comme racine. Soit X une fonction des noeuds de t à valeurs réelles et telle que $X(r) = 0$. Pour un noeud $\sigma \in t$, la valeur de $X(\sigma)$ représente le déplacement de σ par rapport à son ancêtre direct. La paire (t, X) est appelé *arbre marqué.* Soit Ω l'espace des arbres marqués. La tribu engendrée par les n premières générations des arbres marqués dans Ω est noté par $\mathcal{F}_n$.

Similairement à l'espace des arbres, l'ensemble $[t, X]_n := \{(t', X') : t' \in [t]_n; \forall \sigma \in t, X(\sigma) = X'(\sigma)\} \subseteq \Omega$ est l'ensemble des arbres marqués (t', X') tel que les n premiers niveaux sont identiques en terme de progéniture et de déplacement.

À partir d'une promenade aléatoire avec branchement, avec loi de reproduction L et loi de déplacement $\mathcal{L}$, la mesure (aléatoire) μ sur Ω est la mesure canonique induite par cette promenade aléatoire avec branchement. Comme en 2.1.2 où nous avons introduit μ_L dans le cas d'un processus de branchement sans promenade, μ est une mesure de probabilité donnant à un ensemble mesurable la probabilité qu'une promenade aléatoire avec branchement, dictée par L et $\mathcal{L}$, lui appartienne. Notons par μ_n la restriction de μ sur $\mathcal{F}_n$.

3.4.4. *Espace des arbres marqués avec rayons distingués Ω^* et $\mathcal{F}_n^*$.* L'espace des arbres marqués avec rayon est une extension de l'espace Ω. Soit $(t, X) \in \Omega$ ayant r comme racine. En partant de r, une ligne de descendance infinie est appelée un rayon, noté par ξ. L'individu de ξ situé sur la n-ième génération est noté par ξ_n. Nous avons donc que $\xi_0 = r$. Un triplet (t, X, ξ) est appelé *arbre marqué avec rayon distingué* : l'espace de ces éléments sera dénotée par Ω^*. Sur cet espace, nous notons par $\mathcal{F}_n^*$ la tribu engendrée par les n premières générations.

Pour un arbre marqué $(t, X, \xi) \in \Omega^*$ avec un noeud $\xi_k \in \xi$, $(t, X, \xi)_{\xi_k}$ désigne le sous-arbre marqué avec rayon formé par le sous-arbre t_{ξ_k} muni des mêmes déplacements $X(\sigma)$ pour $\sigma \in t_{\xi_k}$, $\sigma \neq \xi_k$, et du sous-rayon de ξ commençant par ξ_k. Puisque ce sous-arbre marqué avec rayon a comme racine ξ_k, il est implicite que $X(\xi_k) = 0$.

De façon analogue aux autres espaces, l'ensemble $[t, X, \xi]_n := \{(t', X', \xi') : (t', X') \in [t, X]_n; \forall k \leq n, \xi_k = \xi'_k\} \subseteq \Omega^*$ est l'ensemble des arbres marqués avec rayons distingués (t', X', ξ') tel que les n premiers niveaux sont identiques.

L'espace des arbres marqués avec rayons distingués est une extension de l'espace des arbres marqués mentionné en 3.4.3. Ainsi, étant donné une fonction $\mathcal{F}_n^*$-mesurable f, il est possible de représenter f à l'aide de fonctions f_σ, toutes $\mathcal{F}_n$-mesurables par :

$$f(t, X, \xi) = \sum_{|\sigma|=n} f_\sigma(t, X) 1_{\xi_n=\sigma},$$

où

$$1_{\xi_n=\sigma} \quad := \quad \begin{cases} 1 & \text{si } \xi_n = \sigma \\ 0 & \text{sinon} \end{cases}.$$

Il suffit de définir les fonctions $f_\sigma(t, X) := f(t, X, \xi_\sigma)$ où ξ_σ est un rayon distingué passant par σ à la n-ième génération. Notons que les fonctions f_σ sont bien définies, puisqu'il n'existe qu'un seul rayon ξ_σ dans $\mathcal{F}_n$ passant par σ, et qu'elles sont effectivement $\mathcal{F}_n$-mesurables.

Sur l'espace des arbres marqués avec rayons distingués, soit μ^* une mesure telle que la restriction μ_n^* sur $\mathcal{F}_n^*$ est une mesure de dénombrement des ensembles $\{\sigma : |\sigma| = n\}$, et telle que la mesure μ^* reste cohérente avec μ lorsque projetée sur $\mathcal{F}_n$. Autrement dit, pour toute fonction non-négative $\mathcal{F}_n^*$-mesurable f, μ^* est la mesure telle que

$$\int f(t, X, \xi) d\mu_n^* = \int \sum_{|\sigma|=n} f_\sigma(t, X) d\mu_n.$$

Il est clair que, puisque μ_n^* dénombre les individus à la génération n, μ_n^* n'est pas une mesure de probabilité, contrairement à μ_n.

Une façon intuitive de voir μ_n^* est la suivante : pour un ensemble $\mathcal{F}_n^*$-mesurable de l'espace des arbres marqués avec rayons distingués, μ_n^* compte les individus des arbres de la n-ième génération en prenant en considération la probabilité des arbres marqués avec rayon dictée par les variables L et $\mathcal{L}$, tout comme le fait μ_n. Autrement dit, μ_n^* attribue la mesure 1 à chacun des rayons passant par les individus de la génération n, tout en restant cohérant avec la mesure μ_n sur l'arbre marqué : pour (t, X) un arbre marqué, nous avons

$$\mu_n^*(\{(t', X', \xi) : (t, X) = (t', X')\}) = Z_n(t)\mu_n([t, X])$$

3.4.5. *Promenade aléatoire avec branchement : une nouvelle construction.* Nous cherchons à définir une mesure $\hat{\mu}^*$, sur Ω^*, tel que, si $\hat{\mu}_n^*$ est la restriction de $\hat{\mu}^*$ sur $\mathcal{F}_n^*$, alors

$$\frac{d\hat{\mu}_n^*}{d\mu_n^*}(t, X, \xi) = \frac{e^{-\alpha S(\xi_n)}}{m(\alpha)^n}.$$

En projetant sur l'espace des arbres marqués, l'équation

$$\frac{d\hat{\mu}}{d\mu}(t, X) = \frac{\sum_{|\sigma|=n} e^{-\alpha S(\sigma)}}{m(\alpha)^n} = W_n(\alpha)$$

est alors obtenue, où $\hat{\mu}$ est la projection de $\hat{\mu}^*$ sur Ω. Pour y parvenir, une nouvelle façon de construire des promenades aléatoires avec branchement est nécessaire.

La restriction de μ à $\mathcal{F}_1$, noté par μ_1, est la mesure canonique induite par $\mathcal{L}$, sur les arbres de hauteur 1. μ_1 ne mesure que les familles de 1 génération. Soit $\hat{\mathcal{L}}$ un élément aléatoire tel que, si $\hat{\mu}_1$ est la mesure canonique induite par $\hat{\mathcal{L}}$ sur $\mathcal{F}_1$, alors

$$\frac{d\hat{\mu}_1}{d\mu_1} = \frac{\langle \alpha, \mathcal{L} \rangle}{m(\alpha)}.$$

Une observation importante est que la progéniture générée à partir de copies de $\hat{\mathcal{L}}$ est toujours de taille supérieure à 0. En effet, pour t un arbre tel que la racine r n'a aucun enfant,

$$\begin{aligned}
\hat{\mu}_1([t, X]_1) &= \frac{\langle \alpha, \mathcal{L}(r) \rangle}{m(\alpha)} \cdot \mu_1([t, X]_1) \\
&= \frac{\sum_i^{L=0} e^{-\alpha X_i}}{m(\alpha)} \cdot \mu_1([t, X]_1) \\
&= 0
\end{aligned}$$

La promenade est initiée avec une première particule v_0 située à l'origine. À v_0 est associée une progéniture avec déplacements $\hat{\mathcal{L}}_1$, une copie i.i.d. à $\hat{\mathcal{L}}$. Parmi les enfants de v_0, un individu σ est choisi aléatoirement, avec probabilité proportionnelle à $e^{\alpha X(\sigma)}$. Cet individu particulier sera appelé v_1. Pour les autres enfants de v_0, la promenade se poursuit avec des copies i.i.d. à $\mathcal{L}$, alors que v_1 se reproduit avec une progéniture avec déplacements $\hat{\mathcal{L}}_2$, i.i.d. à $\hat{\mathcal{L}}$. À la génération suivante, un enfant de v_1, appelé v_2, est choisi avec une probabilité proportionnelle à $e^{\alpha X(v_2)}$, et ainsi de suite. Comme les copies de $\hat{\mathcal{L}}$ donnent toujours un nombre d'enfants supérieur à 0, la promenade aléatoire obtenue ne s'arrête jamais. Cette nouvelle construction sera appelé une *promenade biaisée par la taille.*

3.4.6. *$\hat{\mu}^*$ et $\hat{\mu}$.* Soit $\hat{\mu}^*$ la mesure canonique sur Ω^* induite par la promenade biaisée. Soit $\hat{\mu}_n^*$ la restriction de $\hat{\mu}^*$ à $\mathcal{F}_n^*$.

Soit (t, X, ξ) un arbre marqué avec rayon, et soit k le nombre d'enfants de la racine. La promenade aléatoire sans biais commence avec k enfants ayant des déplacement X_i. À la deuxième génération, (t, X, ξ) se construit de la même façon pour tous les noeuds. Donc,

$$\mu_n^*([t, X, \xi]_n) = \mu_1([t, X]_1) \cdot \mu_{n-1}^*([(t, X, \xi)_{\xi_1}]_{n-1}) \cdot \prod_{|\sigma|=1, \sigma \neq \xi_1} \mu_{n-1}([t_\sigma, X_\sigma]_{n-1}).$$

La promenade biaisée, quant à elle, commence avec le choix d'un individu ξ_1 avec probabilité $\frac{e^{-\alpha X(\xi_1)}}{\langle \alpha, \mathcal{L}(\xi_0) \rangle}$ parmi k enfants ayant fait des déplacements X_i. À la deuxième génération, (t, X, ξ) se construit selon la promenade biaisée pour t_{ξ_1}, et selon une promenade aléatoire sans biais sur les autres sous-arbres t_σ. De la même façon,

$$\hat{\mu}_n^*([t, X, \xi]_n) = \hat{\mu}_1([t, X]_1) \cdot \frac{e^{-\alpha X(\xi_1)}}{\langle \alpha, \mathcal{L}(\xi_0) \rangle} \cdot \hat{\mu}_{n-1}^*([(t, X, \xi)_{\xi_1}]_{n-1}) \cdot \prod_{|\sigma|=1, \sigma \neq \xi_1} \mu_{n-1}([t_\sigma, X_\sigma]_{n-1}).$$

Puisque

$$\frac{d\hat{\mu}_1}{d\mu_1} = \frac{\langle \alpha, \mathcal{L} \rangle}{m(\alpha)},$$

nous avons que

$$\hat{\mu}_1([t, X]_1) = \frac{\langle \alpha, \mathcal{L}(\xi_0) \rangle}{m(\alpha)} \cdot \mu_1([t, X]_1).$$

En substituant dans l'équation ci-haut, on obtient que

$$\begin{aligned}\hat{\mu}_n^*([t,X,\xi]_n) &= \frac{\langle\alpha,\mathcal{L}(\xi_0)\rangle}{m(\alpha)}\cdot\mu_1([t,X]_1)\cdot\frac{e^{-\alpha X(\xi_1)}}{\langle\alpha,\mathcal{L}(\xi_0)\rangle}\cdot\hat{\mu}_{n-1}^*([(t,X,\xi)_{\xi_1}]_{n-1})\\ &\cdot\prod_{|\sigma|=1,\sigma\neq\xi_1}\mu_{n-1}([t_\sigma,X_\sigma]_{n-1})\\ &= \frac{e^{-\alpha X(\xi_1)}}{m(\alpha)}\cdot\frac{\hat{\mu}_{n-1}^*([(t,X,\xi)_{\xi_1}]_{n-1})}{\mu_{n-1}^*([(t,X,\xi)_{\xi_1}]_{n-1})}\cdot\mu_n^*([t,X,\xi]_n).\end{aligned}$$

Ensuite, en itérant avec les $\hat{\mu}_{n-i}^*([(t,X,\xi)_{\xi_i}]_{n-i})$,

$$\hat{\mu}_n^*([t,X,\xi]_n) = (\prod_{i=1}^n\frac{e^{-\alpha X(\xi_i)}}{m(\alpha)})\cdot\mu_n^*([t,X,\xi]_n),$$

d'où

$$\begin{aligned}\frac{d\hat{\mu}_n^*}{d\mu_n^*} &= \prod_{i=1}^n\frac{e^{-\alpha X(\xi_i)}}{m(\alpha)}\\ &= \frac{e^{-\alpha\sum_{i=1}^n X(\xi_i)}}{m(\alpha)^n}\\ &= \frac{e^{-\alpha S(\xi_n)}}{m(\alpha)^n}.\end{aligned}$$

Soit $\hat{\mu}$ la projection de $\hat{\mu}^*$ sur l'espace des arbres marqués et soit $\hat{\mu}_n$ la restriction de $\hat{\mu}$ à $\mathcal{F}_n$. Il est clair que la densité de $\hat{\mu}_n$ par rapport à μ_n est celle recherchée :

$$\begin{aligned}\frac{d\hat{\mu}_n}{d\mu_n}(t,X) &= \frac{\sum_{|\sigma|=n}e^{-\alpha S(\sigma)}}{m(\alpha)^n}\\ &= W_n(t,X).\end{aligned}$$

3.4.7. *Lemme dichotomique.* Le comportement asymptotique de W_n sous la mesure $\hat{\mu}$ permettra d'obtenir le théorème de Biggins. Ainsi, par le lemme suivant, la convergence de la martingale W_n sous μ, la promenade aléatoire initiale, est en relation directe avec son comportement sous $\hat{\mu}$, la promenade biaisée.

Lemme. *Soit μ une mesure finie, et ν une mesure de probabilité sur une tribu $\mathcal{F}$. Soit $\mathcal{F}_n$ une filtration telle que $\cup\mathcal{F}_n = \mathcal{F}$ et telle que μ_n est absolument continue par rapport à ν_n, où μ_n et ν_n sont respectivement la restriction de μ et de ν à $\mathcal{F}_n$. Soit $X_n := \frac{d\mu_n}{d\nu_n}$ et soit $X := \limsup_{n\to\infty} X_n$. Alors*

$$\mu \ll \nu \iff X < \infty\,\mu\text{-p.p.} \iff \int X d\nu = \int d\mu$$

et

$$\mu\perp\nu \iff X = \infty\,\mu\text{-p.p.} \iff \int X d\nu = 0.$$

La démonstration est donnée dans l'annexe 4.1.

En appliquant le lemme à $\hat{u}$ et μ tel que défini dans le cas de la promenade, les deux équivalences suivantes sont obtenues :

$$W = \infty\,\hat{\mu}\text{-p.p.} \iff W = 0\,\mu\text{-p.s.}$$

et

$$W < \infty\ \hat{\mu}\text{-p.p.} \iff \int W d\mu = \int d\hat{\mu} = 1.$$

Le comportement de W par rapport à la mesure $\hat{\mu}$ donne donc les conditions dictées par le théorème de Biggins, et c'est d'ailleurs cette dichotomie qui permettra de terminer la démonstration du théorème.

3.4.8. *L'implication 2) ⇒ 4).* Il faut démontrer que

$$P[W(\alpha) = 0] < 1$$

implique

$$E[(\alpha, \mathcal{L}) \log (\alpha, \mathcal{L})] < \infty \text{ et } \alpha m'(\alpha)/m(\alpha) < \log m(\alpha).$$

Ceci est fait par la contrapositive. Nous commençons en premier lieu par certaines observations.

Soit $\alpha \in \mathbb{R}$ fixé. Soit (t, X) un arbre marqué, soit $(v_0, v_1, \ldots)$ un rayon distingué de (t, X) tel que utilisé dans la promenade biaisée, et soit

$$W_n(t, X) := \frac{\sum_{|\sigma|=n} e^{-\alpha S(\sigma)}}{m(\alpha)^n}.$$

Considérons

$$\begin{aligned} W_{n+1}(t, X) &= \frac{\sum_{|\sigma|=n+1} e^{-\alpha S(\sigma)}}{m(\alpha)^{n+1}} \\ &= \frac{\sum_{|\sigma|=n} (e^{-\alpha S(\sigma)} (\alpha, \mathcal{L}_\sigma))}{m(\alpha)^{n+1}} \\ &\geq \frac{e^{-\alpha S(v_n)} (\alpha, \mathcal{L}_{v_n})}{m(\alpha)^{n+1}} \end{aligned} \tag{3.1}$$

Notons qu'au numérateur, les deux termes sont $\hat{\mu}^*$−indépendants.

Pour tout $k \geq 1$, v_k est choisi parmi les enfants de v_{k-1} avec probabilité proportionnelle à $e^{-\alpha X_{v_k}}$. Soit $L(v_{k-1})$ le nombre d'enfants de v_{k-1}. Donc, la $\hat{\mu}$-espérance du déplacement de v_k est :

$$\begin{aligned} \int X(v_k)\, d\hat{\mu} &= \int \Big(\sum_{i=1}^{L(v_{k-1})} X_i \frac{e^{-\alpha X_i}}{\langle \alpha, \mathcal{L}(v_{k-1}) \rangle} \Big) \cdot d\hat{\mu} \\ &= \int \Big(\sum_{i=1}^{L(v_{k-1})} X_i \frac{e^{-\alpha X_i}}{\langle \alpha, \mathcal{L}(v_{k-1}) \rangle} \Big) \cdot \frac{d\hat{\mu}_1}{d\mu_1} \langle [t_{v_{k-1}}, X]_1 \rangle \cdot d\mu \\ &= E\Big[\Big(\sum_{i=1}^{L(v_{k-1})} X_i \frac{e^{-\alpha X_i}}{\langle \alpha, \mathcal{L}(v_{k-1}) \rangle} \Big) \cdot \frac{d\hat{\mu}_1}{d\mu_1} ([t_{v_{k-1}}, X]_1)\Big] \\ &= E\Big[\sum_{i=1}^{L(v_{k-1})} X_i \frac{e^{-\alpha X_i}}{\langle \alpha, \mathcal{L}(v_{k-1}) \rangle} \cdot \frac{\langle \alpha, \mathcal{L}(v_{k-1}) \rangle}{m(\alpha)} \Big] \\ &= E\Big[\frac{\sum_{i=1}^{L} X_i e^{-\alpha X_i}}{m(\alpha)} \Big] \\ &= \frac{-m'(\alpha)}{m(\alpha)}. \end{aligned}$$

Par la loi des grands nombres, on obtient que $\frac{S(v_n)}{n} \rightarrow \frac{-m'(\alpha)}{m(\alpha)}$ $\hat{\mu}$-p.p. Supposons que 4) ne soit pas vrai. Supposons d'abord que

$$\frac{\alpha m'(\alpha)}{m(\alpha)} > \log m(\alpha)$$

d'où

$$\begin{aligned} -\frac{\alpha S(v_n)}{n} &\rightarrow \frac{\alpha m'(\alpha)}{m(\alpha)} \\ &> \log m(\alpha) \ \hat{\mu}\text{-p.p.} \end{aligned}$$

et donc $e^{-\alpha S(v_n)}$ croît de façon surexponentielle, d'où

$$\limsup_{n\to\infty} \frac{e^{-\alpha S(v_n)}}{m(\alpha)^n} = \infty.$$

Pour tout n, il a été observé avec l'inégalité 3.1 que

$$W_n(t, X) \geq \frac{e^{-\alpha S(v_n)}\langle \alpha, \mathcal{L}_{v_n}\rangle}{m(\alpha)^{n+1}}.$$

Or, $\frac{\langle \alpha, \mathcal{L}_{v_n}\rangle}{m(\alpha)} > 0$ $\hat{\mu}$-p.p., puisque $\mathcal{L}_{v_n}$ est une copie de $\hat{\mathcal{L}}$ sous la mesure $\hat{\mu}$, et donc,

$$W(t, X) := \limsup_{n\to\infty} W_n(t, X) = \infty \ \hat{\mu}\text{-p.p.}$$

En appliquant le lemme dichotomique à la martingale $W_n(t, X)$, les deux équivalences :

$$W(t, X) = \infty \ \hat{\mu}\text{-p.p.} \Longleftrightarrow W(t, X) = 0 \ \mu\text{-p.s.}$$

et

$$W(t, X) < \infty \, \hat{\mu}\text{-p.p.} \Longleftrightarrow \int W(t, X) d\mu = 1.$$

Ainsi, $W(t, X) = 0$ μ-p.s, et $P[W = 0] = 1$, contredisant 2).

Toujours en admettant que 4) ne soit pas vrai, supposons maintenant que

$$\frac{\alpha m'(\alpha)}{m(\alpha)} < \log m(\alpha).$$

Alors

$$E[\langle \alpha, \mathcal{L}\rangle \log\langle \alpha, \mathcal{L}\rangle] = \infty.$$

Donc,

$$\begin{aligned} E_{\hat{\mu}}[\log\langle \alpha, \mathcal{L}_{v_n}\rangle] &= E_{\hat{\mu}_1}[\log\langle \alpha, \mathcal{L}_{v_n}\rangle] \\ &= \int \log\langle \alpha, \mathcal{L}_{v_n}\rangle d\hat{\mu}_1 \\ &= \int \frac{d\hat{\mu}_1}{d\mu_1} \log\langle \alpha, \mathcal{L}_{v_n}\rangle \cdot d\mu_1 \\ &= \int \frac{\langle \alpha, \mathcal{L}_{v_n}\rangle \log\langle \alpha, \mathcal{L}_{v_n}\rangle}{m(\alpha)} \cdot d\mu_1 \\ &= \frac{E_\mu[\langle \alpha, \mathcal{L}\rangle \log\langle \alpha, \mathcal{L}\rangle]}{m(\alpha)} \\ &= \infty, \end{aligned}$$

puisque les $\mathcal{L}_{v_n}$ sont des copies de $\mathcal{L}$ sous la mesure μ.

Lemme. *Soit X, X_1, X_2, ... des variables aléatoires i.i.d. non-négatives i.i.d. Alors, p.s.*

$$\limsup_{n\to\infty} \frac{1}{n} X_n = \begin{cases} 0 & if\, E[X] < \infty \\ \infty & if\, E[X] = \infty \end{cases}$$

La démonstration de ce lemme est donnée dans l'annexe en 4.1. Puisque $E_{\hat{\mu}}[\log\langle\alpha, \mathcal{L}_{v_n}\rangle] = \infty$, on obtient, par le lemme,

$$\limsup_{n\to\infty} \frac{1}{n} \log\langle\alpha, \mathcal{L}_{v_n}\rangle = \infty \ \hat{\mu}\text{-p.p.}$$

En reprenant l'inégalité 3.1,

$$W_{n+1}(t, X) \geq \frac{e^{-\alpha S(v_n)}}{m(\alpha)^{n+1}} \langle\alpha, \mathcal{L}_{v_n}\rangle \ \hat{\mu}\text{-p.p.},$$

on observe que le premier terme décroit de façon exponentielle, alors que le deuxième terme explose de façon sur-exponentielle, $\hat{\mu}$-p.p. De nouveau, $W(t, X) = \infty$, $\hat{\mu}$-p.p.

3.4.9. *L'implication 4) $\Rightarrow$ 3).* Nous cherchons à démontrer que

$$E[\langle\alpha, \mathcal{L}\rangle \ \log\langle\alpha, \mathcal{L}\rangle] < \infty \text{ et } \alpha m'(\alpha)/m(\alpha) < \log m(\alpha)$$

implique

$$E[W(\alpha)] = 1.$$

Soit $\alpha \in \mathbb{R}$, fixé. Soit (t, X) un arbre marqué, soit $(v_0, v_1, ...)$ un rayon distingué de (t, X) tel qu'utilisé dans la promenade biaisée et soit $W_n(t, X)$ tel que défini plus haut en 3.4.8. Considérons $E_{\hat{\mu}^*}[W_n(t, X)|\mathcal{G}]$, où $\mathcal{G}$ est la tribu engendrée par

$\{\hat{\mathcal{L}}_k\}_{k\geq 1}$. L'égalité suivante est obtenue :

$$\begin{aligned}
E_{\hat{\mu}}[W_n(t,X)|\mathcal{G}] &= E_{\hat{\mu}}[\frac{\sum_{|\sigma|=n} e^{-\alpha S(\sigma)}}{m(\alpha)^n}|\mathcal{G}] \\
&= E_{\hat{\mu}}[\frac{1}{m(\alpha)^n}[\sum_{|\sigma|=n,\,\sigma>v_{n-1}} e^{-\alpha S(\sigma)} + \sum_{|\sigma|=n,\,\sigma\not>v_{n-1}} e^{-\alpha S(\sigma)}]|\mathcal{G}] \\
&= E_{\hat{\mu}}[\frac{1}{m(\alpha)^n}[e^{-\alpha S(v_{n-1})}\langle\alpha,\mathcal{L}_{v_{n-1}}\rangle + \sum_{|\sigma|=n-1,\,\sigma\neq v_{n-1}} e^{-\alpha S(\sigma)}\langle\alpha,\mathcal{L}_\sigma\rangle]|\mathcal{G}] \\
&= \frac{e^{-\alpha S(v_{n-1})}\langle\alpha,\mathcal{L}_{v_{n-1}}\rangle}{m(\alpha)^n} + \frac{1}{m(\alpha)^n}E_{\hat{\mu}}[\sum_{|\sigma|=n-1,\,\sigma\neq v_{n-1}} e^{-\alpha S(\sigma)}\langle\alpha,\mathcal{L}_\sigma\rangle]|\mathcal{G}] \\
&= \frac{e^{-\alpha S(v_{n-1})}\langle\alpha,\mathcal{L}_{v_{n-1}}\rangle}{m(\alpha)^n} + \frac{1}{m(\alpha)^{n-1}}E_{\hat{\mu}}[\sum_{|\sigma|=n-1,\,\sigma\neq v_{n-1}} e^{-\alpha S(\sigma)}|\mathcal{G}] \\
&= \frac{e^{-\alpha S(v_{n-1})}\langle\alpha,\mathcal{L}_{v_{n-1}}\rangle}{m(\alpha)^n} + \frac{1}{m(\alpha)^{n-1}}E_{\hat{\mu}}[\sum_{|\sigma|=n-1} e^{-\alpha S(\sigma)} - e^{-\alpha S(v_{n-1})}]|\mathcal{G}] \\
&= \frac{e^{-\alpha S(v_{n-1})}\langle\alpha,\mathcal{L}_{v_{n-1}}\rangle}{m(\alpha)^n} - \frac{e^{-\alpha S(v_{n-1})}}{m(\alpha)^{n-1}} + E_{\hat{\mu}}[\frac{\sum_{|\sigma|=n-1} e^{-\alpha S(\sigma)}}{m(\alpha)^{n-1}}]|\mathcal{G}] \\
&= \frac{e^{-\alpha S(v_{n-1})}\langle\alpha,\mathcal{L}_{v_{n-1}}\rangle}{m(\alpha)^n} - \frac{e^{-\alpha S(v_{n-1})}}{m(\alpha)^{n-1}} + E_{\hat{\mu}}[W_{n-1}(t,X)|\mathcal{G}] \\
&= \sum_{k=0}^{n-1} \frac{e^{-\alpha S(v_k)}}{m(\alpha)^{k+1}}\langle\alpha,\mathcal{L}_{v_k}\rangle - \sum_{k=1}^{n-1}\frac{e^{-\alpha S(v_k)}}{m(\alpha)^k}.
\end{aligned}$$

La dernière égalité est obtenue par récursion sur les $E_{\hat{\mu}}[W_{n-i}(t,X)|\mathcal{G}]$. Puisque $\alpha m'(\alpha)/m(\alpha) < \log m(\alpha)$, et puisque

$$\frac{S(v_n)}{n} \to \frac{-m'(\alpha)}{m(\alpha)}\ \hat{\mu}\text{-p.p.},$$

on obtient que

$$\begin{aligned}
\frac{-\alpha S(v_n)}{n} &\to \frac{\alpha m'(\alpha)}{m(\alpha)} \\
&< \log m(\alpha)\ \hat{\mu}\text{-p.p.}
\end{aligned}$$

d'où

$$\liminf_{n\to\infty}\frac{e^{-\alpha S(v_n)}}{m(\alpha)^n} < 1.$$

Donc, $\frac{e^{-\alpha S(v_n)}}{m(\alpha)^n}$ décroît exponentiellement $\hat{\mu}$-p.p. En procédant comme en 3.4.8., par hypothèse,

$$\begin{aligned}
E_{\hat{\mu}}[\log\langle\alpha,\mathcal{L}_{v_n}\rangle] &= \frac{E_\mu[\langle\alpha,\mathcal{L}\rangle\log\langle\alpha,\mathcal{L}\rangle]}{m(\alpha)} \\
&< \infty.
\end{aligned}$$

Par le lemme 3 de la sous-section 4.1, nous obtenons

$$\limsup_{n\to\infty}\frac{1}{n}\log\langle\alpha,\mathcal{L}_{v_n}\rangle = 0\ \hat{\mu}\text{-p.p.}$$

ce qui implique que $\langle \alpha, \mathcal{L}_{v_n} \rangle$ croît au plus de façon sous-exponentielle, $\hat{\mu}$-p.p. Ainsi, les deux séries

$$\sum_{k=0}^{n-1} \frac{e^{-\alpha S(v_k)}}{m(\alpha)^{k+1}} \langle \alpha, \mathcal{L}_{v_k} \rangle$$

et

$$\sum_{k=1}^{n-1} \frac{e^{-\alpha S(v_k)}}{m(\alpha)^{k}}$$

convergent et donc $E_{\hat{\mu}}[W_n(t,X)|\mathcal{G}] < \infty$. Par le lemme de Fatou,

$$\begin{aligned} E_{\hat{\mu}}[\liminf_{n\to\infty} W_n(t,X)] &\leq \liminf_{n\to\infty} E_{\hat{\mu}}[W_n(t,X)] \\ &< \infty \end{aligned}$$

d'où $\liminf_{n\to\infty} W_n(t,X) < \infty$ $\hat{\mu}$-p.p.

Puisque

$$\frac{d\hat{\mu}_n}{d\mu_n}(t,X) = W_n(t,X),$$

$\{\frac{1}{W_n(t,X)}\}$ est une $\hat{\mu}$-martingale non-négative. Par le théorème de convergence des martingales, $\{\frac{1}{W_n(t,X)}\}$ converge $\hat{\mu}$-p.p., et donc, $\{W_n(t,X)\}$ converge $\hat{\mu}$-p.p. vers $\liminf_{n\to\infty} W_n(t,X) < \infty$, donc $W(t,X) < \infty$ $\hat{\mu}$-p.p. Par le lemme dichotomique, $\int W(t,X)\, d\mu = 1$, ou $E[W(t,X)] = 1$ d'où 3) suit.

Corollaire. *Avec $\alpha = 0$, nous obtenons le théorème de Kesten-Stigum.*

4. Annexe

4.1. Lemmes.

Lemme 1. *Étant donné la non-extinction de l'arbre, toute propriété héritée a une probabilité de 0 ou de 1.*

Démonstration. Soit A l'ensemble des arbres ayant la propriété héritée. Soit t un arbre, avec k enfants pour la racine de t. Soit $t_1, ..., t_k$ les sous-arbres descendant de la racine de t. Alors, par la définition de la propriété héritée, et par indépendence des t_i,

$$\begin{aligned} P[A] &= E[P[t \in A|Z_1]] \\ &\leq E[P[t_1 \in A, ..., t_k \in A|Z_1]] \\ &= E[P[A]^{Z_1}]. \end{aligned}$$

Soit $f(s) := E[s^{Z_1}]$ la fonction génératrice de probabilité, et donc

$$E[P[A]^{Z_1}] = f(P[A])$$

d'où $P[A] \leq f(P[A])$. De plus, puisque tous les arbres finis possèdent la propriété héritée, nous avons que $P[A] \geq q$, où q est la probabilité d'extinction de l'arbre. Par le fait que $f(s)$ est une fonction convexe sur $[0,1]$, et par inspection du graphe de f (voir [9]), il se trouve que les seuls choix pour $P[A]$ sont $P[A] = 1$ si $E[Z_1] \leq 1$, et $P[A] = q$ ou 1 si $E[Z_1] > 1$. Étant donné la non-extinction, $P[A] \in \{0,1\}$. $\square$

Lemme 2. [7] *Soit μ une mesure finie, et ν une mesure de probabilité sur une tribu $\mathcal{F}$. Soit $\mathcal{F}_n$ une filtration telle que $\cup\mathcal{F}_n = \mathcal{F}$ et telle que μ_n est absolument continue par rapport à ν_n, où μ_n et ν_n sont respectivement la restriction de μ et de ν à $\mathcal{F}_n$. Soit $X_n := \frac{d(\mu|\mathcal{F}_n)}{d(\nu|\mathcal{F}_n)}$ et soit $X := \limsup_{n\to\infty} X_n$. Alors*

$$\mu \ll \nu \iff X < \infty\, \mu\text{-p.p.} \iff \int X d\nu = \int d\mu$$

et

$$\mu \perp \nu \iff X = \infty\, \mu\text{-p.p.} \iff \int X d\nu = 0$$

Démonstration. X_n est une ν-martingale non-négative. Elle converge donc vers X ν-p.s. et $X < \infty$ ν-p.s. Nous cherchons à obtenir la décomposition de μ en une partie ν-absolument continue et ν-singulière. Autrement dit, pour tout $A \in \mathcal{F}$,

$$\mu(A) = \int_A X d\nu + \mu(A \cap \{X = \infty\}).$$

Supposons en premier lieu que $\mu \ll \nu$. Soit $\tilde{X} := \frac{d\mu}{d\nu}$. Ainsi, $X_n = E_\nu[\tilde{X}|\mathcal{F}_n]$. Par le théorème de convergence des martingales, $X_n \to \tilde{X}$ ν-p.s. Par la définition de X, $X = \tilde{X} = \frac{d\mu}{d\nu}$. La décomposition voulue suit de la définition de la dérivée de Radon-Nykodym.

Dans le cas général, posons $C := \int d(\mu+\nu)$, et soit $\rho := (\mu+\nu)/C$. Il est clair que $\mu, \nu \ll \rho$. Soit $U_n := \frac{d(\mu|\mathcal{F}_n)}{d(\rho|\mathcal{F}_n)}$ et $V_n := \frac{d(\nu|\mathcal{F}_n)}{d(\rho|\mathcal{F}_n)}$, et soit $U := \limsup_{n\to\infty} U_n$ et $V := \limsup_{n\to\infty} V_n$. Il est à noter que $\frac{U_n}{V_n} = \frac{d(\mu|\mathcal{F}_n)}{d(\nu|\mathcal{F}_n)} = X_n$.

On obtient

$$\begin{aligned}
\int (U_n + V_n) d(\rho|\mathcal{F}_n) &= \int [\frac{d(\mu|\mathcal{F}_n)}{d(\rho|\mathcal{F}_n)} + \frac{d(\nu|\mathcal{F}_n)}{d(\rho|\mathcal{F}_n)}] d(\rho|\mathcal{F}_n) \\
&= \int d(\mu + \nu|\mathcal{F}_n) \\
&= C.
\end{aligned}$$

Donc, $U_n + V_n = C$ ρ-p.s. d'où $\rho(\{U = 0, V = 0\}) = 0$. Ainsi,

$$\frac{U}{V} = \frac{\lim_{n\to\infty} U_n}{\lim_{n\to\infty} V_n} = \lim_{n\to\infty} (\frac{U_n}{V_n}) = \lim X_n = X\ \rho\text{-p.s.}$$

Pour tout $A \in \mathcal{F}$,

$$\begin{aligned}
\mu(A) &= \int_A U d\rho \\
&= \int_A 1_{\{V\neq 0\}} U d\rho + \int_A 1_{\{V=0\}} U d\rho \\
&= \int_A 1_{\{V\neq 0\}} X V d\rho + \int_A 1_{\{V=0\}} U d\rho \\
&= \int_A X d\nu + \mu(A \cap \{X = \infty\}),
\end{aligned}$$

ce qui donne la décomposition voulue de μ en une partie ν-singulière et ν-absolument continue.

La démonstration du lemme découle de la décomposition. Si $\mu \ll \nu$, $\mu(A \cap \{X = \infty\}) = 0$ pour tout $A \in \mathcal{F}$, et donc $X < \infty$, μ-p.p. Supposons ensuite que

$X < \infty$ μ-p.p., alors $\int_A X\,d\nu = \mu(A) = \int_A d\mu$. Finalement, si $\int X\,d\nu = \int d\mu$, par la décomposition, on obtient que $X < \infty$ μ-p.p. et $\mu \ll \nu$.

Pour la deuxième équivalence, si $\mu \perp \nu$, il existe $A \in \mathcal{F}$ tel que $\mu(A^c) = 0$ et $\nu(A) = 0$. Or, si $\nu(A) = 0$, nous avons $\int_A X\,d\nu = 0$ et donc $\mu(A) = \mu(A \cap \{X = \infty\})$, ou $X = \infty$ μ-p.p. Inversement, si $X = \infty$ μ-p.p., alors

$$\begin{aligned} \mu(A) &= \int_A X\,d\nu + \mu(A \cap \{X = \infty\}) \\ &= \int_A X\,d\nu + \mu(A) \end{aligned}$$

d'où $\int X\,d\nu = 0$. Finalement, par la même décomposition, si $\int X\,d\nu = 0$, alors $X = \infty$ μ-p.p., et donc $\mu \perp \nu$. □

Lemme 3. *Soit X, X_1, X_2, ... des variables aléatoires i.i.d. non-négatives. Alors, p.s.*

$$\limsup_{n\to\infty} \frac{1}{n} X_n = \begin{cases} 0 & if\, E[X] < \infty \\ \infty & if\, E[X] = \infty \end{cases}$$

Démonstration. Nous démontrons premièrement

$$E[X] = \int_0^\infty P(X > x)\,dx$$

pour $X \geq 0$. Par le théorème de Fubini, puisque $P(X > x) = \int_\Omega 1_{\{\omega\,:\,X(\omega)>x\}}(\omega)\,dP(\omega)$,

$$\begin{aligned} \int_0^\infty P(X > x)\,dx &= \int_0^\infty \int_\Omega 1_{\{\omega\,:\,X(\omega)>x\}}(\omega)\,dP(\omega)\,dx \\ &= \int_\Omega \int_0^\infty 1_{\{\omega\,:\,X(\omega)>x\}}(\omega)\,dx\,dP(\omega) \\ &= \int_\Omega \int_0^{X(\omega)} dx\,dP(\omega) \\ &= \int_\Omega X(\omega)\,dP(\omega) \\ &= E[X]. \end{aligned}$$

Le lemme suit directement : pour $X \geq 0$,

$$\begin{aligned} E[X] &= \int_0^\infty P(X > x)\,dx \\ &= A\int_0^\infty P(X > Ax)\,dx \\ &< A\sum_{n=0}^{\infty} P(X_n > An). \end{aligned}$$

Dans le cas $E[X] = \infty$, alors $\sum_{n=0}^{\infty} P(X_n > An)$ diverge pour tout A, alors, par le lemme de Borel-Cantelli, $P(\limsup_{n\to\infty} \frac{X_n}{n} > A) = 1$ ce qui implique que

$\limsup_{n\to\infty} \frac{X_n}{n} = \infty$ p.s. Maintenant, si $E[X] < \infty$, alors

$$\begin{aligned} E[X] &= \int_0^\infty P(X > x)\,dx \\ &= A\int_0^\infty P(X > Ax)\,dx \\ &> A\sum_{n=1}^{\infty} P(X_n > An). \end{aligned}$$

Pour un suffisamment petit A, $\sum_{n=0}^{\infty} P(X_n > An)$ converge et $P(\limsup_{n\to\infty} \frac{X_n}{n} < A) = 1$ par le lemme de Borel-Cantelli. Donc, $\limsup_{n\to\infty} \frac{X_n}{n} = 0$ p.s. □

4.2. **Théorème de convergence des martingales.**

Théorème. [9] *Soit W_n une martingale telle que*

$$sup_n E[W_n] < \infty.$$

Alors, p.s., sa limite

$$W := lim_{n\to\infty} W_n$$

existe et est finie.

RÉFÉRENCES

[1] R. Durrett, Probability : Theory and Examples. 1991, Wadsworth, Pacific Grove, CA.

[2] C. C. Heyde et E. Seneta, I. J. Bienaymé : Statistical Theory Anticipated. 1977, Springer-Verlag, New York, NY.

[3] A. Joffe, A New Martingale in Branching Random Walk, Ann. Applied Proba., Vol. 3, No 4 (1993) 1145-1159.

[4] A. Joffe et A. R. Moncayo, Random Variables, Trees, and Branching Random Walks, Adv. Math. Vol. 10 (1973) 401-416.

[5] A. Joffe et W. A. O'N. Waugh, Exact distributions of kin numbers in a Galton-Watson process, J. Appl. Prob., Vol. 19 (1982) 767-775.R. Lyons, A Simple Path to Biggins' Martingale Convergence for Branching Random Walk, Classical and Modern Branching Processes (Minneapolis, MN, 1994), IMA Vol. Math. Appl., Vol. 84, Springer, New York, 1997, 217-221.

[6] R. Lyons, R. Pemantle et Y. Peres, Conceptual Proofs of L log L Criteria for Mean Behavior of Branching Processes, Ann. Proba., Vol. 23, No. 3 (1995) 1125-1138

[7] R. Lyons et Y. Peres, Probability on Trees and Networks, Cambridge University Press, In preparation.

[8] J. Neveu, Arbres et processus de Galton-Watson, Ann. Inst. H. Poincaré.

[9] D. Williams, Probability with Martingales. 1991, Cambridge University Press, Cambridge.