Processus de branchement et populations structurées

Vincent Bansaye.

Notes provisoires.

Table des matières

1 Arbre de Galton Watson **3**
1.1 Processus de Galton-Watson . . . 3
1.1.1 Introduction . . . 3
1.1.2 Probabilité d'extinction . . . 5
1.1.3 Transformation de Lamperti . . . 5
1.2 Arbre aléatoire et construction spinale . . . 7
1.2.1 Définitions . . . 7
1.2.2 Construction spinale . . . 8
1.2.3 Résultats . . . 8
1.3 Processus de Galton-Watson avec immigration . . . 11
1.3.1 Croissance dans le cas surcritique . . . 12
1.3.2 Les régimes critiques et sous critiques : $m \leq 1$. . . . 13
1.4 Croissance malthusienne et vitesse d'exinction . . . 14
1.4.1 Croissance dans le cas surcritique . . . 14
1.4.2 Vitesse d'extinction dans le cas sous-critique . . . 17
1.5 Galton-Watson en environnement aléatoire . . . 18

2 Populations structurées : le cas neutre **21**
2.1 Chaîne de Markov sur un arbre de Galton-Watson . . . 21
2.1.1 Définition du modèle et exemples . . . 21
2.1.2 Propriété de branchement et semigroupe . . . 24
2.1.3 Formule *tous-pour-un* et fourches . . . 26
2.1.4 Loi des grands nombres . . . 32
2.2 Transitions de phase et division cellulaire . . . 34
2.2.1 Processus autoregressif bifurquant . . . 34
2.2.2 Infection dans le modèle de Kimmel . . . 35
2.2.3 Contamination et division cellulaire . . . 38
2.3 Vitesse d'invasion dans un modèle spatial . . . 38
2.3.1 Préliminaires de grandes déviations . . . 40
2.3.2 Vitesse d'invasion . . . 41

3 Chaîne de Markov branchante pour des populations structurées **46**
3.1 Définition du modèle et exemple . . . 46
3.2 Semigroupe et formule tous-pour-un . . . 49
3.3 Fonction harmonique et espace de traits fini . . . 51
3.4 Conditions de Doeblin et uniforme ergodicité . . . 53

4 Approximation en grande population **57**
4.1 Modèle et premiers résultats 57
4.1.1 Définition 57
4.1.2 Décomposition en semi-martingale 58
4.2 Convergence vers une EDP 60
4.2.1 Préliminaires 61
4.2.2 Tension de Y^K 64
4.2.3 Identification de la limite et convergence 65

Chapitre 1

Arbre de Galton Watson

L'objectif de ce chapitre est de décrire les propriétés importantes du modèle de base en processus de branchement : le modèle de Galton-Watson, ou Bienaymé Galton Watson. Pour l'instant, tous les individus ont les mêmes caractéristiques et évoluent indépendamment. Dans un premier temps, nous revenons sur les propriétés importantes de ce processus et précisons son comportement en temps long. Dans un deuxième temps nous introduirons une construction fondamentale pour la suite, la construction spinale. Elle nous renseignera sur la lignée ancestrale d'un individu typique (ou échantillon) et nous permettra d'établir le théorème de renormalisation des processus de Galton Watson sous les hypothèses optimales.

1.1 Processus de Galton-Watson

1.1.1 Introduction

Considérons une loi de reproduction p sur les entiers naturels $\mathbb{N}$ et L une v.a. (variable aléatoire) associée :

$$\mathbb{P}(L = k) = p_k, \qquad \sum_{k\geq 0} p_k = 1.$$

Considérons une famille $(L_{i,n})_{i\geq 1, n\in\mathbb{N}}$ de v.a. i.i.d. (indépendantes et identiquement distribuées) de loi p et une taille initiale $Z_0 \in \mathbb{N}$, indépendante de $(L_{i,n})_{i\geq 1, n\in\mathbb{N}}$.

Définition 1. *Le processus de Galton Watson $Z = (Z_n)_{n\geq 0}$ de loi de reproduction p est défini récursivement par $Z_0 \in \mathbb{N}$ et*

$$Z_{n+1} = \sum_{i=1}^{Z_n} L_{i,n}$$

pour tout $n \geq 0$.

Sauf mention explicite du contraire, nous supposerons par la suite que

$$Z_0 = 1 \quad \text{p.s.}$$

Nous supposerons dans tout le chapitre que le nombre moyen d'enfants est fini :

$$m = \mathbb{E}(L) = \sum_{k\geq 0} k p_k < \infty.$$

On remarque que la définition de Z implique immédiatement :

$$\mathbb{E}(Z_{n+1} \mid \mathcal{F}_n) = \mathbb{E}(Z_{n+1} \mid Z_n) = \sum_{i=1}^{Z_n} \mathbb{E}(L_{i,n}) = Z_n m,$$

en notant $\mathcal{F}_n = \sigma(Z_i : i \leq n)$ la filtration engendrée par le processus Z. On pose ainsi

$$W_n = \frac{Z_n}{m^n}$$

et le calcul précédent assure que

Lemme 2. *Le processus $(W_n)_{n\in\mathbb{N}}$ est une martingale positive. Ce processus converge presque-surement vers une limite $W \in [0, \infty)$ quand n tend vers l'infini.*

Lorsque la population s'éteint en temps fini, W_n atteint 0 et reste alors nul :

$$\{\exists n \in \mathbb{N},\, Z_n = 0\} = \{\exists n \in \mathbb{N},\, \forall p \geq 0,\, Z_{n+p} = 0\} \subset \{W = 0\}.$$

Il est naturel de s'intéresser à la réciproque, est ce que W est strictement positive sur l'événement de survie, ce qui impliquerait que Z_n croît alors bien comme sa moyenne le prédit, c'est à dire comme m^n, à une constante aléatoire près. Nous reviendrons sur cette question que nous résoudrons complètement en Section 1.4. Le cas ou la loi de reproduction a un moment d'ordre 2 fini est plus accessible, et l'objet de l'exercice suivant.

Exercice. On suppose que $m > 1$ et $\mathbb{E}(L^2) < \infty$. Montrer que la martingale $(W_n)_{n\in\mathbb{N}}$ est bornée dans L^2. En déduire que $\mathbb{P}(W > 0) > 0$ puis que $\mathbb{P}(W = 0) = \mathbb{P}(\exists n \in \mathbb{N},\, Z_n = 0)$.

Avant cela, nous allons déjà determiner quand le processus survit. Un outil pertinent pour ces questions, permettant plus généralement de caractériser la somme de v.a. i.i.d. est la fonction génératrice. Posons ainsi pour $s \in [0, 1]$ et $n \geq 0$,

$$f(s) = \mathbb{E}\left(s^L\right) = \sum_{k\geq 0} p_k s^k, \quad f_n(s) = \mathbb{E}\left(s^{Z_n}\right),$$

avec f_0 l'identité. On va montrer que f_n est la composée n-ième de f.

Proposition 3. *Pour tout $n \geq 1$,*

$$f_n = f \circ \ldots \circ f = f_{n-1} \circ f = f \circ f_{n-1}$$

Démonstration. La preuve se fait par récurrence :

$$\begin{aligned} f_{n+1}(s) &= \mathbb{E}\left(s^{Z_{n+1}}\right) = \mathbb{E}\left(\mathbb{E}\left(s^{\sum_{i=1}^{Z_n} L_{i,n}} \,\middle|\, Z_n\right)\right) \\ &= \mathbb{E}\left(\Pi_{i=1}^{Z_n} \mathbb{E}\left(s^{L_{i,n}}\right)\right) = \mathbb{E}\left(f(s)^{Z_n}\right) = f_n(f(s)). \end{aligned}$$

On obtient $f_n = f \circ \ldots \circ f$ et on peut arranger les compositions par associativité. □

1.1.2 Probabilité d'extinction

Définition 4. $p_{ext} = \mathbb{P}(\exists n \in \mathbb{N},\, Z_n = 0)$.

La probabilité d'extinction va être identifiée comme le premier point fixe de la fonction génératrice :

Lemme 5. $p_{ext} = \inf\{s \in [0,1] :\, f(s) = s\}$

Démonstration. On peut poser $u_n = \mathbb{P}(Z_n = 0) = f_n(0)$ et remarquer déjà que u_n est suite croissante qui tend vers p_{ext}. De plus $u_{n+1} = f(u_n)$ et f continue donc u_n tend vers un point fixe de f. Posant $\ell = \inf\{s \in [0,1] :\, f(s) = s\}$, on utilise que $[0,\ell]$ est stable par la fonction croissante f pour obtenir que u_n et donc sa limite p_{ext} appartiennent à $[0,\ell]$. Le fait que ℓ soit le plus petit point fixe garantit que $\ell = p_{ext}$. □

On en déduit le résultat de classification suivant

Proposition 6. *Si $m < 1$ (cas sous critique), $p_{ext} = 1$.*
Si $m > 1$ (cas surcritique), $p_{ext} > 1$.
Si $m = 1$ (cas critique) et $p_1 < 1$, alors $p_{ext} = 1$.

Le cas critique est intéressant : en dehors du cas dégénéré où la population reste constante égale à un individu pour tout temps et p.s., la population s'éteint. Or sa moyenne est préservée : $\mathbb{E}(Z_n) = 1$ pour tout $n \geq 0$. Une explication ? La preuve ci dessous passe par la convexité des fonctions génératrices. La transformée de Lamperti, donnée juste après, offrira un éclairage plus probabiliste sur ce point.

Démonstration. Le point clef est que $m = f'(1)$ et ensuite on dessine la fonction f et la première bissetrice en prêtant attention à ce qui se passe au point $s = 1$. Comme f est convexe, f est au dessus de sa tangente Δ sur $[0,1]$ et en dessous de sa corde C sur $[p_{ext}, 1]$. Le fait d'être au dessus de sa tangente assure que $p_{ext} = 1$ pour $m < 1$. Le fait d'avoir une derivée plus grande que 1 en $s = 1$ pour $m > 1$ assure que $p_{ext} < 1$ par le théorème des valeurs intermédiaires. Si $m = 1$ et $p_{ext} < 1$, Δ et C coincident on $[p_{ext}, 1]$, et doivent donc être égales à f. Par analycité de f, elles coincident sur $[0,1]$ car $f - id$ vaut 0 sur un ensemble qui admet un point d'accumulation. Ce qui signifie $p_1 = 1$. On peut éviter l'analyse complexe et remarquer sinon que f n'est alors pas strictement convexe au voisinage de 1, donc $f''(s) = \sum_{k\geq 2} p_k k(k-1)s^k$ s'annule au voisinage de 1 et p_k est nulle pour $k \geq 2$. □

1.1.3 Transformation de Lamperti

Définissons pour $n \geq 0$,

$$A_n = \sum_{k=0}^{n-1} Z_k,$$

avec la convention $A_0 = 0$, l'effectif cumulé de la population sur n générations. Cette quantité est croissante et on pose

$$A_\infty = \lim_{n\to\infty} A_n \in \mathbb{N} \cup \{+\infty\}$$

la population cumulée. On peut déjà noter que l' extinction en temps fini de la population coincide avec l'événement A_∞ est fini.

Définition 7. *On définit le processus $S = (S_n)_{n\in\mathbb{N},\, n\leq A_\infty}$ à valeurs entières de la façon suivante : $S_0 = 1$ et pour tout $n \geq 0$ et tout $k \in [\![A_n, A_{n+1} - 1]\!]$:*

$$S_{k+1} = S_k + L_{k-A_n+1,n} - 1.$$

La construction itérative de S ci dessus définit bien S de façon univoque car les intervalles successifs $[\![A_n, A_{n+1} - 1]\!]$ forment un recouvrement de $\{n \in \mathbb{N}, n < A_\infty\}$ par des ensembles disjoints.

Théorème 8. *Pour tout $n \geq 0$,*

$$S_{A_n} = Z_n.$$

De plus, le processus S est une marche aléatoire sur $\mathbb{Z}$, de pas $L - 1$, arrêtée à son premier passage en 0 : $A_\infty = \inf\{n \geq 0 : S_n = 0\}$.

Démonstration. Naturellement, $S_0 = 1 = Z_0$. D'autre part, par récurrence :

$$S_{A_{n+1}} = S_{A_n} + \sum_{k=A_n}^{A_{n+1}-1} S_{k+1} - S_k = S_{A_n} + \sum_{k=1}^{Z_n}(L_{k,n} - 1) = S_{A_n} + (Z_{n+1} - Z_n) = Z_{n+1}\,.$$

Le premier point est ainsi vérifié.

On peut vérifier que S se comporte bien comme une marche aléatoire de pas $L - 1$ jusqu'au rang A_∞. L'indépendance et l'identique distribution des pas $(S_{k+1} - S_k)$ pour $0 \leq k < A_\infty$ est une conséquence directe de celles de la famille $(L_{i,n})_{i\geq 1, n\in\mathbb{N}}$.
Il reste à justifier que A_∞ est le premier passage en 0 de S. Sur l'événement $\{A_\infty < +\infty\}$, on note que $N = \inf\{n \in \mathbb{N},\, A_{n+1} = A_n\} = \inf\{n \in \mathbb{N},\, Z_n = 0\}$ et $A_N = A_{N+1} = \ldots = A_\infty$. Donc $S_{A_\infty} = S_{A_N} = Z_N = 0$. Ceci implique que S est en 0 au temps A_∞.
Pour conclure, il faut montrer que S ne passe par 0 avant le temps A_∞. Pour $n < N$, $S_{A_n} = Z_n > 0$ par définition de N. En outre, entre les temps A_n et A_{n+1}, S part de la valeur Z_n et effectue $A_{n+1} - A_n = Z_n$ étapes. Or S diminue d'au plus un cran à chaque étape, donc S ne peut pas atteindre 0 avant l'instant A_{n+1}. Donc S atteint 0 en un temps qui appartient à $\{A_n : n \geq 0\}$. Alors Z s'annule et A est constant égal à A_∞. □

On obtient directement

Corollaire 9. *Les événements suivant coincident p.s.*

$$\{\exists n \in \mathbb{N} :\, Z_n = 0\} = \{A_\infty < \infty\} = \{\exists n \leq A_\infty : S_n = 0\}.$$

De plus la taille totale de la population $\sum_{n\geq 0} Z_n$ est distribuée comme le temps d'atteinte de 0 par une marche aléatoire de pas $L - 1$, issue de 1.

Dans le cas $m = \mathbb{E}(L) = 1$, le processus S est une marche aléatoire centrée arrêtée quand elle touche 0. De plus elle est non constante dès que $\mathbb{P}(L = 1) \neq 1$. Or une marche aléatoire centrée non constante oscille entre $+\infty$ et $-\infty$. Comme les pas négatifs de la marche ne peuvent valoir que -1, elle touche forcément en temps fini. On retrouve ici que l'extinction de la population est p.s. dans le cas critique. De plus, le deuxième point du corollaire nous donne une façon de contrôler la distribution de la taille totale de la population. Ainsi,

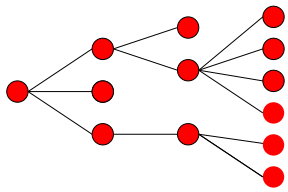


FIGURE 1.1 – Arbre $t = \{\varnothing, 1, 2, 3, 11, 12, 31, 121, 122, 123, 124, 311, 312\}$. Il est associé à la réalisation $L(\varnothing) = 3, L(1) = 2, L(2) = 0, L(3) = 1, L(11) = 0, L(12) = 4, L(31) = 2$ de la famille $(L(u))_{u\in T}$.

$\mathbb{P}(\sum_{n\geq 0} Z_n = k) = \mathbb{P}(A_\infty = k) = \mathbb{P}(\inf\{n \geq 0 : S_n = 0\} = k)$. On peut alors utiliser la formule de Kemperman : $\mathbb{P}(\inf\{n \geq 0 : S_n = 0\} = k) = \mathbb{P}(Y_k = -1)/k$, où Y est une marche aléatoire de pas $L - 1$ issue de 0. Et on obtient que

$$\mathbb{P}(\sum_{n\geq 0} Z_n = k) = \mathbb{P}(Y_k = -1)/k$$

Notons que ce terme peut devenir explicite pour certaines loi de reproduction.

La transformation de Lamperti peut etre opérée en temps et même en espace continu pour les processus de branchement [21]. C'est un outil puissant permettant d'étudier les approximations en grande population des processus de branchement, en les voyant à travers la convergence des marches aléatoires vers les processus de Lévy [20, 9]. Elle permet aussi d'analyser le comportement asymptotique des processus de branchement en espace continu, notamment de savoir si le processus tends vers 0 sans le toucher, ou est absorbé en 0 [7].

1.2 Arbre aléatoire et construction spinale

1.2.1 Définitions

On va utiliser les notations de Ulam Harris Neveu :

$$\mathcal{U} = \bigcup_{n\in\mathbb{N}} (\mathbb{N}^*)^n = \{\varnothing\} \cup \mathbb{N}^* \cup (\mathbb{N}^*)^2 \cup \dots$$

est l'ensemble de mots dont chaque lettre est un entier. Cet ensemble permet de donner un label à chaque individu d'un arbre : ui sera le $i$ème enfant de l'individu u. Un individu u de $\mathcal{U}$ est ancêtre d'un élément v de $\mathcal{U}$ si c'est un préfixe, i.e. si il existe $u_i \in \mathbb{N}^*$ tel que $v = uu_1 \dots u_n$, On note alors $u \preccurlyeq v$.

Définition 10 (Arbre)**.** *Un arbre t est une partie de $\mathcal{U}$ telle que*
- *si $v \in t$ et $u \preccurlyeq v$, alors $u \in t$ [stabilité par ancêtre],*
- *si $uk \in t$ pour un entier $k \in \mathbb{N}^*$, alors pour tout $1 \leq k' \leq k$, $uk' \in t$ [ordonnancement des frères et soeurs].*

On note $\mathbb{T}$ l'ensemble des arbres, $\mathbb{T}_n$ l'ensemble des arbres de hauteur n (exactement) et $\mathbb{T}_f = \cup_{n\in\mathbb{N}}\mathbb{T}_n$ l'ensemble des arbres finis. Ces ensembles sont dénombrables. La longueur du mot est notée $|u|$: $|u| = k$ ssi il exite $u_i \in \mathbb{N}^*$ tel que $u = u_1 \dots u_k$. On

dit que $|u|$ est la génération de l'individu u. On appelle hauteur de l'arbre la longueur maximale des mots de l'arbre.

Considérons $(L(u))_{u\in T}$ une famille de variable alétaoire i.i.d. de loi p.

Définition 11. *On appelle arbre de Galton Watson associé à la famille $(L(u))_{u\in\mathcal{U}}$ la variable aléatoire*

$$\mathcal{T} = \cup_{n\in\mathbb{N}}\mathbb{G}_n$$

à valeurs dans $\mathbb{T}$ définie récursivement (génération par génération) de la façon suivante

— $\mathbb{G}_0 = \{\varnothing\}$

— $\mathbb{G}_{n+1} = \{uk : u \in \mathbb{G}_n \text{ et } k \in [\![1, L(u)]\!]\}$.

Par construction, $Z_n = \#\mathbb{G}_n$ est un processus de Galton-Watson, tel qu'on l'a défini dans la section précédente.

Question : quand j'échantillonne un individu uniformément au hasard au sein de la génération n, combien a-t-il de frères et soeurs en moyenne ?

1.2.2 Construction spinale

L'objectif de cette construction est de faire émerger un individu qui va jouer le rôle d'individu typique ou échantillon (uniforme), en réalisant la construction conjointe de l'arbre de Galton Watson et de cet individu de façon markovienne, de sorte à pouvoir utiliser les notions, méthodes et résultats valables pour les processus de Markov.

L'idée de cette construction est d'isoler, à chaque étape du processus, un individu qui se reproduira selon une loi $\widehat{L}$ biaisée par le nombre de descendants.

Définition 12. *Soit p une probabilité sur $\mathbb{N}$. La loi p biasée par la taille est la loi de probabilité $\widehat{p}$ sur $\mathbb{N}$ définie par*

$$\widehat{p}_k = \frac{kp_k}{m} \qquad (k \in \mathbb{N}).$$

Soit $(\widehat{L}_n)_{n\in\mathbb{N}}$ une suite de v.a. i.i.d. de loi $\widehat{p}$, indépendante de $(L(u))_{u\in\mathcal{U}}$. Nous allons distinguer un individu E_n parmi les individus $\mathbb{V}_n$ de la génération n au sein d'un arbre aléatoire $\mathcal{A} = \cup_{n\in\mathbb{N}}\mathbb{V}_n$. Pour cela nous introduisons une suite de v.a. indépendantes, indépendante de $(L(u))_{u\in\mathcal{U}}$, avec K_n uniformément distribuée sur $[\![1, \widehat{L}_n]\!]$ (et indépendante du reste). La construction se fait à nouveau par récurrence :

– $\mathbb{V}_0 = \{\varnothing\}, E_0 = \varnothing$.

– $\mathbb{V}_{n+1} = \displaystyle\bigcup_{v\in\mathbb{V}_n-\{E_n\}} \{vk : 1 \le k \le L(v)\} \bigcup \{E_n k : 1 \le k \le \widehat{L}_n\}, \quad E_{n+1} = E_n K_n.$

Ainsi, la construction de l'arbre $\mathcal{A}$ est similaire à celle de l'arbre de Galton Watson $\mathcal{T}$, à ceci près que l'individu distingué (épine) E se reproduit suivant la loi biasée par la taille.

1.2.3 Résultats

On introduit maintenant les arbres tronqués à la génération n

$$\mathcal{T}_n = \bigcup_{k=0}^{n} \mathbb{G}_k, \qquad \mathcal{A}_n = \bigcup_{k=0}^{n} \mathbb{V}_k.$$

Notons $g_n(t) = \{u \in t : |u| = n\}$ la génération n de l'arbre t.

Lemme 13. *Pour tous $n \in \mathbb{N}$, $t \in \mathbb{T}_n$ et $e \in g_n(t)$,*

$$\mathbb{P}(\mathcal{A}_n = t, E_n = e) = \frac{1}{m^n}\mathbb{P}(\mathcal{T}_n = t).$$

Démonstration. Montrons le résultat par récurrence sur la hauteur de l'arbre. Naturellement, au rang $n = 0$, les deux termes sont égaux à 1 si $t = \{\varnothing\}$ (et 0 sinon). Supposons le résultat vrai au rang n, et fixons un arbre t' de hauteur $n+1$, et $e' \in g_{n+1}(t')$. Soient $t = t' \cap \big(\cup_{k=0}^{n} \mathbb{N}^k\big)$ la restriction de t' aux n premières générations, et $e \in g_n(t) = g_n(t')$ la maman de e'. Par construction de $(\mathcal{A}, E)$, on a alors :

$$\begin{aligned}\mathbb{P}\big(\mathcal{A}_{n+1} = t', E_{n+1} = e'\big) &= \mathbb{P}(\mathcal{A}_n = t, E_n = e) \times \mathbb{P}\big(\mathcal{A}_{n+1} = t', E_{n+1} = e' \big| \mathcal{A}_n = t, E_n = e\big) \\ &= \frac{\mathbb{P}(\mathcal{T}_n = t)}{m^n}\mathbb{P}\big(\mathcal{A}_{n+1} = t', E_{n+1} = e' \big| \mathcal{A}_n = t, E_n = e\big).\end{aligned}$$

Reste à étudier le terme de droite, ce qui revient à calculer la probabilité que chaque élément $u \in g_n(t')$ ait le nombre $k(u)$ d'enfants souhaités :

$$\begin{aligned}\mathbb{P}\big(\mathcal{A}_{n+1} = t', E_{n+1} = e' \big| A_n = t, E_n = e\big) &= \mathbb{P}\big(\mathcal{A}_{n+1} = t' \big| \mathcal{A}_n = t, E_n = e\big) \\ &\qquad \times \mathbb{P}\big(E_{n+1} = e' \big| \mathcal{A}_{n+1} = t', E_n = e\big) \\ &= \prod_{u \in g_n(t),\, u \neq e} p_{k(u)} \times \widehat{p}_{k(e)} \times \frac{1}{k(e)} \\ &= \frac{1}{m} \prod_{u \in g_n(t)} p_{k(u)} \\ &= \frac{1}{m}\mathbb{P}\big(\mathcal{T}_{n+1} = t' \big| \tau = t\big)\,,\end{aligned}$$

ce qui donne le résultat voulu au rang $n+1$. □

Soit $n \in \mathbb{N}$ et U_n une variable aléatoire qui, conditionnellement à $\mathcal{A}_n$ et $\mathbb{G}_n \neq \varnothing$, est distribuée uniformément sur $\mathbb{G}_n$:

$$\forall t \in \mathbb{T}_n,\, \forall\, e \in g_n(t), \quad \mathbb{P}(U_n = e | \mathcal{T}_n = t) = \frac{1}{\#g_n(t)}.$$

Théorème 14. *Pour tous $n \in \mathbb{N}$ et $F : \mathbb{T}_n \times \mathbb{N}^n \to \mathbb{R}$ fonction bornée :*

$$\mathbb{E}[F(\mathcal{T}_n, U_n)W_n] = \mathbb{E}[F(\mathcal{A}_n, E_n)]\,,$$

où on rappelle que $W_n = Z_n/m^n$ est la martingale introduite précédemment.

Démonstration. La formule de transfert et le lemme précédent assurent que

$$\begin{aligned}
\mathbb{E}[F(\mathcal{T}_n, U_n)W_n] &= \sum_{t\in\mathbb{T}_n,\, e\in g_n(t)} \mathbb{P}(\mathcal{T}_n = t,\, U_n = e) \times F(t,e)\frac{\#g_n(t)}{m^n} \\
&= \sum_{t\in\mathbb{T}_n,\, e\in g_n(t)} \mathbb{P}(\mathcal{T}_n = t) \times \frac{1}{\#g_n(t)} \times F(t,e)\frac{\#g_n(t)}{m^n} \\
&= \sum_{t\in\mathbb{T}_n,\, e\in g_n(t)} \frac{1}{m^n}\mathbb{P}(\mathcal{T}_n = t) \times F(t,e) \\
&= \sum_{t\in\mathbb{T}_n,\, e\in g_n(t)} \mathbb{P}(\mathcal{A}_n = t,\, E_n = e)F(t,e) \\
&= \mathbb{E}[F(\mathcal{A}_n, E_n)]\,.
\end{aligned}$$

□

Donnons une première application de ce résultat et de la construction spinale sur la lignée ancestrale d'un échantillon uniforme. Plaçons nous par souci de simplicité dans le cas où

$$\mathbb{P}(W > 0) = 1$$

ce qui, comme on le verra dans la section suivante, revient à supposer que $p_0 = 0$ (pas de mort) et $\mathbb{E}(L\log L) < \infty$. Considérons

$$F_k^n = \frac{1}{n}\#\{u \prec U_n,\, L(u) = k\}.$$

la proportion d'ancêtres de U_n qui ont eu k enfants.

Corollaire 15. *Pour tout $k \in \mathbb{N}$, $F_k^n \underset{n\to\infty}{\longrightarrow} \widehat{p}_k$ en probabilité.*

Démonstration. Soient $\varepsilon > 0$, $k \in \mathbb{N}$ et F la fonction définie par

$$F_{k,\varepsilon}(t,e) = \mathbb{1}_{\left|\frac{1}{n}\#\{u\prec e:\, k_t(u)=k\} - \widehat{p}_k\right| \geq \varepsilon},$$

avec t un arbre et $k_t(u)$ le nombre d'enfants de u dans t. On remarque que

$$\frac{1}{n}\#\{u \prec E_n :\, k_{\mathcal{A}_n}(u) = k\} = \frac{1}{n}\#\Big\{i < n,\, \widehat{L}_i = k\Big\}.$$

Comme $(\widehat{L}_i)_{i\in\mathbb{N}}$ est une suite iid de variables, la loi des grands nombres appliquée aux variables i.i.d. intégrables $(1_{\widehat{L}_i=k})_{i\in\mathbb{N}}$ assure que cette fréquence tend p.s. (et en proba) vers $\mathbb{E}(1_{\widehat{L}=k}) = \widehat{p}_k$. Donc $F_{k,\varepsilon}(\mathcal{A}_n, E_n)$ tend vers 0 dans L^1.
Le théorème précédent implique alors que $F_{k,\varepsilon}(\mathcal{T}_n, U_n)W_n$ tend vers 0 dans L^1, et donc en probabilité. Or par hypothèse $W_n \to W$ en probabilité quand n tend vers l'infini et W est non nulle. On obtient que $F_{k,\varepsilon}(\mathcal{T}_n, U_n)$ tend vers 0 en probabilité. En particulier,

$$\mathbb{P}(F_{k,\varepsilon}(\mathcal{T}_n, U_n) = 1) = \mathbb{P}\left(\left|\frac{1}{n}\#\{u \in \mathcal{T}_n,\, u < U_n,\, L(u) = k\} - \widehat{p}_k\right| \geq \varepsilon\right) \longrightarrow 0\,,$$

d'où la convergence annoncée. □

1.3 Processus de Galton-Watson avec immigration

Considérons maintenant des v.a. i.i.d $(I_n)_{n\geq 1}$, à valeurs entières, indépendantes des $(L_{i,n})_{i\geq 1, n\in\mathbb{N}}$. On note q la loi commune des I_n et I une v.a. associée.

Définition 16. *Le processus de Galton-Watson avec immigration Y, de loi de reproduction p et loi d'immigration q est défini itérativement par $Y_0 \in \mathbb{N}$ et pour $n \geq 0$*

$$Y_{n+1} = \sum_{i=1}^{Y_n} L_{i,n} + I_{n+1}.$$

Sauf mention explicite du contraire, nous partirons de $Y_0 = 0$ individus p.s.

Exercice. On note $\mathcal{F}_n = \sigma(Y_0, \ldots, Y_n)$.
1) Calculer $\mathbb{E}(Y_{n+1} \mid \mathcal{F}_n)$.
Montrer alors que $u_n = \mathbb{E}(Y_n)$ est une suite récurrente arithmético géométrique (et donc que l'on peut la calculer !).
2) On introduit la fonction génératrice de Y_n :

$$f_n(s) = \mathbb{E}(s^{Y_n}) \qquad (s \in [0,1]).$$

Calculer $\mathbb{E}(s^{Y_{n+1}} \mid \mathcal{F}_n)$ et en déduire une relation entre f_{n+1} et f_n.
3) Dans le cas où Y_n converge en loi vers une variable Y_∞, caractériser la loi de Y_∞ via une équation fonctionnelle satisfaite par son fonction génératrice.

Représentons le processus de Galton Watson avec immigration en utilisant les différentes vagues migratoires. Considérons pour cela une famille $\big((Z_{j,n}^{(x)})_{j\leq n}, x \in \mathbb{N}, j \geq 1\big)$ de processus de Galton-Watson indépendants et indépendants de $(I_n)_{n\geq 1}$, avec $Z_{j,n}^{(x)}$ la taille au temps d' un processus de Galton Watson issu de x individus au temps j.

Proposition 17. *Le processus de Galton Watson avec immigration Y est égal en loi au processus*

$$\left(\sum_{j=1}^{n} Z_{j,n}^{(I_j)}\right)_{n\in\mathbb{N}}.$$

Notons que par couplage des v.a. reproductions, on pourrait construire les processus $\big((Z_{j,n}^{(x)})_{j\leq n}, x \in \mathbb{N}, j \geq 1\big)$ de telle sorte que l'égale précédente soit presque sure, et pas seulement en loi.

Démonstration. Il n'y personne dans la population Y à la génération 0 et chaque individu de la génération n sera le descendant d'un migrant arrivé à une génération $i \leq n$. Le point clef est la propriété de branchement : l'indépendance et l'identique distribution des reproductions dans le processus Y permet de décrire la population engendrée par les I_j migrants arrivés à la génération j par les processus $(Z_{j,n}^{(x)})_{j\leq n}$, indépendants les uns des autres. La preuve consiste alors simplement à écrire à chaque temps n la population Y_n comme le cumul de ces différentes contributions, i.e. sommer $Z_{j,n}^{(I_j)}$ pour $j \leq n$. □

Dans le cas surcritique $m > 1$, la probabilité de survie d'une vague migratoire est positive, et quand $n \to \infty$, le processus tend vers p.s. vers l'infini.

Par ailleurs, en renversant le temps, et on notant $(\mathcal{Z}_j)_{j\geq 0}$ une suite de v.a. indépendante avec $\mathcal{Z}_j$ distribuée comme $Z_{0,j}^{(I)}$, on déduit que pour tout $n \geq 0$,

$$Y_n \overset{loi}{=} \sum_{j=0}^{n-1} \mathcal{Z}_j$$

Or la suite $\sum_{j=1}^{n} \mathcal{Z}_j$ est p.s. croissante quand n augmente et donc convergeant vers

$$Y_\infty := \sum_{j\geq 0} \mathcal{Z}_j.$$

De plus $\mathbb{E}(\mathcal{Z}_j) = \mathbb{E}(I)m^j$ et on obtient que si $\mathbb{E}(I)$ est fini et $m < 1$, $(Y_n)_n$ converge en loi vers une variable aléatoire intégrable. Pas de convergence p.s., la population peut s'éteindre mais redémarrer avec une immigration, et elle fluctue. Nous allons considérer maintenant les cas $m > 1$ et $m = 1$ maintenant. Nous reviendrons enfin rapidement sur le cas $m < 1$ en montrant que la convergence en loi de $(Y_n)_n$ est assure dès que $\mathbb{E}(\log(I)) < \infty$, la variable limite n'étant plus forcément intégrable.

1.3.1 Croissance dans le cas surcritique

On se place dans le cas où $m > 1$ et on a vu qu'alors Y_n tends p.s. vers l'infini. Nous cherchons maintenant à savoir à quelle vitesse.

Proposition 18. *Conditionnellement à $(I_n)_{n\geq 1}$, le processus $\left(Y_n/m^n\right)_{n\in\mathbb{N}}$ est une sous-martingale. De plus*

$$\mathbb{E}\left[\frac{Y_n}{m^n} \,\middle|\, I\right] = \sum_{k=1}^{n} \frac{I_k}{m^k}.$$

Remarquons que ce résultat reste valide si si les termes de I (et donc de Y) ne sont pas intégrables.

Démonstration. Remarquons déjà que

$$\mathbb{E}\left[Z_{j,n+1}^{(I_j)} \middle| \mathcal{F}_n, I\right] = \mathbb{E}\left[Z_{j,n+1}^{(I_j)} \middle| Z_{j,n}^{(I_j)}, I_j\right] = Z_{j,n}^{(I_j)}.m$$

car $(Z_{j,n}^{(x)})_{n\geq j}$ est un processus de Galton Watson et nous avons vu dans la première section que $\mathbb{E}(Z_{j,n+1}^{(x)}|(Z_{j,k}^{(x)} : k \leq n)) = Z_{j,n+1}^{(x)}m$. Les deux identités précédentes assurent

$$\mathbb{E}[Y_{n+1}|\mathcal{F}_n, I] = I_{n+1} + \mathbb{E}\left[\sum_{j=1}^{n} Z_{j,n+1}^{(I_j)} \middle| \mathcal{F}_n, I\right] = I_{n+1} + m\sum_{j=1}^{n} Z_{n,j}^{(I_j)} = I_{n+1} + mY_n\,,$$

ce qui donne

$$\mathbb{E}\left[\frac{Y_{n+1}}{m^{n+1}} \middle| \mathcal{F}_n, I\right] = \frac{I_{n+1}}{m^{n+1}} + \frac{Y_n}{m^n} \geq \frac{Y_n}{m^n},$$

et conclut la preuve. □

Lemme 19. *Soient $(A_n)_{n\geq 0}$ des variables positives i.i.d.*

i) Si $\mathbb{E}(A) < \infty$, alors $\overline{\lim}_{n\to\infty} \frac{A_n}{n} = 0$ p.s. et $\forall c \in (0,1)$, $\sum_{n\in\mathbb{N}} c^n e^{A_n} < \infty$ p.s.

ii) Si $\mathbb{E}[A] = \infty$, alors $\overline{\lim}_{n\to\infty} \frac{A_n}{n} = \infty$ p.s., et $\forall c > 0$, $\sum_{n\in\mathbb{N}} c^n e^{A_n} = \infty$ p.s.

Démonstration. Pour le premier point, remarquons que $\mathbb{E}[A/\varepsilon] < \infty$, donc $\sum_{n\in\mathbb{N}} \mathbb{P}\left(\frac{A}{\varepsilon} \geq n\right) = \sum_{n\in\mathbb{N}} \mathbb{P}\left(\frac{A_n}{\varepsilon} \geq n\right) < \infty$. Par Borel-Cantelli, on en déduit que $\overline{\lim} \frac{A_n}{n} \leq \varepsilon$. En particulier, avec ε assez faible, $ce^{\varepsilon} < 1$ et $e^{A_n} \leq (e^{\varepsilon})^n$ partir d'un rang. On montre le deuxième point de façon analogue. Cette fois, $\sum_{n\geq 1} \mathbb{P}(A_n/K \geq n) = \infty$ et Borel Cantelli implique $\mathbb{P}(\overline{\lim}_{n\to\infty} A_n/n \geq K) = 1$ et $c^n e^{A_n}$ est p.s. non bornée. □

On rappelle qu'on travaille ici dans le cas $m > 1$, et on énonce le résultat de renormalisation du à Seneta (1970).

Proposition 20. *Si $\mathbb{E}[\log I] < \infty$, alors Y_n/m^n converge p.s. vers une variable finie quand n tend vers l'infini.*
Sinon, $\overline{\lim}_{n\to\infty} Y_n/m^n = \infty$ p.s.

On verra dans la partie suivante sous quelle condition la limite de Y_n/m^n est bien strictement positive, en utilisant justement le fait qu'elle est finie.

Démonstration. Dans le premier cas, on utilise le lemme précédent avec $A_i = \log_+(I_i)$ et $c = 1/m \in (0,1)$. Comme $\mathbb{E}(A_i) < \infty$, $i)$ assure que $\sum_{k=1}^{\infty} I_k/m^k < \infty$ p.s. La proposition 18 entraine alors que conditionnellement à I, le processus $(Y_n/m^n)_n$ est une sous-martingale bornée dans L^1, d'où la convergence p.s. vers une variable finie.

On peut montrer l'autre cas de figure en utilisant que $Y_n \geq I_n$ et $\overline{\lim}_{n\to\infty} I_n/m^n = \infty$ puisque le lemme 19 ii) guarantit $\overline{\lim}_{n\to\infty} \log(I_n)/n = \infty$. □

1.3.2 Les régimes critiques et sous critiques : $m \leq 1$.

Le cas critique $m = 1$ est intéressant car plusieurs comportements sont possibles et la variance de la loi de reproduction joue un rôle dans le comportement asymptotique. On suppose ainsi que

$$\sigma^2 = \mathbb{E}((L-1)^2) < \infty$$

On peut d'abord remarquer que le processus croit en moyenne vers l'infini.

$$\mathbb{E}(Y_n) = n\mathbb{E}(I).$$

Si les fluctuations sont suffisamment importantes, le processus sera récurrent (nul), sinon il sera transient et tend p.s. vers l'infini. Dans le premier cas, le processus tend en proba vers l'infini, mais pas p.s. : la population revient dans les compact une infinité de fois. En effet, en posant

$$f(x) = \mathbb{E}(\log(Y_{n+1})|Y_n = x) - \log(x).$$

on peut vérifier que $\lim_{x\to\infty} xf(x) = \mathbb{E}(I) - \sigma^2/2$.
Ainsi, quand $\sigma^2/2 < \mathbb{E}(I)$, $(\mathbb{E}(\log(Y_n)))_n$ est bornée et par le lemme de Fatou $\mathbb{E}(\underline{\lim}_{n\to\infty} Y_n) < \infty$, ce qui implique

$$\underline{\lim}_{n\to\infty} Y_n < \infty \quad \text{p.s.}$$

On renvoie au travail de Kesting [16, 15] sur les modèles de croissance pour une preuve pour les deux régimes, qui étend par ailleurs le critère à des modèles plus généraux que les processus de branchement avec immigration. Pour des référencs bibliographiques et des extensions au cadre multitype, on pourra consulter la thèse d'Etienne d'Adam. Notons aussi que Y_n/n converge en loi quand n tend vers l'infini vers une loi Gamma, ce qui peut être prouvé en utilisant les fonctions génératrices par exemples.

Enfin, revenons sur le cas sous critique $m < 1$. On rappelle que $(Y_n)_n$ connverge en loi vers $Y_\infty := \sum_{j\geq 0} \mathcal{Z}_j \in [0, \infty]$. Si on regarde l'espérence conditionnelle

$$\mathbb{E}\left(Y_\infty \,|\, I\right) = \sum_{n\geq 1} I_n m^n$$

et le lemme 19 assure que $\mathbb{E}(Y_\infty \,|\, I) < \infty$ p.s. si et seulement si $\mathbb{E}(\log(I)) < \infty$. Ceci entraine que $Y_\infty < \infty$ sous la condition $\mathbb{E}(\log(I)) < \infty$. Sous la condition $\mathbb{E}(\log(I)) = \infty$, le lemme 19 assure que

$$\overline{\lim_{n\to\infty}} \frac{I_n}{(1-p_0)^n} = \infty \quad \text{p.s.}$$

Or $\mathbb{P}_i(Z_n = 0) = \mathbb{P}_1(Z_n = 0)^i$ et la probabilité de survie au temps n est au moins la probabilité qu'à chaque génération, un individu donné survive $\mathbb{P}_1(Z_n > 0) \geq (1-p_0)^n$. Du coup

$$\mathbb{P}(\mathcal{Z}_n \geq 1 | I) \geq 1 - (1 - (1-p_0)^n)^{I_n}$$

a une limsup égale à 1 p.s., ce qui implique Y_∞ p.s. finie par Borel Cantelli. On obtient le résultat suivant du à Heathcote (1966) :

Proposition 21. *Si $\mathbb{E}(\log(I)) < \infty$, alors $(Y_n)_n$ tend en loi vers une variable aléatoire finie p.s. Sinon, $(Y_n)_n$ tend en loi vers l'infini.*

1.4 Croissance malthusienne et vitesse d'exinction

1.4.1 Croissance dans le cas surcritique

On suppose ici $m > 1$ et la population survit avec probabilité positive. Revenons sur une question évoquée dans la première section, concernant la martingale limite $W = \lim_{n\to\infty} Z_n/m^n \in [0, \infty)$. On avait remarqué que l'inclusion des événements $\{\exists n \in \mathbb{N}, Z_n = 0\} \subset \{W = 0\}$ et laissée ouverte la réciproque. A-t-on toujours $\{W = 0\} \subset \{\exists n \in \mathbb{N}, Z_n = 0\}$, c'est à dire $\{\forall n \in \mathbb{N}, Z_n > 0\} \subset \{W > 0\}$? Dit autrement, lorsque la population survit, sa taille croît-elle bien comme m^n (croissance Malthusienne de la moyenne) ou de manière moindre ? Nous allons ici répondre par l'affirmative sous une condition de moment finie $\mathbb{E}(L\log(L)) < \infty$, qui va s'avérer être une condition nécessaire et suffisante. La preuve peut se faire via les fonctions génératrices introduites en première section et le lecteur pourra consulter [1]. Nous allons ici utiliser la construction spinale, qui va permettre de comprendre cette condition de moment et d'établir une connection intéressante avec les processus de Galton Watson avec immigration, due à [22]. Cette approche permet d'obtenir aussi la vitesse d'extinction du processus et se généralise à des processus de branchement plus complexe, notamment en environnement aléatoire ou multitype, comme nous auront l'occasion de le voir.

Nous revenons maintenant à la construction spinale de la section précédente. On rappelle que la génération n des individus dans cette construction est noté $\mathbb{V}_n$ et on note

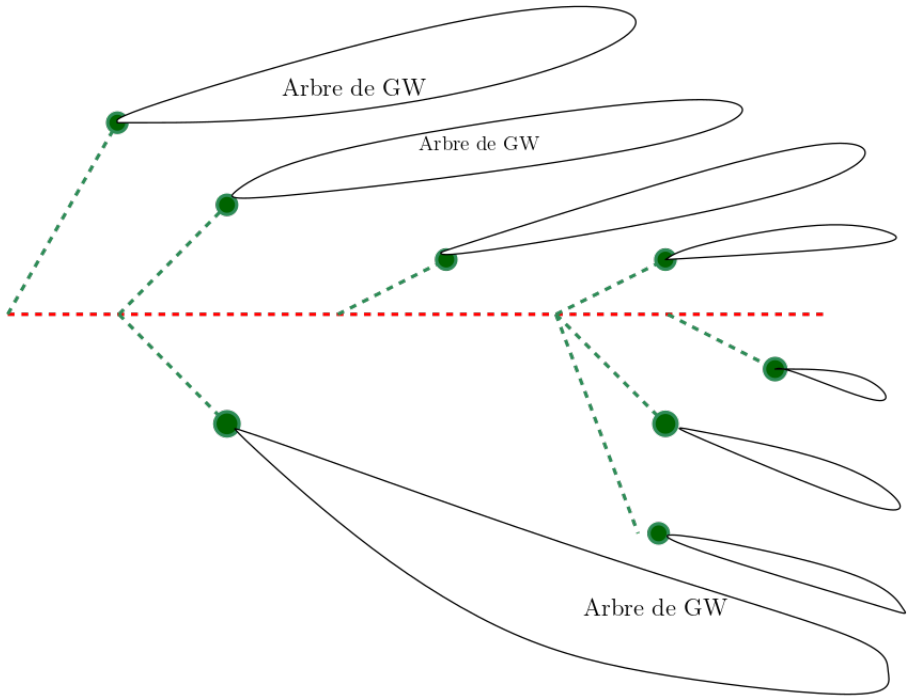
$$X_n = \#\mathbb{V}_n, \qquad Y_n = \#(\mathbb{V}_n - \{E_n\}) = X_n - 1$$

respectivement les tailles de la population à l'instant n avec ou sans l'épine. Le résultat suivant est la clef de cette partie : il permet de décrire la taille de la population dans la construction spinale comme un processus avec immigration et de faire le lien avec la martingale limite W du processus de Galton Watson original.

Proposition 22. *Le processus Y est un processus de Galton Watson avec immigration, de loi de reproduction p et d'immigration distribuée comme $\widehat{L} - 1$. De plus*

$$\mathbb{E}[W] = \mathbb{P}\left(\sup_{n\in\mathbb{N}} \frac{Y_n}{m^n} < \infty\right).$$

Démonstration. Pour la première partie, on peut s'appuyer sur une des deux définitions équivalentes des processus de Galton Watson avec immigration. Le plus simple au vu de la construction spinale est surement d'utiliser les vagues migratoires (proposition précédente) et de faire un dessin :



L'épine se reproduit suivant la loi biaisée $\widehat{L}$. Si on met à part l'individu spinal E (en rouge) mais pas ses descendants (en vert), la contribution que l'épine laisse à chaque génération n correspond à $I_n = \widehat{L}_n - 1$ individus (les noeuds verts), qui sont des v.a. indépendantes et identiquement distribuées. Ces individus constituent nos migrants et les processus qu'ils engendrent les différentes vagues migratoires, que l'on somme à l'instant n. On somme alors des processus de Galton Watson indépendants de loi p car dans la construction spinale, tous les individus autre que l'épine se reproduisent indépendamment

suivant la loi p. Nous avons alors obtenu la taille de la population dans la construction spinale, sans l'épine, soit $X - 1 = Y$.

Pour l'identité qui fait le lien entre $\mathbb{E}(W)$ et la construction spinale Y, on va exploiter le théorème 14 et passer à la limite pour faire apparaître W dans le terme de gauche, tandis que le terme de droite concerne la construction spinale et va faire sortir Y. Pour $n \geq 0$, on introduit la fonction $F_n : \mathbb{T}_n \to [0,1]$ définie par

$$F_n(t) = \mathbb{1}_{\{\sup_{k\leq n} \#g_k(t)/m^k \leq K\}}.$$

qui indique si la population renormalisée dépasse K avant la génération n. Le théorème 14 assure alors que $\mathbb{E}[F(\mathcal{T}_n)W_n] = \mathbb{E}[F(\mathcal{A}_n)]$. On peut le réécrire :

$$\mathbb{E}\Big[\mathbb{1}_{\{\sup_{k\leq n} W_k \leq K\}} W_n\Big] = \mathbb{P}\left(\sup_{k\leq n} \frac{X_k}{m^k} \leq K\right).$$

Quand n augmente, le terme dans l'espérance de gauche est dominé par K par construction, tandis que les événements à droite décroissent. En passant à la limite, on obtient :

$$\mathbb{E}\Big[\mathbb{1}_{\{\sup_{k\geq 0} W_k \leq K\}} W\Big] = \mathbb{P}\left(\sup_{k\geq 0} \frac{X_k}{m^k} \leq K\right).$$

On fait maintenant tendre K vers l'infini et par limite croissante on obtient, comme W est p.s. fini :

$$\mathbb{E}[W] = \mathbb{P}\left(\sup_{k\geq 0} \frac{X_k}{m^k} < \infty\right).$$

ce qui conclut la preuve car X et Y diffèrent de 1 et $m > 1$. □

En exploitant maintenant le lien entre Y et W obtenu dans la proposition précédente et le comportement asymptotique de Y_n/m^n determiné dans la Proposition 20, on obtient le résultat suivant. Il permet de caractériser quand la martingale limite W est dégénérée :

Corollaire 23. *Si* $\mathbb{E}\Big[\log \widehat{L}\Big] < \infty$, *alors* $\mathbb{E}[W] = 1$ *et* $\mathbb{P}(W > 0) > 0$.
Sinon, $\mathbb{E}[W] = 0$ *et* $W = 0$ *p.s.*

Notons que

$$\mathbb{E}\Big[\log \widehat{L}\Big] = \sum_{k=1}^{\infty} \widehat{p_k} \log k = \frac{1}{m} \sum_{k=1}^{\infty} k \log(k) p_k = \frac{1}{m} \mathbb{E}[L \log(L)].$$

et le critère obtenu implique donc un moment de la loi de reproduction p très légèrement supérieur à 1.

On est en mesure de montrer le résultat suivant donnant les conditions pour lesquelles la croissance de la taille de population est la croissance malthusienne attendue, à savoir une croissance géométrique de raison m, à un préfacteur aléatoire. Ce préfacteur W est la martingale limite, non nulle et finie sur l'évément de survie. Son caractère aléatoire est du aux premières générations en petite population, avant que la loi des grands nombres n'opère et confère à Z_n une croissance déterministe. La preuve originale est due à Kesten et Stigum en 1966 et repose sur l'utilisation des fonctions génératrices. La preuve établie ici par la construction spinale est due à Lyons Pemantle et Peres [22].

Théorème 24 (Kesten-Stigum, 1966)**.** *Les assertions suivantes sont équivalentes :*

i) $\{W = 0\} = \{\exists n \in \mathbb{N},\, Z_n = 0\}$ *p.s.,*

ii) $\mathbb{E}[W] = 1$,

iii) $\mathbb{E}[L\log(L)] < \infty$.

Démonstration. Si $\mathbb{E}[L\log(L)] < \infty$, le corollaire 23 implique ii) et pour montrer i), il suffit de montrer que $\mathbb{P}(W = 0) = p_{ext}$ puisque $\{\exists n \in \mathbb{N},\, Z_n = 0\} \subset \{W = 0\}$ p.s.
Si $\mathbb{E}[L\log(L)] = \infty$, le corollaire 23 implique aussi que i) et ii) ne sont pas des assertions valides.

Pour justifier que $\mathbb{P}(W = 0) = p_{ext}$, on va montrer qu'en géneral $\mathbb{P}(W = 0) \in \{p_{ext}, 1\}$ et utiliser que $\mathbb{P}(W = 0) < 1$ quand $\mathbb{E}[L\log(L)] < \infty$. Pour cela, montrons que $\mathbb{P}(W = 0)$ est un point fixe de la fonction génératrice f, qui est une fonction strictement convexe. On décompose la population en fonction de leur ancêtre à la génération 1 et on utilise la propriété de branchement (indépendance entre les sous populations engendrées par les individus de la génération 1). On obtient

$$Z_{n+1} = \sum_{i=1}^{Z_1} Z_n^{(i)},$$

avec $Z_n^{(i)} = \#\{u \in \mathbb{G}_{n+1} : u \geq i\}$ des processus de Galton-Watson indépendants et identiquement distribués. Comme $W = \lim_{n\to\infty} Z_{n+1}/m^{n+1}$,

$$\{W = 0, Z_1 = k\} = \{Z_1 = k\} \cap_{i=1}^{k} \left\{\lim_{n\to\infty} Z_n^{(i)}/m^n = 0\right\}$$

et donc

$$\mathbb{P}(W = 0, Z_1 = k) = \mathbb{P}(Z_1 = k)\mathbb{P}(W = 0)^k.$$

En sommant sur k, on obtient $\mathbb{P}(W = 0) = f(\mathbb{P}(W = 0))$. □

1.4.2 Vitesse d'extinction dans le cas sous-critique

On considère ici dans le le cas $m < 1$ et on rappelle qu'alors $p_{ext} = 1$. La suite $(\mathbb{P}(Z_n > 0))_{n\in\mathbb{N}}$ tend donc vers 0 et nous cherchons à savoir à quelle vitesse Notons que la suite $\left(\frac{\mathbb{P}(Z_n>0)}{m^n}\right)$ est décroissante. En effet, la convexité de f implique

$$\frac{f(1) - f(f^n(0))}{1 - f^n(0)} \leq m$$

en utilisant que la pente de cette corde est inférieure à la dérivée de f en 1. Donc

$$\frac{1 - f^{n+1}(0)}{m^{n+1}} \leq \frac{1 - f^n(0)}{m^n}$$

et $\mathbb{P}(Z_{n+1} > 0)/m^{n+1} \leq \mathbb{P}(Z_n > 0)/m^n$. Par monotonie, cette quantité admet une limite et on défini

$$c = \lim_{n\to\infty} \frac{\mathbb{P}(Z_n > 0)}{m^n} \in [0, 1].$$

La vitesse d'extinction est bien m^n dans le cas sous critique si $c > 0$

Théorème 25. *Si* $\mathbb{E}[L\log(L)] < \infty$, *alors* $c > 0$. *Sinon* $c = 0$

Le critère pour évaluer la vitesse d'extinction par la moyenne, c'est à dire comparer $\mathbb{P}(Z_n > 0)$ à $\mathbb{E}(Z_n)$, fait à nouveau intervenir la critère $L\log L$ comme pour la croissance malthusienne. La preuve peut se faire aussi en exploitant la construction spinale ou les fonctions génératrices et nous renvoyons respectivement à [22] et [1]. Mentionnons enfin que dans le cas critique la vitesse d'extinction est moindre. Sous la condition $\mathbb{E}(L^2) < \infty$, elle est équivalente à σ/n, avec σ l'écart type de L.

1.5 Galton-Watson en environnement aléatoire

Une généralisation importante des processus de Galton-Watson est la prise en compte des variations et changements de l' environnement. Cela peut être causé par les changements métérologiques, sanitaires, les modifications du milieu, de la pression de sélection et de la prédation... Une façon de le modéliser est d'autoriser la loi de reproduction à changer à chaque génération.

On considère du coup maintenant une famille d'environnement E et la famille de loi de reproduction associée, c'est à dire des lois de probabilités sur $\mathbb{N}$ notées $(p^e)_{e\in E}$. Notons $\mathbf{e} = (e_0, e_1, \dots e_n, \dots)$ une suite d'environnements, avec e_n l'environnement dans lequel la population évolue entre la génération n et $n+1$.

Définition 26. *Le processus de Galton Watson en environnement* $\mathbf{e}$ *et de loi de reproduction* $(p^e)_{e\in E}$ *est le processus* $Z = (Z_n^{\mathbf{e}})_{n\geq 0}$ *défini récursivement par* $Z_0 = 1$ *et pour tout* $n \geq 0$,

$$Z_{n+1}^{\mathbf{e}} = \sum_{i=1}^{Z_n^{\mathbf{e}}} L_{i,n}^{e_n},$$

avec $(L_{i,n}^e)_{i\geq 1, n\geq 0, e\in E}$ *v.a. indépendantes et* $L_{i,n}^e$ *de loi* p^e.

Notons que ce processus vérifie encore la *propriété de branchement* : deux sous populations disjointes évoluent indépendamment.
On suppose dorénavant que pour tout $e \in E$,

$$m(e) = \sum_{k\geq 0} k p_k^e = \mathbb{E}(L^e) \in (0, \infty)$$

pour éviter les cas dégénérés et on a le résultat suivant.

Proposition 27. *i) En notant* $s_n^{\mathbf{e}} = \sum_{i=0}^{n-1} \log(m(e_i))$ *(avec* $s_0^{\mathbf{e}} = 0$*), on a*

$$\mathbb{E}(Z_n^{\mathbf{e}}) = \exp(s_n^{\mathbf{e}})$$

et $\left(\exp(-s_n^{\mathbf{e}}) Z_n^{\mathbf{e}}\right)_{n\in\mathbb{N}}$ *est une martingale.*
ii) De plus, en notant, $\ell_n^{\mathbf{e}} = \min\{s_i^{\mathbf{e}} : 0 \leq i \leq n\}$ *(avec* $\ell_0^{\mathbf{e}} = 0$*),*

$$\mathbb{P}(Z_n^{\mathbf{e}} > 0) \leq \exp(\ell_n^{\mathbf{e}})$$

Démonstration. On montre avec la définition de $Z^{\mathbf{e}}$ que

$$\mathbb{E}(Z^{\mathbf{e}}_{n+1} \,|\, \mathcal{F}_n) = m(e_n) Z^{\mathbf{e}}_n$$

et on raisonne par récurrence pour i).
Pour ii), on remarque que $\mathbb{P}(Z^{\mathbf{e}}_n > 0)$ est décroissante, ainsi $\mathbb{P}(Z^{\mathbf{e}}_n > 0) \leq \min\{\mathbb{P}(Z^{\mathbf{e}}_i > 0) : 0 \leq i \leq n-1\}$. Puis on applique l'inégalité de Markov, de sorte à majorer $\mathbb{P}(Z^{\mathbf{e}}_i > 0)$ par $\exp(s^{\mathbf{e}}_i)$. □

Exercice. En notant $f_e(s) = \mathbb{E}(s^{L^e})$ et pour $i \leq j$,

$$f_{i,j} = \mathbb{E}\left(s^{Z^{\mathbf{e}}_j} \,\middle|\, Z^{\mathbf{e}}_i = 1\right)$$

montrer que $f_{i,j} = f_{e_i} \circ \ldots f_{e_{j-1}} \circ f_{e_j}$.

Quittes à augmenter l'espace de probabilité, on considère maintenant un environnement aléatoire $\mathcal{E} = (\mathcal{E}_0, \mathcal{E}_1, \ldots, \mathcal{E}_n \ldots)$ et on suppose ici que les v.a. $\mathcal{E}_i (i \geq 1)$ sont i.i.d. à valeurs dans E et indépendantes de $(L^e_{i,n})_{i \geq 1, n \geq 0, e \in E}$.

Définition 28. *Le processus de Galton Watson en environnement aléatoire $\mathcal{E}$ et de loi de reproduction $(p^e)_{e \in E}$ est le processus $Z = (Z^{\mathcal{E}}_n)_{n \geq 0}$.*

Notons que ce processus ne vérifie plus la propriété de branchement, l'environnement commun corrélant deux sous populations du processus. Par contre conditionnellement à l'environnement, on obtient un processus de Galton-Watson dans un environnement fixé pour lequel la propriété est vrai. On parle de la propriété de branchement *quenched*.

Etablissons le critère d'extinction de la population, en évitant le cas dégénéré où la loi de reproduction est restreinte à 1 et la population reste constante p.s. Un premier calcul sur la moyenne (utiliser le lemme précédent i)) donne

$$E(Z_n) = \mathbb{E}(\exp(S_n)) = \mathbb{E}(m(\mathcal{E}_0)^n)$$

avec $S_n = \sum_{i=0}^{n-1} \log(m(\mathcal{E}_i))$ marche aléatoire de pas $\log(m(\mathcal{E}_0))$.

Théorème 29. *Supposons que $\mathbb{E}(|\log(m(\mathcal{E}_0))|) < \infty$ et $\mathbb{P}(Z_1 = 1) < 1$.*

$$\mathbb{P}(\exists n \geq 0,\, Z_n = 0) = 1 \quad \textit{si et seulement si} \quad \mathbb{E}[\log(m(\mathcal{E}_0))] \leq 0$$

Esquisse de preuve. Considérons le cas $\mathbb{E}[\log(m(\mathcal{E}_0))] \leq 0$ et rappelons que S est une marche aléatoire de pas $\log(m(\mathcal{E}_0))$. Alors, quand $n \to \infty$, S_n dérive linéairement vers $-\infty$ (si $\mathbb{E}(\log(m(\mathcal{E}_0))) < 0$) ou oscille entre $+\infty$ et $-\infty$ (si $\mathbb{E}(\log(m(\mathcal{E}_0))) = 0$). Dans les deux cas, $I_n = \min\{S_i : 0 \leq i \leq n\}$ tend vers $-\infty$ et la proposition 27 assure que $\mathbb{P}(Z_n > 0)$ tend vers 0 quand $n \to \infty$. La réciproque est plus délicate et on renvoie à [2] pour la preuve. □

De plus $\exp(-S_n) Z_n$ est une martingale positive qui converge vers une v.a. finie positive. Lorsque $p_{ext} < 1$, sur l'événement de survie, on montre facilement que la population tend vers l'infini car 0 est absorbant et accessible depuis tout compact quand $\mathbb{P}(Z_1 = 0) > 0$ et sinon le processus est strcitement croissant. Si de plus $\mathbb{E}\big(\log\big(\widehat{Z_1}\big)\big) < \infty$,

alors comme dans la section précédente on peut montrer que $W > 0$ sur l'evénement de survue et obtenir la croissance de Z_n via celle de S_n.

Complétons ces résultats par la vitesse d'extinction de ces processus. Nous supposons ici que la loi de reproduction a un second moment bornée :

$$\sup_{e \in E} \mathbb{E}((L^e)^2) < \infty$$

et nous renvoyons à [13] pour la preuve et des résultats sous des conditions de moments plus faibles

— *Cas fortement sous critique.* Si $\mathbb{E}(m(\mathcal{E}_0)\log(m(\mathcal{E}_0))) < 0$, alors il existe $c \in (0, \infty)$ tel que
$$\mathbb{P}(Z_n > 0) \sim_{n\to\infty} c\mathbb{E}(m(\mathcal{E}_0))^n$$

— *Cas intermédiaire sous critique.* Si $\mathbb{E}(m(\mathcal{E}_0)\log(m(\mathcal{E}_0))) = 0$, alors il existe $c \in (0, \infty)$ tel que
$$\mathbb{P}(Z_n > 0) \sim_{n\to\infty} cn^{-1/2}\mathbb{E}(m(\mathcal{E}_0))^n$$

— *Cas faiblement sous critique ;* Si $\mathbb{E}(m(\mathcal{E}_0)\log(m(\mathcal{E}_0))) > 0$ et $\mathbb{E}(\log(m(\mathcal{E}_0))) < 0$, alors il existe $c \in (0, \infty)$ tel que
$$\mathbb{P}(Z_n > 0) \sim_{n\to\infty} cn^{-3/2}\gamma^n,$$
avec
$$\gamma := \inf_{0\le s\le 1} \mathbb{E}[m(\mathcal{E}_0)^s] < \mathbb{E}[m(\mathcal{E}_0)].$$

— *Cas critique.* Si $\mathbb{E}(\log(m(\mathcal{E}_0))) = 0$, alors
$$\mathbb{P}(Z_n > 0) \sim_{n\to\infty} cn^{-1/2}.$$

Notons que le cas fortement sous critique généralise le cas Galton-Watson (sans environnement changeant) avec un équivalent de la probabilité de survie donné par le nombre moyen d'individu (la majoration s'obtient d'ailleurs immédiatement par l'inégalité de Markov comme vu précédemment). Dans le régime faiblement sous critique, la vitesse est plus lente, c'est du à la possibilité que des environnements particulièrement favorables est lieu au d´but de la dynamique, permettant à la population de survivre. Si cet événement reste peu probable, il l'est bien plus qu'une survie de la population dans des environnements défavorables comme dans le régime fort. Dans le cas critique, la probabilité de survie est supérieure au cas sans environnements changeant pour les mèmes raisons. Plus généralement, les processus de branchement en environnement aléatoire offrent un cadre intéressant où se mêle la stochasticité démorgraphique et environnementale, notions importantes en biologie et écologie, renvoyons à [18, 19] pour des références.

Chapitre 2

Populations structurées : le cas neutre

Les individus sont dorénevant caractérisés par un trait. Commençons par le cadre neutre, c'est à dire que le trait n'affecte pas la loi de reproduction mais il peut évoluer et se transmettre avec une transition possible au moment de la reproduction. Comme la loi de reproduction est fixe, la généalogie est donnée par un arbre de Galton Watson. Lévolution des traits va être modélisée par une chaîne de Markov sur cet arbre

2.1 Chaîne de Markov sur un arbre de Galton-Watson

2.1.1 Définition du modèle et exemples

Nous allons considérer un arbre aléatoire $\mathcal{T}$ pour décrire la généalogie d'une population sans interaction, avec une reproduction de loi p, en prenant donc l' arbre de de Galton-Watson construit dans le chapitre précédent. Tel un sapin à Noel, cet arbre va maintenant être décoré des traits des individus que les noeuds représentent. De façon informelle, chaque individu de trait x va indépendamment donner naissance à un nombre aléatoire L d'individus, de traits $\Theta_x = (X_i^x : i = 1, \dots, L)$.

Plus formellement, on appelle $\mathcal{X}$ l'espace des traits, espace mesurable muni de la tribu $\mathcal{B}_{\mathcal{X}}$, typiquement $\mathbb{N}$ ou $\mathbb{N}^k$ muni de l'ensemble des parties, ou $\mathbb{R}^d$, $(\mathbb{R}_+)^d$ munis des boréliens.
On pose $\mathbb{X} = \cup_{k\geq 0}\mathcal{X}^k$ que l'on munit de la tribu engendrée par $\cup_{k\geq 0}(\mathcal{B}_{\mathcal{X}})^k$. La transmission du trait aux enfants est modélisée par la famille de processus ponctuel $(\Theta_x)_{x\in\mathcal{X}}$, avec pour tout $x \in \mathcal{X}$, $\Theta_x = (X_i^x : i = 1, \dots, L)$ v.a. à valeurs dans $\mathbb{X}$ et $\Theta : \mathcal{X} \times \Omega \to \mathbb{X}$ mesurable avec $\mathcal{X} \times \Omega$ muni de la tribu produit. Notons $\Theta_x(u) = (X_i^x(u) : i = 1, \dots, L(u))$ où $(Theta_x(u))_{x\in\mathcal{X}}$ famille de v.a. indépendantes et identiquement distribuées pour $u \in \mathcal{U}$, comme $(\Theta_x)_{x\in\mathcal{X}}$. En particulier les v.a. $(L(u))_{u\in\mathcal{U}}$ sont i.i.d. et distribuées comme L.

Définition 30. *La chaîne $X = \{(u, X(u)) : u \in \mathcal{T}\}$ issue d'un individu de trait x_0 et de transitions $\big((\Theta_x(u))_{x\in\mathcal{X}} : u \in \mathcal{U}\big)$ est définie récursivement par $\mathcal{T} = \cup_{n\in\mathbb{N}}\mathbb{G}_n$ et*

$$\mathbb{G}_0 = \{\varnothing\}, \qquad X(\varnothing) = x_0$$

et pour tout $n \geq 0$,

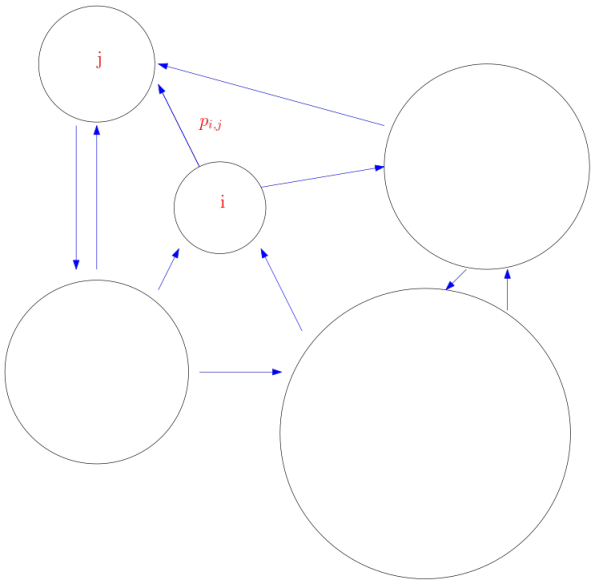
$$\mathbb{G}_{n+1} = \{uk : u \in \mathbb{G}_n,\, k \in [\![1, L(u)]\!]\}, \qquad X(uk) = X_k^{X(u)}(u)$$

pour tous $u \in \mathbb{G}_n$ *et* $k \in [\![1, L(u)]\!]$.

La loi de X est déterminée par la loi de $(\Theta_x)_{x\in\mathcal{X}}$ et x_0. Notons que cette définition s'étend immédiatement à une condition initiale avec plusieurs individus et des traits aléatoires. Remarquons aussi que la taille $Z_n = \#\mathbb{G}_n$ de la population à l'instant n est un processus de Galton-Watson.

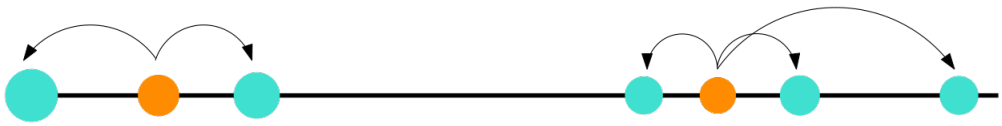
Donnons ici quelques exemples de modèles de populations, sur lesquels nous reviendrons, pour les étudier.

Graphe fini irréductible et métapopulations. On considère un ensemble de traits fini $\mathcal{X}$ et un une matrice de transition $(p_{x,y})_{x,y\in\mathcal{X}}$ sur $\mathcal{X}$ irréductible. C'est à dire que $p_{x,y} \in [0,1]$, $\sum_{y\in\mathcal{X}} p_{x,y} = 1$ pour tout $x \in \mathcal{X}$, et pour tous $x, y \in \mathcal{X}$, on peut trouver un chemin dans $\mathcal{X}$ reliant x à y par des probabilités $p_{x_k,x_{k+1}}$ strictement positives. On s'intéressera alors à $\Theta_x = (X_i^x : i = 1, \dots, L)$ avec X_i^x identiquement distribués de loi $p^x = (p_{x,y} : y \in \mathcal{X})$, i.e. $\mathbb{P}(X_i^x = y) = p_{x,y}$.
Ce modèle permet d'appréhender différentes situations en écologie, biologie, ou épidémiologie. Typiquement, des métapopulations, où le trait x sera un site géographique (bois, champ, ville, ferme, ...) et les sites sont connectés par des arêtes. Une large littérature en écologie s'interesse à ces modèles, motivés notamment par la gestion de l'éspace et la préservation de la biodiversité (corridors, modèles puits-sources...). On pourra consulter par exemple [6, 24] et les références qui s'y trouvent.



L'espace des traits peut contenir plus de caractéristiques individuelles, par exmple un couple $x = (x_1, x_2)$, avec x_1 la composante spatiale correspondant à un premier graphe, et x_2 une caractéristique phénotypique ou génotypique, les transitions correspondant alors aux mutations, recombinaisons etc...
Les cas réductible est également pertinent dans certains modèles, par exemple la différentiation cellulaire, où la cellule change de type en se différentiant, sans retour en arrière possible. D'un point de vue des processus de branchement, on est ici dans un processus de branchement multitype, où le nombre de types est fini.

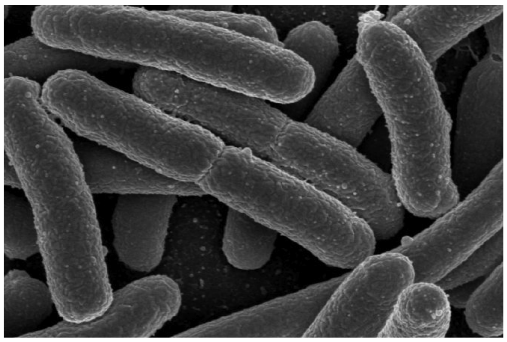
Marche aléatoire branchante et reproduction dispersion. On s'intéresse ici à une population qui se reproduit et les enfants sont dispersés autour des parents, typiquement une population de plantes avec essaimage par le vent ou les pollénisateurs. On voudra notamment comprendre comment une espèce invasive (ou une épidémie) progresse dans l'espace. On prendra ici un espace de traits $\mathcal{X} = \mathbb{R}^d$ et $\Theta_x = (x + X_i : i = 1, \dots, L)$, i.e. $\Theta_x = x + \Theta_0$.



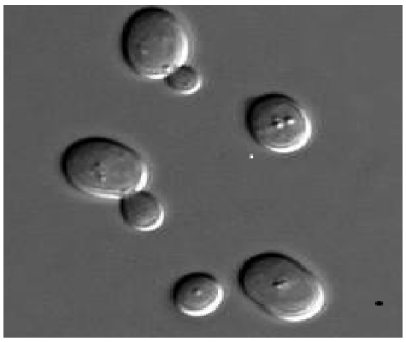
Processus autorégressif bifurcant et division cellulaire. On s'intéressera ici à la transmission d'une caractéristique d'une cellule en division, comme son taux de croissance. La division cellulaire étant binaire, on se focalisera pour cet exemple sur le cas $L = 2$ p.s. et on modélisera la transmission par un processus autorégressif (affine avec pertubation gaussienne) :

$$X_i^x = \alpha_i x + \mathcal{E}_i, \qquad i = 1, 2,$$

avec $\alpha = (\alpha_1, \alpha_2) \in \mathbb{R}^2$ et $\mathcal{E} = (\mathcal{E}_1, \mathcal{E}_2)$ vecteur gaussien centré de matrice de covariance Γ. Notons que dans la division cellulaire, la loi de transmission peut être assymétrique $\alpha_1 \neq \alpha_2$. C'est le cas par exemple des divisions de la levure où une cellule mère donne naissance à une petite cellule fille qui bourgeonne. Cela peut être le cas aussi lorsque la cellule semble se diviser suivant son axe médian, comme pour la bactérie E-Coli, en lien notamment avec des phénomènes de polarisation ou de vieillissement (il y a un vieux pôle et un jeune pôle nouvellement crée au niveau de la membrane à chaque division), on renvoie à [10, 26, 14] pour des détails. Dans ce modèle ou voudra aussi faire dépendre la division du trait, et pour cela on renvoie au chapitre suivant.



E-COLI division



YEAST division

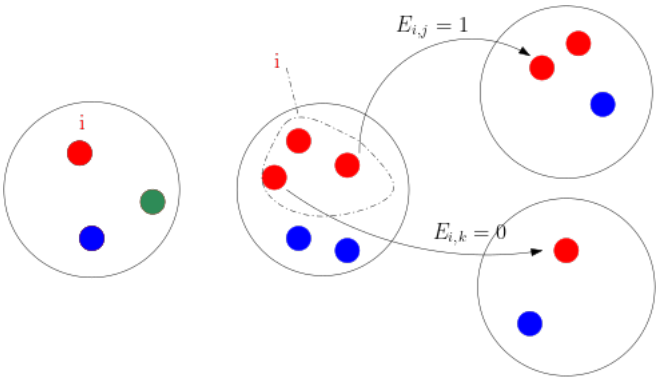
Modèle de Kimmel et infection dans une cellule en division. Dans ce modèle, on suit la croissance d'une infection (pathogènes ; bactériophages,...) à l'intérieur d'une cellule qui se divise. Contrairement au cas précédent, on autorise ici la cellule à mourrir ou ne pas se diviser, cad $p_0 + p_1 + p_2 = 1$ et $L \in \{0, 1, 2\}$ p.s.

Focalisons nous sur le mécanisme le plus simple, déjà riche. Les pathogènes se reproduisent à l'intérieur de la cellule avec la loi μ de moyenne $\overline{\mu}$, pour une génération de cellules. Lorsque la cellule se divise, une répartition aléatoire binomiale de paramètre $q \in (0, 1)$ les affecte dans les deux cellules filles. Autrement dit, chaque pathogène se retrouve dans la première fille avec probabilité q, indépendamment. Concrètement on

introduit une famille de v.a. i.i.d. $(P_i)_{i\in\mathbb{N}}$ de loi μ et une famille $(E_{i,j})_{i,j\in\mathbb{N}}$ v.a. i.i.d. de Bernoulli de paramètre q. La loi de Θ_x est alors définie pour $x \in \mathbb{N}$ et $(y,z) \in \mathbb{N}^2$ par

$$\begin{aligned}
&\mathbb{P}(\Theta_x = \varnothing) = p_0 \\
&\mathbb{P}(\Theta_x = (y)) = p_1 \mathbb{P}\left(\sum_{i=1}^{x} P_i = y\right) \\
&\mathbb{P}(\Theta_x = (y,z)) = p_2 \mathbb{P}\left(\sum_{i=1}^{x}\sum_{j=1}^{P_i} E_{i,j} = y, \sum_{i=1}^{x}\sum_{j=1}^{P_i} 1 - E_{i,j} = z\right)
\end{aligned}$$

En effet, quand une cellule contient x pathogènes, pendant le temps d'une génération cellulaire, ces pathogènes vont se reproduire (et peut être aussi mourir) et leur nombre devient $\sum_{i=1}^{x} P_i$. Si la cellule meurt, ces pathogènes disparaissent. Si la cellule survit sans se diviser, tous les pathogènes restent dans cette même cellule. Si la cellule se divise en deux, le $j$ème pathogène issu du i ème pathogène de la génération précédente se retrouve dans la première cellule fille si $E_{i,j} = 1$, et dans la deuxième sinon.



Ce processus est de type croissance fragmentation. Plus généralement, on peut modéliser la croissance du matériel cellulaire, comme l'ADN extrachromosomales, les plasmides (l'intérêt initial de Marek Kimmel), les mitochondries, et sa transmission au cours de la division.

2.1.2 Propriété de branchement et semigroupe

L'indépendance des v.a. impliquées dans l'évolution de deux individus différents pris à un temps donné et le caractère Markovien de l'évolution des traits assurent la propriété de branchement suivante. On note

$$X^{(u)} = \{(v, X(uv)) : v \in \mathcal{U}, uv \in \mathcal{T}\}$$

la chaîne enracinée en le noeud u (qui est vide si u n'appartient pas à $\mathcal{T}$).

Proposition 31. *Soit $(F_u)_{u\in\mathcal{U}}$ une collection de fonctions mesurables bornées. Pour*

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

$$\mathbb{E}_{(\varnothing,x_0)}\left[\prod_{u\in\mathbb{G}_n} F_u\left(X^{(u)}\right)\middle|\mathcal{F}_n\right] = \prod_{u\in\mathbb{G}_n} \mathbb{E}_{(\varnothing,X(u))}[F_u(X)]\,,$$

Considérons maintenant l'espérance de la mesure empirique.

Définition 32. *On note pour tout $x \in \mathcal{X}$, f mesurable bornée et $n \geq 0$:*

$$M_n(x,f) = \mathbb{E}_{(\varnothing,x)}\left[\sum_{u\in\mathbb{G}_n} f(X(u))\right]$$

La mesure associée est notée $M_n(x,dy)$ et définie par $M_n(x,A) = M_n(x,\mathbf{1}_A)$ pour A mesurable.

On remarque que $M_0(x,f) = f(x)$ et $M_n(x,\mathbf{1}) = \mathbb{E}(\#\mathbb{G}_n) = m^n$. On notera

$$M(x,f) = M_1(x,f) = \mathbb{E}\left[\sum_{i=1}^{L} f(X_i^x)\right].$$

par commodité et on montre que $(M_n)_{n\geq 0}$ est un semi groupe (positif).

Proposition 33. *Pour tous $x \in \mathcal{X}$ et f mesurable bornée et $n \geq 0$,*

$$M_{n+1}(x,f) = M_n(x,M(\cdot,f)) = M(x,M_n(\cdot,f)).$$

En itérant cette relation, on obtient ici $M_n = M^{\circ n}$ et $M_{n+p} = M_n \circ M_p$ avec $\circ$ la règle de composition obtenue dans la proposition.
Dans le cas d'un espace de trait $\mathcal{X}$ fini, en notant $M = (M_{i,j})_{i,j\in\mathcal{X}}$ la matrice définie par

$$M_{i,j} = M(i,\mathbf{1}_j) = \mathbb{E}_{(\varnothing,i)}(\#\{u \in \mathbb{G}_1 : X(u) = j\}),$$

avec $M_{i,j}$ qui donne ainsi le nombre moyen d'individus de type j qui sont enfants d'un individu de type i. On obtient donc $M_n(i,\mathbf{1}_j) = (M^n)_{i,j}$, avec M^n la puissance $n$ième de la matrice M pour le produit usuel.

Démonstration. Calculons

$$\begin{aligned}\mathbb{E}\left(\sum_{u\in\mathbb{G}_{n+1}} f(X(u))\,\middle|\,\mathcal{F}_n\right) &= \mathbb{E}\left(\sum_{u\in\mathbb{G}_n, v\in\mathbb{G}_{n+1},\, v\succcurlyeq u} f(X(v))\,\middle|\,\mathcal{F}_n\right)\\ &= \sum_{u\in\mathbb{G}_n}\mathbb{E}\left(\sum_{v\in\mathbb{G}_{n+1},\, v\succcurlyeq u} f(X(v))\,\middle|\,\mathcal{F}_n\right).\end{aligned}$$

La propriété de branchement [1] assure que

$$\mathbb{E}\left(\sum_{v\geq 1,\, uv\in\mathbb{G}_{n+1}} f(X(uv))\,\middle|\,\mathcal{F}_n\right) = M(X(u),f)$$

1. appliquée à $F_w = 1$ pour $w \neq u$ et $F_u(t) = \sum_{u\in g_1} f(x(u))$ avec $(x(u) : u \in g_1)$ la collection des traits des individus de la premiere génération de l'arbre t

et donc

$$\begin{aligned} M_{n+1}(x,f) &= \mathbb{E}_{(\varnothing,x)}\left[\mathbb{E}\left(\sum_{u\in\mathbb{G}_{n+1}} f(X(u)) \,\middle|\, \mathcal{F}_n\right)\right] \\ &= \mathbb{E}_{(\varnothing,x)}\left[\sum_{u\in\mathbb{G}_n} M(X(u),f)\right] = M_n(x, M(.,f)). \end{aligned}$$

L'autre identité se prouve de façon analogue en conditionnant par $\mathcal{F}_1$, ou bien en utilisant l'associativité de la composition. □

Illustrons cette notion avec la marche aléatoire branchante. Dans ce cas, $M(x,f) = \mathbb{E}\left[\sum_{i=1}^L f(x+X_i)\right]$. En particulier, si les v.a. X_i ont des moments exponentiels,

$$M(x, \exp(\rho.)) = \mathbb{E}\left[\sum_{i=1}^L e^{\rho x}.e^{\rho X_i}\right] = \lambda(\rho)e^{\rho x}$$

avec $\lambda(\rho) = \mathbb{E}\left[\sum_{i=1}^L \exp(\rho X_i)\right]$. Les fonctions exponentielles sont des fonctions propres de M, ce que nous utiliserons par la suite pour comprendre le front de propagation.

2.1.3 Formule *tous-pour-un* et fourches

Définition 34. *Pour $x \in \mathcal{X}$ et f mesurable bornée, on note*

$$Q(x,f) = \frac{1}{m}M(x,f)$$

La mesure associée $Q(x,dy)$ est définie par $Q(x,A) = Q(x,\mathbf{1}_A)$ pour A mesurable.

L'opérateur Q est un noyau de transition, c'est à dire que pour tout $x \in \mathcal{X}$, $Q(x,\mathcal{X}) = Q(x,\mathbf{1}) = 1$ et $Q(x,.)$ est une loi de probabilité. Il se réécrit

$$Q(x,A) = \sum_{k\geq 0}\frac{p_k}{m}\sum_{1\leq i\leq k}\mathbb{P}(X_i^x \in A | L=k) = \sum_{k\geq 0}\widehat{p}_k Q^{(k)}(x,A)\,,$$

où $\widehat{p}$ est la loi de reproduction biaisée par la taille et $Q^{(k)}(x,\cdot)$ est la loi de probabilité donnant le trait d'un enfant tiré uniformément au hasard quand le parent a le trait x et k enfants : $Q^{(k)}(x,\cdot) = k^{-1}\sum_{1\leq i\leq k}\mathbb{P}(X_i^x \in \cdot | L=k)$.

On considère une chaîne de Markov Y de noyau Q, c'est-à-dire que

$$\mathbb{P}_{y_0}(Y_1 \in A_1, \ldots, Y_n \in A_n) = \int_{A_1\times\ldots\times A_n}\Pi_{i=0}^{n-1}Q(y_i, dy_{i+1})$$

Proposition 35 (Tous pour un)**.** *Soit $F : \mathcal{X}^{n+1} \to \mathbb{R}$ mesurable bornée et $n \geq 0$.*

$$\mathbb{E}_{(\varnothing,x)}\left[\sum_{u\in\mathbb{G}_n} F(X(v),\, v \preccurlyeq u)\right] = m^n\mathbb{E}_x[F(Y_0,\ldots,Y_n)].$$

Ce résultat permet de réduire, en moyenne, la mesure empirique des trajectoires des individus à une seule trajectoire, réalisée par une à une chaîne de Markov. Nous verrons que ce résultat se généralise (notamment ici à des modèles non neutres, mais plus généralement à des environnements changeants ou à certains modèles avec interaction).
Remarquons déjà que ce résultat a un interêt numérique quand $m > 1$ puisqu'il permet de ramener certaines questions à la simulation de l'évolution du trait d'un seul individu, plutôt que toute une population qui croit exponentiellement. Le processus Y qui permet cette analyse du "trait typique" a été construit à partir des noyaux de transitions $Q^{(k)}$, avec un biais par la taille qui vient de la loi de reproduction de l'épine obtenue dans le chapitre précdent. En particulier, le processus Y est en général bien différent (en loi) du processus obtenu en suivant un individu dans le sens usuel du temps, en choisissant en chaque génération un enfant uniformément au hasard parmi la descendance de cet individu.
Par exemple, en prenant $F(x_i, i \leq n) = \mathbb{1}_A(x_n)$, on obtient le nombre moyen d'individus dont le trait appartient à A à la génération n :

$$\mathbb{E}_{(\varnothing,x)}[\#\{u \in \mathbb{G}_n : X(u) \in A\}] = m^n \mathbb{P}_x(Y_n \in A)$$

ce qui va nous être utile pour analyser la répartition des traits dans la population ou le front de propagation.

Démonstration. On se ramène aux fonctions $F_n(x_0, \ldots, x_n) = \Pi_{j=0}^n f_j(x_j)$ avec f_j fonctions mesurables bornées en utilisant le théorème des classes monotones. Ainsi, pour une telle fonction, en utilisant à nouveau la propriété de branchement, et oubliant la condition initiale dans la notation (car elle est fixée) :

$$\begin{aligned}
\mathbb{E}\left[\sum_{u\in\mathbb{G}_{n+1}} F_{n+1}(X(v),\, v \preccurlyeq u)\right] &= \mathbb{E}\left[\sum_{w\in\mathbb{G}_n} \mathbb{E}\left[\sum_{u\in\mathbb{G}_{n+1},\, u\succcurlyeq w} F_{n+1}(X(v), v \preccurlyeq u)\middle|\mathcal{F}_n\right]\right] \\
&= \mathbb{E}\left[\sum_{w\in\mathbb{G}_n} F_n(X(v),\, v \preccurlyeq w)\mathbb{E}\left[\sum_{u\in\mathbb{G}_{n+1},\, u\succcurlyeq w} f_{n+1}(X(u))\middle|\mathcal{F}_n\right]\right] \\
&= \mathbb{E}\left[\sum_{w\in\mathbb{G}_n} F_n(X(v),\, v \preccurlyeq w)\, M(X(w), f_{n+1})\right] \\
&= \mathbb{E}\left[\sum_{w\in\mathbb{G}_n} \widetilde{F}_n(X(v),\, v \preccurlyeq w)\right],
\end{aligned}$$

avec

$$\widetilde{F}_n(x_i, i \leq n) = F_n(x_i, i \leq n)\, M(x_n, f_{n+1}) = m\, F_n(x_i, i \leq n)\, Q(x_n, f_{n+1}).$$

On peut alors utiliser l'hypothèse de récurrence au rang n, ce qui nous donne

$$\mathbb{E}\left[\sum_{w\in\mathbb{G}_n} \widetilde{F}_n(X(w))\right] = m^n \mathbb{E}(\widetilde{F}_n(Y_0, \ldots, Y_n))$$

et on conclut avec la définition de Y :

$$\begin{aligned}
\mathbb{E}_x(\widetilde{F}_n(Y_0, \ldots, Y_n)) &= m \int_{\mathcal{X}^n} F_n(y_i, i \leq n) \left(\int_{\mathcal{X}} Q(y_n, dy_{n+1}) f_{n+1}(y_{n+1})\right) \Pi_{i=0}^{n-1} Q(y_i, dy_{i+1}) \\
&= m\mathbb{E}_x[F_{n+1}(Y_0, \ldots, Y_{n+1})].
\end{aligned}$$

en notant $y_0 = x$ et en rassemblant les intégrales par Fubini. □

Exercice. Proposer une autre preuve de cette formule en utilisant la construction spinale du chapitre précédent. *Indication : on pourra conditionner par l'arbre de Galton-Watson sur lequel vit la chaîne de Markov.*

Illustrons ce résultat avec les modèles de dynamique des populations introduits dans la première partie.
Pour le graphe fini et les métapoputions, pour tous $x, y \in \mathbb{X}$,

$$Q(x, \{y\}) = Q^{(k)}(x, \{y\}) = p_{xy}$$

Pour le deuxième exemple, la marche aléatoire branchante :

$$Q^{(k)}(x, x + dy) = \frac{1}{k} \sum_{i=1}^{k} \mathbb{P}(X_i \in dy \,|\, L = k)$$

Pour le processus autoregressif bifurcant,

$$Q(x, \{dy\}) = Q^{(2)}(x, \{y\}) = \frac{1}{2} \left(\mathbb{P}(\alpha_1 x + \mathcal{E}_1 \in dy) + \mathbb{P}(\alpha_2 x + \mathcal{E}_2 \in dy) \right)$$

avec $\mathcal{E}_i$ v.a. gaussienne centrée de variance Γ_{ii}. Le processus Y est alors un processus autoregressif (transformation affine aléatoie), où le coefficient est tiré aléatoirement à chaque pas de temps, unfirmomément parmi $\{\alpha_i, \alpha_2\}$. Nous reviendrons dessus.
Enfin considérons le modèle de Kimmel. Dans ce cas, $L \in \{0, 1, 2\}$. Alors $m = p_1 + 2p_2$, $\widehat{p}_1 = \frac{p_1}{m}$ et $\widehat{p}_2 = \frac{2p_2}{m}$. Le noyau de transition de la chaîne Y est donné par

$$Q^{(1)}(x, A) = \mathbb{P}\left(\sum_{i=1}^{x} P_i \in A \right)$$

$$Q^{(2)}(x, A) = \frac{1}{2} \left(\mathbb{P}\left(\sum_{i=1}^{x} \sum_{j=1}^{P_i} E_{i,j} \in A \right) + \mathbb{P}\left(\sum_{i=1}^{x} \sum_{j=1}^{P_i} 1 - E_{i,j} \in A \right) \right).$$

en rappelant que $P_i = P$ (en loi) est la loi de reproduction des pathogènes (sur une génération). L'évolution Y du nombre de pathogènes est donc un processus de branchement en environnement aléatoire. La division cellulaire induit un environnement aléatoire sur la population de pathogènes, à la fois dû au fait que la cellule peut mourir, se diviser ou survivre. Quand elle se divise et que $q \neq 1$, la répartition est asymétrique entre deux cellules filles. Au final, soit la cellule meurt et tue la sous-population de pathogènes meurt avec, soit elle n'évolue pas et l'environnement est neutre, soit elle se divise et crée deux sous-populations et alors dans une lignée une seule des sous populations est conservée, l'autre disparaît.

Donnons dès maintenant une conséquence de la formule "tous-pour-un" précédente, qui sera utile par la suite. Il s'agit de baiser les traits des individus échantillonnés. En appliquant la formule à $F(x_1, \ldots, x_n) = h(x_n)G(x_1, \ldots x_n)$, on obtient

$$\mathbb{E}_{(\varnothing, x)} \left[\sum_{u \in \mathbb{G}_n} h(X(u)) G(X(v) : v \preccurlyeq u) \right] = m^n \mathbb{E}_x[h(Y_n) G(Y_0, \ldots, Y_n)].$$

et on en déduit le résultat suivant en prenant une fonction propre h du semigroupe.

Corollaire 36. *Supposons qu'il existe une fonction mesurable positive h sur $\mathcal{X}$ telle que $Mh = \lambda h$ pour un certain $\lambda \geq 0$. Alors pour tous $n \geq 0$ et G mesurable positive sur $\mathcal{X}^n$ et pour tout $x \in \mathcal{X}$:*

$$\mathbb{E}_{(\varnothing,x)}\left[\sum_{u\in\mathbb{G}_n} h(X(u))G(X(v),\, v \preccurlyeq u)\right] = \lambda^n h(x)\mathbb{E}_x\Big[G\Big(\widetilde{Y}_0,\ldots,\widetilde{Y}_n\Big)\Big],$$

avec $(\widetilde{Y}_n)_{n\geq 0}$ une chaine de Markov de noyau de transition

$$\widetilde{Q}(x,dy) = \frac{mh(y)}{\lambda h(x)}Q(x,dy) = \frac{h(y)}{\lambda h(x)}M(x,dy)$$

Démonstration. On est ramené au calcul

$$\begin{aligned}\mathbb{E}_x\Big[G\Big(\widetilde{Y}_0,\ldots,\widetilde{Y}_n\Big)\Big] &= \int_{\mathcal{X}^n} G(y_0,\ldots,y_n)\Pi_{i=0}^{n-1}\frac{mh(y_{i+1})}{\lambda h(y_i)}Q(y_i,dy_{i+1})\\ &= \frac{m^n}{\lambda^n}\int_{\mathcal{X}^n} G(y_0,\ldots y_n)\frac{h(y_n)}{\lambda h(x)}\Pi_{i=0}^{n-1}Q(y_i,dy_{i+1})\\ &= \frac{m^n}{\lambda^n.h(x)}\mathbb{E}_x[h(Y_n)G(Y_0,\ldots,Y_n)].\end{aligned}$$

□

Cette formule permet de considérer la population totale $\mathbb{G}_n$ avec un biais h sur le trait, et de ramener sa moyenne à l'étude d'un processus de Markov homogène $\widetilde{Y}$. Le noyau de cette chaîne est la h transformée du noyau Q (ou M). Pour l'étude de la propagation du front dans la marche branchante, nous utiliserons la fonction propre $h(x) = \exp(\rho x)$ qui privéligiera dans la somme les positions les plus grandes. Notons que nous n'avons pas imposé h borné dans la formule precedente, la formule peut se montrer pour les fonctions h et G positive par limite croissance.

On a pour l'instant établi le comportement moyen d'un individu typique, correspondant à un échantillon dans la population. On va vouloir maintenant comprendre ce qui se passe quand on échantillonge deux individus. Les traits de deux individus se retrouvent coréllés par le trait de leur ancêtre commun. Cette question va intervenir en particulier pour obtenir un effet de loi des grands nombres et décrire le profil des traits dans la population, en vérifiant que la décorrélation ne se produit pas trop tard pour assurer une limite déterministe. En choisissant deux individus et en traçant leurs lignées ancestrales, on observe une forme de *fourche*.

On se place dans le reste de ce chapitre dans le cas d'une variance finie pour la loi de reproduction :

$$\mathbb{E}\big[L^2\big] < \infty.$$

Pour une famille de fonctions $(f_i)_{i\in\mathbb{N}}$ (resp. $(g_i)_{i\in\mathbb{N}}$) sur $\mathcal{X}$, on note

$$F_{i,j}(x) = \prod_{k=0}^{j-i} f_{i+k}(x_k), \qquad \text{resp.} \quad G_{i,j}(x) = \prod_{k=0}^{j-i} g_{i+k}(x_k)$$

pour $x = (x_0,\ldots,x_{j-i}) \in \mathcal{X}^{j-i+1}$.

Proposition 37 (Fourches)**.** *Soient* $(f_i)_{i\in\mathbb{N}}$ *et* $(g_i)_{i\in\mathbb{N}}$ *des fonctions mesurables bornées sur* $\mathcal{X}$*. Pour tout* $x \in \mathcal{X}$ *et* $n \in \mathbb{N}$ *:*

$$\mathbb{E}_{(\varnothing,x)}\left[\sum_{u\neq v\in\mathbb{G}_n} F_{0,n}(X(w),\, w \preccurlyeq u)\, G_{0,n}(X(w),\, w \preccurlyeq v)\right]$$
$$= \sum_{p=0}^{n-1} m^{2n-p-2}\mathbb{E}_x[F_{0,p}G_{0,p}(Y_i,\, i \leq p)\, H_{p,n}(Y_p)]\,,$$

avec $H_{p,n}$ *définie par*

$$H_{p,n}(y) = \mathbb{E}_{(\varnothing,y)}\left[\sum_{a\neq b\in\mathbb{G}_1} Q(X(a), F_{p+1,n})\, Q(X(b), G_{p+1,n})\right].$$

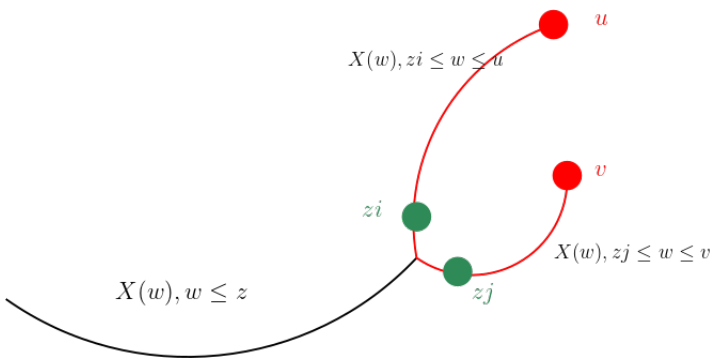
pour $y \in \mathcal{X}$ *et pour* $F : (y_0, \dots y_n) \in \mathcal{X}^{n+1} \to F(y_0, \dots y_n)$ *mesurable bornée*

$$Q(y, F) = \mathbb{E}_y[F(Y_i, 0 \leq i \leq n)]$$

La preuve va utiliser une décomposition de la généalogie en fonction du premier (plus récent) ancêtre commun z d'un couple d'individu (u, v).

Démonstration. Soit $n \in \mathbb{N}$ et $p \leq n-1$. Pour tout $z \in \mathbb{G}_p$ et $i \neq j \in \mathbb{N}$, on peut factoriser pour $u \succcurlyeq zi, v \succcurlyeq zj$,

$$\begin{aligned} &F_{0,n}(X(w), w \preccurlyeq u)\, G_{0,n}(X(w), w \preccurlyeq v) \\ &\quad = F_{0,p}G_{0,p}(X(w), w \preccurlyeq z)\, F_{p+1,n}(X(w), zi \preccurlyeq w \preccurlyeq u) G_{p+1,n}(X(w), zj \preccurlyeq w \preccurlyeq v). \end{aligned}$$



Du coup pour tout $z \in \mathbb{G}_p$ et $i \neq j \in \mathbb{N}$,

$$\mathbb{E}\left[\sum_{\substack{u,v\in\mathbb{G}_n \\ zi\preccurlyeq u, zj\preccurlyeq v}} F_{0,n}(X(w), w \preccurlyeq u)\, G_{0,n}(X(w), w \preccurlyeq v)\,\middle|\mathcal{F}_{p+1}\right] = F_{0,p}G_{0,p}(X(w), w \preccurlyeq z)\, B^{i,j}_{p,n}(z)$$

avec, en utilisant Fubini (somme rectangulaire) puis la propriété de branchement

$$B^{i,j}_{p,n}(z) = \mathbb{E}\left[\left(\sum_{\substack{u\in\mathbb{G}_n\\ zi\preccurlyeq u}} F_{p+1,n}(X(w) : zi \preccurlyeq w \preccurlyeq u)\right)\left(\sum_{\substack{v\in\mathbb{G}_n\\ zj\preccurlyeq v}} G_{p+1,n}(X(w), zj \preccurlyeq w \preccurlyeq v)\right) \Big|\mathcal{F}_{p+1}\right]$$
$$= \mathbb{E}\left[\sum_{\substack{u\in\mathbb{G}_n\\ zi\preccurlyeq u}} F_{p+1,n}(X(w), zi \preccurlyeq w \preccurlyeq u)\,\Big|X(zi)\right]\mathbb{E}\left[\sum_{\substack{v\in\mathbb{G}_n\\ zj\preccurlyeq v}} G_{p+1,n}(X(w), zj \preccurlyeq w \preccurlyeq v)\,\Big|X(zj)\right].$$

De plus, en utilisant la formule "tous pour un"

$$\mathbb{E}\left[\sum_{\substack{u\in\mathbb{G}_n\\ zi\preccurlyeq u}} F_{p+1,n}(X(w), zi \preccurlyeq w \preccurlyeq u)\,\Big|X(zi)\right] = m^{n-p-1}\mathbb{E}_{X(zi)}[F_{p+1,n}(Y_i, i \le n-p-1)]$$
$$= m^{n-p-1}Q(X(zi), F_{p+1,n})$$

On combine les identités obtenues et on somme sur les générations p pour obtenir

$$\mathbb{E}\left[\sum_{u\neq v\in\mathbb{G}_n} F_{0,n}(X(w) : w \preccurlyeq u)\, G_{0,n}(X(w) : w \preccurlyeq v)\right]$$
$$= \mathbb{E}\left[\sum_{p=0}^{n-1}\sum_{z\in\mathbb{G}_p, i\neq j\in\mathbb{N}} \mathbb{E}\left[\sum_{\substack{u,v\in\mathbb{G}_n\\ zi\preccurlyeq u, zj\preccurlyeq v}} F_{0,n}(X(w) : w \preccurlyeq u)\, G_{0,n}(X(w) : w \preccurlyeq v)\,\Big|\,\mathcal{F}_{p+1}\right]\right]$$
$$= \sum_{p=0}^{n-1}\mathbb{E}\left[\sum_{\substack{z\in\mathbb{G}_p, i\neq j,\\ zi, zj\in\mathbb{G}_{p+1}}} m^{2n-2p-2}F_{0,p}G_{0,p}(X(w), w \preccurlyeq z)\, Q(X(zi), F_{p+1,n})Q(X(zj), G_{p+1,n})\right]$$
$$= \sum_{p=0}^{n-1} m^{2n-2p-2}\mathbb{E}\left[\sum_{z\in\mathbb{G}_p} F_{0,p}G_{0,p}(X(w), w \preccurlyeq z)\, H_{p,n}(X(z))\right]$$

et on conclut avec une nouvelle utilisation de la formule tous-pour-un appliquée à la génération p et à la fonction $F(x_0, \dots, x_p) = F_{0,p}G_{0,p}(x_0, \dots, x_p)\, H_{p,n}(x_p)$. □

Exerice. Ecrire la fonction h_{n+p} en fonction de $(\Theta_x)_{x\in\mathcal{X}}$.

En prenant dans la formule pour les fourches $f_i = g_i = 1$ pour $i \le n-1$, de sorte à se focaliser sur les traits au temps n, on obtient :

Corollaire 38. *Pour toute fonction f mesurable bornée sur $\mathcal{X}$ et $n \in \mathbb{N}$,*

$$\mathbb{E}_{(\varnothing,x)}\left[\sum_{u\neq v\in\mathbb{G}_n} f(X(u))f(X(v))\right] = \sum_{p=0}^{n-1} m^{2n-p-2}\mathbb{E}_x[h_{n,p}(Y_p)]\,,$$

avec

$$h_{n,p}(y) = \mathbb{E}_{(\varnothing,y)}\left[\sum_{a\neq b\in\mathbb{G}_1} Q_{n-p-1}(X(a),f)\,Q_{n-p-1}(X(b),f)\right], \quad Q_n(x,f) = \mathbb{E}_x[f(Y_n)].$$

2.1.4 Loi des grands nombres

Dans cette partie, on suppose $m > 1$, de telle sorte à ce que la population survive avec probabilité positive.

Théorème 39. *On suppose $\mathbb{E}(L^2) < \infty$ et on considère f une fonction mesurable bornée. Supposons qu'il existe $\pi(f) \in \mathbb{R}$ tel que pour tout $x \in \mathcal{X}$,*

$$\mathbb{E}_x[f(Y_n)] \overset{n\to\infty}{\longrightarrow} \pi(f).$$

Alors

$$\frac{\mathbf{1}_{\mathbb{G}_n\neq\varnothing}}{\#\mathbb{G}_n}\sum_{u\in\mathbb{G}_n} f(X(u)) \overset{n\to\infty}{\longrightarrow} \mathbf{1}_{\forall n\geq 0,\,\mathbb{G}_n\neq\varnothing}\,\pi(f),$$

en probabilité et dans L^2.

Notons que f est fixé dans l'énoncé et l'hypothèse que l'on fait est de type ergodique : la quantité $\mathbb{E}_x[f(Y_n)]$ doit converger vers une valeur qui ne depend pas de x. La limite obtenue pour la mesure empirique dans la population est déterministe. Cela vient d'un effet de loi de grands nombres quand le temps évolue (on rappelle que la population totale tend vers l'infini). L'exemple $f = \mathbf{1}_A$ permet d'évaluer la proportion asymptotique d'individus de traits appartenant à A.
Le choix de la notation $\pi(f)$ pour la limite n'est pas anodin. En effet, $(Y_n)_n$ est une chaîne de Markov, et si on montre qu'elle converge en loi vers une probabilité stationnaire π, indépendante de sa condition initiale, alors pour tout fonction f continue bornée, l'hypothèse sera satisfaite. On obtient alors la convergence de la distribution des traits de la population vers une distribution déterministe π. En particulier, pour le modèle de graphe fini irréductible (métapopulation), si il y a en plus apériodicité, il y a alors convergence en loi de Y_n vers l'unique mesure stationnaire $\pi = (\pi_k)_{k\in\mathcal{X}}$ et

$$\mathbf{1}_{\mathbb{G}_n\neq\varnothing}\,\frac{\#\{u\in\mathbb{G}_n : X(u) = k\}}{\#\mathbb{G}_n} \overset{n\to\infty}{\longrightarrow} \mathbf{1}_{\forall n\geq 0,\,\mathbb{G}_n\neq\varnothing}\,\pi_k.$$

Démonstration. Étudions désormais

$$\Delta_n = \frac{1}{m^n}\sum_{u\in\mathbb{G}_n} g(X(u)),$$

avec $g(x) = f(x) - \pi(f)$, encore mesurable bornée, qui recentre f. En développant,

$$\mathbb{E}\big[\Delta_n^2\big] = \frac{1}{m^{2n}}\mathbb{E}\left[\sum_{u\in\mathbb{G}_n} g(X(u))^2 + \sum_{u\neq v\in\mathbb{G}_n} g(X(u))g(X(v))\right].$$

En utilisant que

$$\mathbb{E}\left[\sum_{u\in\mathbb{G}_n} g(X(u))^2\right] \leq \mathbb{E}(\#\mathbb{G}_n) \parallel g \parallel_\infty = m^n \parallel g \parallel_\infty$$

on obtient que le premier terme est écrasé par $1/m^{2n}$ quand $n \to \infty$. De plus, le corollaire précédent (fourches) assure

$$\mathbb{E}\big[\Delta_n^2\big] = o(1) + \sum_{p=0}^{n-1} m^{-p-2}\mathbb{E}_x[h_{n,p}(Y_p)]\,,$$

avec

$$h_{n,p}(y) = \mathbb{E}_{(\varnothing,y)}\left[\sum_{a\neq b\in\mathbb{G}_1} Q_{n-p-1}(X(a),g)\,Q_{n-p-1}(X(b),g)\right]$$

Notons que $\mathbb{E}[|h_{n,p}(Y_p)|] \leq \|g\|_\infty^2 \mathbb{E}\big[L^2\big] < \infty$ car $|Q_n(.,g)| \leq |g|$, donc

$$\sum_{p=K}^{n-1} m^{-p-2}\mathbb{E}[h_{n,p}(Y_p)] \leq \|g\|_\infty^2 \mathbb{E}\big[L^2\big]\, m^{-K}/(1-m)$$

tend vers 0 quand K tend vers l'infini (uniformément en $n \geq K$).
Enfin pour chaque $p \leq K$, pour tout $y \in \mathcal{X}$, $Q_{n-p-1}(y,g) = Q_{n-p-1}(y,f) - \int_{\mathcal{X}} f(x)\pi(dx) \to 0$ quand $n \to \infty$. En utilisant que $Q_.(.,g)$ est borné, on obtient que $\mathbb{E}_x[h_{n,p}(Y_p)]$ tend vers 0 quand n tend vers l'infini, par convergence dominée, pour chaque $p \leq K$.

Ceci conclut la preuve de la convergence de Δ_n vers 0 dans L^2. Et donc en probabilité. En ajoutant que $m^n/\#\mathbb{G}_n = 1/W_n$ tend vers une v.a. finie p.s. sur l'événement de survie, on obtient que

$$\mathbf{1}_{\mathbb{G}_n\neq\varnothing}\,\Delta_n\,\frac{m^n}{\#\mathbb{G}_n} = \mathbf{1}_{\mathbb{G}_n\neq\varnothing}\left(\frac{\sum\limits_{u\in\mathbb{G}_n} f(X(u))}{\#\mathbb{G}_n} - \int_{\mathcal{X}} f(x)\pi(dx)\right)$$

tend vers 0 en probabilité, ce qui est le résultat voulu. En effet, la convergence L^2 suit car la variable est bornée. □

Le résultat précédent donne une convergence en probabilité via une convergence L^2. On peut souvent obtenir une convergence p.s., soit par un argument de Borel Cantelli, soit en exploitant le calcul L^2 avec une vitesse suffisament rapide. En effet si $\sum_{n\geq 0}\mathbb{E}(X_n^2)^{1/2} < \infty$, alors par Cauchy Schwarz $\sum_{n\geq 0}\mathbb{E}(|X_n|) = \mathbb{E}(\sum_{n\geq 0}|X_n|) < \infty$ et $\sum_{n\geq 0}|X_n| < \infty$ p.s.. Ceci assure la convergence p.s. de X_n vers 0. Plus généralement la preuve précédente peut permettre de quantifier la vitesse de convergence de la distribution des traits dans la population, en utilisant la vitesse de convergence de la chaine de Markov Y_n vers sa probabilité stationnaire. Pour cette dernière question, on rappelle que de nombreuses techniques existent, notamment par des méthodes de type Doeblin-Lyapunov, la théorie spectrale et le couplage. On note enfin que cette méthode s'étend au cas d'environnement changeant.

2.2 Transitions de phase et division cellulaire

2.2.1 Processus autoregressif bifurquant

On rappelle que dans ce cas $L = 2$ p.s. (division binaire) et donc $\#\mathbb{G}_n = 2^n$. La transmission est affine avec pertubation gaussienne :

$$X_i^x = \alpha_i x + \mathcal{E}_i, \qquad i = 1, 2$$

avec $\alpha = (\alpha_1, \alpha_2) \in \mathbb{R}^2$ et $\mathcal{E} = (\mathcal{E}_1, \mathcal{E}_2)$ vecteur gaussien centré de matrice de covariance Γ. Pour simplifier suppons $Var(\mathcal{E}_1) = Var(\mathcal{E}_2) = \sigma^2$, soit

$$\mathcal{E}_1 \overset{loi}{=} \mathcal{E}_2 \overset{loi}{=} \mathcal{E}$$

loi gaussienne centrée de variance σ^2. Dans ce cas le noyau de la chaîne de Markov Y_n est

$$Q(x, dy) = Q^{(2)}(x, dy) = \frac{1}{2}\left(\mathbb{P}(\alpha_1 x + \mathcal{E} \in dy) + \mathbb{P}(\alpha_2 x + \mathcal{E} \in dy)\right)$$

Proposition 40. *La chaine $(Y_n)_n$ vérifie pour tout $n \geq 0$*

$$Y_{n+1} = \theta_{n+1} Y_n + \varepsilon_{n+1}$$

avec $(\theta_i)_{i\geq 1}$ et $(\varepsilon_i)_{i\geq 1}$ suites de v.a. i.i.d., indépendantes entre elles et indépendantes de Y_0 et

$$\mathbb{P}(\theta_i = \alpha_1) = \mathbb{P}(\theta_i = \alpha_2) = 1/2, \qquad \varepsilon_i \overset{loi}{=} \mathcal{N}(0, \sigma^2)$$

De plus

$$Y_n = \sum_{k=1}^{n} \varepsilon_k \prod_{j=k+1}^{n} \theta_j + Y_0 \prod_{j=1}^{n} \theta_j.$$

Considérons le cas $Y_0 = 0$ (le cas général se traitera pareil) et exploitons la stabilité de la loi gaussienne par sommation indépendante. Conditionnellement aux variables $(\theta_i)_{1\leq i\leq n}$ Y_n est distribué suivant une loi gaussienne centrée de variance

$$A(\theta_1, \ldots, \theta_n) = \sigma^2 \sum_{k=1}^{n} \prod_{j=k+1}^{n} \theta_j^2.$$

De plus, la magie de l'invariance en loi par retournement du temps dans une suite i.i.d. peut opérer

$$A(\theta_1, \ldots, \theta_n) \overset{loi}{=} A(\theta_n, \ldots, \theta_1) = \sigma^2 \sum_{k=0}^{n-1} \prod_{j=1}^{k} \theta_j^2.$$

Cela nous permet de récupérer une structure monotone :

$$A(\theta_n, \ldots, \theta_1) \uparrow_{n\to\infty} A_\infty = \sigma^2 \sum_{k=0}^{\infty} \prod_{j=1}^{k} \theta_j^2 = \sigma^2 \sum_{k=0}^{\infty} e^{S_k} \qquad \text{p.s.}$$

en introduisant la marche aléatoire $S_n = 2\sum_{j=1}^{n} \log(\theta_j)$. On obtient alors la classification suivante

Lemme 41. *(i) Si $\alpha_1\alpha_2 < 1$, alors $S_n \to -\infty$ linéairement p.s. et $A_\infty < \infty$. De plus Y_n converge en loi quand $n \to \infty$ vers un mélange de gaussienne π défini par*

$$\pi[a,b] = \int_{[a,b]\times\mathbb{R}_+} \frac{e^{-\frac{y^2}{2v}}}{\sqrt{2\pi v}}\,\mathrm{d}y\mathbb{P}(A_\infty \in dv).$$

(ii) Si $\alpha_1\alpha_2 \geq 1$, alors Y_n tend en probabilité quand $n \to \infty$ vers $+\infty$.

La loi des grands nombre obtenue dans la section précédente s'applique et on obtient

Théorème 42. *Si $\alpha_1\alpha_2 < 1$, alors pour tout $a < b$,*

$$\lim_{n\to\infty} \frac{\#\{u \in \mathbb{G}_n,\, X(u) \in [a,b]\}}{2^n} = \pi([a,b])$$

dans L^2. Sinon, pour tout $a \in \mathbb{R}_+$,

$$\lim_{n\to\infty} \frac{\#\{u \in \mathbb{G}_n,\, |X(u)| \leq a\}}{2^n} = 0$$

dans L^2.

2.2.2 Infection dans le modèle de Kimmel

Pour le modèle d'infection de Kimmel, rappelons que $L \in \{0,1,2\}$, $m = p_1 + 2p_2$, $\widehat{p}_1 = \frac{p_1}{m}$ et $\widehat{p}_2 = \frac{2p_2}{m}$. Le noyau de transition de la chaîne Y est donné par

$$Q(x,A) = \widehat{p}_1 Q^{(1)}(x,A) + \widehat{p}_2 Q^{(2)}(x,A)$$

avec $Q^{(1)}(x,A) = \mathbb{P}\Big(\sum_{i=1}^x P_i \in A\Big)$ et

$$Q^{(2)}(x,A) = \frac{1}{2}\left(\mathbb{P}\Big(\sum_{i=1}^x \sum_{j=1}^{P_i} E_{i,j} \in A\Big) + \mathbb{P}\Big(\sum_{i=1}^x \sum_{j=1}^{P_i} 1 - E_{i,j} \in A\Big)\right).$$

en rappelant que les P_i sont i.i.d. de loi μ, avec μ la loi de reproduction des pathogènes (sur une génération). Autrement dit

$$Q(x,A) = \frac{p_1}{m}\mathbb{P}\left(\sum_{i=1}^x P_i \in A\right) + \frac{p_2}{m}\mathbb{P}\left(\sum_{i=1}^x Q_i \in A\right) + \frac{p_2}{m}\mathbb{P}\left(\sum_{i=1}^x R_i \in A\right).$$

avec

$$Q_i = \sum_{j=1}^{P_i} E_{i,j}, \qquad R_i = \sum_{j=1}^{P_i} 1 - E_{i,j}.$$

On considère l'ensemble

$$E = \{1, q, 1-q\}$$

formé de 3 environnements (excluons $q \in \{0,1,1/2\}$ mais ce cas se traiterait pareil) et l'environnement aléatoire $\mathcal{E} \in E$ dont la loi est donnée par

$$\mathbb{P}(\mathcal{E} = 1) = \frac{p_1}{m}, \quad \mathbb{P}(\mathcal{E} = q) = \frac{p_2}{m} \quad \mathbb{P}(\mathcal{E} = 1-q) = \frac{p_2}{m}.$$

Les lois de reproduction associées sont données par p^e loi binomiale dont le premier paramètre est aléatoire distribué comme μ et le second vaut $e \in E$, i.e.

$$p_k^e = \sum_{n\geq 0} \mu_n \mathbb{P}(\mathrm{Bin}(n,e) = k) = \sum_{n\geq 0} \mu_n \binom{k}{n} e^k (1-e)^{n-k}.$$

Proposition 43. *La chaîne de Markov $(Y_n)_{n\geq 0}$ est un processus de Galton-Watson en environnement aléatoire $\mathcal{E}$ et loi de reproduction $(p^e)_{e\in E}$.*

Nous connaissons le comportement en temps long des processus de branchement en environnement aléatoire. On sait que si $\mathbb{E}(\log(m(\mathcal{E}))) \leq 0$ alors $(Y_n)_n$ atteint 0 p.s. en temps fini. Sinon, le processus survit avec probabilité positive. Or la formule tous-pour-un assure que le nombre moyen de cellules infectées par un nombre k de parasites est donné par la loi d'un Galton Watson en environnement aléatoire :

$$\mathbb{E}(\#\{u \in \mathbb{G}_n : X(u) = k\}) = \mathbb{E}(\mathbb{G}_n)^n \mathbb{P}(Y_n = k).$$

La loi des grands nombres sur les abres de Galton Watson (Theorème 39) nous assure alors la transition de phase suivante. Elle fait intervenir les paramètres de la division cellulaire (nombre moyen de cellules m et probabilité de division p_2), le paramètre de croissance de l'infection (via la moyenne $\overline{\mu}$ de la loi de reproduction des pathogènes) et le paramètre de Bernoulli q qui détermine la répartition.

Théorème 44. *Si*

$$\log(\overline{\mu}) \leq \frac{p_2}{m} \log\left(\frac{1}{q(1-q)}\right),$$

alors

$$\lim_{n\to\infty} \mathbf{1}_{\mathbb{G}_n \neq \varnothing} \frac{\#\{u \in \mathbb{G}_n : X(u) > 0\}}{\#\mathbb{G}_n} = 0$$

en probabilité. Sinon, pour tout $K > 0$,

$$\varliminf_{n\to\infty} \mathbb{E}\left(1_{\mathbb{G}_n \neq \varnothing} \frac{\#\{u \in \mathbb{G}_n : X(u) > K\}}{\#\mathbb{G}_n}\right) > 0,$$

$$\lim_{n\to\infty} \mathbf{1}_{\mathbb{G}_n \neq \varnothing} \frac{\#\{u \in \mathbb{G}_n : 0 < X(u) < K\}}{\#\mathbb{G}_n} = 0.$$

Dans le premier cas (sous-critique), la proportion de cellules infectées tend vers 0 : la division régule l'infection. Dans le deuxième (surcritique), une proportion positive de cellules devient très infectée lorsque l'infection ne s'éteint pas globalement. A nouveau, on peut aller plus loin et obtenir des limites p.s. On peut aussi obtenir la croissance exponentielle de l'infection au sein de chaque cellule infectée dans le deuxième régime en utilisant la croissance de Y_n.

Démonstration. On calcule

$$\mathbb{E}(\log(m(\mathcal{E}))) = \frac{p_1}{m}\log(m(1)) + \frac{p_2}{m}\log(m(q)) + \frac{p_2}{m}\log(m(1-q)).$$

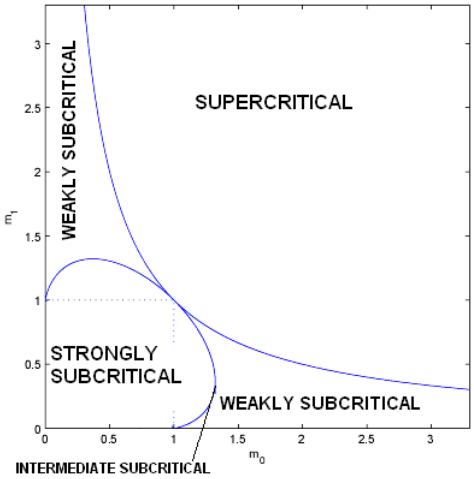


FIGURE 2.1 – Représentation des différents régimes pour le nombre de cellules infectées, dans le cas $p_2 = 1$, avec en abscisse $m_0 = q\bar{\mu}$ et en ordonée $m_1 = (1-q)\bar{\mu}$.

avec $m(e) = \mathbb{E}(\mathrm{Bin}(P, e)) = \overline{\mu}e$. On utilise alors le premier chapitre qui donne le comportement en temps long d'un processus de Galton Watson en environnement aléatoire en fonction du signe de $\mathbb{E}(\log(m(\mathcal{E})))$. Si cette quantité est négative alors $\mathbb{P}_k(Y_n > 0)$ tend vers 0 quand $n \to \infty$, et ce pour tout $k \geq 0$. C'est à dire que $\mathbb{E}_k(f(Y_n)) \to 0$ quand $n \to \infty$ pour tout $k \geq 0$ avec $f(x) = \mathbf{1}_{x>0}$. On peut appliquer la loi des grands nombres (Theorème 39) pour obtenir la première partie du théorème. Notons qu'ici la la distribution se concentre en 0, et on pourrait conclure directement à partir de la formule tous pour un.

Pour l'autre cas, on utilise que Y_n tend vers l'infini sur l'événement de survie et donc pour tout $k \geq 0$, $\mathbb{P}(0 < Y_n < K)$ tend vers 0 quand n tend vers l'infini. On conclut avec la loi des grands nombres en prenant ici $f(x) = \mathbf{1}_{x\in(0,K)}$ □

D'autres transitions de phases sont intéressantes peuvent être observées dans ce modèle, en exploitant le comportement asymptotique des processus de branchement en environnement aléatoire. En particulier, à propos du nombre de cellules infectées, la formule tous pour un donne

$$\mathbb{E}\left(\#\{u \in \mathbb{G}_n : X(u) > 0\}\right) = m^n \mathbb{P}(Y_n > 0)$$

Si dans le premier régime, $\mathbb{P}(Y_n > 0)$ tend vers 0 et que la proportion de cellules infectées est négligeable devant la population totale de cellules, le nombre de cellules infectées moyen $m^n\mathbb{P}(Y_n > 0)$ croit en général exponentiellemment. Cette croissance dépend de la vitesse d'extinction de Y_n. Or cette vitesse $\mathbb{P}(Y_n > 0)$ dépend du signe de

$$\kappa = \mathbb{E}(m(\mathcal{E})\log(m(\mathcal{E}))),$$

comme on l'a vu au premier chapitre. Si $\kappa < 0$ (régime fortement sous-critique), on obtient un nombre de cellules infectées du même ordre que le nombre de pathogènes. Si $\kappa > 0$ (régime faiblement sous-critique), il y a bien moins de cellules infectées que de pathogènes, car les pathogènes s'accumulent dans des mêmes cellules.

2.2.3 Contamination et division cellulaire

On se propose d'étendre le modèle de Kimmel binaire symétrique pour l'infection en prenant en compte une contamination extérieure de pathogènes. C'est à dire que :

— les celules se divisent en deux à chaque étape et la génération n est constituée de 2^n cellules labellisées par $\mathbb{G}_n = \{1,2\}^n$;

— les pathogènes se reproduisent suivant la loi de reproduction μ et se répartitissent suivant une loi binomiale de paramètre $q = 1/2$ au moment de la division ;

— en plus de la prolifération du pathogène au sein de la cellule chacune des deux cellules filles hérite à la génération suivante d'un nombre aléatoire de pathogènes, distribué comme la variable I, indépendamment pour chaque cellule et à chaque génération.

Le mécanisme de transmission du trait (ici un nombre de pathogènes) devient donc

$$\Theta_x = \left(\sum_{i=1}^{x} \sum_{j=1}^{P_i} E_{i,j} \, + \, I_1, \sum_{i=1}^{x} \sum_{j=1}^{P_i} 1 - E_{i,j} \, + \, I_2 \right)$$

avec $(P_i)_{i\geq 1}$ v.a. i.i.d. de loi μ et $E_{i,j}$ v.a. i.i.d. de loi de Bernoulli de paramètre $1/2$ et I_1, I_2 i.i.d. distibuées comme I, toutes ces v.a. étant par ailleurs indépendantes entre elles.

La formule tous pour un donne

$$\mathbb{E}_{(\varnothing,x)} \left[\sum_{u \in \mathbb{G}_n} F(X(v),\, v \preccurlyeq u) \right] = 2^n \mathbb{E}_x[F(Y_0, \ldots, Y_n)].$$

avec Y un processus de Galton Watson de loi de reproduction $\mu^{1/2}$ définie par

$$\mu_k^{1/2} = \sum_{i \geq k} \mu_i \mathbb{P}(Bin(i, 1/2) = k)$$

et de loi d'immigration donnée par I. Le comportement asymptotique de Y dépend du fait que

$$\overline{\mu^{1/2}} = \overline{\mu}/2$$

est plus grande ou plus petite que 1. En effet, la section 1.3 guarantit que si $\mathbb{E}(I) < \infty$ et $\overline{\mu} < 2$ alors $(Y_n)_n$ converge en loi quand n tend vers l'infini vers une v.a. intégrable, indépendante de la condition initiale. La loi des grands nombres (Théorème 44) guarantit alors que la proportion de cellules infectées par un nombre donné de parasites converge en probabilité (on peut montrer aussi une convergence p.s.) vers la loi limite de Y.On obtient ainsi un nouveau régime ou les cellules ne se débarassent pas de l'infection (à cause de la contamination) mais l'infection reste modérée.

Si $\overline{\mu} > 2$, l'infection prend et la proportion de cellules dont l'infection est en dessous de K tend vers 0 en probabilité, et ce pour tout $K \geq 0$.

Notez que deux régimes apparaissent dans le cas critique $\overline{\mu} = 2$. Pour compléter ces résultats, notamment au cas asymétrique et au cas où la contamination dépend du fait que la cellule est déjà infectée ou non, on pourra consulter [5].

2.3 Vitesse d'invasion dans un modèle spatial

On considère dans cette partie l'exemple de la marche aléatoire branchante, dans le cas où la population survit avec probabilité positive :

$$m = \mathbb{E}[L] \in (1, \infty).$$

On considère la variable de déplacement D à valeurs dans $\mathbb{R}$, qui quantifie à quel point chaque descendant s'éloigne de son parent. On recentre le processus de sorte à se ramener au cas où

$$\mathbb{E}[D] = 0,$$

et on suppose que

$$\mathrm{Var}(D) = \sigma^2 \in (0, \infty).$$

Pour simplifier, on suppose que les enfants atterissent à des distances i.i.d. du parent et on a alors la loi de transmission

$$\Theta^x = (x + D_i)_{1 \leq i \leq L},$$

avec D_i v.a. i.i.d. et $D_i \overset{loi}{=} D$. De plus on suppose que $(D_i)_{i \geq 1}$ est indépendant de la variable L, c'est-à-dire que les déplacements ne dépendent as du nombre d'enfants. La formule tous pour un permet d'éclairer ce qui se passe en moyenne pour la mesure empirique :

$$\mathbb{E}_{(\varnothing,0)}(\#\{u \in \mathbb{G}_n : X(u) \in A\}) = m^n \mathbb{P}(Y_n \in A)$$

où $(Y_n)_{n \geq 0}$ est la marche aléatoire définie par $Y_0 = 0$ et

$$Y_n = \sum_{i=1}^{n} D_i.$$

Le théorème central limite assure que

$$\mathbb{P}(Y_n \in [\sqrt{n}a, \sqrt{n}b]) \overset{n\to\infty}{\longrightarrow} \mathbb{P}\big(\mathcal{N}\big(0, \sigma^2\big) \in [a, b]\big)$$

Commençons par un premier résultat qui est une conséquence des résultats que nous avons obtenus. Il donne la distribution des traits dans la population. En adaptant un chouia la preuve la loi des grands nombres obtenues dans la partie précédente pour permettre une renormalisation en $\sqrt{n}$ du trait des individus, nous obtenons

Proposition 45.

$$\frac{1}{\#\mathbb{G}_n} \#\big\{u \in \mathbb{G}_n,\, a\sqrt{n} \leq X(u) \leq b\sqrt{n}\big\} \underset{n\to\infty}{\overset{L^2}{\longrightarrow}} \mathbb{P}\big(\mathcal{N}\big(0, \sigma^2\big) \in [a, b]\big)\,.$$

On s'intéresse maintenant au front d'invasion, c'est à dire à l'individu le plus en avant

$$R_n = \max\{X(u) : u \in \mathbb{G}_n\}$$

Notons que la formule tous-pour-un donne pour tourt $v \geq 0$,

$$N_n(v) := \mathbb{E}[\#\{u \in \mathbb{G}_n,\, X(u) \geq nv\}] = m^n \mathbb{P}(Y_n \geq nv).$$

En invoquant des résultats de grandes déviations qu'on va détailler ci-après, nous allons pouvoir évaluer la probabilité de droite :

$$\mathbb{P}(Y_n \geq nv) \approx e^{-\psi(v)n},$$

avec ψ une fonction croissante, et $\approx$ signifera que les quantités sont équivalentes à l'échelle logarithmique. Ainsi si $\psi(v) > \log(m)$, alors $N_n(v) \ll 1$. Tandis que si $\psi(v) < \log(m)$, alors $N_n(v) \gg 1$. Un seuil avec une vitesse critique apparait, la vitesse $v^\star$ solution de

$$\psi(v^\star) = \log(m).$$

C'est elle qui va nous donner la vitesse du front et un équivalent pour R_n quand n tend vers l'infini.

2.3.1 Préliminaires de grandes déviations

Nous allons utiliser quelques éléments de grandes déviations que nous donnons ici. Pour une référence sur le sujet, le lecteur pourra consulter le livre de Dembo Zeitouni [11]. Nous supposons ici que la fonction

$$\Lambda(\theta) = \log\Big(\mathbb{E}\Big[e^{\theta D}\Big]\Big)$$

est bien définie pour tout $\theta \in \mathbb{R}$, c'est à dire que $\mathbb{E}\big[e^{\theta D}\big] < \infty$. Cette hypothèse est par exemple vérifiée pour une v.a. D bornée. On peut en partie relâcher cette hypothèse mais il faudrait alors faire attention à l'ensemble de définition qui peut intervenir pour la description du front.
La fonction Λ est *positive* car $e^x \geq 1 + x$ et elle est dérivable (autant de fois qu'on veut) sur $\mathbb{R}$ et

$$\Lambda''(\theta) = \frac{\mathbb{E}\big[D^2 e^{\theta D}\big]\mathbb{E}\big[e^{\theta D}\big] - \mathbb{E}\big[De^{\theta D}\big]^2}{\mathbb{E}[e^{\theta D}]^2} > 0,$$

par l'inégalité de Cauchy Schwarz (le cas d'égalité est exclu car D a une variance non nulle). Ainsi

$$\Lambda'(\theta) = \frac{\mathbb{E}\big[De^{\theta D}\big]}{\mathbb{E}[e^{\theta D}]}$$

est strictement croissante et elle admet une limite

$$\gamma = \lim_{\theta\to\infty} \Lambda'(\theta) \in \mathbb{R}_+^* \cup \{\infty\}$$

Définissons la transformée de Fenchel-Legendre pour $x \geq 0$:

$$\psi(x) = \sup\{\theta x - \Lambda(\theta),\, \theta \in \mathbb{R}\} = \sup\{\theta x - \Lambda(\theta),\, \theta \in \mathbb{R}_+\}$$

car pour $\theta \leq 0$, $\Lambda(\theta) \geq 0$ et $\theta x - \Lambda(\theta) \leq 0$ et $\Lambda(0) = 0$. La fonction ψ est convexe croissante sur $\mathbb{R}^+$, car c'est le supremum de telles fonctions. Notons que

$$\frac{\partial}{\partial\theta}[\theta x - \Lambda(\theta)] = x - \Lambda'(\theta) = 0 \qquad \text{ssi} \quad x = \Lambda'(\theta).$$

Lemme 46. *i) Si $x \in [0, \gamma)$, alors il existe un unique $\theta(x) \geq 0$ tel que $\Lambda'(\theta(x)) = x$. De plus $\psi(x) = \theta(x)x - \Lambda(\theta(x))$.*
De plus ψ est finie et strictement croissante sur $[0, \gamma)$.
ii) Si $x > \gamma$, alors $\psi(x) = +\infty$

Par exemple, lorsque D est équiprobable sur $\{-1, 1\}$: $\mathbb{P}(D = 1) = \mathbb{P}(D = -1) = 1/2$, alors $\Lambda(\theta) = \log\Big(\frac{e^\theta + e^{-\theta}}{2}\Big)$, et $\theta(x) = \frac{1}{2}\log\Big(\frac{1+x}{1-x}\Big)$. On a alors

$$\psi(x) = \frac{1}{2}((1 + x)\ln(1 + x) + (1 - x)\log(1 - x))$$

lorsque $x \in [0, 1]$. On notera que $\psi(1) = \log(2)$, et $\psi(x) = \infty$ si $x > 1$.

Théorème 47. *Pour tout $x \in [0, \gamma)$, alors*

$$\lim_{n\to\infty} \frac{1}{n}\log(\mathbb{P}(Y_n \geq nx)) = -\psi(x).$$

Démonstration. Par inégalité de Markov, pout tout $\theta \geq 0$,

$$\mathbb{P}(Y_n \geq nx) \leq \mathbb{E}\Big[e^{\theta Y_n - \theta n x}\Big] = \mathbb{E}\Big[e^{\theta D - \theta x}\Big]^n = e^{-n(\theta x - \Lambda(\theta))},$$

et en optimisant selon θ, on obtient

$$\mathbb{P}(Y_n \geq nx) \leq e^{-n\psi(x)}.$$

Pour la borne inférieure, considérons l'unique θ tel que

$$\Lambda'(\theta) = x.$$

On introduit la variable aléatoire $\widetilde{D}$ dont la loi est donnée par :

$$\mathbb{P}(\widetilde{D} \in dy) = \frac{e^{\theta y}}{\mathbb{E}[e^{\theta D}]}\mathbb{P}(D \in dy).$$

Ce choix permet d'avoir la dérive voulue :

$$\mathbb{E}\Big[\widetilde{D}\Big] = \frac{\mathbb{E}\big[De^{\theta D}\big]}{e^{\theta D}} = \Lambda'(\theta) = x\,.$$

De plus, en prenant $(\widetilde{D}_j)_{j\geq 1}$ des v.a. i.i.d. on remarque que

$$\mathbb{P}(D_1 \in dy_1, \dots D_n \in dy_n) = \prod_{j=1}^n \mathbb{P}(D_j \in dy_j) = \mathbb{E}\Big[e^{\theta D}\Big]^n e^{-\theta \sum_{j=1}^n y_j} \prod_{j=1}^n \mathbb{P}(\widetilde{D}_j \in dy_j)$$

Alors

$$\begin{aligned}
\mathbb{P}(nx \leq Y_n \leq n(x+\delta)) &= \int_{\mathbb{R}^n} \mathbb{1}_{nx \leq \sum_{j=1}^n y_j \leq n(x+\delta)}\, \mathbb{P}(D_1 \in dy_1, \dots, D_n \in dy_n) \\
&= \mathbb{E}\Big[e^{\theta D}\Big]^n \int_{\mathbb{R}^n} \mathbb{1}_{nx \leq \sum\limits_{j=1}^n y_j \leq n(x+\delta)} e^{-\theta \sum\limits_{j=1}^n y_j}\, \mathbb{P}(\widetilde{D}_1 \in dy_1, \dots, \widetilde{D}_n \in dy_n) \\
&= \mathbb{E}\Big[e^{\theta D}\Big]^n \mathbb{E}\Big[e^{-\theta \widetilde{Y}_n} \mathbb{1}_{nx \leq \widetilde{Y}_n \leq n(x+\delta)}\Big] \\
&\qquad\qquad \geq e^{n\Lambda(\theta) - \theta n(x+\delta)}\, \mathbb{E}\Big[\mathbb{1}_{nx \leq \widetilde{Y}_n \leq n(x+\delta)}\Big],
\end{aligned}$$

avec $\widetilde{Y}$ une marche aléatoire de pas $\widetilde{D}$ cette fois-ci, c'est à dire de pas moyen x. On obtient

$$\frac{1}{n}\log\left(\mathbb{P}(nx \leq Y_n \leq n(x+\delta))\right) \geq -\psi(x) - \delta\theta(x) + \frac{1}{n}\log\Big(\mathbb{P}\Big(x \leq \widetilde{Y}_n/n \leq x+\delta\Big)\Big).$$

Or la probabilité de droite converge vers $1/2$ en vertu du thèorème central limite. En faisant tendre δ vers 0 à droite on obtient la minoration voulue. □

2.3.2 Vitesse d'invasion

On rappelle que

$$R_n = \max\{X(u),\, u \in \mathbb{G}_n\}$$

est la position du *front d'invasion* au temps n. Pour simplifier la discussion, on se place dans le cas où il existe $v^\star \in (0, \gamma)$ tel que

$$\psi(v^\star) = \log(m).$$

Commençons par la borne supérieure.

Proposition 48. $\overline{\lim}_{n\to\infty} \frac{R_n}{n} \leq v^\star$ *presque surement*

Démonstration. Soit $v \in (v^\star, \gamma)$, de sorte que $\psi(v) > \psi(v^\star) = \log(m)$ car ψ est strictement croissante. On a donc :

$$\begin{aligned} \mathbb{P}(R_n \geq vn) &= \mathbb{P}(\#\{u \in \mathbb{G}_n : X(u) \geq vn\} \geq 1) \\ &\leq \mathbb{E}[\#\{u \in \mathbb{G}_n : X(u) \geq vn\}] = m^n \mathbb{P}(Y_n \geq vn) \end{aligned}$$

en utilisant la formule "tous-pour-un". Le théorème 47 assure que $\mathbb{P}(Y_n \geq vn)$ decroit exponentiellement au taux $\psi(v) > \log(m)$. On obtient que $\mathbb{P}(R_n \geq vn)$ décroît exponentiellement vite vers 0, car $\log(m) - \psi(v) < 0$ et donc

$$\sum_{n\in\mathbb{N}} \mathbb{P}(R_n \geq nv) < \infty.$$

Ainsi, par le lemme de Borel-Cantelli, presque-sûrement, à partir d'un rang, on a $R_n < nv$, d'où $\overline{\lim}_{n\to\infty} \frac{R_n}{n} \leq v$ p.s. En faisant tendre v par $v^\star$, on a le résultat annoncé. □

Il faut maintenant montrer que cette vitesse $v^\star$ est bien réalisée par quelqu'un, sur l'événement où la population survit : $\mathcal{S} = \{\forall n \in \mathbb{N},\ \#\mathbb{G}_n > 0\}$;

Proposition 49.

$$\underline{\lim}_{n\to\infty} \mathbf{1}_{\mathcal{S}} \frac{R_n}{n} \geq \mathbf{1}_{\mathcal{S}}\, v^\star \qquad p.s.$$

Nous allons donner deux preuves de ce résultat, chacune éclairante et généralisable à d'autres modèles, la première est due à Biggins.

Preuve de la proposition 49 : un processus de Galton-Watson au front. Considérons cette fois $v < v^\star$ et posons

$$Z^{(p)} = \#\{u \in \mathbb{G}_p,\ X(u) \geq pv\}.$$

le nombre d'individus ayant dévié d'au moins pv depuis l'instant initial. La formule tous-pour-un et le théorème de grandes déviations garantissent cette fois

$$\mu_p = \mathbb{E}\Big[Z^{(p)}\Big] = m^p \mathbb{P}(Y_p \geq pv) \approx \exp(p(\log(m) - \psi(v)))$$

qui tend vers l'infini quand $p \to \infty$ car $\log(m) > \psi(v)$. En particulier, on peut se fixer un rang p tel que $\mu_p > 1$.

Construisons la sous population qui va reproduire cette déviation toutes les p générations. Concrètement $\mathbb{G}_0^{(p)} = \{\varnothing\}$, $\mathbb{G}_1^{(p)} = \{u \in \mathbb{G}_p,\ X(u) \geq pv\}$ et plus généralement pour $k \geq 0$

$$\mathbb{G}_{k+1}^{(p)} = \Big\{u \in \mathbb{G}_{(k+1)p} : \exists w \in \mathbb{G}_k^{(p)},\ u \succcurlyeq w \ \text{ et } \ X(u) \geq X(w) + pv\Big\}.$$

Par construction, tout $u \in \mathbb{G}_k^{(p)}$ vérifie $X(u) \geq kpv$. De plus le processus

$$\mathcal{Z}_k = \#\mathbb{G}_k^{(p)}$$

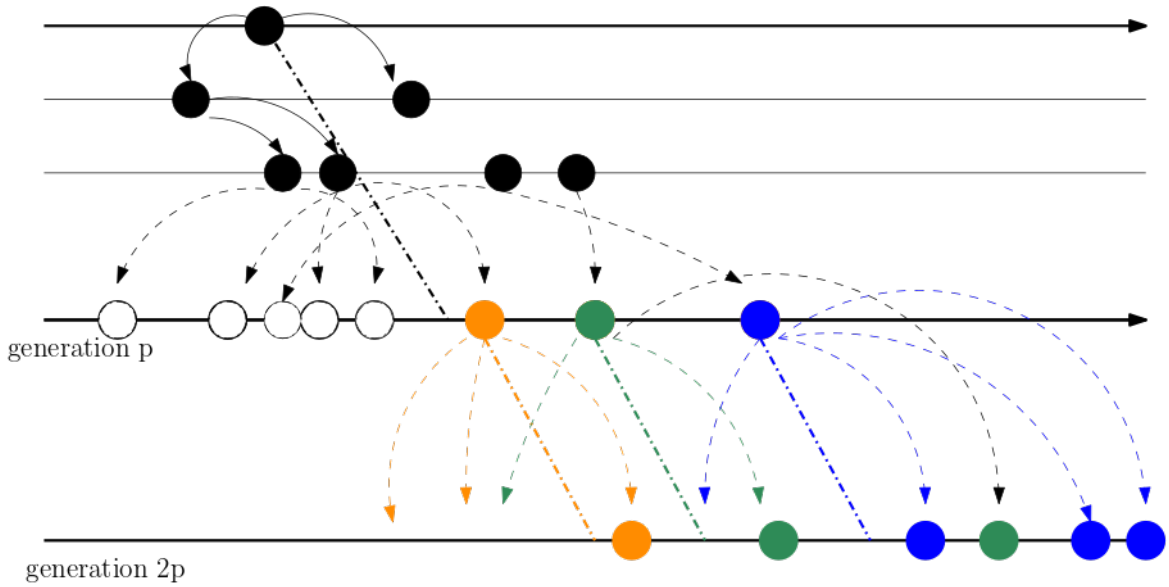


FIGURE 2.2 – Subpopulations (colored balls) created by keeping individuals which have moved to the right enough with respect to the positions of their ancestors p generations before ; the other individuals are forgotten.

est un processus de Galton Watson de loi de reproduction $Z^{(p)}$. Comme $\mu_p > 1$, le processus survit avec probabilité positive et sur cet événement de survie

$$R_{kp} \geq v.kp$$

On obtient pour ces temps $n = kp$ et sur cet événement une vitesse d'au moins v. Il reste à vérifier que ce qui se passe entre des temps kp et $(k+1)p$ ne fait pas diminuer cette vitesse, et de mener le raisonnement sur l'événement $\mathcal{S}$, plus gros a priori que l'événement de survie de $\mathcal{Z}$. Pour le premier point, on peut mener un raisonnement analogue en translatant juste les temps et en considérant les générations $i + kp$ avec i fixé et $k \geq 0$. Pour le second point, on démarre plusieurs constructions indépendantes de sorte à réduire le risque d'extinction. Pour cela on utilise le fait que sur l'événement $\mathcal{S}$ la population totale $\#\mathbb{G}_n$ tend vers l'infini et on attend une génération où ce nombre est assez grand. □

Nous allons maintenant donner les grandes lignes d'une autre preuve inspirée des travaux de Lyons et du monographe de Zhan Shi sur les marches branchantes. Cette preuve va exploiter la formule tous pour un en incluant un poids qui permet de se focaliser sur les individus en avant du front dans cette formule. Remarquons que la fonction exponentielle $x \to \exp(\rho.)$ est un vecteur propre du semigroupe du premier moment, c-a-d :

$$M(x, \exp(\rho.)) = \lambda(\rho) \exp(\rho x)$$

pour $x \in \mathbb{R}$, avec

$$\lambda(\rho) = \mathbb{E}\left(\sum_{i=1}^{L} e^{\rho D_i}\right) = m\mathbb{E}\left(e^{\rho D}\right).$$

On pondère alors dans la formule "tous pour un" par ce vecteur propre et on obtient le résultat suivant, dont la première partie est une conséquence directe du corollaire 36.

Lemme 50. *Pour toute fonction F mesurable positive et $\rho > 0$ et $x \in \mathbb{R}$,*

$$\mathbb{E}_{(\varnothing,x)}\left(\sum_{u\in\mathbb{G}_n} e^{\rho X(u)} F(X(v) : v \leq u)\right) = e^{\rho x}\lambda(\rho)^n \mathbb{E}_x\left(F(Y_i^{(\rho)} : 0 \leq i \leq n)\right)$$

avec $Y^{(\rho)}$ marche aléatoire de pas $D^{(\rho)}$, dont la loi est définie par

$$\mathbb{P}(D^{(\rho)} \in dx) = \frac{e^{\rho x}}{\mathbb{E}(e^{\rho D})}\mathbb{P}(D \in dx).$$

De plus

$$W_n^{(\rho)} = \sum_{u\in\mathbb{G}_n} e^{\rho X(u) - n\log\lambda(\rho)}$$

est une martingale positive convergente vers une variable aléatoire $W^{(\rho)} \in [0,\infty)$ p.s.

Démonstration. Pour la deuxième partie,

$$\begin{aligned}\mathbb{E}\left(W_{n+1}^{(\rho)}|\mathcal{F}_n\right) &= \sum_{u\in\mathbb{G}_n} \mathbb{E}\left(\sum_{v\in\mathbb{G}_{n+1}:v\geq u} e^{\rho X(v)-(n+1)\log\lambda(\rho)}|X(u)\right)\\ &= \sum_{u\in\mathbb{G}_n} e^{\rho X(u)-(n+1)\log\lambda(\rho)}\mathbb{E}\left(\sum_{i=1}^{L} e^{\rho D_i}\right) = W_n^{(\rho)}\end{aligned}$$

car $X(v) = X(ui) = X(u) + D_i(u)$ pour $i \leq L(u)$. □

On admet maintenant le critère $L\log L$ suivant de non dégénérescence de la martingale limite. La preuve se fait avec une construction spinale comme dans le cours et on renvoie par exemple à la section 4.8. du polycopié de Zahn Shi [2].

Lemme 51. *Les trois assertions suivantes sont équivalentes :*
i) $\mathbb{E}(W^{(\rho)}) = 1$.
ii) $W^{(\rho)} > 0$ *sur l'événement de survie* $\mathcal{S} = \{\forall n \in \mathbb{N}, \#\mathbb{G}_n > 0\}$.
iii) $\mathbb{E}(W_1^{(\rho)}\log_+ W_1^{(\rho)}) < \infty$ *et* $\rho\lambda'(\rho) < \lambda(\rho)\log\lambda(\rho)$.

Preuve de la proposition 49 sous la condition $\mathbb{E}(W_1^{(\rho)}\log_+ W_1^{(\rho)}) < \infty$ pour tout $\rho > 0$. On pose $\rho = \theta(v)$ avec $\psi(v) < \log(m)$, de sorte que $v = \lambda'(\rho)/\lambda(\rho)$ et on choisit $\varepsilon > 0$ tel que

$$\rho(v+\varepsilon) \leq \log(\lambda(\rho))$$

ce qui est possible car $\rho v - \log\lambda(\rho) = \theta(v)v - \log\mathbb{E}(e^{\rho D}) - \log(m) = \psi(v) - \log(m) < 0$. Donc $\rho\lambda'(\rho) = \rho v\lambda(\rho) < \lambda(\rho)\log\lambda(\rho)$ et le lemme précédent assure que $W^{(\rho)} > 0$ sur l'événement $\mathcal{S}$. De plus

$$W_n^{(\rho)} = \varphi_n + \psi_n$$

avec

$$\begin{aligned}\varphi_n &= \sum_{u\in\mathbb{G}_n, |X(u)-nv|\leq n\varepsilon} e^{\rho X(u)-n\log\lambda(\rho)}\\ \psi_n &= \sum_{u\in\mathbb{G}_n, |X(u)-nv|>n\varepsilon} e^{\rho X(u)-n\log\lambda(\rho)}\end{aligned}$$

2. https ://www.lpsm.paris/pageperso/zhan/pdffile/brw.pdf

Or $\rho X(u) - n \log \lambda(\rho) \leq 0$ quand $|X(u) - nv| \leq n\varepsilon$ donc

$$\varphi_n \leq \#\{u \in \mathbb{G}_n : X(u) \geq n(v - \varepsilon)\}.$$

De plus

$$\mathbb{E}(\psi_n) = e^{\rho x} \mathbb{P}(|Y_n^{(\rho)} - nv| > n\varepsilon)$$

grâce à la formule "tous pour un" du premier du lemme avec $G(x_i, i \leq n) = e^{\rho x_n} 1_{|x_n - nv| > \varepsilon n}$. Comme $Y^{(\rho)}$ est une marche aléatoire de pas moyen

$$\mathbb{E}(D^{(\rho)}) = \lambda'(\rho)/\lambda(\rho) = v,$$

le théorème de grandes déviations assure que $\mathbb{P}(|Y_n^{(\rho)} - nv| > n\varepsilon)$ decroît exponentiellement vite et donc $\sum_n \mathbb{E}(\psi_n) = \mathbb{E}(\sum_n \psi_n) < \infty$, ce qui prouve que $\sum_n \psi_n < \infty$ et ψ_n tend vers 0 presque surement. On obtient

$$\varliminf_{n\to\infty} \#\{u \in \mathbb{G}_n : X(u) \geq n(v - \varepsilon)\} \geq \varliminf_{n\to\infty} \varphi_n \geq \varliminf_{n\to\infty} \varphi_n + \psi_n = \varliminf_{n\to\infty} W_n^{(\rho)} > 0$$

sur $\mathcal{S}$, ce qui achève la preuve. $\square$

Chapitre 3

Chaîne de Markov branchante pour des populations structurées

Nous souhaitons maintenons prendre en compte le fait que le trait des individus affecte leur loi de reproduction. Nous perdons donc l'hypothse de neutralité faite dans la partie précédente. En particulier la taille totale de la population ne suit plus un processus de Galton-Watson. On obtient ici ce qu'on appelle un processus de Galton Watson multitype, avec un ensemble qui peut etre infini et non discret. On parle aussi de chaine de Markov branchante, pour souligner que la transmission des traits va être un processus de Markov.

3.1 Définition du modèle et exemple

De façon informelle, la variable de reproduction L dépdend maintenant de $x \in \mathcal{X}$. Chaque individu de trait x va indépendamment donner naissance à un nombre aléatoire L^x d'individus, de traits $\Theta_x = (X_i^x,\, i = 1, \dots, L^x)$.

Plus formellement, la transmission du trait aux enfants est modélisée par la famille de processus ponctuel $(\Theta_x)_{x\in\mathcal{X}}$, avec pour tout $x \in \mathcal{X}$, $\Theta_x = (X_i^x : i = 1, \dots, L^x)$ v.a. à valeurs dans $\mathbb{X}$ et $\Theta : \mathcal{X} \times \Omega \to \mathbb{X}$ mesurable où $\mathcal{X} \times \Omega$ est muni de la tribu produit. Introduisons maintenant un processus pour chaque label : $\big((\Theta_x(u))_{x\in\mathcal{X}},\, u \in \mathcal{U}\big)$ famille de v.a. indépendantes et identiquement distribuées comme $(\Theta_x)_{x\in\mathcal{X}}$. On note $\Theta_x(u) = (X_i^x(u),\, i = 1, \dots, L^x(u))$.

Définition 52. *La chaîne $X = \{(u, X(u)) : u \in \mathcal{T}\}$ issue d'un individu de trait x_0 et de transitions $\big((\Theta_x(u))_{x\in\mathcal{X}} : u \in \mathcal{U}\big)$ est définie récursivement par $\mathcal{T} = \cup_{n\in\mathbb{N}}\mathbb{G}_n$ et*

$$\mathbb{G}_0 = \{\varnothing\}, \qquad X(\varnothing) = x_0$$

et pour tout $n \geq 0$,

$$\mathbb{G}_{n+1} = \left\{uk : u \in \mathbb{G}_n,\, k \in \left[\!\left[1, L^{X(u)}(u)\right]\!\right]\right\}, \qquad X(uk) = X_k^{X(u)}(u)$$

pour tous $u \in \mathbb{G}_n$ et $k \in \left[\!\left[1, L^{X(u)}(u)\right]\!\right]$.

A nouveau, la loi de X est déterminée par la loi de $(\Theta_x)_{x\in\mathcal{X}}$ et x_0. Notons que cette définition s'étend immédiatement à une condition initiale avec plusieurs individus et des traits aléatoires. Remarquons aussi que la taille $Z_n = \#\mathbb{G}_n$ de la population à la

génération n n'est plus un processus de Galton-Watson.

De nombreux modèles en biologie et écologie font appel à cette extension qui relâche la neutralité. Chacun des quatre exemples du chapitre précédent illustre cet aspect. Pour les modèles spatiaux, la loi de reproduction peut dépendre de la localisation à travers la natalité ou la mortalité sensible aux quantités de ressources disponibles, à l'environnement (exposition à la lumière, température...), à la présence de prédateurs... De la même façon pour la division cellulaire, la taille, l' âge ou la charge en pathogène de la cellule peut affecter sa survie et sa division.

Exemple d'une structuration en âge. Introduisons ici un premier exemple de modèle de populations avec une structure en âge, qui affecte sa loi de reproduction. L'âge est vu en un sens large, cela peut être le nombre de générations depuis la naissance (nombre de saisons pour une plante ou divisions pour une cellule), ou le nombre d'unités de temps écoulées pour un nouvel infecté dans une épidémie. Pour simplifier, nous restreignons le trait considéré à l'âge, mais on pourra généraliser en couplant avec une variable spatiale ou phénotypique. Pour simplifier aussi, on reste dans un cadre d'une reproduction avec un seul enfant (comme pour la division cellulaire), à savoir $L^a \in \{0, 1, 2\}$ (mort, survie ou reproduction) avec une loi $p(a) = (p_k(a), 0 \leq k \leq 2)$ et un trait $a \in \mathcal{X} = \mathbb{N}$ qui est l'âge :

$$\begin{aligned}
&\mathbb{P}(\Theta_a = \varnothing) = p_0(a)\\
&\mathbb{P}(\Theta_a = (a+1)) = p_1(a)\\
&\mathbb{P}(\Theta_a = (a+1, 0)) = p_2(a).
\end{aligned}$$

Nous allons nous intéresser à la distribution des âges dans la population et à la croissance globale de la population. La question a priori plus simple de la survie possible de la population n'est ici pas triviale, car la taille totale n'est pas décrite par un processus de Galton-Watson. Nous verrons ensuite via le semigroupe des résulats généraux qui lieront cette question de la survie à la position de la valeur propre maximale par rapport à 1. Cette approche générale est souvent non explicite et ne fournit pas facilement un critère de survie en terme des paramètres du modèle. Mais la structure de branchement permet un autre point de vue, qui peut être plus explicite, à travers la structure régénérative liée à l'espace des traits. C'est sur ce point de vue qu'est basé la notion de R_0 en épidémiologie : on regarde ce qu'un individu laisse comme descendance globalement avant de mourrir. Si c'est plus que 1 en moyenne, la population va pouvoir se développer, sinon non. Mais l'hétérogénéité (dépendance de la loi de reproduction par rapport aux traits) fait qu'il faut faire attention à ce qu'on compte. Si un individu laisse en moyenne 3 enfants incapables de se reproduire, la survie est compromise... Il faut typiquement mettre en avant une structure de Galton-Watson (simple type) à l'intérieur de notre population structurée, ou plus généralement une structure régénérative. La structuraction en âge est bon exemple pour voir comment cela fonctionne, voir aussi [6] pour une structure spatiale décrite par un graphe. Posons

$$\mu = \sum_{a \geq 0} p_2(a) \prod_{b=0}^{a-1} (1 - p_0(b)).$$

Ce paramètre permet de classifier le comportement en temps long de la population structurée en âge.

Proposition 53. *Si $\mu < 1$, alors l'effectif de la population reste borné :*

$$\mathbb{P}\left(\sup_{n\in\mathbb{N}} \#\mathbb{G}_n < \infty\right) = 1.$$

Si $\mu > 1$, avec probabilité positive, la population survit et tend alors vers l'infini :

$$\mathbb{P}(\lim_{n\to\infty} \#\mathbb{G}_n = \infty) > 0$$

La preuve éclaire ces questions en établissant que la sous population des nouveaux nés forme un processus de Galton-Watson de moyenne de reproduction μ. Ainsi quand $\mu < 1$ (et même quand $\mu = 1$ mais alors il faut exclure le cas dégénéré ou un individu se reproduit exactement une fois au cours de sa vie), la population totale de nouveau nés cumulée au cours du temps (et donc le nombre cumulé d'invividu) est fini p.s.

Démonstration. Considérons un individu d'âge 0 et notons T le nombre de génération que cet individu va vivre avant de mourir. Dit autrement, si la racine a l'age 0, on regarde

$$T = \max\{a \geq 0 : \mathbf{1}_a \in \mathbb{G}_a\},$$

en notant $\mathbf{1}_a = (1, \ldots, 1) \in \mathcal{U}$ le label de l'individu avec a fois 1. En effet, nous pouvons choisir dans notre labelisation qu'un individu u qui survit est noté $u1$ et que son enfant est noté $u2$. Or

$$\mathbb{P}(T \geq a+1 | T \geq a) = 1 - p_0(a)$$

et donc

$$\mathbb{P}(T \geq a) = \mathbb{P}(\mathbf{1}_a \in \mathbb{G}_a) = \prod_{b=0}^{a-1}(1 - p_0(b))$$

Comptons maintenant le nombre de fois où l'individu a donné naissance avant de mourir :

$$N = \#\{a \geq 0 : (\mathbf{1}_a, 2) \in \mathbb{G}_{a+1}\}.$$

En moyenne,

$$\mathbb{E}(N) = \sum_{a\geq 0} \mathbb{P}((\mathbf{1}_a, 2) \in \mathbb{G}_{a+1}) = \sum_{a\geq 0} \mathbb{P}(\mathbf{1}_a \in \mathbb{G}_n).p_2(a)$$

Suivons maintenant la sous population d'âge 0. La racine, avant de mourrir, va fabriquer un nombre N d'enfants d'âge 0, qui vont perpétuer ce cycle de la vie, indépendamment pour chaque nouveau né, et suivant la même loi, du fait de l'indépendance entre individus dans le modèle (propriété de branchment) ... Le comptage du nombre d'individus d'âge 0 dans la population est ainsi réalisé par un processus de Galton Watson (qui ne suit plus les générations initiales, mais les générations de nouveaux nés). On rappelle que l'extinction de ce processus est donné par le critère $\mathbb{E}(N) \leq 1$, sauf dans le cas dégénéré $\mathbb{P}(N = 1) = 1$. Or quand la population d'âge 0 s'éteint, la population ne plus croitre et elle reste bornée. Et réciproquement. $\square$

3.2 Semigroupe et formule tous-pour-un

La définition du semigroupe du premier moment est inchangée. Pour simplifier les discussions et parce que c'est raisonnable d'un point de vue modélisation, nous supposerons dans tout ce chapitre que le nombre moyen d'enfants par individu est borné sur l'espace des traits

$$\sup_{x\in\mathcal{X}} \mathbb{E}(L^x) < \infty.$$

Ceci nous permet en particulier d'introduire le semigroupe sur l'espace des fonctions continues bornées.

Définition 54. *Pour tout $x \in \mathcal{X}$ et f fonction mesurable bornée sur $\mathcal{X}$,*

$$M(x,f) = \mathbb{E}_{(\varnothing,x)}\left[\sum_{u\in\mathbb{G}_1} f(X(u))\right]$$

est bien défini et fini et il en va de même plus généralement pour $n \geq 0$ de

$$M_n(x,f) = \mathbb{E}_{(\varnothing,x)}\left[\sum_{u\in\mathbb{G}_n} f(X(u))\right].$$

En effet, l'hypothèse de premier moment borné nous permet de dire que M agit sur les fonctions bornées :

$$M(x,f) \leq \| f \|_\infty \sup_{x\in\mathcal{X}} \mathbb{E}(L^x).$$

On peut souvent étendre l'espace sur lequel le semigroupe agit en considérant en considérant par exemple les fonctions dominées par une fonction, notamment une fonction propre ou fonction de Lyapunov pour avoir une propriété de stabilité.

La propriété de semigroupe montrée dans le cas neutre au chapitre précédent sétend directement :

Proposition 55. *Pour tous $x \in \mathcal{X}$ et f mesurable bornée et $n \geq 0$,*

$$M_{n+1}(x,f) = M_n(x, M(\cdot,f)) = M(x, M_n(\cdot,f)).$$

Les mesures associées sont notées $M(x,dy)$ et $M_n(x,dy)$.

Nous pouvons également réduire la mesure empirique moyenne à la dynamique du trait d'un seul individu. La dynamique en question est par contre plus compliqée avec une inhomogénéité en temps pour la propriété de Markov. Nous allons aussi faire intervenir une fonction ψ mesurable (strictement) positive sur $\mathcal{X}$, qui va jouer un rôle important par la suite. On pose ainsi pour $i < n$,

$$Q_i^{(n)}(x,A) = \frac{1}{M_{n-i}(x,\psi)} \int_A M(x,dy) M_{n-i-1}(y,\psi).$$

pour A mesurable de $\mathcal{X}$, ce qui définit la mesure de probabilité $Q_i^{(n)}(x,.)$

Proposition 56. *Soient $n \geq 0$, F mesurable positive sur $\mathcal{X}^{n+1}$ et ψ mesurable bornée strictement positive sur $\mathcal{X}$. Alors :*

$$\mathbb{E}_{(\varnothing,x)}\left[\sum_{u\in\mathbb{G}_n}\psi(X(u))F(X(v),\, v\preccurlyeq u)\right] = M_n(x,\psi)\,\mathbb{E}_x\left[F\left(Y_i^{(n)},\, i\leq n\right)\right],$$

où $\left(Y_i^{(n)}\right)_{i\leq n}$ chaîne de Markov inhomogène définie pour A ensemble mesurable de $\mathcal{X}$ et $0\leq i\leq n-1$ par

$$\mathbb{P}\left(Y_{i+1}^{(n)}\in A\middle|Y_i^{(n)}=x\right) = Q_i^{(n)}(x,A)\,.$$

Démonstration. La preuve est proche du cas neutre. On va ainsi établir le résultat sur les fonctions cylindriques $F_n(x)=\prod_{j=0}^{n} f_j(x_j)$ par récurrence, en omettant la condition initiale $(\varnothing,x)$ dans la notation, et conclure par un lemme de classes monotones.

$$\begin{aligned}
&\mathbb{E}\left[\sum_{u\in\mathbb{G}_{n+1}}\psi(X(u))F_{n+1}(X(v),\, v\preccurlyeq u)\right]\\
&\quad=\mathbb{E}\left[\sum_{w\in\mathbb{G}_n}F_n(X(v),\, v\preccurlyeq w)\times\mathbb{E}\left[\sum_{u\in\mathbb{G}_{n+1},\,u\succcurlyeq w}\psi.f_{n+1}(X(u))\middle|\mathcal{F}_n\right]\right]\\
&\quad=\mathbb{E}\left[\sum_{w\in\mathbb{G}_n}F_n(X(v),\, v\preccurlyeq w)\times M(X(w),\psi f_{n+1})\right]\\
&\quad=\mathbb{E}\left[\sum_{w\in\mathbb{G}_n}\psi(X(w))\widetilde{F_n}(X(v),\, v\preccurlyeq w)\right]
\end{aligned}$$

avec

$$\widetilde{F_n}(y)=\widetilde{F_n}(y_0,y_1,\ldots,y_n)=F_n(y)\frac{M(y_n,\psi f_{n+1})}{\psi(y_n)}.$$

On utilise alors l'hypothèse de récurrence, ce qui donne, en notant $y_0=x$,

$$\begin{aligned}
&\mathbb{E}\left[\sum_{u\in\mathbb{G}_{n+1}}\psi(X(u))F_{n+1}(X(v),\, v\leq u)\right]\\
&\quad=M_n(x,\psi)\mathbb{E}_x\left[\widetilde{F_n}\left(Y_i^{(n)},\, i\leq n\right)\right]=M_n(x,\psi)\int_{\mathcal{X}^n}\widetilde{F_n}(y)\prod_{i=0}^{n-1}Q_i^{(n)}(y_i,dy_{i+1})\\
&\quad=M_n(x,\psi)\int_{\mathcal{X}^n}F_n(y)\frac{M(y_n,\psi f_{n+1})}{\psi(y_n)}\prod_{i=0}^{n-1}\frac{M_{n-i-1}(y_{i+1},\psi)}{M_{n-i}(y_i,\psi)}M(y_i,\mathrm{d}y_{i+1})\\
&\quad=\int_{\mathcal{X}^n}F_n(y)M(y_n,\psi f_{n+1})\prod_{i=0}^{n-1}M(y_i,\mathrm{d}y_{i+1})\\
&\quad=\int_{\mathcal{X}^{n+1}}F_n(y)\,\psi f_{n+1}(y_{n+1})\prod_{i=0}^{n}M(y_i,\mathrm{d}y_{i+1})
\end{aligned}$$

grâce au théoreme de Fubini. Or

$$\begin{aligned}\mathbb{E}_{y_0}\left[F_{n+1}\left(Y_i^{(n+1)},\, i\leq n+1\right)\right] &= \int_{\mathcal{X}^{n+1}} F_{n+1}(y_i,\, i\leq n+1)\prod_{i=0}^{n} Q_i^{(n)}(y_i,\mathrm{d}y_{i+1})\\ &= \int_{\mathcal{X}^{n+1}} F_{n+1}(y_i,\, i\leq n+1)\frac{\psi(y_{n+1})}{M_{n+1}(y_0,\psi)}\prod_{i=0}^{n} M(y_i,\mathrm{d}y_{i+1})\end{aligned}$$

ce qui prouve que les deux termes coincident au facteur $M_{n+1}(y_0,\psi)$ près et la récurrence est établie. □

Exercice. Obtenir une formule pour les fourches, étendant le cadre neutre du chapitre précédent.

3.3 Fonction harmonique et espace de traits fini

Un cadre va jouer un rôle important, celui où on peut exploiter l'existence d'une fonction propre strictement positive pour le semigroupe. De nombreux résultats en mathématiques (théorie spectrale ou contraction) assurent l'existence de ces éléments propres et l'unicité du vecteur propre associé à la valeur maximale et le caractère positif de la fonction. En dimension finie, le théorème de Perron Frobenius assure un tel résultat dès que la matrice M est apériodique et irréductible (ou, ce qui est équivalement, primitive). Ce cas permet de traiter par exemple divers modèles structurés avec un espace de trait fini, notamment les modèles de graphes introduits dans le chapitre précédent dans un cadre irréductible apériodique. Différentes extensions existent en dimension infinie, en particulier avec la thèorème de Krein Rutman.

Proposition 57. *Supposons qu'il existe une fonction h strictement positive et bornée sur $\mathcal{X}$ telle que pour tous $x\in\mathcal{X}$, $M(x,h)=\lambda h(x)$, avec $\lambda>0$. Alors*

$$\mathbb{E}_{(\varnothing,x)}\left[\sum_{u\in\mathbb{G}_n} h(X(u))F(X(v),\, v\preccurlyeq u)\right] = \lambda^n\, h(x)\,\mathbb{E}_x[F(Y_i,\, i\leq n)]\,,$$

avec cette fois-ci Y chaîne de Markov homogène, de noyau de transition

$$Q(x,dy)=\frac{h(y)}{\lambda h(x)}M(x,dy).$$

De plus $\lambda^{-n}\sum_{u\in\mathbb{G}_n} h(X(u))$ est une martingale positive, qui converge p.s. vers $W\in[0,\infty)$.

Remarquons qu'on peut relâcher ici la bornitude de $\psi=h$, comme il s'agit d'une fonction propre. On pourrait aussi relâcher aussi la stricte positivité en définissant la chaine de Markov hors des points où h s'annule.

Démonstration. Dans le cas où ψ est une fonction propre, $M_{n-i}(x,h)=\lambda^{n-i}h(x)$ et le noyau $Q_i^{(n)}=Q$ ne dépend plus de i ni de n, ce qui simplifie le résultat de la formule

tous-pour-un de la proposition précédente. Ceci prouve la première partie du résultat. Pour la deuxième, on écrit

$$\mathbb{E}\left(\sum_{u\in\mathbb{G}_{n+1}} \psi(X(u))\,\middle|\mathcal{F}_n\right) = \sum_{v\in\mathbb{G}_n} \mathbb{E}\left(\sum_{u\in\mathbb{G}_{n+1},\, u\succcurlyeq v} h(X(u)) \mid \mathcal{F}_n\right)$$
$$= \sum_{v\in\mathbb{G}_n} M(X(v),\psi) = \lambda \sum_{v\in\mathbb{G}_n} h(X(v))$$

et on conclut en divisant de part et d'autre par λ^{n+1}. □

Une question délicate se pose alors, notre martingale limite est elle dégénérée ou non ? Si on bénéficie de domination des moments d'ordre 2 dans la loi de reproduction, on peut montrer (via une formule pour les fourches qui généralise celle prouvée dans le cas neutre par exemple) que la martingale est bornée dans L^2. Elle est donc convergente dans L^2 et dans L^1, c'est à dire que $\mathbb{P}(W > 0) > 0$. Or ceci implique en général que W est positive sur l'événement de survie, en s'inspirant de la démonstration faite pour le processus de Galton Watson. En effet, quand la population survit, on montre souvent que la taille de population tend vers l'infini, car la taille doit sortir de tout compact pour éviter l'extinction quand 0 est accessible. Si on peut prouver alors que le nombre d'invividus avec un trait donné tend vers l'infini, on est revenu au cas simple type. Sinon, les preuves peuvent se compliquer car il faut a priori controler la probabilité que $W = 0$ en fonction du trait de l'individu de départ. Le résultat peut etre mis en défaut si les traits évoluent vers un trait pour lequel la probabilité tend trop vite vers 0.

La non dégénérescence de la martingale limite W avec des conditions de moments plus larges, notamment de type $L \log L$ comme on a pu le montrer pour les processus de Galton Watson, peut être appréhendée par une approche spinale, en suivant les idées du premier chapitre. On pourra consulter notamment [17] pour le cas d'un espace de traits $\mathcal{X}$ fini (et une loi des grands nombres pour les proportions) et [3] pour des cas infinis.

Espace de trait fini : $\mathcal{X} < \infty$ Donnons les grandes lignes de l'extension de la construction spinale au cas multitype, en suivant [17]. On se place dans un cas où la matrice $M = (M_{i,j})_{i,j\in\mathcal{X}}$ associée à l'opérateur linéaire M (i.e. $M_{i,j} = M(i, \mathbf{1}_j)$) est irréductible apériodique. Le théorème de Perron Frobenius guarantit alors l'existence et l'unicité d'un triplet positif (λ, h, π) déléments propres positifs pour M :

$$Mh = \lambda h, \quad \pi M = \lambda \pi.$$

On peut identifier une fonction $h : \mathcal{X} \to \mathbb{R}_+$ et le vecteur $h = (h(i))_{i\in\mathcal{X}}$, et ainsi le vecteur propre h correspond à la fonction propre $M(., h) = \lambda h$.
Notons aussi que le théorème de Perron Frobenius assure aussi l' unicité de ces éléments propres sous les conditions de normalisation $\sum_{i\in\mathcal{X}} \pi_i = \sum_{i\in\mathcal{X}} \pi_i h(i) = 1$.

Construisons un nouvel arbre aléatoire, avec un individu distingué, comme dans le premier chapitre. Mais cette fois les individus ont un trait qui se transmet et l'épine va héritér d'une loi de transmission particulière, en plus du biais par la taille sur le nombre d'enfants. L'individu distingué est labelisé par E_n à la génération n et il crée des individus

dont les traits sont donnés par la v.a. $\widehat{\Theta}_x$ pour $x \in \mathcal{X}$. La loi de cette v.a. est donnée par

$$\mathbb{P}(\widehat{\Theta}_x = (x_1, \dots, x_k)) = \mathbb{P}(\Theta_x = (x_1, \dots, x_k)) \frac{\sum_{i=1}^k h(x_i)}{\lambda h(x)}.$$

pour $k \geq 0$ et $x_i \in \mathcal{X}$. Notons que c'est parce que h est fonction propre que l'on définit ici bien une loi de probabilité. Autrement dit,

$$\mathbb{E}(F(\widehat{\Theta}_x)) = \frac{1}{\lambda h(x)} \mathbb{E}\left(F(\Theta_x) \sum_{i=1}^{L^x} h(X_i^x) \right).$$

L'épine est donné par la racine à la génération 0 : $E_0 = \varnothing$ et son trait vaut $X(\varnothing) = x_0$. On choisit l'épine à la génération $n+1$ parmi les enfants de lépine à la génération n de la façon suivante : pour tout $k \geq 1$, $1 \leq i \leq k$, $e \in \mathcal{U}$, $x_j \in \mathcal{X}$:

$$\mathbb{P}(E_{n+1} = ei \,|\, E_n = e,\, \Theta_x(e) = (x_1, \dots, x_k),\, X(e) = x) = \frac{h(x_i)}{\sum_{j=1}^k h(x_j)}.$$

Comme pour la construction spinale pour les arbres de Galton-Watson, les individus non distingués se reproduisent indépendamment et suivant la loi originale Θ_x. On note encore $\mathcal{T}_n$ (resp. $\mathcal{A}_n$) l'arbre aléatoire donnant les individus en vie jusqu'à l'instant n dans le processus original (resp. dans la construction spinale) et $(X(u) : u \in \mathcal{T}_n)$ (resp. $(\widehat{X}(u) : u \in \mathcal{A}_n)$) les traits des individus à la génération n.
Pour tous $n \in \mathbb{N}$ et Ffonction bornée à valeurs réelles :

$$\mathbb{E}_{(\varnothing,x)}[F((X(u), u \in \mathcal{T}_n), U_n) W_n] = h(x) \mathbb{E}_{(\varnothing,x)}\Big[F\Big((\widehat{X}(u), u \in \mathcal{A}_n), E_n \Big) \Big],$$

avec $W_n = \sum_{u \in \mathbb{G}_n} h(X(u))/\lambda^n$.
Notons en particulier, en suivant toujours [17], que l'on peut prouver en utilisant cette construction que la martingale limite W est positive sur l'événement de survie si et seulement les moments $L \log(L)$ sont finis pour les différents types d'individus. En effet, comme dans le premier chapitre, la population dans la construction spinale peut etre vue comme le processus orginal plus une immigration.

3.4 Conditions de Doeblin et uniforme ergodicité

Pour f fonction mesurable positive ou bornée et μ mesure de probabilité et $(P(x, dy))_{x \in \mathcal{X}}$ famille de mesure de probabilité telle que pour tout A mesurable, $x \to P(x, A)$ soit mesurable, on rappelle les notations suivantes. Le réel $\mu(f)$ et la mesure de probabilité μP sont définis par

$$\mu(f) = \int_{\mathcal{X}} f(x) \mu(dx), \qquad \mu P(A) = \int_{\mathcal{X}} \mu(dx) P(x, A)$$

On introduit la norme en variation totale définie par

$$\| \mu \|_{V.T.} = \sup_{\|f\|_\infty \leq 1} |\mu(f)|.$$

Rappelons tout d'abord le résultat de contraction pour une chaine de Markov sous la condition de Doeblin, qui correspond à un couplage des trajectoires des processus de Markov.

Lemme 58. *Supposons qu'il existe $c \in (0,1]$ et ν mesure de probabilité sur $\mathcal{X}$ tels que*

$$\forall x \in \mathcal{X}, \quad P(x,.) \geq c\nu(.)$$

Alors pour toutes mesures de probabilité μ_1, μ_2 sur $\mathcal{X}$,

$$\| \mu_1 P - \mu_2 P \|_{V.T.} \leq (1-c) \| \mu_1 - \mu_2 \|_{V.T.}$$

Démonstration. Posons

$$R(x,dy) = \frac{P(x,dy) - c\nu(dy)}{1-c}.$$

$R(x,.)$ est une mesure de probabilité et $R(x,f) \leq \| f \|_\infty$. Alors pour $\| f \|_\infty \leq 1$, en écrivant $P(x,f) = (1-c)R(x,f) + c\nu(f)$,

$$\mu_1 Pf - \mu_2 Pf = (1-c)(\mu_1 Rf - \mu_2 Rf) \leq (1-c) \| \mu_1 - \mu_2 \|_{V.T.}$$

ce qui implique le résultat en prenant la borne supérieure à gauche. □

Montrons maintenant la conséquence suivante sur la chaine $Y^{(n)}$.

Proposition 59. *On suppose qu'il existe $c \in (0,1]$ et ν mesure de probabilité sur $\mathcal{X}$ tels que pour tout $i \leq n-1$,*

$$Q_i^{(n)}(x,.) \geq c\nu(.).$$

Alors il existe une mesure de probabilité π sur $\mathcal{X}$ telle que

$$\sup_{x \in \mathcal{X}, |f| \leq 1} \left| \mathbb{E}_x(f(Y_n^{(n)})) - \pi(f) \right| \leq (1-c)^n.$$

Démonstration. On note $PQ(x,.) = \int_{\mathcal{X}} P(x,dy)Q(y,.)$ de sorte que $\mu PQ = (\mu P)Q = \mu(PQ)$. En utilisant que $\mu Q_0^{(n)} \dots Q_{n-1}^{(n)} = (\mu Q_0^{(n)} \dots Q_{n-2}^{(n)})Q_{n-1}^{(n)}$ et en itérant la contraction de Doeblin on obtient

$$\| \mu_1 Q_0^{(n)} \dots Q_{n-1}^{(n)} - \mu_2 Q_0^{(n)} \dots Q_{n-1}^{(n)} \|_{V.T.} \leq (1-c)^n \| \mu_1 - \mu_2 \|_{V.T.}$$

Notons

$$V_n(x,f) = \mathbb{E}_x(f(Y_n^{(n)})) = \delta_x Q_0^{(n)} \dots Q_{n-1}^{(n)}(f)$$

et utilisons que

$$Q_{i+1}^{(n+1)}(x,dy) = \frac{M(x,dy)M_{(n+1)-1-(i+1)}(y,\psi)}{M_{(n+1)-(i+1)}(y,\psi)} = Q_i^{(n)}(x,dy).$$

Donc

$$V_{n+1}(x,f) = (\delta_x Q_0^{(n+1)})Q_0^{(n)} \dots Q_{n-1}^{(n)}(f)$$

et la contraction donne avec $\mu_1 = \delta_x Q_0^{(n+1)}$ et $\mu_2 = \delta_x$

$$|V_{n+1}(x,f) - V_n(x,f)| \leq (1-c)^n$$

pour $|f| \leq 1$. $(V_n(f))_n$ est donc une suite de Cauchy et converge vers une valeur notée $\pi(f)$. Or de façon similaire

$$|V_{n+p}(x,f) - V_n(x,f)| \leq (1-c)^n$$

et donc

$$|V_n(x,f) - \pi(f)| \leq (1-c)^n$$

et π est alors une mesure de probabilité, ce qui conclut la preuve. □

Passons à l'application dans l'exemple de de la structure d'âge : $\mathbb{P}(\Theta^{(a)} = \varnothing) = p_0(a)$, $\mathbb{P}(\Theta^{(a)} = (a+1)) = p_1(a)$ et $\mathbb{P}(\Theta^{(a)} = (a+1,0)) = p_2(a)$ avec $p_0(a)+p_1(a)+p_2(a) = 1$.

Corollaire 60. *Supposons que les fonctions p_1 et p_2 sont décroissantes (en âge) et strictement positives et qu'il existe $d \geq 0$ tel que pour tout $a \geq 0$, $p_1(a) \leq dp_2(a)$. Alors il existe $c > 0$ tel que pour tout $x \in \mathbb{N}$,*

$$\sum_{a\geq 0} \left| \frac{\mathbb{E}_{(\varnothing,x)}(\#\{u \in \mathbb{G}_n : X(u) = a\})}{\mathbb{E}_{(\varnothing,x)}(\#\mathbb{G}_n)} - \pi(a) \right| \leq (1-c)^n$$

Ce résultat se généralise, notamment en montrant la condition de Doeblin sur b générations plutot que sur 1. En particulier la condition de domination de p_2 par p_1 peut être relachée.

Démonstration. On commence par montrer qu'il existe $c > 0$ tel que que pour tous $i \leq n-1$ et $a \geq 0$,

$$Q_i^{(n)}(a, \{0\}) \geq c.$$

On remarque d'abord que $M_k(a,1) \leq M_k(0,1)$ par décroissance de p_2 et p_1 (le nombre moyen total fabriqué à un instant donné est décroissant en a, par récurrence)

$$\begin{aligned} M_{k+1}(a,1) &= p_2(a)(M_k(0,1) + M_k(a+1,1)) + p_1(a)M_k(a+1,1) \\ &\leq p_2(a)(2+d)M_k(0,1) \end{aligned}$$

soit $p_2(a)M_k(0,1) = M(a,\{0\})M_k(0,1) \geq cM_{k+1}(a,1)$ avec $c = 1/(2+d)$. On peut donc appliquer la proposition précédente à $Y^{(n)}$ et en notant μ_n la loi de $Y_n^{(n)}$:

$$\mu_n^x = \mathbb{E}_x(f(Y_n^{(n)})) = \frac{\mathbb{E}_{(\varnothing,x)}(\sum_{u\in\mathbb{G}_n} f(X(u))}{\mathbb{E}_{(\varnothing,x)}(\#\mathbb{G}_n)},$$

et on obtient l'existence d'une probabilité π sur $\mathbb{N}$ telle que

$$\sup_{x\in\mathcal{X}, |f|\leq 1} \left|\mu_n(f) - \pi(f)\right| \leq (1-c)^n.$$

On utilise maintenant que la norme en variation totale correspond à la norme L^1 pour conclure. Plus précisément pour $\mathcal{X}$ dénombrable

$$\sup_{|f|\leq 1} \left|\mu_n(f) - \pi(f)\right| = \sup_{|f|\leq 1} \left|\sum_{x\in\mathcal{X}} f(x)(\mu_n(\{x\}) - \pi(\{x\}))\right| = \sum_{x\in\mathcal{X}} \left|\mu_n(\{x\}) - \pi(\{x\})\right|$$

en utilisant la fonction $f(x) = signe(\mu_n(x) - \pi(x))$ pour $x \in \mathcal{X}$. □

On obtient ici une convergence du nombre moyen d'individus d'âge a sur le nombre moyen d'individus total. On peut ensuite montrer une loi des grands nombres, c'est à dire une convergence en probabilité ou p.s. de la proposition d'individus d'âge a dans la population en adaptant la démonstration dans le cas neutre du chapitre précédent (l'ergodicité du trait de l'individu ayant bien été obtenu). On peut également obtenir l'equivalent du nombre d'individus total en utilisant la martingale associé à la fonction propre.

Plus généralement, si la condition de Doeblin n'est pas satisfaite sur tout l'espace, on peut généraliser la contraction précédente en supposant une condition de Doeblin locale, plus une condition de Lyapounov. La méthode de contraction sur le processus $Y_i^{(n)}$ permet d'ailleurs de prouver l'existence d'éléments propres pour M, fournissant des outils complémentaires aux techniques spectrales.

Chapitre 4

Un exemple d'approximation en grande population

Nous allons dans ce chapitre décrire et approcher une population dans une limite d'un grand nombre d'individus et un déplacement rapide (diffusion). Nous nous concentrons sur le cadre de la marche aléatoire branchante.

4.1 Modèle et premiers résultats

4.1.1 Définition

On considère un nombre fini d'individus initiaux et on les labelise par l'ensemble

$$I = \bigsqcup_{n\geq 1} \mathbb{N}^n.$$

On considère un paramètre échelle $K \geq 1$, associée à une variable aléatoire de reproduction intérgrable L^K à valeurs dans $\mathbb{N}$. On note alors pour $k \geq 0$,

$$p_k^K = \mathbb{P}\big(L^K = k\big)$$

et on suppose que la moyenne m^K de cette loi est finie :

$$m_K = \sum_{k\geq 0} k p_k^K < \infty.$$

Dans ce cas, la loi de reproduction dispersion devient

$$x \mapsto \Theta_x^K = \big(x + D_1^K, \ldots, x + D_{L^K}^K\big),$$

avec $D_i^K \overset{\text{iid}}{\sim} D/K$ à valeurs réelles, et D centrée de variance finie :

$$\mathbb{E}(D) = 0, \qquad \sigma^2 = \mathbb{E}(D^2) < \infty.$$

Considérons une population initiale labélisée par $\mathbb{G}_0^K \subset I$, avec $Z_0^K = \#\mathbb{G}_0^K$ individus. Pour simplifier, supposons

$$\mathbb{E}(\#\mathbb{G}_0^K) < \infty$$

On introduit une famille de v.a. $(L^K(u))_{u\in I}$ i.i.d. distribuées comme L^K, et, indépendamment, les variables $\big(D_i^K(u)\big)_{u\in I, i\geq 1}$ i.i.d distribué comme D/K.

Définition 61. *Le processus* $X^K = (X^K(u),\, u \in \mathcal{T}^K)$ *est défini récursivement par*

$$\mathbb{G}_{n+1}^K = \big\{uk :\, u \in \mathbb{G}_n^K,\, 1 \le k \le L^K(u)\big\}$$

et pour tout $u \in \mathbb{G}_n^K$, $1 \le k \le L^K(u)$,

$$X^K(uk) = X^K(u) + D_k^K(u).$$

L'objet qui nous intéresse est la mesure empirique

$$X_n^K = \frac{1}{K} \sum_{u \in \mathbb{G}_n^K} \delta_{X^K(u)}$$

sur $\mathbb{R}$. Elle rend compte de la distribution des positions des différents individus. Avec les bons changements d'échelle, nous allons prouver une convergence en loi du processus vers un objet limite déterministe.

4.1.2 Décomposition en semi-martingale

Commençons par un énoncé général, qui reprend la décomposition habituelle en semi-martingale. On note $(\mathcal{F}_n^K)_{n\in\mathbb{N}}$ la filtration du processus, c'est à dire que $\mathcal{F}_n^K$ est la tribu engendrée par les variables aléatoires $\Theta^K(u)$ jusqu'à l'instant n et la condition initiale $(X^K(u), u \in \mathbb{G}_0^K)$.

Proposition 62. *Soit* $f : \mathbb{R} \to \mathbb{R}$ *mesurable bornée et* $n \ge 0$. *Alors*

$$X_n^K(f) = V_n^K(f) + M_n^K(f)$$

où $V_0^K(f) = X_0^K(f)$, $V_n^K(f)$ *est une variable* $\mathcal{F}_{n-1}^K$*-mesurable pour* $n \ge 1$ *et* $\big(M_n^K(f)\big)_{n\in\mathbb{N}}$ *une* $\big(\mathcal{F}_n^K\big)$*-martingale. En outre,*

$$V_n^K(f) = X_0^K(f) + \sum_{k=1}^n \Delta_k^K(f)$$

avec $\Delta_k^K(f) = \mathbb{E}\big[X_k^K(f) - X_{k-1}^K(f) \big| \mathcal{F}_{k-1}^K\big]$. *Enfin*

$$M_n^K = \sum_{k=1}^n \delta_k^K(f),$$

avec $\delta_k^K(f) = X_k^K(f) - X_{k-1}^K(f) - \Delta_k^K(f)$.

Démonstration. Ce résultat est classique et pas spécifique au cadre considéré ici. On peut vérifier aisément que, par définition, V^K est prévisible et M^K est une martingale. L'intégrabilité de M^K est assurée par l'inégrabilité de L^K et la bornitude de f.
Pour trouver la décomposition (et au passage obtenir d'ailleurs son unicité), on pouvait raisonner par analyse, c'est à dire supposer une décomposition en un processus V^K prévisible plus une martingale M^K. Alors

$$\mathbb{E}\big[X_n^K(f) \big| \mathcal{F}_{n-1}^K\big] = V_n^K(f) + \mathbb{E}\big[M_n^K(f) \big| \mathcal{F}_{k-1}^K\big] = V_n^K(f) + M_{n-1}^K(f)$$

et donc

$$\mathbb{E}\big[X_n^K(f) - X_{n-1}^K(f)\big|\mathcal{F}_{n-1}^K\big] = V_n^K(f) - V_{n-1}^K(f).$$

On voit ainsi comment définir V^K en sommant ses incréments puis on définit M^K par différence. □

Pour notre exemple, introduisons le semigroupe du premier moment

$$M^K(x,f) = K\mathbb{E}_{(1,x)}\big[X_1^K(f)\big] = m^K\mathbb{E}\left[f\left(x + \frac{D}{K}\right)\right]$$

avec $m^K = \mathbb{E}\big[L^K\big]$. On en déduit :

$$\Delta_k^K(f) = \frac{1}{K}\sum_{u\in\mathbb{G}_{k-1}^K} M^K\big(X^K(u), f\big) - f\big(X^K(u)\big)\,. \tag{4.1}$$

On suppose dans le reste du chapitre que

$$\mathbb{E}\Big[\big(L^K\big)^2\Big] < \infty.$$

Proposition 63. *Supposons f mesurable bornée. Alors*

$$M_n^K(f)^2 = \frac{1}{K^2}\sum_{k=1}^n \sum_{u\in\mathbb{G}_{k-1}^K} G_f^K\big(X^K(u)\big) + \widetilde{M}_n^K\,,$$

avec $\widetilde{M}^K$ une $\mathcal{F}^K$-martingale issue de 0 et :

$$G_f^K(x) = \mathbb{E}_{(1,x)}\left[\left(\sum_{w\in\mathbb{G}_1^K} f\big(X^K(w)\big) - M^K(x,f)\right)^2\right].$$

Démonstration. On peut faire la même décomposition que dans la proposition précédente, cette fois pour $(M^K)^2$ plutôt que X. En oubliant K et f dans les notations pour alléger :

$$M_n^2 = \sum_{k=1}^n \widetilde{\Delta_k} + \widetilde{M_n}\,,$$

avec $\widetilde{\Delta_k} = \mathbb{E}\big[M_k^2 - M_{k-1}^2\big|\mathcal{F}_{k-1}\big]$. Pour le terme de gauche, remarquons que $M_k^2 = (M_{k-1} + \delta_k)^2$, donc en développant puis en passant à l'espérance :

$$\widetilde{\Delta_k} = \mathbb{E}\big[2M_{k-1}\delta_k + \delta_k^2\big|\mathcal{F}_{k-1}\big] = \mathbb{E}\big[\delta_k^2\big|\mathcal{F}_{k-1}\big]\,.$$

car $\mathbb{E}(\delta_k\,|\,\mathcal{F}_{k-1}) = 0$. En outre,

$$\delta_k^2 = (X_k - X_{k-1} - \Delta_k)^2 = (X_k - \mathbb{E}[X_k|\mathcal{F}_{k-1}])^2,$$

et, en notant $R_k(u) = \sum_{w\in\mathbb{G}_k, w\geq u} f(X(w)) - M(X(u), f)$,

$$X_k - \mathbb{E}[X_k|\mathcal{F}_{k-1}] = \frac{1}{K}\sum_{u\in\mathbb{G}_{k-1}} R_k(u)$$

Donc

$$\delta_k^2 = \frac{1}{K^2}\sum_{u,v\in\mathbb{G}_{k-1},\, u\neq v} R_k(u)R_k(v) + \frac{1}{K^2}\sum_{u\in\mathbb{G}_{k-1}} R_k(u)^2$$

En utilisant que $\mathbb{E}(R_k(u)|\mathcal{F}_{k-1}) = 0$ et la propriété de branchement [1], on obtient

$$\widetilde{\Delta_k} = \frac{1}{K^2}\sum_{u\in\mathbb{G}_{k-1}} \mathbb{E}(R_k(u)^2|\mathcal{F}_{k-1})$$

On conclut en utilisant que $\mathbb{E}(R_k(u)^2|\mathcal{F}_{k-1}) = G_f^K(X(u))$. □

4.2 Convergence vers une EDP

Avec la formule *tous-pour-un*, on obtient :

$$\mathbb{E}_{(1,0)}\left[X_n^K(f)\right] = \frac{1}{K}\,\mathbb{E}\left[L^K\right]^n \mathbb{E}_0\left[f\left(S_n^K\right)\right],$$

avec $\left(S_n^K\right)_{n\in\mathbb{N}}$ une marche aléatoire de pas D/K, i.e.

$$S_n^K = \frac{1}{K}\sum_{i=1}^n D_i,$$

où $(D_i)_{i\geq 1}$ est une suite de v.a. i.i.d. de même loi que D. Par linéarité, en partant de K individus on a donc :

$$\mathbb{E}_{(1,0),\ldots,(K,0)}\left[X_n^K(f)\right] = \mathbb{E}\left[L^K\right]^n \times \mathbb{E}_0\left[f\left(S_n^K\right)\right].$$

Le théorème central limite assure que $S_n^K/\sqrt{n}$ converge en loi vers une variable aléatoire N_σ gaussienne centrée de variance σ^2/K^2. Ceci suggère une accélération en temps avec un facteur K^2 :

Définition 64. *Soit $(Y_t^K)_{t\in\mathbb{R}_+}$ le processus défini pour $t \geq 0$ par*

$$Y_t^K = X_{[K^2 t]}^K.$$

On suppose dans tout le reste de ce chapitre que $m_K = 1 + \frac{\alpha}{K^2} + o\left(\frac{1}{K^2}\right)$, i.e.

$$\lim_{K\to\infty} K^2(m_K - 1) = \alpha \in \mathbb{R}. \tag{4.2}$$

1. La propriété de branchement est pratique pour simplifier l'expression, mais on pourrait s'en passer et ajouter une densité dependance

Proposition 65. *Pour tout $t \geq 0$,*

$$\mathbb{E}_{\delta_0}\left[Y_t^K(f)\right] \underset{K\to\infty}{\longrightarrow} e^{\alpha t} \int_{-\infty}^{\infty} f(x) \frac{e^{-x^2/(2\sigma^2 t)}}{\sqrt{2\pi\sigma^2 t}} dx.$$

Démonstration. On utilise que $\mathbb{E}_{\delta_0}\left[Y_t^K(f)\right] = m_K^{[K^2 t]} \times \mathbb{E}_0\left[f\left(S_{[K^2 t]}^K\right)\right]$ puis

$$\lim_{K\to\infty} m_K^{[K^2 t]} = e^{\alpha t}, \qquad S_{[K^2 t]}^K \Rightarrow \mathcal{N}(0, \sigma^2 t)$$

en loi, car $S_{[K^2 t]}^K = K^{-1} \sum_{i=1}^{[K^2 t]} D_i$. □

On aimerait montrer un principe de loi des grands nombres, c'est à dire retirer l'espérance dans le terme de droite et montrer que le processus $\left(Y_t^K\right)_{t\geq 0}$ converge vers une équation de réaction diffusion, c'est à dire l'équation de la chaleur pour décrire la dispersion plus un terme de croissance démographique donné par α. Le principe de loi des grands nombres ici va reposer sur le grand nombre d'individus present dans la population au départ, et pas sur la propriété de branchement. L'approche se généralise à des modèles ou la propriété de branchement n'est plus vérifiée, autrement dit on peut prendre en compte des interactions comme la compétition, la coopération, l'intéraction avec les proies et prédateurs ...etc.

4.2.1 Préliminaires

On note $\mathcal{C}_b^{2,+}$ l'ensemble des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, de classe $\mathcal{C}^2$, bornées, dont les deux premières dérivées sont bornées, et dont la dérivée seconde est Lipschitz.

Lemme 66. *Pour tout $f \in \mathcal{C}_b^{2,+}$,*

$$\sup_{x\in\mathbb{R}} \left|K^2\left(M^K(x,f) - f(x)\right) - \mathcal{L}f(x)\right| \overset{K\to\infty}{\longrightarrow} 0,$$

avec

$$\mathcal{L}f(x) = \alpha f(x) + \frac{\sigma^2}{2} f''(x).$$

Démonstration. En utilisant que $\mathbb{E}(D) = 0$, la formule de Taylor Lagrange assure qu'il existe $v_{x,D}^K \in [x - |D|/K, x + |D|/K]$ tel que

$$\begin{aligned} K^2(M^K(x,f) - f(x)) &= K^2(m_K \mathbb{E}(f(x + |D|/K)) - f(x)) \\ &= K^2\left(\left(1 + \frac{\alpha}{K^2} + \varepsilon_K/K^2\right)\left(f(x) + \mathbb{E}\left(f''(v_{x,D}^K)\frac{D^2}{2K^2}\right)\right) - f(x)\right) \\ &= \alpha f(x) + \mathbb{E}(f''(v_{x,D}^K)D^2)/2 + \varepsilon'_K \end{aligned} \tag{4.3}$$

avec $\varepsilon'_K \to 0$ quand $K \to \infty$ (uniformément en x). Or f'' est bornée et Lipschitzienne donc il existe $L \geq 0$ tel que pour tout $x \in \mathbb{R}$,

$$\mathbb{E}\left(\sup_{v\in[x-|D|/K, x+|D|/K]} \left|f''(v) - f''(x)\right| D^2\right) \leq L\mathbb{E}\left(D^2\left(\frac{|D|}{K} \wedge 1\right)\right).$$

De plus $\mathbb{E}\left(D^2\left(\frac{|D|}{K}\wedge 1\right)\right)$ tend vers zero quand K tend vers l'infini par convergence dominée et $\sigma^2=\mathbb{E}(D^2)$ donc

$$\sup_{x\in\mathbb{R}}|\mathbb{E}(f''(v^K_{x,D})D^2)-f''(x)\sigma^2|\overset{K\to\infty}{\longrightarrow}0$$

et la preuve est complète. □

On suppose maintenant que

$$\sup_K K^2\mathbb{E}(L^K(L^K-1))<\infty \tag{4.4}$$

Lemme 67. *Il existe $C\geq 0$ tel que pour tout $f\in\mathcal{C}_b^{2,+}$, pour tout $k\geq 0$,*

$$i)\,|\Delta_k^K(f)|\leq C\,\frac{1+\parallel f\parallel_\infty+\parallel f''\parallel_\infty}{K^2}\,\frac{\#\mathbb{G}^K_{k-1}}{K}$$

$$ii)\sum_{u\in\mathbb{G}^K_{k-1}}|G^K_f(X^K(u))|\leq C\,\frac{1+\parallel f\parallel_\infty^2+\parallel f'\parallel_\infty^2+\parallel f''\parallel_\infty^2}{K}\,\frac{\#\mathbb{G}^K_{k-1}}{K}$$

Démonstration. Tout d'abord, (4.1) assure que

$$|\Delta_k^K|\leq\frac{1}{K}\#\mathbb{G}^K_{k-1}\parallel M^K(.,f)-f(.)\parallel_\infty$$

et on peut utiliser la preuve précédente, à savoir (4.3), qui permet de contrôler $K^2\parallel M^K(.,f)-f(.)\parallel_\infty$. Cela donne $i)$.

Ensuite, en notant $v_K=\mathbb{E}(L^K(L^K-1))$ et en développant

$$\begin{aligned}G^K_f(x)&=\mathbb{E}\left(\left(\sum_{u\in\mathbb{G}^K_1}f(X(u))\right)^2\right)-M^K(x,f)^2\\&=m_K\mathbb{E}\left(f^2(x+D/K)\right)+v_K\mathbb{E}(f(x+D/K))^2-m_K^2\mathbb{E}(f(x+D/K))^2.\end{aligned}$$

Or $v_K=\mathcal{O}(1/K^2)$ par (4.4) et $m_K=1+\mathcal{O}(1/K^2)$ et $m_K^2=1+\mathcal{O}(1/K^2)$ et

$$|\mathbb{E}\left(f^2(x+D/K)\right)-f^2(x)|\leq(\parallel f\parallel_\infty\parallel f''\parallel_\infty+\parallel f'\parallel_\infty^2)/K^2$$

car à nouveau par une inégalité de Taylor :

$$|f^2(x+D/K)-f^2(x)-2f(x)f'(x)D/K|\leq\sup|f'^2+ff''|\,\frac{D^2}{K^2}.$$

Enfin, $|\mathbb{E}(f(x+D/K))-f(x)|\leq\parallel f''\parallel_\infty/K^2$, ce qui donne le résultat en remarquant que $|ff''|\leq|f|^2+|f''|^2$. □

On suppose de plus maintenant que $\#\mathbb{G}^K_0/K$ converge en loi quand K tend vers l'infini vers la v.a. $Z_0\in[0,\infty)$. On va montrer un résulat de bornitude en loi, permettant de "localiser le processus" dans les compacts, puis nous montrerons que la partie martingale tend vers 0.

Lemme 68. *Soit*

$$\tau_N^K = \inf\left\{t \geq 0 : \#\mathbb{G}^K_{[K^2t]}/K \geq N\right\}$$

Alors pour tout $T \geq 0$,

$$\sup_K \mathbb{P}(\tau_N^K \leq T) \overset{N\to\infty}{\longrightarrow} 0$$

et pour tout $N \geq 0$ *et* $f \in \mathcal{C}^{2,+}$,

$$\mathbb{E}\left(\sup_{t\leq T} M^K_{[K^2t]\wedge[K^2\tau_N^K]}(f)^2\right) \overset{K\to\infty}{\longrightarrow} 0.$$

Démonstration. Dans cette preuve, c désignera une constante, qui peut changer mais ne dépendra ni de K, ni de N. On pose $\sigma_N^K = \inf\left\{n \geq 0 : \#\mathbb{G}_n^K/K \geq N\right\}$. $(\#\mathbb{G}_n^K)_n$ est une processus de Galton-Watson et $W_n^K = \#\mathbb{G}_n^K/m_K^n$ est une martingale. En arrêtant cette martingale au temps d'arrêt σ_N^K, on obtient

$$\mathbb{E}\left(W^K_{\sigma_N^K\wedge[K^2T]} \,\middle|\, \#\mathbb{G}_0^K\right) = \#\mathbb{G}_0^K$$

Or sur l'événement $\{\sigma_N^K \leq [K^2T]\}$, $\#\mathbb{G}^K_{\sigma_N^K\wedge[K^2T]} \geq KN$ et $m_K^{\sigma_N^K\wedge[K^2T]} \leq (1+c/K^2)^{[K^2T]} \leq c\exp(cT)$. Donc

$$\mathbf{1}_{\{\sigma_N^K\leq[K^2T]\}} \leq \frac{\#\mathbb{G}^K_{\sigma_N^K\wedge[K^2T]}}{KN}.\frac{c'\exp(c''T)}{m_K^{\sigma_N^K\wedge[K^2T]}} = W^K_{\sigma_N^K\wedge[K^2T]}\frac{c\exp(cT)}{KN}$$

et en prenant l'espérance conditionnelle de part et d'autre

$$\mathbb{P}\left(\sigma_N^K \leq [K^2T] \,\middle|\, \#\mathbb{G}_0^K\right) \leq c_T\frac{\#\mathbb{G}_0^K}{KN}$$

ce qui assure que $\sup_K \mathbb{P}(\sigma_N^K \leq K^2T)$ tend vers 0 quand N tend vers l'infini, en utilisant que $\#\mathbb{G}_0^K/K$ est borné en loi. Ceci prouve la première convergence du lemme. Pour la deuxième, on utilise l'inégalité de Doob dans L^2 pour la martingale arrêtée $\left(M^K_{n\wedge\sigma_N^K}(f)\right)_n$. On obtient

$$\mathbb{E}\left(\sup_{n\leq[K^2T]}\left(M^K_{n\wedge\sigma_N^K}(f)\right)^2\right) \leq c\mathbb{E}\left(\left(M^K_{[K^2T]\wedge\sigma_N^K}(f)\right)^2\right)$$

De plus

$$\mathbb{E}\left(\left(M^K_{[K^2T]\wedge\sigma_N^K}(f)\right)^2\right) = \frac{1}{K^2}\mathbb{E}\left(\sum_{k=1}^{[K^2T]\wedge\sigma_N^K}\sum_{u\in\mathbb{G}_{k-1}^K} G_f^K(X(u))\right)$$

en utilisant la decomposition semimartingale de M^2 obtenue en Proposition 63. Enfin le lemme précédent et la definition du temps σ_N^K assurent que

$$\frac{1}{K^2}\sum_{k=1}^{K^2T\wedge\sigma_N^K}\sum_{u\in\mathbb{G}_{k-1}^K} G_f^K(X(u)) \leq \frac{c}{K^3}\sum_{k=1}^{K^2T\wedge\sigma_N^K}\frac{\#\mathbb{G}_{k-1}^K}{K} \leq \frac{c}{K}.T.N$$

ce qui conclut la preuve. □

4.2.2 Tension de Y^K

On rappelle que

$$Y_t^K = X_{[K^2t]}^K = \frac{1}{K} \sum_{u \in \mathbb{G}_{[K^2t]}^K} X^K(u).$$

L'objectif est de montrer la convergence de Y^K dans l'espace $\mathbb{D}([0,\infty), \mathcal{M})$ des processus càdlàg à valeurs mesures, où $\mathcal{M}$ est l'espace des mesures positives que l'on va munir de la topologie vague : $(\mathcal{M}, v)$. C'est à dire de la topologie la moins fine, rendant continue les applications $\mu \to \mu(f)$, pour $f \in \mathcal{C}_c$. On parlera de convergence étroite dans un second temps, qui est un résultat plus fort, et plus compliqué à vérifier. Pour cela on commence par vérifier la tension, qui donnera par le théorème de Prokhorov de la relative compacité. La tension du processus à valeurs mesures va s'obtenir en vérifiant la tension du processus à valeurs réelles par projection avec des fonction tests. Pour la tension dans $(\mathcal{M}, v)$, il suffit de considérer les fonctions à support compact, et en fait seulement un ensemble dense pour la norme infinie. Comme sous espace dense, nous pourrons considérer l'espace $\mathcal{C}_c^{2,+} = \mathcal{C}_b^{2,+} \cap \mathcal{C}_c$ des fontions de classe C^2 à support compact dont la dérivée second est Lipschitz.

Proposition 69. *Pour tout $T > 0$ et $f \in \mathcal{C}_b^{2,+}$, $(Y^K(f))_K$ est tendue dans $\mathbb{D}([0,T], \mathbb{R})$.*

Démonstration. On applique le critère d'Aldous, qui donne une façon efficace de ramener la tension au contrôle des variations entre deux temps d'arrêt. La décomposition comme somme d'un processus prévisible et d'une martingale donne pour $t \leq T$,

$$|Y_t^K(f)| \leq |Y_0^K(f)| + \sum_{k=1}^{[K^2T]} |\Delta_k^K(f)| + \sup_{t \leq T} |M_{[K^2t]}^K(f)|.$$

Le lemme 67 *i)* et le fait que $\#\mathbb{G}_k^K \leq KN$ avant le temps d'arrêt τ_N^K assure que

$$1_{\tau_N^K > T} \sum_{k=1}^{[K^2T]} |\Delta_k^K(f)| \leq C'.T.N.$$

De plus par l'inégalité de Cauchy Schwarz,

$$\mathbb{E}\left(1_{\tau_N^K > T} \sup_{t \leq T} |M_{[K^2t]}^K(f)|\right) \leq \sqrt{\mathbb{E}\left(\sup_{t \leq T} M_{[K^2t] \wedge [K^2\tau_N^K]}^K(f)^2\right)}$$

qui tend vers 0 quand K tend vers l'infini grâce au lemme 68. Enfin $1_{\tau_N^K > T}|Y_0^K(f)| \leq N \parallel f \parallel_\infty$ et on obtient que pour tout N,

$$\sup_K \mathbb{P}\left(1_{\tau_N^K > T}|Y_t^K(f)| \geq M\right) \overset{M \to \infty}{\longrightarrow} 0.$$

De plus la première partie du Lemme 3 assure que $\sup_K \mathbb{P}\left(1_{\tau_N^K \leq T}|Y_t^K(f)| \geq M\right) \leq \sup_K \mathbb{P}\left(\tau_N^K \leq T\right) \to 0$ quand N tend vers l'infini et on obtient

$$\sup_K \mathbb{P}\left(\sup_{t \leq T} |Y_t^K(f)| \geq M\right) \overset{M \to \infty}{\longrightarrow} 0.$$

De même

$$Y^K_{t+s}(f) - Y^K_t(f) = \sum_{k=[K^2t]+1}^{[K^2(t+s)]} \Delta^K_k(f) + M^K_{[K^2(t+s)]}(f) - M^K_{[K^2t]}(f)$$

et on montre de façon similaire que pour tout $\varepsilon > 0$,

$$\sup_{K, S \leq S' \leq S+\delta} \mathbb{P}\left(|Y^K_{S'}(f) - Y^K_S(f)| \geq \varepsilon\right) \xrightarrow{\delta \to 0} 0,$$

où les temps S et S' sont des temps d'arrêt à valeurs dans $[0, T]$. □

Le résultat précédent suffit à montrer la tension de Y^K dans $\mathbb{D}([0,\infty), (\mathcal{M}, v))$, où $\mathcal{M}$ est dans l'espace des mesures positives que l'on va munir de la topologie vague. Pour cela, on réfère en particulier au critère de tension de Roelly Copoletta [25]. La preuve de la tension de Y^K quand l'espace d'arrivée $(\mathcal{M}, e)$ est muni de la convergence étroite demande plus de travail, et nous reviendrons dessus ensuite. De façon très informelle, pour s'assurer de la convergence étroite, on peut contrôler les grandes valeurs des positions pour ramener l'espace des positions $\mathbb{R}$ à un borné donc à la convergence vague. On pourra consulter [23] pour un critère. Plus généralement pour des conditions suffisantes de de tension dans un espace de dimension infinie (comme l'espace des mesures), on pourra consulter le livre de Ethier et Kurtz [12]. Deux ingrédients sont en jeu en général, comme ici : verifier que notre objet reste dans les compacts (compact containment, qui pose des difficultés en particulier quand on munit l'espace des mesures de la topologie etroite sur un espace non compact), et une condition d'équicontinuité (en réduisant l'espace d'arrivée via des fonctions tests).

4.2.3 Identification de la limite et convergence

Maintenant que l'on sait que l'on peut extraire des sous suites convergentes de Y^K grâce au résultat de tension, il s'agit de vérifier qu'il n'y a qu'une seule limite possible pour ces sous suites.

La décomposition comme somme d'un processus prévisible et d'une martingale donne

$$\begin{aligned}
Y^K_t(f) &= Y^K_0(f) + \sum_{k=1}^{[K^2t]} \Delta^K_k(f) + M^K_{[K^2t]}(f) \\
&= Y^K_0(f) + \sum_{k=1}^{[K^2t]} \int_{\mathbb{R}} Y^K_{k/K^2}(dx) \left(M^K(x,f) - f(x)\right) + M^K_{[K^2t]}(f) \\
&= Y^K_0(f) + \int_0^t ds\, Y^K_s \left(K^2 \left(M^K(x,f) - f(x)\right)\right) + R^K_t(f)
\end{aligned}$$

avec pour tout t,

$$\begin{aligned}
|R^K_t(f)| &\leq \| M(.,f) - f \|_\infty \sup_{s \leq t} Y^K_s(\mathbb{R}) + |M^K_{[K^2t]}(f)| \\
&\leq c\left(\frac{N}{K^2} 1_{\tau^K_N > t} + 1_{\tau^K_N \leq t} \sup_{s \leq t} Y^K_s(\mathbb{R})\right) + |M^K_{[K^2t]}(f)|
\end{aligned}$$

On obtient alors en utilisant le lemme 68 que $\sup_{t\leq T}|R_t^K(f)|$ tend vers 0 en probabilité quand K tend vers l'infini. Reste à gérer la convergence de l'intégrale. Grâce au théorème de représentation de Skorokhod, on peut considérer des sous suites convergeant p.s., identifier la loi de leur limite permettra de conclure à une convergence en loi. Munissons ici l'espace des mesures de la convergence vague. Rappelons que l'espace $\mathcal{C}_c^{2,+}$ est formé des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, de classe $\mathcal{C}^2$ à support compact et dont la dérivée seconde est Lipschitz.

Lemme 70. *Supposons que Y^K converge p.s. vers Y dans $\mathbb{D}([0,\infty),(\mathcal{M},v))$, alors*

$$\sup_{t\leq T}\left|\int_0^t ds\, Y_s^K\left(K^2\left(M(.,f)-f\right)\right)-\int_0^t ds\, Y_s(\mathcal{L}f)\right| \overset{K\to\infty}{\longrightarrow} 0$$

en probabilité, pour tout $f\in\mathcal{C}_c^{2,+}$.

Démonstration. Une bonne vieille inégalité triangulaire nous permet d'écrire la différence à faire converger comme $A_t^K+B_t^K$ avec

$$\begin{aligned}A_t^K &= \int_0^t ds\, Y_s^K\left(K^2\left(M(.,f)-f\right)-\mathcal{L}f\right)\\ B_t^K &= \int_0^t ds\,\left(Y_s^K(\mathcal{L}f)-Y_s(\mathcal{L}f)\right).\end{aligned}$$

Le lemme 66 assure la convergence en probabilité vers 0 de $\sup_{t\leq T}|A_t^K|$ quand $K\to\infty$, en utilisant aussi le lemme 68 pour dominer $Y_s^K(\mathbb{R}^+)$ en faisant intervenir le temps d'arrêt τ_N^K :

$$|A_t^K|\leq T\parallel\left(K^2\left(M(.,f)-f\right)-\mathcal{L}f\right)\parallel_\infty N+\mathbf{1}_{\tau_N^K\leq T}A_t^K.$$

La convergence p.s. de Y^K vers Y dans $\mathbb{D}([0,\infty),(\mathcal{M},v))$ permet d'obtenir la convergence p.s. de $\sup_{t\leq T}|B_t^K|$ vers 0 car $\mathcal{L}f$ est continue et tend vers 0 à l'infini et donc

$$(\mu_t)_{t\in[0,T]}\to\left(\int_0^t ds\mu_s(\mathcal{L}f)\right)_{0\leq t\leq T}$$

est continue de $\mathbb{D}([0,\infty),(\mathcal{M},v))$ dans $(\mathcal{C}([0,T],\mathbb{R}),\parallel.\parallel_\infty)$ (on pourra consulter [8]). □

On peut remarquer qu'un résultat analogue vaudrait pour l'espace des mesures muni de la toplogie étroite.

Exploitons la tension dans $\mathbb{D}([0,T],(\mathcal{M},v))$ obtenue dans la section précédente, de sorte que Y^K a des sous suites convergeant en loi dans cet espace. Une représentation de Skorokhod permet de travailler avec une convergence p.s. et les résultats de cette section assurent que chaque terme de

$$Y_t^K(f)=Y_0^K(f)+\int_0^t ds\int_{\mathbb{R}}Y_s^K(dx)K^2\left(M(x,f)-f(x)\right)+R_t^K(f)$$

converge en probabilité si la condition initiale converge. Alors, en appelant Y_0 la limite faible de la condition initiale, le processus limite Y d'une telle sous suite est forcement solution de l'équation

$$Y_t(f)=Y_0(f)+\int_0^t ds\, Y_s(\mathcal{L}f),$$

pour tout $f \in \mathcal{C}_c^{2,+}$. Considérons le cas où $Y_0 = \mu_0$ est déterministe, ce qui revient à conditionner par l'état initial. Nous avons obtenu que les valeurs d'adhérences de Y^K and $\mathbb{D}([0,T],(\mathcal{M},v))$ sont solutions de l'équation

$$\mu_t(f) = \mu_0(f) + \int_0^t ds \int_{\mathbb{R}} \mu_s(\mathcal{L}f)(x). \tag{4.5}$$

pour $f \in \mathcal{C}_c^{2,+}$. En particulier, cette équation admet bien une solution dans $\mathbb{D}([0,T],(\mathcal{M},v))$.
L'objectif est maintenant de vérifier qu'il y a une seule valeur d'adhérence et obtenir ainsi une convergence. De plus, nous allons vouloir renforcer le résultat en munissant l'espace des mesures de la convergence étroite : $(\mathcal{M},e)$. D'ailleurs, si la question de la tension se complique alors, cela donne une façon de simplifier le problème d'unicité en étendant l'espace des fonctions tests pour laquelle l'équation est satisfaite. Ainsi

Lemme 71. *Soit μ_0 une mesure de masse finie sur $\mathbb{R}$. Il existe un unique $\mu \in \mathbb{D}([0,T],(\mathcal{M},e))$ qui vérifie (4.5) sur $\mathcal{C}_b^{2,+}$ et a pour valeur au temps 0 la mesure μ_0.*

Démonstration. On commence par vérifier qu'une telle solution μ vérifie plus généralement

$$\mu_t(f_t) = \mu_0(f_0) + \int_0^t ds\, \mu_s\left(\mathcal{L}f_s + \partial_s f_s\right), \tag{4.6}$$

pour $f \in \mathcal{C}_b^{2,2,+}$, qui est l'ensemble des fonctions continues de $[0,\infty)\times\mathbb{R}$ dans $\mathbb{R}$ telle que pour tout $t \geq 0$, $f_t(.) \in \mathcal{C}_b^{2,+}$ et pour tout $x \in \mathbb{R}$, $f_.(x)$ est dérivable et $\partial_t f_t(x)$ est continue et bornée sur $[0,\infty)\times\mathbb{R}$. On peut établir cela soit en reprenant la décomposition en processus prévisible et martingale mais en intégrant maintenant le temps, soit en exploitant la bilinéarité de $(\mu,f) \to \mu(f)$ pour montrer que

$$\partial_t \mu_t(f_t) = [\partial_t \mu_t](f_t) + \mu_t(\partial_t f_t).$$

en calculant le taux d'accroissement et en utilisant la bilinéarité de $(\mu,f) \to \mu(f)$.

Ensuite on considère le semigroupe du mouvement brownien B, $P_t f(x) = \mathbb{E}_x(f(B_t))$, qui peut d'ailleurs se calculer explicitement ici (mais la méthode est générale). Le générateur du mouvement brownien vaut $\mathcal{A}f(x) = \sigma^2 f''(x)/2$ et en posant $f_s(x) = P_{t-s}f(x)$, on obtient $\partial_s f_s(x) = -\mathcal{A}f_s(x)$ et donc $\mathcal{L}f_s + \partial_s f_s = \alpha f_s$. De plus pour $f \in \mathcal{C}_0^{2,b}$, $f_. \in \mathcal{C}_0^{2,2,+}$. En considérant deux solutions μ et ν de (4.6) appliquée à cette fonction, la différence $\gamma = \mu - \nu$ vérifie

$$\gamma_t(f_t) = \gamma_t(f) = \alpha \int_0^t ds\, \gamma_s(f_s).$$

Par ailleurs, introduisons la distance en variation totale

$$\| \gamma \|_{V.T.} = \sup_{\|f\|_\infty \leq 1} |\gamma(f)| = \sup_{f \in \mathcal{C}_b^{2,+}, \|f\|_\infty \leq 1} |\gamma(f)|.$$

En utilisant que f_s est borné par $\| f \|_\infty$ et en optimisant sur $f \in \mathcal{C}_0^{2,+}, \| f \|_\infty \leq 1$, on obtient

$$\| \gamma_t \|_{V.T.} \leq \alpha \int_0^t ds \, \| \gamma_s \|_{V.T.} \, .$$

Le lemme de Gronwall assure que $\| \gamma_t \|_{V.T.} = 0$, ce qui achève la preuve de l'unicité. □

Pour conclure à la convergence, il reste donc à démontrer que Y^K est tendue dans $\mathbb{D}([0,T],(\mathcal{M},e))$ et invoquer le résultat d'unicité précédent. Pour cela, on peut utiliser le théorème 3 de [23]. Ainsi, on sait déjà que que $(Y^K_{t\in[0,T]})-k$ est tendue $\mathbb{D}([0,T],(\mathcal{M},v))$ et on peut extraire des sous suites convergentes pour cette topologie. Pour s'assurer que la convergence a bien lieu dans $\mathbb{D}([0,T],(\mathcal{M},e))$, ce théorème guarantit qu'il suffit alors de montrer que

— $(Y^K(\mathbb{R})_{t\in[0,T]})_K$ (ou une sous suite) converge dans $\mathbb{D}([0,T],\mathbb{R})$

— les points limites de Y^K dans $\mathbb{D}([0,T],(\mathcal{M},v))$ sont dans $\mathbb{C}([0,T],(\mathcal{M},v))$

Le premier point se démontre avec les mêmes techniques (tension, identification). On peut utiliser les estimations précédentes, en remarquant que le problème est plus simple car nous ne sommes plus dans l'espaces des mesures, mais dans $\mathbb{R}$. On peut d'ailleurs profiter ici en plus du fait que $Y^K(\mathbb{R})$ est un processus de branchement classique simple type et directement identifier la convergence fini dimensionnelle, qui est une consésquence de la loi des grands nombres, en exploitant l'indépendance des différentes sous populations initiales. Le deuxième point demande de controler les grandes valeurs que peuvent prendre les positions des individus, ce que l'on peut faire en faisant des calculs de moments. On renvoie par exemple à la preuve du Theorème 5.4 dans [7] pour des détails.

Théorème 72. *Supposons que Y^K_0 converge étroitement vers μ_0 mesure de masse finie quand $K\to\infty$ et que les hypothèses de moment (4.2) et (4.4) sont satisfaites. Alors Y^K converge en loi dans l'espace $\mathbb{D}([0,T],(\mathcal{M},e))$ vers l'unique solution Y de*

$$Y_t(f)=Y_0(f)+\int_0^t ds\, Y_s(\mathcal{L}f) \qquad (f\in\mathcal{C}_b^{2,+}).$$

Notons que quand Y_t est à densité (ce que l'on peut montrer quand la condition initiale est à densité), i.e. $Y_t(dx)=u_t(x)dx$, u_t est alors solution d'une EDP (réaction diffusion linéaire), que l'on obtient en faisant des intégrations par parties dans l'équation ci dessus :

$$\partial_t u_t=\alpha u_t+\sigma u_t''$$

Bibliographie

[1] K. B. Athreya, P. E. Ney (2004). *Branching processes.* Dover Publications, Inc., Mineola, NY.

[2] K. B. Athreya, S. Karlin (1971). On branching processes with random environments I : Extinction probability. *Ann. Math. Stat.* 42. 1499-1520.

[3] K.B. Athreya (2000). Change of Measures for Markov Chains and the LlogL Theorem for Branching Processes. *Bernoulli* 6(2). DOI : 10.2307/3318579.

[4] V. Bansaye (2008). Proliferating parasites in dividing cells : Kimmel's branching model revisited. *Annals of Applied Probability*, Volume 18, Number 3, 967-996.

[5] V. Bansaye. Cell contamination and branching process in random environment with immigration, *Advances in Applied Probability*, Volume 41, Number 4, 2009.

[6] V. Bansaye, A. Lambert (2013) New approaches to sourcesink metapopulations decoupling demography and dispersal. *Theoret. Popul. Biol.* 88 31-46.

[7] V. Bansaye, S. Méléard. Some stochastic models for structured populations : scaling limits and long time behavior. Springer, book available via https ://arxiv.org/abs/1506.04165.

[8] P. Billingsley. Weak convergence of probability measures.

[9] M. E. Caballero, A. Lambert and G. Uribe Bravo. Proof(s) of the Lamperti representation of continuous-state branching processes *Probab. Surveys* 6 62–89, 2009.

[10] Cowan, R., and Staudte, R. G. (1986). The bifurcating autoregression model in cell lineage studies, Biometrics, 42, 769-783.

[11] A. Dembo, O. Zeitouni (1998). Large deviations techniques and applications. Applications of Mathematics (New York) 38 (Second edition ed.).

[12] S. N. Ethier and T. G. Kurtz. Markov processes. Characterization and convergence. Wiley Series in Probability and Mathematical Statistics : Probability and Mathematical Statis- tics, New York, 1986.

[13] J. Geiger, G. Kersting, V. A. Vatutin (2003). Limit theorems for subcritical branching processes in random environment. *Ann. Inst. Henri Poincarè* 39, no. 4. 593-620.

[14] J. Guyon (2007). Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging, Ann. Appl. Probab., 17, 15381569.

[15] G. Keller, G. Kersting, and U. Rosler. On the asymptotic behaviour of discrete time stochastic growth processes. Ann. Probab., 15(1) :305-343, 1987.

[16] G. Kersting. On recurrence and transience of growth models. J. Appl. Probab., 23(3) :614-625, 1986.

[17] T. Kurtz, R. Lyons, R. Pemantle, Y. Peres (1997). A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes. In *Classical and Modern Branching Processes*, ed. K. B. Athreya and P. Jagers. Springer, New York. 181-185.

[18] R. Lande (1993). Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *The American Naturalist* 911–927.

[19] R. Lande, S. Engen, B-E Saether (2003). *Stochastic Population Dynamics in Ecology and Conservation.* Oxford University Press.

[20] J. Lamperti. Continuous state branching processes. *Bull. Amer. Math. Soc.*, 73, 382–386, 1967.

[21] J. Lamperti. The limit of a sequence of branching processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 7 :271–288, 1967.

[22] R. Lyons, R. Pemantle, Y. Peres (1995). Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Probab.* 23 (3), 1125-1138.

[23] S. Méléard and S. Roelly. Sur les convergences étroite ou vague de processus ‘a valeurs mesures. *C. R. Acad. Sci.* Paris Sér. I Math. 317, 785-788, 1993.

[24] B. Porro, N. Alloncle, N. Bierne. Sophie Arnaud-Haond. Connectivité et protection de la biodiversit marine : Dynamique spatiale des organismes marins.*https ://hal.archives-ouvertes.fr/hal-02404006.*

[25] S. Roelly-Coppoletta. A criterion of convergence of measure-valued processes : application to measure branching processes. *Stochastics*, 17(1- 2) :4365, 1986.

[26] Stewart, E. J., Madden, R., Paul, G., and Taddéi, F. (2005). Aging and death in an organism that reproduces by morphologically symmetric division, PLoS Biol, 3, 2 : e45.