

Processus de branchement pour des populations structurées

Vincent Bansaye.
Notes provisoires.

Table des matières

1	Arbre de Galton-Watson	6
1.1	Processus de Galton-Watson	6
1.1.1	Extinction	7
1.1.2	Transformation de Lamperti	8
1.2	Arbre aléatoire et construction spinale	10
1.2.1	Construction spinale	11
1.2.2	Changement de probabilité	12
1.3	Processus de Galton-Watson avec immigration	14
1.3.1	Croissance dans le cas surcritique	16
1.3.2	Les régimes critique et sous-critique	17
1.4	Croissance malthusienne et vitesse d'extinction	18
1.4.1	Croissance dans le cas surcritique	18
1.4.2	Vitesse d'extinction dans le cas sous-critique	21
1.5	Processus de Galton-Watson en environnement aléatoire	22
1.5.1	Critère d'extinction	23
1.5.2	Effet de l'environnement sur la vitesse d'extinction	24
2	Populations structurées : le cas neutre	25
2.1	Chaîne de Markov sur un arbre de Galton-Watson	25
2.1.1	Définition du modèle et exemples	25
2.1.2	Propriété de branchement et semigroupe	28
2.1.3	Formule tous-pour-un et fourches	30
2.1.4	Loi des grands nombres	36
2.2	Transitions de phase et division cellulaire	38
2.2.1	Processus autorégressif bifurquant	38
2.2.2	Infection dans le modèle de Kimmel	40
2.2.3	Contamination et division cellulaire	43
2.3	Vitesse d'invasion dans un modèle spatial	44
2.3.1	Préliminaires de grandes déviations	45
2.3.2	Vitesse d'invasion	47
3	Branchement pour des populations structurées	51
3.1	Chaîne de Markov branchante	51
3.1.1	Définition	51
3.1.2	Exemple structuré en âge et régénération	52
3.2	Semigroupe et formules tous-pour-un	55
3.2.1	Formule de Feynman Kac et principe variationnel	56

3.2.2	Lignée typique et processus inhomogène	58
3.2.3	Fourches pour échantillonner deux individus	58
3.3	Utilisation de fonctions harmoniques	60
3.3.1	Exemple de modèle structuré en âge	62
3.3.2	Exemple de modèle structuré en taille	63
3.4	Espace de trait fini et construction spinale	66
3.4.1	Temps long et vitesse de convergence avec variance finie	68
3.4.2	Temps long et construction spinale avec un moment $L \log L$	70
3.5	Ergodicité et temps long	71
3.5.1	Ergodicité et fonction de Lyapunov	72
3.5.2	Conditions de Doeblin et uniform ergodicité	73
4	Approximation en grande population	77
4.1	Généralités et décomposition semi-martingale	77
4.1.1	Définitions	77
4.1.2	Décomposition en semi-martingale	78
4.1.3	Lignées ancestrales	81
4.2	Convergence vers une EDP pour un modèle de reproduction dispersion	82
4.2.1	Préliminaires	83
4.2.2	Tension de la suite de processus	86
4.2.3	Identification de la limite et convergence	88

Introduction

This course focuses on branching processes, which are random models used to describe various population dynamics when individuals evolve independently of each other. The historical, prototype model is the Galton-Watson process, where individuals are interchangeable and evolve in a fixed environment. In this monograph, we will study extensions of this model where the environment can change, including randomly (Chapter 1), and where individuals are characterized by traits (Chapters 2,3,4). These traits can be complex and account for the position of the individual, its size, age, phenotype, genotype, etc. They are transmitted to their descendants and can evolve deterministically or randomly between the successive generations. In the case of a neutral model (Chapter 2), the trait does not influence survival or reproduction, but many questions involve structuring the population to take into account traits that affect survival or reproduction (Chapter 3). We will also study (Chapter 4) the approximation of these models in a large population regime. This allows to make emerge macroscopic simpler object, in particular deterministic approximations by PDE.

Thus, we will generally assume the *branching property*. It requires, once the population is structured, that individuals do not interact and that their evolution depends only on their traits and the environment. This regime appears especially when resources are abundantly available, for the study of invasive systems (epidemics, cancer, the emergence of an adapted mutant...), and more generally in certain regimes where, locally, the branching approximation is relevant. Nevertheless, some results are generalized. If the complications can go beyond technical details, the consideration of density dependence (competition, cooperation, infection...) is possible. It will be discussed, especially for large population approximations in the last chapter or for the different notions introduced on ancestral lineages before. We refer to [10, 46] for recent monographs in this vein.

The models and processes are defined here in discrete time, describing the state of the population across successive generations. In biology, generation time can have various interpretations, such as the lifespan of an organism, the time required for cell division, annual cycles for seasonal processes, circadian rhythms, or the timeframe of infections. Working in discrete time allows us to define general concepts and examples recursively and straightforwardly. It will allow us to introduce the key concepts, to establish connections between objects, to explore various phenomena and to provide complete proofs without requiring a strong background in continuous or discontinuous stochastic calculus. While the key concepts presented in this monograph are simpler in discrete time and can be extended to continuous time, the latter introduces also new mathematical challenges and behaviors, see in particular [39] for branching processes. Furthermore, continuous time models are also very significant in mathematical biology and applications. In the last chapter, the approximation results are leading to continuous

time models. We will also make more comments and references to continuous time counterparts along the monograph.

The literature on branching processes is very rich, and we begin by citing some reference books. The books by Athreya and Ney [2] and Harris [27] and Asmussen and Hering [1] provide a nice introduction and clear view of the long-term analysis of Galton-Watson processes. They also contain the first important results on population structuring through generalizations to multitype processes. For a general overview on branching process with a finite number of types, let us also mention the book of Mode [44]. Besides, these references address the transition to continuous time or consideration of a changing environment. For initial questions such as criterion for survival and growth on the survival event, branching processes are often guided by their average behavior, that is, the first moment semigroup. Roughly, it boils down the study of these questions to a linear operator and martingale arguments. The study of the linear operator can be direct, for example, for the Galton-Watson process, or inherited from analysis, such as the Perron-Frobenius theorem for a finite number of types. To go farther, in particular to characterize the survival probability or some limiting distributions, the theory was essentially based on the generating function of the number of individuals living in a given generation. Indeed, the branching property is reflected in a composition of generating functions, which allows to study the distribution of the number of individuals in successive generations.

Starting from the 1980s, the literature increased significantly, and for an overview of the results obtained, one can refer to the reviews by Vatutin and Zubkov [50, 51], which point to a large bibliography for proofs and additional results, at the end of 80s. At the same time, the applications of branching processes in biology and ecology multiplied : cellular dynamics and DNA replication, population survival, fitness and evolution, viruses and resistance, epidemiology, immunology, PCR... We particularly invite the reader to consult Kimmel and Axelrod [31] and Haccou, Jagers, and Vatutin [25].

There is no longer a global, exhaustive reference on branching processes, but several nice books focus on a part or a viewpoint. Without seeking to be exhaustive, we mention Li [39] for continuous-state and time branching processes (CSBP), measured valued branching Markov process and superprocesses. We refer to Kersting and Vatutin for single type branching processes in random environment [30]. Let us also mention Shi [48] for branching random walks and front advancement, Bertoin [14] for fragmentation processes, Lambert for genealogies of branching processes [34]. For neutral branching process in continuous time and space and the approximation of structured population with interaction, let us refer to [10], which shows applications to evolution and cell division models.

This course focuses on two interconnected aspects of branching processes that have significantly evolved : the long-time study of populations structured by an infinite set of traits and the characterization of ancestral lineages within these populations. The transition to an infinite trait space brings difficulties and interesting phenomena. Indeed, the evolution of the trait can be recurrent or transient, the convergence rates of distributions sub-exponential, and the semigroup of the mean no longer operates in finite dimension. Ancestral lineages focus on the evolution of individuals present within the population and touch on important notions in biology : life history, common ancestor, pedigree, reproductive value...The description of ancestral lineages is also valuable for

several reasons in mathematical analysis, including sampling, inference, and model simulation. Moreover it offers a very fruitful entry point into the long-term analysis of the population : growth rate of the total population, emerging profile of traits in the population, extreme individuals... As we will see in this course, the analysis of the process (and its mean operator) in infinite dimension and the study of ancestral lineages turn out to be intimately linked.

The mathematical description of ancestral lineages exploits one or more individuals who survive and can account for the overall population. These individuals are often referred to as the spine, around which the rest of the population can be constructed. These spinal techniques rely on the fact that the process describing the spine's dynamics is forward, Markovian, allowing us to exploit results on stability and coupling, fluctuations, large deviations, or even symmetry principles of these Markovian objects. These processes live in a much simpler space (the trait space) than the initial Markov process, which is a random collection of traits, possibly represented by a measured valued process. These techniques also rely on the fact that we can move from the spine to the original model to obtain interesting results on the latter. We will see different applications, illustrating both various spinal constructions and these different aspects.

These approaches provide insights into the genealogical structure of individuals and pathwise results that complements the use of generating functions and has interesting biological interpretations. They happen to be particularly fruitful for the study in infinite dimension (infinite trait space), since techniques relying on generating functions are difficult to generalize.

Let us conclude this introduction by further motivating the framework of this manuscript from the point of view of models in biology. Firstly, regarding the interest in an infinite trait space framework. In my experience, one can often describe the biological problem by limiting it to a finite space. But an infinite space proves interesting in various situations. Already when the natural trait space is continuous, for example, a size, a position. Or simply because we want to use a continuous random variable, for example, the Gaussian framework, and we will see an example with autoregressive models. Moreover, when the trait space is large at the scales considered, it becomes infinite in the model. Finally, well-chosen infinite spaces can bring structure to objects and spinal construction ; we will see an interesting example with the infection model and, more generally, branching models within branching.

We have chosen here to illustrate concepts and techniques with simple examples that highlight different phenomena and various fields of biology : spatial models on unbounded domains or graphs, cellular dynamics of a population structured in size or infection load and age structuring. The technics can be used with more complex trait spaces, in particular following several traits can be relevant for applications.

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

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

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

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 supposons par la suite que

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

Nous supposons 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-sûrement 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 où 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 $\mathbb{P}(W = 0) = \mathbb{P}(\exists n \in \mathbb{N}, Z_n = 0)$.

1.1.1 Extinction

Avant cela, nous allons déjà déterminer 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 $n \geq 0$,

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

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

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

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

Démonstration. La preuve se fait par récurrence :

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

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

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

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

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

Démonstration. On pose $u_n = \mathbb{P}(Z_n = 0) = f_n(0)$ et on remarque 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 $f : f(p_{ext}) = p_{ext}$. En 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 $\ell = p_{ext}$. $\square$

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 p.s. Or sa moyenne est préservée : $\mathbb{E}(Z_n) = 1$ pour tout $n \geq 0$. Une explication ? Il y a une question de convergence (non) dominée derrière. 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 bissectrice 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 dérivée 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 coïncident on $[p_{ext}, 1]$, et doivent donc être égales à f . Par analyticit  de f , elles coïncident 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  viter l'analyse complexe et remarquer que si $p_{ext} < 1$, alors f n'est alors pas strictement convexe au voisinage de 1, donc $f''(s) = \sum_{k \geq 2} p_k k(k-1)s^{k-2}$ s'annule au voisinage de 1 et p_k est nulle pour $k \geq 2$. $\square$

1.1.2 Transformation de Lamperti

D finissons pour $n \geq 0$,

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

avec la convention $A_0 = 0$, l'effectif cumul  de la population sur n g n rations. Cette quantit  est croissante et on pose

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

la population cumul e. On peut d j  noter que l'extinction en temps fini de la population co incide avec l' v nement A_∞ est fini.

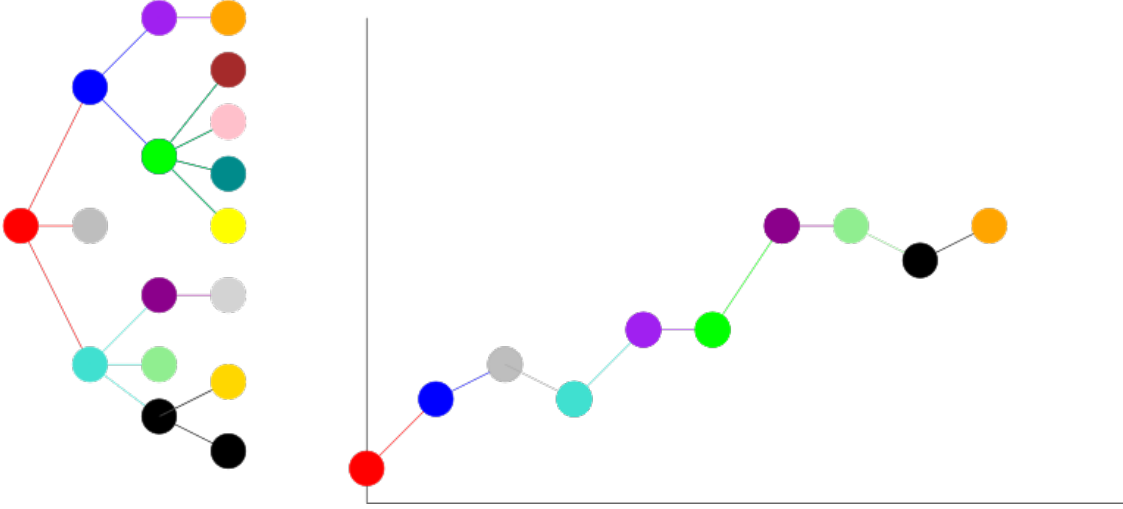


FIGURE 1.1 – Transformée de Lamperti associée à un arbre.

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 $k \in \llbracket A_n, A_{n+1} - 1 \rrbracket$:

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

La construction itérative de S ci dessus définit bien S de façon univoque car les intervalles successifs $\llbracket A_n + 1, A_{n+1} \rrbracket$ forment un recouvrement de $\{n \geq 1, n \leq A_\infty\}$ 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 $A_N = A_{N+1} = \dots = A_\infty$. 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_∞ . $\square$

On obtient directement

Corollaire 9. *Les événements suivants coïncident 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, $\mathbb{P}(\sum_{n \geq 0} Z_n = k) = \mathbb{P}(A_\infty = k) = \mathbb{P}(\inf\{n \geq 0 : S_n = 0\} = k)$. 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 être opérée en temps et même en espace continu pour les processus de branchement [36]. Elle permet d'étudier les approximations en grande population des processus de branchement, en les ramenant à la convergence des marches aléatoires vers les processus de Lévy [35, 17]. 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 [10].

1.2 Arbre aléatoire et construction spinale

On va utiliser les notations de Ulam Harris Neveu :

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

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 $u \preceq v$.

Définition 10 (Arbre). *Un arbre t est une partie de $\mathcal{U}$ telle que*

— *si $v \in t$ et $u \preceq v$, alors $u \in t$ [stabilité par ancêtre],*

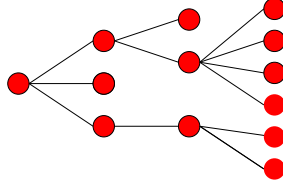


FIGURE 1.2 – Arbre $t = \{\emptyset, 1, 2, 3, 11, 12, 31, 121, 122, 123, 124, 311, 312\}$. Il est associé à la réalisation $L(\emptyset) = 3, L(1) = 2, L(2) = 0, L(3) = 1, L(11) = 0, L(12) = 4, L(31) = 2$ de la famille $(L(u))_{u \in \mathcal{U}}$.

- si $uk \in t$ avec $k \in \mathbb{N}^*$ et $u \in \mathcal{U}$, alors pour tout $1 \leq k' \leq k$, $uk' \in t$ [ordonnancement des frères et soeurs].

On appelle hauteur de l'arbre la longueur maximale des mots de l'arbre. On note $\mathbb{T}$ l'ensemble des arbres, $\mathbb{T}_n$ l'ensemble des arbres finis et 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 existe $u_i \in \mathbb{N}^*$ tel que $u = u_1 \dots u_k$. On dit que $|u|$ est la génération de l'individu u .

Considérons $(L(u))_{u \in \mathcal{U}}$ une famille de variables aléatoires 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

- $\mathbb{G}_0 = \{\emptyset\}$
- $\mathbb{G}_{n+1} = \{uk : u \in \mathbb{G}_n \text{ et } k \in \llbracket 1, L(u) \rrbracket\}$ pour $n \geq 0$.

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 ?

Attention aux glissades, pensez par exemple à une population où le nombre moyen d'enfants par couple est de 0.9... La suite va expliquer les biais d'échantillonnage associés à ce type de questions.

1.2.1 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. On peut ainsi 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 $\hat{L}$ biaisée par le nombre de descendants. Ces idées remontent au moins à Kesten, et la construction que l'on fait ici est due à Lyons, Pemantle et Peres [41].

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

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

Soit $(\hat{L}_n)_{n \in \mathbb{N}}$ une suite de v.a. i.i.d. de loi $\hat{p}$, 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 $(K_n)_{n \in \mathbb{N}}$, indépendante de $(L(u))_{u \in \mathcal{U}}$, avec K_n uniformément distribuée sur $\llbracket 1, \hat{L}_n \rrbracket$. Plus précisément on demande à ce que pour tout $n \in \mathbb{N}$, $\hat{L}_n$ est indépendante de $(\mathcal{A}_n, E_n)$ et

$$\mathbb{P}(K_n = k | \hat{L}_n) = \frac{\mathbf{1}_{1 \leq k \leq \hat{L}_n}}{\hat{L}_n}.$$

La construction se fait à nouveau par récurrence pour $n \geq 0$

$$\begin{aligned} - \mathbb{V}_0 &= \{\emptyset\}, E_0 = \emptyset. \\ - \mathbb{V}_{n+1} &= \bigcup_{v \in \mathbb{V}_n - \{E_n\}} \{vk : 1 \leq k \leq L(v)\} \cup \{E_n k : 1 \leq k \leq \hat{L}_n\}, \quad E_{n+1} = E_n K_n. \end{aligned}$$

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 biaisée par la taille.

1.2.2 Changement de probabilité

On introduit maintenant les arbres tronqués à la génération n

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

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 $t = \{\emptyset\}$ (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 $t = t' \cap (\cup_{k=0}^n (\mathbb{N}^*)^k)$ la restriction de t' aux n premières générations, et $e \in g_n(t) = g_n(t')$ la maman de e' . Par hypothèse de récurrence et construction de $(\mathcal{A}, E)$, on a alors :

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

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 puis choisir le bon parmi ces

enfants avec un choix uniforme :

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

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

Soit $n \in \mathbb{N}$ et U_n une variable aléatoire, qui est un choix uniforme parmi les individus de la génération n . Conditionnellement à $\mathcal{T}_n$ et $\mathbb{G}_n \neq \emptyset$, elle est ainsi distribuée uniformément sur $\mathbb{G}_n$, i.e.

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

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

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

où on rappelle que $(W_n)_n = (Z_n/m^n)_n$ est une martingale.

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

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

$\square$

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 $k \in \mathbb{N}$, $F_k^n \xrightarrow[n \rightarrow \infty]{} \hat{p}_k$ en probabilité.*

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

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

avec t un arbre de taille n et $k_t(u)$ le nombre d'enfants de u dans t et $e \in g_n(t)$. On remarque que

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

Comme $(\hat{L}_i)_{i \in \mathbb{N}}$ est une suite i.i.d. de variables aléatoires, la loi des grands nombres appliquée aux variables i.i.d. intégrables $(1_{\hat{L}_i=k})_{i \in \mathbb{N}}$ assure que cette fréquence tend p.s. (et en proba) vers $\mathbb{E}(1_{\hat{L}=k}) = \hat{p}_k$. Donc $F_{k,\varepsilon}(\mathcal{A}_n, E_n)$ tend vers 0 dans L^1 .

Le théorème précédent implique alors que $\mathbb{E}(F_{k,\varepsilon}(\mathcal{T}_n, U_n)W_n) = \mathbb{E}(F_{k,\varepsilon}(\mathcal{A}_n, E_n))$. Donc $F_{k,\varepsilon}(\mathcal{T}_n, U_n)W_n$ tend vers 0 dans L^1 et en probabilité. Or $W_n \rightarrow W$ p.s. et donc 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,

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

d'où la convergence annoncée. $\square$

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 $Y_0 = 0$ individu p.s.

Exercice. Soit $n \geq 0$ et $\mathcal{F}_n = \sigma(Y_0, \dots, Y_n)$.

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 :

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

Calculer $\mathbb{E}(s^{Y_{n+1}} | \mathcal{F}_n)$ 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_∞ quand n tend vers l'infini, 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 $((Z_{j,n}^{(x)})_{j \leq n}, x \in \mathbb{N}, j \geq 1)$ de processus de Galton-Watson indépendants et indépendants de $(I_n)_{n \geq 1}$, avec $Z_{j,n}^{(x)}$ la taille au temps n 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. de reproduction, on pourrait construire les processus $((Z_{j,n}^{(x)})_{j \leq n}, x \in \mathbb{N}, j \geq 1)$ de telle sorte que l'égalité précédente soit presque sûre, et pas seulement en loi.

Démonstration. Il n'y a 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 d'écrire 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 à écrire à chaque temps n la population Y_n comme le cumul de ces différentes contributions. On somme donc $Z_{j,n}^{(I_j)}$ pour $j \leq n$. $\square$

Dans le cas surcritique $m > 1$, la probabilité de survie d'une vague migratoire est positive, et quand $n \rightarrow \infty$, 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. indépendantes définies par

$$\mathcal{Z}_j \stackrel{\text{a.s.}}{=} Z_{0,j}^{(I_j)} \stackrel{\text{law}}{=} Z_{n-j,n}^{(I_{n-j})},$$

on déduit par retournement du temps que pour tout $n \geq 0$,

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

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 assurée 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 tend p.s. vers l'infini. Nous cherchons maintenant à savoir à quelle vitesse.

Proposition 18. *Conditionnellement à $(I_k)_{k \geq 1}$, le processus $(Y_n/m^n)_{n \in \mathbb{N}}$ est une sous-martingale. De plus pour tout $n \geq 0$,*

$$\mathbb{E} \left[\frac{Y_n}{m^n} \mid (I_k)_{k \geq 1} \right] = \sum_{k=1}^n \frac{I_k}{m^k}.$$

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

Démonstration. Nous pouvons utiliser la définition du processus par récurrence : et en notant $\mathcal{I} = (I_k)_{k \geq 1}$,

$$\mathbb{E}(Y_{n+1} \mid \mathcal{F}_n, \mathcal{I}) = I_{n+1} + \sum_{i=1}^{Y_n} \mathbb{E}(L_{i,n} \mid \mathcal{F}_n, I) = I_{n+1} + mY_n.$$

Nous obtenons

$$\mathbb{E} \left(\frac{Y_{n+1}}{m^{n+1}} \mid \mathcal{F}_n, \mathcal{I} \right) = \frac{I_{n+1}}{m^{n+1}} + \frac{Y_n}{m^n} \geq \frac{Y_n}{m^n}.$$

ce qui prouve le résultat □

Notons qu'on pourrait aussi faire la preuve en utilisant la représentation de la Proposition 17, puisque

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

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

- i) *Si $\mathbb{E}(A) < \infty$, alors $\overline{\lim}_{n \rightarrow \infty} \frac{A_n}{n} = 0$ 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 $\overline{\lim}_{n \rightarrow \infty} \frac{A_n}{n} = \infty$ 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 $\sum_{n \in \mathbb{N}} \mathbb{P}(A/\varepsilon \geq n) = \sum_{n \in \mathbb{N}} \mathbb{P}(A_n/n \geq \varepsilon) < \infty$. Par Borel-Cantelli, on en déduit que $\lim_n A_n/n \leq \varepsilon$. En particulier, avec ε assez petit, $ce^\varepsilon < 1$ et $e^{A_n} \leq (e^\varepsilon)^n$ à partir d'un certain rang. On montre le deuxième point de façon analogue. Cette fois, $\sum_{n \geq 1} \mathbb{P}(A_n/K \geq n) = \infty$ pour tout $K > 0$ et Borel-Cantelli implique $\mathbb{P}(\overline{\lim}_{n \rightarrow \infty} A_n/n \geq K) = 1$ 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 dû à Seneta (1970), avec $\log^+(x) = \log(x)$ pour $x \geq 1$ et $\log^+(x) = 0$ pour $x < 1$.

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, $\lim_{n \rightarrow \infty} Y_n/m^n = \infty$ 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 entraîne 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 $\lim_{n \rightarrow \infty} I_n/m^n = \infty$ puisque le lemme 19 ii) garantit $\lim_{n \rightarrow \infty} \log(I_n)/n = \infty$. $\square$

1.3.2 Les régimes critique et sous-critique

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 croît 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

$$f(x) = \mathbb{E}(\log(Y_{n+1})|Y_n = x) - \log(x) = \mathbb{E}\left(\log\left(1 + \sum_{i=1}^x (L_{i,n} - 1)/x + I_n/x\right)\right).$$

on peut vérifier que

$$\lim_{x \rightarrow \infty} xf(x) = \mathbb{E}(I) - \sigma^2/2$$

en utilisant un développement limité de $\log(1 + u)$ à l'ordre 2.

Ainsi, quand $\sigma^2/2 < \mathbb{E}(I)$, $(\mathbb{E}(\log(Y_n)))_n$ est bornée et par le lemme de Fatou $\mathbb{E}(\lim_{n \rightarrow \infty} \log Y_n) < \infty$, ce qui implique

$$\varliminf_{n \rightarrow \infty} Y_n < \infty \quad \text{p.s.}$$

On renvoie au travail de Kesting [29, 28] 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 références 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 exemple.

Enfin, revenons sur le cas sous critique $m < 1$. On rappelle que $(Y_n)_n$ converge en loi vers

$$Y_\infty := \sum_{j \geq 0} Z_j \in [0, \infty].$$

Si on regarde l'espérance conditionnelle

$$\mathbb{E}(Y_\infty | (I_k)_{k \geq 1}) = \sum_{k \geq 1} I_k m^k$$

et le lemme 19 assure que $\mathbb{E}(Y_\infty | (I_k)_{k \geq 1}) < \infty$ p.s. si et seulement si $\mathbb{E}(\log^+(I)) < \infty$. Ceci entraîne 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

$$\overline{\lim}_{n \rightarrow \infty} (1 - p_0)^n I_n = \infty \quad \text{p.s.}$$

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

$$\mathbb{P}(Z_n \geq 1 | (I_k)_{k \geq 1}) \geq 1 - (1 - (1 - p_0)^n)^{I_n}$$

a une limsup égale à 1 p.s., ce qui implique $\sum_{j \geq 0} Z_j$ et donc Y_∞ p.s. infinie par le lemme de 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'extinction

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 $W = \lim_{n \rightarrow \infty} Z_n / m^n \in [0, \infty)$. 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 [2]. 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 [41]. Cette approche permet d'obtenir aussi la vitesse d'extinction du processus et se généralise à des processus de branchement plus complexes, 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

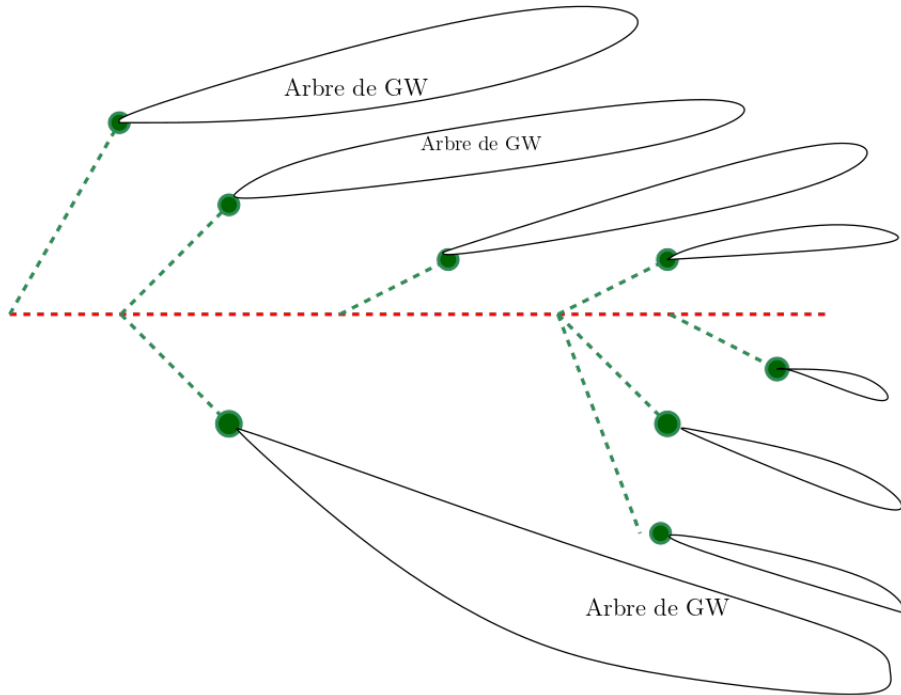
$$X_n = \#\mathbb{V}_n, \quad Y_n = \#(\mathbb{V}_n - \{E_n\}) = X_n - 1$$

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 $\hat{L} - 1$. 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 vue de la construction spinale est sûrement d'utiliser les vagues migratoires (proposition précédente) et de faire un dessin :



L'épine se reproduit suivant la loi biaisée $\hat{L}$. 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 $I_n = \hat{L}_n - 1$ 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 $F_n : \mathbb{T}_n \rightarrow [0, 1]$ 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 :

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

en rappelant que $W_k = Z_k/m^k$. 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 :

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

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 $m > 1$. $\square$

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

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

Notons que

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

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'événement 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 [41].

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 général $\mathbb{P}(W = 0) \in \{p_{ext}, 1\}$ 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 regroupe les individus de 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 $W = \lim_{n \rightarrow \infty} Z_{n+1}/m^{n+1}$,

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

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 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 $(\mathbb{P}(Z_n > 0)/m^n)_n$ 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 définit

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

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 [41] et [2]. 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 Processus de 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 météorologiques, 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 $(\exp(-s_n^{\mathbf{e}}) Z_n^{\mathbf{e}})_{n \in \mathbb{N}}$ 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^e que

$$\mathbb{E}(Z_{n+1}^e | \mathcal{F}_n) = m(e_n)Z_n^e$$

et on raisonne par récurrence pour i .

Pour ii), on remarque que $\mathbb{P}(Z_n^e > 0)$ est décroissante, ainsi $\mathbb{P}(Z_n^e > 0) \leq \min\{\mathbb{P}(Z_i^e > 0) : 0 \leq i \leq n-1\}$. Puis on applique l'inégalité de Markov, de sorte à majorer $\mathbb{P}(Z_i^e > 0)$ par $\exp(s_i^e)$. $\square$

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

$$f_{i,j} = \mathbb{E}\left(s^{Z_j^e} \mid Z_i^e = 1\right)$$

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

1.5.1 Critère d'extinction

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

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 $Z^\mathcal{E} = (Z_n^\mathcal{E})_{n \geq 0}$. We write it $Z = (Z_n)_{n \geq 0} = Z^\mathcal{E}$ for short.

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

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

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 $\mathbb{P}(Z_1 = 1) < 1$.

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

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 $n \rightarrow \infty$, 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 $n \rightarrow \infty$. La réciproque est plus délicate et on renvoie à [3] pour la preuve. $\square$

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 strictement croissant. Si de plus $\mathbb{E}(\log(\widehat{Z}_1)) < \infty$, alors comme dans la section précédente on peut montrer que $W > 0$ sur l'événement de survie et obtenir la croissance de Z_n via celle de S_n .

1.5.2 Effet de l'environnement sur la vitesse d'extinction

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 [23] 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

$$\mathbb{P}(Z_n > 0) \sim_{n \rightarrow \infty} c \mathbb{E}(m(\mathcal{E}_0))^n$$

- *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

$$\mathbb{P}(Z_n > 0) \sim_{n \rightarrow \infty} cn^{-1/2} \mathbb{E}(m(\mathcal{E}_0))^n$$

- *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

$$\mathbb{P}(Z_n > 0) \sim_{n \rightarrow \infty} cn^{-3/2} \gamma^n,$$

avec

$$\gamma := \inf_{0 \leq s \leq 1} \mathbb{E}[m(\mathcal{E}_0)^s] < \mathbb{E}[m(\mathcal{E}_0)].$$

- *Cas critique.* Si $\mathbb{E}(\log(m(\mathcal{E}_0))) = 0$, alors

$$\mathbb{P}(Z_n > 0) \sim_{n \rightarrow \infty} cn^{-1/2}.$$

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 dû à la possibilité que des environnements particulièrement favorables est lieu au début 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 mêmes raisons. Plus généralement, les processus de branchement en environnement aléatoire offrent un cadre intéressant où se mêle la stochasticité démographique et environnementale, notions importantes en biologie et écologie, renvoyons à [37, 38] pour des références.

Chapitre 2

Populations structurées : le cas neutre

Les individus sont dorénavant caractérisés par un trait. Commençons par les modèles neutres, c'est à dire que le trait n'affecte pas la loi de reproduction. Mais le trait peut évoluer, avec une transmission aléatoire d'une génération à la suivante (hérédité et variabilité). Comme la loi de reproduction est fixe, la généalogie est donnée par un arbre de Galton-Watson. L'évolution des traits est ainsi 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 Noël, 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 $\mathbb{X}$, qui donne les traits des enfants d'un individu de trait x . On suppose $\Theta : \mathcal{X} \times \Omega \rightarrow \mathbb{X}$ mesurable avec $\mathcal{X} \times \Omega$ muni de la tribu produit. On définit enfin $\Theta_x(u) = (X_i^x(u) : i = 1, \dots, L(u))$ où $(\Theta_x(u))_{x \in \mathcal{X}}$ famille de v.a. indépendantes et identiquement distribuées pour $u \in \mathcal{U}$, comme $(\Theta_x)_{x \in \mathcal{X}}$. La variable $\Theta_x(u)$ donne les traits de l'individu quand ce dernier a le trait 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 $((\Theta_x(u))_{x \in \mathcal{X}} : u \in \mathcal{U})$ est définie récursivement par $\mathcal{T} = \cup_{n \in \mathbb{N}} \mathbb{G}_n$ et

$$\mathbb{G}_0 = \{\emptyset\}, \quad X(\emptyset) = x_0$$

et pour tout $n \geq 0$,

$$\mathbb{G}_{n+1} = \{uk : u \in \mathbb{G}_n, k \in \llbracket 1, L(u) \rrbracket\}, \quad X(uk) = X_k^{X(u)}(u)$$

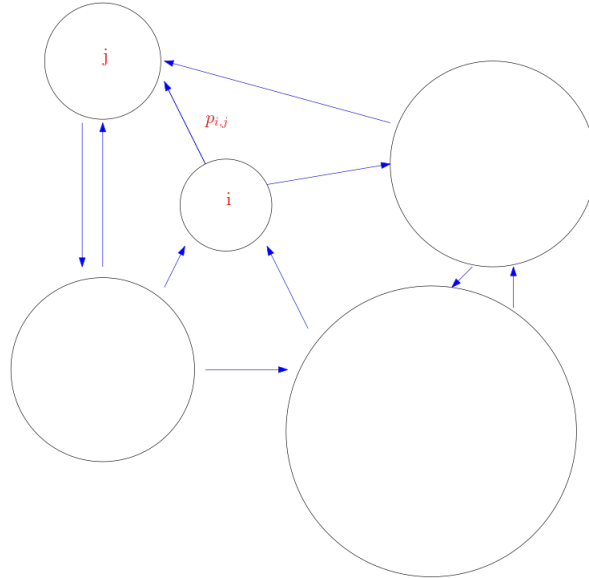
pour tous $u \in \mathbb{G}_n$ et $k \in \llbracket 1, L(u) \rrbracket$.

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 une matrice de transition $(p_{x,y})_{x,y \in \mathcal{X}}$ sur $\mathcal{X}$, qui est 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 de longueur fini $(x_k)_k$ 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'intéresse à ces modèles, motivés notamment par la gestion de l'espace et la préservation de la biodiversité (corridors, modèles puits-sources...). On pourra consulter par exemple [9, 45] et les références qui s'y trouvent.



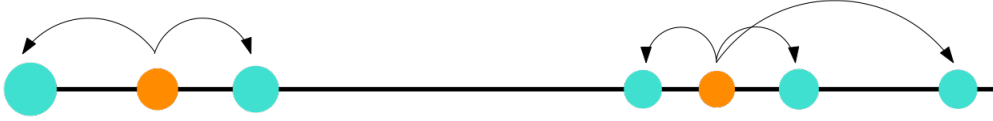
L'espace des traits peut contenir plus de caractéristiques individuelles, par exemple un couple $x = (x_1, x_2)$, avec x_1 la composante spatiale correspondant à un premier graphe, et x_2 une caractéristique phénotypique ou génotypique, les transitions correspondant alors aux mutations, recombinaisons etc...

Les cas réductible est également pertinent dans certains modèles, par exemple la

différentiation cellulaire, où la cellule change de type en se différenciant, sans retour en arrière possible.

Un tel processus est appelé processus de branchement (ou ou Galton-Watson) multitype, avec un nombre de types fini.

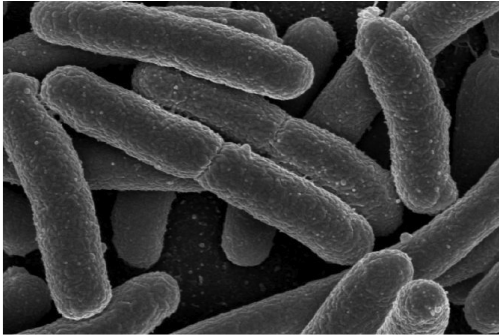
Marche aléatoire branchante et reproduction dispersion. On s'intéresse ici à une population qui se reproduit et les enfants sont dispersés autour des parents, typiquement une population de plantes avec essaimage par le vent ou les pollénisateurs. On voudra notamment comprendre comment une espèce invasive (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. $\Theta_x = x + \Theta_0$.



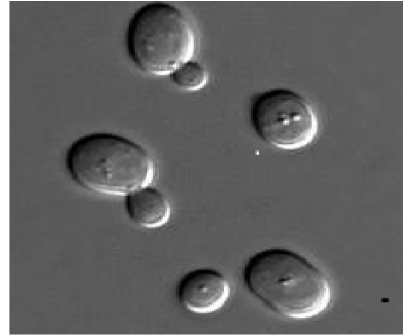
Processus autorégressif bifurcant et division cellulaire. On s'intéressera ici à la transmission d'une caractéristique d'une cellule en division, comme son taux de croissance. La division cellulaire étant binaire, on se focalisera pour cet exemple sur le cas $L = 2$ p.s. et on modélisera la transmission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

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

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 asymétrique $\alpha_1 \neq \alpha_2$. 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 se divise 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 créé au niveau de la membrane à chaque division), on renvoie à [18, 49, 24] pour des détails. Dans ce modèle on voudra aussi faire dépendre la division du trait, et pour cela on renvoie au chapitre suivant.



E-COLI division



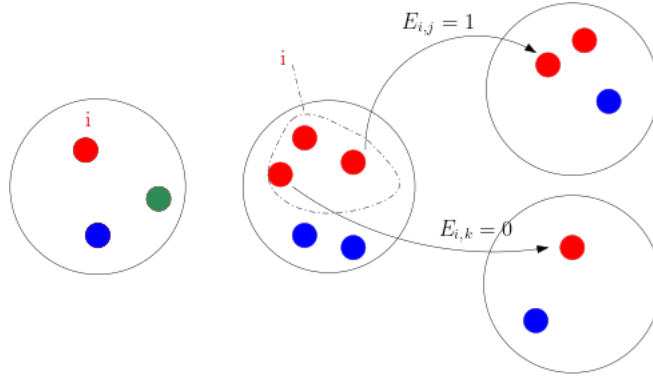
YEAST division

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 $\bar{\mu}$, 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

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

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 et sa transmission au cours de la division, comme l'ADN extrachromosomal, les plasmides (l'intérêt initial de Marek Kimmel), les mitochondries...

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 $n \in \mathbb{N}$,

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

Considérons maintenant l'espérance de la mesure empirique.

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

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

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. 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 $M_n = M^{on}$ 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, on peut noter $M = (M_{i,j})_{i,j \in \mathcal{X}}$ la matrice définie par

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

Ainsi $M_{i,j}$ donne 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 ème de la matrice M pour le produit matriciel usuel.

Démonstration. Calculons

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

La propriété de branchement ¹ assure que

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

et donc

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

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

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,

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

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 $Q(x, \cdot)$ est une loi de probabilité. Il se réécrit

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

où $\hat{p}$ 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 :

$$Q^{(k)}(x, \cdot) = k^{-1} \sum_{1 \leq i \leq k} \mathbb{P}(X_i^x \in \cdot \mid L = k).$$

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

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

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 première génération de l'arbre t

Proposition 35 (Tous pour un). Soient $F : \mathcal{X}^{n+1} \rightarrow \mathbb{R}$ mesurable bornée, $x \in \mathcal{X}$ et $n \geq 0$.

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

Ce résultat permet de réduire, en moyenne, la mesure empirique des trajectoires des individus à une seule trajectoire, réalisée par 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 intérêt 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 croît 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 précédent. 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 :

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

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 $F_n(x_0, \dots, x_n) = \prod_{j=0}^n f_j(x_j)$ 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 en oubliant la condition initiale dans la notation (car elle est fixée) :

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

avec

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

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

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

et on conclut avec la définition de Y :

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

en notant $y_0 = Y_0 = x$ et en rassemblant les intégrales par Fubini. $\square$

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 métapopulations, pour tous $x, y \in \mathbb{X}$,

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

Pour le deuxième exemple, la marche aléatoire branchante :

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

Pour le processus autorégressif bifurcant,

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

avec $\mathcal{E}_i$ v.a. gaussienne centrée de variance Γ_{ii} . Le processus Y est alors un processus autorégressif (transformation affine aléatoire), où le coefficient est tiré aléatoirement à chaque pas de temps, uniformément parmi $\{\alpha_i, \alpha_2\}$. Nous reviendrons dessus.

Enfin considérons le modèle de Kimmel. Dans ce cas, $L \in \{0, 1, 2\}$. Alors $m = p_1 + 2p_2$, $\hat{p}_1 = \frac{p_1}{m}$ et $\hat{p}_2 = \frac{2p_2}{m}$. Le noyau de transition de la chaîne Y est donné par

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

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 sera ainsi donnée par 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 biaiser les traits des individus échantillonnés. En appliquant la formule à $F(x_1, \dots, x_n) = h(x_n)G(x_1, \dots, x_n)$, on obtient

$$\mathbb{E}_{(\emptyset, x)} \left[\sum_{u \in \mathbb{G}_n} h(X(u))G(X(v) : v \preceq u) \right] = m^n \mathbb{E}_x [h(Y_n)G(Y_0, \dots, Y_n)] \quad (2.1)$$

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 $\mathcal{X}$ et $\lambda \geq 0$ telle que $M(x, h) = \lambda h(x)$ pour tout $x \in \mathcal{X}$. Alors pour tous $n \geq 0$ et G mesurable positive sur $\mathcal{X}^n$ et pour tout $x \in \mathcal{X}$:*

$$\mathbb{E}_{(\emptyset, x)} \left[\sum_{u \in \mathbb{G}_n} h(X(u))G(X(v), v \preceq u) \right] = \lambda^n h(x) \mathbb{E}_x \left[G(\tilde{Y}_0, \dots, \tilde{Y}_n) \right],$$

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

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

Démonstration. On est ramené au calcul

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

et on obtient plus l'expression voulue en rappelant (2.1). On conclut la preuve en remarquant que $Q(x, \cdot)$ est bien une mesure de probabilité, définissant ainsi les transitions de la chaîne Y . C'est à cet endroit que le fait que h est fonction propre intervient. En effet, le fait que pour tout $x \in \mathcal{X}$, $Q(x, \mathcal{X})$ est équivalent au fait que tout $x \in \mathcal{X}$, $M(x, h) = \lambda h(x)$. $\square$

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 $\tilde{Y}$. 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 privilégiera dans la somme les positions les plus grandes. Notons que nous n'avons pas imposé h borné dans la formule précédente, 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 échantillonne deux individus. Les traits de deux individus se retrouvent corréllés 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 :

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

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

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

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

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 $n \in \mathbb{N}$:*

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

avec $H_{p,n}$ définie par

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

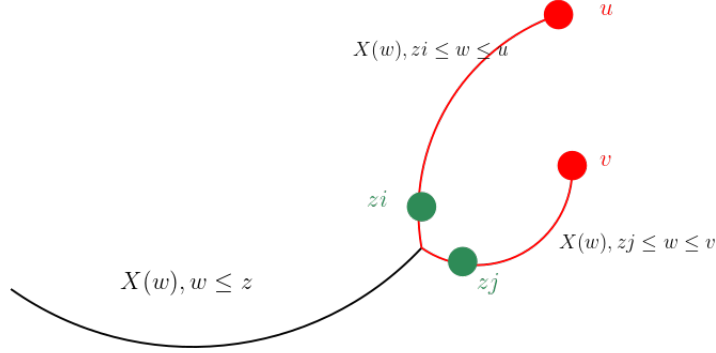
pour $y \in \mathcal{X}$ et pour $F : (y_0, \dots, y_n) \in \mathcal{X}^{n+1} \rightarrow F(y_0, \dots, y_n)$ mesurable bornée :

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

La preuve repose sur 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 $u \succcurlyeq zi, v \succcurlyeq zj$,

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



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

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

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

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

De plus, en utilisant la formule "tous pour un" pour $zi \in \mathbb{G}_{p+1}$,

$$\begin{aligned} \mathbb{E} \left[\sum_{\substack{u \in \mathbb{G}_n \\ zi \preccurlyeq u}} F_{p+1,n}(X(w), zi \preccurlyeq w \preccurlyeq u) \mid X(zi) \right] &= m^{n-p-1} \mathbb{E}_{X(zi)}[F_{p+1,n}(Y_i, i \leq n-p-1)] \\ &= m^{n-p-1} Q_{n-p-1}(X(zi), F_{p+1,n}). \end{aligned}$$

On combine les identités obtenues et on decompose

$$\begin{aligned} \{(u, v) \in \mathbb{G}_n^2 : u \neq v\} &= \bigsqcup_{p \leq n-1} \{(u, v) \in \mathbb{G}_n^2 : u \wedge v \in \mathbb{G}_p\} \\ &\quad \bigsqcup_{p \leq n-1, z \in \mathbb{G}_p} \{(u, v) \in \mathbb{G}_n^2 : zi \preccurlyeq u, zj \preccurlyeq v, i \neq j\} \end{aligned}$$

pour obtenir

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

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

On peut écrire 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 $i \leq n-1$, 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 $n \in \mathbb{N}$,*

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

avec

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

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 $x \in \mathcal{X}$,*

$$\mathbb{E}_x[f(Y_n)] \xrightarrow{n \rightarrow \infty} \pi(f).$$

Alors

$$\frac{\mathbf{1}_{\mathbb{G}_n \neq \emptyset}}{\#\mathbb{G}_n} \sum_{u \in \mathbb{G}_n} f(X(u)) \xrightarrow{n \rightarrow \infty} \mathbf{1}_{\forall n \geq 0, \mathbb{G}_n \neq \emptyset} \pi(f),$$

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 dépend 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

$$\mathbf{1}_{\mathbb{G}_n \neq \emptyset} \frac{\#\{u \in \mathbb{G}_n : X(u) = k\}}{\#\mathbb{G}_n} \xrightarrow{n \rightarrow \infty} \mathbf{1}_{\forall i \geq 0, \mathbb{G}_i \neq \emptyset} \pi_k.$$

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,

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

En utilisant que

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

on obtient que le premier terme est écrasé quand $n \rightarrow \infty$. De plus, le corollaire précédent (fourches) permet de gérer le deuxième terme et on obtient que

$$\mathbb{E}[\Delta_n^2] = \mathcal{O} \left(\frac{1}{m^n} \right) + \sum_{p=0}^{n-1} m^{-p-2} \mathbb{E}_x[h_{n,p}(Y_p)],$$

avec

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

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

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

tend vers 0 quand K tend vers l'infini (uniformément en $n \geq K$).

Enfin pour chaque $p \leq K$, pour tout $y \in \mathcal{X}$, $Q_{n-p-1}(y, g) = Q_{n-p-1}(y, f) - \int_{\mathcal{X}} f(x)\pi(dx) \rightarrow 0$ quand $n \rightarrow \infty$, par hypothèse. En utilisant que $Q(\cdot, g)$ est borné, on obtient que $h_{n,p}(y)$ tend vers 0 quand n tend vers l'infini, pour tout $y \in \mathcal{X}$. De plus $h_{n,p}$ et une nouvelle convergence dominée implique $\mathbb{E}_x[h_{n,p}(Y_p)]$ 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

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

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

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 suffisamment 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 chaîne 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 autorégressif 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 perturbation gaussienne :

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

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 supposons $\text{VAR}(\mathcal{E}_1) = \text{VAR}(\mathcal{E}_2) = \sigma^2$, soit

$$\mathcal{E}_1 \stackrel{\text{loi}}{=} \mathcal{E}_2 \stackrel{\text{loi}}{=} \mathcal{E}$$

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

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

Proposition 40. La chaîne $(Y_n)_n$ 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

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

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

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

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

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

Cela nous permet de récupérer une structure monotone :

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

en introduisant la marche aléatoire $S_n = 2 \sum_{j=1}^n \log(\theta_j)$. Si $\mathbb{E}(\log(\theta_1)) = \log(\alpha_1 \alpha_2)/2 < 0$, alors S_n tend de façon linéaire vers $-\infty$ par la loi des grands nombres et $A_\infty < \infty$ p.s. Sinon $A_\infty = \infty$ p.s. On obtient alors la classification suivante

Lemme 41. (i) Si $\alpha_1 \alpha_2 < 1$, alors $A_\infty < \infty$ et Y_n converge en loi quand $n \rightarrow \infty$ vers un mélange de gaussienne π défini par

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

(ii) Si $\alpha_1 \alpha_2 \geq 1$, alors Y_n tend en probabilité quand $n \rightarrow \infty$ 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 $a < b$,

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

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

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

dans L^2 .

2.2.2 Infection dans le modèle de Kimmel

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

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

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

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

avec

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

en rappelant que les P_i sont i.i.d. de loi μ , avec μ la loi de reproduction des pathogènes (sur une génération) et $E_{i,j}$ v.a. de Bernoulli de paramètre q . Autrement dit

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

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 i.i.d à valeurs dans E $\mathcal{E} = (\mathcal{E}_0, \mathcal{E}_1, \dots)$, dont la loi est donnée par

$$\mathbb{P}(\mathcal{E}_0 = 1) = \frac{p_1}{m}, \quad \mathbb{P}(\mathcal{E}_0 = q) = \frac{p_2}{m}, \quad \mathbb{P}(\mathcal{E}_0 = 1 - q) = \frac{p_2}{m}.$$

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.

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

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 $\mathbb{E}(\log(m(\mathcal{E}))) \leq 0$ (et $\mathbb{P}_1(Y_1 = 1) < 1$, ce qui est le cas ici) 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 arbres de Galton-Watson (Théorème 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 $\bar{\mu}$ de la loi de reproduction des pathogènes) et le paramètre de Bernoulli q qui détermine la répartition.

Théorème 44. *Si*

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

alors

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

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

$$\lim_{n \rightarrow \infty} \mathbb{E} \left(\mathbf{1}_{\mathbb{G}_n \neq \emptyset} \frac{\#\{u \in \mathbb{G}_n : X(u) > K\}}{\#\mathbb{G}_n} \right) > 0,$$

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

en probabilités.

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 . Probablement que la bonne façon de réécrire ce critère, pour interpréter les différents termes et généraliser à des mécanismes de répartition plus complexe est la suivante :

$$\log(\bar{\mu}) \leq \hat{p}_2 \mathbb{E}(\log(1/F)), \quad \hat{p}_2 = \frac{2p_2}{m},$$

avec $\hat{p}_2$ la probabilité de se diviser dans une lignée typique et $\mathbb{E}(\log(1/F))$ l'effet moyen de la division sur la masse de pathogène (moyenne géométrique). Ici, on a symétrisé la v.a. donnant le paramètre de la loi binomiale pour la répartition : $F \stackrel{d}{=} 1 - F$. Ici $\mathbb{P}(F = q) = \mathbb{P}(F = 1 - q) = 1/2$ mais le paramètre q pourrait être lui même aléatoire, cad que la fraction moyenne attendue pour la cellule moins chargée suit elle même une distribution (entre $(0, 1/2]$), plutôt que d'être fixée égale à q .

Démonstration. On calcule

$$\mathbb{E}(\log(m(\mathcal{E}_0))) = \frac{p_1}{m} \log(m(1)) + \frac{p_2}{m} \log(m(q)) + \frac{p_2}{m} \log(m(1 - q)).$$

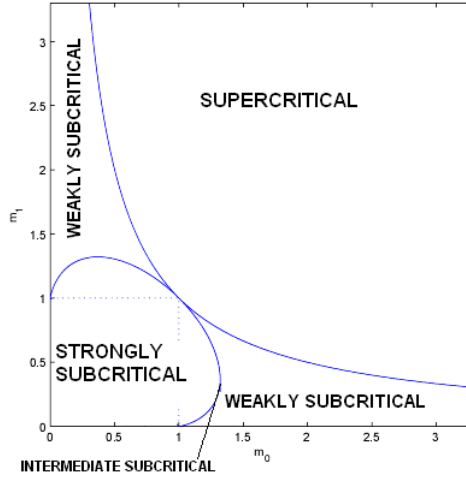


FIGURE 2.1 – Représentation des différents régimes pour le nombre de cellules infectées, dans le cas $p_2 = 1$, avec en abscisse $m_0 = q\bar{\mu}$ et en ordonnée $m_1 = (1 - q)\bar{\mu}$.

avec $m(e) = \mathbb{E}(\text{Bin}(P, e)) = \bar{\mu}e$. 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 $\mathbb{E}(\log(m(\mathcal{E}_0)))$. Si cette quantité est négative alors $\mathbb{P}_k(Y_n > 0)$ tend vers 0 quand $n \rightarrow \infty$, et ce pour tout $k \geq 0$. C'est à dire que $\mathbb{E}_k(f(Y_n)) \rightarrow 0$ quand $n \rightarrow \infty$ 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 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 $f(x) = \mathbf{1}_{x \in (0, K)}$. $\square$

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

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

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 exponentiellement. Cette croissance dépend de la vitesse d'extinction de Y_n . Or cette vitesse $\mathbb{P}(Y_n > 0)$ dépend du signe de

$$\kappa = \mathbb{E}(m(\mathcal{E}_0) \log(m(\mathcal{E}_0))),$$

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 cellules 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 répartissent 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. distribuées comme I , toutes ces v.a. étant par ailleurs indépendantes entre elles.

La formule nous donne pour un

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

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

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

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

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

est plus grande ou plus petite que 1. En effet, la section 1.3 garantissant que si $\mathbb{E}(I) < \infty$ et $\bar{\mu} < 2$ 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) garantissant 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 où les cellules ne se débarrassent pas de l'infection (à cause de la contamination) mais l'infection reste modérée.

Si $\bar{\mu} > 2$, 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 $\bar{\mu} = 2$. 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 [8].

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

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

Pour simplifier, on suppose que les enfants atterrissent 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 $D_i \stackrel{\text{loi}}{=} D$. 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 pas du nombre d'enfants. La formule tous pour un permet d'éclairer ce qui se passe en moyenne pour la mesure empirique :

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

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

$$\mathbb{P}(Y_n \in [\sqrt{n}a, \sqrt{n}b]) \xrightarrow{n \rightarrow \infty} \mathbb{P}(\mathcal{N}(0, \sigma^2) \in [a, b]).$$

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. *Pour tout $a \leq b$ réels*

$$\frac{1_{\mathbb{G}_n \neq \emptyset}}{\#\mathbb{G}_n} \#\{u \in \mathbb{G}_n, a\sqrt{n} \leq X(u) \leq b\sqrt{n}\} \xrightarrow[n \rightarrow \infty]{L^2} 1_{\forall n, \mathbb{G}_n \neq \emptyset} \mathbb{P}(\mathcal{N}(0, \sigma^2) \in [a, b]).$$

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 tout $v \geq 0$,

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

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 $\approx$ signifiera 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 apparaît, la vitesse v^* solution de

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

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 [19]. Nous supposons ici que la fonction

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

est bien définie pour tout $\theta \in \mathbb{R}$, c'est à dire que $\mathbb{E}[e^{\theta D}] < \infty$. 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

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

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

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

est strictement croissante et elle admet une limite

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

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 enfin que

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

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 $\psi(x) = \theta(x)x - \Lambda(\theta(x))$ et 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 $\Lambda(\theta) = \log\left(\frac{e^\theta + e^{-\theta}}{2}\right)$, et $\theta(x) = \frac{1}{2} \log\left(\frac{1+x}{1-x}\right)$. On a alors

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

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

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

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

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

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

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 $\theta = \theta(x)$ tel que

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

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

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

Ce choix permet d'avoir la dérivée voulue :

$$\mathbb{E}[\tilde{D}] = \frac{\mathbb{E}[D e^{\theta D}]}{\mathbb{E}(e^{\theta D})} = \Lambda'(\theta) = x.$$

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

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

Alors

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

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

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

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

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 $v^* \in (0, \gamma)$ tel que

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

Commençons par la borne supérieure.

Proposition 48. $\overline{\lim}_{n \rightarrow \infty} \mathbf{1}_{\mathbb{G}_n \neq \emptyset} \frac{R_n}{n} \leq v^* \text{ p.s.}$

Démonstration. Soit $v \in (v^*, \gamma)$, de sorte que $\psi(v) > \psi(v^*) = \log(m)$ 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 $\mathbb{P}(Y_n \geq vn)$ décroît 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 $\overline{\lim}_{n \rightarrow \infty} \frac{R_n}{n} \leq v$ p.s. En faisant tendre v vers v^* , on a le résultat annoncé. $\square$

Il faut maintenant montrer que cette vitesse v^* 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.

$$\lim_{n \rightarrow \infty} \mathbf{1}_{\mathcal{S}} \frac{R_n}{n} \geq \mathbf{1}_{\mathcal{S}} v^* \quad \text{p.s.}$$

Nous allons donner deux preuves de ce résultat, chacune éclairante et généralisable à d'autres modèles. La première est due à Biggins et repose sur la construction d'un processus de Galton-Watson imbriqué dans la population structurée et avançant à la vitesse voulue.

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

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

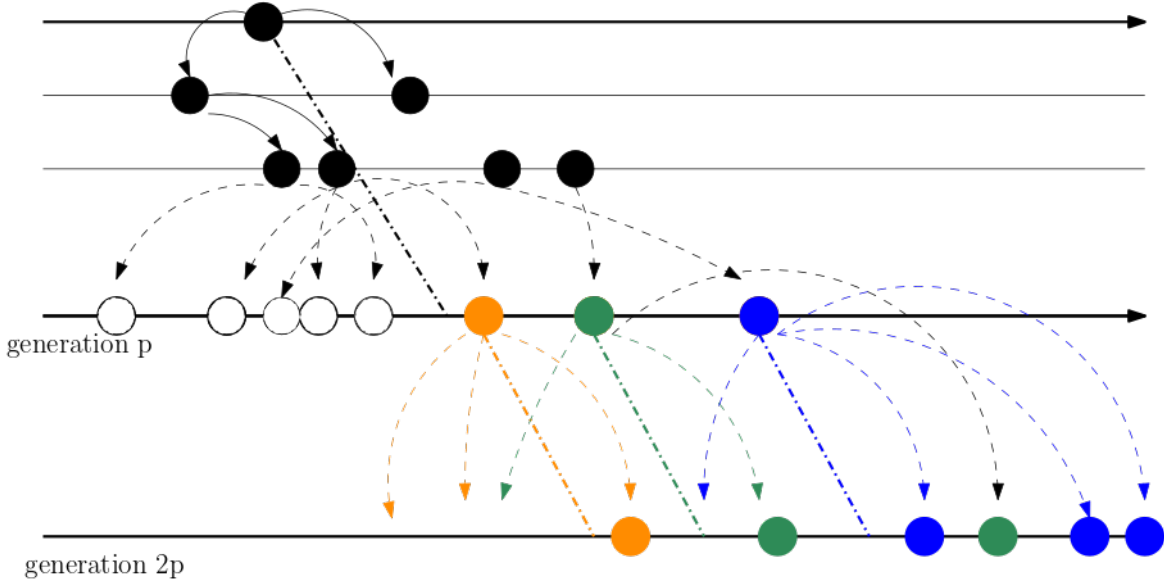


FIGURE 2.2 – Subpopulations (colored balls) created by keeping individuals which have moved to the right enough with respect to the positions of their ancestors p generations before ; the other individuals are forgotten.

qui tend vers l'infini quand $p \rightarrow \infty$ 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 $\mathbb{G}_0^{(p)} = \{\emptyset\}$, $\mathbb{G}_1^{(p)} = \{u \in \mathbb{G}_p, X(u) \geq pv\}$ et plus généralement pour $k \geq 0$

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

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

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

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. $\square$

Nous allons maintenant donner les grandes lignes d'une autre preuve inspirée des travaux de Lyons et du monographe de Zhan Shi [48] 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 $x \rightarrow \exp(\rho \cdot)$ est un vecteur propre du semigroupe du premier moment, c-a-d :

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

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

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

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 $x \in \mathbb{R}$,*

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

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 Zhan 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 $\rho \lambda'(\rho) < \lambda(\rho) \log \lambda(\rho)$.

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

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

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

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 $\mathbb{P}(|Y_n^{(\rho)} - nv| > n\varepsilon)$ décroît 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 sûrement. On obtient

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

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

Chapitre 3

Processus de branchement pour des populations structurées

Nous souhaitons maintenant prendre en compte le fait que le trait des individus affecte leur loi de reproduction. Nous perdons donc l'hypothèse 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 être infini et non discret. On parle aussi de chaîne de Markov branchante, pour souligner que la transmission des traits va être un processus de Markov.

3.1 Chaîne de Markov branchante

3.1.1 Définition

De façon informelle, la variable de reproduction L dépend maintenant de $x \in \mathcal{X}$. Chaque individu de trait x va indépendamment donner naissance à un nombre aléatoire fini 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 $\mathbb{X} = \cup_{k \geq 0} \mathcal{X}^k$ que l'on munit de la tribu engendrée par $\cup_{k \geq 0} (\mathcal{B}_{\mathcal{X}})^k$. Introduisons maintenant un processus pour chaque label : $((\Theta_x(u))_{x \in \mathcal{X}}, u \in \mathcal{U})$ famille de v.a. indépendantes et identiquement distribuées comme $(\Theta_x)_{x \in \mathcal{X}}$. On note $\Theta_x(u) = (X_i^x(u), i = 1, \dots, L^x(u))$. Pour les questions de mesurabilité, on suppose que $\Theta : \mathcal{X} \times \Omega \rightarrow \mathbb{X}$ mesurable où $\mathcal{X} \times \Omega$ est muni de la tribu produit.

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

$$\mathbb{G}_0 = \{\emptyset\}, \quad X(\emptyset) = x_0$$

et pour tout $n \geq 0$,

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

pour tous $u \in \mathbb{G}_n$ et $k \in \llbracket 1, L^{X(u)}(u) \rrbracket$.

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 l'hypothèse de neutralité de la loi de reproduction. 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.

3.1.2 Exemple structuré en âge et régénération

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. Toujours pour simplifier, restons dans le 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 :

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

On démarre d'un individu d'âge 0 : $X(\emptyset) = 0$ et on considère la chaîne de Markov branchante $((u, X(u)) : u \in \mathcal{T})$ associée.

La question 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 résultats 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 basée la notion de R_0 en épidémiologie : on regarde ce qu'un individu laisse comme descendance globalement avant de mourir. 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 qui ne peuvent pas se reproduire, la survie est compromise... Une approche classique est ainsi de mettre en avant une structure de Galton-Watson (simple type) à l'intérieur de notre population structurée, ou plus généralement une structure régénérative. La structuration en âge est bon exemple pour voir comment cela fonctionne, voir par exemple aussi [9] pour une structure spatiale

décrite par un graphe.

Commençons par exclure des cas dégénérés, qui corresponde à des cas particulier du régime critique. La probabilité qu'un individu issu de l'âge 0 survive pendant n générations vaut :

$$\prod_{i=0}^{n-1} (1 - p_0(i)).$$

Si ce produit converge vers une valeur strictement positive, un nouveau né pourra vivre indéfiniment. Nous allons exclure ce cas et il finira toujours par mourir. On veut également empêcher qu'un nouveau né produise toujours exactement 1 enfant au cours de sa vie. Cela revient à empêcher la situation suivante : un nouveau né survit jusqu'à l'âge a presque sûrement et se reproduit alors p.s., puis ne se reproduit plus jamais :

$$p_1(i) = 1 \quad (i < a); \quad p_2(a) = 1, \quad p_2(i) = 0 \quad (i > a).$$

Ainsi nous travaillerons sous l'hypothèse :

$$\prod_{i=0}^{\infty} (1 - p_0(i)) = 0, \quad \forall a \geq 0, \quad \sum_{i=0}^{a-1} (1 - p_0(i)) + (1 - p_2(a)) + \sum_{i=a+1}^{\infty} p_2(a) > 0. \quad (3.1)$$

Introduisons maintenant

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

Ce paramètre, qui donne le nombre d'enfants moyen produit au cours de la vie d'un nouveau né, permet de classifier le comportement en temps long de la population structurée en âge, en terme d'extinction, explosion ou bornitude.

Proposition 53. *On suppose (3.1). Si $\mu \leq 1$, alors le nombre de nouveaux nés est fini et la population s'éteint p.s. :*

$$\#\{u \in \mathcal{T} : X(u) = 0\} < \infty, \quad \exists n \geq 0, \quad \#\mathbb{G}_n = 0 \quad \text{p.s.}$$

Si $\mu > 1$, avec probabilité positive, la population de nouveaux nés est infinie :

$$\mathbb{P}(\#\{u \in \mathcal{T} : X(u) = 0\} = \infty) > 0.$$

Traiter les cas dégénérés exclus par l'hypothèse (3.1) est assez facile. Aller plus loin dans le régime $\mu > 1$, pour savoir quand la population à la génération n tend bien vers l'infini demande plus de travail, et nous renvoyons à [40], chapitre 2, pour aller plus loin.

La preuve de cette proposition repose sur le fait que la sous population des nouveaux nés forme un processus de Galton-Watson de moyenne de reproduction μ . Elle informe d'ailleurs au delà des résultats énoncés dans cette proposition.

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'âge 0, on regarde

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

en notant $\mathbf{1}_a = (1, \dots, 1) \in \mathcal{U}$ le label de l'individu avec a fois 1. En effet, dans notre labelisation, un individu u qui survit est noté $u1$ et 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,

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

Suivons maintenant la sous population d'âge 0. La racine, avant de mourir, 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 branchement) ... 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).

Si $\mu \leq 1$, la population de nouveaux nés s'éteint et leur nombre total est fini. En effet, nous savons depuis le premier chapitre que dans le cas sous critique $\mu < 1$, l'extinction a lieu p.s. Dans le cas critique $\mu = 1$, elle a lieu seulement quand un individu est remplacé par un exactement, ce que nous avons exclu avec la deuxième partie de (3.1). Comme le nombre de nouveaux nés est fini, la population finit par s'éteindre, car chaque nouveau né finit par mourir avec la première partie de l'hypothèse (3.1).

Si $\mu > 1$, avec probabilité positive le processus de Galton-Watson survit et le nombre total de nouveaux nés tend vers l'infini. $\square$

L'idée de régénération pour caractériser notamment la survie ou l'explosion se généralise. Notamment, si on part d'un individu de trait $x_0 \in \mathcal{X}$, on peut lui associer l'ensemble de ses lignées avant le premier instant où le trait des descendants revient en x_0 :

$$\mathcal{C}_\emptyset = \{u \in \mathcal{T} : X(v) \neq x_0 \text{ pour tous } \emptyset \prec v \prec u\}$$

On note $U_1 \subset \mathcal{T}$ la ligne d'arrêt des premiers retours en x_0 , c'est à dire l'ensemble des individus $u \in \mathcal{T}$ tels $X(u) = x_0$ et $X(v) \neq x_0$ pour $\emptyset \prec v \prec u$. On peut itérer la procédure et associer à chaque u de U_1 l'ensemble des individus $\mathcal{C}_u$ qui sont ses descendants qui ont évité x_0 dans leur lignée. On constitue ainsi des générations successives U_n d'individus de trait x_0 qui forment une partition de $\{u \in \mathcal{T} : X(u) = x_0\}$. Par ailleurs on a une partition de l'arbre complet :

$$\mathcal{T} = \cup_{n \in \mathbb{N}} \cup_{u \in U_n} \mathcal{C}_u.$$

Et une structure de branchement simple type (processus de Galton-Watson) pour décrire les générations $(U_n)_{n \in \mathbb{N}}$ successives. Pour cette approche par ligne d'arrêt, on pourra consulter en particulier [15].

3.2 Semigroupe et formules tous-pour-un

La définition du semigroupe du premier moment est inchangée. Mais l'espace de fonctions où il est bien fini peut être plus compliqué à caractériser que dans le cas neutre.

Pour simplifier les discussions et parce que c'est (plus que) raisonnable d'un point de vue modélisation, nous supposons aussi dans tout ce chapitre que le nombre moyen d'enfants par individu est fini, quelque soit le trait :

$$(A) \quad \forall x \in \mathcal{X}, \quad \mathbb{E}(L^x) < \infty.$$

Définition 54. Pour tout $x \in \mathcal{X}$ et f fonction mesurable positive sur $\mathcal{X}$, on définit

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

qui est positif, et peut être infini. Il en va de même plus généralement pour $n \geq 0$ de

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

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

L'hypothèse (A) assure que pour toute fonction positive bornée, $M(\cdot, f)$ est fini. Quand une fonction $f = f_+ - f_-$ se décompose en des parties positive $f_+(x) = f(x)1_{f(x) \geq 0}$ et négative $f_-(x) = -f(x)1_{f(x) \leq 0}$ telles $M(\cdot, f_+)$ et $M(\cdot, f_-)$ sont finis, alors on peut définir

$$M(x, f) = M(x, f_+) - M(x, f_-)$$

qui est fini également. Du coup, $M(\cdot, f)$ est bien défini et fini pour toute fonction mesurable bornée. Mais elle n'est pas forcément bornée. Ainsi, pour que $M_n(\cdot, f)$ soit bien défini et fini, c'est un peu plus délicat, même avec f bornée. Par ailleurs, l'utilisation de fonctions tests non bornées f sera intéressante en soi. Tout cela motive l'introduction d'un espace fonctionnel stable par M :

Définition 55. On considère un sous espace vectoriel $\mathcal{B}_s(\mathcal{X})$ des fonctions mesurables finies de $\mathcal{X}$ dans $\mathbb{R}$ tel que

$$\forall f \in \mathcal{B}_s(\mathcal{X}), \quad M(\cdot, f_+) \in \mathcal{B}_s(\mathcal{X}), \quad M(\cdot, f_-) \in \mathcal{B}_s(\mathcal{X}).$$

Pour $f \in \mathcal{B}_s(\mathcal{X})$, $M(\cdot, f) = M(\cdot, f_+) - M(\cdot, f_-)$ est alors bien défini et appartient aussi à $\mathcal{B}_s(\mathcal{X})$.

Par exemple, dans le cas où le nombre moyen d'enfants est borné : $\sup_{x \in \mathcal{X}} \mathbb{E}(L^x) < \infty$, on peut prendre pour $\mathcal{B}_s(\mathcal{X})$ l'espace des fonctions mesurables bornées car

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

Un autre exemple important qui va jouer un rôle dans la suite est l'espace dominé par une fonction de Lyapounov :

$$\mathcal{B}_s(\mathcal{X}) = \left\{ f \in \mathcal{B}(\mathcal{X}) : \sup_{x \in \mathcal{X}} \frac{|f(x)|}{V(x)} < \infty \right\}$$

où V une fonction positive mesurable telle que $M(x, V) \leq CV$ pour tout $x \in \mathcal{X}$.

On travaille dorénavant sur l'espace de fonctions $\mathcal{B}_s(\mathcal{X})$. La propriété de semigroupe montrée dans le cas neutre au chapitre précédent s'étend directement :

Proposition 56. *Pour tous $x \in \mathcal{X}$ et $f \in \mathcal{B}_s(\mathcal{X})$ et $n \geq 0$,*

$$M_n(\cdot, f) = M_n(\cdot, f_+) - M_n(\cdot, f_-) \in \mathcal{B}_s(\mathcal{X})$$

est bien définie et vérifie

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

3.2.1 Formule de Feynman Kac et principe variationnel

Comme nous l'avons fait dans le cas neutre, nous pouvons également réduire la mesure empirique moyenne à la dynamique du trait d'un seul individu. Plusieurs options sont possibles. On commence par remarquer que

$$\begin{aligned} \mathbb{E} \left(\sum_{u \in \mathbb{G}_{n+1}} F(X(v), v \preccurlyeq u) \mid \mathcal{F}_n \right) &= \sum_{w \in \mathbb{G}_n} \mathbb{E} \left(\sum_{u \in \mathbb{G}_{n+1}: u \succcurlyeq w} F(X(v), v \preccurlyeq w, X(u)) \mid \mathcal{F}_n \right) \\ &= \sum_{w \in \mathbb{G}_n} G(X(v), v \preccurlyeq w) \end{aligned}$$

grâce à la propriété de branchement. De plus, la fonction G correspond à intégrer sur une génération, en fixant la trajectoire passée. Par définition de M et propriété de Markov, G est donnée par

$$G(x_0, \dots, x_n) = M(x_n, F(x_0, \dots, x_n, \cdot)) = \int_{\mathcal{X}} M(x_n, dx_{n+1}) F(x_0, \dots, x_n, x_{n+1}).$$

On peut alors justifier par récurrence que pour $n \geq 0$, F mesurable positive sur $\mathcal{X}^{n+1}$:

$$\mathbb{E}_{(\emptyset, x_0)} \left[\sum_{u \in \mathbb{G}_n} F(X(v), v \preccurlyeq u) \right] = \int_{\mathcal{X}^n} F(x_0, \dots, x_n) \Pi_{i=0}^{n-1} M(x_i, dx_{i+1}), \quad (3.2)$$

car en prenant l'espérance dans le calcul précédant et en procédant par récurrence :

$$\begin{aligned} \mathbb{E}_{(\emptyset, x_0)} \left(\sum_{u \in \mathbb{G}_{n+1}} F(X(v), v \preccurlyeq u) \right) &= \mathbb{E}_{(\emptyset, x_0)} \left(\sum_{u \in \mathbb{G}_n} G(X(v), v \preccurlyeq u) \right) \\ &= \int_{\mathcal{X}^n} G(x_0, \dots, x_n) \Pi_{i=0}^{n-1} M(x_i, dx_{i+1}) \\ &= \int_{\mathcal{X}^n} \int_{\mathcal{X}} F(x_0, \dots, x_n, x_{n+1}) M(x_n, dx_{n+1}) \Pi_{i=0}^{n-1} M(x_i, dx_{i+1}). \end{aligned}$$

En introduisant maintenant le noyau de Markov :

$$Q(x, dy) = \frac{M(x, dy)}{m(x)} \quad \text{avec } m(x) = M(x, \mathbf{1}) = \mathbb{E}_{(\emptyset, x)}(\#\mathbb{G}_n)$$

on peut lire alors

$$\mathbb{E}_{(\emptyset, x)} \left[\sum_{u \in \mathbb{G}_n} F(X(v), v \preccurlyeq u) \right] = \int_{\mathcal{X}^n} F(x_0, \dots, x_n) \Pi_{i=0}^{n-1} m(x_i) Q(x_i, dx_{i+1})$$

et on obtient la représentation de type Feynman Kac suivante

Proposition 57. *Soient $n \geq 0$, F mesurable positive sur $\mathcal{X}^{n+1}$. Alors :*

$$\mathbb{E}_{(\emptyset, x)} \left[\sum_{u \in \mathbb{G}_n} F(X(v), v \preccurlyeq u) \right] = \mathbb{E}_x [F(V_i, i \leq n) \Pi_{i=0}^{n-1} m(V_i)],$$

où V est la chaîne de Markov sur $\mathcal{X}$ de noyau Q .

Remarquons que dans le cas où $(X_i^x)_{i \geq 1}$ est une suite de variable identiquement distribuée et indépendante de L^x , alors $Q(x, dy) = \mathbb{P}(X_1^x \in dy)$.

La chaîne de Markov V est ici "pénalisée" trajectoriellement par le terme

$$S_n = \sum_{i=0}^{n-1} \log(m(V_i)) = n \int_{\mathcal{X}} \log(m(v)) \mu_n(dv)$$

avec μ_n la mesure d'occupation de l'espace des traits :

$$\mu_n = \frac{1}{n} \sum_{i=0}^{n-1} \delta_{V_i},$$

i.e. $\mu_n(A) = \#\{i \leq n-1 : V_i \in A\}/n$ donne la proportion du temps passé en A par la chaîne Y avant n .

Un principe variationnel émerge alors pour quantifier la taille de la population et la trajectoire typique. En effet, on a obtenu en particulier ici

$$\mathbb{E}_{(\emptyset, x)}(\#\mathbb{G}_n) = \mathbb{E}_x(e^{S_n})$$

et pour comprendre le comportement en temps long, on est ramené à un problème qui rappelle la méthode de Laplace, ou plutôt ici dans ce contexte de grande déviation à un lemme de Varadhan. De façon très informelle, les déviations peuvent aider à quantifier la probabilité que μ_n charge un certain domaine de l'espace des fréquences, et on obtient ainsi une approximation de la loi de $S_n = n \int_{\mathcal{X}} \log(m(v)) \mu_n(dv)$:

$$\mathbb{P}(\mu_n \approx \mu) = \mathbb{P}\left(S_n \approx n \int_{\mathcal{X}} \log(m) \mu\right) \asymp e^{-nI(\mu)}$$

où $\asymp$ est un équivalent des logarithmes et I une fonction de taux positive qu'on ne détaillera pas ici. Ainsi à l'échelle $\log$, $\mathbb{E}_{(\emptyset, x)}(\#\mathbb{G}_n)$ est donné par une quantité de la forme $\int \exp(n(\int_{\mathcal{X}} \log(m) \mu - I(\mu)))$. Le principe de Laplace nous invite à chercher le maximum λ de la fonction $\mu \rightarrow \int_{\mathcal{X}} \log(m) \mu - I(\mu)$ pour obtenir la croissance malthusienne λ de la population. Là où le maximum est atteint, on retrouve les fréquences d'occupations des lignées typiques, voir [9, 11]. Pour des références sur cette approche variationnelle des processus de branchement, on renvoie à [6].

3.2.2 Lignée typique et processus inhomogène

D'autres normalisations de M que celle introduite dans la partie précédente en divisant par la masse $M(\cdot, \mathbf{1})$ sont possibles pour exhiber une chaîne de Markov et décrire les lignées ancestrales. En particulier on peut vouloir se débarrasser de la pénalisation trajectorielle et séparer la masse (le nombre total d'individus) des traits (l'évolution du type dans la lignée ancestrale). Pour cela, a priori, il n'y a pas beaucoup de choix, ça va passer par un argument télescopique et des quantités inhomogènes.

Nous allons ainsi faire intervenir une fonction particulière $\psi \in \mathcal{B}_s(\mathcal{X})$ mesurable (strictement) positive, qui va jouer un rôle important par la suite. On suppose que

$$\forall x \in \mathcal{X}, \forall n \in \mathbb{N}, M_n(x, \psi) > 0.$$

On introduit ainsi pour $i < n$ la mesure de probabilité

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

et cette renormalisation permet d'obtenir

$$\Pi_{i=0}^{n-1} Q_i^{(n)}(x_i, dx_{i+1}) = \frac{\psi(x_n)}{M_n(x_0, \psi)} \Pi_{i=0}^{n-1} M(x_i, dx_{i+1})$$

En rappelant (3.2) et en introduisant la chaîne de Markov inhomogène $(Y_i^{(n)})_{i \leq n}$ associée à $Q_i^{(n)}$:

$$\mathbb{P}(Y_{i+1}^{(n)} \in A \mid Y_i^{(n)} = x) = Q_i^{(n)}(x, A)$$

pour tous $0 \leq i \leq n-1$, nous obtenons le résultat suivant

Proposition 58. Soient $n \geq 0$, F mesurable positive sur $\mathcal{X}^{n+1}$ et $\psi \in \mathcal{B}_s(\mathcal{X})$ strictement positive sur $\mathcal{X}$. Alors :

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

Le cas $\psi = 1$ est naturel : il donne une formule non biaisée et permet, comme dans le chapitre précédent, d'attraper la mesure empirique et la lignée ancestrale d'un individu tiré uniformément au hasard. En particulier, dans ce cas, en regardant seulement les traits à la dernière génération :

$$M_n(x, f) = M_n(x, 1) \mathbb{E}_x(f(Y_n^{(n)})). \quad (3.3)$$

3.2.3 Fourches pour échantillonner deux individus

A nouveau, on peut chercher à capter deux individus, et leur lignée ancestrale et proposer des formules pour les fourches. Travaillons avec des fonctions positives pour éviter pour l'instant les discussions sur les moments : soient $(f_i)_{i \in \mathbb{N}}$ et $(g_i)_{i \in \mathbb{N}}$ des fonctions mesurables positives sur $\mathcal{X}$. On note

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

et on définit aussi la famille de mesure $\kappa(x, dydz)$ sur $\mathcal{X}^2$ par

$$\kappa(x, f) = \int_{\mathcal{X}^2} \kappa(x, dydz) f(y, z) = \mathbb{E}_{(\emptyset, x)} \left(\sum_{u \neq v \in \mathbb{G}_1} f(X(u), X(v)) \right)$$

pour f fonction mesurable positive de $\mathcal{X}^2$ dans $\mathbb{R}$ et

$$M_n(x, F) = \mathbb{E}_{(\emptyset, x)} \left[\sum_{u \in \mathbb{G}_n} F(X(u), v \preceq u) \right]$$

pour F fonction mesurable positive de $\mathcal{X}^{n+1}$ dans $\mathbb{R}$ et $n \geq 0$ et

$$M_n(x, F, G) = \int_{\mathcal{X}^2} \kappa(x, dydz) M_{n-1}(y, F) M_{n-1}(z, G)$$

pour F, G fonction mesurables positives de $\mathcal{X}^n$ dans $\mathbb{R}$. En suivant la démonstration de la Proposition 37 (jusqu'au moment d'utiliser une formule tous pour un) on obtient

$$\mathbb{E}_{(\emptyset, x)} \left[\sum_{u \neq v \in \mathbb{G}_n} F_{0,n}(X(u), w \preceq u) G_{0,n}(X(v), w \preceq v) \right] = \sum_{p=0}^{n-1} M_p(x, [F, G]_{p,n})$$

avec

$$[F, G]_{p,n}(x_0, \dots, x_p) = F_{0,p} \cdot G_{0,p}(x_0, \dots, x_p) M_{n-p}(x_p, F_{p+1,n}, G_{p+1,n}).$$

En particulier tout s'écrit en fonction du premier moment M et du noyau de corrélation (sur une reproduction) κ . En écrivant M avec une formule tous pour un grâce à la proposition précédente on obtient un analogue de la Proposition 37. En prenant $F(x_0, \dots, x_n) = f(x_n)$ pour $f : \mathcal{X} \rightarrow \mathbb{R}_+$ mesurable, on obtient

$$[F, F]_{0,n}(x) = \int_{\mathcal{X}^2} \kappa(x, dydz) M_{n-1}(y, f) M_{n-1}(z, f) = \kappa(x, f_n^M)$$

en notant

$$f_n^M(x, y) = M_{n-1}(y, f) M_{n-1}(z, f)$$

Nous en déduisons

$$\mathbb{E}_{(\emptyset, x)} \left[\sum_{u \neq v \in \mathbb{G}_n} f(X(u)) f(X(v)) \right] = \sum_{p=0}^{n-1} M_p(x, \kappa(\cdot, f_{n-p}^M)) \quad (3.4)$$

avec

$$M_p(x, \kappa(\cdot, f_{n-p}^M)) = \int_{\mathcal{X}} M_p(x, dy) \int_{\mathcal{X}^2} \kappa(y, dzdw) M_{n-p-1}(z, f) M_{n-p-1}(w, f)$$

qui est utile pour le calcul des corrélations dans l'approche L^2 de la loi des grands nombres.

3.3 Utilisation de fonctions harmoniques

L'existence et la connaissance de fonctions propres strictement positive pour le semi-groupe joue un rôle important dans l'étude des processus, comme on a pu le voir déjà dans le cas neutre. Ici, quand ψ est une fonction propre, le noyau $Q_i^{(n)}$ se simplifie et devient homogène. Le noyau Markovien ainsi obtenu est la h transformée de l'opérateur M . On obtient ainsi une chaîne de Markov homogène, indépendante de l'instant final n .

Proposition 59. *Supposons qu'il existe une fonction h strictement positive sur $\mathcal{X}$ telle que pour tout $x \in \mathcal{X}$, $M(x, h) = \lambda h(x)$, avec $\lambda > 0$. Alors pour tout $n \geq 0$ et toute fonction F mesurable positive ou bornée sur $\mathcal{X}^{n+1}$,*

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

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

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

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 que le fait que h soit fonction propre assure que $M_n(x, h)$ est bien fini et que pour fonction F bornée (pas nécessairement positive), l'espérance ci dessus est bien définie et finie. On pourrait aussi relâcher la stricte positivité en définissant la chaîne de Markov hors des points où h s'annule.

De nombreux résultats en mathématiques (via la théorie spectrale ou une contraction de l'opérateur) garantissent 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 équivalent, 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. Nous y reviendrons dans la section suivante. Différentes extensions existent en dimension infinie, en particulier avec la théorème de Krein Rutman. Terminons par une dernière remarque avant la preuve. Cette h transformée permet d'écrire

$$M_n(x, f) = \lambda^n h(x) Q_n(f/h) = \lambda^n h(x) \mathbb{E}_x(f(Y_n)) \quad (3.5)$$

et ainsi de ramener l'analyse de M_n quand n tend vers l'infini à celle de la chaîne de Markov Y .

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

$$\begin{aligned} \mathbb{E} \left(\sum_{u \in \mathbb{G}_{n+1}} h(X(u)) \mid \mathcal{F}_n \right) &= \sum_{v \in \mathbb{G}_n} \mathbb{E} \left(\sum_{u \in \mathbb{G}_{n+1}, u \succ v} h(X(u)) \mid \mathcal{F}_n \right) \\ &= \sum_{v \in \mathbb{G}_n} M(X(v), h) = \lambda \sum_{v \in \mathbb{G}_n} h(X(v)) \end{aligned}$$

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

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'individus 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 contrôler la probabilité que $W = 0$ en fonction du trait de l'individu de départ. Le résultat peut être 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 [33] pour le cas d'un espace de traits $\mathcal{X}$ fini (et une loi des grands nombres pour les proportions) et [4] pour des cas infinis.

Pour gérer les corrélations et obtenir des lois des grands nombres pour décrire la distribution des traits dans la population, on peut encore utiliser les calculs L^2 . Utilisons à nouveau notre fonction harmonique h pour simplifier les expressions : $x \in \mathcal{X}$, $M(x, h) = \lambda h(x)$ avec $h > 0$. Soient $(f_i)_{i \in \mathbb{N}}$ et $(g_i)_{i \in \mathbb{N}}$ des fonctions positives sur $\mathcal{X}$. On note

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

et on définit $H_{p,n}$ par

$$H_{p,n}(y) = \frac{1}{h(y)} \mathbb{E}_{(\emptyset, y)} \left(\sum_{a \neq b \in \mathbb{G}_1} h(X(a)) h(X(b)) Q_{n-p-1}(X(a), F_{p+1,n}) Q_{n-p-1}(X(b), G_{p+1,n}) \right),$$

pour $y \in \mathbb{N}$ et $F : (y_0, \dots, y_n) \in \mathbb{N}^{n+1} \rightarrow F(y_0, \dots, y_n)$ bornée, en rappelant que $Q_n(y, F) = \mathbb{E}_y(F(Y_i, 0 \leq i \leq n))$.

Pour tous $x \in \mathcal{X}$ et $n \in \mathbb{N}$:

$$\begin{aligned} \mathbb{E}_{(\emptyset, x)} \left(\sum_{u \neq v \in \mathbb{G}_n} h(X(u)) h(X(v)) F_{0,n}(X(w), w \preccurlyeq u) G_{0,n}(X(w), w \preccurlyeq v) \right) \\ = h(x) \sum_{p=0}^{n-1} \lambda^{2n-p-2} \mathbb{E}_x (F_{0,p} G_{0,p}(Y_i, i \leq p) H_{p,n}(Y_p)). \end{aligned} \quad (3.6)$$

3.3.1 Exemple de modèle structuré en âge

Illustrons sur un modèle en âge l'existence et la forme des éléments propres et en particulier de la fonction harmonique h . Ce modèle généralise celui qui a été considéré en Section 3.1.2, en autorisant des naissances multiples.

Chaque individu est caractérisé par un âge $a \in \mathbb{N}$ et, indépendamment, d'une génération sur l'autre :

- meurt (sans descendance) avec probabilité $1 - q(a)$.
- survit avec probabilité $q(a)$ et a alors k enfants avec probabilité $p_k(a)$, pour $k \geq 0$.

Plus formellement, un individu d'âge a est remplacé par la collection d'individus Θ_a telle que

$$\begin{aligned}\mathbb{P}(\Theta_a = (\emptyset)) &= 1 - q(a) \\ \mathbb{P}(\Theta_a = (a+1, \mathbf{0}_k)) &= q(a)p_k(a)\end{aligned}$$

avec $\mathbf{0}_k = (0, \dots, 0) \in \mathbb{R}^k$, $q(a), p_k(a) \in [0, 1]$ et $\sum_{k \geq 0} p_k(a) = 1$ pour tout $a \in \mathbb{N}$. On suppose que $q(0) > 0$ et que pour tout $a \in \mathbb{N}$,

$$m(a) = q(a) \sum_{k \geq 0} k p_k(a) < \infty.$$

On suppose enfin qu'il existe $z_0 > 0$ tel que

$$1 \leq \sum_{a \geq 0} z_0^{a+1} m(a) \bar{q}(a) < \infty, \quad \text{avec} \quad \bar{q}(a) = \prod_{i=0}^{a-1} q(i), \quad \bar{q}(0) = 1.$$

Ici la matrice (infinie) M définie par $M_{i,j} = M(i, \mathbf{1}_j) = \mathbb{E}_{(\emptyset, i)}(\{\#\{u \in \mathbb{G}_1 : X(u) = j\}\})$ pour $i, j \in \mathbb{N}$, donne le nombre moyen d'individus d'âge j engendrés par un individu d'âge i . Notons que pour $a \geq 0$,

$$M_{a,a+1} = q(a), \quad M_{a,0} = m(a).$$

On peut commencer par remarquer qu'il existe un unique couple (λ, π) avec $\lambda > 0$ et $\pi = (\pi(a) : a \in \mathbb{N})$ probabilité tel que

$$\pi M = \lambda \pi.$$

En effet, si π satisfait cette équation, alors $\pi_a = \pi_0 \bar{q}(a) / \lambda^a$, π_0 est non nulle. De plus on peut caractériser λ par l'équation suivante

$$\sum_{a \geq 0} \left(\frac{1}{\lambda}\right)^{a+1} m(a) \bar{q}(a) = 1.$$

Cette équation permet notamment une CNS pour savoir quand $\lambda > 1$ (régime surcritique) :

$$\mu = \sum_{a \geq 0} m(a) \bar{q}(a) > 1.$$

On notera au passage qu'on retrouve bien le μ de l'exemple de la première section. On peut montrer ensuite qu'il existe une unique fonction $h = (h(a) : a \in \mathbb{N})$ telle que

$$M(., h) = \lambda h, \quad h(0) = 1, \quad \forall a > 0, h(a) > 0.$$

On a en effet $h(a+1) = (\lambda h(a) - M(a)h(0))/q(a)$, et avec $h(0) = 1$,

$$h(a) = \frac{\lambda^a}{\bar{q}(a)} - \sum_{b=0}^{a-1} m(b) \frac{\bar{q}(b)}{\bar{q}(a)} \lambda^{a-1-b}$$

et on peut diviser par λ^{a-1} et multiplier par $\bar{q}(a)$ pour voir le signe positif, en utilisant que $z = 1/\lambda$ solution de l'équation précédente.

Pour appliquer les résultats de convergence que nous avons mis en place dans ce cours, il faut étudier maintenant la chaîne de Markov $(Y_n)_n$, où on rappelle que Y est la chaîne pour un individu typique associée au noyau h transformé $Q(\cdot, f) = M(\cdot, fh)/(\lambda h)$. Plusieurs techniques de stabilité de chaîne de Markov peuvent être invoquées ici. En particulier la condition de Doeblin (voir Section 3.5) offre une condition naturelle pour garantir une convergence en loi géométrique, en exploitant le rôle spécial de l'âge 0, accessible depuis les autres âges directement par naissance. Notons que la fonction harmonique h est bornée, mais elle peut s'approcher de 0. Nous renvoyons à [40] pour des résultats fins dans cette direction, et une extension à l'environnement aléatoire.

3.3.2 Exemple de modèle structuré en taille

Considérons un autre exemple, où la fonction harmonique sera simple, mais le fait qu'elle soit non bornée va introduire de nouvelles difficultés et phénomènes. Le processus de branchement sera ici structuré en taille pour une population de cellules. Chaque individu (cellule) est caractérisée par une taille $x \in \mathbb{R}_+^*$. On suppose que la probabilité de se diviser dépend de la taille de la cellule et que la cellule ne meurt pas. A chaque génération, indépendamment, chaque cellule de taille x

- survit sans se diviser avec probabilité $p_1(x)$ et alors sa masse x est multipliée par une v.a. distribuée comme $D \in \mathbb{R}_+^*$.
- se divise en deux cellules filles avec probabilité $p_2(x)$, et sa masse x est multipliée puis coupée en deux et les tailles des deux cellules filles sont $F Dx$ et $(1-F) Dx$, avec $F \in (0, 1)$ indépendante de D .

Plus formellement, on considère $p_1(x), p_2(x) \in [0, 1]$ tels que

$$p_1(x) + p_2(x) = 1, \quad m(x) = p_1(x) + 2p_2(x) > 0.$$

Une cellule de taille x est ainsi remplacée par la collection aléatoire de cellules Θ_x à valeurs dans $\mathbb{R}_+^* \cup (\mathbb{R}_+^*)^2$ dont la loi est donnée par

$$\mathbb{P}(\Theta_x \in dz) = 1_{\mathbb{R}_+^*}(z) \mathbb{P}(Dx \in dz) p_1(x) + 1_{(\mathbb{R}_+^*)^2}(z) \mathbb{P}((F Dx, (1-F) Dx) \in dz) p_2(x),$$

avec F et D v.a. indépendantes et indépendantes de x . On suppose que

$$1 < \mathbb{E}(D) < \infty, \quad \mathbb{E}(D^2) < \infty, \quad \mathbb{P}(D \neq \mathbb{E}(D)) > 0$$

et que F v.a. distribuée comme $1 - F$.

Enfin, on suppose $x \rightarrow p_2(x)$ est croissante sur $\mathbb{R}_+$ et on note $\bar{p}_2$ sa limite :

$$\bar{p}_2 = \lim_{x \rightarrow \infty} p_2(x).$$

En effet, plus la cellule est grosse, plus elle est a priori prête à se diviser. Ici

$$M(x, f) = p_1(x) \mathbb{E}(f(xD)) + 2p_2(x) \mathbb{E}(f(xFD))$$

et l'identité $\text{Id}(x) = x$ est une fonction propre

$$M(., \text{Id}) = \mathbb{E}(D) \text{Id}.$$

Ainsi, par la Proposition 59,

$$\mathbb{E}(D)^{-n} \sum_{u \in \mathbb{G}_n} X(u)$$

est une martingale positive, et on peut remarquer que $\sum_{u \in \mathbb{G}_n} X(u)$ est la taille cumulée de la population cellulaire. De plus, pour $x \in \mathbb{R}_+^*$,

$$\mathbb{E}_{(\emptyset, x)} \left(\sum_{u \in \mathbb{G}_n} X(u) F(X(v), v \preceq u) \right) = x \mathbb{E}(F(Y_i : i \leq n)),$$

avec $(Y_n)_{n \in \mathbb{N}}$ la chaîne de Markov dont le noyau de transition Q est donné par

$$Q(x, dy) = \frac{M(x, dy)y}{x \mathbb{E}(D)}.$$

En notant $\hat{F}$ la v.a. F biaisée par la taille, i.e. $\mathbb{P}(\hat{F} \in dy) = 2y \mathbb{P}(F \in dy)$, on peut donner le comportement en temps long de Y :

Proposition 60. Si $\mathbb{E}(\log(D)) < \overline{p}_2 \mathbb{E}(\log(1/\hat{F}))$, alors

$$\lim_{n \rightarrow \infty} Y_n < \infty \quad p.s.$$

Si $\mathbb{E}(\log(D)) > \overline{p}_2 \mathbb{E}(\log(1/\hat{F}))$, alors

$$\lim_{n \rightarrow \infty} Y_n = \infty \quad p.s.$$

Démonstration. On peut écrire Y de la façon suivante

$$Y_{n+1} = Y_n D_n \left(1_{U_n \leq p_1(Y_n)} + 1_{U_n > p_1(Y_n)} \hat{F}_n \right)$$

avec $n \geq 0$, $(D_n)_n$, $(\hat{F}_n)_n$, $(U_n)_n$ v.a. indépendantes, $(D_n)_n$ toutes distribuées comme D , $(\hat{F}_n)_n$ toutes distribuées comme $\hat{F}$, et $(U_n)_n$ v.a. uniformes sur $[0, 1]$. On peut alors comparer Y_n à l'exponentielle d'une marche aléatoire. Si $\mathbb{E}(\log(D)) < \overline{p}_2 \mathbb{E}(\log(1/\hat{F}))$, il existe x_1 tel que

$$\mathbb{E}(\log(D)) < p_2(x_1) \mathbb{E}(\log(1/\hat{F}))$$

Comme p_2 est croissant, tant que $(Y_n)_n$ est plus grand x_1 , Y_n est inférieure l'exponentielle d'une marche aléatoire de pas $\log(D) + 1_{U \leq p_2(x_1)} \log(\hat{F})$, et dont la moyenne vaut $\mathbb{E}(\log(D)) + p_2(x_1) \mathbb{E}(\log(\hat{F})) < 0$. La loi des grands nombre permet de conclure pour le premier cas et l'autre se traite de façon similaire. $\square$

Remarquons que $(Y_n)_n$ consiste à choisir une cellule biaisée par la taille dans la population et on obtient ici pour une telle cellule un critère de récurrence. Ce critère ne coïncide pas forcément avec celui qui consiste à choisir une cellule uniformément. On peut s'en rendre compte en considérant le cas neutre, quand la division ne dépend pas de la taille :

$p_1(x) = p_1; p_2(x) = p_2$. Dans ce cas, on peut utiliser la formule tous pour un de la Proposition 35, c'est à dire exploiter le fait que dans ce cas $h = 1$ est également fonction propre de M . Alors le processus de Markov $(Z_n)_n$ par cette formule est récurrent si $\mathbb{E}(\log(D)) + 2p_2\mathbb{E}(\log(F))/m < 0$, et tend vers l'infini si cette quantité est positive. On voit donc qu'il y a un régime où la taille d'une cellule tirée uniformément est récurrente mais une cellule biaisée par la taille tend vers l'infini.

Continuons l'exploration de ce modèle, en précisant qu'on ne donnera que des résultats préliminaires. Malgré leur simplicité apparente, le comportement de ces objets reste en partie à comprendre et l'objet de travaux en cours. On se place dans un cas favorable, ou la taille biaisée Y est bien régulée :

$$(\star) \quad \mathbb{E}(D) (1 - \overline{p_2} + \overline{p_2} \mathbb{E}(\hat{F})) < 1$$

Alors il existe $\alpha < 1$ tel que pour y assez grand

$$\mathbb{E}_y(Y_1) \leq \alpha y.$$

On en déduit que $(\mathbb{E}(Y_n))_{n \geq 0}$ est une suite bornée.

En introduisant pour f bornée de $\mathbb{R}_+$ dans $\mathbb{R}_+$ et $y \in \mathbb{R}_+^*$,

$$H_k(y, f) = yp_2(y) \mathbb{E}_{(\varnothing, y)} \left(D^2 F(1 - F) Q_{k-1}(yDF, f) Q_{k-1}(yD(1 - F), f) \right).$$

la formule pour les fourches (3.6) devient

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

pour tous $x \in \mathbb{R}_+^*$ et $n \in \mathbb{N}$. Puis, en utilisant que $(\mathbb{E}_x(Y_n))_n$ est bornée, on obtient que

$$\mathbb{E}(D)^{-2n} \mathbb{E} \left(\sum_{u \neq v \in \mathbb{G}_n} X(u) X(v) \right)$$

est bornée. De même, en notant $\cdot^2$ la fonction carré et en rappelant (3.5),

$$\mathbb{E}_{(\varnothing, x)} \left(\sum_{u \in \mathbb{G}_n} X(u)^2 \right) = M_n(x, \cdot^2) = x \mathbb{E}(D)^n Q_n(x, \text{ID}) = x \mathbb{E}(D)^n \mathbb{E}_x(Y_n)$$

et sous la condition $(\star)$, on obtient

$$\lim_{n \rightarrow \infty} \mathbb{E}(D)^{-2n} \mathbb{E} \left(\sum_{u \in \mathbb{G}_n} X(u)^2 \right) = 0.$$

En rappelant la Proposition 59, la martingale $\lambda^{-n} \sum_{u \in \mathbb{G}_n} h(X(u))$ est bornée et donc convergente dans L^2 et sa limite W est positive p.s.

Pour décrire la distribution des tailles au sein de la population, on peut aller plus loin en restant sous la condition $(\star)$. Celle ci donne des conditions de Lyapunov et sous des et

on peut utiliser un critère de type Foster, Foster Lyapounov pour garantir la recurrence postive de la chaine. Avec des hypothèses peu contraignantes sur les lois de D ou F en lien avec l'irréductibilité et l'apériodicité, on obtient la convergence de $(Y_n)_{n \in \mathbb{N}}$ vers une loi stationnaire π , indépendamment de la valeur de départ. La condition $(\star)$ donne même ici de conditions de dérive qui assurent une convergence exponentielle, voir [26]. On obtient alors qu'il existe une probabilité π sur $\mathbb{R}_+^*$ tel que pour tous $y \in \mathbb{R}_+^*$,

$$\lim_{n \rightarrow \infty} Q_n(y, f) = \pi(f).$$

Comme on a pu le faire dans le chapitre précédent, on peut alors montrer que pour toute fonction f bornée de $\mathbb{R}_+^*$ telle que $\pi(f) = 0$,

$$\mathbb{E}(D)^{-2n} \mathbb{E} \left(\left(\sum_{u \in \mathbb{G}_n} X(u) f(X(u)) \right)^2 \right)$$

tend vers 0 quand n tend vers l'infini. Ainsi pour toute fonction f bornée de $\mathbb{R}_+^*$,

$$\lim_{n \rightarrow \infty} \frac{\sum_{u \in \mathbb{G}_n} X(u) f(X(u))}{\sum_{u \in \mathbb{G}_n} X(u)} = \pi(f)$$

dans L^2 . Plusieurs problèmes restent ici, qui ont à voir avec la discussion précédente au sujet du biais par la taille. En particulier, la fonction harmonique $h(x) = x$ induit un biais par la taille. On obtient ainsi un contrôle sur la population de cellule qui reviendrait à échantillonner, choisir, une cellule u proportionnellement à sa taille $X(u)$. Concrètement la régulation de la taille de la cellule ici ne correspond plus à une régulation d'une cellule choisie uniformément, c'est à dire au contrôle de la mesure empirique. Pour aller au delà de cette difficulté, il faut aller un plus loin, soit en maîtrisant bien le biais introduit pour "le corriger", soit en travaillant un peu différemment sur le semigroupe du premier moment pour ne pas introduire ce biais. Nous reviendrons la dessus un peu plus loin.

3.4 Espace de trait fini et construction spinale

Considérons le cas où l'espace des traits est fini :

$$d = \#\mathcal{X} < \infty.$$

Ce cas est plus simple mathématiquement pour diverses raisons : compacité, bornitude, uniformité seront vérifiées gratuitement sur $\mathcal{X}$. La question de l'existence et de l'unicité d'éléments propres positifs sont bien connus, et les éléments propres en question automatiquement bornés. Quand $\mathcal{X}$ est fini, l'étude asymptotique du semigroupe du premier moment revient à prendre les puissances positives d'une matrice de taille finie. Notons $M = (M_{i,j})_{i,j \in \mathcal{X}} \in \mathcal{M}_d(\mathbb{R}_+)$ associée à l'opérateur linéaire M , i.e. $M_{i,j} = M(i, \mathbf{1}_j)$. Supposons que M est irréductible apériodique, ce qui donne déjà de nombreux cas intéressants. On verra plus loin comment relâcher l'irréductibilité, par exemple avec la condition Doeblin. L'hypothèse d'apériodicité semble naturelle pour commencer et simplifie les énoncés mais n'est pas cruciale non plus.

Dans ce cadre irréductible apériodique, le théorème de Perron Frobenius garantit l'existence et l'unicité d'un triplet positif d'éléments propres pour M . Plus précisément, il existe un unique triplet $(\lambda, h, \pi) \in \mathbb{R}_+^* \times \mathcal{M}_{d,1}(\mathbb{R}_+^*) \times \mathcal{M}_{1,d}(\mathbb{R}_+^*)$ tel que

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

avec

$$\sum_{i \in \mathcal{X}} \pi_i = \sum_{i \in \mathcal{X}} \pi_i h(i) = 1.$$

On identifie ici la fonction $h : \mathcal{X} \rightarrow \mathbb{R}_+$ et le vecteur ligne $h = (h(i))_{i \in \mathcal{X}} \in \mathcal{M}_{d,1}(\mathbb{R}_+^*)$. Ainsi le vecteur propre à droite h pour M correspond à la fonction propre $M(\cdot, h) = \lambda h$ de la section précédente. Remarquons aussi que $\pi = (\pi_i)_{i \in \mathcal{X}}$ est une loi de probabilité sur $\mathcal{X}$. Le théorème de Perron Frobenius implique aussi qu'il existe $C > 0$ et $\eta \in (0, 1)$ tels que pour tous $x, y \in \mathcal{X}$, $n \in \mathbb{N}$ et

$$|\lambda^{-n}(M^n)_{i,j} - h(i)\pi_j| \leq C\eta^n$$

avec M^n la matrice M à la puissance n , c'est à dire $(M^n)_{i,j} = M_n(i, 1_j)$. Ainsi pour tous $n \in \mathbb{N}$ et $f : \mathcal{X} \rightarrow \mathbb{R}_+$,

$$\|\lambda^{-n}M_n(\cdot, f) - \pi(f)h\|_\infty \leq C\eta^n \|f\|_\infty \quad (3.7)$$

avec $\pi(f) = \sum_{i \in \mathcal{X}} \pi_i f(i)$. Les éléments propres donnent donc le comportement en temps long du semigroupe du premier moment, avec une convergence géométrique uniforme. On verra dans le chapitre suivant que cela se complique en dimension infinie.

Nous pouvons déjà donner des résultats sur l'extinction ou la survie en remarquant que pour tout $x \in \mathcal{X}$,

$$\mathbb{E}_{(\emptyset, x)}(\#\mathbb{G}_n) = M_n(x, 1) \sim_{n \rightarrow \infty} \lambda^n h(x).$$

Comme par l'inégalité de Markov $\mathbb{P}_{\delta_x}(\mathbb{G}_n \neq \emptyset) \leq \mathbb{E}_{\delta_x}(\#\mathbb{G}_n)$, la probabilité de survie tend vers 0 et la population s'éteint p.s. en temps fini dès que $\lambda < 1$.

Dans le cas $\lambda = 1$, le lemme de Fatou implique

$$\mathbb{E}_{(\emptyset, x)}\left(\liminf_{n \rightarrow \infty} \#\mathbb{G}_n\right) \leq \liminf_{n \rightarrow \infty} \mathbb{E}_{(\emptyset, x)}(\#\mathbb{G}_n) < \infty$$

et donc $\liminf_{n \rightarrow \infty} \#\mathbb{G}_n < \infty$ p.s. On peut alors montrer que la population s'éteint dès que 0 est accessible depuis tout point.

Pour le cas $\lambda > 1$, il faut travailler un peu plus pour montrer que la survie a lieu avec probabilité positive. On peut s'inspirer des idées du cas Galton Watson simple type et on renvoie au premier chapitre ou à [2, 27]. On va en réalité montrer un peu plus que la survie, avec une condition de moment un peu plus forte sur la loi de reproduction. On va chercher à établir la croissance géométrique de la taille de la population et la convergence de la distribution des types dans la population, sur l'événement de survie de la population. Rappelons en particulier pour cela que, grâce à la section précédente, on sait que

$$\lambda^{-n} \sum_{u \in \mathbb{G}_n} h(X(u)) = \sum_{x \in \mathcal{X}} h(x) \#\{u \in \mathbb{G}_n : X(u) = x\}$$

est une martingale positive convergeant vers une v.a. finie, positive. Le paramètre λ est bien le candidat pour le taux de croissance malthusien de la population, dès que la limite de la martingale n'est pas nulle.

3.4.1 Temps long et vitesse de convergence avec variance finie

Nous supposons dans cette partie que nous sommes dans le régime surcritique

$$\lambda > 1.$$

Nous supposons également que les lois de reproduction ont un second moment fini :

$$\forall x \in \mathcal{X}, \quad \mathbb{E}_{(\emptyset, x)} (\#\mathbb{G}_1^2) = \mathbb{E}((L^x)^2) < \infty.$$

Lemme 61. *Soit $f : \mathcal{X} \rightarrow \mathbb{R}$ et $(\emptyset, x)$ la condition initiale, avec $x \in \mathcal{X}$. Le processus*

$$\left(\lambda^{-n} \sum_{u \in \mathbb{G}_n} f(X(u)) \right)_{n \in \mathbb{N}}$$

est borné dans L^2 . De plus, la martingale

$$\left(\lambda^{-n} \sum_{u \in \mathbb{G}_n} h(X(u)) \right)_{n \in \mathbb{N}}$$

converge p.s. et dans L^2 vers une v.a. finie W telle que

$$\mathbb{E}_{(\emptyset, x)}(W) = h(x), \quad \{W > 0\} = \{\forall n \geq 0, \#\mathbb{G}_n > 0\} \text{ p.s.}$$

Démonstration. En particulier la dernière partie assure que le processus survit avec probabilité positive et qu'alors il croît comme λ^n .

Considérons une fonction $f : \mathcal{X} \rightarrow \mathbb{R}$ et utilisons (3.4) qui donne

$$\mathbb{E}_{(\emptyset, x)} \left[\sum_{u \neq v \in \mathbb{G}_n} f(X(u))f(X(v)) \right] = \sum_{p=0}^{n-1} \sum_{y \in \mathcal{X}} (M^p)_{x,y} \sum_{z,w \in \mathcal{X}} \kappa_{y,z,w} M_{n-p-1}(z, f) M_{n-p-1}(w, f)$$

avec

$$\begin{aligned} \kappa_{y,z,w} &= \mathbb{E}_{(\emptyset, y)} \left(\sum_{u \neq v \in \mathbb{G}_1} 1_{X(u)=z, X(v)=w} \right) \\ &= \mathbb{E}_{(\emptyset, y)} (\#\{(u, v) \in \mathbb{G}_1^2 : u \neq v, X(u) = z, X(v) = w\}). \end{aligned}$$

Notons que sous nos hypothèses de moment d'ordre 2, la fonction κ est partout finie et donc bornée puisqu'elle vit sur un espace fini. Alors

$$\mathbb{E}_{(\emptyset, x)} \left[\sum_{u \neq v \in \mathbb{G}_n} f(X(u))f(X(v)) \right] \leq C' \sum_{p=0}^{n-1} \sum_{y \in \mathcal{X}} (M^p)_{x,y} \sum_{z,w \in \mathcal{X}} |M_{n-p-1}(z, f) M_{n-p-1}(w, f)| \quad (3.8)$$

pour une constante $C' > 0$. Utilisons (3.7) pour borner $M_p(x, 1)$ par $\lambda^p(h(x) + C\eta^p)$ et faisons de même pour $M_{n-p-1}(\cdot, f)$, on obtient avec $\lambda > 1$ et une nouvelle constante C'' ,

$$\mathbb{E}_{(\emptyset, x)} \left[\sum_{u \neq v \in \mathbb{G}_n} f(X(u))f(X(v)) \right] \leq C'' \|f\|_\infty^2 \sum_{p=0}^{n-1} \lambda^p \cdot \lambda^{2n-2p-2} = \mathcal{O}(\lambda^{2n}).$$

Ceci prouve la première partie du résultat et le reste en découle en utilisant la proposition 59. En effet la convergence L^2 implique la convergence L^1 et donc que

$$\mathbb{E}(W) = \lim_{n \rightarrow \infty} \mathbb{E}_{(\emptyset, x)} \left(\lambda^{-n} \sum_{u \in \mathbb{G}_n} h(X(u)) \right) = \mathbb{E}_{(\emptyset, x)} \left(\lambda^{-0} \sum_{u \in \mathbb{G}_0} h(X(u)) \right) = h(x) > 0.$$

On peut alors raisonner comme en population simple type, dans le chapitre 1, pour conclure. Le fait que la survie de la population coïncide avec l'événement $\{W > 0\}$ est un peu plus délicat. On peut commencer par montrer que si la population survit, elle tend forcément vers l'infini, en distinguant le cas où 0 est accessible et le cas où il ne l'est pas. Quand 0 est accessible, la population s'éteint forcément sur l'événement $\lim_{n \rightarrow \infty} \#\mathbb{G}_n < \infty$. Puis on justifie que si la population devient grande, alors forcément un des individus va croître comme λ^n puisque la probabilité que ça se produise pour un individu est uniformément positive en fonction du trait dans le cas $d < \infty$. $\square$

Pour étudier la convergence de la distribution des traits vers π , posons maintenant $g(x) = f(x) - \pi(f)$, de sorte qu'on travaille avec g telle que

$$\pi(g) = 0.$$

L'estimation (3.7) devient,

$$\|M_n(\cdot, g)\|_\infty \leq C(\lambda\eta)^n \|g\|_\infty.$$

L'estimation (3.8) donne alors

$$\mathbb{E}_{(\emptyset, x)} \left[\sum_{u \neq v \in \mathbb{G}_n} g(X(u))g(X(v)) \right] \leq C''' \|g\|_\infty^2 \eta^{2n} \sum_{p=0}^{n-1} \lambda^{2p} \eta^{-2p} M_p(x, 1).$$

Utilisons à nouveau (3.7) pour majorer $M_p(x, 1)$ par $\lambda^p(h(x) + C\eta^p)$ puis pour $\lambda > 1$,

$$\eta^{2n} \sum_{p=0}^{n-1} \eta^{-2p} \lambda^p \leq \sum_{p=0}^{\infty} \eta^{-2p} \cdot \lambda^n = \mathcal{O}(\lambda^n).$$

Nous obtenons

$$\mathbb{E}_{(\emptyset, x)} \left[\sum_{u \neq v \in \mathbb{G}_n} g(X(u))g(X(v)) \right] \leq C'''' \|g\|_\infty^2 \lambda^n.$$

Nous aboutissons au résultat suivant, qui assure une convergence L^2 à vitesse géométrique :

Lemme 62. *Soit $g : \mathcal{X} \rightarrow \mathbb{R}$ telle que $\pi(g) = 0$. Alors il existe une constante C telle que pour tous $x \in \mathcal{X}$,*

$$\mathbb{E}_{(\emptyset, x)} \left[\left(\lambda^{-n} \sum_{u \in \mathbb{G}_n} g(X(u)) \right)^2 \right] \leq \frac{C}{\lambda^n}.$$

Ceci entraîne la convergence suivante p.s. quand n tend vers l'infini :

$$R_n = \lambda^{-n} \sum_{u \in \mathbb{G}_n} g(X(u)) \longrightarrow 0,$$

En effet, on peut le justifier utilisant l'inégalité de Cauchy Schwarz pour établir que $\mathbb{E}(\sum_{n \geq 1} |R_n|) < \infty$. Ainsi pour tout f ,

$$\lambda^{-n} \sum_{u \in \mathbb{G}_n} f(X(u)) - \lambda^{-n} \# \mathbb{G}_n \pi(f) \xrightarrow{n \rightarrow \infty} 0 \quad \text{p.s. et dans } L^2.$$

En prenant $f = h$ dans cette limite et en rappelant que le premier terme devient alors une martingale convergeant vers W (Lemme 61), on obtient successivement

Proposition 63. *Les converges suivantes ont lieu p.s. et dans L^2 :*

$$\frac{\# \mathbb{G}_n}{\lambda^n} \xrightarrow{n \rightarrow \infty} W$$

et pour toute fonction $f : \mathcal{X} \rightarrow \mathbb{R}$

$$\frac{\sum_{u \in \mathbb{G}_n} f(X(u))}{\lambda^n} \xrightarrow{n \rightarrow \infty} \pi(f) W$$

et pour tout $i \in \mathcal{X}$

$$1_{\mathbb{G}_n \neq \emptyset} \frac{\{u \in \mathbb{G}_n : X(u) = i\}}{\# \mathbb{G}_n} \xrightarrow{n \rightarrow \infty} 1_{W > 0} \pi_i.$$

En utilisant les estimations précédentes, ce résultat peut être complété par des vitesses de convergence géométrique dans L^2 . On peut relâcher l'hypothèse de second moment sur la loi de reproduction, de différentes façons. Par des arguments de troncations par exemple [1]. Suivons ici l'approche du premier chapitre, exploitant une construction spinale.

3.4.2 Temps long et construction spinale avec un moment $L \log L$

La section précédente a montré comment mener efficacement l'analyse en temps long en utilisant les calculs L^2 du chapitre précédent. Cela permet en particulier d'estimer les vitesses de convergence et d'étendre les résultats, notamment à des modèles en environnement changeant. Pour relâcher cette hypothèse de moment, on peut utiliser des techniques de troncation, citons [1]. Comme dans le chapitre 1.4.1, une autre façon de s'en affranchir est l'utilisation de l'approche spinale, qui permet de capter plus globalement des informations généalogiques, en particulier un échantillonnage uniforme.

Suivant [33] et le premier chapitre, construisons un nouvel arbre aléatoire, avec un individu distingué. Cette fois, les individus ont un trait qui se transmet et l'épine va hériter 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. $\hat{\Theta}_x$ pour $x \in \mathcal{X}$. La loi de cette v.a. est donnée par

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

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, pour G mesurable positive ou bornée :

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

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

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

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. $(Y(u) : u \in \mathcal{A}_n)$) les traits des individus à la génération n .

On note pour $n \geq 0$,

$$\bar{X}_n = \{(u, X(u)), u \in \mathcal{T}_n\}, \quad \bar{Y}_n = \{(u, Y(u)), u \in \mathcal{T}_n\}$$

les processus ainsi construits jusqu'à la génération n , c'est à dire les individus u présents dans les n premières générations, et leur trait. Pour tous $n \in \mathbb{N}$ et F fonction bornée à valeurs réelles, on peut montrer par récurrence, en s'inspirant du chapitre 1.2.2 et en y ajoutant les valeurs des traits des individus

$$\mathbb{E}_{(\emptyset, x)} [F(\bar{X}_n, U_n) W_n] = h(x) \mathbb{E}_{(\emptyset, x)} [F(\bar{Y}_n, E_n)],$$

avec

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

la martingale qui converge vers W . En supposant maintenant,

$$\forall x \in \mathcal{X}, \quad \mathbb{E}_{(\emptyset, x)} (\#\mathbb{G}_1 \log(\#\mathbb{G}_1)) < \infty,$$

on peut montrer que la martingale limite W est positive avec probabilité positive, puis qu'elle est positive positive sur l'événement de survie de la population. En effet, comme dans le premier chapitre, la population dans la construction spinale peut être vue comme le processus original plus une immigration faisant intervenir une loi biaisée. On peut ensuite invoquer un résultat de loi des grands nombres pour obtenir la même conclusion qu'en Proposition 63. Nous renvoyons à l'article original [33] pour tous les détails.

3.5 Ergodicité et temps long

Dans la partie précédente, nous avons vu que les éléments propres (λ, h, π) associés au semigroupe du premier moment M : guidaient le comportement asymptotique de M et plus globalement du processus de branchement quand on ajoutait des conditions de

moments. Concrètement, en notant h la fonction positive sur $\mathcal{X}$ et π probabilité sur π telle que $Mh = \lambda h(x)$, $\pi M = \lambda \pi$, l'indépendance venant du branchement fait reposer l'analyse sur la moyenne et

$$M_n(x, f) \sim_{n \rightarrow \infty} \lambda^n h(x) \pi(f)$$

pour tout $x \in \mathcal{X}$ et f fonction sur $\mathcal{X}$. En réalité, dans les preuves, on a aussi exploité le fait que quand $\mathcal{X}$ est fini, ces convergences simples sont automatiquement uniformes :

$$\sup_{x \in \mathcal{X}, |f| \leq 1} \left| \frac{M_n(x, f)}{\lambda^n h(x) \pi(f)} - 1 \right| \xrightarrow{n \rightarrow \infty} 0. \quad (3.9)$$

En dimension infinie $\#\mathcal{X} = \infty$, on peut étendre les résultats de la section précédente sous la condition (3.9). Cette extension est assez directe pour l'approche L^2 et le chapitre 3.4.1 peut être adapté essentiellement en étant un peu moins brutal dans la domination du noyau de covariance. Le relachement du moment d'ordre 2 et l'obtention de conditions $L \log L$ est plus difficile, et on renvoie en particulier à [1] pour un beau résultat donnant une version infinie dimensionnelle du théorème de Kesten Stigum et les lois fortes des grands nombres pour les processus de branchement sous (3.9).

3.5.1 Ergodicité et fonction de Lyapunov

En dimension infinie, l'uniformité de la condition (3.9) ne va pas souvent être vraie. Elle n'est d'ailleurs pas forcément facilement vérifiable même quand elle l'est, sans une bonne connaissance spectrale de l'opérateur M . En terme spectrale, il faut en particulier que la première fonction propre h domine les autres fonctions propres. En terme plus probabiliste, si on prend par exemple le modèle neutre comme dans le chapitre précédent (la loi de reproduction p ne dépend pas de x), h constante est fonction propre et il faut alors que la convergence du processus de Markov le long d'une lignée typique (l'épine) vers sa mesure stationnaire π soit uniforme en la position de départ, c'est à dire un phénomène de descente de l'infini pour la dynamique du trait en temps continu.

Plusieurs autres modes de convergence issues de la théorie spectrale ou de l'étude des chaînes de Markov peuvent être exploitées. Une hypothèse plus souvent vérifiée et vérifiable que l'uniforme convergence (3.9) est l'ergodicité suivante. Il existe $\lambda > 0$ et $\eta \in (0, 1)$, des fonctions mesurables $h : \mathcal{X} \rightarrow (0, \infty)$, $V : \mathcal{X} \rightarrow [1, \infty)$ et une probabilité π sur $\mathcal{X}$ telle que pour tous $x \in \mathcal{X}$,

$$\sup_{|f| \leq V} \left| \frac{M_n(x, f)}{\lambda^n} - h(x) \pi(f) \right| \leq CV(x) a_n, \quad (3.10)$$

avec a_n suite qui tend vers 0 quand n tend vers l'infini. Quand $M(\cdot, \mathcal{X}) = 1$, $M = P$ est un noyau markovien sur $\mathcal{X}$. En prenant $a_n = \eta^n$ avec $\eta \in [0, 1)$, on est alors ramené un résultat de stabilité exponentielle de la chaîne de Markov associée à P . Dans ce cas, $\lambda = 1$ et $h = 1$ et la condition (3.10) devient

$$\sup_{|f| \leq V} \left| P_n(x, f) - \pi(f) \right| \leq CV(x) \eta^n, \quad (3.11)$$

avec π est la loi stationnaire de la chaîne de Markov. On renvoie à [42, 20, 26] pour des livres références sur le sujet. Ils donnent des conditions sur la chaîne de Markov associée

à P pour garantir (3.11). La condition clef qui fait apparaître la fonction de Lyapunov V est $PV \leq \eta V$ en dehors d'un ensemble compact, assurant un retour vers les compacts suffisamment rapide.

Quand M n'est pas un noyau markovien, on peut utiliser les idées des sections précédentes. En particulier quand on dispose d'une fonction harmonique, on peut passer de M à un noyau Markovien P par conjugaison :

$$Pf = M(fh)/h.$$

On renvoie au chapitre 3.3 pour des détails, avec deux exemples illustratifs, et à [21] pour des branchements diffusions. Plus globalement, en utilisant les résultats chapitre 3.2 sur les lignées typiques, on peut lier l'analyse de l'opérateur M itéré n fois à une chaîne de Markov sur $\mathcal{X}$ de noyau P . Ceci donne un chemin pour obtenir (3.10) à partir de (3.11). Avec l'idée d'exploiter les trajectoires pénalisées et la formule de Feynman Kac du chapitre, on pourra consulter [32], avec des conditions impliquant des bornes uniformes sur la survie et le nombre moyen d'enfants, qui passent par une trajectoire Markovienne pénalisée. Pour relâcher notamment ces hypothèses de bornitude, citons [12] qui exploite la lignée inhomogène d'un échantillon uniforme présentée dans le chapitre. Enfin, pour justifier des lois grands nombres fortes et la non dégénérescence sous la condition (3.10), nous renvoyons à [13]

Nous ne nous lancerons pas dans ces travaux généraux, mais nous pouvons présenter un cadre intéressant d'un point de vue des applications, pour lequel l'analyse mathématique est plus directe. Cela correspond à prendre $V = 1$ au dessus et donne ainsi une uniformité en $x \in \mathcal{X}$ dans les convergences.

3.5.2 Conditions de Doeblin et uniform ergodicité

Nous allons étudier un cas où les traits sont "rapidement mélangés". 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, $x \rightarrow P(x, A)$ soit mesurable, on rappelle tout d'abord les notations suivantes. Le réel $\mu(f)$ et la mesure de probabilité μP sont définis par

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

On introduit la norme en variation totale définie par

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

Commençons par le résultat de contraction pour une chaîne de Markov sous la condition de Doeblin, qui correspond à un couplage des trajectoires des processus de Markov.

Lemme 64. *Supposons qu'il existe $c \in (0, 1]$ et ν mesure de probabilité sur $\mathcal{X}$ tels que*

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

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

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

$R(x, \cdot)$ 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. $\square$

Montrons maintenant la conséquence suivante sur la chaîne de Markov inhomogène $Y^{(n)}$ de noyau $(Q_i^{(n)} : i \leq n - 1)$.

Proposition 65. *On suppose qu'il existe $c \in (0, 1]$ et ν mesure de probabilité sur $\mathcal{X}$ tels que pour tout $i \leq n - 1$,*

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

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

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

Démonstration. On note $PQ(x, \cdot) = \int_{\mathcal{X}} P(x, dy)Q(y, \cdot)$ 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 $|f| \leq 1$. La suite $(V_n(f))_n$ est donc 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. $\square$

Rappelant (3.3), nous obtenons ainsi la convergence exponentielle pour la distribution des traits :

$$\mathbb{E}_x(f(Y_n^{(n)})) = \frac{M_n(x, f)}{M_n(x, 1)}$$

Appliquons ce résultat au modèle structuré en âge introduit en début de chapitre : $\mathbb{P}(\Theta^{(a)} = \emptyset) = p_0(a)$, $\mathbb{P}(\Theta^{(a)} = (a+1)) = p_1(a)$ et $\mathbb{P}(\Theta^{(a)} = (a+1, 0)) = p_2(a)$ avec $p_0(a) + p_1(a) + p_2(a) = 1$.

Corollaire 66. *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 $x \in \mathbb{N}$,*

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

Ce résultat correspond à prendre $f = 1_a$ et donne la proportion moyenne d'individus d'âge a . Il se généralise, notamment en montrant la condition de Doeblin sur b générations plutôt que sur 1. En particulier la condition de domination de p_2 par p_1 peut être relâchée.

Démonstration. On commence par montrer qu'il existe $c > 0$ tel 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 $Y_n^{(n)}$:

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

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

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

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

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

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

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 proportion 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'équivalent du nombre d'individus total en utilisant la martingale associée 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 Lyapunov. 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

Approximation en grande population

Dans ce chapitre, on s'intéresse à un régime où la taille de la population est grande et on cherche approcher la dynamique en faisant converger le processus convenablement renormalisé. La limite va être continu en temps et la taille de la population limite n'est plus un nombre entier fini et pour décrire dans un même espace le processus individu centré et sa limite continue, nous allons dans l'ensemble des mesures. Cet espace a aussi l'avantage d'avoir des propriétés agréables de compacité. Concrètement nous allons donc remplacer une collection de traits par la somme des masses Dirac en les traits en question. Dans un premier temps, nous allons donner des résultats généraux, avec en particulier la décomposition du processus renormalisé en une partie prévisible et une partie martingale, qui est la clef pour l'approximation. La partie prévisible donnera la tendance déterministe du système et la partie stochastique les fluctuations et, si elles disparaissent, on obtient une approximation déterministe. Dans un deuxième temps, nous montrerons ainsi la convergence vers une EDP en grande population pour un exemple de reproduction dispersion (marche branchante) dans de grande échelle de temps (ou déplacement rapide). Ces techniques présentées ici dans un cadre de processus de branchement et un exemple spatial sont utiles plus généralement et on renvoie notamment à [10] et les références à l'intérieur.

4.1 Généralités et décomposition semi-martingale

4.1.1 Définitions

On considère maintenant un nombre fini d'individus initiaux et on labelise la population grâce à l'ensemble

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

Introduisons un paramètre échelle $K \geq 1$, qui a vocation typiquement à donner l'ordre de grandeur des tailles de populations impliquées, ce qui peut aller d'une centaine d'individus ($K = 10^2$) pour certaines espèces dans un écosystème, à des milliards ($K = 10^9$) pour des bactéries ou globules rouges. Ce paramètre va affecter à la fois la loi de reproduction et la transmission des traits.

Plaçons nous dans le cas d'un modèle neutre pour simplifier : la variable aléatoire de reproduction L^K à valeurs dans $\mathbb{N}$ dépend de K mais pas du trait. On la suppose aussi intégrable :

$$m_K = \mathbb{E}(L^K) = \sum_{k \geq 0} k p_k^K < \infty,$$

avec $p_k^K = \mathbb{P}(L^K = k)$. Un individu de trait $x \in \mathcal{X}$ est remplacé par une collection $\Theta_x^K = (X_1^{x,K}, \dots, X_{L^K}^{x,K})$ d'individus.

Considérons une population initiale labélisée par $\mathbb{G}_0^K \subset \mathbb{N}$, avec $Z_0^K = \#\mathbb{G}_0^K$ individus et supposons

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

Chaque individu de la population initiale a un trait et on se donne $(X(u) : u \in \mathbb{G}_0^K)$. On introduit une famille de v.a. $(\Theta_x^K(u))_{x \in \mathcal{X}} = (X_1^{x,K}(u), \dots, X_{L^K}^{x,K}(u))_{x \in \mathcal{X}}$ i.i.d. pour $u \in I$ et distribuées comme $(\Theta_x^K)_{x \in \mathcal{X}}$.

Définition 67. Le processus $X^K = (X^K(u), u \in \mathcal{T}^K)$ est défini récursivement par

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

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

$$X^K(uk) = X^{X^K(u),K}(u).$$

Enfin, on définit $\mathcal{T}^K = \cup_{n \in \mathbb{N}} \mathbb{G}_n^K$.

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, en ramenant les mesures à des processus à valeurs réelles :

$$X_n^K(f) = \int_{\mathcal{X}} f(x) X_n^K(dx) = \frac{1}{K} \sum_{u \in \mathbb{G}_n^K} f(X^K(u)).$$

4.1.2 Décomposition en semi-martingale

Commençons par un énoncé général classique. Lorsqu'un processus $(X_n)_{n \in \mathbb{N}}$ réel et intégrable est adapté à une filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$, on peut le décomposer en

$$X = V + M$$

avec $(V_n)_n$ processus prévisible, c'est à dire V_n est $\mathcal{F}_{n-1}$, et $(M_n)_n$ martingale issue de 0. Cette décomposition est unique. En effet, si V est prévisible alors

$$\mathbb{E}[X_n | \mathcal{F}_{n-1}] = V_n + \mathbb{E}[M_n | \mathcal{F}_{n-1}] = V_n + M_{n-1}$$

et donc

$$\mathbb{E}[X_n | \mathcal{F}_{n-1}] - X_{n-1} = V_n - V_{n-1}.$$

Ainsi on obtient V_n comme somme de ses incréments, et en déduit M_n qui vaut $X_n - V_n$,

$$V_n = X_0 + \sum_{k=1}^n (\mathbb{E}[X_k | \mathcal{F}_{k-1}] - X_{k-1}), \quad M_n = \sum_{k=1}^n (X_k - \mathbb{E}[X_k | \mathcal{F}_{k-1}]).$$

On note maintenant $(\mathcal{F}_n^K)_{n \in \mathbb{N}}$ 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 $(X^K(u), u \in \mathbb{G}_0^K)$. On introduit à nouveau l'opérateur pour la moyenne :

$$M^K(x, f) = K \mathbb{E}_{(1,x)}[X_1^K(f)] = \mathbb{E}_{(1,x)} \left[\sum_{u \in \mathbb{G}_1^K} f(X^K(u)) \right] = m^K Q^K(x, f)$$

L'intégrabilité de L^K et la bornitude de f assure l'intégrabilité de $(X_n^K(f))_n$. La décomposition en semimartingale s'écrit alors de la façon suivante ;

Proposition 68. *Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ mesurable bornée et $n \geq 0$. 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 $n \geq 1$ et $(M_n^K(f))_{n \in \mathbb{N}}$ une $(\mathcal{F}_n^K)_{n \in \mathbb{N}}$ -martingale issue de 0. En outre,

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

avec

$$\Delta_k^K(f) = X_{k-1}^K(M^K(., f) - f) \quad \delta_k^K(f) = X_k^K(f) - X_{k-1}^K(M^K(., f)). \quad (4.1)$$

Démonstration. On utilise la décomposition introduite avec comme accroissement pour V ici :

$$\begin{aligned} \Delta_k^K(f) &= \mathbb{E}[X_k^K(f) | \mathcal{F}_{k-1}^K] - X_{k-1}^K(f) \\ &= \frac{1}{K} \mathbb{E} \left[\sum_{u \in \mathbb{G}_{k-1}^K} \sum_{v \in \mathbb{G}_k^K, v \succ u} f(X^K(v)) | \mathcal{F}_{k-1}^K \right] - \frac{1}{K} \sum_{u \in \mathbb{G}_{k-1}^K} f(X^K(u)) \\ &= \frac{1}{K} \sum_{u \in \mathbb{G}_{k-1}^K} M^K(X^K(u), f) - f(X^K(u)) = \frac{1}{K} \sum_{u \in \mathbb{G}_{k-1}^K} g(X^K(u)), \end{aligned}$$

avec $g = M^K(., f) - f(.)$ et on reconnaît $X_{k-1}^K(g)$. □

On suppose dans le reste du chapitre que

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

ce qui assure que $\mathbb{E}((\mathbb{G}_n^K)^2)$ fini pour tout n . On peut ainsi faire une décomposition analogue pour la carré de la martingale.

Proposition 69. *Supposons f mesurable bornée. Alors*

$$M_n^K(f)^2 = \frac{1}{K} \sum_{k=1}^n X_{k-1}^K(G_f^K) + \widetilde{M}_n^K,$$

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

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

Démonstration. On peut utiliser à nouveau la décomposition en semimartingale car $M_n^K(f)^2$ est intégrable, puisque $x \rightarrow M(x, f) - f(x)$ est bornée et $\mathbb{E}((\mathbb{G}_n^K)^2)$ fini. En oubliant K et f dans les notations pour alléger, on obtient

$$M_n^2 = \sum_{k=1}^n \widetilde{\Delta}_k + \widetilde{M}_n,$$

avec $\widetilde{\Delta}_k = \mathbb{E}[M_k^2 - M_{k-1}^2 | \mathcal{F}_{k-1}]$ et $(\widetilde{M}_n)_n$ martingale. 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 :

$$\widetilde{\Delta}_k = \mathbb{E}[2M_{k-1}\delta_k + \delta_k^2 | \mathcal{F}_{k-1}] = \mathbb{E}[\delta_k^2 | \mathcal{F}_{k-1}]$$

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 branchement¹, 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 $\mathbb{E}(R_k(u)^2 | \mathcal{F}_{k-1}) = G_f^K(X(u))$, qui donne

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

avec $\widetilde{M}^K$ martingale issue de 0. □

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

4.1.3 Lignées ancestrales

Utilisons la formule tous pour un ici et composons par une fonction qui permet de renormaliser le processus :

$$Q^K(x, f) = \frac{1}{m_K} M^K(x, f) = \sum_{k \geq 0} \widehat{p}_k^K Q^{(K,k)}(x, f),$$

où $\widehat{p}_k^K = (kp_k^K/m_K)_{k \geq 0}$ est la loi de reproduction biaisée par la taille et $Q^{(K,k)}(x, \cdot)$ est défini par

$$Q^{(K,k)}(x, \cdot) = k^{-1} \sum_{1 \leq i \leq k} \mathbb{P}(X_i^{K,x} \in \cdot | L^K = k).$$

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

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

Soient $F : \mathcal{X}^{n+1} \rightarrow \mathbb{R}$ mesurable bornée et $n \geq 0$.

$$\mathbb{E}_{(1,x)} \left[\sum_{u \in \mathbb{G}_n^K} F(X^K(u), u \preceq u) \right] = m_K^n \mathbb{E}_x [F(Y_0^K, \dots, Y_n^K)].$$

On peut composer F pour normaliser en espace et changer l'échelle de temps. On note pour $u \in \mathbb{G}_n$ et $t \leq n$,

$$X_t^K(u) = X^K(u_t)$$

avec u_t l'ancêtre de u en vie à la génération $[t]$. $X_t^K(u)$ est donc le trait de l'ancêtre de u au temps t . On obtient pour toute fonction $F : \mathbb{D}([0, \infty), \mathbb{R}) \rightarrow \mathbb{R}$ continue bornée et $t \geq 0$,

$$\mathbb{E}_{(1,x)} \left[\sum_{u \in \mathbb{G}_{[v_K t]}^K} F(\varphi_K(X_{v_K s}^K(u)), s \leq t) \right] = m_K^{[v_K t]} \mathbb{E}_x \left[F(\varphi_K(Y_{[v_K s]}^K), s \leq t) \right]$$

avec φ_K une fonction permettant une renormlisation des traits (on peut prendre $\varphi_K = \text{Id}$ pour commencer). Ainsi, si $(\varphi_K(Y_{[v_K t]}^K))_{t \geq 0}$ converge en loi dans $\mathbb{D}([0, \infty), \mathbb{R})$ vers le processus V , v_K tend vers l'infini et

$$v_K(m_K - 1) \rightarrow \alpha$$

quand K tend vers l'infini, alors

$$\mathbb{E}_{(1,x)} \left[\sum_{u \in \mathbb{G}_{[v_K t]}^K} F(\varphi_K(X_{v_K s}^K(u)), s \leq t) \right] \xrightarrow{K \rightarrow \infty} e^{\alpha t} \mathbb{E}_x [F(V_s, s \leq t)].$$

Nous verrons dans la section suivante, un exemple d'une telle convergence avec comme processus limite V le mouvement brownien.

4.2 Convergence vers une EDP pour un modèle de reproduction dispersion

Considérons maintenant un exemple pour spécifier les échelles en jeu et la limite. Le modèle décrit la dispersion et reproduction d'une espèce dans un grand domaine, unidimensionnel, homogène avec des déplacements indépendants pour simplifier :

$$\mathcal{X} = \mathbb{R}, \quad \Theta_x^K = x + \Theta_0^K = \left(x + \frac{D_i}{K} : 1 \leq i \leq L^K \right)$$

avec D_i v.a. i.i.d. à valeurs réelles, distribuées comme D , et D v.a. centrée de variance finie :

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

Dans ce cas,

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

avec toujours $m_K = \mathbb{E}[L^K]$. La formule *tous-pour-un*, nous donne aussi :

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

avec $(S_n^K)_{n \in \mathbb{N}}$ 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 :

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

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 70. 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 $m_K = 1 + \frac{\alpha}{K^2} + o(\frac{1}{K^2})$, i.e.

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

Proposition 71. Pour tout $t \geq 0$,

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

Démonstration. On utilise que

$$\mathbb{E}_{\delta_0}[Y_t^K(f)] = \mathbb{E}_{\delta_0}[X_{[K^2t]}^K(f)] = m_K^{[K^2t]} \times \mathbb{E}_0[f(S_{[K^2t]}^K)]$$

puis en rappelant que $S_{[K^2t]}^K = K^{-1} \sum_{i=1}^{[K^2t]} D_i$,

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

avec à droite une convergence en loi quand K tend vers l'infini, via le théorème central limite. $\square$

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 $(Y_t^K)_{t \geq 0}$ 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'interaction 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 72. *Pour tout $f \in \mathcal{C}_b^{2,+}$,*

$$\sup_{x \in \mathbb{R}} |K^2 (M^K(x, f) - f(x)) - \mathcal{L}f(x)| \xrightarrow{K \rightarrow \infty} 0,$$

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

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

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

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

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

$$\sup_{x \in \mathbb{R}} |\mathbb{E}(f''(v_{x,D}^K) D^2) - f''(x) \sigma^2| \xrightarrow{K \rightarrow \infty} 0$$

et la preuve est complète. $\square$

On suppose maintenant que

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

Lemme 73. *Il existe $C \geq 0$ tel que pour tout $f \in \mathcal{C}_b^{2,+}$, pour tout $k \geq 1$,*

$$\begin{aligned} i) \quad & |X_{k-1}^K(M^K(., f) - f)| \leq C \frac{1 + \|f\|_\infty + \|f''\|_\infty}{K^2} \frac{\#\mathbb{G}_{k-1}^K}{K}. \\ ii) \quad & X_{k-1}^K(G_f^K) \leq C \frac{1 + \|f\|_\infty^2 + \|f'\|_\infty^2 + \|f''\|_\infty^2}{K^2} \frac{\#\mathbb{G}_{k-1}^K}{K}. \end{aligned}$$

Démonstration. Tout d'abord,

$$|X_{k-1}^K(M^K(., f) - f)| \leq X_{k-1}^K(\mathcal{X}) \|M^K(., f) - f(\cdot)\|_\infty \quad \text{et} \quad X_{k-1}^K(\mathcal{X}) = \frac{\#\mathbb{G}_{k-1}^K}{K}.$$

On peut alors utiliser la preuve précédente, à savoir (4.3), qui permet de contrôler $K^2 \|M^K(., f) - f(\cdot)\|_\infty$. Cela donne i).

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

$$\begin{aligned} G_f^K(x) &= \mathbb{E} \left(\left(\sum_{u \in \mathbb{G}_1^K} f(X(u)) \right)^2 \right) - M^K(x, f)^2 \\ &= m_K \mathbb{E}(f^2(x + D/K)) + v_K \mathbb{E}(f(x + D/K))^2 - m_K^2 \mathbb{E}(f(x + D/K))^2. \end{aligned}$$

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

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

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, $|ff''| \leq |f|^2 + |f''|^2$ et $|\mathbb{E}(f(x + D/K)) - f(x)| \leq \|f''\|_\infty / K^2$, ce qui implique

$$\|G_f^K\|_\infty \leq C \frac{1 + \|f\|_\infty^2 + \|f'\|_\infty^2 + \|f''\|_\infty^2}{K^2}.$$

On conclut avec $X_{k-1}^K(G_f^K) \leq X_{k-1}^K(\mathcal{X}) \|G_f^K\|_\infty$. $\square$

On suppose de plus maintenant que la condition initiale converge en loi dans l'espace $(\mathcal{M}, e)$ des mesures positives finies sur $\mathbb{R}$ muni de la topologie étroite, c'est à dire de la topologie la moins fine, rendant continue les applications $\mu \rightarrow \mu(f)$, pour f continue bornée :

$$X_0^K \Longrightarrow \mu_0, \quad \text{quand } K \rightarrow \infty \quad \text{dans } (\mathcal{M}, e). \quad (4.5)$$

En particulier $\#\mathbb{G}_0^K/K$ converge en loi quand K tend vers l'infini vers la v.a. $\mu_0(\mathcal{X}) \in [0, \infty)$. On va montrer un résultat de bornitude en loi, permettant de "localiser le processus" dans les compacts, puis nous montrerons que la partie martingale tend vers 0.

Lemme 74. *Soit*

$$\sigma_N^K = \inf \{n \geq 0 : \#\mathbb{G}_n^K/K \geq N\}.$$

Alors pour tout $T \geq 0$,

$$\sup_K \mathbb{P}(\sigma_N^K \leq [K^2 T]) \xrightarrow{N \rightarrow \infty} 0$$

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

$$\mathbb{E} \left(\sup_{t \leq T} M_{\sigma_N^K \wedge [K^2 t]}^K (f)^2 \right) \xrightarrow{K \rightarrow \infty} 0.$$

Démonstration. Dans cette preuve, c désignera une constante, qui peut changer mais ne dépendra ni de K , ni de N . $(\#\mathbb{G}_n^K)_n$ est un 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

$$\mathbb{E} \left(W_{\sigma_N^K \wedge [K^2 T]}^K \mid \#\mathbb{G}_0^K \right) = \#\mathbb{G}_0^K$$

Or sur l'événement $\{\sigma_N^K \leq [K^2 T]\}$,

$$\#\mathbb{G}_{\sigma_N^K \wedge [K^2 T]}^K \geq KN, \quad m_K^{\sigma_N^K \wedge [K^2 T]} \leq (1 + c/K^2)^{[K^2 T]} \leq \exp(cT)$$

par concavité de $\log(1+x)$. Donc

$$\mathbf{1}_{\{\sigma_N^K \leq [K^2 T]\}} \leq \frac{\#\mathbb{G}_{\sigma_N^K \wedge [K^2 T]}^K}{KN} \cdot \frac{\exp(cT)}{m_K^{\sigma_N^K \wedge [K^2 T]}} = W_{\sigma_N^K \wedge [K^2 T]}^K \frac{\exp(cT)}{KN}$$

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

$$\mathbb{P}(\sigma_N^K \leq [K^2 T] \mid \#\mathbb{G}_0^K) \leq c_T \frac{\#\mathbb{G}_0^K}{KN}$$

ce qui assure que $\sup_K \mathbb{P}(\sigma_N^K \leq [K^2 T])$ tend vers 0 quand N tend vers l'infini, en utilisant que $\#\mathbb{G}_0^K/K$ est stochastiquement borné car convergeant 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 $(M_{n \wedge \sigma_N^K}^K(f))_n$. On obtient

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

De plus

$$\mathbb{E} \left(\left(M_{[K^2 T] \wedge \sigma_N^K}^K(f) \right)^2 \right) = \frac{1}{K} \mathbb{E} \left(\sum_{k=1}^{[K^2 T] \wedge \sigma_N^K} X_{k-1}^K(G_f^K) \right)$$

en utilisant la decomposition semimartingale de M^2 obtenue en Proposition 69. Enfin le lemme précédent et la définition du temps σ_N^K assurent que

$$\frac{1}{K} \sum_{k=1}^{[K^2 T] \wedge \sigma_N^K} X_{k-1}^K(G_f^K) \leq \frac{c}{K^3} \sum_{k=1}^{[K^2 T] \wedge \sigma_N^K} \frac{\#\mathbb{G}_{k-1}^K}{K} \leq \frac{c}{K} \cdot T \cdot N$$

ce qui conclut la preuve. $\square$

4.2.2 Tension de la suite de processus

On rappelle que

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

L'objectif est de montrer la convergence de Y^K dans l'espace $\mathbb{D}([0, \infty), (\mathcal{M}, e))$ des processus càdlàg à valeurs mesures, où on rappelle que $\mathcal{M}$ est l'espace des mesures positives finies et la topologie étroite la topologie la moins fine, rendant continue les applications $\mu \rightarrow \mu(f)$, pour $f \in \mathcal{C}_b$. 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.

Proposition 75. *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. Pour la tension, il s'agit d'obtenir de la bornitude et le contrôle des variations avec grande probabilité uniformément en K . On applique ici 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 $t \leq T$,

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

Le lemme 73 i) et le fait que $\#\mathbb{G}_k^K \leq KN$ avant le temps d'arrêt σ_N^K assure que

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

De plus par l'inégalité de Markov,

$$\mathbb{P} \left(1_{\sigma_N^K > [K^2 T]} \sup_{t \leq T} |M_{[K^2 t]}^K(f)| \geq M \right) \leq \frac{1}{M^2} \mathbb{E} \left(\sup_{t \leq T} M_{[K^2 t] \wedge \sigma_N^K}^K(f)^2 \right)$$

tend vers 0 quand K tend vers l'infini grâce au lemme 74. Enfin $1_{\sigma_N^K > T} |Y_0^K(f)| \leq N \|f\|_\infty$ et on obtient que pour tout N ,

$$\sup_K \mathbb{P} \left(1_{\sigma_N^K > [K^2 T]} |Y_t^K(f)| \geq 3M \right) \xrightarrow{M \rightarrow \infty} 0.$$

De plus la première partie du Lemme 3 assure que

$$\sup_K \mathbb{P} \left(1_{\sigma_N^K \leq [K^2 T]} |Y_t^K(f)| \geq M \right) \leq \sup_K \mathbb{P} (\sigma_N^K \leq [K^2 T]) \rightarrow 0$$

quand N tend vers l'infini et on obtient

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

De même

$$Y_{t+s}^K(f) - Y_t^K(f) = \sum_{k=[K^2 t]+1}^{[K^2(t+s)]} \Delta_k^K(f) + M_{[K^2(t+s)]}^K(f) - M_{[K^2 t]}^K(f)$$

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

$$\sup_{K, S \leq S' \leq S+\delta} \mathbb{P} (|Y_{S'}^K(f) - Y_S^K(f)| \geq \varepsilon) \xrightarrow{\delta \rightarrow 0} 0,$$

où les temps S et S' sont des temps d'arrêt à valeurs dans $[0, T]$. On a ainsi vérifiés conditions d'Aldous et on peut conclure. $\square$

On peut maintenant justifier la tension dans l'espace $(\mathcal{M}, v)$ des mesures finies munie de la convergence vague, c'est à dire la topologie la moins fine rendant continues les applications $\mu \rightarrow \mu(f)$ pour toutes les fonctions f continues à support compact $\mathcal{C}_c$. Il suffit en fait de considérer un ensemble dense pour la norme infinie. Comme sous espace dense, prenons l'espace $\mathcal{C}_c^{2,+} = \mathcal{C}_b^{2,+} \cap \mathcal{C}_c$ des fonctions de classe C^2 à support compact dont la dérivée seconde est Lipschitz.

Corollaire 76. *La suite $(Y^K)_K$ est tendue dans $\mathbb{D}([0, T], (\mathcal{M}, v))$.*

Démonstration. Grâce à la proposition 75, on peut appliquer le critère de tension de Roelly-Coppoletta [47], avec l'espace $\mathcal{C}_c^{2,+}$ $\square$

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. De façon 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. Pour cela, on peut utiliser le théorème 3 de [43]. Ainsi, on sait déjà que que $(Y_{t \in [0, T]}^K)_K$ 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 garantit 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 sont continus.

Le premier point vient de la proposition 75 avec $f = 1$. On peut aussi 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 conséquence de la loi des grands nombres, en exploitant l'indépendance des différentes sous populations initiales. Le deuxième point peut se traiter à nouveau avec la décomposition en semimartingale et le lemme 73. On obtient

Corollaire 77. *La suite $(Y^K)_K$ est tendue dans $\mathbb{D}([0, T], (\mathcal{M}, e))$.*

Plus généralement pour des conditions suffisantes de tension dans un espace de dimension infinie (comme l'espace des mesures), on pourra consulter le livre de Ethier et Kurtz [22]. Deux ingrédients sont en jeu en général, comme ici : vérifier 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 étroite 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 le résultat suivant

Proposition 78. *Pour tout $K \geq 1, t \geq 0$,*

$$Y_t^K(f) = Y_0^K(f) + \int_0^t ds Y_s^K (K^2 (M^K(x, f) - f(x))) + R_t^K(f),$$

avec les convergences en probabilité suivante quand K tend vers l'infini :

$$\begin{aligned} \sup_{t \leq T} |R_t^K(f)| &\rightarrow 0, \\ \sup_{t \leq T} \left| \int_0^t ds Y_s^K (K^2 (M^K(., f) - f)) - \int_0^t ds Y_s^K(\mathcal{L}f) \right| &\rightarrow 0. \end{aligned}$$

Démonstration. En rappelant la proposition 68,

$$\begin{aligned} Y_t^K(f) &= Y_0^K(f) + \sum_{k=1}^{[K^2 t]} \Delta_k^K(f) + M_{[K^2 t]}^K(f) \\ &= Y_0^K(f) + \sum_{k=1}^{[K^2 t]} Y_{k/K^2}^K (M^K(., f) - f) + M_{[K^2 t]}^K(f) \\ &= Y_0^K(f) + \int_0^t ds Y_s^K (K^2 (M^K(., f) - f)) + R_t^K(f), \end{aligned}$$

avec pour tout $t, K, f \in \mathcal{C}_b^{2,+}$,

$$|R_t^K(f)| \leq \|M^K(., f) - f\|_\infty \sup_{s \leq t} Y_s^K(\mathbb{R}) + |M_{[K^2 t]}^K(f)|.$$

Donc

$$1_{\sigma_N^K > [K^2 T]} |R_t^K(f)| \leq \|M^K(\cdot, f) - f\|_\infty KN + |M_{\sigma_N^K \wedge [K^2 t]}^K(f)|$$

On utilise alors le lemme 72 pour majorer $\|M^K(\cdot, f) - f\|_\infty$ par $1/K^2$ à une constante près et le lemme 74 pour la martingale. On obtient que $1_{\sigma_N^K > [K^2 T]} \sup_{t \leq T} |R_t^K(f)|$ tend vers 0 en probabilité quand K tend vers l'infini. En utilisant cette fois la première partie du lemme 74, $\sup_{t \leq T} |R_t^K(f)|$ tend vers 0. Ensuite

$$1_{\sigma_N^K > [K^2 T]} \sup_{t \leq T} \left| \int_0^t ds Y_s^K (K^2 (M^K(\cdot, f) - f) - \mathcal{L}f) \right| \leq T K^2 \|M(\cdot, f) - f\|_\infty N$$

qui tend vers 0 grâce au lemme 72. On conclut à nouveau avec la première partie du lemme 74. $\square$

Nous allons maintenant pouvoir identifier la limite. Grâce au théorème de représentation de Skorokhod, on peut considérer des sous suites convergeant p.s. De plus on peut supposer la limite continue, ce qui permet de s'appuyer sur de la convergence uniforme en temps. Identifier la loi de leur limite permettra de conclure à une convergence en loi.

Lemme 79. *Supposons que Y^K converge p.s. vers Y dans $\mathbb{D}([0, \infty), (\mathcal{M}, e))$ (resp. $\mathbb{D}([0, \infty), (\mathcal{M}, v))$), alors*

$$Y_t(f) = Y_0(f) + \int_0^t ds Y_s(\mathcal{L}f) \quad p.s.$$

pour tout $f \in \mathcal{C}_b^{2,+}$ (resp $f \in \mathcal{C}_c^{2,+}$).

Démonstration. On utilise le lemme précédent pour montrer que

$$\sup_{t \leq T} \left| \int_0^t ds Y_s^K (K^2 (M(\cdot, f) - f)) - \int_0^t ds Y_s(\mathcal{L}f) \right| \xrightarrow{K \rightarrow \infty} 0$$

en probabilité. Pour cela il reste seulement à justifier que

$$\sup_{t \leq T} \left| \int_0^t ds (Y_s^K(\mathcal{L}f) - Y_s(\mathcal{L}f)) \right|$$

tend vers 0 en probabilité quand K tend vers l'infini. C'est une conséquence de la convergence p.s. de Y^K vers Y car $g = \mathcal{L}f$ est continue à support compact pour $f \in \mathcal{C}_c^{2,+}$ (et g continue bornée pour $f \in \mathcal{C}_b^{2,+}$) et

$$(\mu_t)_{t \in [0, T]} \rightarrow \left(\int_0^t ds \mu_s(g) \right)_{0 \leq t \leq T}$$

est continue de $\mathbb{D}([0, \infty), (\mathcal{M}, v))$ (resp. $\mathbb{D}([0, \infty), (\mathcal{M}, e))$) dans $(\mathcal{C}([0, T], \mathbb{R}), \|\cdot\|_\infty)$ pour $f \in \mathcal{C}_c^{2,+}$ (resp $f \in \mathcal{C}_b^{2,+}$). Pour cela on pourra consulter [16]. Reste à invoquer la convergence ponctuelle en temps, qui découle de la convergence uniforme (notons que cette dernière permet de simplifier la convergence précédente), pour conclure. $\square$

L'objectif est maintenant de vérifier qu'il y a une seule valeur d'adhérence Y et obtenir ainsi une convergence.

Lemme 80. 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

$$\mu_t(f) = \mu_0(f) + \int_0^t ds \mu_s(\mathcal{L}f) \quad (4.6)$$

pour tout $f \in \mathcal{C}_b^{2,+}$.

Démonstration. On commence par vérifier qu'une telle solution μ vérifie plus généralement

$$\mu_t(f_t) = \mu_0(f_0) + \int_0^t ds \mu_s(\mathcal{L}f_s + \partial_s f_s), \quad (4.7)$$

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 $t \geq 0$, $f_t(\cdot) \in \mathcal{C}_b^{2,+}$ et pour tout $x \in \mathbb{R}$, $f_t(x)$ 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 $(\mu, f) \rightarrow \mu(f)$ pour montrer que

$$\partial_t \mu_t(f_t) = [\partial_t \mu_t](f_t) + \mu_t(\partial_t f_t).$$

Ensuite on considère le semigroupe du mouvement brownien B , $P_t f(x) = \mathbb{E}_x(f(\sigma B_t))$, qui peut d'ailleurs se calculer explicitement ici (mais la méthode est générale). Le générateur du mouvement brownien vaut $\mathcal{A}f(x) = \partial_t P_t f(x)|_{t=0} = \sigma^2 f''(x)/2$ 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 $f \in \mathcal{C}_0^{2,b}$, $f_t \in \mathcal{C}_0^{2,2,+}$. En considérant deux solutions μ et ν de (4.7) appliquée à cette fonction, la différence $\gamma = \mu - \nu$ vérifie pour tout $t \geq 0$,

$$\gamma_t(f_t) = \gamma_t(f) = \alpha \int_0^t ds \gamma_s(f_s).$$

Par ailleurs, introduisons la distance en variation totale (on pourrait ici procéder sans, mais pour les généralisations à des modèles plus compliqués cette approche est intéressante)

$$\|\gamma\|_{V.T.} = \sup_{\|f\|_\infty \leq 1} |\gamma(f)| = \sup_{f \in \mathcal{C}_b^{2,+}, \|f\|_\infty \leq 1} |\gamma(f)|.$$

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. $\square$

En combinant les résultats de tension, identification et unicité, nous avons ainsi obtenu :

Théorème 81. Supposons que Y_0^K converge étroitement vers μ_0 mesure de masse finie quand $K \rightarrow \infty$ 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

$$Y_t(f) = Y_0(f) + \int_0^t ds Y_s(\mathcal{L}f) \quad (f \in \mathcal{C}_b^{2,+}).$$

Notons que quand Y_t est à densité (ce que l'on peut montrer dès l'instant initial 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] S. Asmussen, and H. Hering (1976). Strong limit theorems for general supercritical branching processes with applications to branching diffusions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 36 195-212.
- [2] K. B. Athreya, P. E. Ney (2004). *Branching processes*. Dover Publications, Inc., Mineola, NY.
- [3] K. B. Athreya, S. Karlin (1971). On branching processes with random environments I : Extinction probability. *Ann. Math. Stat.* 42. 1499-1520.
- [4] K.B. Athreya (2000). Change of Measures for Markov Chains and the LlogL Theorem for Branching Processes. *Bernoulli* 6(2). DOI : 10.2307/3318579.
- [5]
- [6] J.-B. Baillon, P. Clément, A. Greven, F.A. den Hollander (1993). Variational approach to branching random walk in random environment. *Ann. Probab.* 21, no. 1, 290-317.
- [7] V. Bansaye (2008). Proliferating parasites in dividing cells : Kimmel's branching model revisited. *Annals of Applied Probability*, Volume 18, Number 3, 967-996.
- [8] V. Bansaye. Cell contamination and branching process in random environment with immigration, *Adv. Appl. Probab.*, Volume 41, Number 4, 2009.
- [9] V. Bansaye, A. Lambert (2013) New approaches to sourcesink metapopulations decoupling demography and dispersal. *Theoret. Popul. Biol.* 88 31-46.
- [10] V. Bansaye, S. Méléard (2015). Some stochastic models for structured populations : scaling limits and long time behavior. Springer, MBI. Arxiv version via <https://arxiv.org/abs/1506.04165>.
- [11] V. Bansaye. Ancestral lineage and limit theorems for branching Markov chains. *Journal of Theoretical Probability*, 2018.
- [12] V. Bansaye, B. Cloez, P. Gabriel and A. Marguet. A non-conservative Harris ergodic theorem. *Journal London Math. Soc.* 2, 2022, 1-52.
- [13] V. Bansaye, T. Berah, B. Cloez (2024+). On the strong law of large numbers and $L \log L$ condition for general branching processes. *In preparation*.
- [14] J. Bertoin (2006). Random fragmentation and coagulation processes. Cambridge Studies in Advanced Mathematics, 102. Cambridge University Press, Cambridge, 2006.
- [15] B. Chauvin (1988). Arbres et processus de branchement. Thèse.
- [16] P. Billingsley (1968). Weak convergence of probability measures. Wiley Series in Probability and Mathematical Statistics.

- [17] 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.
- [18] Cowan, R., and Staudte, R. G. (1986). The bifurcating autoregression model in cell lineage studies, *Biometrics*, 42, 769-783.
- [19] A. Dembo, O. Zeitouni (1998). Large deviations techniques and applications. Applications of Mathematics (New York) 38 (Second edition ed.).
- [20] R. Douc, E. Moulines, P. Priouret, P. Soulier (2018). Markov chains. Springer.
- [21] J. Engländer, S. C. Harris, and A. E. Kyprianou. Strong law of large numbers for branching diffusions. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(1) :279298, 2010.
- [22] S. N. Ethier and T. G. Kurtz. Markov processes. Characterization and convergence. Wiley Series in Probability and Mathematical Statistics : Probability and Mathematical Statistics, New York, 1986.
- [23] J. Geiger, G. Kersting, V. A. Vatutin (2003). Limit theorems for subcritical branching processes in random environment. *Ann. Inst. Henri Poincaré* 39, no. 4. 593-620.
- [24] J. Guyon (2007). Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging, *Ann. Appl. Probab.*, 17, 15381569.
- [25] P. Haccou, P. Jagers, V.A. Vatutin. *Branching Processes : Variation, growth and extinction of populations*, Cambridge University Press, 2005.
- [26] M. Hairer and J. C. Mattingly. Yet another look at Harris ergodic theorem for Markov chains. *Seminar on Stochastic Analysis, Random Fields and Applications VI*, volume 63 of *Progr. Probab.*, pages 109117. Birkhäuser/Springer Basel AG, Basel, 2011.
- [27] T. E. Harris (1958). The theory of branching processes. *Springer-Verlag Berlin and Heidelberg GmbH and Co. K*
- [28] G. Keller, G. Kersting, and U. Rosler. On the asymptotic behaviour of discrete time stochastic growth processes. *Ann. Probab.*, 15(1) :305-343, 1987.
- [29] G. Kersting. On recurrence and transience of growth models. *J. Appl. Probab.*, 23(3) :614-625, 1986.
- [30] G. Kersting, V. A. Vatutin (2017). Discrete Time Branching Processes in Random Environment. Wiley.
- [31] M. Kimmel, D. E. Axelrod, 2002. Branching processes in biology. Volume 19, Interdisciplinary Applied Mathematics. New York : Springer-Verlag. *Branching Processes in Biology*, Springer, 2002.
- [32] I. Kontoyiannis and S. P. Meyn (2012). Geometric ergodicity and the spectral gap of non-reversible Markov chains. *Probab. Theory Related Fields*, 154(1-2) :327339.
- [33] 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.
- [34] A. Lambert (2008). Population Dynamics and Random Genealogies. Stochastic models, Vol. 24.
- [35] J. Lamperti. Continuous state branching processes. *Bull. Amer. Math. Soc.*, 73, 382–386, 1967.
- [36] J. Lamperti. The limit of a sequence of branching processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 7 :271–288, 1967.

- [37] R. Lande (1993). Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *The American Naturalist* 911–927.
- [38] R. Lande, S. Engen, B-E Saether (2003). *Stochastic Population Dynamics in Ecology and Conservation*. Oxford University Press.
- [39] Z. Li, 2022. Measure-Valued Branching Markov Processes (Probability Theory and Stochastic Modelling, 103) 2nd ed.
- [40] Maxime Ligonnière. Thesis.
- [41] 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.
- [42] S. Meyn and R. L. Tweedie (2009). Markov chains and stochastic stability. Cambridge University Press, Cambridge, second edition.
- [43] S. Méléard and S. Roelly. Sur les convergences étroite ou vague de processus à valeurs mesures. *C. R. Acad. Sci. Paris Sér. I Math.* 317, 785–788, 1993.
- [44] C. J. Mode. *Multitype Branching Processes : Theory and Applications* (1971). Modern analytic and computational methods in science and mathematics. American Elsevier Publishing Company.
- [45] B. Porro, N. Alloncle, N. Bierne. Sophie Arnaud-Haond. Connectivité et protection de la biodiversité marine : Dynamique spatiale des organismes marins. <https://hal.archives-ouvertes.fr/hal-02404006>.
- [46] E. Pardoux (2016). Probabilistic Models of Population Evolution : Scaling Limits, Genealogies and Interactions. Springer, MBI.
- [47] S. Roelly-Coppoletta. A criterion of convergence of measure-valued processes : application to measure branching processes. *Stochastics*, 17(1- 2) :4365, 1986.
- [48] Z. Shi (2008). *Random walks and trees*. Lecture notes, Guanajuato, Mexico, November 3-7.
- [49] 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.
- [50] V. A. Vatutin, A. M. Zubkov. (1987). Branching processes. I. *Journal of Soviet Mathematics*. Vol. 39, pages 2431-2475.
- [51] V. A. Vatutin, A. M. Zubkov. (1993). Branching processes. II. *Journal of Soviet Mathematics*. Vol. 63, pages 3407-3485.