Chapitre 6
Arbres aléatoires et processus de branchement

1 Motivations. Premiers arbres

Qu'elles viennent des mathématiques ou d'autres sciences que l'informatique, les raisons et motivations pour étudier les arbres sont nombreuses, citons parmi elles :

- le fait que les arbres constituent une simplification par rapport à une structure de réseau (verres de spin en physique) ou par rapport à un graphe (notion d'arbre couvrant), tout en restant une structure riche en dépendances.

- une motivation historiquement importante qui fut l'étude des modèles de ramification ou de reproduction, par Bienaymé, Galton et Watson au $XIX^{\text{ème}}$ siècle. Ces questions se posent maintenant naturellement en biologie ou en physique.

- le *codage* d'ensembles compliqués de $\mathbb{R}^{\mathrm{d}}$, par un arbre, simplifie souvent la compréhension de la géométrie de l'ensemble. La théorie des fractales aléatoires ressort de ce domaine.

Un peu de vocabulaire sera nécessaire, pour être sûrs de ce dont on parle. Il s'agit souvent du langage issu de l'étude des graphes et de certaines approches spécifiques aux graphes. On pourra se reporter à [S] ou [B] pour en savoir plus sur ce sujet.

1.1 Arbre comme graphe particulier

Un **graphe** $\mathcal{G}$ (fini ou infini) sur un ensemble N est un couple (N, E) où E est un ensemble de paires de N. Les éléments de N s'appellent les **sommets** du graphe et les paires en sont les **arêtes**.

Le nombre d'arêtes contenant le sommet x s'appelle le **degré** de x. Un **sommet terminal** ou **feuille** est un sommet de degré 1.

Un **chemin** est une suite d'arêtes adjacentes. Un **cycle** est un chemin fermé.

Un graphe est **connexe** si deux sommets peuvent être joints par au moins un chemin.

Un **arbre**, au sens général venant de la théorie des graphes, appelé aussi **arbre libre** (dans [10] ou [8]) est un **graphe connexe sans cycle**.

On appelle **distance** de deux sommets a et b d'un arbre la longueur du seul chemin joignant a et b.

On appelle **arbre couvrant** (en anglais spanning tree) d'un graphe un graphe partiel qui est un arbre.

Les sommets d'un arbre s'appellent aussi les **noeuds** de l'arbre.

<u>Exercice</u>

Tout arbre fini a au moins deux feuilles.

<u>Démonstration</u> : prendre les extrémités d'un chemin maximal de l'arbre. Elles sont de degré 1 car n'ont pas d'autre voisin que leur voisin immédiat sur le chemin car sinon il y aurait un cycle ou bien le chemin ne serait pas maximal.

Proposition 1 *Deux quelconques des trois conditions suivantes entrainent la troisième et signifient que le graphe $\mathcal{G}$ sur N de cardinal n est un arbre.*

(i) $\mathcal{G}$ est connexe

(ii) $\mathcal{G}$ est sans cycle

(iii) $\mathcal{G}$ possède $n-1$ arêtes.

<u>Démonstration</u> : par récurrence sur n.

<u>Exercice</u> (Théorème de Cayley (1889))

Le nombre d'arbres sur un ensemble N de cardinal n vaut n^{n-2}.

<u>Démonstration</u> : de type combinatoire (voir par exemple Comtet [7] p. 79.).

Un **graphe orienté** ou **digraphe** sur un ensemble fini N est un couple (N, E) où E est un ensemble de paires ordonnées de N. Les arêtes deviennent orientées (et s'appellent alors aussi des arcs). Un **arbre orienté** ou **arbre enraciné** (rooted tree) est un graphe orienté qui possède un sommet spécifique, la racine, tel que

(i) tout sommet différent de la racine est relié à la racine par un unique arc

(ii) la racine n'est l'origine d'aucun arc.

Un arbre est **ordonné** quand on distingue l'ordre des sous-arbres d'un sommet. On parle aussi d'arbre **planaire**. Il y a ainsi pour l'instant trois sortes d'arbres : libre, enraciné (=orienté = non ordonné) et ordonné. Sans mention explicite, on appellera arbre la dernière catégorie. C'est ce que l'on fera tout au long de ce chapitre et en particulier dès que l'on fera des probabilités.

En résumé, pour nous, et dans la littérature en général ([10], [9] par exemple), arbre = arbre général = arbre planaire = arbre enraciné ordonné = arbre probabiliste.

Une **forêt** est un ensemble ordonné de zéro ou plusieurs arbres disjoints. Ainsi l'ensemble des descendants de n'importe quel nœud d'un arbre est une forêt.

Exercice : faire un dessin avec trois arbres ordonnés différents, deux arbres enracinés (orientés) différents et un seul arbre libre.

Un **arbre binaire**, appelé aussi arbre de Catalan, est soit vide, soit constitué d'une racine et de deux arbres binaires appelés sous-arbre gauche et sous-arbre droit. Cette définition est récursive. Notez bien qu'*elle ne produit pas un arbre binaire comme un cas particulier d'arbre...* Néanmoins nous verrons de forts liens entre les deux (le "lemme de correspondance" plus bas). Un arbre binaire est *complet* si chaque sommet possède 0 ou 2 descendants.

Exercice (Théorème de Konig (1926)) : Soit A un arbre enraciné infini dont tout sommet est de degré fini. Alors il existe un chemin infini partant de la racine.
Démonstration : évident en construisant ce chemin.

1.2 Lire un arbre

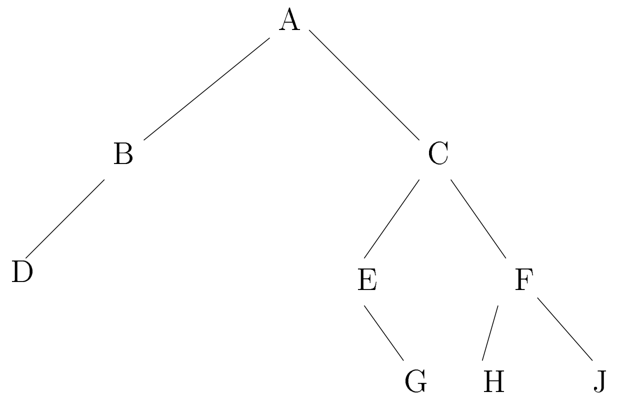
Il y a plusieurs façons de lire un arbre, c'est-à-dire d'énumérer les sommets d'un arbre qui fournissent des bijections entre une structure bi-dimensionnelle, celle de l'arbre et une structure unidimensionnelle, celle de l'énumération. Pour un arbre binaire, on définit

- l'ordre préfixe : d'abord la racine puis lecture en ordre préfixe du sous-arbre gauche puis lecture en ordre préfixe du sous-arbre droit. Ce parcours de l'arbre s'appelle parcours en profondeur (depth first search).

- l'ordre suffixe : lecture en ordre suffixe du sous-arbre gauche, lecture en ordre suffixe du sous-arbre droit, puis racine.

- l'ordre symétrique (inorder en anglais) : ordre symétrique sur le sous-arbre gauche, la racine, ordre symétrique sur le sous-arbre droit.

- l'ordre hiérarchique : la racine, puis la première génération de gauche à droite, puis la seconde génération de gauche à droite, etc... Ce parcours de l'arbre s'appelle parcours en largeur (breadth first search).
Exemple :



ordre préfixe : ABDCEGFHJ
ordre suffixe : DBGEHJFCA
ordre symétrique : DBAEGCHFJ
ordre hiérarchique : ABCDEFGHJ

Pour lire un arbre *non binaire*, les différents types de lecture se généralisent clairement, sauf l'ordre symétrique.

1.3 Dénombrement. Arbres de Catalan

La méthode classique pour compter des objets combinatoires, ici des arbres, utilise les fonctions génératrices ; voyons sur l'exemple des arbres de Catalan comment fonctionne cette méthode.

Soit C_n le nombre d'arbres binaires à n sommets, et soit

$$C(z) := \sum_{n \geq 0} C_n z^n$$

la série génératrice associée.

La méthode consiste à écrire une relation de récurrence sur les C_n et à en déduire une équation fonctionnelle sur $C(z)$. Puis il restera à extraire de

$C(z)$ le coefficient du terme de degré n. Ici, $C_0 = 1$ et en raisonnant sur les sous-arbres gauche et droit, on a

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k} \ ,$$

par conséquent

$$\begin{aligned} C(z) &= 1 + \sum_{n \geq 0} C_{n+1} z^{n+1} \\ &= 1 + z \sum_{n \geq 0} \sum_{k=0}^{n} C_k C_{n-k} z^k z^{n-k} \\ &= 1 + zC(z)^2 \end{aligned}$$

Cette équation du second degré en $C(z)$ se résoud formellement en

$$C(z) = \frac{1}{2z}(1 - \sqrt{1 - 4z})$$

(c'est la racine avec le signe moins qui a été gardée, pour tenir compte du fait que $C(0) = 1$)). Maintenant, grâce à la formule de Taylor, on peut extraire le coefficient de z^n dans $C(z)$ (ce coefficient est noté $[z^n]C(z)$) :

$$[z^n]C(z) = C_n = \frac{(2n)!}{n!(n+1)!} = \frac{1}{n+1}\binom{2n}{n}.$$

Ces nombres C_n s'appellent **nombres de Catalan**. Leur comportement asymptotique est donné par la formule de Stirling $n! \sim \sqrt{2\pi n}(n/e)^n$, et on trouve

$$C_n \sim \frac{4^n}{n\sqrt{\pi n}}$$

Proposition 2 *le nombre d'arbres binaires à n sommets, $C_n = \frac{1}{n+1}\binom{2n}{n}$, est égal au nombre d'arbres ordonnés (= planaires) à $n+1$ sommets.*

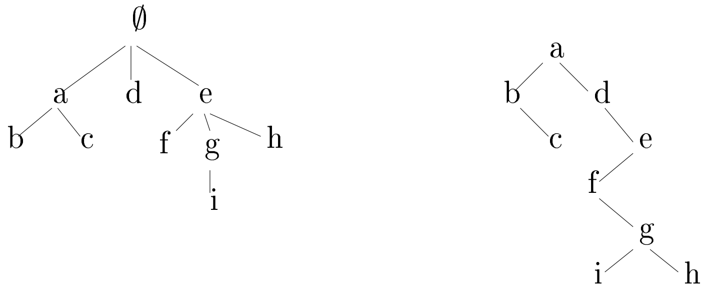
L'explication est simple, c'est qu'il y a une bijection entre les deux types d'arbres, donnée par le lemme suivant.

Lemme 1 *Lemme de correspondance naturelle*

Il existe une bijection entre forêts et arbres binaires qui consiste à associer à tout sommet le fils ainé comme sous-arbre gauche et son frère immédiatement plus jeune comme sous-arbre droit. Elle s'appelle aussi correspondance par rotation et plus explicitement :
si $F = (T_1, T_2, \ldots, T_n)$ *est une forêt,* $B(F)$ *est défini par*

- *Si* $n = 0$*, alors* $B(F)$ *est vide*
- *Si* $n > 0$*, la racine de* $B(F)$ *est la racine de* T_1 *; le sous-arbre gauche de* $B(F)$ *est* $B(T_{11}, T_{12}, \ldots, T_{1m})$*, où* $T_{11}, T_{12}, \ldots, T_{1m}$ *sont les sous-arbres de la racine de* T_1 *; le sous-arbre droit de* $B(F)$ *est* $B(T_2, \ldots, T_n)$.

Exemple :



Remarquons que la lecture de l'arbre binaire $B(F)$ en ordre préfixe (cf 1.2) correspond à l'ordre dynastique.

Remarque-Exercice : On dit qu'un arbre binaire est *complet* si chaque nœud a exactement 0 ou 2 descendants. Appelons a_n le nombre d'arbres binaires complets à n feuilles (appelées parfois aussi sommets externes). En utilisant la même méthode que ci-dessus pour les arbres binaires on trouve que ce nombre est égal à C_{n-1}. Ceci est une conséquence directe du fait que le nombre de feuilles d'un arbre binaire complet est égal au nombre de sommets internes +1 (voir pourquoi). Il y a encore une bijection (la trouver) entre les arbres binaires complets à n feuilles et les arbres binaires à $n-1$ sommets (on en reparlera au paragraphe suivant, cf jolidessin).

1.4 Quelques autres interprétations des nombres de Catalan

Les ensembles suivants contiennent le même nombre d'éléments et ce nombre est C_n. Les bijections sont faciles à identifier et il suffit ainsi de compter le nombre d'éléments d'un de ces ensembles. Cette liste n'est absolument pas exhaustive.

(1) les arbres binaires à n sommets

(2) les arbres (planaires) à $n+1$ sommets (la bijection entre (1) et (2) est donnée par le lemme de correspondance). Ceci explique pourquoi, dans le chapitre 3 précédent, on a trouvé le nombre de Catalan à un près quand on a énuméré les arbres planaires généraux.

(3) les arbres binaires complets à $2n+1$ sommets (la bijection entre (1) et (3) consiste à compléter un arbre binaire avec des feuilles)

(4) les suites de $2n$ parenthèses bien formées. Un suite de parenthèses est dite bien formée si à tout endroit de la suite il y a au moins autant de parenthèses gauches que droites et si au total il y a autant de parenthèses gauches que droites. Ces suites sont évidemment en bijection avec les mots à deux lettres A et B tels que le nombre de A est supérieur ou égal au nombre de B à toute coupure du mot et si le nombre total de A dans le mot est égal au nombre total de B. Ces mots s'appellent mots de Dyck et sont en bijection avec les chemins de Dyck de longueur $2n$ (que l'on reverra au paragraphe ??).

La bijection entre (2) et (4) se construit par récurrence en considérant les sous-arbres issus de la racine.

(5) les problèmes de scrutin à $2n$ bulletins. Il y a deux candidats et le résultat du vote est l'égalité. On cherche le nombre de dépouillements dans lesquels le candidat A est toujours en tête (au sens large). La bijection avec (4) est évidente.

2 Excursions de marches aléatoires. Chemins de Dyck

2.1 Définitions

Définition 1 *On appelle* **chemin de Dyck** *un chemin w de $\mathbb{N} \times \mathbb{N}$ tel que $w = (s_0, s_1, \ldots, s_{2n})$ avec $s_0 = (0,0)$, $s_{2n} = (2n, 0)$ et pour tout $i \leq 2n-1$, si $s_i = (x,y)$, alors $s_{i+1} = (x+1, y+1)$ ou $(x+1, y-1)$.*

Un **mot de Dyck** *est un mot w sur l'alphabet à deux lettres $\{a, b\}$ (ou $\{g, d\}, \{+1, -1\}, \ldots$) tel que*

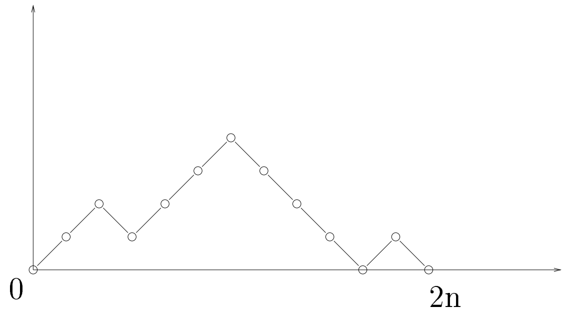
(i) $|w|_a = |w|_b$ et

(ii) pour toute factorisation $w = uv$, alors $|u|_a \geq |u|_b$.
($|w|_a$ désigne le nombre d'occurrence de a dans le mot w).

Il est clair que l'on associe (et l'on ne les distinguera plus dans la suite) les chemins de Dyck de longueur $2n$ et les mots de Dyck de longueur $2n$ par la bijection qui associe à tout pas (s_{i-1}, s_i) la lettre a si $y_i = y_{i-1} + 1$ (pas nord-est) et la lettre b si $y_i = y_{i-1} - 1$ (pas sud-est).

Si l'on autorise dans ce qui précède des paliers horizontaux, on obtient des chemins dits de Motzkin et des mots de Motzkin.

Exemple : mot de Dyck : aabaaabbbbab



2.2 Dénombrement des chemins de Dyck

Soit a_n le nombre de chemins de Dyck de longueur $2n$. Considérons le premier instant de retour à 0, disons $2p$. Si $p = n$, comme le premier et le dernier pas sont imposés, le nombre de chemins de Dyck de cette catégorie est a_{n-1}. Sinon, entre 0 et $2p$, comme le premier et le dernier pas sont imposés, il y a a_{p-1} chemins possibles ; et entre $2p$ et $2n$ il y a a_{n-p} chemins possibles. Finalement on obtient la relation de récurrence

$$a_n = a_{n-1} + \sum_{p=1}^{n-1} a_{p-1} a_{n-p}$$

qui s'écrit encore

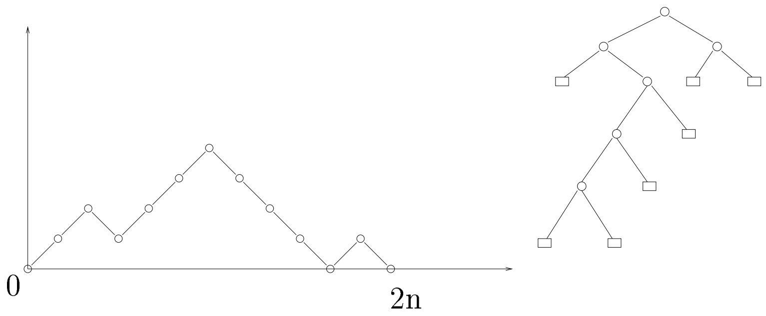
$$a_n = \sum_{k=0}^{n-1} a_k a_{n-1-k}$$

ce qui avec la convention $a_0 = 1$ donne comme solution les nombres de Catalan. On a ainsi montré le

Théorème 1 *Le nombre de chemins de Dyck de longueur $2n$ est égal au nombre de Catalan $C_n = \frac{1}{n+1}\binom{2n}{n}$.*

Ainsi, une bijection existe entre chemins de Dyck et arbres binaires complets à n sommets internes (à $2n+1$ sommets, à $n+1$ feuilles) ; effectivement on montre que l'application qui à tout arbre A binaire complet à n sommets internes associe le mot obtenu en parcourant l'arbre A en ordre préfixe en mettant la lettre a sur un sommet interne et la lettre b sur une feuille, définit une bijection entre les arbres binaires complets à n sommets internes et les mots de Dyck de longueur $2n$.

Exemple : codage : aabaaabbbbab(b)



Cette bijection peut aussi s'exprimer en termes d'excursions d'une marche aléatoire, voyons comment.

2.3 Lien avec une marche aléatoire simple.

Etant donnée une marche aléatoire simple $X = (X_n, n \geq 0)$ sur $\mathbb{R}$, issue de 0, (ce qui signifie que $X_n = X_0 + Y_1 + \cdots + Y_n$, où les variables Y_j sont indépendantes et indépendantes de X_0, et $\mathbb{P}(Y_j = 1) = \mathbb{P}(Y_j = -1) = \frac{1}{2}$), appelons T le premier instant d'atteinte de $]-\infty, 0[$

$$T = \inf\{n \geq 1, X_n < 0\}$$

de sorte que pour un entier n, $T = 2n+1$ et la marche entre 0 et $T-1$ est un chemin de Dyck de longueur $2n$.
La fonction génératrice de T peut se calculer de deux façons, soit par une méthode combinatoire qui devrait être en rapport avec le paragraphe précédent, soit par une méthode probabiliste (à la Le Gall) qui va fournir le même arbre en bijection et qui a pour avantage de permettre d'autres développements.

Méthode combinatoire :

Soit $\tau(s)$ la fonction génératrice de T.

$$\tau(s) = \sum_{n=1}^{\infty} \mathbb{P}(T = n)s^n = \sum_{n=0}^{\infty} \mathbb{P}(T = 2n+1)s^{2n+1}.$$

En se rappelant le dénombrement ci-dessus des chemins de Dyck,

$$\mathbb{P}(T = 2n+1) = \frac{a_n}{2^{2n+1}}$$

et donc

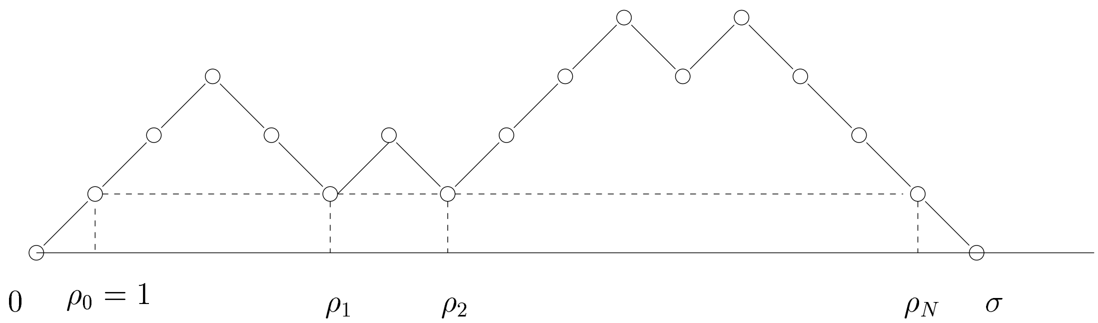
$$\tau(s) = \sum_{n=0}^{\infty} \frac{a_n}{2^{2n+1}} s^{2n+1} = \frac{s}{2} C(\frac{s^2}{4})$$

où C est la fonction génératrice des a_n et on sait que $C(s) = \frac{1-\sqrt{1-4s}}{2s}$. D'où

$$\tau(s) = \frac{1 - \sqrt{1 - s^2}}{s}$$

et l'arbre associé bijectivement à la marche est précisé dans le paragraphe précédent.

Méthode probabiliste :



Etant donnée la marche aléatoire simple $S = (S_n, n \geq 0)$ sur $\mathbb{R}$, issue de 0, appelons σ le premier instant d'atteinte de 0 et considérons les excursions positives de la marche. Elles sont clairement en bijection avec les chemins de Dyck de longueur $\sigma - 2$ correspondant à la marche entre les instants 1 et $\sigma - 1$. De sorte que si on appelle $\mathbb{P}_+$ la loi de la marche $X_{n\wedge\sigma}/X_1 > 0$ c'est à dire la loi d'une excursion positive, nous remarquons que la loi de $\sigma - 1$ sous $\mathbb{P}_+$ est la même que la loi de T. Ainsi, la fonction génératrice de σ est donnée par la fonction génératrice de T, elle-même donnée par la méthode combinatoire :

$$s\mathbb{E}_+(s^{\sigma-1}) = s\mathbb{E}(s^T)$$

$$\mathbb{E}_+(s^\sigma) = 1 - \sqrt{1-s^2}$$

Voyons si l'on peut retrouver cette fonction génératrice avec les excursions. Appelons $\rho_0 = 1, \rho_1, \ldots, \rho_N = \sigma - 1$ les instants de passage en 1. N est aléatoire, c'est le nombre de retours en 1 de l'excursion.

Lemme 2 *La loi de N est donnée par $\mathbb{P}_+(N=p) = \frac{1}{2^{p+1}}$. Sachant $N = p$, les excursions $\tau_1, \tau_2 \ldots, \tau_p$ au-dessus de 1 sont indépendantes et de même loi $\mathbb{P}_+$.*

Démonstration : utiliser la construction et la propriété de Markov forte. □

La correspondance entre arbres et excursions positives de la marche simple se déduit alors de la façon suivante grâce au théorème :

Théorème 2 *Il existe une unique application J de l'ensemble E_+ des excursions positives sur l'ensemble Ω des arbres, bijective, vérifiant les deux conditions*

(i) si $N = 0$, alors $J(\omega) = \{\emptyset\}$ (arbre réduit à sa racine)

(ii) si $N \geq 1$, alors $N_\emptyset(J(\omega)) = N(\omega)$ et $\forall p = 1, 2, \ldots, N(\omega)$, $J(\tau_p\omega) = T_p(J(\omega))$.

De plus J conserve la mesure, c'est-à-dire que $J(\mathbb{P}_+) = \mathbb{P}_{g\acute{e}o}$, où $\mathbb{P}_{g\acute{e}o}$ est la loi d'un processus de Galton-Watson de loi géométrique ($\mathbb{P}(N_\emptyset = k) = \frac{1}{2^{k+1}}$), en particulier ce processsus est critique.

Démonstration : Par récurrence sur σ, on construit J explicitement. □

Remarquons que l'arbre obtenu par cette dernière bijection est un arbre au sens habituel (arbre ordonné du paragraphe 1.1). Grâce au lemme de correspondance 1, on en déduit un arbre binaire, puis un arbre binaire complet et il suffit ensuite de lui appliquer la lecture en préordre pour en déduire un codage qui est un mot de Dyck, donc un chemin de Dyck.

Un exemple permet de visualiser ces bijections.

Par construction de la bijection J, il est clair que sous la loi $\mathbb{P}_+$, le nombre de montées de la marche entre 0 et σ est égal au cardinal de l'arbre $J(\omega)$

donc au cardinal d'un arbre de Galton-Watson de loi géométrique. De plus σ est égal à deux fois le nombre de montées de la marche, par conséquent le calcul de la fonction génératrice de σ se déduira du lemme suivant

Lemme 3 *Pour un arbre de Galton-Watson critique de fonction génératrice f, appelons $I(s)$ la fonction génératrice du cardinal de l'arbre. Alors $I(s)$ est l'unique solution de l'équation*

$$\frac{I(s)}{f \circ I(s)} = s \tag{1}$$

<u>Démonstration</u> :

$$I(s) = \mathbb{E}(s^{|\omega|})$$

et

$$|\omega| = 1 + \sum_{j=1}^{N} |T_j(\omega)|,$$

donc

$$s^{|\omega|} = s \prod_{j=1}^{N} s^{|T_j(\omega)|},$$

et en appliquant la propriété de branchement (**??**),

$$I(s) = s\mathbb{E}(I(s)^N) = sf \circ I(s)$$

d'où l'équation. De plus, pour un processus de branchement critique, la fonction $\frac{u}{f(u)}$ est croissante (calculer la dérivée et remarquer que $\sum_{k \geq 0} p_k(1 - k)u^k$ est un barycentre) et vaut 0 en 0 et 1 en 1, d'où l'unicité. □

Appliquons ceci à la loi géométrique pour laquelle $f(s) = \frac{1}{2-s}$. L'équation 1 est du second degré en $I(s)$ et fournit

$$I(s) = 1 - \sqrt{1-s}.$$

Puisque $\sigma(\omega) = 2|J(\omega)|$, on a bien

$$\mathbb{E}_+(s^\sigma) = 1 - \sqrt{1-s^2}\ .$$

3 Arbres de branchement

3.1 Modèle d'arbre pour les processus de branchement

Le modèle le plus simple que l'on cherche à construire correspond à la situation où l'on se donne une loi de probabilité $(p_k)_{k\in\mathbb{N}}$ sur $\mathbb{N}$, et l'on souhaite étudier le processus Z_n défini par $Z_0 = 1$ (un seul ancêtre) et l'ancêtre a k descendants avec probabilité p_k. Les descendants se reproduisent avec la même loi, indépendamment les uns des autres et de leur ascendant.

Nous allons pour cela définir un espace probabilisé d'arbres $(\Omega, \mathcal{F}, \mathbb{P})$, où $\mathbb{P}$ sera calibré en fonction du phénomène aléatoire que l'on souhaite étudier. Partons de U l'espace des suites finies $u = j_1...j_n$ d'entiers strictement positifs. La suite notée $\emptyset$ (parfois aussi notée $*$) appartient à U. Ainsi

$$U = \{\emptyset\} \cup \bigcup_{n\geq 1} \mathbb{N}^{*n}$$

U est l'ensemble d'"étiquettes" qui nommeront canoniquement les noeuds de l'arbre. La longueur d'une suite est notée $|u|$; la suite obtenue par concaténation de u et v se note uv. $|\emptyset| = 0$.

Définition 2 *Un <u>arbre</u> ω est un sous-ensemble de U qui vérifie les trois conditions suivantes*

(i) $\emptyset \in \omega$ (un arbre n'est jamais vide, il contient au moins la racine)

(ii) $uv \in \omega \Rightarrow u \in \omega$ (on peut remonter le long d'une branche vers la racine)

(iii) $u \in \omega \Rightarrow \exists N_u(\omega) \in \mathbb{N}$ tel que $uj \in \omega, j \geq 1 \Leftrightarrow 1 \leq j \leq N_u(\omega)$ (on peut descendre vers les feuilles)

On voit que la donnée de la suite des $N_u, u \in U$ (nombre de descendants de chaque nœud u), assortie des conditions précédentes, suffit à définir un arbre.

Autrement dit, un arbre "probabiliste" est un arbre planaire, fini ou non, dont les noeuds sont étiquetés par des suites finies d'entiers indiquant la place du noeud dans l'arbre. Dans la suite de ce chapitre, on appellera "arbre" un élément défini comme ci-dessus.

On notera plutôt N au lieu de $N_\emptyset$. N est l'arité de la racine. L'ensemble Ω des arbres est muni de la tribu $\mathcal{F}$ engendrée par les $\Omega_u, u \in U$ où $\Omega_u = \{\omega \in$

$\Omega, u \in \omega\}$. Ainsi les N_u sont des variables aléatoires définies sur Ω_u et N_u désigne le nombre de descendants du noeud u. Pour tout entier n la n-ième génération est

$$z_n(\omega) = \{u \in \omega, |u| = n\}$$

et

$$Z_n(\omega) = Card\ z_n(\omega)$$

(on a $Z_1 = N$). Cette variable aléatoire va beaucoup nous intéresser dans ce chapitre.
Attention : notez bien qu'ici n désigne le niveau courant, le numéro de génération considéré et pas du tout le nombre total de noeuds (comme dans les chapitres précédents), qui peut d'ailleurs être infini.

L'espace des arbres Ω est alors muni d'une filtration $\mathcal{F}_n$ engendrée par les $z_m, 0 \leq m \leq n$ et de translations

$$\begin{array}{llll} T_u : & \Omega & \rightarrow & \Omega \\ & \omega & \mapsto & T_u(\omega) \quad = \{v, v \in U, uv \in \omega\} \end{array}$$

de sorte que T_u désigne le sous-arbre dont la racine est u. Par exemple, $T_1, T_2, \ldots, T_N$ sont les sous-arbres fils de la racine, ils constituent une forêt. On voit facilement que pour tous n et k dans $\mathbb{N}$,

$$Z_{n+k} = \sum_{u \in z_n} Z_k \circ T_u \tag{2}$$

qui est une relation typique et fondamentale dans l'étude des processus de branchement. Un arbre peut être fini ou infini. La hauteur d'un arbre ω est

$$h(\omega) = sup\{|u|, u \in \omega\}.$$

Dans la suite, définir un processus de branchement reviendra à probabiliser l'espace $(\Omega, \mathcal{F})$. Si l'on veut par exemple mettre des "clés" à chaque noeud de l'arbre, il faudra que chaque noeud de l'arbre soit muni d'une "marque" disons X_u qui peut être aussi compliquée et riche que nécessaire (X_u pourra être une variable aléatoire), c'est la notion d'"arbre marqué" que l'on verra au paragraphe 4.1.

3.2 Croissance et régularité d'un arbre (infini). Nombre de branchement

Il est possible de définir la croissance d'un arbre en considérant la n-ième génération : Posons

$$\underline{gr}(\omega) = liminf\{|z_n|^{\frac{1}{n}}\}$$

$$\overline{gr}(\omega) = limsup\{|z_n|^{\frac{1}{n}}\}$$

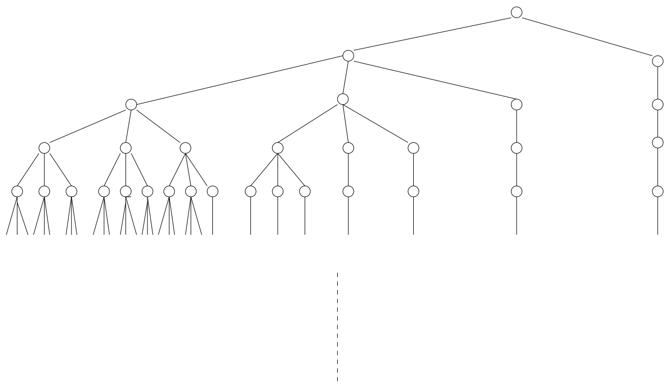
et si $\underline{gr}(\omega) = \overline{gr}(\omega)$, disons que c'est $gr(\omega)$ et appelons cette quantité le <u>taux de croissance</u> de l'arbre.

<u>Exemple 1</u> : l'arbre binaire : gr = 2 (on appelle "arbre binaire" dans ce chapitre l'arbre binaire infini, c'est-à-dire encore l'arbre dans lequel pour tout noeud u de l'arbre $N_u = 2$)

<u>Exemple 2</u> : l'arbre de Galton-Watson qui a en moyenne m enfants ; on verra que gr = m

<u>Exemple 3</u> : l'arbre 3-1 (exemple d'arbre pathologique)

Cet arbre est défini par : à chaque génération n, la moitié gauche de z_n a 3 enfants par noeud et la moitié droite de z_n a 1 enfant par noeud.



On montre facilement par récurrence que $|z_n| = 2^n$ et donc le taux de croissance est $gr(\omega) = 2$. Mais on sent bien que cet arbre est très différent d'un arbre binaire qui a le même taux de croissance. En particulier, dès qu'on n'est pas sur la branche la plus à gauche, on va tomber pour n assez grand, sur des noeuds qui ont toujours un seul descendant.

C'est pourquoi on a besoin d'une notion plus fine que $gr(\omega)$, qui rende compte de la structure de l'arbre. C'est le <u>nombre de branchement</u> que l'on va définir un peu plus loin.

3.2.1 Cutsets

Pour cela, définissons d'abord un <u>cutset</u> ou ensemble de coupure (à rapprocher de la notion de "ligne d'arrêt" qui sera vue plus tard), qui est une notion intuitive et très utile.

Définition 3 *Soit G un graphe connexe, fini, non orienté et soient deux sommets particuliers s et t appelés respectivement source et bord. Un cutset π est un ensemble d'arêtes du graphe G tel que s et t soient dans deux composantes connexes différentes de $G - \pi$.*

Par extension, la source peut être un ensemble de sommets, de même que le bord.

Dans le cas particulier d'un arbre fini, la source s peut être la racine et le bord t peut être l'ensemble des feuilles de l'arbre.

Pour un arbre infini, un cutset π est un ensemble d'arêtes tel que tout chemin infini issu de la racine (toute branche infinie) passe par une arête de π. De plus, sur un arbre, toute arête est de la forme $(u, \text{père de } u)$. Par conséquent, on parle indifféremment d'ensemble d'arêtes ou d'ensemble de sommets. Par exemple, pour tout n, $z_n(\omega)$ est un cutset de ω.

3.2.2 Flots

La notion de flot sur un réseau est utile naturellement pour modéliser l'écoulement d'un liquide à travers des tuyaux, mais aussi un courant électrique à travers un circuit ou la circulation d'information à travers un réseau de communication.

Pour u et v sommets du graphe, (u, v) est un conduit possible, et chaque arête est munie de capacités $C(u, v)$, réelles positives. Par convention, si (u, v) n'est pas une arête du graphe, on dira que $C(u, v) = 0$. Les contraintes sur un tel réseau sont ainsi, d'une part les capacités, d'autre part la loi de conservation qui exprime qu'à travers chaque sommet, le flot entrant doit être égal au flot sortant (c'est la loi de Kirchoff pour les réseaux électriques). La question est : quel est le flot maximal qui peut passer à travers un réseau donné avec des capacités $C(u, v)$ données ?

Définition 4 *Un flot à travers un réseau G de source s et de bord t, est une application f de $G \times G$ dans $\mathbb{R}$, qui à tout couple de sommets (u, v) associe $f(u, v)$ qui vérifie*

- *la contrainte de capacités :*

$$\forall (u,v) \in G \times G,\ f(u,v) \leq C(u,v)$$

- *l'antisymmétrie :* $f(v,u) = -f(u,v)$
- *la loi de Kirchoff :*

$$\forall u \in G - \{s,t\},\ \sum_{v \in G} f(u,v) = 0$$

Par définition le *<u>débit</u>* *du flot* f *est*

$$\mid f \mid = \sum_{v \in G} f(s,v)\ .$$

3.2.3 Flot maximal. Méthode de Ford et Fulkerson

Proposition 3 *Flots et cutsets.*
Si f *est un flot, alors*

$$\pi \ cutset \iff \sum_{(u,v) \in \pi} f(u,v) = \mid f \mid$$

Par conséquent, pour tout flot f, et pour tout cutset π, on aura toujours

$$\mid f \mid \leq \sum_{(u,v) \in \pi} C(u,v)$$

et comme un flot est contraint de passer par tout cutset, on a le théorème suivant appelé théorème "min flow max cut"

Théorème 3 *Théorème de Ford et Fulkerson.*
Le flot maximal φ *passant dans le réseau* G *de source* s *et de bord* t *a pour débit*

$$\mid \varphi \mid = \inf_{\pi \ cutset} \Big(\sum_{(u,v) \in \pi} C(u,v) \Big)$$

L'algorithme dit de Ford et Fulkerson repose sur le théorème ci-dessus et permet de trouver ce flot maximal.

3.2.4 Nombre de branchement

Nous allons maintenant définir le nombre de branchement en utilisant le formalisme précédent. Donnons-nous un arbre infini ω avec la racine comme source et la frontière de l'arbre (l'ensemble des chemins infinis) comme bord. Imaginons que pour un réel $\lambda \geq 1$ donné, l'on mette des capacités $\frac{1}{\lambda^n}$ sur chaque arête reliant un sommet de longueur n et un sommet de longueur $n+1$. Ces capacités peuvent être vues comme des diamètres de tuyaux dans lesquels on veut faire couler de l'eau depuis la racine jusqu'aux feuilles.

$$\forall u \in \omega, \forall j \in \mathbb{N}, j = 1 \dots N_u, \qquad C(u, uj) = \frac{1}{\lambda^{|u|}}$$

Ces tuyaux sont de plus en plus fins en descendant vers le bas de l'arbre. Il est clair que si λ est trop grand, l'eau ne va pas pouvoir couler. Par exemple, pour l'arbre binaire, si $\lambda > 2$, l'eau ne coule pas. Généralisons cette idée.

Par le théorème de Ford et Fulkerson, le flot maximal admissible est

$$\varphi(\lambda) = \inf_{\pi \text{ cutset}} \sum_{v \in \pi} \frac{1}{\lambda^{|v|}}$$

et voyons comment se comporte la fonction φ en fonction de λ.

$$\varphi(1) = \inf_{\pi \text{ cutset}} \; Card(\pi) \; > 0$$

et la fonction φ décroît avec λ. Il existe une valeur critique de λ à partir de laquelle aucun flot ne peut passer (c'est facile à voir si N est borné ou pour un GW). C'est cette valeur critique que l'on appelle nombre de branchement et que l'on note $br(\omega)$.

$$br(\omega) := \sup\{\lambda \geq 1, \text{tel que } \inf_{\pi \text{ cutset}} \sum_{v \in \pi} \frac{1}{\lambda^{|v|}} > 0\}.$$

On conçoit que dès que λ devient plus grand que le nombre moyen de branches issues d'un noeud, le flot ne peut plus passer. C'est pourquoi, pour les arbres assez "réguliers", par exemple pour les arbres binaires ou pour les arbres de Galton-Watson, on aura $br(\omega) = gr(\omega)$ et on a dans tous les cas $br(\omega) \leq gr(\omega)$ (démonstration dans Peres [16]). En revanche dans le cas de l'arbre 3-1, on devine que sauf sur la branche la plus à gauche, le nombre de branchement est 1 et effectivement on montre que pour cet arbre, $br(\omega) = 1$.

Pour le prouver, on raisonne sur la dimension de Hausdorff de la frontière de cet arbre, qui est nulle, car la frontière est dénombrable. Or il n'est pas trop difficile de voir que pour un arbre infini, la dimension de Hausdorff de la frontière est égale au logarithme du nombre de branchement.

Les notions de flots et de nombre de branchement interviennent également dans la théorie de la percolation et dans la théorie des marches aléatoires sur un arbre.

3.3 Arbre de Galton-Watson

Proposition-Définition 1 *Soit $p = (p_k)_{k\geq 0}$ une loi de probabilité sur $\mathbb{N}$. Alors il existe une unique probabilité $\mathbb{P}_p$ (notée $\mathbb{P}$ quand il n'y a pas d'ambiguité) sur $(\Omega, \mathcal{F})$ telle que*

(i) N a pour loi p.

(ii) conditionnellement sachant $\{N = j\}$, les variables $T_1, T_2, ..., T_j$ sont indépendantes et de même loi $\mathbb{P}_p$.

Par conséquent, pour tout $n \geq 0$, conditionnellement par rapport à $\mathcal{F}_n$, les variables aléatoires $\{T_u, u \in z_n\}$ sont i.i.d. de loi $\mathbb{P}_p$. Autrement dit encore, pour toutes fonctions mesurables positives $f_u, u \in U$

$$\mathbb{E}_p^{\mathcal{F}_n}(\prod_{u\in z_n} f_u \circ T_u) = \prod_{u\in z_n} \mathbb{E}_p(f_u) \tag{3}$$

La suite $(Z_n)_{n\geq 0}$ s'appelle un processus de Galton-Watson de loi p et d'état initial $Z_0 = 1$. L'égalité 3 s'appelle propriété de branchement.

<u>Démonstration</u>

Dans [15], pages 202, 203. C'est une propriété d'existence de processus de Markov. □

Dans la suite, on utilisera beaucoup la fonction génératrice de la loi p :

$$f(s) = \mathbb{E}_p(s^N) = \sum_{k\geq 0} p_k s^k = \mathbb{E}_p(s^{Z_1}), \quad s \in [0,1]$$

On pose

$$m = f'(1) = \sum_{k\geq 0} kp_k = \mathbb{E}(Z_1).$$

Dans la suite on supposera toujours que la moyenne m est finie (attention, cela ne veut pas dire du tout que N est borné) et on déduit immédiatement de la relation 2 et de la propriété de branchement 3 que

$$\mathbb{E}_p(s^{Z_n}) = f^{(n)}(s)$$

où $f^{(n)}(s) := f \circ f \circ \cdots \circ f(s)$, n fois.

La première question qui se pose est celle de l'extinction du processus. Si

$$q = \mathbb{P}_p(\exists n, Z_n = 0)$$

est la probabilité d'extinction et si le temps d'extinction T est le premier entier n tel que $Z_n = 0$, on a

Proposition 4 *Extinction du processus :*
(i) $q = \lim_{n \to +\infty} f^{(n)}(0)$
(ii) $q = 1 \Leftrightarrow m \leq 1$ *(quand* $p_1 \neq 1$*)*
q *est la plus petite racine dans* $[0,1]$ *de l'équation* $f(s) = s$

Le processus de branchement est dit sous-critique (resp critique, surcritique) lorsque $m \leq 1$ (resp $m = 1$, $m > 1$). Remarquons que si $p_0 = 0$, alors le processus est bien entendu surcritique (on supposera dans tout ce chapitre que l'on n'est pas dans le cas trivial où $p_1 = 1$).
La taille de la population est précisée grâce à la proposition suivante

Proposition 5 *Martingale de Galton-Watson.*

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

est une $\mathcal{F}_n$*-martingale d'espérance égale à 1, qui converge p.s. vers* W *quand* $n \to +\infty$*.*

Démonstration : immédiate avec

$$Z_{n+1} = \sum_{u \in z_n} Z_1 \circ T_u \qquad \square$$

Le lemme suivant qui établit une loi 0-1 pour un processus de Galton-Watson sera souvent utile. Il est appliqué tout de suite pour relier l'extinction de l'arbre à W.

Lemme 4 *On dit qu'une propriété est "héritée" quand, lorsqu'un arbre possède cette propriété, les arbres $T_1, T_2, ..., T_N$ la possèdent aussi, et si tout arbre fini la possède. Toute propriété héritée est de probabilité q ou 1. Par conséquent, conditionnellement à la non-extinction, toute propriété héritée est de probabilité 0 ou 1.*

Démonstration : Soit E l'ensemble des arbres possédant une certaine propriété héritée. Par la propriété de branchement, $\mathbb{P}(E) \leq f(\mathbb{P}(E))$ et comme $\mathbb{P}(E) \geq q$, c'est que $\mathbb{P}(E) \in \{q, 1\}$. □

Comme $\{W = 0\}$ est clairement une propriété héritée, on a la

Proposition 6 $\mathbb{P}(W = 0) = q$ *ou* 1
Soit $W = 0$ *p.s. soit* $\{W > 0\} = \{Z_n \to \infty\} = E^c$.

Par conséquent, pour un processus de branchement sous-critique ou critique, $q = 1$ et $W = 0$ p.s. et pour un processus de branchement surcritique, $\mathbb{P}(W = 0) = q$ et $W > 0$ sur l'ensemble de non-extinction.

La relation suivante qui se déduit immédiatement de la propriété de branchement, est fondamentale dans l'étude de la frontière d'un arbre. Appelons $W_u := W \circ T_u = \lim_{n\to\infty} \frac{Z_n \circ T_u}{m^n}$ de sorte que pour tout entier n, p.s.

$$W = \frac{1}{m^n} \sum_{u \in z_n} W_u \tag{4}$$

La question suivante qui se pose est : dans le cas où W_n converge vers une limite strictement positive, c'est-à-dire dans le cas surcritique, y a-t-il convergence dans L^1 ? Cette question est résolue dans le thérème de Kesten-Stigum du paragraphe **??**. Ce théorème peut être démontré de plusieurs manières. Une démonstration récente est basée sur la description d'un arbre remarquable, l'"arbre biaisé" (cf. Lyons, Pemantle et Peres [13]).

3.4 L'arbre biaisé

Par le théorème de Kolmogorov, toute famille compatible de probabilités Q_n définies sur $(\Omega, \mathcal{F}_n)$ induit une unique probabilité sur $(\Omega, \mathcal{F})$ qui prolonge les Q_n. Ceci permet de définir une probabilité $\widehat{\mathbb{P}}$ sur $(\Omega, \mathcal{F})$ telle que

$$\widehat{\mathbb{P}} = \frac{Z_n}{m^n} \,.\, \mathbb{P} \text{ sur } \mathcal{F}_n, \forall n \in \mathbb{N}$$

autrement dit,

$$\forall n \in \mathbb{N}, \forall A \in \mathcal{F}_n, \quad \widehat{\mathbb{P}}(A) = \int_A \frac{Z_n}{m^n}\, d\mathbb{P}$$

On appellera <u>arbre biaisé</u> l'arbre sous $\widehat{\mathbb{P}}$.

<u>Propriétés de $\widehat{\mathbb{P}}$</u>

$$\begin{aligned} \widehat{\mathbb{P}}(T > n) &= \widehat{\mathbb{P}}(Z_n \neq 0) = \int 1_{Z_n \neq 0} \frac{Z_n}{m^n}\, d\mathbb{P} \\ &= \int 1_{Z_n \neq 0} \frac{Z_n}{m^n}\, d\mathbb{P} + \int 1_{Z_n = 0} \frac{Z_n}{m^n}\, d\mathbb{P} = \int \frac{Z_n}{m^n}\, d\mathbb{P} = \mathbb{E}(\frac{Z_n}{m^n}) = 1 \end{aligned}$$

de sorte que pour tout n, $\{T > n\}$ est de $\widehat{\mathbb{P}}$-probabilité égale à 1 et donc $\widehat{\mathbb{P}}$ est portée par $\{T = \infty\}$. Ainsi, si $m \leq 1$, $\mathbb{P}$ et $\widehat{\mathbb{P}}$ sont étrangères.

Dans le cas $m > 1$, un petit raisonnement d'intégration, non détaillé ici (voir Peres [16]) conduit à la dichotomie (*) suivante :

1. soit on est dans le cas

$$W = 0 \ \mathbb{P}p.s. \Leftrightarrow \mathbb{P}(W = 0) = 1 \Leftrightarrow \mathbb{P} \perp \widehat{\mathbb{P}} \Leftrightarrow W = \infty \ \widehat{\mathbb{P}}p.s.$$

2. soit on est dans le cas

$$\mathbb{E}(W) = 1 \Leftrightarrow \widehat{\mathbb{P}} << \mathbb{P} \Leftrightarrow W < \infty \ \widehat{\mathbb{P}}p.s.$$

Revenons aux cas sous-critique ou critique pour lesquels on peut expliquer la loi biaisée de la façon suivante.

Proposition 7 *Dans le cas $m \leq 1$ et pour tout $A \in \bigcup \mathcal{F}_n$,*

$$\widehat{\mathbb{P}}(A) = \lim_{n \to \infty} \mathbb{P}(A / Z_n \neq 0) \tag{5}$$

Ainsi l'arbre biaisé est obtenu en conditionnant un arbre de Galton-Watson sous-critique ou critique à survivre indéfiniment.

<u>Démonstration</u>

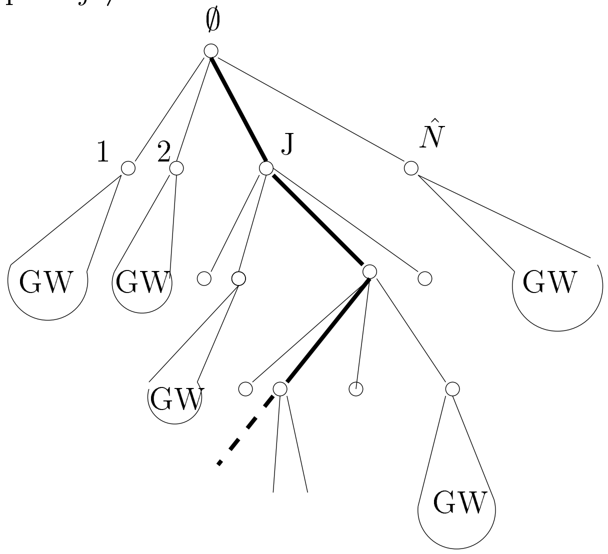
Partir de $\mathbb{P}(A / Z_{n+p} \neq 0)$ pour $A \in \mathcal{F}_p$ et utiliser la propriété de branchement. □

Proposition 8 *Description de l'arbre sous* $\widehat{\mathbb{P}}$: $\forall (A_j)_{j\in\mathbb{N}} \in \mathcal{F}, \forall k \geq 0$,

$$\widehat{\mathbb{P}}(N = k, T_j \in A_j, j = 1, \dots, k) = \frac{kp_k}{m}\,\frac{1}{k}\sum_{j=1}^{k} \mathbb{P}(A_1)\dots\widehat{\mathbb{P}}(A_j)\dots\mathbb{P}(A_k) \tag{6}$$

Démonstration : utiliser la définition de $\widehat{\mathbb{P}}$. □

Autrement dit, pour l'arbre biaisé, la loi de reproduction $\hat{N}$ de l'ancêtre est biaisée et $\mathbb{P}(\hat{N} = k) = \frac{kp_k}{m}$. Puis un descendant, disons J, est choisi au hasard uniformément parmi la première génération et l'arbre T_J sera biaisé alors que les T_j pour $j \neq J$ sont des arbres de Galton-Watson ordinaires.



Il y a donc une branche infinie (une épine dorsale) et des cheveux Galton-Watson qui en partent.

Remarque : interprétation en termes de processus de Galton-Watson avec immigration.

Un processus de G-W avec immigration est caractérisé par une loi de reproduction et une loi d'immigration. A chaque génération il y a immigration de disons Y_n particules où les Y_n sont iid de loi la loi d'immigration. Pendant ce temps, les particules se reproduisent indépendamment les unes des autres avec la loi de reproduction.

On voit ainsi que la description précédente peut aussi se traduire par : $Z_n - 1$ sous $\widehat{\mathbb{P}}$ est un processus de Galton-Watson avec immigration $Y_n = \hat{N}_n - 1$ où $\hat{N}$ a pour loi $\frac{kp_k}{m}$.

3.5 Processus surcritique. Le théorème de Kesten-Stigum

Théorème 4 *Théorème de Kesten-Stigum (1966) :*

Pour un processus de Galton-Watson surcritique (où $Z_0 = 1$ et $1 < m < \infty$), les assertions suivantes sont équivalentes

(i) $\mathbb{P}(W = 0) = q$

(ii) $\mathbb{E}(W) = 1$

(iii) $W_n := \frac{Z_n}{m^n} \xrightarrow{L^1} W$

(iv) (W_n) *est uniformément intégrable*

(v) $\mathbb{E}(\sup_n W_n) < +\infty$

(vi) $\mathbb{E}(Nlog^+ N) < +\infty$*, c'est-à dire* $\sum_{k=2}^{\infty} p_k(klogk) < \infty$

<u>Démonstration</u> :

Les équivalences entre (ii) (iii) et (iv) sont des résultats généraux. (ii) $\Rightarrow$ (i) se déduit facilement du lemme 4 sur les propiétés héritées : puisque $W = 0$ est une propriété héritée, $\mathbb{P}(W = 0) = q$ ou 1. (v) $\Rightarrow$ (iv) est trivial. Montrons (ii) $\Rightarrow$ (v). Cela repose sur le lemme suivant dont la démonstration se trouve dans [1] page 27.

Lemme 5 *Supposons que* $\mathbb{P}(W = 0) = q$*. Posons* $M = sup_n W_n$ *Alors il existe* A *et* $B > 0$ *tels que* $\forall x \geq 1, \forall n \in \mathbb{N}$

$$\mathbb{P}(W \geq Ax) \geq B\mathbb{P}(M \geq x)$$

Avec ce lemme, on a

$$\begin{aligned} \mathbb{E}(M) &= \int_0^{\infty} \mathbb{P}(M \geq x)dx \leq 1 + \frac{1}{B}\int_1^{\infty} \mathbb{P}(W \geq Ax)dx \\ &\leq 1 + \frac{1}{B}\mathbb{E}(\frac{W}{A}) = 1 + \frac{\mathbb{E}(W)}{AB} < \infty \end{aligned}$$

et donc (v) est vérifié.

On va voir que la suite de la démonstration du théorème de Kesten-Stigum se déduit de l'étude de l'arbre biaisé. On a immédiatement que :

$$\text{(vi) } \mathbb{E}(Nlog^+ N) < +\infty \Leftrightarrow \mathbb{E}(log\hat{N}) < \infty \tag{7}$$

où $\hat{N}$ est la v.a. à valeurs dans $\mathbb{N}$ de loi $\mathbb{P}(\hat{N}=k)=\frac{kp_k}{m}$ que l'on a définie dans l'arbre biaisé. Le lemme suivant est un exercice d'intégration et il est nécessaire pour montrer le théorème de Seneta qui suit.

Lemme 6 *Soient $X, X_1, X_2, \ldots$ iid, positives. Alors p.s.*

$$\limsup_{n\to\infty} \frac{1}{n} X_n = 0 \ si\ E(X) < \infty$$

$$\limsup_{n\to\infty} \frac{1}{n} X_n = \infty \ si\ E(X) = \infty$$

Admettons le théorème suivant sur les processus avec immigration, dû à Seneta (1970) (démonstration dans Peres [16], qui utilise le lemme ci-dessus).

Théorème 5 *Théorème de Seneta (1970) :*

Soit Z_n les tailles des générations dans un processus de GW avec immigration Y_n. On suppose que m, moyenne de la loi de reproduction est >1. Si $\mathbb{E}(log^+ Y)<\infty$, alors la limite de Z_n/m^n existe et est finie p.s. Si $\mathbb{E}(log^+ Y)=\infty$, alors $\limsup Z_n/c^n=\infty$ pour toute constante $c>0$.

Il reste à montrer (vi) $\Rightarrow$ (ii) et (i) $\Rightarrow$ (vi).

Supposons (vi). Alors $\mathbb{E}(log\hat{N})<\infty$. Or sous $\widehat{\mathbb{P}}$, $Z_n = 1+(Z_n-1)$ où Z_n-1 est un processus de Galton-Watson avec immigration $\hat{N}-1$. Par le théorème 5 d'immigration de Seneta, on en déduit $W<\infty$ $\widehat{\mathbb{P}}p.s.$, et grâce à la dichotomie (*), donc $\mathbb{E}(W)=1$, c'est-à-dire (ii).

Inversement si l'on n'a pas (vi), alors par le même théorème, $W = \infty$ $\widehat{\mathbb{P}}p.s.$, et donc $\mathbb{P}(W=0)=1\neq q$ puisque $m>1$ donc (i) n'est pas vérifié. Ce qui termine la démonstration du théorème de Kesten-Stigum. □

3.6 Processus sous-critique ou critique

Le processus s'éteint presque sûrement. La question est de savoir à quelle vitesse et quel est son comportement avant l'extinction. Nous donnons ci-dessous les résultats principaux. Les preuves reposent sur l'utilisation de la fonction génératrice f et elles ont aussi été revisitées avec l'arbre biaisé.

3.6.1 Cas sous-critique

Théorème 6 *Théorème de Yaglom (1947, puis améliorations en 1967) :*

Pour un processus de Galton-Watson sous-critique (où $0 < m < 1$), il y a une loi limite du processus conditionné par la non extinction :

$$\mathbb{P}(Z_n = j / Z_n > 0) \overset{n \to \infty}{\longrightarrow} Q(j)$$

où Q est une loi de probabilité sur $\mathbb{N}$ non dégénérée.

On montre aussi que le taux de décroissance de $P(Z_n > 0)$ (qui est majoré par $E(Z_n) = m^n$) est bien de l'ordre de m^n, sous condition de moment.

Théorème 7 *(1967)*

La suite de terme général $P(Z_n > 0)/m^n$ est décroissante. Dans le cas sous-critique $m < 1$, les propriétés suivantes sont équivalentes

(i) $\lim\limits_{n\to\infty} \dfrac{\mathbb{P}(Z_n > 0)}{m^n} > 0$

(ii) $\sup\limits_n \mathbb{E}(Z_n / Z_n > 0) < \infty$

(iii) $\mathbb{E}(N log^+ N) < \infty$.

3.6.2 Cas critique

Théorème 8 *:*

On suppose $m = 1$ et $\sigma^2 \;:= Var(N) \leq \infty$. Alors

$$\lim_{n\to\infty} n\mathbb{P}(Z_n > 0) = \frac{2}{\sigma^2}$$

et si $\sigma^2 < \infty$, on a la loi limite de Yaglom : la loi conditionnelle de $\frac{Z_n}{n}$ converge vers une loi exponentielle de moyenne $\frac{\sigma^2}{2}$. Si $\sigma = \infty$, alors cette loi conditionnelle converge vers $+\infty$.

4 Processus spatiaux

4.1 Arbres marqués

Pour des modèles plus complexes que celui de Galton-Watson, on a besoin de “marquer” les branches de l'arbre par des variables aléatoires.

Définition 5 *Soit Γ un espace mesurable. Un* <u>*arbre marqué*</u> *est un couple*

$$\bar{\omega} = (\omega, (\gamma_u)_{u\in\omega})$$

où ω est un arbre et pour tout $u \in U$, $\gamma_u \in \Gamma$. On notera $\bar{\Omega}$ l'espace des arbres marqués par Γ et comme dans l'espace des arbres Ω, on appellera $\bar{\Omega}_u$ l'ensemble des arbres marqués contenant un nœud u. Les γ_u sont alors des variables aléatoires définies sur $\bar{\Omega}_u$ à valeurs dans Γ. L'espace des arbres marqués est muni de la tribu $\bar{\mathcal{F}}$ engendrée par ces applications. On note sans indice les variables aléatoires attachées à l'ancêtre.

On a l'analogue de la définition-proposition 1 :

Proposition-Définition 2 *Soit Q une loi de probabilité sur $\mathbb{N}\prod\Gamma$, où Γ est un espace mesurable. Alors il existe une unique probabilité $\mathbb{P}_Q$ (notée $\mathbb{P}$ quand il n'y a pas d'ambiguité) sur $(\bar{\Omega}, \bar{\mathcal{F}})$ telle que*

(i) (N, γ) a pour loi Q.

(ii) (propriété de branchement) conditionnellement sachant $\{N = j\}$ et γ, les variables $T_1, T_2, ..., T_j$ sont indépendantes et de même loi $\mathbb{P}$.

Par conséquent, pour tout $n \geq 0$, conditionnellement par rapport à $\bar{\mathcal{F}}_n$ (tribu engendrée par les variables $N_u, |u| < n$ et $\gamma_u, |u| \leq n$), les variables aléatoires $\{T_u, u \in z_n\}$ sont i.i.d. de loi $\mathbb{P}$. Autrement dit encore, pour toutes fonctions mesurables positives $f_u, u \in U$

$$\mathbb{E}^{\bar{\mathcal{F}}_n}(\prod_{u\in z_n} f_u \circ T_u) = \prod_{u\in z_n} \mathbb{E}(f_u) \ . \tag{8}$$

Les marques peuvent être des durées de vie à valeurs dans $\mathbb{R}$; dans ce cas en appelant Z_t le nombre d'individus en vie à l'instant t, on obtient le processus des âges de Bellmann-Harris (voir chapitre IV de [A-N]).

Les marques peuvent être des trajectoires ; prenons pour Γ l'espace des fonctions continues sur un intervalle de la forme $[0, \sigma]$ à valeurs dans $\mathbb{R}$ ou ($\mathbb{R}^{\mathrm{d}}$) et prenons comme loi de γ celle d'un mouvement brownien standard arrêté à un temps exponentiel. Alors la mesure des emplacements des particules à l'instant t est un processus de branchement brownien. Nous allons prendre dans la suite l'analogue discret du branchement brownien : la marche aléatoire avec branchement (branching random walk).

4.2 Marche aléatoire avec branchement

Le phénomène se décrit de la façon suivante : un ancêtre se trouve en 0 (ou en x) à l'instant 0. Ses enfants forment la première génération et leurs positions sont données par un processus ponctuel Z sur $\mathbb{R}$. Ensuite, chaque particule u donne naissance, indépendamment des autres particules et du passé, à des enfants selon un processus ponctuel copie de Z. Ce modèle est discret mais un modèle continu se décrit tout aussi bien avec des durées de vie aléatoires.

L'espace considéré $\bar{\Omega}$ est celui des arbres marqués par les déplacements réels γ_u des particules ($\Gamma = \mathbb{R}$). On définit la position X_u de la particule u par

$$X_u = X_\emptyset + \gamma_{i_1} + \gamma_{i_1 i_2} + \cdots + \gamma_{i_1 \dots i_n}$$

pour $u = i_1 i_2 \dots i_n$. La propriété de branchement suivante s'est enrichie de la composante spatiale.

Proposition-Définition 3 *Pour toute mesure μ de comptage sur $\mathbb{R}$ (qu'on supposera dans la suite de logLaplace finie partout) et pour tout $x \in \mathbb{R}$, il existe une unique probabilité $\mathbb{P}_x$ sur $(\bar{\Omega}, \bar{\mathcal{F}})$ telle que*
(i) $X_\emptyset = x$

Z_1 a pour mesure d'intensité μ, c'est-à-dire que pour toute fonction f mesurable positive

$$\mathbb{E}(Z_1(f)) = \mathbb{E}(\sum_{j=1}^{N} f(X_j)) = \int f(x)\mu(dx)$$

(ii) pour tout $n > 0$, conditionnellement à $\bar{\mathcal{F}}_n$ (tribu engendrée par les variables $N_u, |u| < n$ et $\gamma_u, |u| \leq n$), les variables aléatoires $\{T_u, u \in z_n\}$ sont i.i.d. de loi $\mathbb{P}_{X_u}$. Autrement dit encore, pour toutes fonctions mesurables positives $f_u, u \in U$

$$\mathbb{E}_x^{\bar{\mathcal{F}}_n}(\prod_{u \in z_n} f_u \circ T_u) = \prod_{u \in z_n} \mathbb{E}_{X_u}(f_u) \tag{9}$$

Dans la suite on notera $\mathbb{E}$ pour $\mathbb{E}_0$. On supposera se trouver dans le cas surcritique où $m := \mathbb{E}(Z_1(\mathbb{R})) = \mathbb{E}(\mathrm{N}) = \mu(\mathbb{R}) > 1$.

L'analogue de la martingale Z_n/m^n du processus de Galton-Watson est pour tout θ réel

$$W_n(\theta) = \sum_{u \in z_n} e^{-\theta X_u - n\Lambda(\theta)}$$

où $\Lambda(\theta) = \log m(\theta)$ et

$$m(\theta) = \mathbb{E}(\sum_{j=1}^{N} e^{-\theta X_j}) = \int e^{-\theta x}\mu(dx) = \mathbb{E}(Z_1(e^{-\theta \cdot})).$$

Remarquons qu'en faisant $\theta = 0$, on retrouve le processus de Galton-Watson sous-jacent. Le processus des âges est également un cas particulier de ce modèle en supposant la reproduction et le mouvement indépendants et en supposant que la loi du déplacement est concentrée sur $\mathbb{R}_+$.

La propriété de martingale passe sans difficulté au cas spatial :

Théorème 9 *(Kingman, 1975)*

Pour tout $\theta \in \mathbb{R}$, $W_n(\theta)$ est une $\mathcal{F}_n$-martingale positive. Elle converge presque sûrement vers $W(\theta)$ avec a priori $\mathbb{E}(W(\theta)) \leq 1$.

<u>Démonstration</u> : on applique la propriété de branchement à

$$\begin{aligned} \mathbb{E}_x^{\mathcal{F}_n} W_{n+1}(\theta) = & \ \tfrac{1}{m(\theta)^{n+1}} \sum_{u \in z_n} \mathbb{E}_x^{\mathcal{F}_n} \sum_{j=1}^{N \circ T_u} e^{-\theta X_j \circ T_u} \\ = & \ \tfrac{1}{m(\theta)^{n+1}} \sum_{u \in z_n} \mathbb{E}_{X_u}(\sum_{j=1}^{N} e^{-\theta X_j}) \\ = & \ \sum_{u \in z_n} \frac{e^{-\theta X_u}}{m(\theta)^n} = W_n(\theta) \end{aligned}$$

□

L'analogue du théorème de Kesten-Stigum existe mais un phénomène nouveau apparaît qui sépare l'espace en deux zones.

Théorème 10 *(Biggins 1977)*

$W_n(\theta) \xrightarrow{L^1} W(\theta)$ *et* $\mathbb{E}(W(\theta)) = 1$
si et seulement si

$$\log m(\theta) - \theta \frac{m'(\theta)}{m(\theta)} > 0 \tag{10}$$

et

$$\mathbb{E}(Z_1(\theta) \log Z_1(\theta)) < \infty.$$

Dans le cas contraire, $W_n(\theta) \longrightarrow 0$ presque sûrement.

On voit que (10) définit deux zones de $\mathbb{R}$ limitées par deux valeurs critiques θ_c et θ'_c. Dans la zone $\theta_c < \theta < \theta'_c$ on a convergence L^1 sous "xlog x". Dans la zone $\theta < \theta_c$ ou $\theta > \theta'_c$ on a $W_n(\theta) \longrightarrow 0$.

Démonstration : plusieurs démonstrations existent, dont une utilisant l'extension au cas spatial de l'arbre biaisé (cf. Lyons [12]).

5 Lignes d'arrêt

Définition 6 *On dit qu'un sous-ensemble L de U a la propriété de ligne si et seulement si tout élément de L n'a pas d'ancêtre dans L, autrement dit si et seulement si pour tous éléments u et v de la ligne L on n'a ni $u < v$ ni $v < u$. S'il s'agit d'un sous-ensemble $L(\omega)$ d'un arbre ω ceci est encore équivalent à T_u et T_v disjoints pour tous éléments u et v de la ligne.*

Une ligne $L(\omega)$ définit l'ensemble $D(\omega)$ des descendants de la ligne

$$D(\omega) = \{u \in \omega, \exists v \in L(\omega), v \leq u\}$$

Si l'on considère la tribu engendrée par les N_u (et éventuellement les γ_u) pour tous les $u \notin D(\omega)$, on souhaite écrire une propriété de branchement plus générale. Pour cela on introduit la notion de ligne d'arrêt.

Pour un processus spatial ou plus généralement pour les arbres marqués, on définit $\mathcal{A}_u(s)$, la filtration attachée à chaque branche. Alors une ligne d'arrêt sera une collection de variables $\tau_u, u \in U$, définies chacune sur Ω_u et à valeurs dans $[0, \sigma_u]$ (où σ_u est la durée de vie de u) ou bien $\{0, 1\}$ s'il s'agit d'un processus discret et vérifiant les deux conditions

(i) τ_u est un temps d'arrêt pour $\mathcal{A}_u(s)$,

(ii) l'ensemble $L_\tau(\omega) := \{u \in \omega, 0 \leq \tau_u < \sigma_u\}$ a la propriété de ligne.

Par exemple pour la marche aléatoire avec branchement, pour tout réel positif a, posons $\tau_u = 0$ si $X_u \geq a$ et pour tous les $v < u$, $X_v < a$, $\tau_u = 1$ sinon. Alors $(\tau_u, u \in U)$ est une ligne d'arrêt et $L_a := L_\tau = \{u, X_u \geq a, \forall v < u,\ X_v < a\}$ est la ligne des premières particules à atteindre la droite d'équation $x = a$. On peut bien entendu imaginer toutes sortes de lignes d'arrêt pertinentes suivant le modèle considéré. On vérifie facilement que la n-ième génération est une ligne d'arrêt ainsi que l'ensemble des particules en vie à l'instant t dans un processus continu.

Proposition 9 *Soit* $\tau = \{\tau_u, u \in U\}$ *une ligne d'arrêt et soit* $\mathcal{F}_\tau$ *la tribu associée. Alors conditionnellement à* $\mathcal{F}_\tau$ *les sous-arbres* $T_u, u \in L_\tau$ *sont indépendants et de même loi* $\mathbb{P}_{X_u(\tau_u)}$.

Cette propriété de branchement (analogue de la propriété de Markov forte sur les arbres) et les lignes d'arrêt sont utilisées notamment dans [6] pour un processus de branchement brownien, dans [4] pour une marche aléatoire avec branchement ou dans [5] en théorie du contrôle.

References

[1] S. Asmussen, H. Hering. Branching processes. Birkhauser. 1983

[2] K.B. Athreya, P. Ney. Branching processes. Springer. 1972

[3] J.D.Biggins. Martingale convergence in the branching random walk. Adv in Appl. Probab. 10, 62-84, 1978

[4] J.D.Biggins, A.E. Kyprianou. Seneta-Heyde norming in the branching random walk. Ann. Probab. **25**, no. 1, 337–360, 1997.

[5] C. Ceci, A. Gerardi. Filtering of a branching process given its split times. J. Appl. Prob. **55**, 27-50, 1997.

[6] B. Chauvin. Product martingales ans stopping lines for branching brownian motion. Ann. Probab. **19**, 3, 1195-1205. 1991.

[7] L. Comtet. Advanced Combinatorics. Reidel. Dordrecht. 1974

[8] T.H. Cormen, C.E. Leiserson, R.L. Rivest. *Introduction à l'algorithmique*. Dunod. 1990.

[9] P. Flajolet, R. Sedgewick. An introduction to analysis of algorithms. Addison-Wesley.

[10] D. Knuth. The art of computer programming. Vol. 1. Fundamental Algorithms. Addison-Wesley. Second edition 1973

[11] D. Knuth. The art of computer programming. Vol. 3. Sorting and searching. Addison-Wesley. 1973

[12] R. Lyons. A simple path to Biggins' martingale convergence for branching random walk. Classical and Modern Branching Processes. Ed. Athreya and Jagers. IMA Volumes in Mathematics and its Applications. Vol. 84. Springer. 1997.

[13] R. Lyons, R. Pemantle, Y. Peres. Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Probab.* **23**, 1125-1138, 1995.

[14] H.M. Mahmoud. Evolution of Random Search Trees. Wiley. 1992

[15] J. Neveu. Arbres et processus de Galton-Watson. *Ann. IHP* **22**, 2, 199-207, 1986

[16] Y.Peres. Probability on trees : an introductory climb. Cours Ecole d'été de Saint-Flour 1997. Lecture Notes in Mathematics 1717. Springer 1999.