

Dimension de Hausdorff d'arbres de Galton-Watson

Exposé de maîtrise proposé par Mathilde WEILL

Pierre-Loïc MÉLIOT

Rémy OUDOMPHENG

27 août 2005

Table des matières

Table des matières	1
Introduction	2
1 Arbres aléatoires – Processus de branchement	3
1.1 Codage canonique d'un arbre	3
1.2 Processus et arbres de Galton-Watson	4
1.3 Comportement asymptotique des processus	8
2 Dimension de Hausdorff – Fractales	11
2.1 Dimension de Hausdorff d'un ensemble	11
2.1.1 Définition	11
2.1.2 Exemples	11
2.2 Arbres et fractales	13
2.2.1 Le bord d'un arbre et sa dimension	13
2.2.2 Correspondance entre arbres et ensembles fractals	15
3 Percolation, flots, et nombre de branchement	18
3.1 Notion de nombre de branchement	18
3.1.1 Définitions et premières propriétés	18
3.1.2 Lien avec la dimension de Hausdorff	19
3.2 Percolation et flots	20
3.2.1 Percolation sur un arbre	20
3.2.2 Percolation et nombre de branchement	21
3.2.3 Flots sur un arbre de Galton-Watson	23
4 Applications concrètes	26
4.1 Exemples de fractales aléatoires	26
4.2 Application au mouvement brownien	27
Conclusion	30
A Annexe : théorèmes asymptotiques pour les processus de Galton-Watson	31
A.1 Théorème de Kesten-Stigum	31
A.2 Théorème de Heathcote-Seneta-Vere-Jones	36
Références	38

Introduction

Le but de cet exposé est de tisser un lien entre les objets fractals et les arbres. La structure d'un arbre est très similaire à celle d'une fractale, en ce sens que les sous-arbres d'un arbre ont généralement une structure voisine de celle de l'arbre entier. Il est donc naturel de s'inspirer des concepts utilisés dans l'étude des uns pour avoir un point de vue nouveau sur les autres. On peut ainsi faire émerger certaines propriétés d'un objet fractal en le représentant par un arbre et réciproquement, construire un objet fractal, éventuellement aléatoire, à partir d'un arbre.



FIG. 1 – Cube de Cantor aléatoire de dimension de Hausdorff $\left(4 - \frac{\log 4}{\log 3}\right)$.

1 Arbres aléatoires – Processus de branchement

Une idée naturelle pour construire des arbres aléatoires est de s’inspirer de l’évolution d’une population, où chaque individu a un nombre L de descendants. On peut alors considérer L comme une variable aléatoire. Commençons par rappeler quelques généralités sur les arbres.

1.1 Codage canonique d’un arbre

Définition 1.1 (Graphes, arbres). On appelle *graphe*¹ tout couple $G = (V, E)$, où :

- V est un ensemble a priori quelconque de *sommets* — dans la plupart des cas dénombrable.
- E constitue une partie de l’ensemble $\mathfrak{P}_2(E)$ des parties à deux éléments de E : les *arêtes* de l’arbre.

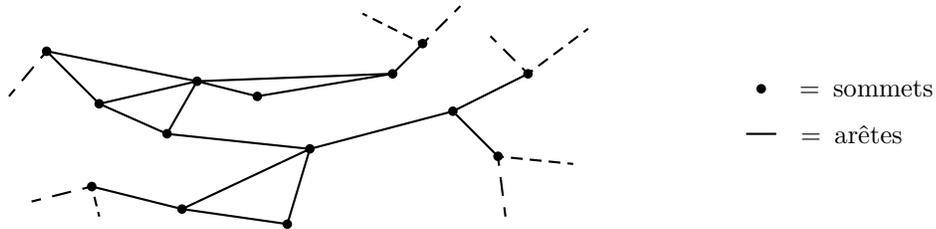


FIG. 2 – Un graphe (V, E) .

Le graphe est dit *localement fini* si pour tout sommet $v \in V$, l’ensemble de ses voisins $N_v = \{x \in V \mid \{x, v\} \in E\}$ est fini. Nous ferons toujours cette hypothèse par la suite. On appelle *chemin* sur le graphe G toute suite de points (finie ou infinie) $(x_i)_{i \in I}$ de V telle que $\forall i \in I, \{x_i, x_{i+1}\} \in E$. Un *cycle* sur le graphe est un chemin fini $(x_i)_{i \in \llbracket 1, m \rrbracket}$ tel que $x_1 = x_m$. On parle de cycle non trivial si de plus :

- l’entier m est supérieur à 3,
- les sommets x_1, x_2, \dots, x_{m-1} sont deux à deux distincts.

Un *arbre* est un graphe T connexe localement fini² sans cycle non trivial. Choissant arbitrairement un sommet du graphe T , on peut toujours représenter T comme sur la figure 3, en plaçant le sommet en question à la racine de l’arbre.

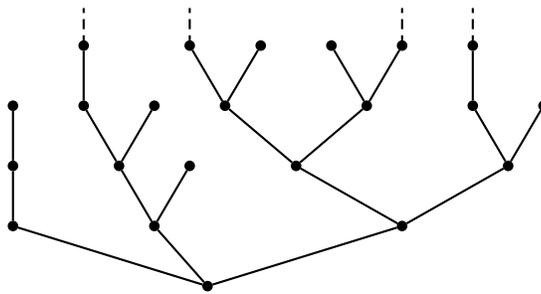


FIG. 3 – Un arbre $T = (V, E)$ est un graphe localement fini sans cycle non trivial.

Étant donné un arbre T muni d’une racine r , la *hauteur* $h(s)$ d’un sommet s de T est la distance le séparant de la racine; et l’on dit qu’un sommet t est le père d’un sommet s — ou encore que s est le fils de t — si $h(s) = h(t) + 1$ et si s et t sont reliés par une arête.

¹On ne s’intéressera ici qu’aux graphes non orientés, de sorte que les éléments de E seront des doublets $\{a, b\}$, et non des couples (a, b) .

²Tout graphe connexe localement fini est dénombrable.

Il peut paraître relativement difficile de définir un processus aléatoire à valeurs dans l'ensemble des arbres, dans la mesure où le moyen le plus simple de se donner un arbre semble être ... de le dessiner. Néanmoins, il existe une façon canonique et relativement intuitive de « coder » un arbre et sa représentation. En effet, chaque sommet d'un arbre est entièrement caractérisé par son père (s'il n'est pas la racine) et par sa position parmi les fils de celui-ci. On peut donc associer à tout sommet s de l'arbre un élément de $\mathcal{F} = \bigsqcup_{n \in \mathbb{N}} (\mathbb{N}^*)^n$ — avec pour convention $(\mathbb{N}^*)^0 = \{\emptyset\}$:

$$\psi : V(T) \rightarrow \mathcal{F}$$

$$s \mapsto \begin{cases} \emptyset & \text{si } s \text{ est la racine de l'arbre} \\ (\psi(t), p) & \text{si } s \text{ est le } p\text{-ième fils de } t \end{cases}$$

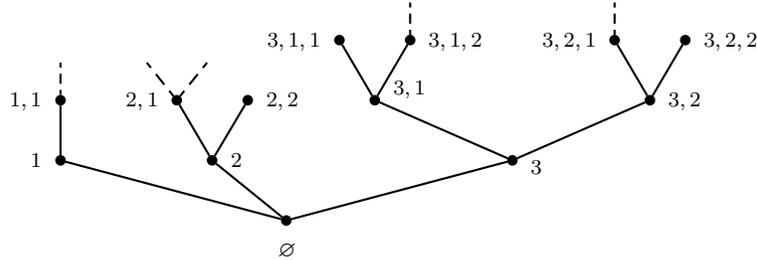


FIG. 4 – Étiquettes des sommets d'un arbre T .

Proposition 1.1 (Codage canonique d'un arbre). L'application ψ induit une correspondance entre les arbres dénombrables localement finis et les parties de \mathcal{F} , et cette correspondance est injective³, de sorte que tout arbre peut être codé par une partie de $\bigsqcup_{n \in \mathbb{N}} (\mathbb{N}^*)^n$.

Démonstration. Il suffit de poser $\psi(T) = \{\psi(s), s \in V(T)\}$. □

C'est précisément cette identification qui va nous permettre de construire rigoureusement des arbres « aléatoires » : en effet, il suffira de construire un processus aléatoire sur $\mathfrak{P}(\mathcal{F})$ dont les valeurs sont p.s. dans $\text{im } \psi$, et d'identifier ces valeurs aux arbres qu'elles codent.

1.2 Processus et arbres de Galton-Watson

Commençons par définir le processus des générations $(Z_n)_{n \in \mathbb{N}}$ d'un arbre aléatoire de Galton-Watson.

Définition 1.2 (Processus de Galton-Watson). Soit p une mesure de probabilité sur \mathbb{N} . On appelle *processus de branchement de Galton-Watson* de loi de reproduction p une chaîne de Markov $(Z_n)_{n \in \mathbb{N}}$ dont l'espace des états est $E = \mathbb{N}$, dont la valeur initiale est déterminée égale à 1, et dont la matrice de transition est :

$$Q(x, y) = p^{*x}[y].$$

En d'autres termes, si $(L_{n,k})_{(n,k) \in \mathbb{N}^2}$ est une famille de variables aléatoires entières indépendantes de même loi p , alors un processus de Galton-Watson s'identifie à une suite $(Z_n)_{n \in \mathbb{N}}$ de v.a. vérifiant la relation de récurrence :

$$\begin{cases} Z_0 = 1 \\ Z_{n+1} = \sum_{k=1}^{Z_n} L_{n,k} \end{cases}$$

Un tel processus peut s'interpréter intuitivement de la façon suivante. On part d'un individu — la racine de l'arbre — qui engendre L descendants, L étant une variable aléatoire de loi p .

³La correspondance n'est en revanche pas surjective, puisqu'il existe des parties de \mathcal{F} qui ne sont pas atteintes, par exemple $\{3\}$.

À la génération suivante, chaque descendant « k » donne à son tour naissance à L_k enfants, indépendamment des autres individus de première génération, et indépendamment de l'individu racine. On réitère ce processus à chaque génération, indépendamment des générations précédentes et des autres individus de la génération. La variable aléatoire Z_n correspond alors au nombre d'individus de n -ième génération.

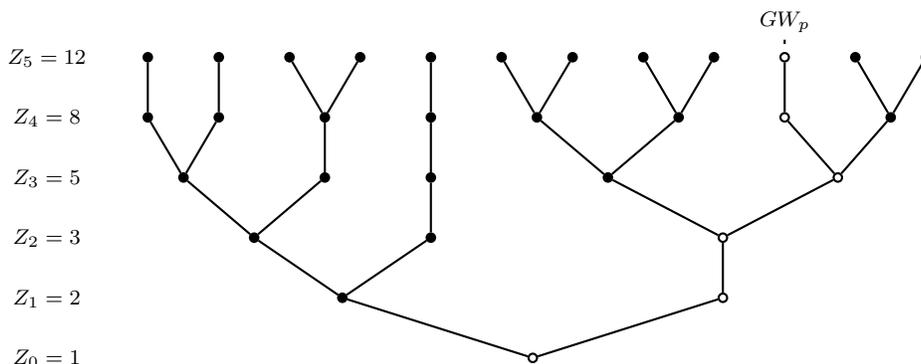


FIG. 5 – Un exemple d'arbre de Galton-Watson de loi de reproduction p , avec $p_1 = p_2 = \frac{1}{2}$.

Un processus de Galton-Watson décrit donc bien l'évolution de la taille d'une population dont chaque individu aurait L descendants. Néanmoins, nous n'avons pas encore construit d'arbre correspondant à ce processus : le processus des générations $(Z_n)_{n \in \mathbb{N}}$ ne donne en effet que la taille des générations, et pas la structure de l'arbre. Le codage canonique des arbres dénombrables permet de pallier⁴ cette lacune.

Définition 1.3 (Arbre de Galton-Watson). Donnons-nous une famille de copies indépendantes de L indexée par l'ensemble dénombrable \mathcal{F} , et posons :

$$T = \{(p_1, p_2, \dots, p_n) \in \mathcal{F} \mid \forall j \in \llbracket 1, n \rrbracket, p_j \leq L_{(p_1, p_2, \dots, p_{j-1})}\}.$$

La partie T de \mathcal{F} est un arbre aléatoire — i.e. est dans $\text{im } \psi$ — et de plus, pour tout $n \in \mathbb{N}$, les cardinaux Z_n des intersections $T \cap (\mathbb{N}^*)^n$ forment un processus de Galton-Watson. On dit que T est un *arbre aléatoire de Galton-Watson*.

Dans toute la suite, on supposera que p admet un moment⁵ d'ordre 1, et l'on notera m l'espérance du nombre de descendants par individus.

$$m = \mathbb{E}[L] = \sum_{k=0}^{\infty} p_k k \quad (1)$$

D'autre part, on notera GW_p la mesure de probabilité sur $\mathfrak{P}(\mathcal{F})$ associée à la loi de reproduction p , i.e. la loi de T — qui ne charge que $\text{im } \psi$.

Nous verrons par la suite que les arbres de Galton-Watson permettent de coder des *fractales aléatoires*, et que certaines propriétés géométriques de ces objets sont fortement reliées au comportement asymptotique de processus de Galton-Watson $(Z_n)_{n \in \mathbb{N}}$. C'est pourquoi l'on se propose d'étudier en détail l'évolution des processus $(Z_n)_{n \in \mathbb{N}}$.

Un processus aléatoire de Galton-Watson est susceptible de donner deux types de résultats :

- *extinction* : la famille aléatoire s'éteint, i.e. il existe un rang N tel que $Z_N = 0$, auquel cas :

$$\forall n \geq N, Z_n = 0.$$

⁴Il convient de noter que la construction grammaticale « pallier à » est incorrecte - bien que des auteurs célèbres, Albert Camus par exemple, y aient recours.

⁵Il s'agit là d'une hypothèse peu restrictive : que dire d'un individu qui a en moyenne une infinité de descendants ?

– *survie* : la famille aléatoire se reproduit indéfiniment, et :

$$\forall n \in \mathbb{N}, Z_n \neq 0.$$

La proposition suivante précise cette alternative entre extinction et survie :

Proposition 1.2 (Comportement limite de Z_n). Soit $(Z_n)_{n \in \mathbb{N}}$ un processus de Galton-Watson.

1. Si $L = 1$ p.s., alors toutes les générations comptent un et un seul individu, et, en particulier, le processus ne s'éteint jamais : $\forall n \in \mathbb{N}, Z_n = 1$.

Ce cas ne présente que peu d'intérêt, et nous l'excluerons systématiquement par la suite en supposant $p_1 < 1$.

2. Sinon, la famille aléatoire s'éteint avec une certaine probabilité P_e , et :

$$\begin{aligned} Z_n &\rightarrow 0 \quad \text{p.s. sur l'événement d'extinction,} \\ Z_n &\rightarrow \infty \quad \text{p.s. sur l'événement de non-extinction.} \end{aligned}$$

Démonstration. Seule la dernière assertion, à savoir que $Z_n \rightarrow \infty$ p.s. sur l'événement de non-extinction, mérite une démonstration. Distinguons deux cas :

- si $p_0 = 0$, alors la suite $(Z_n)_{n \in \mathbb{N}}$ est p.s. croissante, supérieure à 1 pour tout n . De plus, l'événement $Z_{n+1} = Z_n$ ne se produit que si $L_{(n,k)} = 1$ pour tous les individus k de la n -ième génération, donc :

$$\mathbb{P}[Z_{n+1} = Z_n] = \sum_{k=1}^{\infty} (p_1)^k \mathbb{P}[Z_n = k] \leq p_1 < 1.$$

Ainsi, $\mathbb{P}[Z_{n+1} > Z_n] \geq 1 - p_1 > 0$, et le lemme de Borel-Cantelli⁶ permet de conclure que cet événement se produit presque sûrement une infinité de fois. La suite (Z_n) tend donc vers $+\infty$ presque sûrement.

- si $p_0 > 0$, alors le seul état récurrent de la chaîne de Markov $(Z_n)_{n \in \mathbb{N}}$ est 0, puisque c'est un état absorbant — $Q(0,0) = 1$ — tel que :

$$\forall n \in \mathbb{N}, U(n,0) = \sum_{k=0}^{\infty} Q_k(n,0) \geq Q_1(n,0) = p_0^n > 0.$$

Par conséquent, sur l'événement de non-extinction, la chaîne ne passe qu'un nombre fini de fois par chaque état, donc tend presque sûrement vers $+\infty$.

Le résultat est donc bien établi dans les deux cas. □

Deux types de questions se posent alors :

- avec quelle probabilité P_e un processus de Galton-Watson de loi p s'éteint-il ?
- quelle est la taille moyenne de la n -ième génération d'un processus de Galton-Watson ? Quelle est sa taille sachant l'extinction, respectivement la survie du processus ?

Un outil essentiel d'étude d'un processus de Galton-Watson est la *fonction génératrice* de sa loi p :

$$f(x) = \mathbb{E}[x^{Z_1}] = \sum_{k=0}^{\infty} p_k x^k.$$

Il s'agit d'une application de classe C^∞ , croissante et convexe, définie sur l'intervalle $[0, 1]$.

Théorème 1.3 (Probabilité d'extinction et fonction génératrice). *La probabilité d'extinction P_e est le plus petit point fixe de f appartenant à $[0, 1]$. Les assertions suivantes sont donc équivalentes :*

$$P_e < 1 \iff f'(1) > 1 \iff m > 1. \quad (2)$$

Démonstration. Commençons par montrer par récurrence sur $n \in \mathbb{N}$ que :

$$\forall x \in [0, 1], \mathbb{E}[x^{Z_n}] = \underbrace{f \circ f \circ \dots \circ f}_{n \text{ fois}}(x) = f^{\circ n}(x). \quad (3)$$

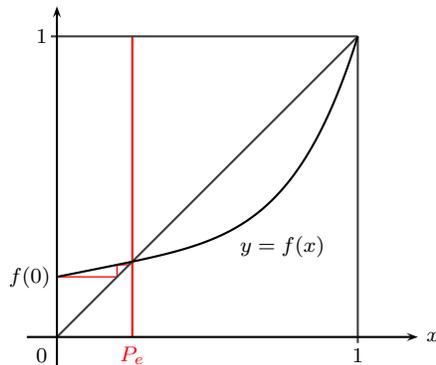


FIG. 6 – La probabilité d’extinction est le plus petit point fixe de la fonction génératrice de p .

Le résultat est clair pour $n = 0$. Supposons-le maintenant établi jusqu’au rang $n - 1 \geq 0$. L’indépendance des variables $L_{n,k}$ et Z_{n-1} nous permet d’écrire :

$$\begin{aligned} \mathbb{E} [x^{Z_n}] &= \mathbb{E} [\mathbb{E} [x^{Z_n} | Z_{n-1}]] = \mathbb{E} \left[\mathbb{E} \left[\prod_{k=1}^{Z_{n-1}} x^{L_{n,k}} \middle| Z_{n-1} \right] \right] \\ &= \mathbb{E} \left[\prod_{k=1}^{Z_{n-1}} \mathbb{E} [x^{L_k}] \right] = \mathbb{E} [f(x)^{Z_{n-1}}] = f^{\circ(n-1)}(f(x)) = f^{\circ n}(x). \end{aligned}$$

D’où le résultat au rang n , et par récurrence pour tout $n \in \mathbb{N}$.

En particulier, $\mathbb{P} [Z_n = 0] = f^{\circ n}(0)$. Comme l’extinction est la réunion croissante des événements $Z_n = 0$, $P_e = \lim_{n \rightarrow \infty} \uparrow \mathbb{P} [Z_n = 0]$, et l’on en déduit :

$$P_e = \lim_{n \rightarrow \infty} \uparrow f^{\circ n}(0).$$

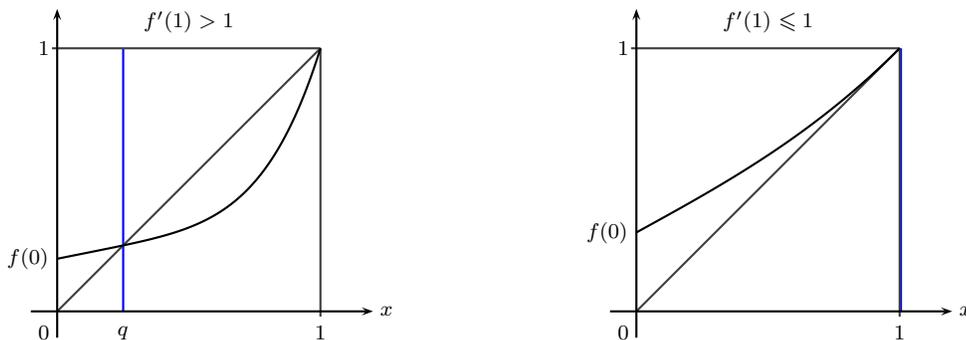


FIG. 7 – Deux types de fonctions génératrices.

La fonction f étant continue croissante, la seule limite possible de $(f^{\circ n}(0))_{n \in \mathbb{N}}$ est le plus petit point fixe de f appartenant à $[0, 1]$. Ainsi, P_e est un point fixe de f .

Or la fonction f est convexe, de sorte que deux cas peuvent se produire :

- si $f'(1) > 1$, alors f a deux points fixes dans $[0, 1]$: $q < 1$ et 1. Le processus de Galton-Watson s’éteint alors avec la probabilité $P_e = q < 1$.
- si $f'(1) \leq 1$, l’unique point fixe de f appartenant à $[0, 1]$ est 1, et le processus s’éteint donc presque sûrement.

Comme $f'(1) = \sum_{k=1}^{\infty} p_k k = m$ d’après (1), on en déduit l’équivalence (2) annoncée. \square

⁶Les variables aléatoires $L_{n,k}$ étant indépendantes, il en va de même des événements $A_n = \{Z_{n+1} > Z_n\}$.

Ce résultat motive les définitions suivantes :

Définition 1.4. Un processus de Galton-Watson de loi p est dit :

- *sous-critique* si $m < 1$. La probabilité d'extinction P_e est égale à 1.
- *critique* si $m = 1$. La probabilité d'extinction P_e est encore égale à 1.
- *sur-critique* si $m > 1$. La probabilité d'extinction P_e est alors strictement inférieure à 1.

Remarque. Il convient de noter que l'espérance m ne détermine pas entièrement P_e : il existe des fonctions génératrices de même dérivée en 1, mais n'ayant pas le même point fixe.

On se propose dans les sections suivantes de déterminer le comportement asymptotique de $(Z_n)_{n \in \mathbb{N}}$ dans chacun de ces cas.

1.3 Comportement asymptotique des processus

Si m est l'espérance du nombre d'enfants par individu, on peut s'attendre à ce que les tailles des générations Z_n soient de l'ordre de m^n . Cette intuition est bien confirmée :

Proposition 1.4. Notons (\mathcal{A}_n) la filtration canonique associée à la chaîne de Markov $(Z_n)_{n \in \mathbb{N}}$, et pour tout $n \in \mathbb{N}$:

$$W_n = \frac{Z_n}{m^n}. \quad (4)$$

Le processus $(W_n)_{n \in \mathbb{N}}$ est une martingale positive pour la filtration $(\mathcal{A}_n)_{n \in \mathbb{N}}$, et converge donc presque sûrement vers une variable aléatoire W positive et finie presque sûrement.

Démonstration. La variable W_n est bien mesurable par rapport à \mathcal{A}_n , et, de plus :

$$\begin{aligned} \mathbb{E}[W_n | \mathcal{A}_{n-1}] &= \mathbb{E}\left[\frac{Z_n}{m^n} \middle| \mathcal{A}_{n-1}\right] = \mathbb{E}\left[\frac{1}{m^n} \sum_{k=1}^{Z_{n-1}} L_{(n,k)} \middle| \mathcal{A}_{n-1}\right] = \mathbb{E}\left[\frac{1}{m^n} \sum_{k=1}^{Z_{n-1}} \mathbb{E}[L] \middle| \mathcal{A}_{n-1}\right] \\ &= \frac{1}{m^{n-1}} \mathbb{E}[Z_{n-1} | \mathcal{A}_{n-1}] = \frac{Z_{n-1}}{m^{n-1}} = W_{n-1}. \end{aligned}$$

Les théorèmes relatifs à la convergence des sur-martingales positives permettent de conclure. \square

Si $W > 0$, le processus de Galton-Watson se comporte donc comme m^n — à un facteur multiplicatif aléatoire près. On ne peut en revanche rien dire — pour l'instant — sur la vitesse de croissance (resp., d'extinction) de $(Z_n)_{n \in \mathbb{N}}$ lorsque $W = 0$. Trois nouvelles questions se posent donc :

- Que dire de la loi de W ? Quelle est son espérance?
- Avec quelle probabilité P_{se} a-t-on une croissance sous-exponentielle — $W = 0$?
- Si $W = 0$, que dire du comportement asymptotique de Z_n sur l'événement de survie? Sur l'événement d'extinction?

La probabilité de croissance sous-exponentielle P_{se} est évidemment supérieure à la probabilité d'extinction P_e : si le processus s'éteint, on a bien $W_n \rightarrow 0$. Nous allons en fait voir que P_{se} ne peut prendre pour valeurs que P_e et 1. Cette propriété surprenante tient au fait que les sous-arbres d'un arbre aléatoire « héritent » toujours de l'événement $W = 0$.

Définition 1.5. Une propriété \mathcal{P} sur les arbres est dite *héréditaire* si tout arbre fini l'a, et si pour tout sous arbre T_2 d'un arbre T_1 :

$$\mathcal{P}(T_1) \Rightarrow \mathcal{P}(T_2).$$

Lemme 1.5. Une propriété héréditaire a pour probabilité 0 ou 1 sur l'événement de survie, et pour probabilité globale P_e ou 1.

Démonstration. Par définition, une propriété \mathcal{P} héréditaire est vraie pour tous les arbres finis, et donc vraie sur l'événement d'extinction. Ainsi, $\mathbb{P}[\mathcal{P}] \geq P_e$. Conditionnellement à Z_1 , l'arbre vérifie \mathcal{P} avec une probabilité $\mathbb{P}[\mathcal{P} | Z_1]$, et dans ce cas chacun des L sous-arbres vérifie \mathcal{P} . Or ces événements sont indépendants, et ont chacun pour probabilité $\mathbb{P}[\mathcal{P}]$. Ainsi :

$$\mathbb{P}[\mathcal{P}] = \mathbb{E}[\mathbb{P}[\mathcal{P} | Z_1]] \leq \mathbb{E}[\mathbb{P}[\mathcal{P}]^L] = f(\mathbb{P}[\mathcal{P}]).$$

Or, entre P_e et 1, la fonction $f(x)$ est en dessous de x — cf figure 7. On a donc égalité, et $\mathbb{P}[\mathcal{P}]$ ne peut valoir que P_e ou 1. Comme $\mathbb{P}[\mathcal{P} \text{ extinction}] = 1$, on en déduit :

$$\mathbb{P}[\mathcal{P} \text{ survie}] = 0 \text{ ou } 1.$$

□

Corollaire 1.6. La probabilité P_{se} vaut P_e ou 1.

Démonstration. La propriété $W = 0$ est héréditaire : si un arbre se développe sous-exponentiellement, il en va de même pour toutes ses branches. □

Lorsque $m \leq 1$, la probabilité P_e est égale à 1, donc, a fortiori, $P_{se} = 1$. Dans le cas où $m \geq 1$, on dispose d'un critère analytique permettant de trancher entre les deux alternatives. Le cas sur-critique est ainsi essentiellement résolu par le théorème suivant :

Théorème de Kesten-Stigum. *Supposons l'espérance m strictement supérieure à 1 — cas sur-critique. Les assertions suivantes sont équivalentes :*

1. La limite W de la martingale $(W_n)_{n \in \mathbb{N}}$ est strictement positive p.s. sur l'événement de non-extinction : $P_{se} = P_e$.
2. L'espérance de W est égale à 1.
3. La variable aléatoire $L \log^+ L$ est intégrable, i.e. $\mathbb{E}[L \log^+ L] < \infty$.

La démonstration du théorème de Kesten-Stigum repose essentiellement sur la notion de *processus de branchement avec immigration*. Compte tenu de la longueur de la preuve, nous renvoyons le lecteur à l'annexe pour de plus amples développements.

Que dire de la croissance de $(Z_n)_{n \in \mathbb{N}}$, lorsque $m > 1$ et $P_{se} = 1$, sur l'événement de non-extinction ? Il s'agit d'une situation qui peut paraître contre-intuitive : comme indiqué précédemment, on s'attend à ce que les générations (Z_n) se développent comme m^n . Le théorème suivant montre que, en un certain sens, l'on est toujours très proche d'une telle évolution.

Théorème de Seneta-Heyde. *Dans le cas sur-critique, la loi p étant fixée, on peut toujours trouver une suite $(c_n)_{n \in \mathbb{N}}$ telle que :*

1. Il existe presque sûrement une limite positive finie V à (Z_n/c_n) .
2. La probabilité $\mathbb{P}[V = 0]$ est égale à P_e , et donc $V > 0$ p.s. sur l'événement de non-extinction.
3. Le rapport c_{n+1}/c_n tend vers m .

Ainsi, si l'on n'a pas forcément $Z_n \simeq m^n$ sur l'événement de non-extinction, on a toujours $Z_{n+1}/Z_n \simeq m$, i.e. la $(n+1)$ -ième génération a approximativement pour taille m fois celle de la n -ième.

Démonstration. Comme $m > 1$, $1 > P_e$ et l'on peut donc choisir $x_0 \in]P_e, 1[$. Il convient alors de noter que f est une bijection continue entre $[P_e, 1]$ et lui-même. Définissons alors la suite $(x_n)_{n \in \mathbb{N}}$ par la relation de récurrence :

$$x_{n+1} = f^{-1}(x_n).$$

Cette suite converge vers le point fixe 1 de f . D'autre part, l'équation (3) se réécrit dans ce contexte :

$$\mathbb{E}[(x_n)^{Z_n} | Z_0, Z_1, \dots, Z_{n-1}] = f(x_n)^{Z_{n-1}} = (x_{n-1})^{Z_{n-1}}.$$

Ainsi, $((x_n)^{Z_n})_{n \in \mathbb{N}}$ est une martingale positive bornée par 1, et elle converge donc p.s. et L^1 vers une certaine variable aléatoire U . Posons alors $c_n = -\frac{1}{\log x_n}$. Pour tout n , $(x_n)^{Z_n} = \exp(-\frac{Z_n}{c_n})$, de sorte que par passage au logarithme, Z_n/c_n converge presque sûrement vers une limite positive V . La règle de l'Hôpital permet en outre d'écrire :

$$\lim_{x \rightarrow 1} \frac{-\log f(x)}{-\log x} = \lim_{x \rightarrow 1} \frac{\frac{f'(x)}{f(x)}}{\frac{1}{x}} = \lim_{x \rightarrow 1} \frac{x f'(x)}{f(x)} = m.$$

Comme $(x_n)_{n \in \mathbb{N}} \rightarrow 1$, on en déduit la convergence et :

$$\lim \frac{c_{n+1}}{c_n} = m.$$

Par conséquent, les événements $\lim Z_n/c_n = 0$ et $\lim Z_n/c_n < \infty$ sont hérités, et ont donc pour probabilité P_e ou 1. Or, $\mathbb{E}[U] = \mathbb{E}[x_0^{Z_0}] = x_0 < 1$, donc nécessairement $\mathbb{P}[V = 0] = P_e$ — sans quoi l'on aurait $\mathbb{E}[U] = 1$. De même, $\mathbb{E}[U] > P_e$, donc $V < \infty$ presque sûrement. Les trois assertions sont donc bien démontrées. \square

Lorsque $m \leq 1$, la probabilité d'extinction P_e est égale à 1. Les deux quantités pertinentes à évaluer sont alors :

- d'une part, la vitesse d'extinction, qui est directement reliée à la proportion $K_n = \mathbb{P}[Z_n > 0]$ d'arbres non-éteints avant la n -ième génération — en quelque sorte, le « taux d'erreur » du processus au temps n .
- d'autre part, la taille des arbres non-éteints à la n -ième génération : $J_n = \mathbb{E}[Z_n | Z_n > 0]$.

Intuitivement, comme $\mathbb{E}[Z_n | Z_n > 0] = m^n$, on peut s'attendre, au moins dans le cas sous-critique, à ce que les variables K_n soient de l'ordre de m^n . On dispose d'ailleurs au moins d'une majoration : l'inégalité de Markov permet en effet d'affirmer que $K_n \leq \mathbb{E}[Z_n] = m^n$. Dans le cas sous-critique, un critère analytique identique à celui du théorème de Kesten-Stigum permet de déterminer dans quelles conditions $K_n \simeq m^n$.

Théorème de Heathcote-Seneta-Vere-Jones. *Supposons l'espérance m strictement inférieure à 1 — cas sous-critique. La suite $(S_n = K_n/m^n)_{n \in \mathbb{N}}$ est décroissante, et les assertions suivantes sont équivalentes :*

1. La limite de $(S_n = K_n/m^n)_{n \in \mathbb{N}}$ est strictement positive.
2. La suite $(J_n)_{n \in \mathbb{N}}$ est bornée.
3. La variable aléatoire $L \log^+ L$ est intégrable, i.e. $\mathbb{E}[L \log^+ L] < \infty$.

La démonstration de ce théorème utilise de nouveau les résultats relatifs aux arbres aléatoires avec immigration⁷, et figure comme précédemment dans l'annexe.

Il ne reste alors que le cas critique des processus de branchements à traiter. Les quantités pertinentes à évaluer sont toujours K_n et J_n , mais les résultats sont sensiblement différents des précédents, dans la mesure où la condition analytique $\mathbb{E}[L \log^+ L] < \infty$ ne suffit plus pour pouvoir dire quelque chose d'intéressant sur le comportement asymptotique de Z_n .

Théorème de Kesten-Ney-Spitzer. *Supposons l'espérance m du nombre de descendants égale à 1 — cas critique — et $\sigma^2 = \text{Var } L < \infty$.*

1. La probabilité K_n est équivalente à $\frac{2}{n \sigma^2}$ lorsque $n \rightarrow \infty$.
2. La loi de Z_n/n conditionnée par l'événement $Z_n > 0$ converge vers la loi exponentielle de paramètre $\sigma^2/2$.

Nous renvoyons le lecteur à [1] pour une preuve complète de ce résultat.

Après ce bref aperçu des propriétés des arbres aléatoires et des processus de Galton-Watson, on va maintenant étudier comment, à partir d'un arbre, on peut construire un objet fractal dont les propriétés seront fortement liées à celle de l'arbre.

⁷Ceci constitue une nouvelle invitation à consulter l'annexe...

2 Dimension de Hausdorff – Fractales

2.1 Dimension de Hausdorff d'un ensemble

2.1.1 Définition

On aimerait définir, pour un objet quelconque, une mesure de sa taille. Mais on peut construire, dans l'espace euclidien usuel, des objets de surface nulle et de périmètre infini (pour une bonne définition de ces notions). On voudrait donc définir une notion de dimension qui tienne compte de ces objets intermédiaires.

Définition 2.1. Soit E un espace métrique, et s un réel positif. La *mesure de Hausdorff* de dimension s d'une partie X de E est :

$$\mathcal{H}_s(X) = \lim_{\varepsilon \rightarrow 0} \inf \left\{ \sum_{i \in I} (\text{diam } X_i)^s, X \subseteq \bigcup_{i \in I} X_i \text{ avec } \text{diam } X_i < \varepsilon \right\}.$$

Proposition 2.1. La fonction \mathcal{H}_s est une mesure extérieure sur l'ensemble des parties de E , i.e. :

- $\mathcal{H}_s(\emptyset) = 0$,
- si $A \subset B$, $\mathcal{H}_s(A) \leq \mathcal{H}_s(B)$,
- $\mathcal{H}_s(\bigcup_n A_n) \leq \sum_n \mathcal{H}_s(A_n)$.

Démonstration. La preuve est immédiate compte tenu de la définition de la mesure de Hausdorff par des recouvrements. \square

Si $s \in \mathbb{N}$, on constate que la mesure extérieure de Hausdorff est équivalente la mesure extérieure de Lebesgue de \mathbb{R}^s à une constante près. En effet, si on peut remplacer chaque E_i par un cube d'arête $\text{diam } E_i$ qui le contient, et on trouve donc $\mathcal{H}_d(E) \leq d^{\frac{d}{2}} \lambda_{\mathbb{R}^d}(E)$.⁸

Proposition 2.2. Soit E un espace métrique. Il existe un unique réel d tel que $\mathcal{H}_s(E) = 0$ si $s > d$ et $\mathcal{H}_s(E) = \infty$ si $s < d$. Ce nombre est appelé *dimension de Hausdorff* de E , que l'on notera dans la suite $\dim E$.

Démonstration. La preuve repose sur l'inégalité suivante :

$$\sum_{i \in I} (\text{diam } E_i)^t < \varepsilon^{t-s} \sum_{i \in I} (\text{diam } E_i)^s \text{ si } \text{diam } E_i < \varepsilon \text{ et } t > s.$$

En effet, si $x < \varepsilon$, $x^t < \varepsilon^{t-s} x^s$. En conséquence, si $\mathcal{H}_s(E) < \infty$ pour un certain s , $\mathcal{H}_t(E)$ est nulle pour les $t > s$, et si $\mathcal{H}_s(E) > 0$, $\mathcal{H}_t(E) = \infty$ si $t < s$. Il existe donc un réel d tel que $\mathcal{H}_s(E) = 0$ si $s > d$ et $\mathcal{H}_s(E) = \infty$ si $s < d$. \square

2.1.2 Exemples

Proposition 2.3. Soit (A_n) une suite de parties d'un espace métrique E . Alors $\dim(\bigcup_n A_n) = \sup_n(\dim A_n)$.

Démonstration. Soit $s = \sup_n(\dim A_n)$.

- **1^{er} cas** : $t > s$. Alors pour tout n , $t > \dim A_n$, donc $\mathcal{H}_t(A_n) = 0$ et $\mathcal{H}_t(\bigcup_n A_n) = 0$.
- **2^e cas** : $t < s$. Par définition de la borne supérieure, il existe un n_0 tel que $\dim A_{n_0} > t$, et $\mathcal{H}_t(A_{n_0}) = \infty$. Donc $\mathcal{H}_t(\bigcup_n A_n) = \infty$.

Le nombre s vérifie la propriété caractéristique de la dimension de Hausdorff, d'où la propriété recherchée.⁹ \square

⁸On a également une inégalité dans l'autre sens.

⁹En considérant des atlas sur une variété riemannienne, on voit ainsi que sa dimension de Hausdorff est bien sa dimension topologique.

Exemple (L'ensemble triadique de Cantor). L'ensemble triadique de Cantor peut être construit par itération. Pour cela, on considère la suite :

- $E_0 = [0, 1]$.
- Si $E_n = \bigsqcup [a_i, b_i]$, on pose $E_{n+1} = \bigsqcup ([a_i, \frac{2a_i+b_i}{3}] \cup [\frac{a_i+2b_i}{3}, b_i])$.

L'ensemble triadique de Cantor est alors $C = \bigcap E_n$.



FIG. 8 – Premières itérations de la construction de l'ensemble triadique de Cantor.

Comme $\lambda(E_{n+1}) = \frac{2}{3}\lambda(E_n)$, $\lambda(C) = 0$. Donc $\dim C < 1$. Par ailleurs, on dispose des recouvrements canoniques de C par les E_n , qui possèdent 2^{-n} segments de longueur 3^{-n} . Par conséquent $\mathcal{H}_{\frac{\log 2}{\log 3}}(C) \leq 1$. Ceci prouve que $\dim C \leq \frac{\log 2}{\log 3}$. Dans le cas général, un recouvrement canonique de l'ensemble permet d'obtenir une inégalité sur la dimension, et l'autre sens est souvent donné par le principe suivant :

Lemme 2.4 (Mass distribution principle). Soit F un espace métrique. On suppose qu'il existe une mesure μ finie et non nulle sur F telle que, pour toute partie mesurable $G \subset F$ de diamètre inférieur à ϵ , l'on ait :

$$\mu(G) \leq c (\text{diam } G)^s.$$

Alors $\mathcal{H}_s(F) \geq \mu(F)/c$ et $s \leq \dim F$.

Démonstration. Soit $(U_i)_{i \in I}$ un recouvrement dénombrable de F par des ouverts.

$$0 < \mu(F) \leq \sum_{i \in I} \mu(U_i) \leq c \sum_{i \in I} (\text{diam } U_i)^s$$

Par passage à la borne inférieure : $0 < \mu(F) \leq c \mathcal{H}_s(F)$. En particulier, la dimension de Hausdorff de F est supérieure à s . \square

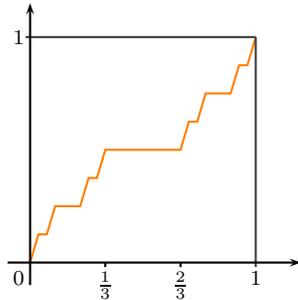


FIG. 9 – Escalier de Cantor.

Ce lemme permet de minorer la dimension de Hausdorff d'un objet fractal, la majoration étant le plus souvent l'inégalité facile à obtenir. Considérons par exemple l'ensemble triadique de Cantor. On a montré précédemment que $\dim(C) \leq \frac{\log 2}{\log 3}$. Réciproquement, soit μ la mesure sur C dont la fonction de répartition est l'escalier de Cantor. Cette mesure ne charge que C , et $\mu(C) = 1$. De plus, les intervalles de longueur 3^{-k} sont tous de masse 2^{-k} .

Soit G une partie de C non négligeable, et k l'unique entier tel que $3^{-(k+1)} \leq \text{diam}(G) < 3^{-k}$.

$$\mu(G) \leq 2^{-k} = (3^{-k})^{\frac{\log 2}{\log 3}} \leq (3 \text{ diam}(G))^{\frac{\log 2}{\log 3}}$$

Ainsi, $\mathcal{H}_{\frac{\log 2}{\log 3}}(C) \geq 3^{-\frac{\log 2}{\log 3}} = \frac{1}{2}$ et $\dim C \geq \frac{\log 2}{\log 3}$. On a donc démontré que la dimension de Hausdorff de C était précisément $\frac{\log 2}{\log 3}$.

2.2 Arbres et fractales

Un arbre infini est proche des objets fractals, puisqu'il est quasiment autosimilaire : le sous-arbre enraciné en l'un des nœuds a ainsi une structure proche de l'arbre entier. Il est donc naturel de chercher à définir la dimension de Hausdorff d'un arbre. On a deux façons de visualiser cette grandeur : soit on définit une métrique sur l'arbre, soit on remplace l'arbre par une partie de \mathbb{R}^d sur laquelle on peut travailler.

2.2.1 Le bord d'un arbre et sa dimension

Définition 2.2. Soit A un arbre. On note ∂A et on appelle *bord de A* l'ensemble des chemins reliant la racine de A à l'infini. Ou encore, l'ensemble des suites de sommets (ξ_n) , avec $\xi_0 = o$ et ξ_n de hauteur n .

Proposition 2.5. Soient α et β deux éléments de ∂A . On note $\alpha \wedge \beta$ le point commun de α et β de hauteur maximale (éventuellement infinie si $\alpha = \beta$).

Alors la fonction $d(\alpha, \beta) = \exp(-|\alpha \wedge \beta|)$ définit une distance sur ∂A .

De plus $d(\alpha, \gamma) \leq \sup\{d(\alpha, \beta), d(\beta, \gamma)\}$ (inégalité ultramétrique).

Démonstration.

- On a bien sûr $d(\alpha, \alpha) = \exp(-\infty) = 0$.
- Par ailleurs $d(\alpha, \beta) = d(\beta, \alpha)$ de manière évidente.
- Soient α, β, γ trois éléments du bord. Les chemins α et β coïncident jusqu'au sommet $\alpha \wedge \beta$, tandis que les chemins β et γ coïncident jusqu'au point $\beta \wedge \gamma$.

Donc α et γ coïncident au moins jusqu'à $\alpha \wedge \beta \wedge \gamma$. D'où :

$$\begin{aligned} d(\alpha, \gamma) &\leq \exp(-|\alpha \wedge \beta \wedge \gamma|) \\ &\leq \max(\exp(-|\alpha \wedge \gamma|), \exp(-|\beta \wedge \gamma|)) \quad (\text{cf. figure}) \\ &\leq \max(d(\alpha, \beta), d(\beta, \gamma)). \end{aligned}$$

(Le sommet $\alpha \wedge \beta \wedge \gamma$ est forcément l'un des points $\alpha \wedge \beta$, $\alpha \wedge \gamma$, $\beta \wedge \gamma$. En effet, si ce n'était pas le cas, ces trois points seraient issus de trois fils différents de $\alpha \wedge \beta \wedge \gamma$ et ils seraient chacun dans deux des chemins sans être dans le troisième.)

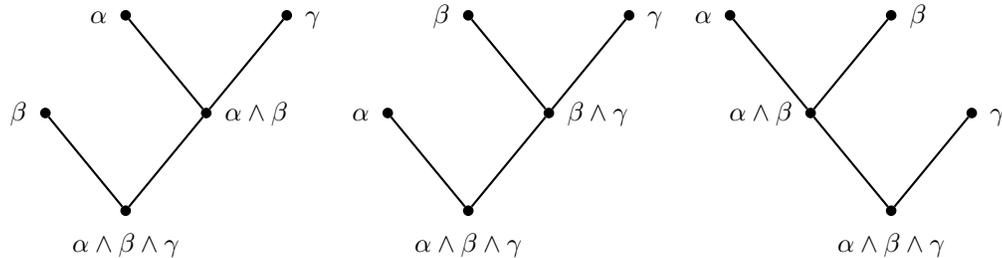


FIG. 10 – Les différents cas à étudier.

□

Proposition 2.6. Le bord d'un arbre localement fini (i.e. tout sommet a un nombre fini de fils), est un espace compact.

Démonstration. Il y a toujours un nombre fini de sommets de hauteur donnée. Par conséquent, on peut toujours extraire d'une suite une sous-suite telle que le sommet de hauteur n soit stationnaire à partir d'un certain rang (l'un d'eux apparaît une infinité de fois).

Par extraction diagonale, on peut extraire une suite convergente. Comme l'espace est métrique, il est compact, puisque séquentiellement compact. □

Cette définition fait du bord de A un espace métrique, et on peut donc parler de sa dimension de Hausdorff. Cependant, dans le cas particulier des arbres, il est commode d'introduire une définition plus simple de la dimension de Hausdorff.

Définition 2.3. Une coupe Π de A est un ensemble de nœuds tels que tout chemin de ∂A passe par un des nœuds de Π .

Proposition 2.7. Tout recouvrement de A par des boules s'identifie à une coupe Π .

Démonstration. Il suffit de vérifier qu'une boule est en fait l'ensemble des chemins passant par un sommet donné. Soit une boule ouverte $B(\xi, \varepsilon)$, avec $\varepsilon \leq 1$ (sinon $B = \partial A$ est la boule issue de o). Soit $n \in \mathbb{N}$ l'unique entier tel que $e^{-n} < \varepsilon \leq e^{-n+1}$ (soit $n = 1 + \lfloor \ln \varepsilon^{-1} \rfloor$). Par construction, $B = B(\xi, e^{-n})$ est la boule issue de ξ_n (le n -ième nœud de ξ). Ceci démontre la proposition. \square

Cette façon d'interpréter les recouvrements comme des coupes permet de simplifier la définition de la mesure de Hausdorff.

Lemme 2.8. Soit \mathcal{C} l'ensemble des coupes de A . Alors :

$$\mathcal{H}_s(\partial A) = \lim_{n \rightarrow \infty} \inf \left\{ \sum_{x \in \Pi} e^{-s|x|} \text{ pour } \Pi \in \mathcal{C} \text{ telle que } \forall x \in \Pi, |x| > n \right\}.$$

Démonstration. On remarque en premier lieu que toute partie E de ∂A est contenue dans une boule de même diamètre que E .

En effet, soit n tel que $e^{-n-1} < \text{diam } E \leq e^{-n}$ et $\xi \in E$. Alors $\forall \psi \in E$, $d(\xi, \psi) \leq \text{diam } E \leq e^{-n}$, donc ψ passe par ξ_n . La boule issue de ξ_n contient donc E , notée B_{ξ_n} . Par ailleurs, si ψ et φ appartiennent à $B = \overline{B}(\xi, e^{-n})$, on a :

$$d(\psi, \varphi) \leq \sup \{d(\xi, \psi), d(\xi, \varphi)\} \leq \text{diam } E.$$

On a donc trouvé une boule B contenant E et de même diamètre.

Soit (E_i) un recouvrement de ∂A . On peut sans changer $\sum_i (\text{diam } E_i)^s$, remplacer (E_i) par des boules. Réciproquement, un recouvrement par des boules est un recouvrement. Les recouvrements s'identifient donc aux coupes.

Le diamètre d'une boule issue d'un nœud x est exactement $e^{-|x|}$, à moins qu'il s'agisse d'un singleton. Mais comme la contribution des singletons est nulle à la limite, on obtient bien la formule désirée. \square

Théorème 2.9. Soit A un arbre et \mathcal{C} l'ensemble de ses coupes. Alors :

$$\dim \partial A = \inf \left\{ s \text{ tel que } \inf_{\Pi \in \mathcal{C}} \left(\sum_{x \in \Pi} e^{-s|x|} \right) = 0 \right\}.$$

Démonstration. On sait déjà que :

$$\dim \partial A = \inf \{s \geq 0 \text{ tel que } \mathcal{H}_s(\partial A) = 0\}$$

$$\text{et } \mathcal{H}_s(\partial A) = \lim_{n \rightarrow \infty} \inf \left\{ \sum_{x \in \Pi} e^{-s|x|} \text{ pour } \Pi \in \mathcal{C} \text{ telle que } \forall x \in \Pi, |x| > n \right\}.$$

Il faut donc montrer que $\mathcal{H}_s(\partial A) = 0$ si et seulement si $C_{\min} = \inf_{\Pi \in \mathcal{C}} (\sum_{x \in \Pi} e^{-s|x|})$ est nul. On sait déjà que si $\mathcal{H}_s(\partial A) = 0$, alors $C_{\min} = 0$.

Réciproquement, supposons $C_{\min} = 0$, et $\mathcal{H}_s(\partial A) > 0$. Il existe un certain rang N tel que si Π est incluse dans l'ensemble des nœuds de rang supérieur à N , $C(\Pi) > e^{-sN}$ (quitte à augmenter N), où $C(\Pi)$ est la somme des $e^{-s|x|}$ pour $x \in \Pi$. Soit une coupe Π_0 de taille inférieure à e^{-sN} . Elle ne peut pas contenir de nœud de hauteur inférieure à N , par conséquent $C(\Pi) > e^{-sN}$, ce qui est absurde. \square

Voyons maintenant le cas particulier des arbres réguliers :

Définition 2.4. Soit A un arbre k -aire. On considère une suite (X_n) de variables indépendantes et uniformes sur $[[1, k]]$, et l'application de l'espace de probabilité sous-jacent (Ω, \mathbb{P}) dans ∂A , qui à une suite (X_n) associe le chemin dont les sommets sont les (X_1, X_2, \dots, X_k) . La mesure de probabilité uniforme sur A est la mesure image de \mathbb{P} par cette application.

Proposition 2.10. Le bord d'un arbre k -aire est de dimension $\log k$.

Démonstration. On considère les coupes naturelles Π_n constituées des nœuds de hauteur n . Leur taille est $k^n e^{-sn}$ qui tend vers zéro pour $s > \log k$. Donc $\dim \partial A_k \leq \log k$.

Considérons la mesure de probabilité uniforme sur ∂A_k , et $s = \log k$. Cette mesure vérifie les hypothèses du lemme 2.4. En effet, soit G une partie de ∂A_k mesurable. Alors, G est contenue dans une boule B de diamètre $\text{diam } G = e^{-|x|}$, où x engendre B , donc sa mesure est au plus $k^{-|x|} = (\text{diam } G)^{(\log k)}$. Donc $\log k \leq \dim \partial A_k$. \square

2.2.2 Correspondance entre arbres et ensembles fractals

Soit E une partie de $[0,1]$. On construit le *codage b -adique de E* comme l'arbre dont les nœuds sont les intervalles b -adiques rencontrant E et dont la relation de filiation est la relation d'inclusion.

Exemple. Le codage dyadique de l'ensemble triadique de Cantor est un arbre binaire.



FIG. 11 – Triangle de Sierpinski.

Pour étudier des objets plus compliqués, comme le triangle de Sierpinski, il est nécessaire d'introduire une notion de codage plus générale :

Définition 2.5. Soit A un arbre. On suppose qu'à chaque nœud v de A , est associé un compact K_v de \mathbb{R}^d , de façon à ce que :

- $K_v = \overline{\text{int } K_v}$.
- $K_u \subseteq K_v$ si u est le père de v .
- K_v et K_w sont d'intérieurs disjoints si v et w sont frères.
- Si $\xi \in \partial A$, $\text{diam } K_{\xi_n} \rightarrow_{n \rightarrow \infty} 0$.
- Il existe des constantes C_l et C_v telles que :

$$\begin{cases} \text{diam } K_v \geq C_l \text{diam } K_w & \text{où } w \text{ est le père de } v \\ \lambda(\text{int } K_v) \geq C_v (\text{diam } K_v)^d. \end{cases}$$

Soit $E = \bigcup_{\xi \in \partial A} \bigcap_n K_{\xi_n}$.

On dispose d'une application canonique de ∂A dans \mathbb{R}^d en envoyant ξ sur $\bigcap_n K_{\xi_n}$ (intersection décroissante de compacts de diamètre tendant vers zéro). On dira que A code E .

Autrement dit, on construit l'ensemble E par itération en remplaçant à chaque fois un morceau par des parties plus petites, représentées dans l'arbre par les fils du nœud correspondant. Cette construction permet de ramener l'étude d'objets fractals à celle d'arbres. Le théorème suivant exprime cette similarité entre les deux objets :

Théorème 2.11. Si E est un ensemble codé par un arbre A , alors :

$$\dim E = \inf \left\{ s \text{ tel que } \inf_{\Pi \in \mathcal{C}} \sum_{x \in \Pi} (\text{diam } K_x)^s = 0 \right\}.$$

Démonstration. Comme pour le théorème 2.9, il suffit de montrer que $\mathcal{H}_s(E) = 0$ si et seulement si $\inf_{\Pi \in \mathcal{C}} \sum_{x \in \Pi} (\text{diam } K_x)^s = 0$.

Supposons $\mathcal{H}_s(E) = 0$. Soit $\varepsilon > 0$ et (E_i) un recouvrement dénombrable de E de taille inférieure à ε , c'est-à-dire tel que $\sum_i (\text{diam } E_i)^s \leq \varepsilon$.

On va recouvrir chaque E_i par des K_v dont la taille sera maîtrisée : pour chaque $\xi \in E_i$ (on identifie le bord de A à E), on considère le premier nœud v tel que $\text{diam } K_v \leq \text{diam } E_i$. L'ensemble de ces nœuds pour ξ parcourant E_i est noté V_i . Alors $\Pi = \bigcup_i V_i$ est une coupe, puisque tout chemin infini est dans un E_i , donc passe par construction par un sommet de V_i . Par ailleurs :

$$\begin{aligned} \sum_{x \in \Pi} (\text{diam } K_x)^s &\leq \sum_i \sum_{x \in V_i} (\text{diam } K_x)^s \\ &\leq \sum_i \sum_{x \in V_i} (\text{diam } E_i)^s \text{ car } \text{diam } K_x \leq \text{diam } E_i \\ &\leq \sup \{ \text{card } V_i \} \sum_i (\text{diam } E_i)^s = \varepsilon \sup \{ \text{card } V_i \}. \end{aligned}$$

Maintenant, il faut majorer le cardinal des V_i , qui sont des recouvrements des E_i par des ensembles d'intérieurs disjoints (deux K_v sont soit inclus l'un dans l'autre, soit d'intérieurs disjoints).

Or si $x \in V_i$, $\text{diam } K_x \leq \text{diam } E_i$, donc les K_x sont inclus dans la boule de centre Ω (point quelconque de E_i), et de rayon $2 \text{diam } E_i$. De plus, par construction, leur volume est minoré par :

$$\lambda(\text{int } K_x) \geq C_v (\text{diam } K_x)^d \geq C_v C_l^d (\text{diam } K_w)^d \geq C_v C_l^d (\text{diam } E_i)^d.$$

Comme les intérieurs sont disjoints, c'est que

$$\text{card } V_i \leq \frac{\lambda(B(0, 2))}{C_v C_l^d}.$$

Il existe donc une coupe de « taille » inférieure à $M\varepsilon$, où M est une constante. Par conséquent : $\inf_{\Pi \in \mathcal{C}} \sum_{x \in \Pi} (\text{diam } K_x)^s = 0$.

Supposons maintenant que $\inf_{\Pi \in \mathcal{C}} \sum_{x \in \Pi} (\text{diam } K_x)^s = 0$ et que $\mathcal{H}_s(E) > 0$. Alors il existe η et ε , tels que, si on recouvre E par des ensembles de diamètre inférieur à η , le recouvrement est de taille $< \varepsilon$. Les seules coupes de petite taille seraient parmi celles possédant un élément v tel que $\text{diam } K_v > \eta$. Bien sûr une telle coupe a une taille d'au moins η^s , ce qui est absurde. \square

Dans les cas simples, comme pour le triangle de Sierpinski ou l'ensemble de Cantor, les K_v sont homothétiques de leur père avec un rapport constant (entier). On a alors une caractérisation très simple.

Proposition 2.12. Supposons que les K_v soient tous homothétiques, avec un rapport b d'une génération à l'autre. Alors, $(\dim K_A)(\log b) = (\dim A)$.

Démonstration. Si $|v| = n$, $\text{diam } K_v = D_0 b^{-n}$, où D_0 est le diamètre de K_o . D'après les théorèmes précédents (2.9 et 2.11) :

$$\begin{aligned} \dim K_A &= \inf \left\{ s \text{ tel que } \inf_{\Pi \in \mathcal{C}} \sum_{x \in \Pi} D_0 b^{-s|x|} = 0 \right\} \\ \text{et } \dim A &= \inf \left\{ s \text{ tel que } \inf_{\Pi \in \mathcal{C}} \sum_{x \in \Pi} e^{-s|x|} = 0 \right\}. \end{aligned}$$

Le résultat est donc clair, puisque $e^{-s(\log b)|x|} = b^{-s|x|}$. \square

Proposition 2.13. Soit E une fractale construite par itération en remplaçant un motif par k motifs b fois plus petits (chaque motif étant pavé par b^d petits motifs).

$$\text{Alors } \dim E = \frac{\log k}{\log b}.$$

Démonstration. L'arbre correspondant est un arbre k -aire. C'est alors évident, puisque la dimension de l'arbre k -aire est $\log k$ (proposition 2.10). \square

Ceci prouve donc que la dimension de l'ensemble de Cantor est $(\log 2)/(\log 3)$, mais aussi que celle du triangle de Sierpinski est de $(\log 3)/(\log 2)$. Mieux encore : si on construit un triangle de Sierpinski en choisissant de façon quelconque le triangle à éliminer, sa dimension est toujours de $(\log 3)/(\log 2)$.

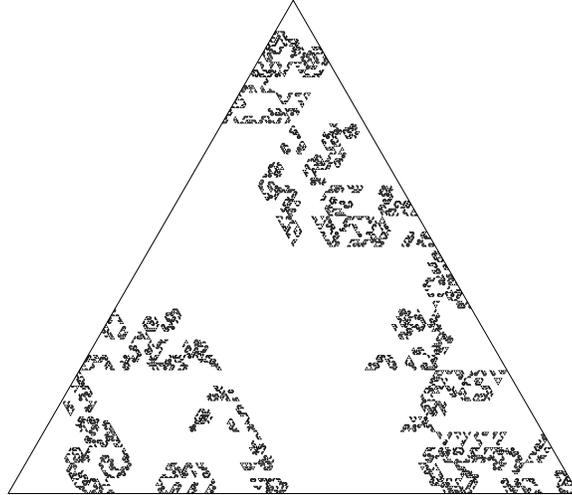


FIG. 12 – Triangle de Sierpinski aléatoire. À chaque itération, on découpe un triangle en quatre, et on choisit aléatoirement un de ces triangles pour l'évincer.

On a ainsi vu comment on pouvait calculer la dimension de Hausdorff d'un arbre (et donc celle des objets fractals associés) en cherchant des coupes minimales. Cependant, on voudrait quantifier plus précisément la valeur de cette dimension, en fonction des paramètres explicites de l'arbre. Et en particulier, pour un arbre de Galton-Watson, on verra que sa dimension de Hausdorff n'est pas aléatoire, mais complètement déterminée par la loi génératrice.

3 Percolation, flots, et nombre de branchement

3.1 Notion de nombre de branchement

3.1.1 Définitions et premières propriétés

L'un des paramètres les plus intuitifs d'un arbre est certainement le nombre moyen de branches partant d'un nœud donné. Pour un arbre régulier k -aire, ce nombre est manifestement k , qui intervient dans la formule de la dimension de Hausdorff de son bord. On souhaite généraliser cette notion à un arbre quelconque, tout en fournissant un paramètre lié de façon simple à la dimension de Hausdorff.

On pourrait par exemple étudier le comportement asymptotique de Z_n , le nombre de sommets de hauteur donnée. Pour un arbre k -aire, $Z_n = k^n$.

Cependant, si on considère l'arbre ci-contre, $Z_n = 2^n$, mais cependant il n'y a qu'un seul chemin infini, les autres étant des culs-de-sac. Le but étant de calculer la dimension de Hausdorff du bord de l'arbre, on comprend que cette notion ne convienne pas.

On va donc définir une notion plus asymptotique, le nombre de branchement : pour cela, il faut voir l'arbre comme le parcours d'un liquide qui se fraie un chemin au hasard dans un milieu. Le flot du liquide doit se diviser pour traverser, et c'est ce facteur de division qui est intéressant. Comme cette vision ignore les branches de longueur finie, elle est beaucoup plus adaptée à l'étude de la structure asymptotique de l'arbre.

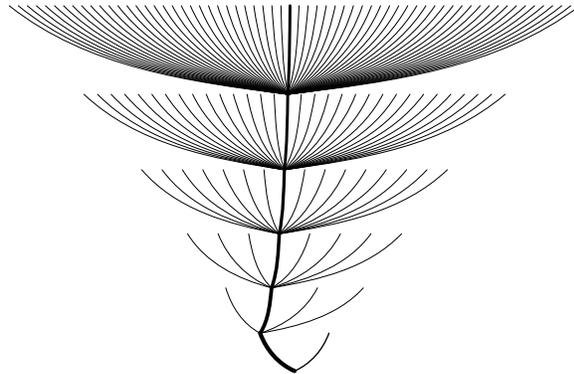


FIG. 13 – Arbre dont le bord est un singleton.

Définition 3.1.

- Un *flot* sur un graphe G est une fonction numérique sur ses arêtes orientées ($f(e^-) = -f(e^+)$) telle que pour tout nœud, la somme des flots arrivant en ce nœud est nulle, sauf à la source s où cette somme est positive et au puits t où elle est négative.
- L'*intensité* du flot f est la somme des flots sortant de la source.
- Une *capacité* sur un graphe est une fonction positive sur les arêtes du graphe. Lorsqu'une capacité est définie, on supposera toujours le flot borné par cette capacité.

Remarque. Pour un arbre, on suppose que la source est la racine o est que le puits est à l'infini (il n'y en a pas). Par ailleurs, les arêtes sont en bijection avec les sommets (sauf o), en associant à une arête son fils, et à un sommet l'arête unique qui s'y termine. On identifiera souvent les deux.

Définition 3.2. Le *nombre de branchement* (*branching number*) de l'arbre A est la borne supérieure des λ tels qu'il existe un flot non nul sur l'arbre A muni de la capacité $\lambda^{-|e|}$, où $|e|$ est la hauteur du nœud fils de e . On le notera $br A$.

Remarque. On peut voir très facilement que les λ tels qu'il existe un flot non nul issu de o vérifiant $f(e) \leq \lambda^{-|e|}$ sont en fait tous les éléments de $]0, br A[$, avec éventuellement l'extrémité.

Remarque. Le mot « nombre de branchement » est une traduction exécrationnelle. L'expression « indice de ramification » est peut-être plus parlant.

Proposition 3.1. Le nombre de branchement d'un arbre k -aire est k .

Démonstration. Soit un flot sur un arbre k -aire valant 1 à la racine. Alors il existe une arête issue de o dont le flot est au moins k^{-1} (sinon la somme ne ferait pas 1).

De même, pour tout n , il existe un x de hauteur n , où le flot est au moins k^{-n} . S'il existait un flot majoré par $\lambda^{-|x|}$, on aurait : $f(o)k^{-n} \leq \lambda^{-n}$, et ce, pour tout n , en choisissant à chaque fois le bon x . Donc quand $|x|$ tend vers ∞ , on trouve $k \geq \lambda$. Par ailleurs, le flot canonique $f(x) = k^{-|x|}$ montre que le nombre de branchement est exactement k . \square

3.1.2 Lien avec la dimension de Hausdorff

Théorème 3.2 (Max flow/Min cut Theorem). *Soit A un arbre et c une fonction de capacité sur les arêtes. Alors l'intensité maximale d'un flot issu de o sur A est :*

$$\inf_{\Pi \in \mathcal{C}} \sum_{x \in \Pi} c(x).$$

Démonstration. Soit une coupe Π et un flot f .

Soit U_Π l'ensemble de tous les ancêtres (stricts) de Π . Par compacité de l'arbre, Π est bornée en hauteur, et U_Π est fini. On peut compter de deux façons différentes l'intensité du flot :

- en sommant en chaque point le flot entrant, ce qui donne $I(f)$, car seul o apporte une contribution,
- en sommant la contribution de chaque arête, qui est non nulle seulement si l'une des extrémités de l'arête est dans U_Π et pas l'autre. Dans ce cas, l'arête est dans Π . En effet, dans le cas contraire, on pourrait sortir de U_Π sans passer par Π , et on est donc dans un cul-de-sac, sinon il y aurait un chemin infini ne passant pas par Π . La somme est donc au plus $\sum_{x \in \Pi} c(x)$.

Donc $I(f) \leq \sum_{x \in \Pi} c(x)$. Ceci donne une première inégalité, car c'est vrai pour tous f et Π .

Soit un entier n et A_n le graphe obtenu en coupant l'arbre à la hauteur n et en reliant tous les sommets de hauteur n à un puits fictif t . Soit Π une coupe sur A_n , c'est-à-dire un ensemble de sommets/arêtes séparant o de t .

Soit F_n un flot d'intensité maximale sur A_n . On suppose qu'il existe un chemin de o à t le long duquel $F_n(e) < c(e)$ (un tel chemin est dit *augmentant*). Alors évidemment en choisissant ε comme le plus petit écart entre $F_n(e)$ et $c(e)$ (les chemins sont finis), on peut augmenter F_n de ε le long de ce chemin, et F_n reste un flot. Donc F_n n'admet pas de chemin augmentant.

Soit X l'ensemble des sommets v tels que l'on puisse trouver un chemin augmentant de o à v . On vient de voir que $t \notin X$, par contre $o \in X$. Donc la frontière de X , i.e. l'ensemble des arêtes ayant une extrémité dans X et l'autre à l'extérieur de X , est une coupe. De plus, $F_n(e) = c(e)$ si $e \in \partial X$. Car si on avait $F_n(e) < c(e)$, on pourrait prendre un chemin augmentant jusqu'à e^- , le prolonger jusqu'à e^+ , et avoir un chemin augmentant de o à $e^+ \notin X$, ce qui est impossible. Donc $I(F_n) = c(\partial X)$. Donc une coupe minimale est inférieure à l'intensité de F_n .

Quand n tend vers l'infini, la taille minimale d'une coupe de A_n notée c_n , tend vers la taille minimale des coupes de A , à savoir $c_\infty = \inf_{\Pi \in \mathcal{C}} c(\Pi)$. En effet, $c_n \geq c_\infty$, car toute coupe de A_n est une coupe de A . Par ailleurs, si Π est une coupe de A de taille inférieure à $c_\infty + \varepsilon$, elle est bornée en hauteur par compacité, donc c'est une coupe d'un des A_n . D'où $c_n < c_\infty + \varepsilon$.

Par conséquent $\lim c_n = c_\infty$. Mais on vient de construire des flots partiels F_n d'intensité c_n . Par extraction diagonale, on peut construire F qui soit limite simple des $F_{\varphi(n)}$. Soit x un sommet de hauteur q . Pour tout n tel que $\varphi(n) > q$, la loi des nœuds est vérifiée en x par F_n , donc par F en passant à la limite. Donc F est un flot, et en passant à la limite en o , $I(F) = c_\infty$. \square

Proposition 3.3. Soit A un arbre. Alors $\dim \partial A = \log \text{br } A$.

Démonstration. D'après le théorème 2.9, on a :

$$\dim \partial A = \inf \left\{ s \text{ tel que } \inf_{\Pi \in \mathcal{C}} \left(\sum_{x \in \Pi} e^{-s|x|} \right) = 0 \right\}.$$

Et par contraposition de la définition :

$$\log \text{br } A = \inf \left\{ s \text{ tel que pour tout flot non nul } f, \exists x f(x) > e^{-s|x|} \right\}.$$

Or d'après le théorème 3.2 (Max flow/Min cut), on sait que les deux conditions sont justement équivalentes, puisque le flux d'entrée maximal pour une capacité de $e^{-s|x|}$ est exactement $\inf_{\Pi \in \mathcal{C}} \left(\sum_{x \in \Pi} e^{-s|x|} \right)$. \square

3.2 Percolation et flots

3.2.1 Percolation sur un arbre

On voudrait maintenant calculer le nombre de branchement d'un arbre de Galton-Watson. Pour cela, nous allons introduire la notion de percolation : comme le suggère son nom, la percolation simule l'écoulement d'un liquide dans un réseau. On suppose que la possibilité pour le liquide d'emprunter une arête du réseau est une variable aléatoire, et qu'elles sont toutes indépendantes. On peut également imaginer un réseau informatique infini, dont les connexions tombent en panne (ou en marche) aléatoirement¹⁰ et on se demande si on a des chances de rester connecté à un nombre fini ou infini de machines.

Dans la suite, on ne travaillera qu'avec des graphes localement finis (i.e. chaque sommet est de degré fini), ce qui est le cas des arbres de Galton-Watson.

Définition 3.3. Soit G_0 un graphe dénombrable, dont les arêtes (e_n) sont indexées par \mathbb{N} , et (B_n) une suite de variables indépendantes de loi de Bernoulli $\mathcal{B}(p)$, sur un espace de probabilité Ω . Soit G le graphe aléatoire constitué des arêtes e_n telles que $B_n = 1$. Le processus aléatoire ainsi défini s'appelle une *percolation de Bernoulli*, de paramètre p .

Proposition 3.4. On travaille dans la tribu qui rend mesurables les événements $\{x \in G\}$ et $\{e \in G\}$ pour tout sommet x et toute arête e du graphe G_0 . Cette mesurabilité est essentielle pour assurer que ce que l'on dit a un sens.

Soient x et e fixés. Les fonctions suivantes sont mesurables pour cette tribu :

- $\mathbb{1}[x \text{ et } e \text{ sont dans la même composante connexe}]$,
- $\mathbb{1}[x \text{ est dans une composante connexe infinie}]$.

Démonstration. L'événement : $\{x \text{ est une extrémité de } e\}$ est mesurable (c'est l'intersection de $\{x \text{ est extrémité de } e \text{ dans } G_0\}$ et de $\{e \in G\}$).

De même, l'événement $\{d(x, e) \leq n\}$ est mesurable, car c'est l'événement « il existe n arêtes reliant x à une extrémité de e ». Or de tels chemins sont en quantité dénombrable. Et pour un chemin donné, l'événement : « chaque arête du chemin est dans G » est mesurable. En prenant la réunion pour tout n , on obtient que $\{d(x, e) < \infty\}$ est mesurable, ce qu'on voulait.

L'événement $\{d(x, e) > n\}$ est mesurable pour tout e , ainsi que $\{x \text{ et } e \text{ sont reliés}\}$. En prenant la réunion pour tous les e , on déduit que l'événement « il existe une arête éloignée de plus de n du point x » est mesurable. Si on prend l'intersection pour $n \in \mathbb{N}$, on trouve que l'événement : « il existe des points reliés à x arbitrairement loin » est mesurable.

Cet événement implique que la composante connexe de x est infinie, et si au contraire les points étaient à distance bornée, cette composante serait finie (par locale finitude). D'où la mesurabilité du second événement. \square

Intuitivement, si on considère une percolation de probabilité de survie $(\text{br } A)^{-1}$ sur un arbre A , il restera en moyenne une seule branche par nœud. En dessous, il y aura moins d'un fils en moyenne, et l'arbre s'éteindra. Au-dessus de ce seuil, il y aura plus d'un fils et l'arbre aura un bord non vide. C'est précisément ce que l'on va prouver. La proposition suivante, évidente, permet de quantifier ce seuil.

Proposition 3.5. On suppose les arêtes de G_0 énumérées par \mathbb{N} .

Soit (U_n) une suite de variables aléatoires indépendantes et de même loi uniforme sur $[0, 1]$. On note $G_p(\omega)$ le graphe obtenu en supprimant de G_0 les arêtes e_n si $U_n < p$. Alors la loi de G_p est exactement \mathbb{P}_p , la loi de la percolation de Bernoulli. Ce couplage est appelé *couplage standard de la percolation de Bernoulli*.

Proposition 3.6. Il existe un réel (unique) p_c tel que la probabilité que $G_p(\omega)$ ait une composante connexe infinie soit 0 pour $p < p_c$ et 1 pour $p > p_c$.

¹⁰On peut par exemple modéliser ce phénomène expérimentalement en confiant à des singes la gestion d'un réseau informatique.

Démonstration. L'événement « G_p a une composante connexe infinie » appartient à la tribu asymptotique (effectuer la percolation sur les n premières arêtes ne change rien quant au résultat). Donc la probabilité vaut 0 ou 1, d'après la loi du tout ou rien de Kolmogorov. Par ailleurs, le couplage précédent prouve clairement que la probabilité de garder une composante infinie est croissante, d'où le résultat. \square

Proposition 3.7. On suppose G_0 connexe. Soit x un sommet quelconque. Alors p_c est aussi la borne inférieure des p tels que x soit dans une composante infinie avec probabilité strictement positive.

Démonstration. Si G_p n'a que des composantes finies p.s., la composante de x est finie p.s.

Si au contraire G_p a une composante infinie p.s., on considère une suite exhaustive Γ_n de sous-graphes connexes de G_0 , contenant x . Il existe alors n tel que $\mathbb{P}[E_n] > 0$, où E_n désigne l'événement : « Γ_n contient un point dans la composante est infinie ». Or cette probabilité est bien sûr inférieure à la probabilité de E_n sachant S_n , où S_n est l'événement : « toutes les arêtes de Γ_n survivent à la percolation ». On a $\mathbb{P}[S_n] = p^N$, où N est le nombre d'arêtes de Γ_n .

Donc $\mathbb{P}[E_n \cap S_n] > 0$; or cet événement implique que x est dans une composante infinie, car il existe alors un point de Γ_n dans une composante infinie, et cette composante est reliée à x par connexité. \square

3.2.2 Percolation et nombre de branchement

La percolation se révèle être un bon moyen d'interpréter le nombre de branchement, et offre des manières parfois inattendues de le calculer. En particulier, elle s'adapte très bien à la nature probabiliste des arbres de Galton-Watson.

Proposition 3.8. Pour une percolation quelconque sur G_0 , et x un sommet de G_0 , la probabilité que x soit dans une composante infinie de $G(\omega)$ est majorée par :

$$\inf_{\Pi \in \mathcal{C}} \left\{ \sum_{e \in \Pi} \mathbb{P}[x \leftrightarrow e] \right\}.$$

où \mathcal{C} est l'ensemble des coupes séparant x de l'infini, c'est-à-dire des ensembles d'arêtes telles que si on les enlève, la composante connexe de x soit finie. $\{x \leftrightarrow e\}$ désigne l'événement : « x reste relié à e » après percolation.

Démonstration. Évident, puisque pour avoir une composante infinie, il faut être relié à l'une des arêtes de Π . \square

Corollaire 3.9. Pour un arbre A , la probabilité critique de la percolation de Bernoulli est supérieure à $\frac{1}{\text{br } A}$.

Démonstration. D'après le théorème 3.2 (Max flow-min cut), on sait que le flot maximal est non nul pour une capacité λ^{-n} si et seulement si :

$$\inf_{\Pi \in \mathcal{C}} \sum_{x \in \Pi} \lambda^{-|x|} > 0 \text{ c'est-à-dire si } \lambda < \text{br } T.$$

Or d'après la proposition précédente, la probabilité que la racine soit dans une composante infinie est majorée par :

$$\inf_{\Pi \in \mathcal{C}} \sum_{x \in \Pi} p^{|x|}.$$

puisque la probabilité que o soit lié à une arête dont le fils est de hauteur n est exactement p^n (il faut laisser toutes les arêtes les reliant).

Ainsi, o est dans une composante finie p.s. si p^{-1} est supérieur à $\text{br } A$. Donc par contraposition, si o est dans une composante infinie p.s., $p \geq \frac{1}{\text{br } A}$, (par la loi du tout ou rien). Le seuil critique est donc $p_c \geq \frac{1}{\text{br } A}$. \square

Il y a en fait égalité, non seulement p.s. pour les arbres de Galton-Watson, mais également pour un arbre quelconque (cas que l'on n'étudiera pas). Ceci permet donc de calculer le nombre de branchement d'un arbre quelconque en déterminant sa probabilité critique de percolation. C'est très difficile pour un graphe quelconque (par exemple, sur les réseaux¹¹ \mathbb{Z}^n). En revanche, il faut remarquer qu'une percolation de Bernoulli de paramètre p transforme un processus de Galton-Watson de moyenne m en un processus de moyenne pm . On voit alors intuitivement que $p_c = 1/m$ et $m = \text{br } A$.

Proposition 3.10. La probabilité critique de percolation d'un arbre engendré par un processus de Galton-Watson de moyenne m est exactement $1/m$ presque sûrement, sur l'événement de non-extinction.

Démonstration. Soit $T'(\omega)$ la composante de la racine dans l'arbre issu de T après percolation. Alors, si T est l'arbre aléatoire du processus de Galton-Watson de moyenne m , la variable aléatoire T' a même loi qu'un certain processus de Galton-Watson de moyenne mp .

En effet, le nombre de fils d'un sommet (après percolation) est une somme de L variables aléatoires indépendantes suivant une loi de Bernoulli, où L est la variable aléatoire correspondant au processus de Galton-Watson d'origine. C'est donc une loi qui ne dépend pas de la position dans l'arbre, et par ailleurs,

$$\begin{aligned} \mathbb{E} \left[\sum_{i=1}^L B_i \right] &= \mathbb{E} \left[\mathbb{E} \left[\sum_{i=1}^L B_i \mid L \right] \right] = \mathbb{E} \left[\sum_{i=1}^L \mathbb{E} [B_i \mid L] \right] \quad \text{car } \sum_{i=1}^L \text{ est une fonction } L\text{-mesurable,} \\ &= \mathbb{E} [pL] = pm \quad \text{car } B_i \text{ et } L \text{ sont indépendantes.} \end{aligned}$$

De plus, on sait que $mp \leq 1$ si et seulement si l'extinction est presque sûre. Par conséquent $\mathbb{P}[|T'| < \infty] = 1$. Par les propriétés de l'espérance conditionnelle, c'est aussi $\mathbb{E}[\mathbb{P}[|T'| < \infty \mid T]]$, ce qui signifie que pour presque tout T , la percolation donne presque sûrement une composante finie à la racine, si $p \leq \frac{1}{m}$.

Inversement, si $p > \frac{1}{m}$, on remarque que la propriété « une percolation sur T laisse une composante finie p.s. » est héréditaire, et on a vu qu'une propriété héréditaire avait pour probabilité 0 ou 1, sachant la non-extinction. Mais ce n'est pas 0, sinon T' serait fini p.s., et on aurait $mp \leq 1$. Donc pour presque tout T , la percolation laisse une composante infinie à la racine.

On a donc bien la dichotomie qui définit la probabilité critique de percolation $\frac{1}{m}$, pour presque tout T sur la non-extinction. \square

Corollaire 3.11. On a presque sûrement sur l'événement de non-extinction :

$$\text{br } T = m \quad \text{et} \quad \dim \partial T = \log m.$$

Démonstration. On sait déjà que $\text{br } T = \exp(\dim \partial T)$ et que $\text{br } T \geq \frac{1}{p_c(T)} = m$.

Par ailleurs, le nombre de branchement est majoré par m , puisque si $\lambda > m$, les capacités des arêtes de la n^{e} génération sont au total majorées par Z_n/λ^n , mais on sait que Z_n/m^n converge presque sûrement (c'est une martingale, comme on l'a vu dans la partie 1), donc Z_n/λ^n tend vers zéro, et il ne peut y avoir de flot non nul. \square

Corollaire 3.12. Soit E un ensemble défini par des ensembles diminuant régulièrement dans un rapport b , codé par un arbre de Galton-Watson de moyenne m .

$$\text{Alors } \dim E = \frac{\log m}{\log b}.$$

¹¹On peut en fait démontrer qu'elle vaut $\frac{1}{2}$ sur \mathbb{Z}^2 , et il n'est pas difficile de voir que cette probabilité est décroissante en d .

3.2.3 Flots sur un arbre de Galton-Watson

Définition 3.4. On considère un processus aléatoire qui engendre des arbres aléatoires, avec des capacités aléatoires, tout en préservant la propriété d'auto-similarité.

Soit L une variable aléatoire, et (A_1, \dots, A_L) un L -uplet de variables aléatoires. On construit récursivement un arbre en donnant L fils à la racine, et on affecte la valeur A_i au sommet i . Ensuite, pour toute feuille y (on reprend le codage des arbres exposé dans la section 1), on considère une copie indépendante de (L, A_1, \dots, A_L) , notée $X_y = (L_y, A_{y1}, \dots, A_{yL_y})$, et on donne L_y fils à y , notés $y1, \dots, yL$, auxquels on affecte la valeur A_{yk} .

À chaque sommet $x = x_1 \dots x_n$ est alors associé la capacité $c(x) = \prod_{k \leq n} A_{x_1 \dots x_k}$ (voir le codage des arbres dans la partie 1). L'arbre aléatoire muni de cette capacité est appelé *réseau de Galton-Watson*.

Les A_i représentent la proportion de la capacité entrante qui est redistribuée au fils i . En particulier, l'arbre régulier est engendré par la variable constante $(k, 1/k, \dots, 1/k)$. On comprend intuitivement que si la somme des A_i est inférieure à 1, la capacité décroît à la génération suivante.

La grandeur cruciale est donc : $\gamma = \mathbb{E} \left[\sum_{i=1}^L A_i \right]$.

Le théorème suivant, dû à Falconer (1986), confirme cette intuition :

Théorème 3.13 (Falconer).

- Si $\gamma \leq 1$, il n'existe presque jamais de flot non nul, sauf si $\sum_{i=1}^L A_i = 1$ p.s.
- Si $\gamma > 1$, il existe un flot non nul p.s. sur l'événement de non-extinction.

Démonstration.

- Soit x un fils de o . Conditionnellement au fait que o ait x pour fils, l'arbre issu de x , noté T_x , normalisé en divisant les capacités par A_x , a même loi que l'arbre entier. Soit I_x l'intensité maximale d'un flot sur T_x .

On voit que si I est l'intensité maximale des flots sur l'arbre entier, $I = \sum_{i=1}^L A_i \min(1, I_i)$.

On peut alors calculer l'espérance :

$$\begin{aligned} \mathbb{E}[I] &= \mathbb{E}[\mathbb{E}[I | X_0]] \\ &= \mathbb{E} \left[\mathbb{E} \left[\sum_{i=1}^L A_i \min(1, I_i) \mid X_0 \right] \right] \\ &= \mathbb{E} \left[\sum_{i=1}^L A_i \mathbb{E}[\min(1, I_i) | X_0] \right] \text{ car } L \text{ et les } A_i \text{ sont } \sigma(X_0)\text{-mesurables,} \\ &= \mathbb{E} \left[\sum_{i=1}^L A_i \mathbb{E}[\min(1, I)] \right] \text{ car les } I_i \text{ ont même loi que } I, \\ &= \gamma \mathbb{E}[\min(1, I)]. \end{aligned}$$

Donc $\mathbb{E}[I] \leq \mathbb{E}[\min(1, I)]$, et $I \leq 1$, donc si $\gamma < 1$, $I = 0$ p.s. Sinon, lorsque $\gamma = 1$, comme les A_i et les I_i sont indépendantes,

$$\|I\|_\infty = \left\| \sum_{i=1}^L A_i I_i \right\|_\infty \leq \left\| \sum_{i=1}^L A_i \right\|_\infty \|I\|_\infty \text{ car } I \text{ et } I_i \text{ ont même loi,}$$

et de plus, si $\varepsilon > 0$,

$$\begin{aligned} \mathbb{P} \left[I_i > \|I\|_\infty - \frac{\varepsilon}{\left\| \sum_{i=1}^L A_i \right\|_\infty} \right] &> 0 \\ \mathbb{P} \left[\sum_{i=1}^L A_i > \left\| \sum_{i=1}^L A_i \right\|_\infty - \frac{\varepsilon}{\|I\|_\infty} \right] &> 0 \end{aligned}$$

$$\mathbb{P} \left[\sum_{i=1}^L A_i I_i = I > \left\| \sum_{i=1}^L A_i \right\|_{\infty} \|I\|_{\infty} - \varepsilon \right] > 0 \quad \begin{array}{l} \text{en prenant l'intersection de ces} \\ \text{événements indépendants.} \end{array}$$

Donc $\|I\|_{\infty} = \left\| \sum_{i=1}^L A_i \right\|_{\infty} \|I\|_{\infty}$. D'où $\left\| \sum_{i=1}^L A_i \right\|_{\infty} = 1$, si I n'est pas p.s. nulle (puisqu'alors $\|I\|_{\infty} = 0$). Comme $\mathbb{E} \left[\sum_{i=1}^L A_i \right] = 1$, on en déduit $\sum_{i=1}^L A_i = 1$ p.s.

- On suppose $\gamma > 1$. On introduit une percolation, c'est-à-dire un processus aléatoire de suppression des arêtes d'un graphe. Pour cela, on considère pour chaque arête des variables aléatoires indépendantes (i.e. conditionnellement à la tribu \mathcal{F} engendrée par les X_y), (p_y) où p_y suit une loi de Bernoulli de paramètre A_y .

On considère alors le sous-arbre constitué des sommets tels que $p_y = 1$ sur tout le chemin de o à y , c'est-à-dire qu'on supprime les arêtes telles que $p_y = 0$, et qu'on regarde la composante connexe de o .

Le flot maximal F est, par le théorème du max flow/min cut, la capacité minimale d'une coupe. Or la capacité d'une arête est le produit des A_y , où y parcourt les ancêtres de l'arête en question. Et A_y est justement la probabilité que y survive à la percolation, et finalement, par construction, une arête e subsiste avec probabilité $c(e)$ (conditionnellement à \mathcal{F}).

De plus, l'arbre percolé est un arbre de Galton-Watson, engendré par la loi $\tilde{L} = \sum_{i=1}^L p_i$, dont l'espérance est justement γ , car p_i est d'espérance A_i . Sa probabilité de non-extinction q est donc non nulle. Par ailleurs la probabilité conditionnelle de non-extinction sachant \mathcal{F} , notée \mathcal{Q} , est d'espérance q .

Conditionnellement à \mathcal{F} , lorsque la percolation laisse une composante infinie, elle laisse obligatoirement une des arêtes d'une coupe, puisqu'elle laisse un chemin infini (qui passe donc par la coupe). Donc pour toute coupe Π , on a :

$$\begin{aligned} \sum_{x \in \Pi} \mathbb{1}[x \text{ survit}] &\geq \mathbb{1}[\text{non-extinction}] \\ \sum_{x \in \Pi} c(x) &\geq \mathcal{Q} \end{aligned} \quad (\text{espérance conditionnelle par rapport à } \mathcal{F}).$$

Le flot maximal F est donc minoré par \mathcal{Q} . Or $\mathbb{E}[\mathcal{Q}] > 0$, donc $\mathbb{P}[F > 0] > 0$. Donc $\{F = 0\}$ n'est pas presque sûr, et c'est une propriété héréditaire (s'il existe un flot non nul, on peut le restreindre à un sous-arbre), donc il est limité à l'événement d'extinction, et $F > 0$ p.s. sur l'événement de non-extinction. □

L'intérêt de ce théorème est de pouvoir définir des fractales aléatoires, toujours par itération, mais avec des rapports d'homothétie aléatoires. On introduit donc une généralisation de la méthode de codage de la section 2.

Définition 3.5. Soit un réseau de Galton-Watson, engendré par le L -uplet (L, A_1, \dots, A_L) de variables aléatoires. On suppose que les A_i sont dans $]0, 1[$ p.s. On associe à chaque sommet de l'arbre un compact $K_v \subset \mathbb{R}^d$ tel que :

- $K_v = \overline{\text{int } K_v}$.
- $K_v \subseteq K_w$ où w est le père de v .
- K_v et K_w sont d'intérieurs disjoints si v et w sont frères.
- Si $\xi \in \partial A$, $\text{diam } K_{\xi_n} \rightarrow_{n \rightarrow \infty} 0$.
- Il existe une constante C telle que $\lambda(\text{int } K_v) \geq C(\text{diam } K_v)^d$.

Mais au lieu de minorer le facteur de réduction entre un père et son fils par une constante, on le définit par une variable aléatoire, à savoir $\text{diam } K_v = c(v)$, la capacité de l'arête en v . L'ensemble codé par le réseau est alors encore $E = \bigcup_{\xi \in \partial A} \bigcap_{n \in \mathbb{N}} K_{\xi_n}$.

Le théorème suivant permet de calculer la dimension de E . Il est dû à Falconer, Mauldin et Williams.

Théorème 3.14. Sur l'événement de non-extinction, la dimension de Hausdorff de E vaut presque sûrement :

$$\min \left\{ s, \mathbb{E} \left[\sum_{i=1}^L A_i^s \right] \leq 1 \right\}.$$

Démonstration. On commence par montrer la continuité à droite de l'espérance indiquée, pour justifier qu'il existe un s minimal. Pour cela, on remarque que si Δ est le diamètre de K_o , les K_x (pour x fils de o) sont de diamètre $A_x \Delta$, donc leur volume est au moins $C A_x^d \Delta^d$. Comme ils sont presque disjoints et dans K_o , $\sum_{i=1}^L C A_x^d \Delta^d \leq \lambda(K_o) \leq \Delta^d$, car on peut inclure K_o dans un cube d'arête Δ (vu que les coordonnées ne peuvent pas varier de plus de Δ).

Donc $\sum_{i=1}^L A_x^d \leq \frac{1}{C}$. De plus, $\sum_{i=1}^L A_i^s$ est une fonction décroissante de s , qui tend vers 0 en $+\infty$ par convergence dominée. On en déduit l'existence d'un s_0 infimum. Enfin, le théorème de convergence dominée permet d'établir la continuité de l'espérance à droite de tout s pour laquelle elle est finie. Ainsi s_0 est un minimum.

1^{er} cas : les A_i sont tous dans $]\varepsilon, 1 - \varepsilon[$ p.s., pour un certain $\varepsilon > 0$. Donc pour tout sommet v de père w , $\text{diam } K_v = A_i \text{ diam } K_w > \varepsilon \text{ diam } K_w$. On retrouve alors la construction de la section 2, et le théorème 2.11 s'applique, et

$$\dim E = \inf \left\{ s \text{ tel que } \inf_{\Pi \in \mathcal{C}} \sum_{x \in \Pi} (\text{diam } K_x)^s = 0 \right\} \text{ mais } \text{diam } K_x = c(x),$$

$$\text{donc } \dim E = \inf \left\{ s \text{ tel que } \inf_{\Pi \in \mathcal{C}} \sum_{x \in \Pi} c(x)^s = 0 \right\}.$$

Il s'agit donc en fait de voir si les coupes d'un réseau engendré par une variable (L, A_1^s, \dots, A_L^s) sont minorées, c'est-à-dire s'il existe un flot non nul : c'est justement l'objet du théorème 3.13 : les s qui empêchent l'existence d'un flot sont ceux tels que $\mathbb{E} \left[\sum_{i=1}^L A_i^s \right] \leq 1$, ce qui donne le résultat. Noter qu'alors, $\inf_{\Pi \in \mathcal{C}} \sum_{x \in \Pi} (\text{diam } K_x)^s = 0$, et que $\mathcal{H}_s(E) = 0$.

2^e cas : on se ramène au cas précédent en considérant les arbres $T(\varepsilon)$, constitués des branches du réseau de départ telles que $A_x \in]\varepsilon, 1 - \varepsilon[$, qui donnent un sous-ensemble $E(\varepsilon)$ de E . Remarquer que $T(\varepsilon)$ est un réseau de Galton-Watson, engendré par la variable

$$\left(\tilde{L} = \sum_{i=1}^L \mathbb{1} [A_i \in]\varepsilon, 1 - \varepsilon[], \text{ les } A_i \text{ qui sont dans }]\varepsilon, 1 - \varepsilon[] \right).$$

Alors, comme $\mathcal{H}_s(E(\varepsilon)) = \inf_{\Pi \in \mathcal{C}} \sum_{x \in \Pi} c(x)^s$, si $\mathbb{E} \left[\sum_{i=1}^L A_i^s \mathbb{1} [\varepsilon < A_i < 1 - \varepsilon] \right] > 1$, la mesure de Hausdorff de $E(\varepsilon)$ est non nulle p.s. (sur la non-extinction). Comme E contient $E(\varepsilon)$, sa mesure est non nulle également, p.s. sur la non-extinction de $T(\varepsilon)$.

– Supposons $\mathbb{E} \left[\sum_{i=1}^L A_i^s \right] > 1$. Par convergence monotone, $\mathbb{E} \left[\sum_{i=1}^L A_i^s \mathbb{1} [\varepsilon < A_i < 1 - \varepsilon] \right] > 1$ pour ε assez petit, et E est de mesure non nulle sur la non-extinction de $T(\varepsilon)$ pour ε assez petit.

Par ailleurs, la loi de reproduction de $T(\varepsilon)$, qui est celle de $\left(\sum_{i=1}^L \mathbb{1} [\varepsilon < A_i < 1 - \varepsilon] \right)$, converge p.s. vers L quand ε tend vers zéro, donc converge en loi, ce qui impose aux fonctions génératrices de converger : la probabilité d'extinction de $T(\varepsilon)$ tend vers celle de T .

Par conséquent, $\mathcal{H}_s(E) > 0$ sur l'intersection des événements de non-extinction des $T(\varepsilon)$, qui est incluse dans la non-extinction de T , et de même mesure.

– Inversement, si $\mathbb{E} \left[\sum_{i=1}^L A_i^s \right] \leq 1$, on a remarqué plus haut que la mesure de E était nulle.

On a donc $\dim E = \min \left\{ s, \mathbb{E} \left[\sum_{i=1}^L A_i^s \right] \leq 1 \right\}$, et de plus $\mathcal{H}_{\dim E}(E) = 0$. □

4 Applications concrètes

4.1 Exemples de fractales aléatoires

Rappelons ici les théorèmes les plus importants :

Théorème. Soit E une fractale construite par itération en remplaçant un motif par k motifs b fois plus petits (chaque motif est pavé par b^d petits motifs).

$$\text{Alors } \dim E = \frac{\log k}{\log b}.$$

Théorème. Soit E une fractale construite par itération en remplaçant un motif par L motifs b fois plus petits (chaque motif est pavé par b^d petits motifs), avec $\mathbb{E}[L] = m$.

$$\text{Alors } \dim E = \frac{\log m}{\log b}.$$

Théorème. Soit E une fractale construite par itération en remplaçant un motif par L motifs dans des rapports A_1, \dots, A_L .

$$\text{Alors } \dim E = \min s, \mathbb{E} \left[\sum_{i=1}^L A_i^s \right] \leq 1.$$

Remarque. Ces 3 théorèmes sont dans un ordre croissant d'inclusion : le premier est un cas particulier du second (avec L constant), et le second un cas particulier du troisième ($A_i = 1/b$).

Exemple 1. Considérons un ensemble triadique de Cantor aléatoire construit de la manière suivante :

- $E_0 = [0, 1]$.
- Étant donné $E_n = \bigcup [\frac{a_n}{3^n}, \frac{a_{n+1}}{3^n}]$, on remplace chacun de ces segments par un nombre aléatoire de segments (en position quelconque, éventuellement aléatoire) de taille $1/3$, de manière indépendante, et selon la loi d'une variable aléatoire N .

Sur l'événement $\{E \neq \emptyset\}$, $E = \bigcap E_n$ est alors presque sûrement un ensemble de dimension de Hausdorff $\frac{\mathbb{E}[N]}{\log 3}$. En particulier, si on choisit à chaque fois 2 segments placés de façon quelconque, ou encore si on enlève chaque segment avec probabilité $1/3$, l'ensemble obtenu aura même dimension que l'ensemble de Cantor usuel.

Remarque. Si N est la variable aléatoire constante 2, on retrouve l'ensemble de Cantor classique si on prend les morceaux de gauche et de droite à chaque fois. Par ailleurs, si on prend les deux morceaux de gauche, on obtient l'ensemble de Cantor réduit d'un facteur 2.

Exemple 2. On peut de la même façon construire un analogue aléatoire du triangle de Sierpinski. Il aura presque toujours (au sens de la mesure de Galton-Watson) une dimension de $\frac{\log 3}{\log 2}$ du moment qu'on garde en moyenne 3 triangles, sur l'événement de non-vacuité.

Remarque. Le triangle de Sierpinski usuel correspond à une variable constante égale à 3, lorsqu'on enlève toujours le triangle central.

Exemple 3. On construit un ensemble de Cantor aléatoire en partant de $[0, 1]$, puis en remplaçant chaque segment récursivement par ses deux extrémités, non pas de longueur $1/3$, mais de longueur aléatoire. Ici, par exemple, on choisit cette longueur selon une loi uniforme sur $[1/4, 1/2]$.

Alors la dimension de Hausdorff est l'unique s tel que $\mathbb{E}[X^s + Y^s] = 1$ (ici c'est strictement décroissant), c'est à dire :

$$4 \int_{\frac{1}{4}}^{\frac{1}{2}} x^s ds = \frac{1}{2}.$$

Autrement dit, $\dim E = s \approx 0,703$, avec

$$\frac{1}{2^{s+1}} - \frac{1}{4^{s+1}} = \frac{s+1}{8}.$$

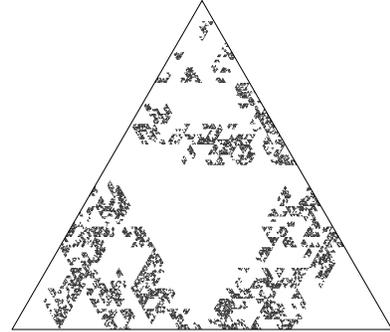


FIG. 14 – Triangle de Sierpinski aléatoire engendré par une loi binomiale de paramètres $(4, 3/4)$.



FIG. 15 – Ensemble de Cantor aux segments de longueur aléatoire.

4.2 Application au mouvement brownien

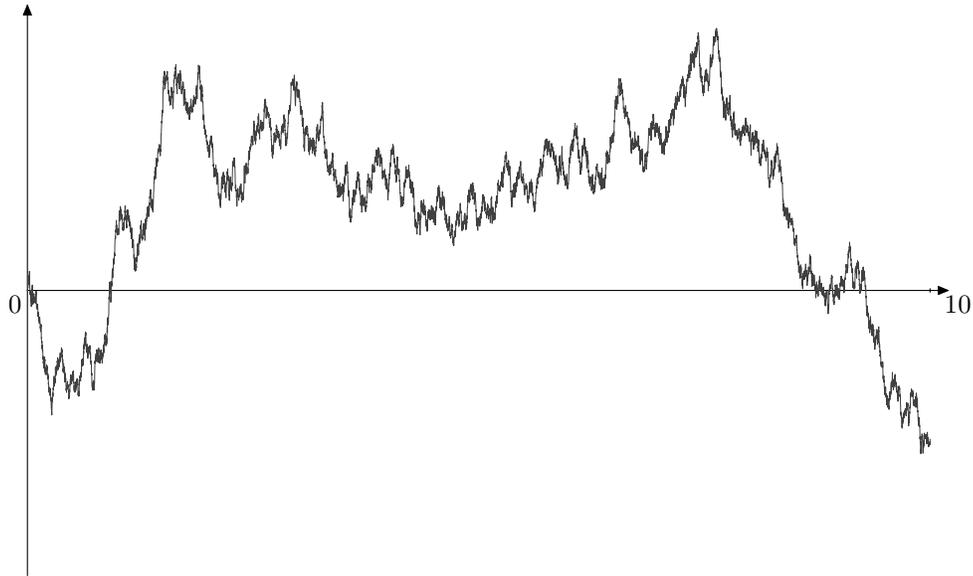


FIG. 16 – Une réalisation du mouvement brownien.

On peut également appliquer le théorème de Falconer à des cas plus surprenants. Ainsi, si on appelle E l'ensemble des zéros d'un mouvement brownien classique, alors $\dim E = 1/2$ p.s. Il convient d'énoncer quelques généralités sur les mouvements browniens :

Définition 4.1.

- Un *mouvement brownien* est une mesure de probabilité sur l'espace $\mathcal{C}_0(\mathbb{R}_+, \mathbb{R}^d)$ des fonctions continues s'annulant en zéro, tel que pour tous $t_1 < \dots < t_p$, la matrice de covariance du vecteur gaussien $(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}})$ soit $\text{diag}(t_1, t_2 - t_1, \dots, t_n - t_{n-1})$.
- Soit (B_t) un mouvement brownien. Le processus aléatoire défini par $U_t = B_t - t B_1$ est appelé *pont brownien*.
- Un *processus gaussien* est une famille (X_t) de variables aléatoires telles que pour toute famille finie (t_1, \dots, t_p) , le vecteur $(X_{t_1}, \dots, X_{t_p})$ soit un vecteur gaussien.

Proposition 4.1. Soit (X_t) un processus gaussien continu ($t \mapsto X_t$ est continue p.s.).

- (X_t) est un mouvement brownien si et seulement si $\forall s, t \in \mathbb{R}_+ \mathbb{E}[X_s X_t] = \min(s, t)$.
- (X_t) est un pont brownien si et seulement si $\forall s, t \in [0, 1] \mathbb{E}[X_s X_t] = \min(s, t) - st$.

Démonstration.

Soit (B_t) un mouvement brownien. Alors $\mathbb{E}[B_s B_t] = \mathbb{E}[\mathbb{E}[B_s B_t | B_s]] = \mathbb{E}[B_s B_s] = s$ si $s < t$. Réciproquement, si $\text{cov}(B_s, B_t) = s \wedge t$, cela fixe la matrice de covariance de chacun des vecteurs

gaussiens $(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}})$, donc leur loi. Par continuité, la loi de B est fixée par une quantité dénombrable de B_t , donc comme les accroissements finis ont même loi que pour un mouvement brownien, (B_t) a la loi d'un mouvement brownien.

Soit (U_t) un pont brownien. On vérifie sans peine que :

$$\mathbb{E}[(B_s - s B_1)(B_t - t B_1)] = s \wedge t - st - st + st = s \wedge t - st.$$

Réciproquement, si un processus gaussien continu vérifie cette propriété, il fixe la covariance, donc la loi du vecteur $(U_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}})$, et la loi de (U_t) est déterminée par ces marginales. \square

Proposition 4.2. Soit (B_t) un mouvement brownien, et (U_t) un pont brownien. Alors

- (kB_{t/k^2}) est un mouvement brownien (invariance d'échelle).
- $(B_{s+t} - B_s)$ est un mouvement brownien (invariance par translation).
- $(tB_{1/t})$ est un mouvement brownien (invariance par inversion temporelle).
- $((1+t)U_{\frac{t}{1+t}})$ est un mouvement brownien.
- U_{1-t} est un pont brownien.

Démonstration.

Il suffit de vérifier la propriété sur les covariances. Par exemple, si $\beta_t = tB_{1/t}$, β_t est gaussienne centrée et

$$\mathbb{E}[\beta_s \beta_t] = \mathbb{E}[stB_{1/s}B_{1/t}] = st \min\left(\frac{1}{s}, \frac{1}{t}\right) = \frac{st}{\max(s, t)} = \min(s, t).$$

On définit un mouvement brownien à partir d'un pont brownien car

$$\begin{aligned} \mathbb{E}\left[(1+s)U_{\frac{s}{1+s}}(1+t)U_{\frac{t}{1+t}}\right] &= (1+s)(1+t) \left(\min\left(\frac{t}{1+t}, \frac{s}{1+s}\right) - \frac{st}{(1+s)(1+t)} \right) \\ &= \min(t(1+s), s(1+t)) - st = \min(s, t). \end{aligned}$$

Enfin, le pont brownien a même loi que son symétrisé, puisque :

$$\begin{aligned} \mathbb{E}[U_{1-s}U_{1-t}] &= \min(1-s, 1-t) - (1-s)(1-t) \\ &= 1 - \max(s, t) - 1 + s + t - st = \min(s, t) - st. \end{aligned}$$

\square

Lemme 4.3. Soit (B_t) un mouvement brownien, et $T = \sup\{t \leq 1, B_t = 0\}$. Alors le processus $(T^{-1/2}B_{tT})$ définit un pont brownien.

Démonstration. Soit $\check{B}_t = tB_{1/t}$ le mouvement brownien renversé, et le temps d'arrêt $U = 1/T = \inf\{t \geq 1, \check{B}_t = 0\}$. Donc $\hat{B}_s = \check{B}_{U+s}$ est un mouvement brownien ($\hat{B}_U = 0$). Dans la suite \bar{B} et X désignent des mouvements browniens qui seront liés à \hat{B} comme indiqué. Alors :

$$\begin{aligned} \frac{1}{\sqrt{T}}B_{tT} &= \frac{1}{\sqrt{T}} \cdot tT \cdot \check{B}_{\frac{1}{tT}} = t\sqrt{T}\check{B}_{\frac{1}{tT}} \\ &= t\sqrt{T}\hat{B}_{\frac{1}{tT}-U} = \frac{t}{\sqrt{U}}\hat{B}_{sU} \text{ où } s = \frac{1}{t} - 1 \\ &= t\bar{B}_s \text{ en loi (invariance d'échelle), car } U \text{ et } \hat{B} \text{ sont indépendants} \\ &= t(X_{\frac{1}{t}} - X_1) \text{ en loi, car } s = \frac{1}{t} - 1 \text{ et par la loi de Markov forte pour } X \\ &= \check{X}_t - t\check{X}_1. \end{aligned}$$

Donc $\frac{1}{\sqrt{T}}B_{tT}$ est un pont brownien. \square

Le fait que l'ensemble des zéros du mouvement brownien est presque sûrement sans points isolés et de mesure nulle, est un résultat classique, qui découle directement du théorème de Fubini-Tonnelli. Il suffit d'écrire $\mathbb{E}[\int \mathbb{1}[B_t = 0] dt] = \int \mathbb{P}[B_t = 0] dt = 0$. On démontre ici un résultat plus précis.

Théorème 4.4 (Dimension de l'ensemble des zéros du mouvement brownien). *La dimension de Hausdorff de $E = \{t \in \mathbb{R}^+, B_t = 0\}$ est égale à $1/2$ presque sûrement.*

Démonstration. Soit t_0 tel que $B_{t_0} = 0$. Par invariance d'échelle, à la fois du mouvement Brownien et de la dimension de Hausdorff, il suffit d'étudier l'intervalle $[0, t_0]$. On choisit par exemple $t_0 = \max\{t \leq 1, B_t = 0\}$. On a alors $t_0 > 0$ p.s., et $U_t = t_0^{-1/2} B_{tt_0}$ est un pont brownien. L'ensemble des zéros d'un mouvement brownien est donc de même dimension que l'ensemble des zéros du pont brownien associé (c'est même presque le même à une homothétie près).

On va coder E de la manière suivante : on part de l'intervalle $E_0 = [0, 1]$, et on le coupe en deux, en posant

$$\begin{aligned} - a &= \max\{t < 1/2, B_t = 0\}, \\ - b &= \min\{t > 1/2, B_t = 0\}. \end{aligned}$$

On pose $E_1 = [0, a] \cup [b, 1]$ et ainsi de suite en recommençant sur les deux petits intervalles. Il est facile de voir que E est toujours contenu dans les E_n . Par ailleurs, on a presque sûrement $B_{k/2^n} \neq 0$ pour tous k et n , donc les composantes connexes de $\tilde{E} = \bigcap E_n$ sont de longueur nulle, et leurs extrémités sont dans E , d'où $E = \tilde{E}$ (puisque les composantes connexes de E_n sont de longueur inférieure à 2^{-n}).

On a ainsi codé E par un arbre binaire, avec des facteurs de taille $X = a$ et $Y = 1 - b$. Par invariance des ponts browniens par changement d'échelle, par translation, par retournement du temps, etc. la loi de X et de Y ne dépend pas de l'intervalle. Donc E est codé par un réseau de Galton-Watson de variable génératrice $(2, X, Y)$. D'après le théorème de Falconer, $\dim E$ est le réel (unique) s tel que $\mathbb{E}[X^s + Y^s] = 1$. On calcule la loi de X en écrivant :

$$\begin{aligned} b &= \inf\{t \in [1/2, 1], U_t = 0\} \\ &= \inf\{t \in [1/2, 1], B_t - tB_1 = 0\} && \text{car } U_t \text{ a même loi que } B_t - tB_1, \\ &= \inf\{t \in [1/2, 1], tB_{1/t} - tB_1 = 0\} && \text{par inversion temporelle,} \\ &= \inf\{t \in [1/2, 1], B_{1/t} - B_1 = 0\} \\ &= \left(\sup\{u \in [1, 2], B_u - B_1 = 0\}\right)^{-1} && \text{en posant } u = 1/t, \\ &= (1 + \sup\{u \in [0, 1], B_u = 0\})^{-1} && \text{en utilisant l'invariance par translation.} \end{aligned}$$

Soit $G = \sup\{u \in [0, 1], B_u = 0\}$. La densité de G est $\frac{1}{\sqrt{x(1-x)}}$, c'est la loi de l'arcsinus. En effet, par inversion temporelle, G a même loi que D^{-1} où $D = \inf\{u \geq 1, B_u = 0\}$. D est liée à B_1 , car on sait que pour $t > 1$ et $B_1 < 0$:

$$\begin{aligned} \mathbb{P}[D > t \mid B_1] &= \mathbb{P}\left[\sup_{1 < s < t} B_s < 0 \mid B_1\right] = \mathbb{P}\left[\sup_{s < t-1} B_s < -B_1 \mid B_1\right] \text{ par translation} \\ &= \mathbb{P}[|B_{t-1}| < |B_1| \mid B_1] \text{ car } \sup_{s < t} B_s \stackrel{\text{loi}}{=} |B_t| \\ &= \mathbb{P}[(t-1)N^2 < B_1^2 \mid B_1] \text{ où } N \text{ est de loi normale indépendante de } B_1 \\ &= \mathbb{P}\left[t < 1 + \frac{B_1^2}{N^2} \mid B_1\right] \end{aligned}$$

Donc D a même loi que $1 + \frac{B_1^2}{N^2}$.

$$\begin{aligned} \mathbb{E}[f(D)] &= \iint f\left(1 + \frac{x^2}{y^2}\right) \frac{1}{2\pi} \exp\left(-\frac{x^2 + y^2}{2}\right) dx dy \\ &= 4 \int_{u=1}^{\infty} \int_{y \geq 0} \frac{f(u)}{2\pi} \exp\left(-\frac{uy^2}{2}\right) \frac{y}{2\sqrt{u-1}} du dy \\ &\text{en posant } u = 1 + \frac{x^2}{y^2}, \quad du dy = \frac{2}{x} y^2 dx dy \end{aligned}$$

$$\begin{aligned}
&= \int_1^\infty \frac{f(u)}{\pi\sqrt{u-1}} du \int_0^\infty \exp\left(-\frac{uy^2}{2}\right) y dy \\
&= \int_1^\infty \frac{f(u)}{\pi\sqrt{u-1}} du \int_0^\infty \exp(-uv) dv \text{ où } v = \frac{y^2}{2} \\
&= \int_1^\infty \frac{f(u)}{\pi u\sqrt{u-1}} du.
\end{aligned}$$

On obtient finalement :

$$\begin{aligned}
P^D(dx) &= \frac{dx}{\pi x\sqrt{x-1}} \\
P^G(dy) &= \frac{y}{\pi\sqrt{1/y-1}} \frac{dy}{y^2} = \frac{dy}{\pi\sqrt{y(1-y)}} = \frac{d(\text{Arcsin}(2y-1))}{\pi}.
\end{aligned}$$

Après avoir utilisé toutes les invariances imaginables des mouvements browniens, et calculé des intégrales ignobles, voyons ce que vaut la dimension de E . C'est l'unique réel tel que $\mathbb{E}[X^s + Y^s] = 1$. Toujours par symétrie des ponts browniens, X et Y ont même loi (mais elles ne sont pas indépendantes). Il suffit donc de vérifier que $\mathbb{E}[\sqrt{Y}] = \frac{1}{2}$. On a vu que $Y = 1-b = 1 - \frac{1}{1+G} = \frac{G}{1+G}$.

$$\mathbb{E}[\sqrt{Y}] = \frac{1}{\pi} \int_0^1 \sqrt{\frac{G}{1+G}} \frac{dG}{\sqrt{G(1-G)}} = \frac{1}{\pi} \int_0^1 \frac{dG}{1-G^2} = \frac{1}{\pi} \text{Arcsin } 1 = \frac{1}{2}$$

Ce qui achève cette démonstration. □

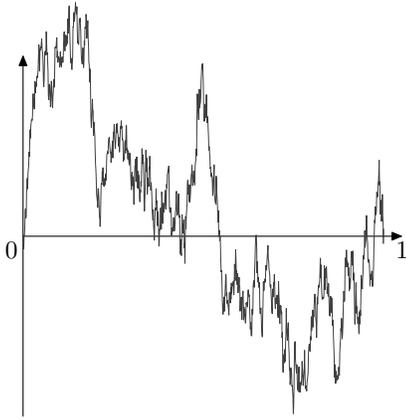


FIG. 17 – Une réalisation du pont brownien.

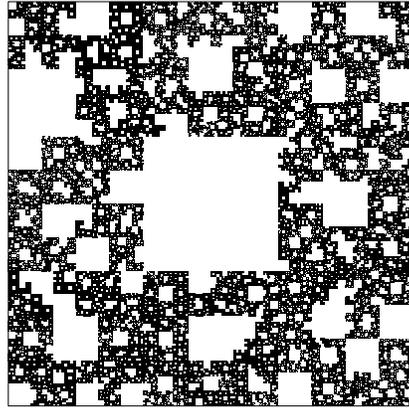


FIG. 18 – Fractale aléatoire obtenue en divisant chaque carré aléatoirement en 4, 9 ou 16 morceaux, et en enlevant respectivement 1, 1, et 4 morceaux. Sa dimension de Hausdorff vaut environ 1,7984.

Conclusion

On voit ainsi comment des concepts de natures apparemment différentes sont en fait très liés : l'étude des arbres apporte un regard nouveau sur l'étude des objets fractals, en permettant de calculer la dimension de Hausdorff d'objets à première vue complexes. Par ailleurs, elle permet également d'étudier des objets aléatoires, comme le pont brownien, dont la structure ne paraît pas simple au premier abord.

A Annexe : théorèmes asymptotiques pour les processus de Galton-Watson

A.1 Théorème de Kesten-Stigum

On se propose ici de démontrer entièrement le théorème de Kesten-Stigum. Commençons par établir deux lemmes indépendants.

Lemme A.1. Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires positives, indépendantes et de même loi.

$$\limsup \frac{X_n}{n} = \begin{cases} 0 & \text{si } \mathbb{E}[X] < \infty \\ \infty & \text{si } \mathbb{E}[X] = \infty \end{cases} \quad \text{p.s.}$$

Démonstration. Supposons dans un premier temps $\mathbb{E}[X] < \infty$. Fixons $\varepsilon > 0$, et notons A_n l'évènement $\{X_n \geq n\varepsilon\}$ et $r = \lceil \frac{1}{\varepsilon} \rceil$. Il suffit de remarquer¹² que :

$$\begin{aligned} \sum_{n=0}^{\infty} \mathbb{P}[A_n] &= \sum_{n=0}^{\infty} \mathbb{P}\left[\frac{X_0}{n} \geq \varepsilon\right] = \sum_{n=0}^{\infty} \mathbb{P}[X_0 \geq n\varepsilon] = \sum_{k=0}^{\infty} \sum_{k \leq n\varepsilon < k+1} \mathbb{P}[X_0 \geq n\varepsilon] \\ &\leq \underbrace{\sum_{k=0}^{\infty} r \mathbb{P}[X_0 \geq k]}_{=r \mathbb{E}[X] < \infty}. \end{aligned}$$

Le lemme de Borel-Cantelli permet de conclure : $\mathbb{P}[\limsup A_n] = \mathbb{P}[\limsup (X_n/n) \geq \varepsilon] = 0$. Faisant tendre ε vers 0, on en déduit la première partie du lemme.

Supposons maintenant $\mathbb{E}[X] = \infty$. Fixons $M > 0$, et notons A_n l'évènement $\{X_n \geq nM\}$. Les évènements A_n sont indépendants, et, notant $r = \lceil M \rceil$:

$$\begin{aligned} \sum_{n=0}^{\infty} \mathbb{P}[A_n] &= \sum_{n=0}^{\infty} \mathbb{P}\left[\frac{X_0}{n} \geq M\right] = \sum_{n=0}^{\infty} \mathbb{P}[X_0 \geq nM] \\ &\geq \frac{1}{r} \sum_{k=r}^{\infty} \mathbb{P}[X_0 \geq k] \geq \infty. \end{aligned}$$

De nouveau, le lemme de Borel-Cantelli permet d'affirmer que

$$\mathbb{P}[\limsup A_n] = \mathbb{P}[\limsup X_n/n \geq M] = 1.$$

Faisant tendre M vers $+\infty$, on en déduit que p.s., $\limsup X_n/n = \infty$. □

Lemme A.2.

Soient μ et ν deux mesures finies sur un espace (Ω, \mathcal{A}) , avec $\nu(\Omega) = 1$. On suppose que :

1. La tribu \mathcal{A} est l'union croissante de sous-tribus $(\mathcal{A}_n)_{n \in \mathbb{N}}$.
2. Les mesures restreintes $\mu_n = \mu|_{\mathcal{A}_n}$ sont absolument continues par rapport aux $\nu|_{\mathcal{A}_n}$.

Notons $X_n = \frac{d\mu_n}{d\nu_n}$ les dérivées de Radon-Nikodym, et $X = \limsup_n X_n$. On dispose des deux chaînes d'équivalence suivantes :

$$\begin{aligned} \mu \ll \nu &\iff X \text{ est finie } \mu\text{-p.s.} \iff \int X \cdot d\nu = \int d\mu \\ \text{et } \mu \perp \nu &\iff X = \infty \quad \mu\text{-p.s.} \iff \int X \cdot d\nu = 0. \end{aligned}$$

¹²Les séries qui suivent sont toutes à termes positifs, ce qui justifie pleinement les manipulations effectuées.

Démonstration. Considérons la suite de variables aléatoires $(X_n)_{n \in \mathbb{N}}$. Chaque variable est \mathcal{A}_n -mesurable, positive, et ν -intégrable. De plus, pour toute fonction g mesurable par rapport à \mathcal{A}_n :

$$\begin{aligned} \mathbb{E}[g X_{n+1}] &= \int_{w \in \Omega} g(w) X_{n+1}(w) \cdot d\nu(w) = \int_{w \in \Omega} g(w) X_{n+1}(w) \cdot d\nu|_{\mathcal{A}_{n+1}}(w) \\ &= \int_{w \in \Omega} g(w) \cdot d\mu|_{\mathcal{A}_{n+1}}(w) = \int_{w \in \Omega} g(w) \cdot d\mu|_{\mathcal{A}_n}(w) \\ &= \int_{w \in \Omega} g(w) X_n(w) \cdot d\nu|_{\mathcal{A}_n}(w) = \int_{w \in \Omega} g(w) X_n(w) d\nu|_{\mathcal{A}_n}(w) = \mathbb{E}[g X_n]. \end{aligned}$$

La propriété caractéristique de l'espérance conditionnelle permet de conclure que $\mathbb{E}[X_{n+1} | \mathcal{A}_n] = X_n$. Ainsi, $(X_n)_n$ est une martingale positive adaptée à la filtration $(\mathcal{A}_n)_{n \in \mathbb{N}}$. Elle converge donc ν -p.s. vers X , et, en particulier, X est ν -p.s. finie.

Montrons alors que pour toute partie mesurable $A \in \mathcal{A}$, l'on a :

$$\mu(A) = \int_A X \cdot d\nu + \mu(A \cap \{X = +\infty\}). \quad (5)$$

Supposons dans un premier temps $\mu \ll \nu$. Le théorème de Radon-Nikodym permet d'exhiber une variable aléatoire \mathcal{A} -mesurable Y telle que $d\mu = Y d\nu$. Alors, pour tout $n \in \mathbb{N}$, X_n est précisément l'espérance conditionnelle de Y par rapport à \mathcal{A}_n . En outre, $Y \in L^1(\Omega, \mathcal{A}, \nu)$. Par conséquent, X_n converge ν -p.s. et L^1 vers¹³ Y . Ainsi, ν -p.s., $X = Y$, et :

$$\int_A X \cdot d\nu = \int_A Y \cdot d\nu = \int_A d\mu = \mu(A).$$

Dans le cas général, posons $\rho = (\nu + \mu)/C$, avec $C = \int d\nu + \int d\mu$. Les mesures ν et μ sont toutes deux absolument continues par rapport à ρ . Définissons alors comme précédemment :

$$\left\{ \begin{array}{l} M_n = \frac{d\mu_n}{d\rho_n} \\ M = \limsup_n M_n \end{array} \right. \quad \text{et} \quad \left\{ \begin{array}{l} N_n = \frac{d\nu_n}{d\rho_n} \\ N = \limsup_n N_n \end{array} \right.$$

Les suites (M_n) et (N_n) convergent ρ -p.s. et L^1 vers M et N . Remarquons en outre que $M_n + N_n = C$ p.s., et donc que l'évènement $\{M = N = 0\}$ est de ρ -probabilité nulle. L'évènement $\{N \neq 0\}$ ne diffère ainsi de $\{X = \infty\}$ que d'un ensemble de ρ -mesure nulle. Sur cet évènement :

$$\frac{M}{N} = \frac{\lim M_n}{\lim N_n} = \lim \frac{M_n}{N_n} = \lim X_n = X \quad \rho\text{-p.s.}$$

Ainsi, si A est une partie mesurable :

$$\begin{aligned} \mu(A) &= \int_A M \cdot d\rho = \int_{A \cap \{N > 0\}} M \cdot d\rho + \int_{A \cap \{N = 0\}} M \cdot d\rho \\ &= \int_{A \cap \{N > 0\}} XN \cdot d\rho + \int_{A \cap \{X = \infty\}} M \cdot d\rho = \int_A XN \cdot d\rho + \int_{A \cap \{X = \infty\}} M \cdot d\rho \\ &= \int_A X \cdot d\nu + \mu(A \cap \{X = +\infty\}). \end{aligned}$$

Il suffit alors de se persuader que les deux chaînes d'équivalences sont des conséquences immédiates de la décomposition (5) démontrée ci-dessus. \square

Remarque. L'équation (5) n'est pas vérifiée dans le cas général — i.e., avec deux mesures finies μ et ν quelconques. En effet, on s'est ici servi du fait que les mesures μ_n étaient absolument continues par rapport aux mesures ν_n .

¹³En effet, on a bien $\mathcal{A} = \bigvee_{n \in \mathbb{N}} \mathcal{A}_n$ par hypothèse.

L'idée est alors d'appliquer le lemme A.2 à la loi $GW = GW_p$ des arbres aléatoires de Galton-Watson, et à une loi \widehat{GW} , correspondant à des arbres aléatoires "déviés en taille".

Commençons par introduire quelques nouvelles notations. Soit T un arbre quelconque — c'est à dire, via l'identification canonique, un élément de $\text{im } \psi$ — et v un sommet de T de hauteur n . On note :

- $Z_n(T)$ le nombre de sommets de hauteur n dans T .
- $W_n(T)$ la quantité $Z_n(T)/m^n$.
- $[T]_n$ l'ensemble des arbres dont les n premiers étages correspondent à ceux de T .
- $[T, v]_n$ l'ensemble des couples (S, C) tels que $S \in [T]_n$ et C est un chemin sur S partant de la racine et passant par v .

Si L est la loi de reproduction d'un processus de Galton-Watson, introduisons également une variable aléatoire \hat{L} définie par :

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

Définition A.1 (Processus et arbres de Galton-Watson avec immigration). Soient $(L_{n,k})_{(n,k) \in \mathbb{N}^2}$ et $(\hat{L}_n)_{n \in \mathbb{N}^*}$ des copies indépendantes de L et \hat{L} . On appelle *processus de branchement de Galton-Watson avec immigration* la donnée de deux suites de variables aléatoires $(Z_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ vérifiant les relations suivantes :

$$\left\{ \begin{array}{l} v_n = X_{Z_n}, \quad \text{avec } X_m \text{ variable aléatoire équadistribuée dans } \llbracket 1, m \rrbracket \\ Z_0 = 1 \\ Z_{n+1} = \hat{L}_n + \sum_{i \in \llbracket 1, Z_n \rrbracket} L_{n,i} \end{array} \right. \quad (6)$$

Un tel processus peut s'interpréter intuitivement de la façon suivante. Partant d'un individu v_0 , on lui attribue un nombre aléatoire de descendants \hat{L}_0 . L'un des descendants v_1 de première génération, choisi aléatoirement et de façon équiprobable, donne ensuite naissance à \hat{L}_1 descendants, et les $\hat{L}_0 - 1$ autres à une famille aléatoire qui suit la loi de reproduction habituelle p . De même, l'un des descendants v_2 de seconde génération engendre \hat{L}_2 descendants, et les $\hat{L}_1 - 1$ autres des familles aléatoires qui suivent la loi de reproduction habituelle p . On réitère le processus indéfiniment¹⁴, et le résultat $((Z_n)_{n \in \mathbb{N}}, (v_n)_{n \in \mathbb{N}})$ s'identifie à la donnée d'un arbre T et d'un chemin infini C tracé sur cet arbre. La variable aléatoire $Z_n - 1$ d'un processus avec immigration peut être interprétée comme la taille de la n -ième génération d'une population dont les membres se reproduisent selon une loi p , et qui accueille de nouveaux membres se reproduisant selon une loi \hat{p} .

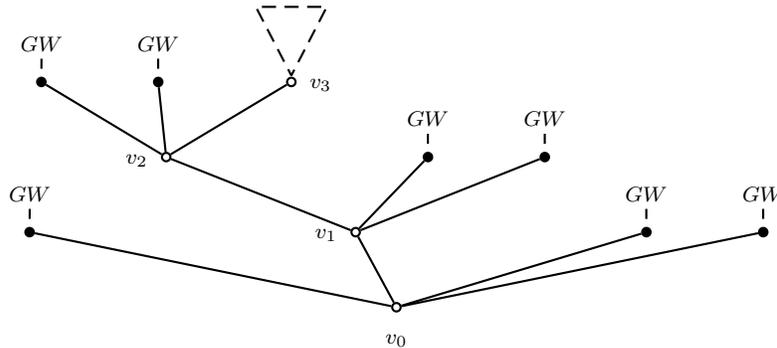


FIG. 19 – Un exemple d'arbre de Galton-Watson avec immigration.

Comme précédemment, il est possible de construire des arbres aléatoires dans $\mathfrak{B}(\mathcal{F})$ qui correspondent à ces processus — i.e. dont les tailles des générations sont des processus de Galton-Watson avec immigration. Il suffit en effet d'utiliser le codage canonique et d'adapter la proposition 1.3.

¹⁴Remarquons que $\hat{L} \geq 1$, et donc que l'on est bien en mesure de poursuivre indéfiniment l'expérience.

Ceci permet de munir l'ensemble des couples (arbre infini S , chemin infini C tracé sur S) d'une loi \widehat{GW}^* , et l'ensemble $\mathfrak{P}(\mathcal{F})$ des arbres de la loi \widehat{GW} correspondant à la projection de \widehat{GW}^* sur les arbres.

Remarque. La loi \widehat{GW}^* décrit donc les arbres généalogiques d'une population accueillant $Y_n = \hat{L}_{n-1} - 1$ immigrés à la n -ième génération, et se reproduisant selon la loi p habituelle.

Proposition A.3. Si k désigne le nombre de descendants de la racine d'un arbre T , $T^{(j)}$ les k sous-arbres issus des descendants, et w un sommet de T de hauteur $n+1$ appartenant au sous-arbre $T^{(i)}$, alors on dispose des égalités suivantes :

$$\widehat{GW}^*([T, w]_{n+1}) = \frac{pk}{m} \widehat{GW}^*([T^{(i)}, w]_n) \prod_{j \neq i} GW([T^{(j)}]_n) \quad (7)$$

$$\text{et } \widehat{GW}^*([T, v]_n) = \frac{1}{m^n} GW([T]_n). \quad (8)$$

Démonstration. Commençons par démontrer (7). L'indépendance des développements des sous-arbres d'un arbre de Galton-Watson permet d'écrire :

$$\begin{aligned} \widehat{GW}^*([T, w]_{n+1}) &= \hat{p}_k && \left\{ \begin{array}{l} \text{probabilité que la racine d'un arbre de Galton-} \\ \text{Watson avec immigration ait } k \text{ descendants} \end{array} \right. \\ &\times \mathbb{P}[w_1 = v_1 \mid Z_1 = k] && \left\{ \begin{array}{l} \text{probabilité que le premier sommet choisi} \\ \text{pour l'immigration soit } v_1, \text{ sachant } Z_1 \end{array} \right. \\ &\times \widehat{GW}^*([T^{(i)}, w]_n) && \left\{ \begin{array}{l} \text{probabilité d'obtenir le sous-arbre } T^{(i)} \text{ par un} \\ \text{processus de Galton-Watson avec immigration} \end{array} \right. \\ &\times \prod_{j \neq i} GW([T^{(j)}]_n) && \left\{ \begin{array}{l} \text{probabilité d'obtenir les autres sous-arbres} \\ \text{par un processus de Galton-Watson classique} \end{array} \right. \\ &= \frac{k p_k}{m} \frac{1}{k} \widehat{GW}^*([T^{(i)}, w]_n) \prod_{j \neq i} GW([T^{(j)}]_n) \\ &= \frac{pk}{m} \widehat{GW}^*([T^{(i)}, w]_n) \prod_{j \neq i} GW([T^{(j)}]_n). \end{aligned}$$

L'équation (8) s'en déduit par récurrence sur n . Si le résultat est établi jusqu'au rang $n-1 \geq 1$, alors :

$$\begin{aligned} \widehat{GW}^*([T, v]_n) &= \frac{pk}{m} \widehat{GW}^*([T^{(i)}, v]_{n-1}) \prod_{j \neq i} GW([T^{(j)}]_{n-1}) \\ &= \frac{pk}{m^n} \prod_{j \in [1, k]} GW([T^{(j)}]_{n-1}) = \frac{1}{m^n} GW([T]_n). \end{aligned}$$

Les deux égalités sont donc bien démontrées. \square

Ces résultats sur les arbres aléatoires avec immigration nous permettent¹⁵ maintenant de démontrer le théorème de Kesten-Stigum :

Théorème A.4 (Seneta, 1970). *Supposons l'espérance m strictement supérieure à 1.*

1. Si $\mathbb{E}[\log^+ Y] < \infty$, la limite W de W_n existe, et est finie presque sûrement.
2. Si $\mathbb{E}[\log^+ Y] = \infty$, alors pour toute constante $c > 0$, $\limsup_n Z_n/c^n = +\infty$ presque sûrement.

¹⁵Il est tout à fait stupéfiant que l'étude d'arbres aléatoires *avec processus d'immigration* donne des résultats sur les arbres de Galton-Watson *sans processus d'immigration*. Une preuve — mathématique! — que l'on a toujours à apprendre de l'autre...

Démonstration. Supposons dans un premier temps $\mathbb{E}[\log^+ Y] < \infty$. Notons $\mathcal{B} = \sigma((Y_n)_{n \geq 1})$, et $Z_{n,k}$ le nombre d'individus de la génération n qui ont immigré à la génération k . On a évidemment $Z_n = \sum_{k=1}^n Z_{n,k}$ — tous les individus ont immigré à un moment $k \leq n$, et ce une seule fois — et :

$$\mathbb{E} \left[\frac{Z_n}{m^n} \middle| \mathcal{B} \right] = \mathbb{E} \left[\frac{1}{m^n} \sum_{k=1}^n Z_{n,k} \middle| \mathcal{B} \right] = \sum_{k=1}^n \frac{1}{m_k} \mathbb{E} \left[\frac{Z_{n,k}}{m^{n-k}} \middle| \mathcal{B} \right].$$

Or, $Z_{n,k}/m^{n-k}$ correspond à la martingale à l'ordre $n-k$ habituellement associée à un processus de Galton-Watson, à ceci près que le nombre d'individus initial est Y_k , qui est précisément une variable \mathcal{B} -mesurable. On en déduit ¹⁶ que $\mathbb{E} [Z_{n,k}/m^{n-k} | \mathcal{B}] = Y_k$, et que :

$$\mathbb{E} \left[\frac{Z_n}{m^n} \middle| \mathcal{B} \right] = \sum_{k=1}^n \frac{Y_k}{m^k}.$$

L'hypothèse $\mathbb{E}[\log^+ Y] < \infty$ implique, d'après le lemme A.1 :

$$\begin{aligned} \limsup \frac{\log^+ Y_n}{n} &= 0, \\ \limsup (Y_n)^{\frac{1}{n}} &\leq 1. \end{aligned}$$

Comme $m > 1$, on en déduit que la série entière $\sum_{k=1}^{\infty} \frac{Y_k}{m^k}$ converge p.s. vers une limite finie. Par le lemme de Fatou, W est intégrable, finie presque sûrement.

Maintenant, si $\mathbb{E}[\log^+ Y] = \infty$, le lemme A.1 permet d'affirmer $\limsup Y_n/c^n = \infty$ presque sûrement. Mais l'on a bien sûr $Z_n \geq Y_n$, donc $\limsup_n Z_n/c^n = +\infty$ presque sûrement. \square

Théorème A.5 (Kesten-Stigum, 1966). *Supposons l'espérance m strictement supérieure à 1. Les assertions suivantes sont équivalentes :*

1. *La limite W de la martingale $(W_n)_{n \in \mathbb{N}}$ est strictement positive p.s. sur l'événement de non-extinction.*
2. *L'espérance de W est égale¹⁷ à 1.*
3. *La variable aléatoire $L \log^+ L$ est intégrable¹⁸ $\mathbb{E}[L \log^+ L] < \infty$.*

Démonstration du théorème de Kesten-Stigum.

Soit \mathcal{A}_n la tribu engendrée par les n premiers étages de l'arbre aléatoire. D'après la proposition A.3, et compte tenu de la définition de \widehat{GW} en fonction de \widehat{GW}^* :

$$\begin{aligned} \widehat{GW}^*([T, v]_n) &= \frac{1}{m^n} GW([T]_n) \\ \widehat{GW}([T]_n) &= Z_n \widehat{GW}^*([T, v]_n) \\ \text{d'où : } \widehat{GW}([T]_n) &= \frac{Z_n(T)}{m^n} GW([T]_n) = W_n(T) GW([T]_n). \end{aligned}$$

Ceci revient à dire que :

$$\frac{d\widehat{GW}|_{\mathcal{A}_n}}{dGW|_{\mathcal{A}_n}} = W_n(T).$$

Le lemme A.2 s'applique donc à l'espace des arbres muni des deux mesures de probabilité GW et \widehat{GW} :

$$\widehat{GW} \ll GW \iff W = \limsup_n W_n \text{ est finie } \widehat{GW}\text{-p.s.} \iff \int W \cdot dGW = 1 \quad (9)$$

$$\text{et } \widehat{GW} \perp GW \iff W = \limsup_n W_n = \infty \widehat{GW}\text{-p.s.} \iff W = 0 \text{ } GW\text{-p.s.} \quad (10)$$

¹⁶Car le développement de la population après immigration est indépendant de \mathcal{B} .

¹⁷C'est le résultat auquel on pourrait s'attendre : $\forall n \in \mathbb{N}$, $\mathbb{E}[W_n] = 1$.

¹⁸Cette condition est bien sûr toujours vérifiée dans le cas de « vrais » arbres généalogiques : L est bornée.

Par construction, $\mathbb{E}[L \log^+ L] = \mathbb{E}[\log^+ \widehat{L}] / m$ est de même nature que $\mathbb{E}[\log^+ Y]$. Par ailleurs, rappelons que $P_{se} \in \{P_e, 1\}$.

La chaîne d'équivalences se démontre alors sans difficulté :

- [3 \Rightarrow 2] : Si $\mathbb{E}[L \log^+ L] < +\infty$, le théorème de Seneta montre que W est finie \widehat{GW} -p.s., et donc $\mathbb{E}[W] = 1$ d'après l'équation (9).
- [2 \Rightarrow 1] : La contraposée est évidente, puisque si $P_{se} \neq P_e$, $P_{se} = 1$ et W est nulle GW -p.s., donc $\mathbb{E}[W] = 0 \neq 1$.
- [1 \Rightarrow 3] : De nouveau par contraposée, alors si $\mathbb{E}[L \log^+ L] = +\infty$, en vertu du théorème A.4, W est infinie \widehat{GW} -p.s., donc W est nulle GW -p.s., i.e. $P_{se} = 1$.

Ce qui achève la démonstration du théorème. \square

A.2 Théorème de Heathcote-Seneta-Vere-Jones

La démonstration du théorème de Heathcote-Seneta-Vere-Jones suit essentiellement le même schéma que celle du théorème de Kesten-Stigum : on commence par démontrer un lemme relatif aux mesures déviées, puis l'on utilise les résultats concernant les arbres aléatoires avec immigration.

Lemme A.6. Soit $(\mu_n)_{n \in \mathbb{N}}$ une suite de mesures de probabilités sur \mathbb{N} admettant des moments d'ordre 1, et $(c_n = \mathbb{E}_{\mu_n}[X])_{n \in \mathbb{N}}$ les espérances associées. On note $\hat{\mu}_n$ la mesure déviée de μ_n , définie par :

$$\hat{\mu}_n(k) = \frac{k \mu_n(k)}{c_n}.$$

1. Si $\lim_{M \rightarrow \infty} (\sup_{n \in \mathbb{N}} \hat{\mu}_n(\llbracket M, \infty \rrbracket)) = 0$, c'est-à-dire, si la famille de mesures $(\hat{\mu}_n)_{n \in \mathbb{N}}$ est étroite, alors $(c_n)_{n \in \mathbb{N}}$ est bornée.
2. Si, pour tout $M \in \mathbb{N}$, $\hat{\mu}_n(\llbracket 0, M \rrbracket) \rightarrow 0$, alors $c_n \rightarrow \infty$.

Démonstration. 1. Fixons M tel que $\sup_{n \in \mathbb{N}} \hat{\mu}_n(\llbracket M, \infty \rrbracket) \leq 1/2$. On dispose de la majoration suivante :

$$c_n = \sum_{k=0}^{\infty} \mu_n(k) k \leq c_n \sum_{k=M}^{\infty} \hat{\mu}_n(k) + \sum_{k=0}^{M-1} k \leq \frac{c_n + M^2}{2}.$$

Donc $\forall n \in \mathbb{N}$, $c_n \leq M^2$, et les moments d'ordre 1 sont bornés.

2. Raisonnons par l'absurde en supposant que $(c_n)_{n \in \mathbb{N}}$ ne tend pas vers ∞ . à extraction près, on peut alors supposer qu'il existe une borne M uniforme des espérances c_n . Remarquons maintenant que :

$$\forall k \in \mathbb{N}, \frac{\mu_n(k) k}{c_n} \rightarrow 0.$$

Comme les espérances $(c_n)_{n \in \mathbb{N}}$ sont bornées, on en déduit que :

$$\forall k \in \mathbb{N}, \mu_n(k) \rightarrow 0.$$

En particulier, $\mu_n(\llbracket 0, 2M \rrbracket) \rightarrow 0$. On aboutit alors à une contradiction :

$$1 \geq \sum_{k=2M+1}^{\infty} \frac{\mu_n(k) k}{c_n} \geq \sum_{k=2M+1}^{\infty} \frac{\mu_n(k) k}{M} \geq 2\mu(\llbracket 2M+1, \infty \rrbracket) \rightarrow 2.$$

Donc les espérances $(c_n)_{n \in \mathbb{N}}$ tendent vers l'infini. \square

Théorème A.7 (Heathcote, 1966). Soit $(Z_n)_{n \in \mathbb{N}}$ la suite des générations d'un processus de Galton-Watson avec immigration. On suppose l'espérance m strictement inférieure à 1.

1. Si $\mathbb{E}[\log^+ Y] < \infty$, alors Z_n converge en loi vers une variable aléatoire Z p.s. finie.
2. Si $\mathbb{E}[\log^+ Y] = \infty$, alors Z_n converge en probabilité vers ∞ .

Démonstration. Considérons comme dans la preuve du théorème A.4 les nombres $Z_{n,k}$ d'individus de la génération n dont les parents ont immigré à la génération k . On a toujours $Z_n = \sum_{k=1}^n Z_{n,k}$, mais, en outre, il convient de noter que la loi de $Z_{n,k}$ ne dépend que de la différence $n - k$. Par conséquent, Z_n a même loi que n'importe quelle somme de variables $Z_{p,k}$ telle que les différences $p - k$ parcourent une fois l'intervalle $\llbracket 1, n \rrbracket$. En particulier :

$$\forall n \in \mathbb{N}, Z_n \text{ et } \sum_{k=1}^m Z_{2k-1,k} \text{ ont même loi.}$$

Maintenant, le point-clef est que la somme $\sum_{k=1}^m Z_{2k-1,k}$ est croissante avec n , et donc converge p.s. vers une certaine variable aléatoire Z . A fortiori, $\sum_{k=1}^m Z_{2k-1,k} \rightarrow Z$, donc $Z_n \rightarrow Z$.

Notons maintenant que l'événement $Z = \infty$ appartient à la tribu asymptotique¹⁹, donc par la loi du tout ou rien de Kolmogorov, est presque sûr ou n'arrive presque jamais. Le théorème de Heathcote sera donc établi si l'on parvient à démontrer l'équivalence :

$$Z < \infty \iff \mathbb{E} [\log^+ Y] < \infty.$$

Raisonnons par double implication. Si $\mathbb{E} [\log^+ Y] < \infty$, alors $\limsup \log^+ Y_n/n = 0$ d'après le lemme A.1, i.e. Y_n est presque sûrement sous-exponentielle en n . Notant comme précédemment \mathcal{B} la tribu engendrée par les variables Y_n , le théorème de convergence monotone nous permet d'écrire :

$$\mathbb{E} [Z | \mathcal{B}] = \sum_{k=1}^{\infty} \mathbb{E} [Z_{2k-1,k} | \mathcal{B}] = \sum_{k=1}^{\infty} \mathbb{E} [Z_{2k-1,k} | \sigma(Y_1, Y_2, \dots, Y_k)] = \sum_{k=1}^{\infty} m^{k-1} Y_k < \infty \text{ p.s.}$$

Le lemme A.1 peut en outre se réécrire de la façon suivante :

$$\begin{cases} \mathbb{E} [X] < \infty & \iff \forall c \in]0, 1[, \sum_n c^n \exp(X_n) \text{ est finie p.s.} \\ \mathbb{E} [X] = \infty & \iff \forall c \in]0, 1[, \sum_n c^n \exp(X_n) \text{ est infinie p.s.} \end{cases}$$

On en déduit que Z est finie presque sûrement. Réciproquement, supposons $Z < \infty$ p.s., et écrivons $Z_{2k-1,k} = \sum_{i=1}^{Y_k} V_{k,i}$, les variables $V_{k,i}$ étant indépendantes et désignant la taille des $(k-1)$ -ièmes générations d'arbres de Galton-Watson issus de Y_k parents.

La somme de variables indépendantes $Z = \sum_{k=1}^{\infty} Z_{2k-1,k} = \sum_{k=1}^{\infty} \sum_{i=1}^{Y_k} V_{k,i}$ est par hypothèse finie p.s., donc les variables $V_{k,i}$ qui sont non nulles sont presque sûrement en nombre fini. Le lemme de Borel-Cantelli permet de conclure

$$\sum_{k=1}^{\infty} Y_k \text{ GW}[Z_{k-1} > 0] \text{ est fini p.s.}$$

A fortiori, $\sum_{k=1}^{\infty} Y_k (p_0)^k$ est fini, puisque $\text{GW}[Z_{k-1} > 0] \geq \mathbb{P}[L > 0]^{k-1} = p_0^{k-1}$. Si $p_0 < 1$, le lemme A.1 permet de conclure que $\mathbb{E} [\log^+ Y] < \infty$. Si $p_0 = 1$, c'est encore plus clair. \square

Le théorème de Heathcote-Seneta-Vere-Jones s'en déduit :

Théorème A.8. *Supposons l'espérance m du nombre de descendants strictement inférieure à 1 — cas sous-critique. La suite $(S_n = K_n/m^n)_{n \in \mathbb{N}}$ est décroissante, et les assertions suivantes sont équivalentes :*

1. *La limite de $(S_n = K_n/m^n)_{n \in \mathbb{N}}$ est strictement positive.*
2. *La suite $(J_n)_{n \in \mathbb{N}}$ est bornée.*
3. *La variable aléatoire $L \log^+ L$ est intégrable.*

Démonstration. Soit μ_n la loi de Z_n conditionnée par l'événement $Z_n > 0$. Pour tout arbre T tel que $Z_n(T) > 0$, notons $k(n)$ l'indice de l'individu de première génération le plus à gauche de l'arbre qui a au moins un descendant dans la n -ième génération.

¹⁹La tribu asymptotique associée aux familles de variables aléatoires $(L_{n,k})_{(n,k) \in \mathbb{N}^2}$ et $(\hat{L}_n)_{n \in \mathbb{N}}$.

