

Quelques contributions aux probabilités et à la théorie ergodique

Document de synthèse présenté pour
l'Habilitation à Diriger des Recherches

par
Élise JANVRESSE

soutenue à l'université de Rouen le 17 février 2012
Discipline : Mathématiques

Rapporteurs

Yves Derriennic	Professeur émérite, Université de Bretagne Occidentale
Emmanuel Lesigne	Professeur, Université François Rabelais
Karl Petersen	Professeur, University of North Carolina, Chapel Hill (USA)

Composition du Jury

Claude Dellacherie	Directeur de Recherches émérite, CNRS - Université de Rouen
Yves Derriennic	Professeur émérite, Université de Bretagne Occidentale
Fabien Durand	Professeur, Université de Picardie Jules Verne
Emmanuel Lesigne	Professeur, Université François Rabelais
Dalibor Volný	Professeur, Université de Rouen.

Avant-Propos

Le présent document présente une synthèse des travaux réalisés depuis ma thèse de doctorat, ainsi qu'un ensemble de pistes de recherches qui leur sont liées. Certains s'inscrivent dans la suite logique de ceux effectués pendant ma thèse qui était consacrée à l'étude des comportements asymptotiques de systèmes de particules, et donc à des problèmes issus de la physique. D'autres, plus récents, concernent la théorie ergodique et sont parfois liés à la théorie des nombres.

Le premier chapitre s'intéresse aux suites de Fibonacci aléatoires, qui font l'objet des articles [53, 52, 39] et [54] (soumis pour publication). Les suites de Fibonacci aléatoires et leurs multiples généralisations constituent un sujet d'investigation que les travaux existants n'ont pas épuisé. Nous nous sommes bien sûr d'abord intéressé aux facteurs de croissance (presque sûr et en moyenne) des suites de Fibonacci aléatoires. L'une des ramifications de ce sujet concerne les fractions continues de Rosen, une généralisation naturelle des fractions continues. Les structures combinatoires qui apparaissent lors de l'étude des suites de Fibonacci aléatoires montrent une parenté forte entre les deux objets. Comprendre l'un doit permettre de mieux comprendre l'autre, c'est pourquoi nous avons étudié une dynamique des fractions continues de Rosen.

Le chapitre 2 concerne la loi de Benford et fait la synthèse des articles [43, 38, 47]. Nous y proposons une interprétation probabiliste de cette loi, expliquant son apparition dans les nombres rencontrés dans la vie quotidienne. Cela nous permet notamment de retrouver et généraliser des résultats plus anciens de B. J. Flehinger sur la convergence vers la loi de Benford, pour lesquels nous déterminons la vitesse de convergence. Notre travail nous a aussi amenés à étudier diverses notions de densités pour des sous-ensembles infinis d'entiers naturels, et à donner une interprétation probabiliste de certaines d'entre elles.

Dans le chapitre 3, nous présentons quelques résultats de théorie ergodique. Les premiers concernent la transformation Pascal-adique [44, 49], dont, malgré la simplicité de la construction, très peu de propriétés sont connues. L'article [45] sur un travail sur les autocouplages 2 à 2 indépendants est résumé dans la Section 3.2. Enfin, nous mentionnons nos travaux [51, 46] sur les suspensions de Poisson et les différentes notions d'entropie pour des systèmes dynamiques en mesure infinie.

Nous avons regroupé dans le chapitre 4 des travaux portant sur quelques problèmes issus de la physique. Les travaux [40, 41, 42] s'intéressent à la convergence de marches aléatoires sur le groupe orthogonal $SO(n)$ et sur la sphère (cette dynamique a notamment été utilisée par Marc Kac pour obtenir un analogue de l'équation de Boltzmann spatialement homogène). Dans [48], nous considérons un modèle simplifié de l'interaction d'un polymère dirigé avec un potentiel dilué attractif et nous caractérisons les environnements qui donnent lieu à accrochage. Dans [50], nous proposons une variante de l'algorithme de *shuffling* adaptée aux situations avec arêtes de poids nul, ce qui permet de générer des pavages aléatoires d'une large classe de sous-graphes pondérés du réseau \mathbb{Z}^2 et d'autres graphes réguliers; cela fait apparaître des exemples de modèles dans lesquels les effets de bord se propagent à une grande distance de la frontière du domaine et sont donc visibles à l'échelle macroscopique. La dernière partie du chapitre 4 donne un très bref aperçu de travaux plus appliqués, [74, 75, 76] et [77] (soumis), portant sur le traitement d'images cérébrales obtenues par SPECT (Single-Photon Emission Computed Tomography).

Remerciements

Je voudrais tout d'abord adresser mes sincères remerciements à Yves Derriennic, Emmanuel Lesigne et Karl Petersen, pour avoir accepté d'être rapporteurs du présent travail. Qu'ils soient assurés de ma plus profonde reconnaissance pour l'attention qu'ils ont portée à ce manuscrit et pour le temps qu'ils ont consacré à son évaluation.

Mes remerciements vont aussi, bien sûr, aux membres du jury de la soutenance de cette habilitation, Claude Dellacherie, Fabien Durand et Dalibor Volný. Ils me font un grand honneur en acceptant de participer à ce jury.

Ces recherches doivent énormément à la rencontre enrichissante d'autres mathématiciens. Sans donner la liste exhaustive de ceux auprès de qui j'ai appris quelque chose, je citerai simplement Thierry de la Rue, Benoît Rittaud et Emmanuel Roy. Je leur exprime ici, outre ma plus sincère amitié, toute ma gratitude pour les collaborations fructueuses que nous avons entretenues et qui, je le souhaite, ne s'arrêteront pas là.

Un grand merci enfin aux membres du Département de Mathématiques de l'Université de Rouen et du Laboratoire de Mathématiques Raphaël Salem pour l'ambiance agréable qui règne sur notre lieu de travail.

Liste de publications

- **First order correction for the hydrodynamic limit of symmetric simple exclusion processes with speed change in dimension $d \geq 3$.**
É. Janvresse.
Annals of Probability, Vol 26 - N°4 (1998), pp 1874–1912
- **Relaxation to equilibrium of conservative dynamics I : Zero range processes.**
É. Janvresse, C. Landim, J. Quastel et H. T. Yau.
Annals of Probability, Vol 27 - N°1 (1999), pp 325–360
- **Spectral gap for Kac’s model of Boltzmann equation.**
É. Janvresse.
Annals of Probability, Vol 29 - N°1 (2001), pp 288–304
- **Note about ”First order correction for the hydrodynamic limit of symmetric simple exclusion processes with speed change in dimension $d \geq 3$ ”.**
É. Janvresse.
Annals of Probability, Vol 30 - N°2 (2002), pp 1–2
- **Approach to Equilibrium for Kac’s Master Equation.**
É. Janvresse.
Markov Processes and Related Fields, Vol 8 - N°2 (2002), pp 221–232.
- **Bounds on Semigroups of Random Rotations on $SO(n)$.**
É. Janvresse.
Theory of Probability and Its Applications, Vol 47 - N°3 (2003), pp. 526-532
- **From Uniform Distributions to Benford’s Law.**
É. Janvresse et T. de la Rue.
Journal of Applied Probability, 41 (2004), pp. 1203-1210.
- **The Pascal Adic Transformation is Loosely Bernoulli.**
É. Janvresse et T. de la Rue.
Annales de l’I.H.P., Probab. Stat., Vol 40 (2004), pp. 133-139.
- **A new method (the random walk) for heterogeneity quantification of brain spect perfusion images.**
R. Modzelewski, T. de la Rue, A. Hitzel, O. Benois, I. Gardin, É. Janvresse, A. Manrique, G. Grancher, R. Fernández, P. Vera.
EANM’04 - Annual Congress of the European Association of Nuclear Medicine (Helsinki), 2004.
- **Self-Similarity in the Corrections to the Ergodic Theorem for the Pascal-Adic Transformation.**
É. Janvresse, T. de la Rue et Y. Velenik.
Stochastics and Dynamics 5 No 1 (2005). p. 1-25.
- **Pinning by a sparse potential.**
É. Janvresse, T. de la Rue et Y. Velenik
Stoch. Proc. Appl. 115 No 8, (2005), pp 1323-1331.
- **A note on domino shuffling.**
É. Janvresse, T. de la Rue et Y. Velenik
Electron. J. Combin. 13 (1), (2006).
- **How do random Fibonacci sequences grow ?**

- É. Janvresse, B. Rittaud et T. de la Rue
Probability Theory and Related Fields 142 No 3-4, (2008), pp 619-648.
- **A Class of pairwise-independent Joinings.**
 É. Janvresse et T. de la Rue
Ergodic Theory and Dynamical Systems 28 No 5, (2008), pp 1545-1557.
 - **Development and validation of the Random Walk Algorithm : Application to the classification of diffuse heterogeneity in brain Spect perfusion images.**
 R. Modzelewski, T. de la Rue, É. Janvresse, A. Hitzel, J.-F. Ménard, I. Gardin, E. Gerardin, A. Manrique, D. Hannequin, S. Lepretre, H. Tilly et P. Vera
Journal of Computer Assisted Tomography, 32 (4) (July/August 2008), pp 651-659.
 - **Brain perfusion heterogeneity measurement based on random walk algorithm : Choice and influence of inner parameters.**
 R. Modzelewski, T. de la Rue, É. Janvresse, et P. Vera
Comput Med Imaging Graph., doi :10.1016/j.compmedimag.2009.11.006., (2009).
 - **Growth rate for the expected value of a generalized random Fibonacci sequence.**
 É. Janvresse, B. Rittaud et T. de la Rue
J. Phys. A : Math. Theor. 085005. 42 (2009).
 - **Poisson suspensions and Entropy for infinite transformations.**
 É. Janvresse, T. Meyerovitch, E. Roy et T. de la Rue
Transactions of the AMS, 362 No 6, (2010), pp 3069-3094. *Arxiv 0801.3155*
 - **Almost-sure growth rate of generalized random Fibonacci sequences.**
 É. Janvresse, B. Rittaud et T. de la Rue
Annales de l'IHP, Probab. Stat., 46 No 1, (2010) pp. 135-158.
 - **A unifying probabilistic interpretation of Benford's Law.**
 R. Giuliano, É. Janvresse
Uniform Distribution Theory., 5 No 2, (2010), pp 169-182.
 - **Zero Krengel Entropy does not kill Poisson Entropy.**
 É. Janvresse, T. de la Rue
 Accepté dans *Annales de l'IHP, Probab. Stat.* (2010).
 - **Averaging along Uniform Random Integers.**
 É. Janvresse, T. de la Rue
 Accepté dans *Uniform Distribution Theory.* (2011).
 - **Dynamics of λ -continued fractions and β -shifts.**
 É. Janvresse, B. Rittaud et T. de la Rue
 Soumis (2011).
 - **Heterogeneity quantification methods for brain spect perfusion images.**
 R. Modzelewski, É. Janvresse, T. de la Rue, and P. Vera
 Soumis (2011).

Table des matières

1	Autour des suites de Fibonacci aléatoires	9
1.1	Suites de Fibonacci aléatoires et généralisations	9
1.2	Exposants de Lyapunov pour $\lambda \geq 2$	11
1.3	Exposants de Lyapunov pour $\lambda < 2$	12
1.3.1	La mesure $\nu_{k,\rho}$ et les intervalles de Stern-Brocot	13
1.3.2	Cas $k = 3$ ($\lambda = 1$)	15
1.3.3	Cas $k > 3$	15
1.3.4	Croissance des exposants de Lyapunov	17
1.4	Comportement “en moyenne”	17
1.5	Pavages	18
1.5.1	Suite de Fibonacci aléatoires et pavages	18
1.5.2	Fractions continues et pavages	20
1.6	Dynamique des fractions continues de Rosen	22
1.6.1	Codage des réels positifs	22
1.6.2	La transformation T_λ ($0 < \lambda < 2$)	27
1.6.3	Lien avec les β -shifts	30
2	La loi de Benford	33
2.1	Introduction	33
2.2	Interprétation probabiliste	35
2.2.1	Heuristique	35
2.2.2	Vitesse de convergence d’une chaîne de Markov	36
2.2.3	Généralisation du résultat de Flehinger	38
2.3	Suites d’entiers aléatoires	38
2.3.1	URI-set	38
2.3.2	Résultats de type <i>loi de Benford</i>	39
2.3.3	Liens entre différentes densités	40
2.3.4	Autres questions	41
3	Quelques problèmes de théorie ergodique	43
3.1	La transformation Pascal-adique	43
3.1.1	Introduction	43
3.1.2	Lâche-bernoullicité	45
3.1.3	Théorèmes limites : corrections au théorème ergodique	46
3.1.4	Autres questions ouvertes et conjectures	48
3.2	Les autocouplages 2 à 2 indépendants	49
3.3	Suspensions de Poisson et entropie en mesure infinie	50
4	Quelques problèmes issus de la physique	53
4.1	Marches aléatoires sur la sphère et sur $SO(n)$	53
4.1.1	Convergence à l’équilibre	53
4.1.2	Trou spectral	53

4.1.3	Inégalités de Sobolev logarithmique	55
4.2	Un problème d'accrochage de polymères	57
4.3	Génération de pavages aléatoires	57
4.4	Traitement d'images médicales	60
	Bibliographie	63

Chapitre 1

Autour des suites de Fibonacci aléatoires

1.1 Suites de Fibonacci aléatoires et généralisations

La suite de Fibonacci doit son nom au mathématicien italien du XIII^e siècle Leonardo Fibonacci qui, dans un problème récréatif posé dans un de ses ouvrages, le Liber Abaci (1202), décrit la croissance d'une population de lapins. Elle est définie de la façon suivante : on fixe les deux premiers termes, puis chaque nouveau terme est obtenu en faisant la somme des deux précédents. Un résultat classique, démontré par Édouard Lucas [66], indique que le rapport de deux termes consécutifs de la suite tend vers le nombre d'or $\phi := (1 + \sqrt{5})/2$, sous réserve que le rapport des deux premiers termes soit différent de $-1/\phi$ (et que, bien sûr, les deux premiers termes de la suite ne soient pas tous les deux nuls).

En collaboration avec Benoît Rittaud (Paris 13) et Thierry de la Rue (Rouen), je me suis intéressée aux suites de Fibonacci aléatoires $(F_n)_{n \geq 1}$ et $(\tilde{F}_n)_{n \geq 1}$, définies par récurrence comme suit : $F_1, F_2, \tilde{F}_1, \tilde{F}_2 > 0$, et pour tout $n \geq 1$,

$$F_{n+2} = \lambda F_{n+1} \pm \mu F_n \quad (\text{cas linéaire}), \quad (1.1)$$

$$\tilde{F}_{n+2} = |\lambda \tilde{F}_{n+1} \pm \mu \tilde{F}_n| \quad (\text{cas non linéaire}), \quad (1.2)$$

où les signes sont choisis selon un processus de Bernoulli de paramètre $p \in [0, 1]$. Notons que ce type de récurrence intéresse particulièrement les physiciens de la matière condensée, qui l'interprètent en termes d'équation de Schrödinger discrète en dimension 1 (voir [23, 93]).

Une première question est bien sûr de savoir si les suites aléatoires croissent elles aussi exponentiellement vite. Il n'est pas difficile de voir qu'un simple argument de renormalisation permet de se ramener au cas où l'un des paramètres (par exemple μ) vaut 1, ce que l'on supposera par la suite. Lorsque chaque signe \pm est choisi indépendamment et vaut $+$ avec probabilité p ($0 < p \leq 1$) ou $-$ avec probabilité $1 - p$, on parle alors de (p, λ) -suite de Fibonacci aléatoire. Cependant, la croissance exponentielle n'est pas préservée. Par exemple, Embree et Trefethen [23] ont étudié le cas linéaire $g_{n+2} = g_{n+1} \pm \beta g_n$ pour $p = 1/2$, ce qui correspond à une $(1/2, 1/\sqrt{\beta})$ -suite de Fibonacci aléatoire. Le taux de croissance exponentielle $\sigma(\beta) := \lim |g_n|^{1/n}$ qu'ils trouvent vérifie alors $\log \sigma(\beta) = \gamma + \log \sqrt{\beta}$, où γ est l'exposant de Lyapunov de la suite de Fibonacci aléatoire correspondante.

Dans le cas $p = 1/2$ et $\lambda = 1$, les deux suites $(|F_n|)$ et (\tilde{F}_n) ont la même distribution que la suite $(|t_n|)$ étudiée par Viswanath dans [100]. Dans ce papier, utilisant la formule de Furstenberg [28] (voir aussi [7], Chap. II), Viswanath prouve qu'avec probabilité 1, $\sqrt[n]{|t_n|} \xrightarrow{n \rightarrow +\infty} 1.13198824\dots$,

où le logarithme de la constante correspond à l'intégrale de la fonction

$$m \mapsto \frac{1}{4} \log \left(\frac{1 + 4m^4}{(1 + m^2)^2} \right)$$

par rapport à une mesure "fractale" explicite ν_f définie inductivement sur les intervalles de Stern-Brocot.

Mais les suites $(|F_n|)$ et (\tilde{F}_n) n'ont en général pas la même distribution. Le cas linéaire peut toujours être vu comme un produit de matrices aléatoires indépendantes et identiquement distribuées (i.i.d.). En effet, la suite $(F_n)_{n \geq 1}$ peut être codée par une suite $(X_n)_{n \geq 3}$ de variables aléatoires i.i.d. à valeurs dans l'alphabet $\{R, L\}$ avec probabilité $(p, 1 - p)$: chaque R correspond au choix du signe $+$ et chaque L correspond au choix du signe $-$, et on peut les interpréter comme la multiplication à droite de (F_{n-1}, F_n) par l'une des matrices suivantes :

$$L := \begin{pmatrix} 0 & -1 \\ 1 & \lambda \end{pmatrix} \quad \text{et} \quad R := \begin{pmatrix} 0 & 1 \\ 1 & \lambda \end{pmatrix}. \quad (1.3)$$

On a alors

$$(F_{n-1}, F_n) = (F_1, F_2) X_3 \dots X_n, \quad \forall n \geq 3.$$

Pour appliquer comme Viswanath la formule de Furstenberg et déterminer le plus grand exposant de Lyapunov, il faudrait arriver à identifier la mesure invariante, ce qui est déjà délicat lorsque $p = 1/2$. Dans le cas non linéaire, à nouveau, chaque R correspond au choix du signe $+$ et peut être interprété comme la multiplication de $(\tilde{F}_{n-1}, \tilde{F}_n)$ par la matrice R ci-dessus. Chaque L correspond au choix du signe $-$, mais il faut tenir compte des valeurs absolues : $X_{n+1} = L$ correspond donc à la multiplication de $(\tilde{F}_{n-1}, \tilde{F}_n)$ ou bien par L ou bien par

$$L' := \begin{pmatrix} 0 & 1 \\ 1 & -\lambda \end{pmatrix}. \quad (1.4)$$

Il s'agit donc aussi d'un produit de matrices aléatoires, mais elles ne sont plus i.i.d. et la théorie ne s'applique plus.

Nous avons donc dû développer une autre méthode pour étudier les (p, λ) -suites de Fibonacci aléatoires ([52, 39]). Il y a une première distinction à faire selon la valeur de λ . Le cas $\lambda \geq 2$ est le plus simple et nous obtenons une formule explicite pour l'exposant de Lyapunov (voir section 1.2). Lorsque $\lambda < 2$, nous avons réussi à traiter une famille dénombrable de paramètres λ , à savoir $\lambda = \lambda_k = 2 \cos \pi/k$ où k est un entier supérieur à 3 (voir Section 1.3). Notre méthode, valable à la fois pour les cas linéaire et non linéaire, consiste à ramener notre problème à l'étude d'une chaîne de Markov d'ordre $(k - 2)$, dont on peut déterminer la croissance exponentielle en utilisant certains développements en λ -fractions continues de Rosen [90]. Cela nous permet de donner une expression de l'exposant de Lyapunov, valable pour tout $k \geq 3$ et tout paramètre $p \in]0, 1]$, sous la forme de l'intégrale de la fonction \log par rapport à une mesure de probabilité sur \mathbb{R}^+ définie récursivement sur des intervalles de Stern-Brocot généralisés. Par une étude fine de cette mesure, nous parvenons à montrer que l'exposant est strictement positif dès que $p > 0$ dans le cas linéaire, alors que dans le cas non linéaire il n'est strictement positif que pour $p > 1/k$, et qu'il est croissant en p . Il est intéressant de noter que les valeurs λ_k sont les seules inférieures à 2 pour lesquelles le groupe des transformations du demi-plan hyperbolique généré par les transformations $z \mapsto -1/z$ et $z \mapsto z + \lambda$ est discret. Nous n'utilisons pas cette propriété, mais cela suggère l'existence d'un lien entre les suites de Fibonacci aléatoires et la géométrie hyperbolique.

Nous nous sommes aussi intéressés au comportement "en moyenne" des (p, λ) -suites de Fibonacci aléatoires. Il s'agit d'étudier la croissance exponentielle de l'espérance de $|F_n|$ et \tilde{F}_n , lorsque n tend vers l'infini. Bien sûr, par l'inégalité de Jensen, on sait que les limites de $n^{-1} \log \mathbb{E}|F_n|$ et $n^{-1} \log \mathbb{E}\tilde{F}_n$ sont supérieures à l'exposant de Lyapunov correspondant. En exploitant à nouveau les réductions qui s'opèrent lorsque $\lambda = \lambda_k$, et en développant des arguments combinatoires, nous montrons dans [53] que $\mathbb{E}\tilde{F}_n$ croît exponentiellement vite si et seulement si $p > p_c := (2 - \lambda_k)/4$.

Comme $p_c < 1/k$, il y a donc des valeurs du paramètre p pour lesquelles p.s. \tilde{F}_n ne croît pas exponentiellement vite, alors que $\mathbb{E}[\tilde{F}_n]$ admet une croissance exponentielle (voir la section 1.4).

Enfin, dans un autre travail [54] qui n'est pas directement liée aux suites de Fibonacci aléatoires, nous avons tenté de mieux comprendre la dynamique liée aux λ -fractions continues (voir la section 1.6). Notre motivation était la suivante : un point essentiel pour traiter le cas des λ_k est la similitude entre la dynamique associée aux λ_k -fractions continues et celle correspondant au développement des réels en base (entière) $k - 1$. Nous nous sommes aperçus qu'en fait cette similitude pouvait se généraliser à tous les λ entre 0 et 2, et nous parvenons à décrire un isomorphisme entre une dynamique associée aux λ -fractions continues et la dynamique du β -shift pour un certain β fonction continue croissante de λ . Nous espérons que cette correspondance générale entre λ -fractions continues et développement en base non entière nous apportera un nouvel angle d'attaque pour traiter le cas des suites de Fibonacci aléatoires pour tout $\lambda \in]0, 2[$.

Dans la section 1.5, nous donnons une interprétation des suites de Fibonacci aléatoires en terme de pavages. Ce point de vue, valable pour tout λ , permet de redémontrer des résultats bien connus sur les continuants et les fractions continues mais nous ne savons pas non plus encore s'il sera utile pour traiter le cas de tous les paramètres $0 < \lambda < 2$.

1.2 Exposants de Lyapunov pour $\lambda \geq 2$

Lorsque $\lambda \geq 2$, les comportements sont identiques dans les cas linéaire et non linéaire. En effet, dans le cas non linéaire, $\mathbb{P}(\tilde{F}_{n+1}/\tilde{F}_n \geq 1 | \tilde{F}_{n-1}, \tilde{F}_n) \geq p$, et si $\tilde{F}_{n+1}/\tilde{F}_n \geq 1$, alors $\tilde{F}_{n+2}/\tilde{F}_{n+1} \geq 1$. Avec probabilité 1, il existe donc un entier N_+ à partir duquel les quotients $\tilde{F}_{n+1}/\tilde{F}_n$ sont tous plus grands que 1. De plus, pour $n \geq N_+$, il n'y a pas besoin de prendre la valeur absolue et la suite se comporte comme dans le cas linéaire.

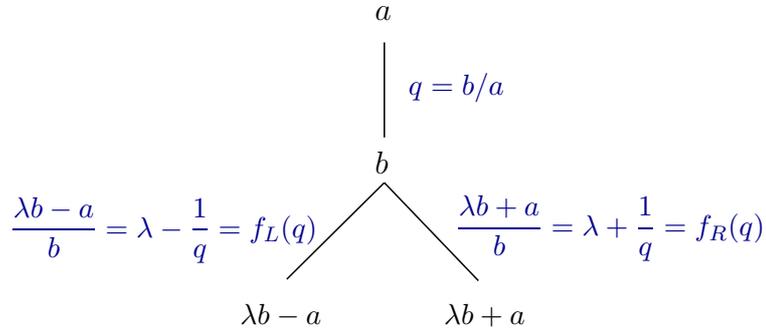


FIGURE 1.1 – Évolution du quotient de deux termes consécutifs d'une (p, λ) -suite de Fibonacci aléatoire (cas linéaire).

La suite de quotients $Q_n := F_n/F_{n-1}$ est une chaîne de Markov à valeurs réelles de probabilité de transition

$$\mathbb{P}\left(Q_{n+1} = f_R(q) \mid Q_n = q\right) = p \quad \text{et} \quad \mathbb{P}\left(Q_{n+1} = f_L(q) \mid Q_n = q\right) = 1 - p,$$

où

$$f_R(q) := \lambda + \frac{1}{q} \geq \lambda \geq f_L(q) := \lambda - \frac{1}{q}$$

(voir Figure 1.1).

Soit $B := \frac{\lambda + \sqrt{\lambda^2 - 4}}{2} \in [1, \lambda]$ le plus grand point fixe de f_L . On définit de façon récursive

des sous-intervalles de \mathbb{R}_+ inclus dans $[B, \lambda + \frac{1}{B}]$ par

$$I_R := f_R([B, \infty]) = \left[\lambda, \lambda + \frac{1}{B} \right] \quad \text{et} \quad I_L := f_L([B, \infty]) = [B, \lambda],$$

et pour toute suite finie X dans $\{R, L\}^*$,

$$I_{XR} := f_R(I_X) \quad \text{et} \quad I_{XL} := f_L(I_X).$$

Comme pour toute suite $(W_i)_{i \geq 1}$ de R et de L , $\bigcap_{n \geq 1} I_{W_n \dots W_1}$ est réduit à un singleton, l'unique mesure de probabilité invariante $\mu_{p,\lambda}$ de la chaîne de Markov $(Q_n) = (F_n/F_{n-1})$ est donnée par

$$\mu_{p,\lambda}(I_W) := p^{|W|_R} (1-p)^{|W|_L} \quad (1.5)$$

pour tout mot fini W dans $\{R, L\}^*$, où $|W|_R$ et $|W|_L$ sont respectivement le nombre de R et de L dans W .

D'après un théorème classique sur la loi des grands nombres d'une chaîne de Markov (voir e.g. [72], Théorème 17.0.1), on obtient dans [39] le résultat suivant :

Théorème 1.1. *Soit $\lambda \geq 2$ et $0 < p \leq 1$. Pour tout $F_1 > 0$ et $F_2 > 0$,*

$$\frac{1}{n} \log |F_n| \xrightarrow[n \rightarrow \infty]{} \gamma_{p,\lambda} := \int_0^\infty \log x \, d\mu_{p,\lambda}(x) \quad p.s.$$

où $\mu_{p,\lambda}$ est donnée par (1.5).

1.3 Exposants de Lyapunov pour $\lambda < 2$

Pour des valeurs bien choisies de la constante λ , $\lambda_k = 2 \cos \pi/k$ ($k \geq 3$), nous avons pu développer une méthode différente de celle de Furstenberg et qui repose sur la réduction des (p, λ_k) -suites de Fibonacci aléatoires. Plus précisément, nous obtenons le théorème suivant (voir [39]) :

Théorème 1.2.

- **Cas linéaire :** pour $p = 0$, la suite $(|F_n|)$ est périodique de période k .
Pour tout $p \in]0, 1[$

$$\frac{1}{n} \log |F_n| \xrightarrow[n \rightarrow \infty]{} \gamma_{p,\lambda_k} := \int_0^\infty \log x \, d\nu_{k,\rho}(x) > 0 \quad p.s.,$$

où

$$\rho := \sqrt[k-1]{1 - p_R}$$

et p_R est l'unique solution strictement positive de

$$\left(1 - \frac{px}{p + (1-p)x} \right)^{k-1} = 1 - x.$$

- **Cas non linéaire :** pour $0 \leq p \leq 1/k$, il existe presque sûrement une sous-suite bornée (\tilde{F}_{n_j}) de (\tilde{F}_n) de densité $(1 - kp)$.
Pour $p \in]1/k, 1[$,

$$\frac{1}{n} \log \tilde{F}_n \xrightarrow[n \rightarrow \infty]{} \tilde{\gamma}_{p,\lambda_k} := \int_0^\infty \log x \, d\nu_{k,\rho}(x) > 0 \quad p.s.,$$

où

$$\rho := \sqrt[k-1]{1 - p_R}$$

et p_R est, pour $p < 1$, l'unique solution strictement positive de

$$\left(1 - \frac{px}{(1-p) + px} \right)^{k-1} = 1 - x.$$

(Pour $p = 1$, $p_R = 1$.)

Dans les deux cas, $\nu_{k,\rho}$ est une mesure de probabilité sur \mathbb{R}_+ définie sur les intervalles de Stern-Brocot lorsque $k = 3$ (voir Figure 1.3) et sur les intervalles de Stern-Brocot généralisés lorsque $k > 3$ (voir Figure 1.4).

Dans le cas non linéaire, la sous-suite bornée mentionnée dans le théorème vérifie $\tilde{F}_{n_{j+1}} = |\lambda_k \tilde{F}_{n_j} - \tilde{F}_{n_{j-1}}|$ pour tout j , ce qui correspond au cas où $p = 0$. Son comportement dépend fortement du choix des valeurs initiales. Ce phénomène était passé inaperçu dans notre premier article [52], car les valeurs initiales étaient fixées à $\tilde{F}_1 = \tilde{F}_2 = 1$. On montre en effet par un argument géométrique (voir [39]) la proposition suivante, valable pour tout λ dans $]0, 2[$:

Proposition 1.3. *Soit (\tilde{F}_n) une $(0, \lambda)$ -suite de Fibonacci dont les deux premiers termes \tilde{F}_1 et \tilde{F}_2 sont strictement positifs. Les propriétés suivantes sont équivalentes :*

1. \tilde{F}_1/\tilde{F}_2 admet un développement fini en λ -fraction continue¹.
2. La suite (\tilde{F}_n) est ultimement périodique.
3. Il existe n tel que $\tilde{F}_n = 0$.

De plus, \tilde{F}_n peut décroître exponentiellement vite vers 0, mais l'exposant dépend alors du ratio \tilde{F}_1/\tilde{F}_2 .

Dans la section suivante, nous expliquons d'où vient la mesure de probabilité $\nu_{k,\rho}$.

1.3.1 La mesure $\nu_{k,\rho}$ et les intervalles de Stern-Brocot

Rappelons que l'on peut coder les suites de Fibonacci par une suite (X_n) de lettres dans l'alphabet $\{R, L\}$ indiquant soit le choix du signe $+$, soit le choix du signe $-$. Selon le contexte, on interprétera une suite finie de R et de L comme le produit de matrices correspondant (voir (1.3) et (1.4)). Pour $0 \leq j \leq k-2$, la matrice RL^j a des coefficients positifs, tandis que

$$RL^{k-1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad RL^{k-2}L' = \text{Id}, \quad RL^{k-1}R = -L \quad \text{et} \quad RL^{k-1}L = -R.$$

Cela signifie que l'on peut toujours retirer les motifs RL^{k-1} dans le processus (X_n) , mais il faut veiller dans le cas linéaire à changer la lettre suivante. Que peut-on alors observer une fois les motifs éliminés? Dans le cas non linéaire, si $p \leq 1/k$, il ne reste presque sûrement plus que des L alors que si $p > 1/k$, il reste une infinité de R . Dans le cas linéaire, il n'y a besoin d'aucune condition sur p pour s'assurer qu'il restera presque sûrement une infinité de R .

Finalement, la suite réduite obtenue en éliminant les motifs RL^{k-1} dans le processus (X_n) ne contient jamais $(k-1)$ L de suite (excepté éventuellement au début). On peut donc la décomposer de manière unique sous la forme $L^s B_1 B_2 B_3 \dots$ où $B_i \in \{R, RL, \dots, RL^{k-2}\}$. La suite de blocs (B_i) est i.i.d. et on calcule facilement la loi d'apparition des blocs :

$$P(B_1 = RL^j) := \frac{\rho^j}{\sum_{m=0}^{k-2} \rho^m}, \quad 0 \leq j \leq k-2,$$

où ρ dépend de k , de p et du cas étudié (linéaire ou non).

La mesure $\nu_{k,\rho}$ sur \mathbb{R}_+ , qui reflète le comportement des quotients F_n/F_{n-1} en fonction de la succession des blocs de la suite réduite est construite récursivement. Voyons quelle est l'action d'un bloc sur le quotient de deux termes consécutifs. Soient $0 \leq j \leq k-2$ et $(a, b) \in \mathbb{R}^* \times \mathbb{R}$. Alors $(a', b') := (a, b)RL^j$ vérifie

$$\frac{b'}{a'} = f_L^j \circ f_R \left(\frac{b}{a} \right),$$

où $f_R(q) := \lambda + 1/q$ et $f_L(q) := \lambda - 1/q$ (voir Figure 1.1). Dans la suite, on notera f_j la fonction $f_L^j \circ f_R$.

1. Pour la définition des développement en λ -fraction continue, voir la Section 1.3.3

Ainsi, on voit qu'après élimination des motifs RL^{k-1} dans $X_1 \dots X_n$, si le dernier bloc est RL^j , $0 \leq j \leq k-2$, alors le quotient F_n/F_{n-1} est dans $I_j := f_j([0, +\infty]) = [b_{j+1}, b_j]$, où

$$b_0 = +\infty, \quad b_1 = \lambda = f_0(+\infty) = f_1(0), \quad b_{j+1} = f_L(b_j) = f_j(+\infty) = f_{j+1}(0).$$

Mais comme $RL^{k-1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, on a $b_{k-1} = f_{k-1}(0) = 0$ et donc $(I_j)_{0 \leq j \leq k-2}$ est une subdivision de $[0, +\infty]$ (voir Figure 1.2). On associe donc à l'intervalle I_j la probabilité $\rho^j / (\sum_{m=0}^{k-2} \rho^m)$ du bloc RL^j correspondant.

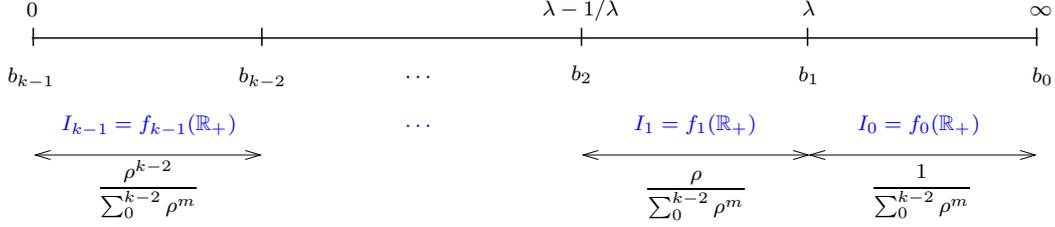


FIGURE 1.2 – Subdivision de \mathbb{R}_+ en $k-1$ intervalles $(I_j)_{0 \leq j \leq k-2}$ auxquels on attribue la probabilité du bloc correspondant.

À l'étape r , on va partitionner \mathbb{R}_+ en fonction des r derniers blocs de la trajectoire réduite. Pour cela, on définit l'intervalle de \mathbb{R}_+ associé à la suite des r blocs $(RL^{j_r}) \dots (RL^{j_1})$:

$$I_{j_1, j_2, \dots, j_r} := f_{j_1} \circ f_{j_2} \circ \dots \circ f_{j_r}([0, +\infty]), \quad \forall (j_1, j_2, \dots, j_r) \in \{0, \dots, k-2\}^r.$$

Pour tout $r \geq 1$, on obtient bien une subdivision $\mathcal{I}(r)$ de $[0, +\infty]$ car

$$I_{j_1, j_2, \dots, j_{r-1}} = \bigcup_{j_r=0}^{k-2} I_{j_1, j_2, \dots, j_r}.$$

On définit alors la mesure $\nu_{k, \rho}$ sur ces $(k-1)^r$ intervalles en posant

$$\nu_{k, \rho}(I_{j_1, j_2, \dots, j_r}) := \frac{\rho^{j_1 + \dots + j_r}}{(\sum_0^{k-2} \rho^m)^r}.$$

(Voir Figures 1.3 et 1.4.)

Voyons quelques cas particuliers : lorsque $p = 1$, c'est-à-dire pour les suites de Fibonacci déterministes, $\rho = 0$ dans les deux cas et $\nu_{k, \rho}$ est la masse de Dirac en $\frac{\lambda_k + \sqrt{\lambda_k^2 + 4}}{2}$ (et on retrouve bien le nombre d'or ϕ pour $\lambda_3 = 1$). Pour $p = 1/2$, on obtient la même valeur pour ρ dans les deux cas, et si de plus $k = 3$ la mesure $\nu_{3, \rho}$ est égale à la mesure de Viswanath conditionnée sur \mathbb{R}^+ . Ce sont les seules valeurs de p pour lesquelles $\gamma_{p, \lambda_k} = \tilde{\gamma}_{p, \lambda_k}$.

Quand $p \rightarrow 0$ dans le cas linéaire, ou $p \rightarrow 1/3$ dans le cas non linéaire, $\rho \rightarrow 1$ et $\nu_{3, \rho} \rightarrow \nu_{3, 1}$ qui est la probabilité sur \mathbb{R}^+ donnant la même masse 2^{-r} à tous les intervalles de Stern-Brocot de rang r . Cette mesure est reliée à la fonction "Question Mark" de Minkowski (voir [17]), qui fait correspondre à tout nombre x dont la représentation en fraction continue est $[a_0; a_1, a_2, \dots]$, le nombre

$$?(x) = a_0 + 2 \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{2^{a_1 + \dots + a_k}}.$$

On a en effet

$$\forall x \in [0, 1], \quad ?(x) = 2 \nu_{3, 1}([0, x]). \quad (1.6)$$

Dans les sections 1.3.2 et 1.3.3, nous explicitons les bornes des intervalles sur lesquelles la mesure $\nu_{k, \rho}$ est définie récursivement. Avant cela, mentionnons une propriété de ces mesures, dont la preuve est fournie dans la section 1.6.1.

Théorème 1.4. *Les mesures $\nu_{k,\rho}$ sont toutes singulières par rapport à la mesure de Lebesgue.*

1.3.2 Cas $k = 3$ ($\lambda = 1$)

À chaque étape de notre procédure, les intervalles sont divisés en $k - 1$ sous-intervalles. Quand $k = 3$, il existe un moyen simple de trouver la nouvelle borne : si $[\frac{a}{b}, \frac{c}{d}]$ est un intervalle de l'étape r , alors $[\frac{a}{b}, \frac{a+c}{b+d}]$ et $[\frac{a+c}{b+d}, \frac{c}{d}]$ sont deux intervalles de l'étape $r + 1$. La nouvelle borne $\frac{a+c}{b+d}$ est appelée *médiane* de $\frac{a}{b}$ et $\frac{c}{d}$.

On voit que si $[\frac{a}{b}, \frac{c}{d}]$ est un intervalle de Stern-Brocot (*i.e.* si a, b, c et d sont des entiers tels que $ad - bc = -1$), alors les deux sous-intervalles ainsi construits le sont encore. Partant de $\mathbb{R}_+ = [\frac{0}{1}, \frac{1}{0}]$, qui peut être vu comme l'intervalle de Stern-Brocot de rang 0, notre construction nous donne donc la subdivision de $[0, +\infty]$ en intervalles de Stern-Brocot. Celle-ci a été découverte indépendamment par le mathématicien Moriz Stern en 1858 et par l'horloger Achille Brocot en 1860. Tout nombre rationnel étant la borne d'un intervalle de Stern-Brocot, notre mesure est bien définie pour $k = 3$.

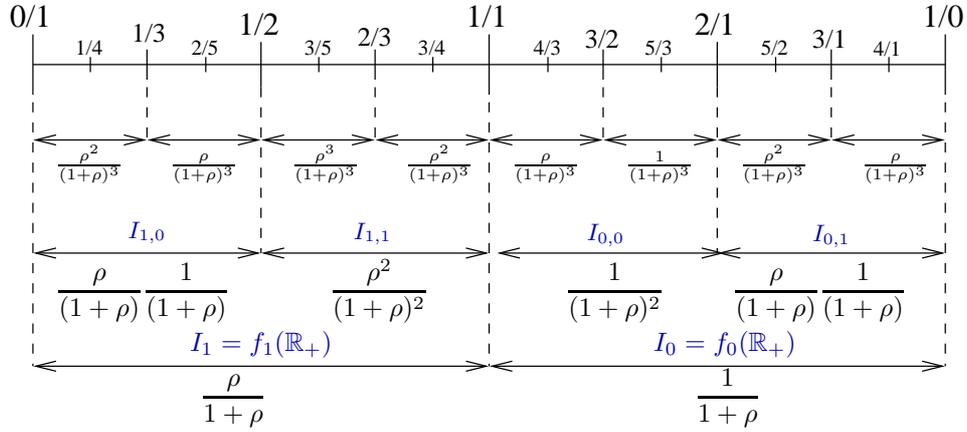


FIGURE 1.3 – Lorsque $k = 3$ ($\lambda = 1$), la mesure $\nu_{3,\rho}$ est définie sur les intervalles de Stern-Brocot classiques (ici de rang 1, 2, 3). On attribue d'abord une masse $\rho/(1 + \rho)$ à l'intervalle $I_1 = [0, 1]$ et $1/(1 + \rho)$ à l'intervalle $I_0 = [1, \infty]$. Une fois $\nu_{3,\rho}$ définie sur un intervalle de Stern-Brocot de rang r , une proportion $1/(1 + \rho)$ de sa masse est donnée au sous-intervalle de gauche (respectivement de droite) de rang $r + 1$ lorsque r est impair (respectivement pair).

1.3.3 Cas $k > 3$

Dans le cas où $k > 3$, on peut aussi expliciter les bornes des intervalles sur lesquelles la mesure $\nu_{k,\rho}$ est définie récursivement, que nous qualifierons d'*intervalles de Stern-Brocot généralisés*. Pour cela, nous avons besoin de rappeler la notion de λ -fraction continue.

Soit $1 \leq \lambda < 2$. Tout nombre réel x peut être écrit sous la forme (voir [90])

$$x = a_0\lambda + \frac{1}{a_1\lambda + \frac{1}{\ddots + \frac{1}{a_n\lambda + \ddots}}}$$

où $(a_n)_{n \geq 0}$ est une suite (finie ou infinie) telle que $a_0 \in \mathbb{Z}$ et $a_n \in \mathbb{Z} \setminus \{0\}$ pour $n \geq 1$. Notons $[a_0, \dots, a_n, \dots]_\lambda$ cette expression, appelée *développement en λ -fraction continue de Rosen* de x ,

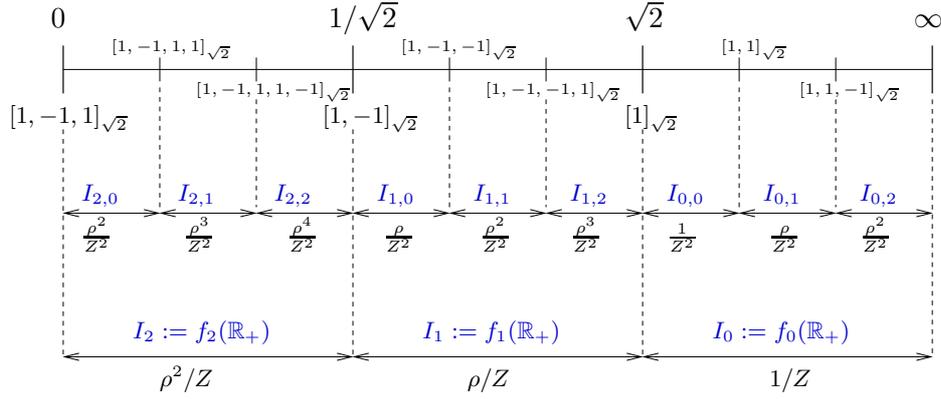


FIGURE 1.4 – La mesure $\nu_{k,\rho}$ sur les intervalles de Stern-Brocot généralisés de rang 1 et 2, dans le cas où $k = 4$ ($\lambda_k = \sqrt{2}$). La constante de normalisation Z est égale à $1 + \rho + \rho^2$. Les bornes des intervalles sont données par leur développement en $\sqrt{2}$ -fraction continue.

qui n'est en général pas unique. Lorsque $\lambda = 1$, on retrouve le développement en fraction continue dans lequel les quotients partiels sont des entiers positifs ou négatifs.

Notons que les fonctions f_j s'expriment facilement à l'aide des développements en λ_k -fraction continue de Rosen. En effet, le développement en λ_k -fraction continue de Rosen de $f_j(q)$ est la concaténation de $(j+1)$ termes ± 1 alternés et du développement de q ou $-q$ (selon la parité de j) :

$$f_j([a_0, \dots, a_n, \dots]_{\lambda_k}) = \begin{cases} \underbrace{[1, -1, 1, \dots, 1, a_0, \dots, a_n, \dots]_{\lambda_k}}_{(j+1) \text{ termes}} & \text{si } j \text{ est pair} \\ \underbrace{[1, -1, 1, \dots, -1, -a_0, \dots, -a_n, \dots]_{\lambda_k}}_{(j+1) \text{ termes}} & \text{si } j \text{ est impair.} \end{cases} \quad (1.7)$$

Les bornes (finies) de la première subdivision $\mathcal{S}(1)$ de notre construction sont donc

$$b_j = f_j(0) = \underbrace{[1, -1, 1, \dots, (-1)^{j-1}]_{\lambda_k}}_{j \text{ termes}} \quad \forall 1 \leq j \leq k-1.$$

De plus, par (1.7), comme $b_{k-1} = 0$, on retrouve bien

$$b_0 = f_0(b_{k-1}) = \underbrace{[1, 1, -1, 1, \dots, \pm 1]_{\lambda_k}}_{k-1 \text{ termes}} = \infty.$$

En itérant (1.7), on voit que pour tout $\ell \geq 1$, les bornes de la subdivision $\mathcal{S}(r)$ peuvent être écrites sous la forme de développements finis en λ_k -fraction continue de Rosen, avec des coefficients ± 1 . Nous montrons par ailleurs que tout λ_k -rationnel positif (les nombres positifs dont un développement en λ_k -fraction continue est fini) est la borne d'un intervalle de Stern-Brocot généralisé. Ainsi, la tribu engendrée par $\mathcal{S}(r)$ converge vers la tribu borélienne sur \mathbb{R}_+ , ce qui montre que la mesure $\nu_{k,\rho}$ est bien définie.

Il serait intéressant de savoir dans quelle mesure le fait qu'on obtienne tous les λ_k -rationnels à partir des intervalles de Stern-Brocot généralisés pourrait aider à leur classification. Actuellement, on connaît les cas suivants : pour $k = 3, 4, 6$, les λ_k -rationnels correspondants sont respectivement \mathbb{Q} , $\sqrt{2}\mathbb{Q}$, $\sqrt{3}\mathbb{Q}$; pour $k = 2^n + 1$, les λ_k -rationnels sont $\mathbb{Q}(\lambda_k)$; enfin, pour $k = 8, 10, 12$, les λ_k -rationnels sont $\lambda_k\mathbb{Q}(\lambda_k^2)$ (voir [5, 6, 63, 64]).

1.3.4 Croissance des exposants de Lyapunov

Nous avons aussi obtenu quelques résultats sur la dépendance en p des plus grands exposants de Lyapunov. Plus précisément, lorsque $k = 3$, nous avons une formule pour leur dérivée par rapport à p , qui fait intervenir la mesure produit $\nu_{3,\rho} \otimes \nu_{3,\rho}$, et qui nous permet de prouver leur croissance (voir [52]). Lorsque $k \geq 3$ est quelconque, nous montrons aussi dans [39] que les exposants sont croissants en p , mais par une méthode de couplage. Contrairement au cas linéaire, le plus grand exposant de Lyapunov n’est pas une fonction analytique de p dans le cas non linéaire. Nous ne sommes néanmoins pas parvenu à établir des résultats de convexité.

Pour $p = 1$, la $(1, \lambda)$ -suite de Fibonacci (déterministe) $(F_n)_n = (\tilde{F}_n)_n$ croît exponentiellement vite et

$$\tilde{\gamma}_{1,\lambda} = \gamma_{1,\lambda} = \log \left(\frac{\lambda + \sqrt{\lambda^2 + 4}}{2} \right), \quad \text{pour tout } \lambda.$$

Donc $\tilde{\gamma}_{1,\lambda} = \gamma_{1,\lambda}$ est croissant en λ . De plus, les simulations numériques que nous avons faites (voir Figure 1.5) nous amènent à faire la conjecture suivante :

Conjecture : pour p fixé, les exposants de Lyapunov γ_{p,λ_k} et $\tilde{\gamma}_{p,\lambda_k}$ sont croissants en k , et $\gamma_{p,\lambda}$ est croissant en λ lorsque $\lambda \geq 2$.

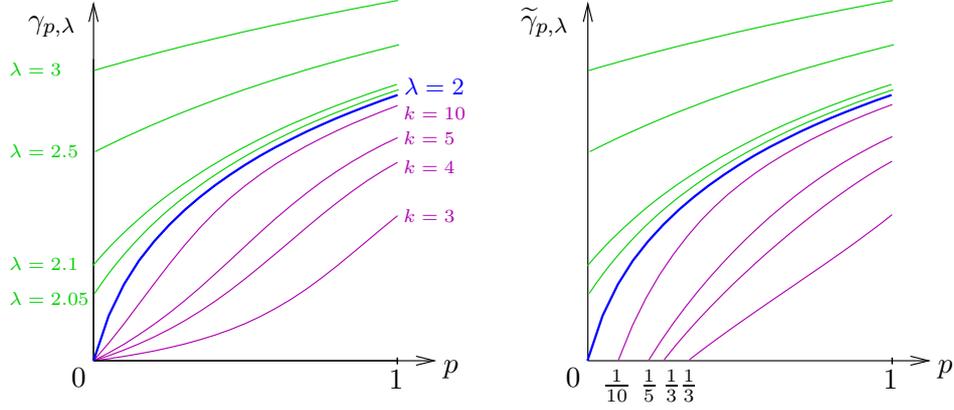


FIGURE 1.5 – Valeurs de γ_{p,λ_k} (cas linéaire, à gauche) et $\tilde{\gamma}_{p,\lambda_k}$ (cas non linéaire, à droite) pour $\lambda = \lambda_k$ ($k = 3, 4, 5, 10$), $\lambda = 2$ (en gras), $\lambda = 2.05, 2.1, 2.5$ et 3 . Les simulations numériques supportent la conjecture selon laquelle γ_{p,λ_k} et $\tilde{\gamma}_{p,\lambda}$ sont croissants en k . Attention, pour $p \in [1/3, 1]$, le graphe de $\tilde{\gamma}_{p,\lambda_3}$ n’est pas un segment de droite : en effet, on a $\frac{d\tilde{\gamma}_p}{dp}(1) = \frac{\log 5}{2}$, tandis que la pente moyenne est $(3/2) \log \phi$.

1.4 Comportement “en moyenne”

Dans cette partie, nous résumons nos résultats concernant le comportement “en moyenne” des suites de Fibonacci aléatoires étudié dans [53]. Il s’agit d’étudier, lorsque n tend vers l’infini, l’espérance du n -ième terme d’une (p, λ) -suite de Fibonacci aléatoire.

Dans le cas linéaire, on montre facilement que $\mathbb{E}(F_n) = \lambda \mathbb{E}(F_{n-1}) + (2p-1) \mathbb{E}(F_{n-2})$ pour tout λ . Le polynôme correspondant à cette récurrence admet deux racines réelles $(\lambda \pm \sqrt{\lambda^2 + 4(2p-1)})/2$, et on obtient donc une formule explicite pour $\mathbb{E}(F_n)$ dépendant des conditions initiales. La question intéressante dans le cas linéaire est donc plutôt le taux de croissance de $\mathbb{E}(|F_n|)$.

Le cas particulier $p = 1/2$ et $\lambda = 1$ a été résolu dans [87], où il est montré que le taux de croissance de l’espérance d’une $(1/2, 1)$ -suite de Fibonacci aléatoire converge vers $\alpha - 1 \approx 1.20556943$, où α est l’unique racine réelle de $\alpha^3 = 2\alpha^2 + 1$. Le théorème suivant, montré dans [53], généralise ce résultat :

Théorème 1.5. *Pour tout n , soit $m_n := \mathbb{E}(\tilde{F}_n)$ l'espérance du n -ième terme d'une (p, λ) -suite de Fibonacci aléatoire (cas non linéaire).*

(1) *Cas $\lambda \geq 2$: si $0 < p \leq 1$, alors*

$$\frac{m_{n+1}}{m_n} \xrightarrow{n \rightarrow \infty} \frac{\lambda + \sqrt{\lambda^2 + 4(2p-1)}}{2}.$$

(2) *Cas $\lambda = \lambda_k$, avec $k \geq 3$:*

- *Si $p > (2 - \lambda_k)/4$, alors*

$$\frac{m_{n+1}}{m_n} \xrightarrow{n \rightarrow \infty} \alpha_k(p) \left(1 + \frac{pq^{k-1}}{\alpha_k(p)^k} \right) > 1,$$

où $q := 1 - p$ et $\alpha_k(p)$ est la seule racine positive du polynôme $P_k(X) := X^{2k} - \lambda_k X^{2k-1} - (2p-1)X^{2k-2} - \lambda_k pq^{k-1} X^{k-1} - p^2 q^{2k-2}$.

- *Si $p = (2 - \lambda_k)/4$, alors $(m_n)_n$ croît au plus linéairement.*

- *Si $p < (2 - \lambda_k)/4$, alors $(m_n)_n$ est bornée.*

Notons que la valeur critique pour laquelle la croissance de $\mathbb{E}(\tilde{F}_n)$ n'est plus exponentielle est $p_c := (2 - \lambda_k)/4 = (1 - \cos(\pi/k))/2$: elle est donc strictement inférieure à la valeur critique $1/k$ pour laquelle l'exposant de Lyapunov est nul (cf. Théorème 1.2).

Comme nous l'a suggéré Steven Finch (INRIA), les simulations numériques concernant le taux de croissance de l'espérance d'une (p, λ) -suite de Fibonacci aléatoire ne sont pas faciles à interpréter. Cela est dû à la différence de comportement entre \tilde{F}_n et $\mathbb{E}(\tilde{F}_n)$. Pour $\lambda = \lambda_k$, on a vu que si $\frac{2-\lambda_k}{4} < p \leq 1/k$, l'espérance d'une (p, λ) -suite de Fibonacci aléatoire croissait exponentiellement vite tandis que la suite contenait presque sûrement une sous-suite bornée. Lorsque $1/k < p < 1$, les estimations numériques du taux de croissance de l'espérance suggèrent qu'il est strictement plus grand que l'exposant de Lyapunov correspondant. Cela impliquerait que la variance de \tilde{F}_n croît exponentiellement vite à un taux au moins deux fois plus grand que celui de l'espérance :

$$\frac{1}{n} \log \mathbb{E}[(\tilde{F}_n - \mathbb{E}(\tilde{F}_n))^2] \geq \frac{1}{n} \mathbb{E}[2 \log(\mathbb{E}(\tilde{F}_n) - \tilde{F}_n)] = \frac{2}{n} \left(\log \mathbb{E}(\tilde{F}_n) + \mathbb{E}[\log(1 - \tilde{F}_n/\mathbb{E}(\tilde{F}_n))] \right).$$

Dans le cas $p = 1/2$ et $\lambda = 1$, il est prouvé [67] que le taux de croissance de la variance est $1 + \sqrt{5}$. Il serait intéressant de connaître la valeur exact du taux de la variance pour tout p , et plus généralement des moments d'ordre supérieur.

Afin d'étudier le comportement d'une (p, λ) -suite de Fibonacci aléatoire pour d'autres valeurs de λ , nous avons étudié le comportement du taux de croissance d'une suite au voisinage de $\lambda = 2$.

Corollaire 1.6. *Soit $0 < p \leq 1$. Supposons que, pour λ dans un voisinage de 2, l'espérance d'une (p, λ) -suite de Fibonacci aléatoire croît exponentiellement vite au taux $\mathcal{G}(\lambda)$. Alors \mathcal{G} ne peut être analytique en $\lambda = 2$.*

Dans le cas particulier $p = 1/2$, on peut montrer la non-analyticité du taux de croissance sur tout voisinage gauche de 2.

1.5 Pavages

1.5.1 Suite de Fibonacci aléatoires et pavages

L'interprétation des termes de la suite de Fibonacci classique ($s_{n+1} = s_n + s_{n-1}$ pour tout $n \geq 1$, $s_0 = s_1 = 1$) en terme de pavages est bien connue : s_n est le nombre de pavages possibles de n cases à l'aide de carrés et de dominos. Cette interprétation a été généralisée dans [2] à des

suites définies par des récurrences linéaires à coefficients entiers positifs en considérant des carrés et des polyominos de différentes couleurs.

Une interprétation en terme de pavages est aussi possible lorsque les coefficients sont des réels, mais nous ne savons pas encore si elle peut être utile pour l'étude des exposants de Lyapunov pour les valeurs de λ différentes des λ_k . Notons que le caractère aléatoire de la suite de Fibonacci ne joue pas ici de rôle particulier, elle pourrait être considérée dans cette section comme une suite de $+$ et de $-$.

Considérons par exemple la suite définie par ses deux premiers termes $G_0 = 1$ et $G_1 = a_1$, et pour tout $n \geq 2$, $G_n = a_n G_{n-1} + b_n G_{n-2}$, où $(a_n)_{n \geq 1}$ et $(b_n)_{n \geq 2}$ sont deux suites de réels. Nous allons aussi nous intéresser aux pavages de cases à l'aide de carrés et de dominos, mais cette fois chaque carré et chaque domino est affecté d'un poids dépendant de la case recouverte : pour tout i , un carré qui recouvre la case i a un poids a_i et un domino qui recouvre les cases $i - 1$ et i a un poids b_i . On définit le *poids d'un pavage* comme le produit des poids des polyominos qui le constituent (voir Figure 1.6).

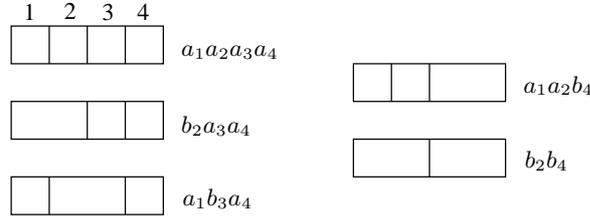


FIGURE 1.6 – Les 5 pavages possibles des cases 1 à 4 avec leur poids respectif. La somme des poids des pavages des 4 cases est donc $\mathscr{W}_1^4 = a_1 a_2 a_3 a_4 + b_2 a_3 a_4 + a_1 b_3 a_4 + a_1 a_2 b_4 + b_2 b_4$.

Soit \mathscr{W}_1^n la somme des poids de tous les pavages possibles des n cases numérotées 1 à n . On a évidemment $\mathscr{W}_1^1 = a_1$ puisque le seul pavage possible de la case 1 est constitué d'un carré. Si $n \geq 2$, tout pavage des cases numérotées 1 à n se décompose ou bien en un pavage des cases 1 à $(n - 1)$ suivi d'un carré sur la case n , ou bien en un pavage des cases 1 à $(n - 2)$ suivi d'un domino recouvrant les cases $n - 1$ et n . On en déduit donc que

$$\mathscr{W}_1^n = a_n \mathscr{W}_1^{n-1} + b_n \mathscr{W}_1^{n-2}, \text{ pour tout } n \geq 2$$

(avec la convention $\mathscr{W}_1^0 = 1$). Finalement, on a montré que $G_n = \mathscr{W}_1^n$. En particulier, on en déduit le théorème suivant :

Théorème 1.7. *Soit $(F_n)_n$ une (p, λ) -suite de Fibonacci aléatoire de premier terme $F_0 = 1$. Alors pour tout n , F_n est la somme des poids des pavages des cases 1 à n avec des dominos et des carrés, où le poids d'un carré sur la case 1 est $a_1 = F_1$, sur la case $i \geq 2$ est $a_i = \lambda$, et les poids des dominos sont des variables aléatoires i.i.d. égales à 1 avec probabilité p ou -1 avec probabilité $1 - p$.*

Corollaire 1.8. *Soit $(F_n)_n$ une (p, λ) -suite de Fibonacci aléatoire de premiers termes $F_0 = 1$ et $F_1 = \lambda$. Alors pour tout n ,*

$$F_n = \sum_{k=0}^{\lfloor n/2 \rfloor} \lambda^{n-2k} \sum_{\substack{(i_\ell)_{\ell \leq k} \in \{2, \dots, n\}^k \\ i_\ell - i_{\ell-1} > 1}} \prod_{\ell=1}^k b_{i_\ell},$$

où les (b_i) sont des variables aléatoires i.i.d. égales à 1 avec probabilité p ou -1 avec probabilité $1 - p$.

Démonstration. Chaque pavage des cases 1 à n est constitué de k dominos affectés d'un poids dépendant de leur place et de $n - 2k$ carrés de poids λ . □

Lorsque $p = 1$ et $a_i = 1$ pour tout i , on retrouve la formule bien connue pour la suite de Fibonacci classique : $s_n = \sum_{k \geq 0} \binom{n-k}{k}$.

Notons que lorsque $b_i = 1$ pour tout i , on peut voir \mathcal{W}_1^n comme un polynôme en les n variables a_1, \dots, a_n . Ces polynômes, appelés *continuants* à cause de leur lien avec les fractions continues, ont été étudiés par Euler (voir [30], chapitre 6.7). Dans la section suivante, nous faisons le lien entre fractions continues et pavages, et redémontrons quelques résultats bien connus sur les fractions continues. En particulier, la formule du miroir (corollaire 1.10) explique pourquoi le début du développement en fraction continue du quotient de termes consécutifs d'une suite de Fibonacci aléatoire dépend des derniers signes ($-$ ou $+$) utilisés lors de la construction de la suite.

1.5.2 Fractions continues et pavages

Pour toutes suites $(a_n)_{n \geq 1}$ et $(b_n)_{n \geq 2}$ de réels, notons

$$[a_1, {}^{b_2} a_2, {}^{b_3} a_3, \dots, {}^{b_n} a_n] := a_1 + \frac{b_2}{a_2 + \frac{b_3}{a_3 + \frac{b_4}{\dots + \frac{b_n}{a_n}}}}$$

Pour $1 \leq j \leq n$, soit \mathcal{W}_j^n la somme des poids de tous les pavages possibles des $(n-j+1)$ cases numérotées j à n , dans lesquels un carré sur la case i est affecté du poids a_i et un domino sur les cases $i-1$ et i est affecté du poids b_i . Si $j > n$, on adopte la convention $\mathcal{W}_j^n := 1$.

Théorème 1.9. *Pour tout $n \geq 1$, on a*

$$[a_1, {}^{b_2} a_2, {}^{b_3} a_3, \dots, {}^{b_n} a_n] = \frac{\mathcal{W}_1^n}{\mathcal{W}_2^n}.$$

De plus, si $a_i \in \mathbb{Z}$ et $b_i = \pm 1$ pour tout i , alors \mathcal{W}_1^n et \mathcal{W}_2^n sont premiers entre eux.

Démonstration. Soit $i \in \{1, \dots, n-1\}$. Tout pavage des cases numérotées i à n se décompose ou bien en un pavage des cases $i+1$ à n précédé par un carré sur la case i , ou bien en un pavage des cases $i+2$ à n précédé par un domino recouvrant les cases i et $i+1$. Donc $\mathcal{W}_i^n = a_i \mathcal{W}_{i+1}^n + b_{i+1} \mathcal{W}_{i+2}^n$ pour tout $1 \leq i < n$. Ainsi, on en déduit que

$$\frac{\mathcal{W}_i^n}{\mathcal{W}_{i+1}^n} = a_i + \frac{b_{i+1}}{\mathcal{W}_{i+1}^n / \mathcal{W}_{i+2}^n}, \quad \text{pour tout } 1 \leq i < n.$$

Comme $\mathcal{W}_n^n = a_n$, cela suffit à montrer la première partie du théorème. En outre, si \mathcal{W}_{i+2}^n et \mathcal{W}_{i+1}^n sont premiers entre eux, alors quand $a_i \in \mathbb{Z}$ et $b_i = \pm 1$, $\mathcal{W}_i^n = a_i \mathcal{W}_{i+1}^n + b_{i+1} \mathcal{W}_{i+2}^n$ et \mathcal{W}_{i+1}^n sont aussi premiers entre eux. \square

Le corollaire suivant, connu sous l'appellation de formule du miroir, explique pourquoi lorsqu'à la section 1.3.1 nous avons étudié les quotients de termes consécutifs de suites de Fibonacci aléatoires, ce sont les derniers blocs qui déterminaient le début du développement en fraction continue de Rosen du quotient. Cette formule nous permet même d'écrire explicitement le développement en λ -fraction continue du quotient de termes consécutifs d'une (p, λ) -suite de Fibonacci aléatoire quel que soit λ .

Corollaire 1.10 (Formule du miroir). *Pour tout $n \geq 1$, on a*

$$[a_n, {}^{b_n} a_{n-1}, \dots, {}^{b_2} a_1] = \frac{\mathcal{W}_1^n}{\mathcal{W}_1^{n-1}}.$$

Démonstration. D'après le théorème 1.9, la fraction continue $[a_n, {}^{b_n} a_{n-1}, \dots, {}^{b_2} a_1]$ est égale au quotient du poids total des pavages des cases n à 1 et du poids total des pavages des cases $n-1$ à 1, donc à $\mathscr{W}_1^n / \mathscr{W}_1^{n-1}$. \square

Théorème 1.11 (Identité d'Euler).

$$\mathscr{W}_1^{m+n} \mathscr{W}_{m+1}^{m+\ell} - \mathscr{W}_1^{m+\ell} \mathscr{W}_{m+1}^{m+n} = (-1)^\ell \prod_{k=1}^{\ell+1} b_{m+k} \mathscr{W}_1^{m-1} \mathscr{W}_{m+\ell+2}^{m+n}$$

$m \geq 1, l \geq 0, n \geq l+1$.

La formule précédente est connue dans le cas où $b_i = 1$ pour tout i sous le nom d'identité d'Euler et est habituellement prouvée par induction. Une autre preuve, proposée dans [96], repose sur le fait que le déterminant d'une matrice antisymétrique de taille paire est le carré d'un certain polynôme. Notons qu'elle permet aussi de retrouver l'identité de Cassini : $s_{n+1}s_{n-1} - s_n^2 = (-1)^n$.

Démonstration. On dit que deux pavages ont une *cassure commune* en i s'ils possèdent tous les deux un pavé (carré ou domino) débutant à la case $i+1$. Notons que $\mathscr{W}_1^{m+n} \mathscr{W}_{m+1}^{m+\ell}$ est la somme, sur tous les couples (P, P') où P est un pavage des cases 1 à $m+n$ et P' est un pavage des cases $m+1$ à $m+\ell$, des produits des poids de P et de P' . De même, $\mathscr{W}_1^{m+\ell} \mathscr{W}_{m+1}^{m+n}$ est la somme, sur tous les couples (\tilde{P}, \tilde{P}') où \tilde{P} est un pavage des cases 1 à $m+\ell$ et \tilde{P}' est un pavage des cases $m+1$ à $m+n$, des produits des poids de \tilde{P} et de \tilde{P}' .

Nous allons établir une bijection entre l'ensemble des tels couples (P, P') tels que P et P' ont une cassure commune et l'ensemble des tels couples (\tilde{P}, \tilde{P}') tels que \tilde{P} et \tilde{P}' ont une cassure commune. En outre, cette bijection conservera le produit des poids des pavages. En effet, puisque P et P' ont au moins une cassure commune, il suffit d'échanger la fin de ces pavages après leur dernière cassure commune pour obtenir un pavage \tilde{P} des cases 1 à $m+\ell$ et un pavage \tilde{P}' des cases $m+1$ à $m+n$ ayant au moins une cassure commune et dont le produit des poids est égal au produit des poids de P et P' (voir Figure 1.7).

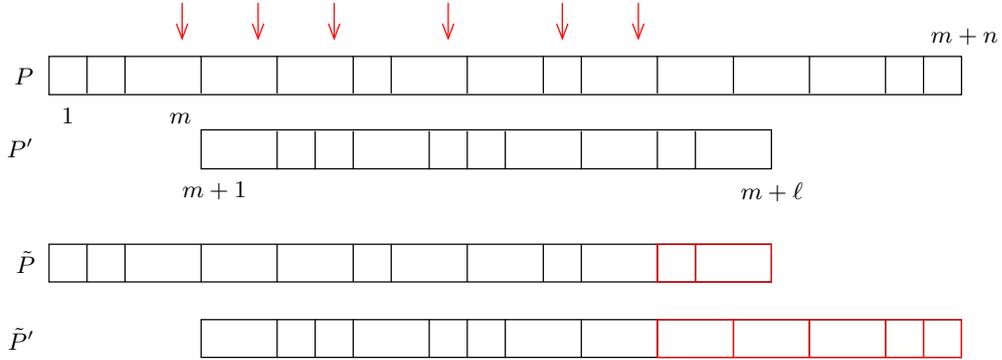


FIGURE 1.7 – En haut, un pavage P des cases 1 à $m+n$ et un pavage P' des cases $m+1$ à $m+l$; les flèches indiquent leurs cassures communes. En échangeant les polyominos après la dernière cassure commune, on obtient un pavage \tilde{P} des cases 1 à $m+l$ et un pavage \tilde{P}' des cases $m+1$ à $m+n$.

Si ℓ est pair, alors la seule possibilité pour que deux pavages P et P' n'aient pas de cassure commune est que P' soit constitué uniquement de dominos et P ait uniquement des dominos sur les cases m à $m+l+1$. Ainsi, le produit des poids de tous les couples (P, P') n'ayant pas de cassure commune est $\prod_{k=1}^{\ell+1} b_{m+k} \mathscr{W}_1^{m-1} \mathscr{W}_{m+\ell+2}^{m+n}$. Par ailleurs, tous les couples de pavages (\tilde{P}, \tilde{P}') ont une cassure commune.

Si ℓ est impair, tous les couples de pavages (P, P') ont une cassure commune et la seule possibilité pour que deux pavages \tilde{P} et \tilde{P}' n'aient pas de cassure commune est que \tilde{P} ait uniquement

des dominos à partir de la case m et \tilde{P}' ait uniquement des dominos jusqu'à la case $m + \ell + 1$. Ainsi, le produit des poids de tous les couples (\tilde{P}, \tilde{P}') n'ayant pas de cassure commune est $\prod_{k=1}^{\ell+1} b_{m+k} \mathcal{W}_1^{m-1} \mathcal{W}_{m+\ell+2}^{m+n}$. \square

Corollaire 1.12. *Pour tout n ,*

$$[a_1, {}^{b_2} a_2, {}^{b_3} a_3, \dots, {}^{b_n} a_n] - [a_1, {}^{b_2} a_2, {}^{b_3} a_3, \dots, {}^{b_{n-1}} a_{n-1}] = \frac{(-1)^n \prod_{i=2}^n b_i}{\mathcal{W}_2^n \mathcal{W}_2^{n-1}}.$$

Démonstration. D'après le théorème 1.9, on peut réécrire la formule à démontrer sous la forme

$$\mathcal{W}_1^n \mathcal{W}_2^{n-1} - \mathcal{W}_1^{n-1} \mathcal{W}_2^n = (-1)^n \prod_{i=2}^n b_i,$$

qui est un cas particulier du théorème 1.11. \square

Lemme 1.13 (Folding lemma). *Pour tout n et tout $c \in \mathbb{R}$, on a*

$$[a_1, {}^{b_2} a_2, {}^{b_3} \dots, {}^{b_n} a_n, {}^1 c, {}^1 -a_n, {}^{b_n} -a_{n-1}, {}^{b_{n-1}} \dots, {}^{b_3} -a_2] = [a_1, {}^{b_2} a_2, {}^{b_3} \dots, {}^{b_n} a_n] + \frac{(-1)^{n+1} \prod_{i=2}^n b_i}{c(\mathcal{W}_2^n)^2}.$$

Démonstration. D'après la formule du miroir (Corollaire 1.10), on a $[a_n, {}^{b_n} a_{n-1}, \dots, {}^{b_3} a_2] = \frac{\mathcal{W}_2^n}{\mathcal{W}_2^{n-1}}$. Comme $[-a_n, {}^{b_n} -a_{n-1}, \dots, {}^{b_3} -a_2] = -[a_n, {}^{b_n} a_{n-1}, \dots, {}^{b_3} a_2]$, on a donc

$$[a_1, {}^{b_2} a_2, {}^{b_3} \dots, {}^{b_n} a_n, {}^1 c, {}^1 -a_n, {}^{b_n} -a_{n-1}, {}^{b_{n-1}} \dots, {}^{b_3} -a_2] = \left[a_1, {}^{b_2} a_2, {}^{b_3} \dots, {}^{b_n} a_n, {}^1 c - \frac{\mathcal{W}_2^{n-1}}{\mathcal{W}_2^n} \right].$$

Le théorème 1.12 nous permet donc d'écrire

$$[a_1, {}^{b_2} a_2, {}^{b_3} \dots, {}^{b_n} a_n, {}^1 c, {}^1 -a_n, {}^{b_n} -a_{n-1}, {}^{b_{n-1}} \dots, {}^{b_3} -a_2] - [a_1, {}^{b_2} a_2, {}^{b_3} \dots, {}^{b_n} a_n] = \frac{(-1)^{n+1} \prod_{i=2}^n b_i}{c \mathcal{W}_2^n \mathcal{W}_2^{n+1}},$$

où le poids a_{n+1} d'un carré sur la case $n + 1$ est égal à $c - \frac{\mathcal{W}_2^{n-1}}{\mathcal{W}_2^n}$. Mais on a $\mathcal{W}_2^{n+1} = a_{n+1} \mathcal{W}_2^n + \mathcal{W}_2^{n-1} = c \mathcal{W}_2^n$, ce qui conclut la preuve du lemme. \square

1.6 Dynamique des fractions continues de Rosen

1.6.1 Codage des réels positifs

L'étude des (p, λ_k) -suites de Fibonacci aléatoires et de la mesure de Stern-Brocot $\nu_{\rho, k}$ nous a amené à introduire une partition de \mathbb{R}_+ en $(k - 1)$ intervalles de Stern-Brocot généralisés (voir Section 1.3.1 et Figure 1.4) de la forme

$$\mathbb{R}_+ = \bigcup_{j=0}^{k-2} f_j(\mathbb{R}_+),$$

et à étudier la transformation $T_k : [0, \infty[\rightarrow [0, \infty[$, correspondant à la suppression du dernier bloc. Plus précisément, T_k est définie par

$$\forall x \in f_j(\mathbb{R}_+), T_k(x) := f_j^{-1}(x).$$

(Voir Figure 1.8.)

La transformation T_k permet de coder tout nombre réel positif x , en indiquant à chaque étape l'intervalle dans lequel le nombre x se trouve. Ce codage rappelle donc le développement

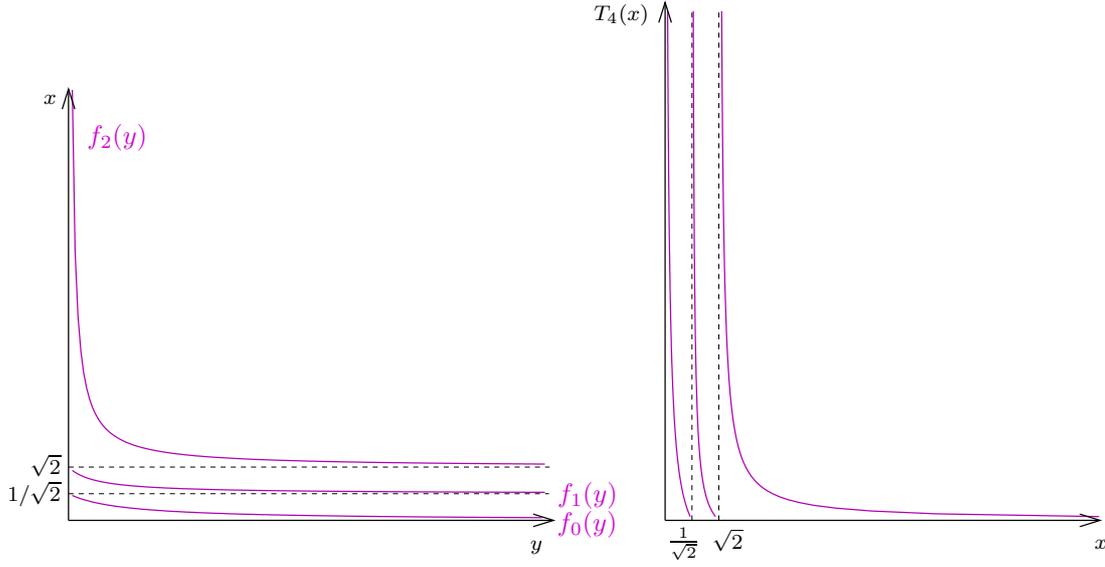


FIGURE 1.8 – Gauche : graphes des fonctions f_j pour $\lambda = \sqrt{2}$ ($k = 4$). Droite : graphes de la transformation T_4 .

en base $k - 1$, qui permet de localiser n'importe quel réel compris entre 0 et 1 : pour obtenir le développement en base $(k - 1)$ d'un réel $t \in [0, 1]$, on part de l'intervalle $[0, 1]$, qui est coupé en $k - 1$ intervalles de longueur $1/(k - 1)$, eux-mêmes coupés en $k - 1$ sous-intervalles de longueur $1/(k - 1)^2$, et ainsi de suite. À la n -ième étape, on définit t_n comme l'entier entre 0 et $(k - 2)$ codant le sous-intervalle auquel le réel t appartient. La suite (t_n) constitue un codage symbolique du réel t qui reçoit une interprétation analytique donnée par l'écriture de x sous la forme de la série suivante :

$$t = \sum_{n \geq 0} \frac{t_n}{(k - 1)^{n+1}}.$$

Tout comme pour le développement en base entière, la suite (x_n) d'entiers entre 0 et $k - 2$ définie pour tout $x \in \mathbb{R}_+$ grâce à la transformation T_k possède une interprétation analytique, donnée par le développement en λ_k -fraction continue de x . Plus précisément, pour $k = 3$ ($\lambda_k = 1$), considérons le réel positif x codé par la suite $(x_n)_{n \geq 1} \in \{0, 1\}^{\mathbb{N}^*}$, et introduisons une "coupure" à l'instant n lorsque $x_n = x_{n+1}$. Soit (n_ℓ) la suite (croissante) des instants de coupure. Alors le développement en fraction continue de x est donné, si $x_1 = 1$ par

$$x = [0, n_1, n_2 - n_1, \dots, n_\ell - n_{\ell-1}, \dots] < 1,$$

sinon par

$$x = [n_1, n_2 - n_1, \dots, n_\ell - n_{\ell-1}, \dots] > 1.$$

Exemple : si x est codé par une suite (x_n) commençant par 00101100, alors on a $n_1 = 1$, $n_2 = 5$ et $n_3 = 7$. Comme $x_1 = 0$, le développement en fraction continue de x est de la forme $[1, 4, 2, \dots]$.

Pour $k > 3$, l'équation (1.7) permet d'obtenir le développement en λ_k -fraction continue de tout $x \in \mathbb{R}_+$ à partir de la suite (x_n) . Notons que ce développement en λ_k -fraction continue utilise uniquement des ± 1 alors que pour $k = 3$ nous arrivons à obtenir une fraction continue classique avec des entiers positifs.

Un gros inconvénient de ce codage est qu'il ne respecte pas l'ordre lexicographique (car les fonctions (f_j) sont décroissantes) et qu'il semble difficile à généraliser à des valeurs $\lambda \neq \lambda_k$. Nous introduisons donc dans la section suivante une nouvelle dynamique T_λ , obtenue en modifiant les fonctions f_j afin de travailler avec des fonctions homographiques croissantes.

Avant cela, notons toutefois quelques propriétés de la transformation T_k .

Proposition 1.14. *La transformation T_k préserve la mesure μ sur \mathbb{R}_+ , de densité $\frac{d\mu}{dx}(x) = 1/x$ par rapport à la mesure de Lebesgue.*

Démonstration. Notons que les fonctions (f_j) sont des fonctions homographiques de la forme

$$f_j(y) = \frac{u_{j+1}y + u_j}{u_jy + u_{j-1}},$$

où $u_{-1} = 0$, $u_0 = 1$ et pour tout $j \geq 0$, $u_{j+1} = \lambda_k u_j - u_{j-1}$. De plus, la forme particulière de λ_k donne $u_{k-2-j} = u_j$ pour tout $j \geq 0$. On en déduit donc pour tout $b \geq a \geq 0$

$$\begin{aligned} \mu(T_k^{-1}([a, b])) &= \sum_{j=0}^{k-2} \int_{f_j(b)}^{f_j(a)} \frac{dx}{x} \\ &= \sum_{j=0}^{k-2} \ln \left(\frac{u_{j+1}a + u_j}{u_ja + u_{j-1}} \right) - \ln \left(\frac{u_{j+1}b + u_j}{u_jb + u_{j-1}} \right) \\ &= \ln \left(\frac{u_{k-1}a + u_{k-2}}{u_0a + u_{-1}} \right) - \ln \left(\frac{u_{k-1}b + u_{k-2}}{u_0b + u_{-1}} \right) \\ &= \ln b - \ln a = \mu([a, b]). \end{aligned}$$

□

Théorème 1.15. *La mesure μ sur \mathbb{R}_+ , de densité $\frac{d\mu}{dx}(x) = 1/x$ par rapport à la mesure de Lebesgue, est ergodique pour T_k .*

Ce théorème nous permet de montrer la singularité des mesures $\nu_{k,\rho}$ (Théorème 1.4).

Démonstration du théorème 1.4. D'après le théorème ergodique de Hopf pour une transformation T ergodique et préservant la mesure ν , on a pour tout $f, g \in L^1(\nu)$ tels que $\int g d\nu \neq 0$,

$$\frac{\sum_{i=0}^n f(T^i x)}{\sum_{i=0}^n g(T^i x)} \xrightarrow{n \rightarrow \infty} \frac{\int f d\nu}{\int g d\nu} \quad \nu\text{-p.s.}$$

Grâce au théorème 1.15, en prenant $f = \mathbb{1}_{[1/2\lambda_k, 1/\lambda_k]}$ et $g = \mathbb{1}_{[0, 1/2\lambda_k]}$, on a alors

$$\frac{\sum_{i=0}^n f(T_k^i x)}{\sum_{i=0}^n g(T_k^i x)} \xrightarrow{n \rightarrow \infty} 0 \quad \mu\text{-p.s.}$$

car $\int g d\mu = \infty$. Mais, T_k étant un décalage de Bernoulli sous $\nu_{\rho,k}$, on a aussi

$$\frac{\sum_{i=0}^n f(T_k^i x)}{\sum_{i=0}^n g(T_k^i x)} \xrightarrow{n \rightarrow \infty} \frac{\nu_{\rho,k}([1/2\lambda_k, 1/\lambda_k])}{\nu_{\rho,k}([0, 1/2\lambda_k])} = \sum_{j=1}^{k-2} \rho^j \neq 0 \quad \nu_{\rho,k}\text{-p.s.}$$

Donc $\nu_{\rho,k}$ et μ sont mutuellement singulières. □

La fin de cette section est consacrée à la démonstration du théorème 1.15. Avant cela, observons d'un peu plus près la dynamique de T_k . Soit $n \geq 1$. Si $x \in \left[\frac{1}{(n+1)\lambda_k}, \frac{1}{n\lambda_k} \right[\subset I_{k-2}$, alors $T_k x = f_{k-2}^{-1}(x) \in](n-1)\lambda_k, n\lambda_k]$. De même, si $x \in]n\lambda_k, (n+1)\lambda_k] \subset I_0$, alors $T_k x = f_0^{-1}(x) \in \left[\frac{1}{n\lambda_k}, \frac{1}{(n-1)\lambda_k} \right[$. Ainsi, pour presque tout $x \in \mathbb{R}_+$, le temps d'entrée

$$\min \left\{ j \geq 1 : T_k^j x \in \left[\frac{1}{2\lambda_k}, 2\lambda_k \right] \right\}$$

est fini. De plus, pour presque tout $x \in \left[\frac{1}{\lambda_k}, \lambda_k\right]$, le temps de sortie

$$\min \left\{ j \geq 1 : T_k^j x \notin \left[\frac{1}{\lambda_k}, \lambda_k\right] \right\}$$

est lui aussi fini. En effet, posons

$$B := \left\{ x \in \left[\frac{1}{\lambda_k}, \lambda_k\right] : \forall j \geq 1, T_k^j x \in \left[\frac{1}{\lambda_k}, \lambda_k\right] \right\}.$$

On a $B \subset T_k^{-1}B$. Si $\mu(B)$ n'était pas nulle, alors on aurait aussi $\mu\left(T_k^{-1}B \cap \left[\frac{1}{2\lambda_k}, \frac{1}{\lambda_k}\right]\right) > 0$, et donc $\mu(T_k^{-1}B) > \mu(B)$. Or T_k préserve la mesure μ d'après la proposition 1.14. Donc $\mu(B) = 0$.

Comme T_k possède la propriété de symétrie suivante

$$T_k\left(\frac{1}{x}\right) = 1/T_k x \text{ pour tout } x \in \mathbb{R}_+,$$

nous allons considérer la transformation $\tilde{T}_k : [0, 1] \rightarrow [0, 1]$ définie par

$$\tilde{T}_k x = \begin{cases} T_k x & \text{si } T_k x < 1, \\ 1/T_k x & \text{sinon.} \end{cases}$$

Cette transformation préserve la mesure μ' de densité dx/x sur $[0, 1]$ (voir proposition 1.14). De plus, d'après ce que nous avons vu précédemment concernant les temps d'entrée, p.s. finis, de $\left[\frac{1}{2\lambda_k}, 2\lambda_k\right]$ pour T_k (la première entrée dans cet intervalle se faisant toujours dans $\left[\frac{1}{2\lambda_k}, \frac{1}{\lambda_k}\right] \cup [\lambda_k, 2\lambda_k]$) et les temps de sortie, p.s. finis, de $\left[\frac{1}{\lambda_k}, \lambda_k\right]$, il découle que pour μ' -presque tout x , le temps d'entrée $r_{\tilde{A}}(x)$ dans $\tilde{A} :=]\frac{1}{2\lambda_k}, \frac{1}{\lambda_k}[$ pour \tilde{T}_k est fini :

$$r_{\tilde{A}}(x) := \min \left\{ j \geq 1 : \tilde{T}_k^j x \in \tilde{A} \right\} < \infty \text{ p.s.}$$

Proposition 1.16. *La transformation $\tilde{T}_{k, \tilde{A}}$ induite par \tilde{T}_k sur \tilde{A} est ergodique.*

Ouvrons une parenthèse pour remarquer que dans le cas $k = 3$ ($\lambda_k = 1$), la transformation $\tilde{T}_{3, \tilde{A}}$ est isomorphe à la transformation de Gauss G associée aux fractions continues usuelles $G : x \in [0, 1] \mapsto \left\{\frac{1}{x}\right\} \in [0, 1]$, qui préserve la mesure de densité $\frac{1}{x+1}$ par rapport à la mesure de Lebesgue. Soit $x = [0, a_1, a_2, a_3, \dots] \in [0, 1]$. Alors

$$\tilde{T}_3 x = \begin{cases} [0, a_1 - 1, a_2, a_3, \dots] & \text{si } a_1 > 1 \text{ (i.e. } x < 1/2), \\ [0, a_2, a_3, \dots] & \text{si } a_1 = 1 \text{ (i.e. } x \geq 1/2). \end{cases}$$

Pour $x = [0, 1, a_2, a_3, \dots] \in \tilde{A} =]1/2, 1[$, on a donc $\tilde{T}_3 \circ \tilde{T}_{3, \tilde{A}} x = [0, a_3, \dots] = G \circ \tilde{T}_3 x$.

Démonstration de la Proposition 1.16. Pour tout $i \geq 1$ et tout entier $\ell_1, \dots, \ell_i \geq 1$, soit $R_{\ell_1, \dots, \ell_i}$ l'ensemble des réels dans $[0, 1]$ dont les i premiers temps de retour dans \tilde{A} sont ℓ_1, \dots, ℓ_i :

$$R_{\ell_1, \dots, \ell_i} := \left\{ x \in [0, 1] : r_{\tilde{A}}(x) = \ell_1, r_{\tilde{A}}(\tilde{T}_{k, \tilde{A}} x) = \ell_2, \dots, r_{\tilde{A}}(\tilde{T}_{k, \tilde{A}}^{i-1} x) = \ell_i \right\}.$$

Notons que $R_{\ell_1, \dots, \ell_i}$ est un ouvert. En effet, pour $i = 1$, on peut réécrire R_ℓ comme

$$\begin{aligned} R_\ell &= \left\{ x \in [0, 1] : \tilde{T}_k^\ell x \in \tilde{A} \text{ et pour tout } j < \ell, \tilde{T}_k^j x \notin \tilde{A} \right\} \\ &= \left\{ x \in [0, 1] : \tilde{T}_k^\ell x \in \tilde{A} \text{ et pour tout } j < \ell, \tilde{T}_k^j x \in [0, 1/2\lambda_k[\cup]1/2\lambda_k, 1] \right\} \end{aligned}$$

car si $\tilde{T}_k^j x \in \{1/2\lambda_k, 1/\lambda_k\}$, alors comme $\tilde{T}_k(1/2\lambda_k) = 1/\lambda_k$ et $\tilde{T}_k(1/\lambda_k) = 0$, on ne peut pas avoir $\tilde{T}_k^\ell x \in \tilde{A}$. Puisque \tilde{A} est un ouvert et que \tilde{T}_k est continue, on en déduit que R_ℓ est un ouvert. Par ailleurs, en remarquant que

$$R_{\ell_1, \dots, \ell_i} = R_{\ell_1, \dots, \ell_{i-1}} \cap \tilde{T}_k^{-(\ell_1 + \dots + \ell_{i-1})}(R_{\ell_i})$$

on obtient en faisant une récurrence sur i que $R_{\ell_1, \dots, \ell_i}$ est un ouvert.

Soit $x \in R_{\ell_1, \dots, \ell_i}$, on note $I^{(i)}(x)$ le plus grand intervalle ouvert contenant x et inclus dans $R_{\ell_1, \dots, \ell_i}$. Alors $\tilde{T}_k^{\ell_1 + \dots + \ell_i} \Big|_{I^{(i)}(x)}$ est une fonction homographique (continue, donc monotone) envoyant $I^{(i)}(x)$ sur \tilde{A} .

Pour $i = 1$, posons $]a, b[:= I^{(1)}(x)$. Il existe forcément un entier $j \leq \ell_1$ tel que $\tilde{T}_k^j(a) \in \{1/2\lambda_k, 1/\lambda_k\}$ (sinon, on pourrait prolonger $I^{(1)}(x)$). Cela implique que $\tilde{T}_k^{\ell_1}(a) \in \{0, 1/2\lambda_k, 1/\lambda_k\}$. Mais comme $\tilde{T}_k^{\ell_1}(a)$ est dans la fermeture de \tilde{A} , on en déduit que $\tilde{T}_k^{\ell_1}(a)$ ne peut prendre qu'une des deux valeurs $1/2\lambda_k$ ou $1/\lambda_k$ (et il en est de même pour $\tilde{T}_k^{\ell_1}(b)$). Par récurrence sur j entre 1 et ℓ_1 , on montre que $\tilde{T}_k^j \Big|_{I^{(1)}(x)}$ est une fonction homographique monotone continue.

On peut alors faire une récurrence sur i , en remarquant que

$$I^{(i)}(x) = I^{(1)}(x) \cap \tilde{T}_k^{-\ell_1} \left(I^{(i-1)}(x) \right)$$

et donc que

$$\tilde{T}_k^{\ell_1 + \dots + \ell_i} \Big|_{I^{(i)}(x)} = \tilde{T}_k^{\ell_2 + \dots + \ell_i} \Big|_{I^{(i-1)}(\tilde{T}_k^{\ell_1} x)} \circ \tilde{T}_k^{\ell_1} \Big|_{I^{(1)}(x)}.$$

Dans le graphe de \tilde{T}_k , chaque branche est la réciproque ou l'inverse de la réciproque d'une fonction f_j , donc la réciproque d'une fonction homographique à coefficients positifs. C'est aussi vrai pour les branches du graphe des itérés de \tilde{T}_k . Étudions une des branches de $\tilde{T}_k^{\ell_1 + \dots + \ell_i}$ sur l'intervalle $I^{(i)}(x)$. Soit $f : y \mapsto \frac{uy+u'}{vy+v'}$ la fonction homographique réciproque à coefficients positifs correspondant à cette branche. Pour tout $y_1, y_2 \in \tilde{A}$, $y_1/y_2 \in [1/2, 2]$. Donc comme $v'/v > 0$,

$$\forall y_1, y_2 \in \tilde{A}, \quad \frac{1}{4} \leq \frac{f'(y_1)}{f'(y_2)} = \left(\frac{y_2 + \frac{v'}{v}}{y_1 + \frac{v'}{v}} \right)^2 \leq 4.$$

On en déduit alors que pour tout $y_1, y_2 \in \tilde{A}$,

$$\frac{\left| \tilde{T}_{k, \tilde{A}}^{-i}([y_1, y_2]) \cap I^{(i)}(x) \right|}{|I^{(i)}(x)|} = \frac{\left| f(y_2) - f(y_1) \right|}{\left| f(1/\lambda_k) - f(1/2\lambda_k) \right|} = \frac{|f'(t_1)||y_2 - y_1|}{|f'(t_2)|1/2\lambda_k} \geq \frac{1}{4} \frac{|y_2 - y_1|}{|\tilde{A}|}$$

Donc pour tout borélien $C \subset \tilde{A}$, pour presque tout x et pour tout $i \geq 1$, on a

$$\frac{\left| \tilde{T}_{k, \tilde{A}}^{-1}(C) \cap I^{(i)}(x) \right|}{|I^{(i)}(x)|} \geq \frac{|C|}{4|\tilde{A}|}. \quad (1.8)$$

Soit $C \subset \tilde{A}$ invariant par $\tilde{T}_{k, \tilde{A}}$ et tel que $|C| > 0$. La formule (1.8) dit alors que la proportion de C dans chaque $I^{(i)}(x)$ est toujours au moins $\frac{1}{4|\tilde{A}|} > 0$. Or les $I^{(i)}(x)$, ($x \in [0, 1]$, $i \geq 1$), engendrent la tribu des boréliens. On en déduit que $|C| = |\tilde{A}|$. Donc $\tilde{T}_{k, \tilde{A}}$ est ergodique. \square

Nous sommes maintenant prêts à démontrer le théorème 1.15.

Démonstration du théorème 1.15. Supposons qu'il existe $E \subset \mathbb{R}_+$, invariant par T_k , tel que $|E| > 0$ et $|E^c| > 0$. Soit alors $E_0 := E \cap [0, 1]$ et $E_1 := E \cap [1, \infty[$.

L'ergodicité de $\tilde{T}_{k, \tilde{A}}$ (Proposition 1.16) implique l'ergodicité de \tilde{T}_k . On en déduit alors que

$$\left| \left\{ x, \frac{1}{x} \right\} \cap E \right| = 1 \text{ pour presque tout } x. \quad (1.9)$$

(car par exemple l'ensemble $\{x \in [0, 1], \{x, \frac{1}{x}\} \cap E = \emptyset\}$ est \tilde{T}_k -invariant.)

Supposons que $|E_1 \cap [\lambda_k, 2\lambda_k]| = 0$. Alors, d'après (1.9), $[\frac{1}{2\lambda_k}, \frac{1}{\lambda_k}] \subset E$ (à un négligeable près), ce qui implique que $T_k^2 \left([\frac{1}{2\lambda_k}, \frac{1}{\lambda_k}] \right) = \mathbb{R}_+ \subset E$, et on arrive à une contradiction. Ainsi $|E_1 \cap [\lambda_k, 2\lambda_k]| > 0$. Avec le même raisonnement, on montre que l'on a aussi $|E_0 \cap [\frac{1}{2\lambda_k}, \frac{1}{\lambda_k}]| > 0$.

Soit maintenant $x \in \tilde{A} =]\frac{1}{2\lambda_k}, \frac{1}{\lambda_k}[$ et soient ℓ_1, \dots, ℓ_i les i premiers temps de retour de x dans \tilde{A} . Alors $T_k^{\ell_1 + \dots + \ell_i}(I^{(i)}(x))$ est soit complètement contenu dans $[0, 1]$, soit complètement contenu dans $[1, \infty[$ (car $\tilde{T}_k^j(I^{(i)}(x))$, $j \leq \ell_1 + \dots + \ell_i$, ne contient aucun point critique). Donc $T_k^{\ell_1 + \dots + \ell_i}(I^{(i)}(x)) \cap E$ est soit égal à $E_0 \cap]\frac{1}{2\lambda_k}, \frac{1}{\lambda_k}[$ soit égal à $E_1 \cap]\lambda_k, 2\lambda_k[$. En appliquant la formule (1.8) avec l'un ou l'autre, on obtient

$$\frac{|T_k^{-(\ell_1 + \dots + \ell_i)}(E) \cap I^{(i)}(x)|}{|I^{(i)}(x)|} \geq c,$$

pour une constante $c > 0$. Ainsi,

$$|E \cap I^{(i)}(x)| \geq c|I^{(i)}(x)|.$$

Comme cela est vrai pour tout x et tout i , E remplit tout \tilde{A} . Mais on a déjà vu que c'était impossible! \square

1.6.2 La transformation T_λ ($0 < \lambda < 2$)

Nous nous intéressons dans cette section à une nouvelle dynamique T_λ , obtenue en modifiant les fonctions (f_j) afin de travailler avec des fonctions homographiques croissantes et obtenir un codage qui respecte l'ordre lexicographique. Cette fois, il n'y a plus de restriction sur λ , qui sera un réel strictement inférieur à 2. Les résultats ci-après font l'objet de l'article [54].

Commençons par définir les fonctions homographiques h et h_0 par

$$h_0(y) := \frac{1}{\lambda + \frac{1}{y}} \quad \text{et} \quad h(y) := \frac{1}{\lambda - y}.$$

Puis, définissons de façon récursive les fonctions homographiques $(h_i)_i$ de la façon suivante : pour tout i tel que $h_i(0) < \lambda$,

$$h_{i+1}(y) := h \circ h_i(y).$$

La dernière fonction ainsi définie, notée h_{i_λ} , admet donc un pôle en un point que l'on notera ℓ_λ . Notons que si $\lambda \in]2 \cos(\pi/(k-1)), 2 \cos(\pi/k)]$ ($\forall k \geq 3$), alors $i_\lambda = k - 2$. On obtient ainsi une nouvelle partition de \mathbb{R}_+ :

$$\mathbb{R}_+ = h_0(\mathbb{R}_+) \cup \dots \cup h_{i_\lambda-1}(\mathbb{R}_+) \cup h_{i_\lambda}([0, \ell_\lambda]) = \bigcup_{i=0}^{i_\lambda} I_i,$$

sur laquelle on définit la transformation $T_\lambda : [0, \infty[\rightarrow [0, \infty[$ par

$$\forall x \in I_i, T_\lambda(x) := h_i^{-1}(x).$$

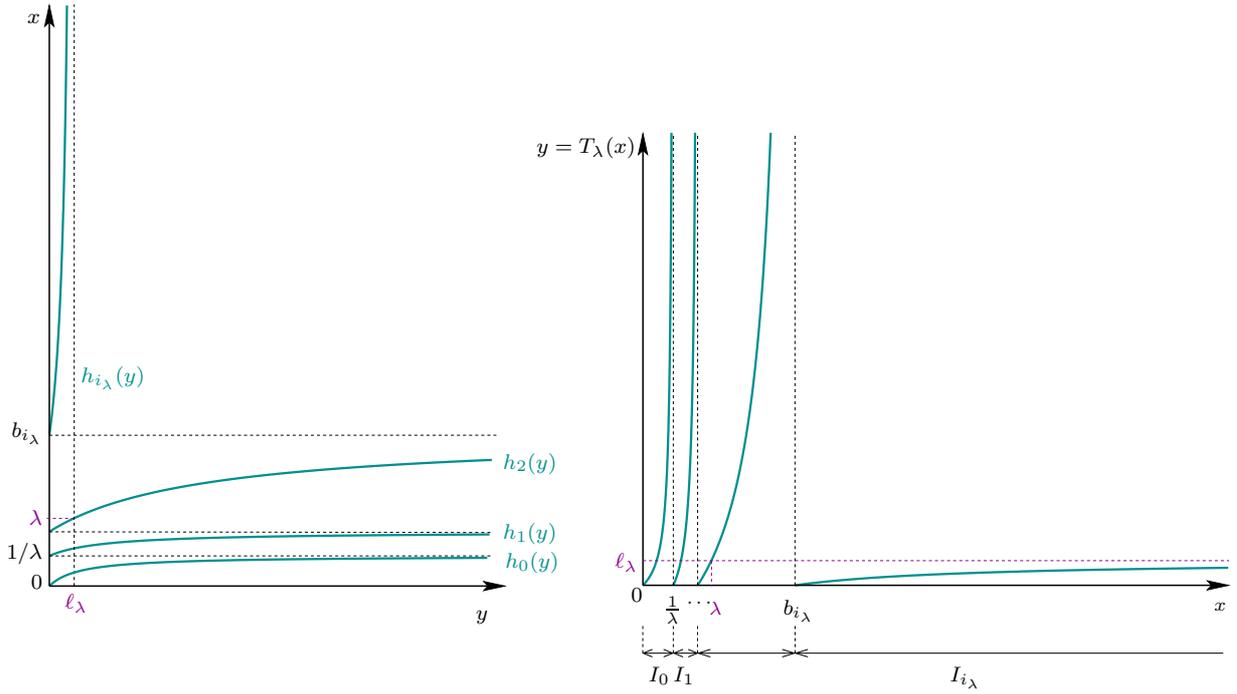


FIGURE 1.9 – À gauche, graphes des fonctions homographiques h_i . À droite, graphes de la transformation T_λ . (Ici $\lambda = 1.5 \in]2 \cos(\pi/4), 2 \cos(\pi/5)[$ et donc $i_\lambda = 3$.)

(Voir Figure 1.9.)

Pour λ de la forme $2 \cos(\pi/k)$, la transformation T_λ admet une unique mesure invariante absolument continue, de densité $\frac{d\mu}{dx}(x) = 1/x$ par rapport à la mesure de Lebesgue. Les mesures invariantes pour les autres valeurs de λ semblent plus difficiles à analyser.

Interprétation géométrique

La transformation T_λ possède une jolie interprétation géométrique qui a été utile pour nos démonstrations et que nous donnons maintenant. Soit θ tel que $\lambda = 2 \cos \theta$.

Soit $x \in \mathbb{R}_+$. Pour trouver $T_\lambda(x)$, écrivons d'abord x sous la forme $x = \frac{t_1}{t_0}$, où $t_1 \geq 0$, $t_0 > 0$. Il est possible de trouver un cercle centré à l'origine et deux points M_0 et M_1 sur ce cercle, d'abscisses respectives t_0 et t_1 , tel que l'angle orienté (OM_0, OM_1) soit égal à θ . Considérons alors l'image de M_1 par la rotation d'angle θ : son abscisse est $t_2 = \lambda t_1 - t_0$ (voir Figure 1.10). On a donc $\frac{t_2}{t_1} = \frac{\lambda t_1 - t_0}{t_1} = h^{-1}\left(\frac{t_1}{t_0}\right)$. En continuant tant que l'abscisse du point obtenu après rotation est positive, on obtient finalement un point M_{i+1} d'abscisse $t_{i+1} < 0$ vérifiant $\frac{t_i}{-t_{i+1}} = h_0^{-1}\left(\frac{t_i}{t_{i-1}}\right) = h_i^{-1}(x) = T_\lambda(x)$.

Pour itérer la transformation, on recommence à placer deux points d'abscisses t_i et $-t_{i+1}$ sur un cercle centré à l'origine et formant un angle θ .

Notons que cette interprétation géométrique nous avait déjà permis dans [39] de montrer la proposition 1.3. Mais un point important pour montrer la monotonie du codage des réels que nous introduisons maintenant (voir Proposition 1.17) est que les rayons des cercles utilisés à chaque itération sont décroissants. C'est aussi ce qui permet d'établir l'inégalité (1.10) ci-après.

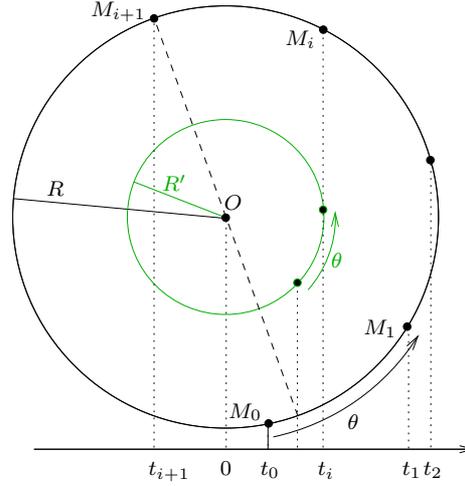


FIGURE 1.10 – Interprétation géométrique de la transformation $T_\lambda : T_\lambda \left(\frac{t_1}{t_0} \right) = \frac{t_i}{-t_{i+1}}$

Codage de l'orbite d'un point $x \in \mathbb{R}_+$

Pour tout $x \in \mathbb{R}_+$, on définit l'orbite de x comme

$$\omega_\lambda(x) := x_0 x_1 x_2 \dots,$$

où x_ℓ est l'unique élément de $\{0, \dots, i_\lambda\}$ tel que $T_\lambda^\ell(x) \in I_{x_\ell}$.

Notre codage symbolique des réels positifs respecte bien l'ordre lexicographique sur $\{0, \dots, i_\lambda\}^{\mathbb{Z}_+}$, et même les distingue bien. Soient $\omega = (\omega_j), \omega' = (\omega'_j) \in \{0, \dots, i_\lambda\}^{\mathbb{Z}_+}$. On note $\omega \prec \omega'$ ou $\omega' \succ \omega$ lorsqu'il existe $\ell \geq 0$ tel que $\omega_\ell < \omega'_\ell$ et $\omega_j = \omega'_j$ pour tout $0 \leq j < \ell$.

Proposition 1.17. *Pour tous $x < x'$, on a $\omega_\lambda(x) \prec \omega_\lambda(x')$.*

La forme particulière des fonctions (h_i) permet donc de trouver un algorithme simple pour déterminer un développement en λ -fraction continue d'un réel positif x , dans lequel le premier coefficient est nul, le second est strictement positif, et les signes des coefficients suivants alternent : écrivons l'orbite $\omega_\lambda(x)$ de x sous la forme $0^{e_0} a_0 0^{e_1} a_1 \dots 0^{e_\ell} a_\ell 0^{e_{\ell+1}} \dots$, où $a_i > 0$ et $e_i \geq 0$ pour tout i . Alors si $e_{\ell+1} = \infty$ (i.e. il existe un entier n_0 tel que $x_n = 0$ pour tout $n \geq n_0$), x est la borne de gauche de l'intervalle $I_{0^{e_0} a_0 0^{e_1} a_1 \dots 0^{e_\ell} a_\ell}^\lambda$ et peut s'écrire

$$x = \llbracket \underbrace{e_0 + 1, 1, \dots, 1}_{a_0 \text{ termes}}, \underbrace{e_1 + 2, 1, \dots, 1}_{a_1 \text{ termes}}, \dots, \underbrace{e_\ell + 2, 1, \dots, 1}_{a_\ell \text{ termes}} \rrbracket_\lambda,$$

où pour toute suite $(b_i)_{i \geq 1}$ d'entiers strictement positifs,

$$\llbracket b_1, b_2, \dots, b_\ell, \dots \rrbracket_\lambda := [0, b_1, -b_2, \dots, (-1)^{\ell-1} b_\ell, \dots]_\lambda = \frac{1}{b_1 \lambda - \frac{1}{\ddots - \frac{1}{b_\ell \lambda}}}.$$

Sinon, il existe une constante $C < 1$ et une suite (s_ℓ) qui tend vers l'infini telles que pour tout ℓ

$$\left| x - \llbracket \underbrace{e_0 + 1, 1, \dots, 1}_{a_0 \text{ termes}}, \underbrace{e_1 + 2, 1, \dots, 1}_{a_1 \text{ termes}}, \dots, \underbrace{e_\ell + 2, 1, \dots, 1}_{a_\ell \text{ termes}} \rrbracket_\lambda \right| \leq C^{s_\ell}. \quad (1.10)$$

On peut donc écrire x sous la forme d'un développement en λ -fraction continue infini :

$$x = \llbracket \underbrace{e_0 + 1, 1, \dots, 1}_{a_0 \text{ termes}}, \underbrace{e_1 + 2, 1, \dots, 1}_{a_1 \text{ termes}}, \dots, \underbrace{e_\ell + 2, 1, \dots, 1}_{a_\ell \text{ termes}}, \dots \rrbracket_\lambda.$$

1.6.3 Lien avec les β -shifts

Nous allons voir dans cette section que les orbites de la transformation T_λ ont une structure similaire aux orbites pour les β -shifts.

Rappels sur les β -shifts

Pour tout $\beta > 1$, Renyi [86] définit le β -développement d'un nombre réel $0 \leq t < 1$ comme la série

$$t = \sum_{n \geq 0} \frac{t_n}{\beta^{n+1}},$$

où les coefficients t_n sont des entiers positifs ou nuls. Considérons la transformation $S_\beta : t \in [0, 1[\rightarrow \beta t \bmod 1$ (voir Figure 1.11). La partition naturelle de $[0, 1[$ associée à cette transformation est constituée d'intervalles de la forme $J_j^\beta := \left[\frac{j}{\beta}, \frac{j+1}{\beta} \right[$, pour $0 \leq j < \beta - 1$, et de $J_{[\beta]}^\beta := \left[\frac{[\beta]}{\beta}, 1 \right[$, où $[\beta]$ est le plus grand entier strictement plus petit que β ; Les coefficients t_n sont donnés par l'orbite de t sous la transformation $S_\beta : t_n = j$ si $S_\beta^n t \in J_j^\beta$.

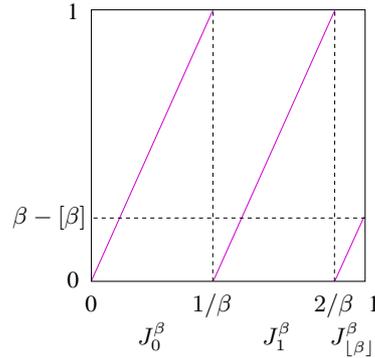


FIGURE 1.11 – Graphe de S_β pour $\beta = 9/4$.

Soit $\mathcal{O}_\beta(t)$ la suite $(t_n)_{n \geq 0}$ codant l'orbite de $t \in [0, 1[$ sous la transformation S_β . Comme dans le cas de T_λ , l'application $t \mapsto \mathcal{O}_\beta(t)$ est strictement croissante pour l'ordre lexicographique. Notons $\mathcal{O}_\beta(1)$ la limite, quand $t \rightarrow 1$, de $\mathcal{O}_\beta(t)$. Pour tout $n \geq 0$, on a $\sigma^n \mathcal{O}_\beta(1) \preceq \mathcal{O}_\beta(1)$, où σ désigne le shift.

Parry caractérise les suites qui codent les orbites pour S_β en utilisant le β -développement de $\beta - [\beta]$. Son résultat peut être traduit en terme de $\mathcal{O}_\beta(1)$ grâce au lemme suivant :

Lemme 1.18. *Soit $(b_i)_{i \geq 0}$ le codage de l'orbite de $\beta - [\beta]$. Si $(b_i)_i$ contient une infinité de termes non nuls, alors $\mathcal{O}_\beta(1) = [\beta]b_0b_1\dots$. Sinon, si β n'est pas un entier, soit b_ℓ le dernier terme non nul. Alors $\mathcal{O}_\beta(1)$ est une répétition périodique du motif $[\beta]b_0\dots b_{\ell-1}(b_\ell - 1)$. Si $\beta \in \mathbb{Z}_+$, $\mathcal{O}_\beta(1) = (\beta - 1)(\beta - 1)\dots$*

Théorème 1.19 (Parry [80], Théorème 3). *Une suite $(t_n)_{n \geq 0}$ est le β -développement d'un nombre réel $t \in [0, 1[$ si et seulement si*

$$t_n t_{n+1} \dots \prec \mathcal{O}_\beta(1), \quad \forall n \geq 0. \quad (1.11)$$

Caractérisation des orbites de T_λ

Pour cela, nous aurons besoin de considérer l'orbite de l'infini, définie comme

$$\omega_\lambda(\infty) := \lim_{x \rightarrow \infty} \uparrow \omega_\lambda(x).$$

Théorème 1.20 (Caractérisation des orbites). $(x_i)_{i \geq 0}$ est l'orbite d'un point $x \in \mathbb{R}_+$ pour la transformation T_λ ssi

$$x_\ell x_{\ell+1} \cdots \prec \omega_\lambda(\infty), \quad \forall \ell \geq 0. \quad (1.12)$$

Conjugaison entre T_λ et un β -shift

Puisque l'ensemble des suites codant les orbites pour T_λ et pour S_β ont la même structure (voir (1.12) et (1.11)), nous avons été amenés à chercher une correspondance entre λ et β : connaissant $\lambda \in]0, 2[$, nous montrons qu'il existe un unique $\beta > 1$ tel que $\mathcal{O}_\beta(1) = \omega_\lambda(\infty)$.

Nous avons vu que $x \mapsto \omega_\lambda(x)$ était une application bijective de $[0, \infty[$ dans l'ensemble des suites d'entiers vérifiant (1.12). De plus, le théorème 1.19 nous dit que $t \mapsto \mathcal{O}_\beta(t)$ est une application bijective de $[0, 1[$ dans l'ensemble des suites d'entiers vérifiant (1.11). Ainsi, si β est tel que $\mathcal{O}_\beta(1) = \omega_\lambda(\infty)$, on obtient par composition une application bijective ϕ_λ envoyant $x \in [0, \infty[$ sur $t \in [0, 1[$, où $\mathcal{O}_\beta(t) = \omega_\lambda(x)$. Le diagramme suivant commute :

$$\begin{array}{ccc} x \in [0, \infty[& \xrightarrow{T_\lambda} & T_\lambda x \in [0, \infty[\\ \phi_\lambda \downarrow & & \downarrow \phi_\lambda \\ t \in [0, 1[& \xrightarrow{S_\beta} & S_\beta t \in [0, 1[\end{array}$$

FIGURE 1.12 –

Théorème 1.21. Pour tout $\lambda \in]0, 2[$, il existe un unique $\beta > 1$ et un homéomorphisme $\phi_\lambda : [0, \infty[\rightarrow [0, 1[$ conjuguant $([0, \infty[, T_\lambda)$ et $([0, 1[, S_\beta)$.

Nous avons déjà vu dans la section 1.3.1 que dans le cas où $\lambda = 1$, il existait un lien entre les suites de Fibonacci aléatoires et la fonction Question Mark de Minkowski (voir (1.6)). Dans notre contexte, ce lien se traduit par

$$\phi_1(x) = \begin{cases} \frac{1}{2}?(x) & \text{si } 0 \leq x \leq 1, \\ \frac{1}{2} + \frac{1}{2}?(1 - 1/x) & \text{sinon.} \end{cases}$$

Comme l'entropie topologique du β -shift est $\log \beta$ (voir *e.g.* [36]), on peut déduire du théorème 1.21 que, pour tout $\lambda \in]0, 2[$, l'entropie topologique de la transformation T_λ est $\log \beta(\lambda)$, où $\beta(\lambda)$ est défini dans le théorème précédent.

Théorème 1.22 (Propriétés de $\lambda \mapsto \beta$). L'application $\lambda \in]0, 2[\mapsto \beta(\lambda) \in]1, \infty[$ est strictement croissante et continue. De plus, $\beta(\lambda_k) = k - 1$ pour tout entier $k \geq 3$ et $\beta(1/\sqrt{k})$ est la plus grande racine de $1 = \frac{1}{\beta} + \frac{1}{\beta^k}$ pour tout entier $k \geq 2$. L'application $\lambda \mapsto \beta(\lambda)$ n'est donc pas analytique.

(Voir Figure 1.13.)

De nombreux travaux se sont intéressés aux connexions entre les propriétés algébriques de β et les propriétés dynamiques du β -shift associé (voir *e.g.* [4]). Nous ne savons pas si elles sont aussi connectées aux propriétés algébriques du λ correspondant. Les valeurs particulières de λ que nous avons étudiées sont algébriques et correspondent à des β algébriques, mais est-ce toujours le cas ?

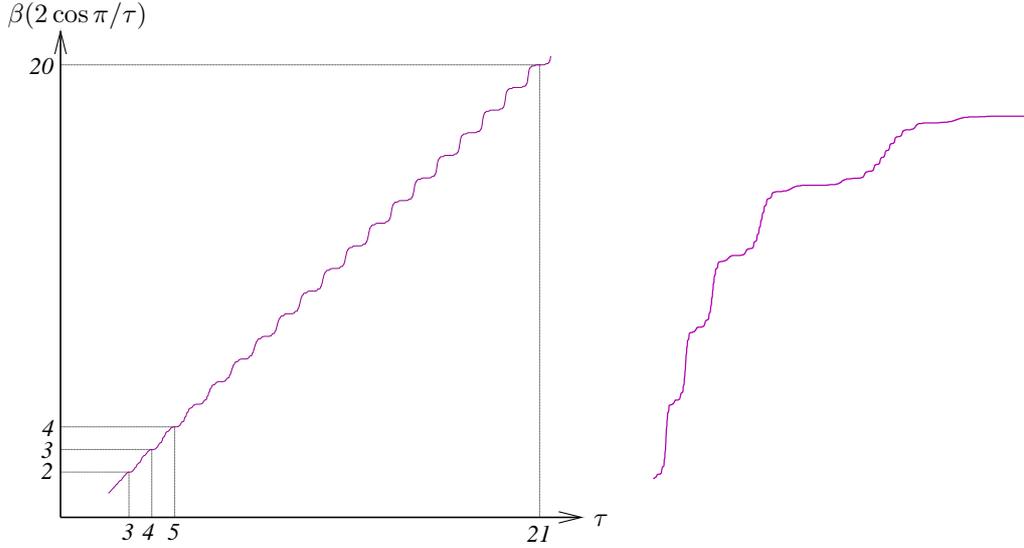


FIGURE 1.13 – À gauche, graphe de $\tau \mapsto \beta(2 \cos(\pi/\tau))$. À droite, détail du graphe pour $10.5 \leq \tau \leq 11$. Il semble que le graphe contienne une certaine auto-similarité qui suggère que l'application n'est pas différentiable.

L'algorithme décrit dans la section 1.6.3 permettant d'obtenir le β -développement d'un réel entre 0 et 1 en itérant S_β est parfois appelé *algorithme glouton*, car à chaque étape il utilise le plus grand entier possible pour le développement. Mais pour $\beta > 1$, il y a en général de multiples façons de développer un réel $t \in [0, 1[$ sous la forme

$$t = \sum_{n \geq 0} \frac{t_n}{\beta^{n+1}},$$

où $t_n \in \{0, 1, \dots, \lfloor \beta \rfloor\}$. On peut par exemple considérer le β -développement de t le plus petit pour l'ordre lexicographique, obtenu en prenant à chaque étape l'entier le plus petit possible et correspondant à l'itération de la "*lazy map*" (voir [11] et les références incluses). Entre ces deux extrêmes, il existe tout un continuum de β -développements, dont les *β -développements aléatoires* étudiés dans [11, 12]. Pouvons nous faire le lien entre ces différents β -développements et des développements en λ -fractions continus particuliers ?

Chapitre 2

La loi de Benford

2.1 Introduction

Étant donné un nombre réel $x > 0$, on définit la *mantisse* de x comme l'unique réel $\mathcal{M}(x) \in [1, 10[$ tel que l'on puisse trouver un entier k vérifiant

$$x = \mathcal{M}(x)10^k.$$

La loi de Benford décrit précisément la distribution statistique de la mantisse des nombres rencontrés dans la vie quotidienne : selon cette loi, la proportion des nombres x vérifiant $\mathcal{M}(x) \in [a, b[$ vaut, pour tous $1 \leq a < b \leq 10$

$$P_{Benford}(\mathcal{M}(x) \in [a, b[) = \log_{10} b - \log_{10} a. \quad (2.1)$$

Cette loi est souvent plus connue dans le cas particulier où $[a, b[= [i, i + 1[$ pour $i \in \{1, \dots, 9\}$: elle nous dit alors que la proportion des nombres dont le premier chiffre significatif est i est donnée par $\log_{10}(1 + \frac{1}{i})$.

Bien sûr le terme « proportion » n'est pas ici mathématiquement bien défini ; il s'agit de décrire la distribution observée dans un grand nombre de données empiriques. Le premier à mettre en évidence ce phénomène, qu'il décrit en 1881 dans [78], est l'astronome Simon Newcomb. Ce dernier s'était étonné que les tables de logarithmes se déchiraient plus vite au niveau des premières pages que des dernières. Le physicien Franck Benford, qui ne connaissait pas l'article de Newcomb, fit la même découverte 57 ans plus tard, et publia lui aussi un article sur le sujet. Ce dernier ayant eu plus de retentissement que celui de Newcomb, cette formule porte désormais le nom de *loi de Benford*.

Il est clair que de nombreuses données empiriques ne suivent pas du tout la loi de Benford. En son temps, Benford avait testé sa loi sur un nombre considérable d'observations, « *aussi large que le temps et l'énergie humainement disponible le permettent* », d'origines diverses : il avait collecté plus de 20 000 nombres provenant aussi bien de résultats de Base-Ball que de relevés d'hydrologie. Dans son article, la moitié des listes considérées s'écartent significativement de la loi prévue. Mais lorsqu'il cumule toutes les données relevées, la ressemblance devient frappante.

La loi de Benford a suscité depuis sa découverte un grand nombre de publications chez les scientifiques. Parmi celles-ci, les articles de mathématiques cherchent essentiellement à répondre à deux questions :

- Quelles conditions générales peuvent expliquer l'apparition de la loi de Benford ?
- Pourquoi la plupart des données empiriques vérifient-elles approximativement cette loi ?

La loi de Benford est par exemple la seule loi sur les mantisses invariante par changement d'échelle [83], ou encore la seule loi sur les mantisses invariante par changement de base [33]. L'article de revue de R. A. Raimi [85] contient des interprétations probabilistes de la loi de Benford : ainsi, la mantisse du produit de variables aléatoires indépendantes converge vers la loi de

Benford (voir aussi [8]). D'autres articles donnent des conditions sur la densité d'une variable aléatoire pour que sa mantisse suive une loi proche de celle de Benford [16].

Une idée plus ancienne, proposée par B.J. Flehinger, consiste à vérifier la loi de Benford pour la série de données la plus naturelle qui soit : l'ensemble des nombres entiers naturels. Ceci n'est pas tout à fait trivial, car la proportion d'entiers entre 1 et N de premier chiffre significatif d n'a pas de limite quand $N \rightarrow \infty$: elle oscille sur des périodes de plus en plus longues entre $d/9$ et $10d/9(d+1)$ (voir Figure 2.1). L'idée de Flehinger fut de chercher une limite en itérant le procédé de moyenne de Cesàro. En définissant récursivement

$$P_N^1(d) := \frac{1}{N} \sum_{n=1}^N \mathbb{1}_{[d, d+1[}(\mathcal{M}(n)) \quad \text{et} \quad P_N^{k+1}(d) := \frac{1}{N} \sum_{n=1}^N P_n^k(d),$$

Flehinger [26] prouve que

$$\lim_{k \rightarrow \infty} \limsup_{N \rightarrow \infty} P_N^k(d) = \lim_{k \rightarrow \infty} \liminf_{N \rightarrow \infty} P_N^k(d) = \log_{10} \left(1 + \frac{1}{d} \right), \quad (2.2)$$

ce qui est la proportion prédite par la loi de Benford. Ce résultat fut ensuite généralisé par D.E. Knuth [58] à la distribution de la mantisse des nombres entiers (c'est-à-dire pour tous les chiffres significatifs).

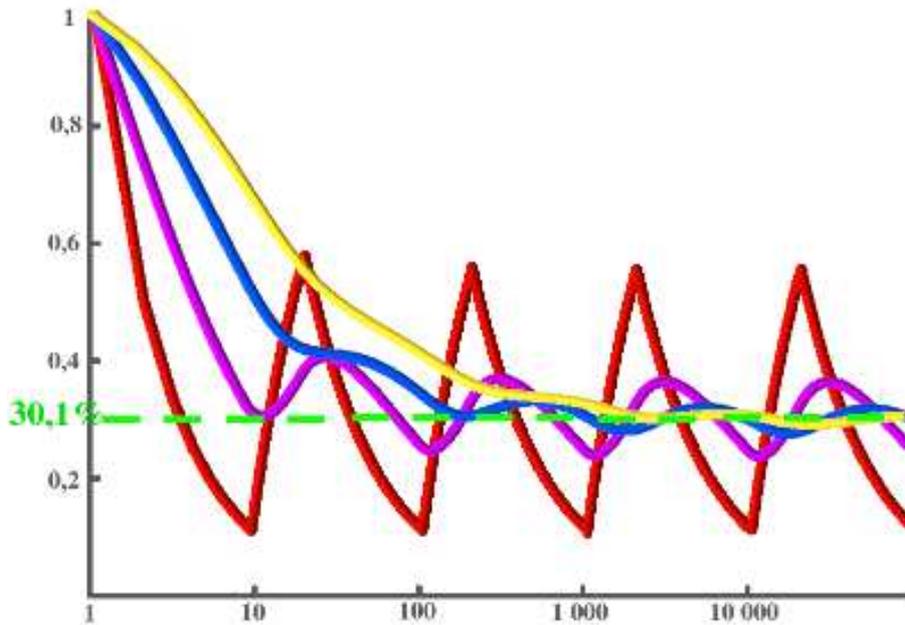


FIGURE 2.1 – La proportion $P_N^1(1)$ d'entiers entre 1 et N de premier chiffre significatif 1 (en rouge) oscille entre $1/9$ et $5/9$. En itérant le procédé de moyenne de Cesàro, on obtient $P_N^2(1)$ (en violet), $P_N^3(1)$ (en bleu) et $P_N^4(1)$ (en jaune), qui oscillent de moins en moins.

En dépit de son titre (*On the probability that a random integer has initial digit A*), l'article de Flehinger ne présente aucun contenu probabiliste : à la fois l'énoncé du théorème principal et les outils utilisés pour le prouver ne relèvent que du cadre déterministe. Une bonne raison pour cela est qu'il n'est pas possible de tirer au hasard un entier de manière uniforme dans l'ensemble de tous les entiers naturels. Notre première motivation fut donc d'interpréter le théorème de Flehinger dans un cadre probabiliste. En collaboration avec Thierry de la Rue [43], nous avons donné une nouvelle justification probabiliste de la loi de Benford, en montrant qu'elle apparaît

naturellement lorsque l'on considère des mélanges de lois uniformes. Cela nous a permis de donner une preuve probabiliste du résultat de Flehinger-Knuth et même la vitesse exponentielle dans la convergence des moyennes itérées. Avec Rita Giuliano (Pise, Italie), nous avons étendu ces résultats en considérant dans [38] des variables aléatoires continues, pas forcément uniformes. Ces travaux sont exposés dans la section 2.2.

Enfin, en nous inspirant des idées précédentes, nous avons réalisé avec Thierry de la Rue un travail qui tente de placer des énoncés de type Flehinger dans un véritable cadre probabiliste [47]. Cela nous a amenés à étudier diverses notions de densités pour des sous-ensembles infinis d'entiers naturels et à donner une interprétation probabiliste de certaines d'entre elles. Nous évoquerons ceci dans la section 2.3.

Notons que tous ces résultats sont donnés en base 10, mais s'adaptent automatiquement à d'autres bases.

2.2 Interprétation probabiliste

2.2.1 Heuristique

Lors de son étude, Benford avait collecté des milliers de nombres d'origines diverses. Selon leur origine, ceux-ci suivaient plus ou moins bien la loi, mais il avait remarqué qu'en les regroupant, les proportions obtenues étaient beaucoup plus proches de la loi prédite. En fait, plus les sources étaient variées et mieux on approchait la loi. Nous avons proposé dans [43] de modéliser les différentes origines des données collectées par des variables aléatoires uniformément distribuées sur $[0, S]$, où le maximum S inconnu dépend de l'origine.

Quelle sera la loi de la mantisse dans un tel modèle? Soit X une variable aléatoire (qui représente les données collectées) uniforme dans $[0, S]$, où $S > 0$ est une variable aléatoire (représentant le maximum dépendant de l'origine des données). On montre facilement que la loi de $\mathcal{M}(X)$ ne dépend que de $\mathcal{M}(S)$. Pour connaître la loi de $\mathcal{M}(X)$, il faut donc décider de la loi de la mantisse de S . Supposons que lorsqu'on collecte un grand nombre de données d'origines diverses, leur mantisse suit une distribution fixée μ . Comme leur maxima S proviennent eux-même de sources variées, on peut penser que leur mantisse suit la même distribution μ . On obtient alors l'équation suivante qui devrait être vérifiée par la loi μ de la mantisse de X : pour tout $t \in [1, 10[$,

$$\mu([1, t]) = \int_1^{10} P(\mathcal{M}(X) \leq t \mid \mathcal{M}(S) = a) d\mu(a), \quad (2.3)$$

Théorème 2.1. *L'équation (2.3) caractérise la loi de Benford. Autrement dit, la loi de Benford est l'unique mesure de probabilité μ sur $[1, 10[$ telle que, si $\mathcal{M}(S)$ suit la loi μ et si X sachant S est une variable aléatoire uniforme sur $[0, S]$, alors $\mathcal{M}(X)$ suit aussi la loi μ .*

Nous avons mentionné la caractérisation de la loi de Benford par l'invariance par changement d'échelle mise en évidence par Roger S. Pinkham [83] en 1961 (voir aussi [85, 33]). L'idée heuristique est la suivante : s'il existe une loi universelle régissant la mantisse des nombres réels S rencontrés dans la vie quotidienne, celle-ci ne doit pas dépendre de l'unité de mesure. On peut donc penser que pour tout $y > 0$, $\mathcal{M}(S)$ et $\mathcal{M}(yS)$ suivent la même loi. Et en effet, la seule loi vérifiant cette propriété est la loi de Benford.

Le théorème 2.1 peut être vu comme une généralisation de cette idée. En effet, une variable aléatoire uniforme sur $[0, S]$ peut être considérée comme le produit de S et d'une variable aléatoire uniforme sur $[0, 1]$. On peut donc traduire le théorème 2.1 sous la forme suivante : soit Y une variable aléatoire uniforme sur $[0, 1]$; si $\mathcal{M}(S)$ et $\mathcal{M}(YS)$ suivent la même loi, alors cette loi est la loi de Benford. Dans [38], nous montrons que cela reste vrai lorsque Y est une variable aléatoire continue sur $[0, 1]$, pas nécessairement uniforme.

2.2.2 Vitesse de convergence d'une chaîne de Markov

Notre approche conduit naturellement à considérer une chaîne de Markov (M_n) à valeurs dans $[1, 10[$, telle que M_{n+1} conditionnée à M_n suit la loi de la mantisse d'une variable aléatoire continue sur $[0, M_n]$ et dont l'unique probabilité invariante est la loi de Benford. Notons que cela revient à considérer une chaîne de Markov définie comme la mantisse d'un produit de variables aléatoires indépendantes à valeurs dans $[0, 1]$ de même loi continue ν . Le problème de la vitesse de convergence vers la loi de Benford de la mantisse d'un produit de variables indépendantes a été intensément étudié (voir e.g. [8], [73], [37] et les références qui s'y trouvent). L'intérêt de la proposition 2.3 ci-dessous est que la vitesse de convergence obtenue est exprimée explicitement en fonction de la loi ν . Mais pour cela, nous avons besoin d'une hypothèse technique sur cette loi : soit

$$H(u, t) := \nu \left(\bigcup_{k \geq 1} \left[\frac{u}{10^k}, \frac{ut}{10^k} \right] \right) \quad \forall u \in [0, 1], \forall t \in [1, 10]. \quad (2.4)$$

Supposons que

$$u \in [0, 1] \mapsto H(u, t) \text{ est strictement croissante pour tout } t \in [1, 10]. \quad (2.5)$$

Alors H est une fonction de répartition dont la première marginale a loi ν et la seconde marginale a la même loi que la mantisse d'une variable aléatoire de loi ν .

Exemple 2.2. Soit ν absolument continue par rapport à la mesure de Lebesgue, de densité $f(x) = \alpha x^{\alpha-1} 1_{[0,1]}(x)$, pour $\alpha > 0$ (le cas $\alpha = 1$ correspond à la mesure uniforme). Alors pour $u \in [0, 1]$ et $t \in [1, 10]$,

$$H(u, t) = \frac{u^\alpha (t^\alpha - 1)}{10^\alpha - 1},$$

qui est strictement croissante en u . De plus, H est la fonction de répartition d'une mesure produit dont la première marginale a densité f , et la seconde $\alpha t^{\alpha-1} / (10^\alpha - 1)$ sur $[1, 10]$. Nous ne savons pas s'il existe d'autres densités f telles que H corresponde à une mesure produit.

En utilisant une méthode développée dans [25], nous construisons explicitement la chaîne de Markov (M_n^α) partant de $M_0^\alpha = a$ et montrons par des techniques de couplages la proposition suivante (qui donne en particulier une autre preuve du théorème 2.1).

Proposition 2.3. Supposons que ν vérifie (2.5). Alors pour tout ensemble mesurable $B \subseteq [1, 10]$,

$$|P(M_n^\alpha \in B) - P_{Benford}(B)| \leq \nu \left(\left[\frac{1}{10}, 1 \right] \right)^n.$$

En particulier, si $\nu \left(\left[\frac{1}{10}, 1 \right] \right) < 1$, M_n^α converge exponentiellement vite vers la loi de Benford.

L'hypothèse (2.5) faite sur ν est cruciale pour la validité de la proposition précédente, même si la mantisse du produit de variables aléatoires i.i.d. de loi continue ν converge vers la loi de Benford. Certaines situations peuvent néanmoins être traitées même lorsque la mesure ν ne vérifie pas (2.5) : par exemple, si ν est la mesure de probabilité uniforme sur $[0, 1/10]$, alors on a bien sûr $\nu([1/10, 1]) = 0$ et $P(M_n^\alpha \in B) \neq P_{Benford}(B)$; mais puisque le produit de variables aléatoires i.i.d. de loi uniforme sur $[0, 1/10]$ a la même mantisse que le produit de v.a. i.i.d. de loi uniforme sur $[0, 1]$ (qui comme on l'a vu dans l'exemple 2.2 vérifie bien (2.5)), on peut quand même utiliser le théorème pour estimer la vitesse de convergence.

Bien sûr, si $\nu([1/10, 1]) = 0$, alors ν ne vérifie pas (2.5), à moins que la mantisse d'une variable aléatoire de loi ν suive la loi de Benford, et dans ce cas $P(M_n^\alpha \in B) = P_{Benford}(B)$ (c'est par exemple le cas de la mesure de densité $\frac{1}{x} 1_{[1/100, 1/10]}$ pour laquelle $H(u, t) = \log_{10}((10u \wedge 1)t) 1_{\{ut > 1/10\}}$). Mais on peut trouver des mesures ν vérifiant (2.5) avec $\nu([1/10, 1])$ arbitrairement proche de 0 et telle que $P(M_n^\alpha \in B) \neq P_{Benford}(B)$: la famille de mesure de l'exemple 2.2 vérifie $\nu([1/10, 1]) = 1 - (1/10)^\alpha$, qui tend vers 0 quand $\alpha \rightarrow 0$.

L'hypothèse (2.5) est utile dans notre démonstration car elle nous a permis de coupler les lois de $\mathcal{M}(mY)$, où Y suit la loi ν , pour tout $m \in [1, 10[$. Dans le cas où ν admet une densité f par rapport à la mesure de Lebesgue, on arrive au même type de résultat en couplant les lois de mY pour tout $m \in [1, 10[$ si on fait par exemple l'hypothèse suivante

$$A_\nu := \int_0^1 \left[\min_{1 \leq m < 10} \frac{1}{m} f\left(\frac{x}{m}\right) \right] dx > 0. \quad (2.6)$$

En effet, pour tout $m \in [1, 10[$, la densité de mY est $f_m(x) := \frac{1}{m} f\left(\frac{x}{m}\right)$, qui peut être décomposée sous la forme

$$f_m(x) = (1 - A_\nu)g_m(x) + A_\nu \frac{\phi(x)}{A_\nu}, \quad \text{où } \phi(x) := \min_{1 \leq m < 10} \frac{1}{m} f\left(\frac{x}{m}\right) \text{ et } g_m(x) := \frac{f_m(x) - \phi(x)}{1 - A_\nu}. \quad (2.7)$$

Comme g_m et $\phi(x)/A_\nu$ sont des densités, nous pouvons aussi construire explicitement la chaîne de Markov $(M_n)_{n \geq 1}$ et obtenir la vitesse de convergence de la chaîne vers sa mesure invariante $P_{Benford}$, de la façon suivante.

Soit $(X_n)_{n \geq 1}$ une suite i.i.d. de variables de Bernoulli de paramètre A_ν et soit $(V_n)_{n \geq 1}$ une suite i.i.d. de variables aléatoires de densité $\frac{1}{A_\nu} \phi$ indépendantes de $(X_n)_{n \geq 1}$. Pour tout m , soit encore $(U_n^{(m)})_{n \geq 1}$ une suite i.i.d. de variables aléatoires de densité g_m , indépendantes de $(X_n, V_n)_{n \geq 1}$. On définit alors $(M_n^a)_{n \geq 0}$ par $M_0^a = a \in [1, 10)$ et

$$\forall n \geq 1, \quad M_n^a := \begin{cases} \mathcal{M}(V_n) & \text{si } X_n \leq A_\nu \\ \mathcal{M}\left(U_n^{(M_{n-1}^a)}\right) & \text{si } X_n > A_\nu. \end{cases}$$

Proposition 2.4. *Supposons que ν admet une densité f par rapport à la mesure de Lebesgue qui vérifie (2.6). Alors pour tout ensemble mesurable $B \subseteq [1, 10)$,*

$$|P(M_n^a \in B) - P_{Benford}(B)| \leq (1 - A_\nu)^n.$$

Démonstration. Nous utilisons la même technique de couplage que pour la proposition 2.3. Considérons deux chaînes de Markov (M_n^a) et (M_n^b) , partant respectivement de a et b , et construites comme décrit précédemment (avec les mêmes suites $(X_n)_{n \geq 1}$, $(V_n)_{n \geq 1}$ et $(U_n^{(m)})_{n \geq 1}$).

Soit $\tau^{a,b}$ le temps de couplage, c'est-à-dire le premier instant où les deux chaînes M_n^a et M_n^b se rencontrent :

$$\tau^{a,b} := \inf\{n \geq 1 | M_n^a = M_n^b\}.$$

On a

$$P(\tau^{a,b} > n) \leq P(X_k > A_\nu, \forall 1 \leq k \leq n) = (1 - A_\nu)^n. \quad (2.8)$$

Donc, pour tout $B \subset [1, 10[$, comme $P_{Benford}$ est invariante pour $(M_n^b)_n$,

$$\begin{aligned} |P(M_n^a \in B) - P_{Benford}(B)| &= \left| \int_1^{10} (P(M_n^a \in B) - P(M_n^b \in B)) dP_{Benford}(b) \right| \\ &\leq \int_1^{10} dP_{Benford}(b) E[|1_{M_n^a \in B} - 1_{M_n^b \in B}|] \\ &\leq \int_1^{10} dP_{Benford}(b) E[1_{\tau^{a,b} > n}] \leq (1 - A_\nu)^n. \end{aligned}$$

□

Lorsque l'application $x \mapsto xf(x)$ est croissante sur $[0, 1]$, les deux hypothèses (2.5) et (2.6) sont vérifiées et les propositions 2.3 et 2.4 donnent le même résultat. En effet, on a dans ce cas, $\phi(x) = \frac{1}{10} f\left(\frac{x}{10}\right) 1_{[0,1]}(x)$ et $1 - A_\nu = 1 - \int_0^{1/10} f(x) dx = \nu([1/10, 1])$.

La proposition précédente permet par contre de traiter le cas de la probabilité uniforme sur $[0, 1/10]$ directement (on obtient $1 - A_\nu = 9/10$), et même certains cas inaccessibles avec la proposition 2.3, comme par exemple la loi de densité $f(x) = 2(1 - x)$ sur $[0, 1]$ qui ne vérifie pas (2.5). Nous n'avons par contre pas trouvé d'exemple de loi à densité vérifiant (2.5) mais pas (2.6) ou vérifiant les deux hypothèses mais donnant des vitesses de convergence différentes.

2.2.3 Généralisation du résultat de Flehinger

En corollaire de la proposition 2.3, nous obtenons une preuve probabiliste simple du théorème de Flehinger (2.2), ainsi qu'une vitesse de convergence dans le théorème. En effet, lorsque ν est la loi uniforme sur $[0, 1]$, $P_n^k(i)$ correspond à k étapes de notre chaîne de Markov, au sens où pour tout $k \geq 1$ et tout $a \in [1, 10[$,

$$Q^k(a, i) := \lim_{j \rightarrow \infty} P_{\lfloor a10^j \rfloor}^k(i) = P(M_k^a < i + 1),$$

où M_k^a est la chaîne de Markov débutant en a . Heuristiquement, cela repose sur l'idée que l'itération des moyennes de Césaire peut être vu comme un tirage récursif de variables aléatoires uniformes. Comme $\liminf_n P_n^k(i) = \min_{1 \leq a < 10} Q^k(a, i)$ et $\limsup_n P_n^k(i) = \max_{1 \leq a < 10} Q^k(a, i)$, la convergence exponentielle de la chaîne vers la loi de Benford nous permet donc de retrouver le théorème de Flehinger et d'obtenir une estimation de la vitesse de convergence dans (2.2). Enfin, une adaptation directe de notre argument donne la vitesse de convergence dans le théorème de Knuth qui concerne la distribution de toute la mantisse des entiers.

La proposition 2.3 nous a permis d'obtenir une généralisation de ce résultat dans laquelle la moyenne de Césaire itérée est remplacée pour tout entier $N \geq 1$ par

$$\bar{g}_0(N) := g(N) \quad \text{et} \quad \forall m \geq 1, \quad \bar{g}_m(N) := \sum_{k=1}^N \bar{g}_{m-1}(k) \nu_m \left(\left[\frac{k-1}{N}, \frac{k}{N} \right] \right),$$

où (ν_m) une suite de distributions à valeurs dans $[0, 1]$ vérifiant (2.5) et g_0 est une fonction continue sur \mathbb{R}_+^* , Riemann intégrable sur $[1, 10]$ telle que

$$g(10x) = g(x), \quad \forall x \in \mathbb{R}_+^*.$$

On montre alors que pour tout $m \geq 1$

$$\left| \liminf_{N \rightarrow \infty} \bar{g}_m(N) - \int_1^{10} g \, dP_{Benford} \right| \leq \prod_{i=1}^m \nu_i \left(\left[\frac{1}{10}, 1 \right] \right),$$

et la même chose est valable lorsque la \liminf est remplacée par la \limsup .

2.3 Suites d'entiers aléatoires

2.3.1 URI-set

En nous inspirant des idées précédentes, nous avons réalisé avec Thierry de la Rue un travail qui tente de placer des énoncés de type Flehinger dans un véritable cadre probabiliste [47]. Une réalisation de la chaîne de Markov décrite dans la section 2.2.2 peut être vue comme un ensemble aléatoire de réels $\mathbf{M} := \{M_n, n \geq 0\}$. Nous souhaitons imiter la structure de cet ensemble mais avec des entiers. Pour tout $t > 0$, $\max(\mathbf{M} \cap]0, t])$ est uniformément distribué dans $]0, t]$ et indépendant de $\mathbf{M} \cap]t, \infty[$ (sachant que $\mathbf{M} \cap]t, \infty[\neq \emptyset$). On cherche donc un ensemble aléatoire \mathbf{E} d'entiers tel que

- (U) $\forall n \geq 1$, $\max(\mathbf{E} \cap \{1, \dots, n\})$ uniformément distribué dans $\{1, \dots, n\}$;
- (I) $\forall n \geq 1$, $\max(\mathbf{E} \cap \{1, \dots, n\})$ indépendant de $\mathbf{E} \cap \{n+1, n+2, \dots\}$.

La solution est d'utiliser des variables aléatoires de Bernoulli indépendantes : pour chaque entier $n \geq 1$, nous décidons avec probabilité $1/n$ (indépendamment de ce qui se passe pour tous les autres entiers) si n doit être placé dans \mathbf{E} . Nous appelons \mathbf{E} un *ensemble aléatoire d'entiers uniformes* (ou encore *URI-set*). La donnée d'un tel ensemble permet de définir des variables aléatoires N_k , $k \geq 1$, en ordonnant les éléments de \mathbf{E} qui sont presque sûrement en nombre infini :

$$\mathbf{E} = \{N_1 = 1 < N_2 < \dots < N_k < \dots\}.$$

2.3.2 Résultats de type *loi de Benford*

Un premier résultat montre que les éléments d'un URI-set satisfont en moyenne la loi de Benford :

Théorème 2.5. *Pour tout $t \in [1, 10[$,*

$$\frac{1}{\ell} \sum_{k=1}^{\ell} \mathbb{1}_{\mathcal{M}(N_k) < t} \xrightarrow[\ell \rightarrow \infty]{p.s.} \log_{10} t.$$

Démonstration. L'idée de la démonstration, que nous esquissons ici, repose sur un couplage de l'URI-set et de son homologue continu. Soit ξ un processus de Poisson sur \mathbb{R}_+ de densité $\frac{dx}{x}$: notons ξ_B le nombre de points dans un borélien B . Cela signifie que

- ξ_B suit une loi de Poisson de paramètre $\int_B \frac{dx}{x}$;
- si $(B_i)_{1 \leq i \leq \ell}$ sont des ensembles 2 à 2 disjoints, $\xi_{B_1}, \dots, \xi_{B_\ell}$ sont indépendants.

Considérons alors l'ensemble

$$\mathbf{E} := \{n \geq 1 : \xi_{]n-1, n]} \geq 1\}.$$

On a $\mathbb{P}(n \in \mathbf{E}) = 1/n$ pour tout n et la propriété d'indépendance. Cet ensemble \mathbf{E} est donc bien un URI-set.

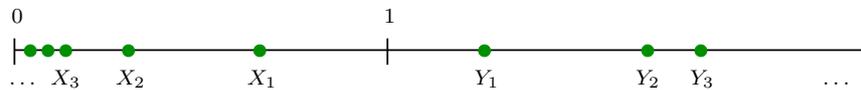


FIGURE 2.2 – Une réalisation du processus de Poisson sur \mathbb{R}_+ de densité $\frac{dx}{x}$.

Notons $\dots < X_n < \dots < X_1 < 1 < Y_1 < \dots < Y_n < \dots$ les points du processus de Poisson (voir Figure 2.3.2). La variable aléatoire X_1 est uniformément distribuée sur $[0, 1]$, et comme $\mathbb{P}(X_{n+1} \leq t | X_n) = \mathbb{P}(\xi_{]t, X_n]} = 0) = t/X_n$, la variable aléatoire X_{n+1} , conditionnée à X_1, \dots, X_n , est uniformément distribuée dans $[0, X_n]$ pour tout n . Donc d'après la proposition 2.3 les mantisses $\mathcal{M}(X_n)$ constituent une chaîne de Markov dont l'unique loi invariante est la loi de Benford. On en déduit que pour tout $t \in [1, 10[$,

$$\frac{1}{\ell} \sum_{k=1}^{\ell} \mathbb{1}_{\mathcal{M}(X_k) < t} \xrightarrow[\ell \rightarrow \infty]{p.s.} \log_{10} t.$$

Par ailleurs, la loi de Benford étant stable par passage à l'inverse, $(Y_n)_{n \geq 1}$ et $(X_n)_{n \geq 1}$ suivent la même loi. Enfin, avec probabilité 1, il y a seulement un nombre fini de n tels que $\xi_{]n-1, n]} \geq 2$. Cela implique qu'il existe p.s. un entier r tel que, pour tout k assez grand, $0 \leq N_k - Y_{k+r} < 1$. Comme $N_k \rightarrow \infty$ p.s., on a $|\mathcal{M}(N_k) - \mathcal{M}(Y_{k+r})| \rightarrow 0$ sauf si N_k est de la forme 10^p , ce qui n'arrive p.s. que pour un nombre fini de k . \square

Nous prouvons également la généralisation du théorème précédent dans le cas où l'on regarde simultanément les mantisses d'un même nombre dans plusieurs bases de numération. Rappelons

que la mantisse en base b d'un réel positif x est l'unique réel $\mathcal{M}_b(x) \in [1, b[$ tel que $x = \mathcal{M}_b(x)b^k$ pour $k \in \mathbb{Z}$. La loi de benford en base b est alors

$$P_{Benford}^b(\mathcal{M}_b(x) \in [1, t]) = \log_b t, \text{ pour tout } 1 \leq t < b.$$

Soit $(b_i)_{1 \leq i \leq \ell}$ des entiers positifs tels que :

$$\frac{s_1}{\ln b_1} + \dots + \frac{s_\ell}{\ln b_\ell} = 0 \text{ avec } (s_i)_{1 \leq i \leq \ell} \subset \mathbb{Z} \implies s_i = 0 \text{ pour tout } i.$$

Alors pour tout $t_i \in [1, b_i[$,

$$\frac{1}{K} \sum_{k=1}^K \mathbb{1}_{\{\mathcal{M}_{b_i}(N_k) \leq t_i, \forall i\}} \xrightarrow[K \rightarrow \infty]{p.s.} \prod_{i=1}^{\ell} \log_{b_i} t_i.$$

Lorsque $\ell = 2$, notre condition est équivalente à l'indépendance multiplicative des bases : $b_1^{s_1} \dots b_\ell^{s_\ell} = 1$ avec $(s_i)_{1 \leq i \leq \ell} \subset \mathbb{Z}$ implique $s_i = 0$ pour tout i . Par contre, pour $\ell > 2$ il n'est pas clair que l'indépendance multiplicative des bases soit une condition suffisante pour obtenir le résultat. Cette question est liée la conjecture de Schanuel en théorie des nombres transcendants (voir [60], p. 30-31 ou [101]).

Un second résultat montre qu'asymptotiquement, le k -ième élément d'un URI-set vérifie la loi de Benford :

Théorème 2.6. *Pour tout $t \in [1, 10[$,*

$$\lim_{k \rightarrow \infty} \mathbb{P}(\mathcal{M}(N_k) < t) = \log_{10} t.$$

2.3.3 Liens entre différentes densités

Les résultats mentionnés ci-dessus conduisent à considérer les propriétés de moyennisation le long des éléments d'un URI-set dans le cadre général de l'étude des diverses notions de densités pour un sous-ensemble A infini d'entiers naturels. La plus simple de ces notions de densité est bien sûr la densité naturelle, définie lorsqu'elle existe par la limite quand $n \rightarrow \infty$ de la proportion d'éléments de A dans $\{1, \dots, n\}$. D'autres densités d'intérêt pour notre étude sont la densité H_∞ , qui est celle utilisée dans le théorème de Flehinger et qui est obtenue par itération du procédé de moyenne de Césàro, ainsi que la densité logarithmique, définie quand elle existe par

$$d_{\log}(A) := \lim_{n \rightarrow \infty} \frac{1}{\ln n} \sum_{k=1}^n \frac{\mathbb{1}_A(k)}{k}.$$

Il est connu (voir [18]) que la densité H_∞ est une généralisation de la densité naturelle, et que la densité logarithmique est elle-même encore plus forte que la densité H_∞ . Nos travaux mettent en avant deux nouvelles notions de densité : l'*URI-densité*, qui est définie par la densité naturelle presque sûre de A dans un URI-set

$$d_{\text{uri}}(A) := \lim_{K \rightarrow \infty} (p.s.) \frac{1}{K} \sum_{k=1}^K \mathbb{1}_A(N_k),$$

et l'*URI-densité locale* définie par

$$d_{\text{uri loc}}(A) := \lim_{k \rightarrow \infty} \mathbb{P}(N_k \in A).$$

Nous prouvons que la notion d'URI-densité est en fait équivalente à la densité logarithmique (ce qui relie un de nos résultats à un travail de R.L. Duncan [21] et fournit une autre démonstration du théorème 2.5), et que l'URI-densité locale généralise la densité naturelle.

Une question importante qui reste ouverte est la place de l'URI-densité locale dans cette hiérarchie de notions de densité. Nous conjecturons que l'URI-densité est plus forte que l'URI-densité locale. Nous ne savons pas non plus si l'URI-densité locale est équivalente à la densité H_∞ utilisée par Flehinger. Cependant, nous avons une interprétation probabiliste de la densité H_∞ en terme d'URI-set. Considérons les éléments de $\mathbf{E}_n := \mathbf{E} \cap \{1, \dots, n\}$ dans l'ordre décroissant :

$$\mathbf{E}_n = \left\{ Z_1^{(n)} > Z_2^{(n)} > \dots > Z_{|\mathbf{E}_n|}^{(n)} = 1 \right\}.$$

Proposition 2.7 (Interprétation probabiliste de la densité H_∞). *Un ensemble $A \subset \mathbb{Z}_+$ admet une H_∞ -densité α si et seulement si*

$$\lim_{k \rightarrow \infty} \liminf_{n \rightarrow \infty} \mathbb{P} \left(Z_k^{(n)} \in A \right) = \lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P} \left(Z_k^{(n)} \in A \right) = \alpha.$$

2.3.4 Autres questions

De nombreux autres problèmes se posent autour de ces notions, notamment sur la bonne façon de définir une URI-densité conditionnelle, une URI-densité itérée, ainsi que sur les corrélations qui existent entre les éléments successifs d'un URI-set.

URI-densité conditionnelle

Soit $\mathbf{E} = \{N_1 < N_2 < \dots\}$ un URI-set et soit $P \subset \mathbb{Z}_+$ tel que $\sum_{p \in P} \frac{1}{p} = \infty$ (ce qui implique que $|P \cap \mathbf{E}| = \infty$ p.s). Comment interpréter l'URI-densité d'un ensemble A d'entiers sachant P ?

Une première possibilité est de faire une moyenne sur $P \cap \mathbf{E}$:

$$\frac{\sum_{k=1}^K 1_{P \cap A}(N_k)}{\sum_{k=1}^K 1_P(N_k)}$$

Une autre est d'ordonner les éléments de $P = \{p_1 < p_2 < \dots\}$, puis de moyenner le long de p_{N_k} :

$$\frac{1}{K} \sum_{k=1}^K 1_A(p_{N_k})$$

Nous ne savons pas si ces notions sont équivalentes.

URI-densité infinie

Nous avons vu que Flehinger s'était intéressée aux moyennes de Cesàro itérées. Plus tard, Diaconis a introduit dans [18] la log-sommation itérée : soit

$$L^1(A, n) := \frac{1}{\ln n} \sum_{j=1}^n \frac{1}{j} 1_A(j),$$

et pour tout $\ell \geq 2$

$$L^\ell(A, n) := \frac{1}{\ln n} \sum_{j=1}^n \frac{1}{j} L^{\ell-1}(A, j).$$

L'ensemble A admet une ℓ -ème log-densité α si la limite de $L^\ell(A, n)$ existe quand $n \rightarrow \infty$ et est égale à α . En fait, cette notion ne donne pas lieu à une nouvelle densité puisque Diaconis a montré que A admettait une ℓ -ème log-densité si et seulement si A admettait une log-densité (et, bien sûr, les deux coïncident). Il a alors proposé de définir la L_∞ -densité, qui étend la log-densité de la même manière que la densité H_∞ étend la densité naturelle : considérons $\lim_{\ell \rightarrow \infty} \liminf_{n \rightarrow \infty} L^\ell(A, n)$ et $\lim_{\ell \rightarrow \infty} \limsup_{n \rightarrow \infty} L^\ell(A, n)$. Si ces deux limites sont égales, leur valeur commune est appelée

L_∞ -densité de A . Comme montré dans [19], la L_∞ -densité est strictement plus forte que la log-densité.

Est-il aussi intéressant de définir une URI-densité itérée? Une définition pourrait être la suivante : soit $(\mathbf{E}^{(\ell)})_{\ell \geq 1}$ une suite d'URI-sets indépendants, et notons $N_1^{(\ell)} = 1 < N_2^{(\ell)} < \dots < N_k^{(\ell)} < \dots$ les éléments aléatoires de $\mathbf{E}^{(\ell)}$. On dira que $A \subset \mathbb{Z}_+$ admet une *URI-densité d'ordre 2* égale à α si

$$\frac{1}{n} \sum_{k=1}^n \mathbb{1}_A \left(N_{N_k^{(2)}}^{(1)} \right) \xrightarrow[n \rightarrow \infty]{p.s.} \alpha.$$

De même, on définit l'*URI-densité d'ordre ℓ* en moyennant le long de la sous-suite $N_{N_k^{(2)}}^{(1)}$. On

peut aussi introduire une URI-densité infinie en considérant les limsup et liminf presque sûre dans l'expression précédente, et en regardant si elle convergent vers la même limite lorsque $\ell \rightarrow \infty$.

Bien que l'URI-densité et la log-densité coïncident, le lien entre l'URI-densité itérée et la log-densité itérée n'est pas clair. L'URI-densité d'ordre fini peut-elle être strictement plus forte que l'URI-densité? Peut-on comparer les itérations infinies de l'URI-densité et de la log-densité?

URI-densité pour plusieurs ensembles

Soit $\mathbf{E} = \{N_1 < N_2 < \dots\}$ un URI-set et soient $(A_i)_{0 \leq i \leq \ell}$ des ensembles d'entiers de densités respectives $d(A_i)$. Que peut-on dire des corrélations qui existent entre les éléments successifs de \mathbf{E} ? A-t-on

$$\frac{1}{K} \sum_{k=1}^K \mathbb{1}_{A_0}(N_k) \dots \mathbb{1}_{A_\ell}(N_{k+\ell}) \xrightarrow[K \rightarrow \infty]{p.s.} \prod_{i=0}^{\ell} d(A_i)?$$

Nous pensons que cela est vrai lorsque d est la densité naturelle, mais cela est faux pour la URI-densité : en effet, si A_i est l'ensemble des entiers de premier chiffre significatif i (qui admet donc une URI-densité mais pas de densité naturelle), alors N_k et N_{k+1} ne sont pas indépendants. Mais que se passe-t-il lorsque par exemple $d(A_1)$ est la densité naturelle de A_1 et $d(A_2)$ est seulement l'URI-densité de A_2 ?

Chapitre 3

Quelques problèmes de théorie ergodique

3.1 La transformation Pascal-adique

3.1.1 Introduction

La notion de transformation adique a été introduite par Vershik ([98], [97]) comme un modèle combinatoire pour les transformations construites par *cutting and stacking* (découpage et empilage), une méthode de construction commune en théorie ergodique. Une transformation adique agit sur des chemins infinis dans certains graphes appelés *diagramme de Bratteli*, de manière transverse au shift usuel : seuls un nombre fini de pas de la trajectoire sont modifiés par la transformation. Tout comme la construction par cutting and stacking, la représentation d'un système comme une transformation adique est universelle, au sens où toute transformation ergodique préservant la mesure sur un espace de Lebesgue est mesurablement isomorphe à une transformation adique qui n'admet qu'une mesure invariante ergodique non atomique.

Des familles importantes de systèmes dynamiques admettent une représentation particulièrement simple sous forme de transformation adique : les odomètres et les substitutions sur un alphabet fini. Ces systèmes se représentent par des adiques correspondant à des diagrammes de Bratteli *stationnaires*, c'est-à-dire pour lesquels, au-delà de la racine, tous les niveaux ont le même nombre de sommets et se connectent de manière identique au niveau suivant.

Le niveau de complexité immédiatement supérieur pour le diagramme de Bratteli est représenté par un exemple extrêmement simple : le triangle de Pascal. Dans ce diagramme, le niveau n comporte $n + 1$ points (numérotés de 0 à n), et les connexions se font d'un sommet k du niveau n aux sommets k et $k + 1$ du niveau $n + 1$. La dynamique de la transformation adique associée, introduite également par Vershik comme l'un des exemples basiques de transformations adiques, se présente de la façon suivante. Ayant codé une trajectoire infinie dans le triangle de Pascal par une suite infinie de 0 et de 1 (voir Figure 3.1.1), la transformation T , appelée *transformation Pascal-adique*, consiste à modifier l'ordre des symboles jusqu'à la première occurrence du motif 01 comme l'indique la formule suivante :

$$T(1^n 0^m 01x) = 0^m 1^n 10x.$$

Il est à noter que la transformation n'est définie que sur $\{0, 1\}^{\mathbb{N}}$ privé d'un nombre dénombrable de points (les trajectoires dans lesquelles le motif 01 n'apparaît pas).

Malgré la simplicité de sa construction, l'étude des propriétés dynamiques de cette transformation pose immédiatement des questions difficiles à résoudre, si bien qu'aujourd'hui encore, très peu de choses sont connues sur la Pascal-adique (voir [69, 71]).

Puisque le nombre de 0 et de 1 dans le préfixe de la suite est conservé par la transformation, il est clair que pour tout p entre 0 et 1, la mesure produit μ_p sous laquelle chaque coordonnée de

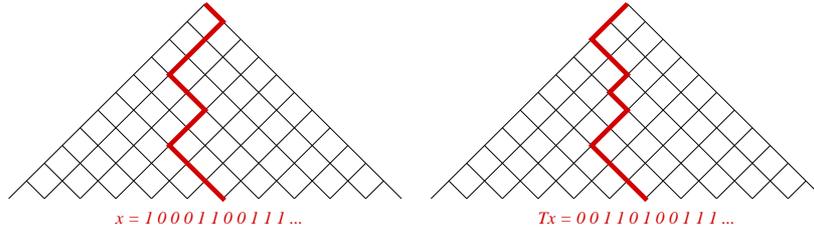


FIGURE 3.1 – Action de la transformation Pascal-adique sur une trajectoire infinie dans le triangle de Pascal. Un pas de la trajectoire est codé par 0 (respectivement 1) si le sommet k du niveau n est relié au sommet k (respectivement $k + 1$) du niveau $n + 1$.

la suite vaut 1 avec probabilité p est invariante sous l'action de T . On peut montrer en fait que ces mesures μ_p sont exactement les composantes ergodiques de T . Il est également connu que la transformation Pascal-adique est d'entropie nulle (voir [69]).

Dans [44], nous avons réussi à montrer que la transformation Pascal-adique était lâchement Bernoulli. Avant d'expliquer cette notion, rappelons la construction par “cutting and stacking” de la transformation Pascal-adique de paramètre p (voir [69]) dont nous nous sommes servi dans nos travaux.

Considérons l'espace $X = [0, 1[$ muni de sa tribu borélienne \mathcal{A} et équipé de la mesure de Lebesgue μ . Soit $0 < p < 1$ fixé. On commence par diviser X en deux sous-intervalles $P_a := [0, p[$ et $P_b := [p, 1[$. Soit $\mathcal{P} := \{P_a, P_b\}$ la partition obtenue lors de cette première étape. P_a et P_b peuvent être vues comme des tours de Rokhlin “dégénérées” de hauteur 1, que l'on note respectivement τ_0^1 et τ_1^1 . À la seconde étape, P_a et P_b sont divisées en proportion $(p, 1 - p)$. La transformation T est définie sur la partie droite de P_a en l'envoyant linéairement sur la partie gauche de P_b (les deux intervalles ont bien la même largeur $p(1 - p)$). On obtient ainsi une collection de 3 tours de Rokhlin notées $\tau_0^2, \tau_1^2, \tau_2^2$, de hauteurs respectives 1, 2, 1 (voir Figure 3.2).

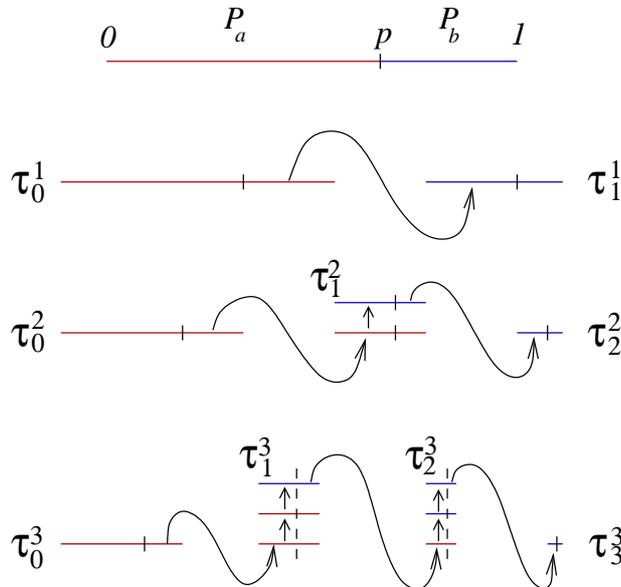


FIGURE 3.2 – Construction par “cutting and stacking” de la transformation Pascal-adique de paramètre p .

Après n étapes, on a $(n + 1)$ tours $\tau_0^n, \dots, \tau_n^n$ de hauteurs respectives $\binom{n}{0}, \dots, \binom{n}{n}$, telles que

τ_k^n est de largeur $p^{n-k}(1-p)^k$. On note F_k^n la base de la tour τ_k^n . À cette étape, la transformation T est définie partout sauf au sommet des tours. On subdivise alors chaque tour en proportion $(p, 1-p)$, et on définit T sur la partie droite du sommet de τ_k^n en l'envoyant linéairement sur la partie gauche de la base F_{k+1}^n de τ_{k+1}^n (qui ont toutes deux largeur $p^{n-k}(1-p)^{k+1}$).

En répétant cette construction, T est finalement définie presque partout, et il est clair que T préserve la mesure μ .

3.1.2 Lâche-bernoullicité

La notion de lâche-bernoullicité a été introduite par Feldman en 1976 ([24]), puis utilisée par Ornstein, Rudolph et Weiss ([79]) pour l'étude de l'équivalence au sens de Kakutani de transformations préservant la mesure. Dans le contexte d'une transformation d'entropie nulle, dire qu'une transformation T est lâchement Bernoulli revient à dire que T est isomorphe à une transformation induite par une rotation irrationnelle (c'est-à-dire équivalente au sens de Kakutani à une rotation irrationnelle).

La caractérisation de lâche-bernoullicité proposée par Feldman utilise la notion de “ \mathcal{P} -nom” d'un point x et la distance \bar{f} entre des mots finis.

Soit $\mathcal{P} = \{P_0, \dots, P_k\}$ une partition finie mesurable de l'espace de probabilité (X, \mathcal{A}, μ) . Pour $x \in X$, on pose $\mathcal{P}(x) := j \in \{0, \dots, k\}$ si $x \in P_j$. Pour $m < n$ dans \mathbb{Z} , le \mathcal{P} -nom de x (de m à n) est donné par

$$\mathcal{P}|_m^n(x) := j_m j_{m+1} \cdots j_n,$$

où, pour tout $m \leq i \leq n$, $j_i := \mathcal{P}(T^i x)$. Le \mathcal{P} -nom (entier) de x correspond à la suite doublement infinie $\mathcal{P}|_{-\infty}^{+\infty}(x)$.

Soit $V = v_1 \cdots v_\ell$ et $w = w_1 \cdots w_\ell$ deux mots de longueur ℓ sur le même alphabet. La distance \bar{f} entre v et w est donnée par

$$\bar{f}(v, w) := \frac{\ell - s}{\ell},$$

où s est le plus grand entier dans $\{0, \dots, \ell\}$ tel que l'on peut trouver $1 \leq i_1 < i_2 < \cdots < i_s \leq \ell$ et $1 \leq j_1 < j_2 < \cdots < j_s \leq \ell$ avec $v_{i_r} = w_{j_r}$ ($r = 1, \dots, s$).

Soit T une transformation préservant la mesure et d'entropie nulle sur (X, \mathcal{A}, μ) , et soit \mathcal{P} une partition finie mesurable de X . Le processus (\mathcal{P}, T) est dit *lâchement Bernoulli* (LB) si pour tout $\epsilon > 0$ et pour tout ℓ suffisamment grand, on peut trouver $A \subset X$ avec $\mu(A) > 1 - \epsilon$ tel que

$$\forall x, y \in A, \quad \bar{f}(\mathcal{P}|_0^\ell(x), \mathcal{P}|_0^\ell(y)) < \epsilon.$$

La transformation T est dite *lâchement Bernoulli* (LB) si (\mathcal{P}, T) est LB pour toute partition \mathcal{P} .

Nous avons montré dans [44] le théorème suivant :

Théorème 3.1. *La transformation Pascal-adique est lâchement Bernoulli.*

Nous allons seulement donner ici les grandes lignes de la démonstration. Notons que cette méthode a pu être adaptée par S. Bailey Frick [27] et X. Méla [70] à d'autres transformations adiques.

La première étape consiste à montrer qu'il suffit à un processus (\mathcal{P}, T) de vérifier la propriété plus faible suivante pour être LB : pour tout $\epsilon > 0$ et $\mu \otimes \mu$ -presque tout $(x, y) \in X \times X$, il existe un entier $\ell(x, y) \geq 1$ tel que

$$\bar{f}(\mathcal{P}|_0^{\ell(x,y)}(x), \mathcal{P}|_0^{\ell(x,y)}(y)) < \epsilon.$$

Enfin, notons qu'il n'y a pas besoin, pour montrer qu'une transformation est lâchement Bernoulli, de vérifier que (\mathcal{P}, T) est LB pour toute partition \mathcal{P} . En effet, pour prouver un tel résultat, il suffit de le vérifier pour une partition \mathcal{P} génératrice, i.e. telle que $\bigvee_{k=-\infty}^{+\infty} T^k \mathcal{P} = \mathcal{A}$.

Nous utilisons dans notre travail la partition $\{P_a, P_b\} = \{[0, p[, [p, 1[$ donnée par la première étape de la construction par “cutting and stacking”. Le fait qu'elle soit génératrice repose sur

l'idée suivante : pour tout $n \geq 1$, on définit les *blocs de base de niveau n* , $(B_{n,k})_{0 \leq k \leq n}$, qui sont des mots sur l'alphabet $\{a, b\}$, définis récursivement par $B_{n,0} := a$, $B_{n,n} := b$, et par la relation de concaténation

$$B_{n,k} := B_{n-1,k-1}B_{n-1,k} \quad \forall 1 \leq k \leq n-1. \quad (3.1)$$

Il est clair que le bloc $B_{n,k}$ est le \mathcal{P} -nom de longueur $\binom{n}{k}$ de tout point x se trouvant dans la base F_k^n de τ_k^n . En fait, le bloc de base $B_{n,k}$ caractérise la base de τ_k^n . Plus précisément, nous montrons que

$$\forall n \geq 2 \text{ et } 1 \leq k \leq n-1, \quad \mathcal{P}|_0^{\binom{n}{k}-1}(x) = B_{n,k} \text{ ssi } x \in F_k^n.$$

Or, pour tout $n \geq 1$ fixé, le \mathcal{P} -nom entier de tout point x est une concaténation de blocs de base de niveau n . Cette décomposition en blocs de base $B_{n,k}$ étant unique, connaître le \mathcal{P} -nom de x nous permet de dire pour tout n dans quelle tour, et dans quel étage de la tour de niveau n , se trouve le point x . Comme la partition \mathcal{Q}_n de X en étages de tours τ_k^n , $0 \leq k \leq n$, est constituée d'intervalles de largeur maximale $\max(p, 1-p)^n$ et que \mathcal{Q}_{n+1} est plus fine que \mathcal{Q}_n , on en déduit que $\bigvee_{n \geq 1} \mathcal{Q}_n = \mathcal{A}$.

Nous montrons ensuite que (\mathcal{P}, T) vérifie bien la propriété annoncée. Pour tout $x \in X$ et $n \geq 1$, notons $k_n(x) \in \{0, \dots, n\}$ le numéro de la tour de niveau n dans laquelle se trouve x ($x \in \tau_{k_n(x)}^n$). Soit $\epsilon > 0$, et r un entier tel que $(1-p)^r < \epsilon/2$. Notons que toute tour τ_k^n est un empilement de 2^r paquets qui sont des morceaux de tours de niveau $n-r$. En considérant $k_n(x)$ comme la somme de n variables aléatoires $(X_m)_{\{1 \leq m \leq n\}}$ i.i.d. de loi de Bernoulli de paramètre p , nous obtenons que pour $\mu \otimes \mu$ -presque tout (x, y) , on a

- $\frac{k_n(x)}{n} \xrightarrow{n \rightarrow +\infty} 1-p$;
- il existe une infinité d'entiers n tels que $k_n(x) = k_n(y)$, et x et y se trouvent dans le premier bloc de niveau $(n-r)$ de $\tau_{k_n(x)}^n$.

Considérons un tel n et notons $k = k_n(x)$. Comme $\frac{\binom{n-r}{k-r}}{\binom{n}{k}} \sim (1-p)^r$ quand $n \rightarrow +\infty$, pour n est assez grand, la hauteur $\binom{n-r}{k-r}$ du premier bloc de niveau $(n-r)$ de τ_k^n est petite par rapport à la hauteur $\binom{n}{k}$ de τ_k^n . Finalement, en posant $l := \binom{n-r}{k-r}$, les \mathcal{P} -noms $\mathcal{P}|_0^l(x)$ et $\mathcal{P}|_0^l(y)$ de x et de y commencent tous deux par un suffixe de $B_{n,k}$ de longueur plus grande que $(1-\epsilon/2)l$. On peut donc trouver une sous-suite commune de $\mathcal{P}|_0^l(x)$ et $\mathcal{P}|_0^l(y)$ de longueur plus grande que $(1-\epsilon)l$, telle que

$$\bar{f}(\mathcal{P}|_0^l(x), \mathcal{P}|_0^l(y)) < \epsilon.$$

3.1.3 Théorèmes limites : corrections au théorème ergodique

La question du rang de la transformation Pascal-adique est reliée à des questions de combinatoire des mots, en particulier à une description fine de la famille de blocs de base $B_{n,k}$ apparaissant naturellement dans l'étude de cette transformation. Rappelons que ces mots sont définis sur l'alphabet $\{a, b\}$ d'une manière similaire aux coefficients binomiaux dans le triangle de Pascal (voir (3.1)) et que $B_{n,k}$ est le \mathcal{P} -nom de longueur $\binom{n}{k}$ de tout point x se trouvant dans la base F_k^n de τ_k^n .

Pour tenter de répondre à ces questions, nous avons été amenés, avec Thierry de la Rue et Yvan Velenik, à réaliser différentes simulations des processus issus de la Pascal-adique, sur lesquelles nous avons pu constater des propriétés remarquables. En effet, une représentation graphique des blocs

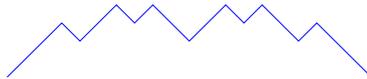


FIGURE 3.3 – Représentation graphique du bloc de base $B_{6,3} = aaabaabbaababbabb$.

de base (un pas vers le haut pour a , un pas vers le bas pour b , voir Figure 3.3) a mis en évidence un phénomène lié aux corrections au théorème ergodique :

Associons à chaque bloc de base $B_{n,k}$, le graphe de la fonction $F_{n,k} : [0, \binom{n}{k}] \rightarrow \mathbb{R}$ définie comme suit.

- $F_{n,k}(0) = 0$;
- si $1 \leq \ell \leq \binom{n}{k}$ est un entier, $F_{n,k}(\ell) = \begin{cases} F_{n,k}(\ell-1) + 1 & \text{si la } \ell\text{-ème lettre de } B_{n,k} \text{ est } a \\ F_{n,k}(\ell-1) - 1 & \text{si la } \ell\text{-ème lettre de } B_{n,k} \text{ est } b; \end{cases}$
- $F_{n,k}$ est interpolée linéairement entre deux entiers consécutifs.

Comme nous allons le voir, le théorème ergodique implique que le graphe de la fonction

$$t \mapsto \frac{1}{\binom{n}{k}} F_{n,k} \left(t \binom{n}{k} \right)$$

converge vers une ligne droite quand $n \rightarrow \infty$ et $k/n \rightarrow p$. Une fois enlevée cette contribution dominante, nous montrons dans [49] qu'après un changement d'échelle, le nouveau graphe converge vers le graphe d'une fonction auto-affine universelle, ne dépendant que de $p = \lim k/n$ (voir Figure 3.4).

Pour $p = 1/2$, il s'agit de la fonction de Van der Waerden (ou Takagi [94], ou Blancmanger), et pour des valeurs de p différentes, on trouve une déformation de celle-ci.

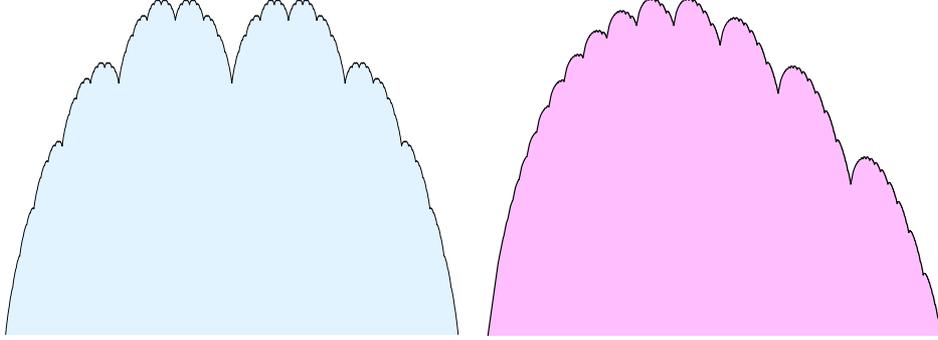


FIGURE 3.4 – Courbe limite correspondant aux blocs de base $B_{n,pn}$ pour $p = 0.5$ (à gauche) et $p = 0.8$ (à droite).

Interprétation ergodique

La fonction $F_{n,k}$ peut être interprétée comme un cas particulier de la situation suivante :

Soit g une fonction à valeurs réelles définie sur un espace de probabilité (X, μ) sur lequel agit une transformation T préservant la mesure. À tout point $x \in X$ et tout entier $\ell \geq 1$ donnés, on associe la fonction continue $F_{x,\ell}^g : [0, \ell] \rightarrow \mathbb{R}$ définie par $F_{x,\ell}^g(0) := 0$; pour tout entier j , $1 \leq j \leq \ell$,

$$F_{x,\ell}^g(j) := \sum_{k=0}^{j-1} g(T^k x); \quad (3.2)$$

et $F_{x,\ell}^g$ est interpolée linéairement entre deux entiers consécutifs.

Si g est intégrable, le théorème ergodique implique que, pour $0 < t < 1$, pour presque tout x ,

$$\lim_{\ell \rightarrow \infty} \frac{1}{\ell} \sum_{0 \leq j < t\ell} g(T^j x) = t \lim_{\ell \rightarrow \infty} \frac{1}{\ell} \sum_{0 \leq j < \ell} g(T^j x).$$

Donc si on fait un changement d'échelle consistant à diviser l'abscisse et l'ordonnée par ℓ , le graphe de $F_{x,\ell}^g$, pour ℓ grand, est proche d'une droite de pente $1/\ell \sum_{0 \leq j < \ell} g(T^j x)$. Pour pouvoir étudier les petites fluctuations dans le théorème ergodique, il est naturel de retirer la contribution

dominante venant de cette ligne droite et de faire un changement d'échelle en ordonnée. Cela mène naturellement à la renormalisation $\varphi_{x,\ell}^g : [0, 1] \rightarrow \mathbb{R}$ donnée par

$$\varphi_{x,\ell}^g(t) := \frac{F_{x,\ell}^g(t\ell) - tF_{x,\ell}^g(\ell)}{R_{x,\ell}^g},$$

où $R_{x,\ell}^g$ est la renormalisation en ordonnée, définie par

$$R_{x,\ell}^g := \begin{cases} \max_{0 \leq t \leq 1} |F_{x,\ell}^g(t\ell) - tF_{x,\ell}^g(\ell)| & \text{si cette quantité n'est pas nulle,} \\ 1 & \text{sinon.} \end{cases} \quad (3.3)$$

Si l'on considère un shift de Bernoulli dans lequel les fonctions $g \circ T^k$ sont des variables aléatoires i.i.d., le principe d'invariance de Donsker dit que ces corrections à la loi des grands nombres suivent un pont brownien renormalisé.

Dans le contexte de la transformation Pascal-adique, puisque le bloc de base $B_{n,k}$ code la trajectoire $(x, Tx, \dots, T^{(\frac{n}{k})-1}x)$ d'un point x de la base de la tour $\tau_{n,k}$ pour la partition $[0, 1/2[$ (pour "a") and $[1/2, 1[$ (pour "b"), la fonction $F_{n,k}$ n'est rien d'autre que $F_{x,(\frac{n}{k})}^g$ où x est dans la base de $\tau_{n,k}$ et g est définie par

$$g := \mathbb{1}_{[0,1/2[} - \mathbb{1}_{[1/2,1[}.$$

Nous pensons que le résultat reste vrai pour une classe plus large de fonctions g , bien qu'il existe des fonctions mesurables pour lesquelles cela est faux. Il serait intéressant de savoir si le même type de comportement apparaît dans d'autres systèmes dynamiques (en particulier des systèmes de rang un), ou s'il est propre à la Pascal-adique.

3.1.4 Autres questions ouvertes et conjectures

La question du rang

La lâche-bernoullicité de la Pascal-adique conduit naturellement à se poser d'autres questions. En général, dans la catégorie des systèmes d'entropie nulle, la propriété de lâche-bernoullicité est connue des transformations de rang un (construites par cutting and stacking en n'utilisant qu'une seule tour à chaque étape), de rang fini (nombre fini de tours) ou de rang un local (une tour occupe une place substantielle de l'espace), ces propriétés étant de moins en moins fortes. On ne sait pas encore aujourd'hui à quel niveau se situe la Pascal-adique dans cette échelle de propriétés, bien que la conjecture la plus naturelle soit de penser qu'elle ne soit même pas de rang un local.

Théorie spectrale, mélange faible

On ne connaît que très peu de résultats concernant la théorie spectrale de cette transformation. La première grande question concerne le mélange faible : il est connu que l'opérateur $U_T : f \mapsto f \circ T$ de $L^2(\mu_p)$ n'admet aucune valeur propre racine de l'unité (hors de 1), donc que l'action de T sur chaque composante ergodique est totalement ergodique. Il est raisonnable de croire que T est faiblement mélangeante, et cette question est reliée à la fois à la théorie des nombres et aux marches aléatoires : en effet, il est prouvé (voir par exemple [82]) qu'une éventuelle valeur propre λ de U_T agissant sur $L^2(\mu_p)$ devrait satisfaire

$$\lambda^{(S_n)} \longrightarrow 1 \quad \text{p.s.}$$

où S_n désigne la somme des n variables de Bernoulli indépendantes de paramètre p (les n premières coordonnées de la suite de 0 et de 1 sur laquelle agit T), et il est conjecturé que cela entraîne $\lambda = 1$. Un angle d'attaque possible pour aborder ce problème serait d'étudier l'éventuelle équidistribution, pour tout $\lambda \neq 1$ de module 1, de la suite $\lambda^{(S_n)}$. Malheureusement, Adams et Petersen [1] puis Berend, Boshernitzan et Kolesnik [3] prouvèrent que, une suite (S_n) étant donnée, la suite $\lambda^{(S_n)}$

n'est pas uniformément distribuée pour un ensemble de λ qui est G_δ dense dans le cercle unité. La question du mélange faible de la Pascal-adique reste donc complètement ouverte, et constitue un défi majeur concernant cette transformation.

Au-delà du problème du mélange faible se posent naturellement les autres questions spectrales du mélange, du spectre de Lebesgue et de la multiplicité spectrale de T , mais qui semblent pour le moment extrêmement difficiles. On ne peut pas exclure cependant que la question du mélange faible puisse être résolue par exemple en prouvant directement le mélange.

Couplages et facteurs

Un terrain totalement vierge reste à défricher : celui de la structure des facteurs et des couplages de la transformation Pascal-adique. On ne sait pas si elle admet des facteurs non triviaux, des auto-couplages non triviaux. Deux composantes ergodiques différentes sont-elles disjointes ?

3.2 Les autocouplages 2 à 2 indépendants

Considérons un processus stationnaire $(\xi_i)_{i \in \mathbb{Z}}$ à valeurs dans un alphabet fini \mathbb{A} . Ce processus est *mélangeant* si deux fenêtres de taille fixée deviennent asymptotiquement indépendantes lorsque la distance qui les sépare tend vers l'infini :

$$\forall \ell \geq 0, \forall E_1, E_2 \subset \mathbb{A}^{\ell+1}, \mathbb{P}\left(\xi_0^\ell \in E_1, \xi_p^{p+\ell} \in E_2\right) - \mathbb{P}\left(\xi_0^\ell \in E_1\right)\mathbb{P}\left(\xi_p^{p+\ell} \in E_2\right) \xrightarrow{p \rightarrow \infty} 0.$$

En 1949, V.A. Rohlin [89] a proposé une généralisation de la définition précédente impliquant plus de deux fenêtres : le processus ξ est *mélangeant à l'ordre 3* si

$$\forall \ell \geq 0, \forall E_1, E_2, E_3 \subset \mathbb{A}^{\ell+1}, \mathbb{P}\left(\xi_0^\ell \in E_1, \xi_p^{p+\ell} \in E_2, \xi_{p+q}^{p+q+\ell} \in E_3\right) - \mathbb{P}\left(\xi_0^\ell \in E_1\right)\mathbb{P}\left(\xi_p^{p+\ell} \in E_2\right)\mathbb{P}\left(\xi_{p+q}^{p+q+\ell} \in E_3\right) \xrightarrow{p, q \rightarrow \infty} 0.$$

Rohlin se demandait alors si tout processus stationnaire mélangeant était mélangeant à l'ordre 3. Cette question, qui est toujours ouverte, est à l'origine de l'intérêt porté aux autocouplages 2 à 2 indépendants.

Rappelons la notion d'autocouplage, introduite par Furstenberg [29] : soit μ la loi du processus stationnaire ξ sur $\mathbb{A}^{\mathbb{Z}}$. Soit ξ' un autre processus défini sur le même espace de probabilité et suivant la même loi μ . Alors (ξ, ξ') est un processus qui prend ses valeurs dans $\mathbb{A} \times \mathbb{A}$. S'il est stationnaire, on dit que sa loi λ sur $\mathbb{A}^{\mathbb{Z}} \times \mathbb{A}^{\mathbb{Z}} \approx (\mathbb{A} \times \mathbb{A})^{\mathbb{Z}}$ est un *autocouplage à l'ordre 2* de ξ . En d'autres termes, un autocouplage à l'ordre 2 de ξ est une probabilité sur $\mathbb{A}^{\mathbb{Z}} \times \mathbb{A}^{\mathbb{Z}}$ dont les marginales sont égales à μ , et qui est invariante par le décalage des coordonnées. On définit de même un autocouplage d'ordre 3 qui fait intervenir trois processus ξ, ξ' et ξ'' de même distribution μ .

Il est bien connu que si ξ est mélangeant mais pas mélangeant à l'ordre 3, alors ξ admet un autocouplage $\lambda \neq \mu \otimes \mu \otimes \mu$ d'ordre 3 à marges 2 à 2 indépendantes. De plus, on sait que s'il existe un contre-exemple à la question de Rohlin, alors on peut en trouver un dans la catégorie des processus d'entropie nulle. Cela nous amène à la question suivante.

Question : Existe-t-il un processus stationnaire ξ , d'entropie nulle, mélangeant, possédant un autocouplage d'ordre 3 à marges 2 à 2 indépendantes mais différent de la mesure produit ?

Cette question est encore sans réponse complète. De nombreux travaux montrent qu'un tel autocouplage ne peut pas exister pour certaines catégories de processus (processus gaussiens [62] [95], à spectre singulier [34], de rang fini [57] [92], etc.) Par ailleurs, Ledrappier [61] a construit un exemple d'entropie nulle dans le cas d'un processus indexé par \mathbb{Z}^2 , mais Thierry de la Rue [15, 14] a montré qu'un tel exemple ne pouvait pas être réalisé en dimension 1.

Dans [45], en collaboration avec Thierry de la Rue, nous nous sommes intéressés à la classe des autocouplages d'ordre 3 à marges 2 à 2 indépendantes pour lesquels ξ'' est une fonction continue de ξ et ξ' . Cela signifie qu'il existe $p \in \mathbb{N}$ tel que ξ''_j est déterminé par $\xi_{j-p}, \dots, \xi_{j+p}$ et $\xi'_{j-p}, \dots, \xi'_{j+p}$.

(Nous dirons dans ce cas que ξ a la propriété PIJ.) Contrairement aux précédents travaux, qui considèrent une certaine classe de processus pour laquelle on montre que les autocouplages d'ordre 3 à marges 2 à 2 indépendantes sont réduits à la mesure produit, l'approche consiste ici à supposer l'existence d'un autocouplage d'ordre 3 à marges 2 à 2 indépendantes vérifiant une certaine propriété, et à analyser les conséquences.

Nous montrons que la propriété PIJ implique que

- ou bien ξ est périodique,
- ou bien ξ est d'entropie positive.

Nous nous intéressons particulièrement au cas où ξ est un sous-shift de type fini, et montrons alors qu'il est shift-équivalent à un Bernoulli uniforme. Par ailleurs, une conjecture raisonnable est que la propriété PIJ ne peut être réalisée que dans la classe des sous-shifts de type fini. En effet, c'est vrai si par exemple on suppose de plus que ξ est lui-même une fonction continue de ξ' et ξ'' .

Tous les résultats obtenus dans cette étude tendent à montrer que les seuls exemples de PIJ possibles sont construits à partir de l'exemple du Bernoulli uniforme par une conjugaison topologique (qui préserve clairement la propriété PIJ).

La prochaine étape, probablement difficile, consiste à lever l'hypothèse de continuité dans la condition PIJ : que se passe-t-il si ξ'' est simplement une fonction mesurable de ξ et ξ' ?

3.3 Suspensions de Poisson et entropie en mesure infinie

Les systèmes dynamiques d'origine probabiliste sont toujours des objets d'étude particulièrement riches, et l'étude de leur structure et de leurs propriétés présente un grand intérêt. Suite aux travaux d'Emmanuel Roy (Paris 13), je me suis intéressée à une classe importante de tels systèmes : les suspensions de Poisson. Ces systèmes offrent le remarquable intérêt de jeter un pont entre la théorie ergodique en mesure infinie et la théorie ergodique en mesure finie.

Partant d'un système dynamique (X, \mathcal{B}, μ, T) pour lequel la mesure μ préservée par la transformation T est infinie (mais néanmoins σ -finie), l'idée est de considérer l'espace X^* des mesures ponctuelles sur (X, \mathcal{B}) , c'est-à-dire les mesures γ définies sur la tribu \mathcal{B} , à valeurs dans l'ensemble des entiers naturels, et vérifiant $\gamma(A) < \infty$ pour chaque A de mesure $\mu(A)$ finie. Cet espace X^* peut lui-même être muni d'une structure d'espace mesurable, en considérant la plus petite tribu \mathcal{B}^* rendant mesurables les applications $\gamma \mapsto \gamma(A)$ pour $A \in \mathcal{B}$. On munit ensuite cet espace mesurable (X^*, \mathcal{B}^*) d'une mesure de probabilité μ^* , qui est la mesure de Poisson d'intensité μ : pour A_1, \dots, A_k deux à deux disjoints, les variables aléatoires $\gamma(A_1), \dots, \gamma(A_k)$ sont indépendantes sous μ^* et suivent k lois de Poisson de paramètres respectifs $\mu(A_1), \dots, \mu(A_k)$. Enfin, on définit sur l'espace probabilisé $(X^*, \mathcal{B}^*, \mu^*)$ la transformation T_* en posant, pour tout $A \in \mathcal{B}$,

$$T_*\gamma(A) := \gamma(T^{-1}A).$$

Heuristiquement, si l'on se représente la mesure aléatoire ponctuelle γ comme un nuage de points, la transformation T_* consiste à faire agir la transformation T individuellement et simultanément sur chacun de ces points. La mesure μ étant invariante par T , il est immédiat que μ^* est invariante par T_* , et l'on obtient ainsi un système dynamique probabilisé $(X^*, \mathcal{B}^*, \mu^*, T_*)$ construit à partir du système de base en mesure infinie.

Dans un premier travail [51] en collaboration avec Tom Meyerovitch (Tel Aviv University), Emmanuel Roy (Paris 13) et Thierry de la Rue, nous nous intéressons aux différentes notions d'entropie pour les systèmes dynamiques en mesure infinie : si cette théorie est parfaitement aboutie en ce qui concerne les transformations préservant une mesure de probabilité, ce n'est pas le cas pour la mesure infinie. Différentes définitions ont été proposées pour l'entropie d'une transformation préservant une mesure infinie, parmi lesquelles celle de Krengel [59], se ramenant à la mesure finie par induction, et celle de Parry [81], qui utilise la notion d'entropie conditionnelle, qui toutes deux généralisent l'entropie de Kolmogorov d'une transformation préservant une mesure de probabilité : plus précisément, l'entropie de Krengel d'une transformation préservant la mesure

(X, \mathcal{B}, μ, T) est définie comme

$$h_{\text{Kr}}(X, \mathcal{B}, \mu, T) := \sup_{A \in \mathcal{F}_+} \mu(A) h(A, \mathcal{B} \cap A, \mu_A, T_A),$$

où \mathcal{F}_+ est la collection des ensembles dans \mathcal{B} de mesure finie strictement positive, μ_A est la mesure de probabilité sur A obtenue par restriction de μ à $\mathcal{B} \cap A$, et $T_A : A \rightarrow A$ est la transformation induite sur A ; l'entropie de Parry est donnée par

$$h_{\text{Pa}}(X, \mathcal{B}, \mu, T) := \sup_{T^{-1}\mathcal{C} \subset \mathcal{C}} H_\mu(\mathcal{C} | T^{-1}\mathcal{C}),$$

où le supremum est pris sur toutes les sous-tribus σ -finies \mathcal{C} de \mathcal{B} telles que $T^{-1}\mathcal{C} \subset \mathcal{C}$. Beaucoup plus récemment, Daniel Rudolph et Alexander Danilenko [13] ont abordé cette question par le biais de l'entropie relative par rapport à un facteur. Emmanuel Roy a proposé de définir l'*entropie de Poisson* d'une transformation préservant une mesure infinie par l'entropie de la suspension de Poisson au-dessus de cette suspension, ce qui est un autre moyen de se ramener à la mesure finie.

Dans [51], nous parvenons à montrer que pour une large classe de systèmes dynamiques, toutes ces notions coïncident. L'introduction de l'entropie de Poisson permet également de montrer des résultats inconnus concernant l'existence d'un facteur de Pinsker pour ce type de transformations. Mais de nombreuses questions restent ouvertes à ce sujet, notamment pour les systèmes dits non quasi-finis.

Dans un second travail en collaboration avec Thierry de la Rue [46], nous parvenons à construire un contre-exemple montrant que l'entropie de Krengel peut être nulle sans que l'entropie de Poisson le soit. Dans ce cas, comme l'entropie de Krengel domine l'entropie de Parry [81], cette dernière est nulle aussi. Notre exemple est obtenu comme une tour au-dessus d'un odomètre, donc son entropie de Krengel s'annule, mais nous montrons qu'il existe des processus d'entropie positive dans la suspension de Poisson associée. La découverte de ce contre-exemple fut une surprise pour nous, et nous montre qu'il nous reste beaucoup de questions à explorer dans ce domaine. On peut notamment se demander si l'entropie de Poisson est toujours supérieure ou égale à celle de Krengel.

Chapitre 4

Quelques problèmes issus de la physique

4.1 Marches aléatoires sur la sphère et sur $SO(n)$

4.1.1 Convergence à l'équilibre

Dans [32], Hastings suggère de générer une rotation aléatoire du groupe orthogonal $SO(n)$ en utilisant une marche aléatoire : On part d'un élément quelconque $M \in SO(n)$. On choisit au hasard deux coordonnées i, j dans $\{1, \dots, n\}$, un angle $\theta \in [0, 2\pi[$ et on multiplie M par $R_{i,j}^\theta$, la rotation d'angle θ dans le plan (i, j) .

On considère alors la chaîne de Markov associée, non pas le long de la suite des temps entiers, mais le long d'une suite de temps donnés par un processus de Poisson (indépendant de la chaîne). Le semi-groupe associé à ce processus de Markov est $P_t := \exp(-t(Id - K))$, où K est l'opérateur de Markov correspondant à l'application d'une rotation aléatoire suivant la procédure indiquée ci-dessus. Son générateur infinitésimal est donné par

$$L_n F(M) = \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \int_0^{2\pi} (F(R_{i,j}^\theta M) - F(M)) \frac{d\theta}{2\pi}.$$

La mesure uniforme μ_n sur $SO(n)$ (appelée mesure de Haar) est invariante et réversible pour L_n (on a pour tout f, g , $\int f L_n g d\mu_n = \int L_n f g d\mu_n$). La question est ici de savoir pendant combien de temps il faut itérer ce procédé pour pouvoir considérer que nous sommes proches de μ_n .

Notons qu'il s'agit d'un exemple d'échantillonnage de Gibbs (voir [88]). Supposons que l'on souhaite simuler un échantillon suivant une loi jointe $Y = (Y_1, \dots, Y_p)$ et que l'on sait simuler les densités conditionnelles correspondante f_1, \dots, f_p . L'algorithme de l'échantillonnage de Gibbs associé à (Y_1, \dots, Y_p) consiste à choisir une coordonnée i au hasard dans $\{1, \dots, p\}$, geler les autres coordonnées et simuler suivant la distribution $f_i(\cdot | (Y_k)_{k \neq i})$. Ainsi, pour générer une rotation aléatoire (par rapport à la mesure uniforme sur la $SO(n)$), on choisit deux coordonnées (i, j) au hasard et on simule suivant la distribution conditionnelle aux autres coordonnées. On retrouve exactement la marche suggérée par Hastings.

4.1.2 Trou spectral

Comme le générateur L_n est un opérateur auto-adjoint de $L^2(\mu_n)$, toutes ses valeurs propres sont réelles. De plus, 0 est une valeur propre associée aux fonctions constantes et on voit en écrivant $\int f L_n f d\mu_n$ sous la forme

$$-\frac{1}{2} \int \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \int_0^{2\pi} (f(R_{i,j}^\theta M) - f(M))^2 \frac{d\theta}{2\pi} d\mu_n(M) \leq 0$$

que les valeurs propres sont négatives. Le théorème spectral nous dit que pour une condition initiale ϕ , la vitesse de convergence vers l'équilibre est donnée par

$$\|P_t\phi - E_{\mu_n}(\phi)\|_2 \leq e^{-\lambda_n t} \|\phi(0) - E_{\mu_n}(\phi)\|_2,$$

où $E_{\mu_n}(\phi)$ est l'espérance de ϕ par rapport à la mesure μ_n et λ_n est la plus petite valeur propre non nulle de $-(L_n + L_n^*)/2$ (ici égal à $-L_n$), appelée aussi trou spectral de L_n .

Ainsi, la vitesse de convergence mesurée en norme L^2 est exponentielle. Mais elle dépend de λ_n et donc de n . Il faut donc estimer λ_n , ou au moins son ordre de grandeur.

Ce problème est en fait assez délicat et le seul résultat disponible à l'époque était le travail de Diaconis et Saloff-Coste [20] dans lequel ils trouvent une borne inférieure pour le trou spectral d'ordre $1/n^3$ (la borne supérieure d'ordre $1/n$ étant facilement obtenue grâce à des fonctions test). Pour cela, ils utilisent une "méthode par comparaison" qui consiste à comparer la forme de Dirichlet de la marche considérée avec celle d'une marche dont le trou spectral est connu (ici, la marche engendrée par des rotations aléatoires d'angle π dans n'importe quel plan étudiée par Rosenthal dans [91]).

J'ai abordé ce problème en étudiant d'abord les propriétés de contraction du semi-groupe considéré sur la sphère $S^{n-1}(\sqrt{E})$ de rayon \sqrt{E} .

$$\mathcal{L}_n f(v) = \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} \int_0^{2\pi} (f(R_{i,j}^\theta v) - f(v)) \frac{d\theta}{2\pi}, \quad (4.1)$$

pour $v \in S^{n-1}(\sqrt{E})$. La mesure S_{n-1} , probabilité uniforme sur $S^{n-1}(\sqrt{E})$, est invariante and réversible pour \mathcal{L}_n . Le cas de la sphère est déjà très intéressant puisque cette dynamique a été utilisée par Marc Kac (voir [55]) pour obtenir l'analogie suivant de l'équation de Boltzmann spatialement homogène :

$$\begin{aligned} \frac{\partial g_t(x)}{\partial t} &= C \int \int_0^{2\pi} \left(g_t(x \cos \theta + y \sin \theta) g_t(-x \sin \theta + y \cos \theta) \right. \\ &\quad \left. - g_t(x) g_t(y) \right) d\theta dy. \end{aligned} \quad (4.2)$$

Les propriétés spectrales de l'opérateur de collision apparaissant dans (4.2) ont une grande importance pour comprendre cette équation et cet opérateur étant obtenu à partir du processus introduit par Kac, il est naturel de chercher à connaître l'ordre de grandeur de son trou spectral. Kac [56] conjectura que celui-ci était d'ordre $1/n$.

Notons que l'on peut voir le trou spectral $SG(\mathcal{L}_n)$ de \mathcal{L}_n comme

$$SG(\mathcal{L}_n) := \inf_{f \in L^2(S^{n-1})} \frac{\mathcal{D}_n(f)}{E_{S_{n-1}}[f; f]},$$

où $E_{S_{n-1}}[f; f]$ est la variance de f par rapport à S_{n-1} et $\mathcal{D}_n(f)$ est la forme de Dirichlet de f , donnée par $\mathcal{D}_n(f) := -\int f \mathcal{L}_n f dS_{n-1}$. Ici,

$$\mathcal{D}_n(f) = \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} \frac{1}{2} \int_0^{2\pi} \int (f(R_{i,j}^\theta v) - f(v))^2 dS_{n-1}(v) \frac{d\theta}{2\pi}.$$

Comme la valeur du trou spectral ne dépend pas de E , nous fixerons $E = 1$. Il est facile d'obtenir une borne supérieure pour le trou spectral en utilisant une fonction test (par exemple $f(v) = v_1$). Pour avoir une borne inférieure, il faut donc prouver une *inégalité de Poincaré*, c'est-à-dire trouver une constante C_n telle que

$$\forall f \in L^2(S^{n-1}), \quad E_{S_{n-1}}[f; f] \leq C_n \mathcal{D}_n(f). \quad (4.3)$$

En utilisant la méthode des martingales ([65]), j'ai réussi dans [40] à établir une borne inférieure d'ordre $1/n$, et donc à prouver la conjecture de Kac.

Théorème 4.1. *Il existe une constante $C > 0$ telle que pour toute fonction $f \in L^2(S_{n-1})$*

$$E_{S_{n-1}}[f; f] \leq C n \mathcal{D}_n(f).$$

Le principe de la méthode des martingales est le suivant : pour tout entier $l \geq 2$, soit a_l la plus petite constante telle que, pour toute fonction $h \in L^2(S_{l-1})$,

$$E_{S_{l-1}}[h; h] \leq a_l l \mathcal{D}_l(h). \quad (4.4)$$

Le but est d'établir une équation réursive pour a_n afin de montrer que a_n est bornée par une constante indépendante de n .

– Décomposition de la variance : La variance peut être réécrite comme

$$E_{S_{n-1}}[f; f] = \frac{1}{n} \sum_{j=1}^n E_{S_{n-1}} [E_{S_{n-1}}[f; f|v_j]] + \frac{1}{n} \sum_{j=1}^n E_{S_{n-1}} [f_j; f_j], \quad (4.5)$$

où f_j est l'espérance de f sachant v_j ($f_j(v_j) = E_{S_{n-1}}[f|v_j]$) et $E_{S_{n-1}} [f; g|\mathcal{F}]$ est la covariance de f et g par rapport à la mesure S_{n-1} conditionnée à \mathcal{F} :

$$E_{S_{n-1}} [f; g|\mathcal{F}] = E_{S_{n-1}} [fg|\mathcal{F}] - E_{S_{n-1}} [f|\mathcal{F}] E_{S_{n-1}} [g|\mathcal{F}].$$

- Induction : La mesure S_{n-1} conditionnée à v_j est la mesure uniforme sur $S^{n-2}(\sqrt{1-v_j^2})$, la sphère de dimension $(n-2)$ et de rayon $\sqrt{1-v_j^2}$. On peut donc utiliser (4.4) pour borner $E_{S_{n-1}}[f; f|v_j]$.
- Variance sur une coordonnée : $E_{S_{n-1}}[f_j; f_j]$ est la variance par rapport à S_{n-1} d'une fonction ne dépendant que d'une coordonnée. La partie difficile, et souvent technique, de la méthode des martingales consiste à trouver une borne fine pour ces termes.

Depuis, Carlen, Carvalho and Loss [9] ont réussi à montrer que le trou spectral de \mathcal{L}_n était exactement égal à $(n+2)/2n(n-1)$.

J'ai alors cherché à améliorer le résultat de Diaconis et Saloff-Coste qui était aussi valable pour la marche aléatoire sur le groupe orthogonal $SO(n)$. Pour cela, j'ai utilisé une décomposition de la mesure de Haar $\mu_n^{\mathcal{H}}$ sur $SO(n)$ et le trou spectral sur la sphère : la marginale de $\mu_n^{\mathcal{H}}$ sur la première ligne $L_1 = (M_{11}, \dots, M_{1n})$ de M est la mesure uniforme S_{n-1} sur la sphère $S^{n-1}(1)$ de rayon 1, et la probabilité conditionnelle de $\mu_n^{\mathcal{H}}$ par rapport à L_1 peut être identifiée à la mesure de Haar $\mu_{n-1}^{\mathcal{H}}$ sur $SO(n-1)$. En d'autres termes, l'application suivante est isométrique :

$$\begin{aligned} SO(n) &\longrightarrow S^{n-1} \times SO(n-1) \\ M &\longrightarrow (L_1, \tilde{Q}_{n-1}(M)) \end{aligned}$$

On peut maintenant utiliser la méthode des martingales : on décompose la variance en introduisant l'espérance de $F(M)$ sachant la ligne L_i de M . Le Théorème 4.1 nous donne une estimée de la variance sur L_i . Ce terme est d'ordre n , donc on obtient une borne inférieure d'ordre n^{-2} pour le trou spectral sur $SO(n)$ ([41, 42]).

Cette borne améliore le résultat de Diaconis et Saloff-Coste [20] d'un facteur n , mais n'est pas optimale. En effet, D. Maslen a prouvé plus tard dans [68] que le trou spectral sur $SO(n)$ était le même que sur la sphère, *i.e.* $(n+2)/2n(n-1)$.

4.1.3 Inégalités de Sobolev logarithmique

Lorsque le trou spectral est connu, on cherche généralement à prouver une inégalité de Sobolev logarithmique (LSI), *i.e.* à trouver une constante K_n telle que

$$\forall f, \quad H(f) := \int f \log \frac{f}{\int f dS_{n-1}} dS_{n-1} \leq K_n \mathcal{D}_n(\sqrt{f}), \quad (4.6)$$

Gross [31] a montré que la LSI était équivalente à l'hypercontractivité du semigroupe.

L'inégalité (4.6) est plus forte que l'inégalité de Poincaré (4.3). En effet, en prenant $f = 1 + \epsilon g$ où $\int g dS_{n-1} = 0$ dans (4.6) et en faisant tendre ϵ vers 0, on trouve que $E_{S_{n-1}}[f; f] \leq K_n/2 \mathcal{D}_n(f)$, et donc que le trou spectral est plus grand que $2/K_n$.

En fait, l'inégalité de Sobolev logarithmique n'est pas vérifiée pour le modèle de Kac : Pour tout $\epsilon > 0$, on peut trouver un ensemble A_ϵ de probabilité ϵ . Or, l'entropie de $f := \mathbb{1}_{\{A_\epsilon\}}$ est égale à $H_{S_{n-1}}(f) - \epsilon \log \epsilon$ et $\mathcal{D}_n(\sqrt{f}) \leq \epsilon$. On ne peut donc pas borner l'entropie par la forme de Dirichlet de \sqrt{f} .

L'inégalité de Sobolev logarithmique modifiée (MLSI) est un intermédiaire entre la LSI et l'inégalité de Poincaré (4.3). Si on peut trouver une constante M_n vérifiant

$$\forall f \text{ telle que } \int f dS_{n-1} = 1, \quad H(f) \leq M_n \int f \mathcal{L}_n(\log f) dS_{n-1}, \quad (4.7)$$

alors $K_n/4 \geq M_n \geq C_n/2$.

De plus, comme la dérivée en temps de $H(P_t f)$ est égale à $\int P_t f \mathcal{L}_n(\log P_t f) dS_{n-1}$ (la dissipation d'entropie), la MLSI est équivalente à la décroissance exponentielle de l'entropie.

Villani ([99], Annexe III, Appendix B) a obtenu un résultat dans cette direction pour la marche sur la sphère avec $M_n = O(n^2)$. Sa méthode est une variante de celle qu'il a utilisée avec Toscani pour obtenir ce type d'inégalité pour l'équation de Boltzmann. Ce résultat ne semble pas optimal puisque cette méthode permet seulement d'obtenir une borne inférieure pour le trou spectral d'ordre n^{-2} (au lieu de n^{-1}). Cependant, il n'est pas évident qu'il soit possible d'obtenir une constante d'ordre n . En effet, cela impliquerait que la caricature de l'équation de Boltzmann (4.2) vérifierait une inégalité de la forme "Entropie $\leq C$ dissipation d'entropie" pour une constante C . Or, Bobylev and Cercignani ont montré qu'une telle inégalité n'était pas satisfaite pour l'équation de Boltzmann.

On pourrait aussi essayer d'estimer M_n en utilisant la méthode des martingales, qui consiste ici à décomposer l'entropie sous la forme

$$\begin{aligned} H(f) &= \frac{1}{n} \sum_{j=1}^n E_{S_{n-1}} \left[E_{S_{n-1}} \left[f \log \frac{f}{f_j} \middle| v_j \right] \right] + \frac{1}{n} \sum_{j=1}^n H(f_j) \\ &= \frac{1}{n} \sum_{j=1}^n E_{S_{n-1}} [H(f|v_j)] + \frac{1}{n} \sum_{j=1}^n H(f_j), \end{aligned} \quad (4.8)$$

où $f_j(v_j) = E_{S_{n-1}}[f|v_j]$, et à utiliser une hypothèse d'induction pour le premier terme. Là encore, la partie difficile est d'avoir une bonne estimée pour le second terme. Par contre, en utilisant le résultat de Villani et la décomposition de la mesure de Haar $\mu_n^{\mathcal{H}}$, on obtient facilement par la méthode des martingales une MLSI sur $SO(n)$.

Théorème 4.2. *Il existe une constante $C > 0$ telle que pour toute fonction lisse F sur $SO(n)$,*

$$H_n(F) \leq Cn^3 \int F \mathcal{L}_n(\log F) d\mu_n^{\mathcal{H}}.$$

4.2 Un problème d'accrochage de polymères

L'étude des polymères dirigés en milieu aléatoire est un problème important de physique qui a suscité récemment un grand intérêt de la part de la communauté mathématique (voir [10]). Le comportement de ces polymères est modélisé par une perturbation exponentielle de la loi d'une marche aléatoire d -dimensionnelle, qui dépend d'un environnement aléatoire. La perturbation favorise les visites de la marche dans les régions où l'environnement prend des valeurs positives, tandis que les régions de signe opposé sont pénalisées.

Un exemple de telle perturbation est le suivant : considérons une trajectoire $Y : \mathbb{N} \rightarrow \mathbb{Z}^d$ fixée, et perturbons la loi d'une marche aléatoire X en récompensant chaque intersection entre les trajectoires de X et de Y . La question est alors de comprendre quelles propriétés de la trajectoire Y entraîne l'accrochage du polymère, *i.e* une densité positive d'intersections.

Ioffe et Louidor [35] ont considéré la situation où Y est elle-même une marche aléatoire et montré que l'accrochage avait lieu pour presque toutes les réalisations de Y . Mais leur résultat présente l'inconvénient de ne pas expliciter l'ensemble des trajectoires qui ne donnent pas lieu à accrochage.

Dans un travail en collaboration avec Thierry de la Rue et Yvan Velenik [48], nous analysons une situation plus simple pour laquelle nous caractérisons les environnements qui provoquent l'accrochage. Le potentiel d'accrochage est restreint à une droite, et l'aléa vient du fait que ce potentiel est dilué. Plus précisément, soit $\omega \in \{0, 1\}^{\mathbb{N}}$, $\eta > 0$, et notons \mathbb{P}_0 la loi d'une marche aléatoire apériodique symétrique sur \mathbb{Z}^d ($d = 1$ ou 2), issue de 0 , dont les accroissements sont de variance finie. Nous nous intéressons à la perturbation de \mathbb{P}_0 donnée par

$$\mathbb{P}_{N,\eta}^\omega(X) := (Z_{N,\eta}^\omega)^{-1} \exp\left(\eta \sum_{i \in \Lambda_N} \mathbb{1}_{(X_i=0)} \omega_i\right) \mathbb{P}_0(X), \quad (4.9)$$

où $\Lambda_N := \{1, \dots, N\}$ et $Z_{N,\eta}^\omega$ est la constante de normalisation. Cette perturbation modélise l'interaction d'un polymère dirigé interagissant avec un potentiel dilué attractif.

Nous obtenons qu'il y a accrochage, c'est-à-dire

$$\liminf_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}_{N,\eta}^\omega \left(\sum_{i=1}^N \mathbb{1}_{X_i=0} \right) > 0,$$

si et seulement si l'environnement ω vérifie

$$\liminf_{N \rightarrow \infty} |\Lambda_N|^{-1} \sum_{i \in \Lambda_N} \omega_i > 0.$$

Ce résultat reste valable dans le cas d'une surface aléatoire attirée par un potentiel dilué dans un plan.

4.3 Génération de pavages aléatoires

La physique statistique s'intéresse à l'étude à grande échelle de systèmes formés d'un nombre immense de composantes microscopiques. Une question essentielle est de savoir comment cette multitude d'éléments s'organise en fonction des contraintes qui leur sont imposées : quel comportement observe-t-on à l'échelle macroscopique? Expérimenter dans ce domaine nécessite de pouvoir générer des configurations typiques à l'aide d'algorithmes simulant les effets du hasard.

L'exemple que nous considérons ici est celui des pavages, avec des dominos de taille 2×1 , d'une région du plan appelée diamant aztèque. Le diamant aztèque d'ordre $n \geq 1$ est un graphe planaire qui peut être vu comme un sous-ensemble du réseau \mathbb{Z}^2 ; l'ensemble de ses sommets est

$$A_n = \left\{ (x, y) \in \mathbb{Z}^2 : \left| x - \frac{1}{2} \right| + \left| y - \frac{1}{2} \right| \leq n \right\},$$

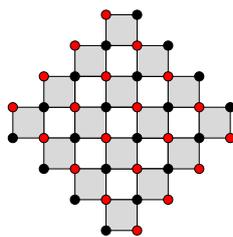


FIGURE 4.1 – Le diamant aztèque d'ordre 4.

et une arête connecte chaque paire de sommets distants de 1 unité (voir Figure 4.1). Un pavage par dominos est un sous-ensemble des arêtes du graphe tel que chaque sommet appartient à une et une seule des arêtes choisies. L'algorithme de *shuffling* fut introduit dans [22] pour compter le nombre de pavages par dominos du diamant aztèque d'ordre n . Peu après, on s'aperçut que cet algorithme permettait aussi de générer un tel pavage aléatoire de manière uniforme. On découvrit ainsi le phénomène suivant : pour n grand, si l'on tire au hasard l'un des pavages du diamant aztèque, on verra apparaître avec une probabilité proche de 1 le fameux *cercle arctique* ; à l'extérieur du cercle inscrit, dans chaque coin du carré, les dominos sont tous gelés dans le même sens, alors qu'à l'intérieur du cercle le désordre règne (voir Figure 4.2). Ainsi, on voit que dans ce modèle, les effets de bord se propagent à une grande distance de la frontière du domaine et sont donc visibles à l'échelle macroscopique.

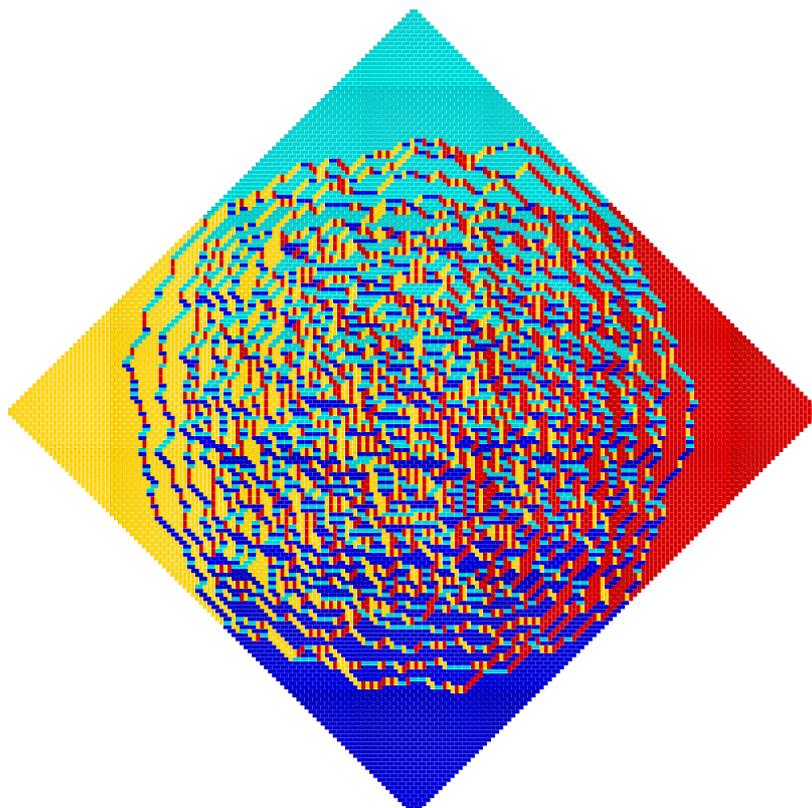


FIGURE 4.2 – Le cercle arctique dans un diamant aztèque d'ordre 100.

Dans [84], James Propp a introduit une version généralisée de l'algorithme de *shuffling* per-

mettant de générer efficacement des pavages aléatoires du diamant avec des arêtes pondérées. Malheureusement cette méthode ne s'applique pas toujours lorsque certaines arêtes sont de poids nul. Pourtant ce cas est important, car il permet, en interdisant certaines arêtes, d'obtenir des pavages d'une grande variété de graphes planaires.

Dans un travail en collaboration avec Thierry de la Rue et Yvan Velenik [50], nous proposons une variante de l'algorithme de shuffling généralisé adaptée aux situations avec arêtes de poids nul. Nous pouvons ainsi générer des pavages aléatoires d'une large classe de sous-graphes pondérés du réseau \mathbb{Z}^2 en les plongeant dans un diamant aztèque assez grand, et nous donnons des bornes sur la taille minimale requise pour le diamant dans lequel on plonge le graphe. Cela fonctionne même avec des graphes a priori différents. Par exemple, le graphe associé au réseau hexagonal peut s'identifier à un sous-graphe d'un diamant aztèque assez grand (voir Figure 4.3 pour d'autres exemples). Une fois généré un pavage du sous-graphe d'un diamant aztèque, on lui associe un pavage de l'hexagone de la façon décrite dans la Figure 4.4. Cette fois les pavés utilisés ne sont plus des dominos rectangulaires, mais des losanges.

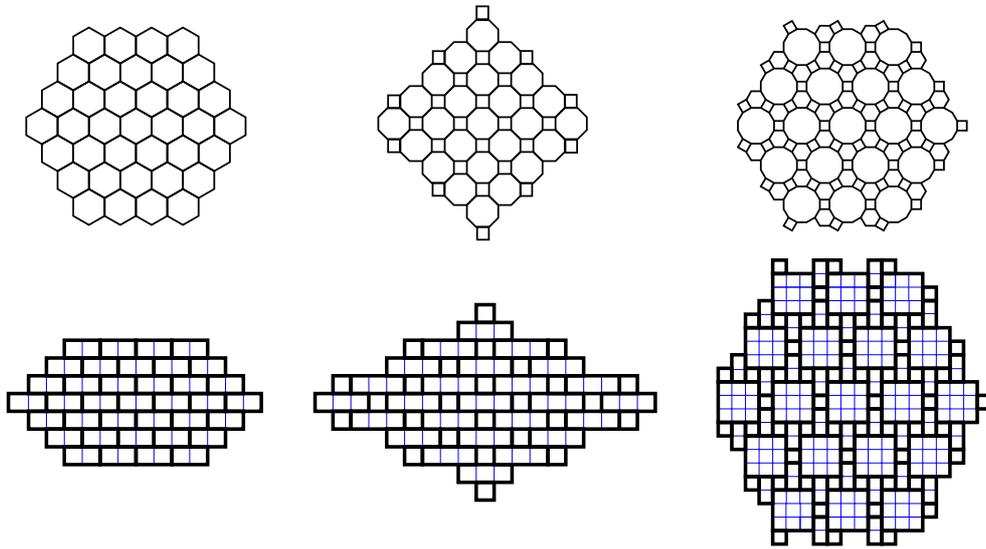


FIGURE 4.3 – Les pavages de type 6-6-6 (hexagonal), 4-8-8 et 4-6-12, et leur plongement dans \mathbb{Z}^2 .

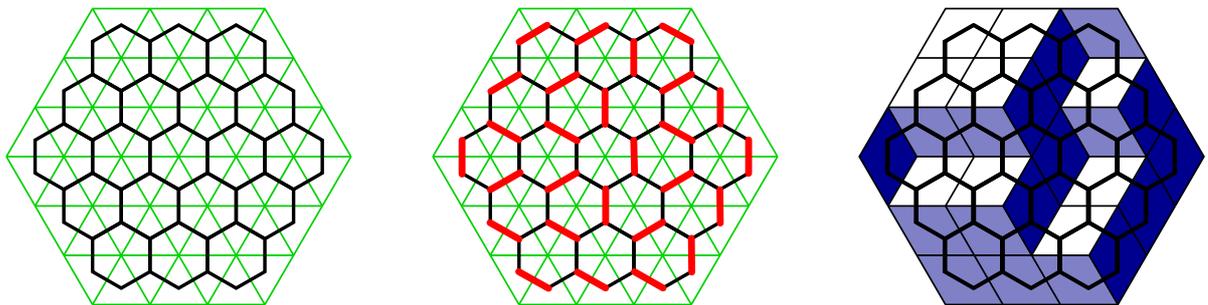


FIGURE 4.4 – À gauche, le pavage hexagonal et son dual. Au milieu, un pavage possible par des dominos. À droite, le pavage correspondant du dual avec des losanges.

Les Figures 4.5 et 4.6 montrent des pavages aléatoires du dual des graphes de la Figure 4.3 obtenus par notre variante de l'algorithme de shuffling. Ces pavages typiques font apparaître un

phénomène semblable au cercle arctique dans le diamant aztèque, mais on ne sait pas encore identifier la forme de la partie gelée.

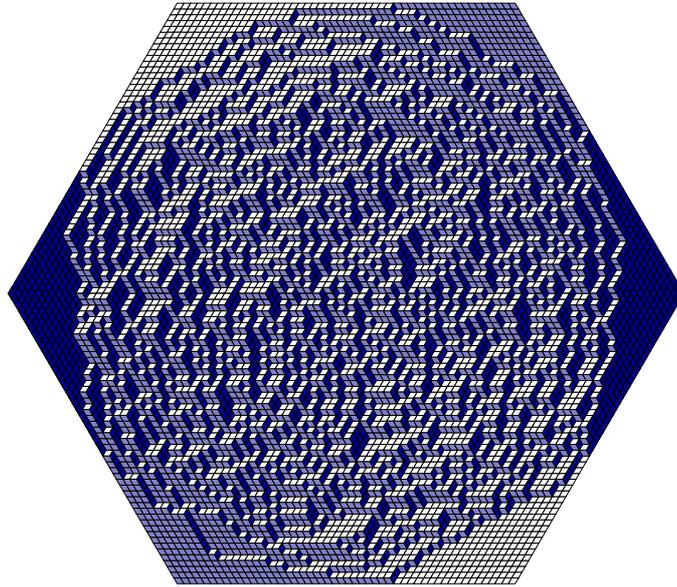


FIGURE 4.5 – Pavage du dual du pavage hexagonal.

4.4 Traitement d'images médicales

Il s'agit d'un programme de travail qui a son origine dans une problématique posée par des médecins du département de médecine nucléaire du centre Henri Becquerel de Rouen (CHU Charles Nicolle).

Le problème consistait à trouver un algorithme efficace pour une détermination automatique, à partir d'images cérébrales obtenues par SPECT (Single-Photon Emission Computed Tomography), de la gravité de l'état du malade, ou de la toxicité de certains traitements par chimiothérapie. La difficulté essentielle réside dans la faible qualité des images à traiter : il s'agit d'images à faible résolution et fortement bruitées.

L'une des propositions faites par notre équipe pour mesurer l'hétérogénéité de l'image du cortex observée est fondée sur des méthodes stochastiques ; elle consiste à analyser le comportement d'une marche aléatoire dont la loi dépend de la densité des pixels mesurée. La loi de la marche aléatoire et les paramètres ont été choisis afin que celle-ci fasse de grandes excursions dans les régions homogènes du cortex et reste confinée (c'est-à-dire repasse de nombreuses fois sur les mêmes pixels) dans les régions hétérogènes.

Les premiers résultats de cette méthode ont été doublement validés : d'une part en testant l'algorithme sur des fantômes simulant des niveaux dégressifs d'hétérogénéité corticale ; d'autre part en comparant le classement obtenu sur des images réelles de patients avec celui établi par des médecins experts. Nous avons aussi étudié l'influence des différents paramètres de la marche. Enfin, nous avons comparé cette méthode à d'autres méthodes proposées dans la littérature (dimension fractale, matrices de co-occurrence, ondelettes...). Ces résultats ont donné lieu à un acte de colloque et deux publications [74, 75, 76] et à un article actuellement soumis [77].

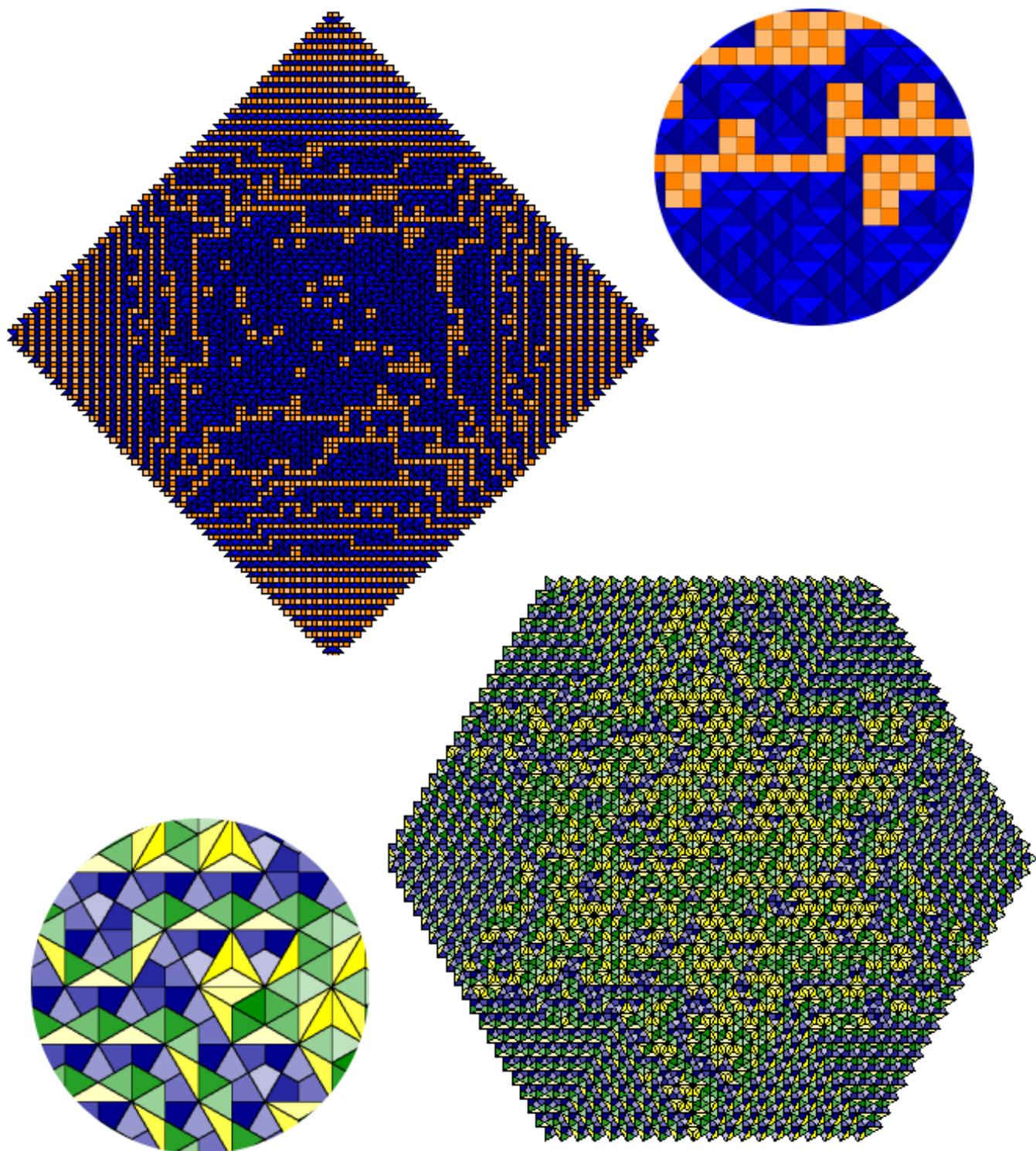


FIGURE 4.6 – Pavage du dual des graphes de type 4-8-8 (en haut) et 4-6-12 (en bas).

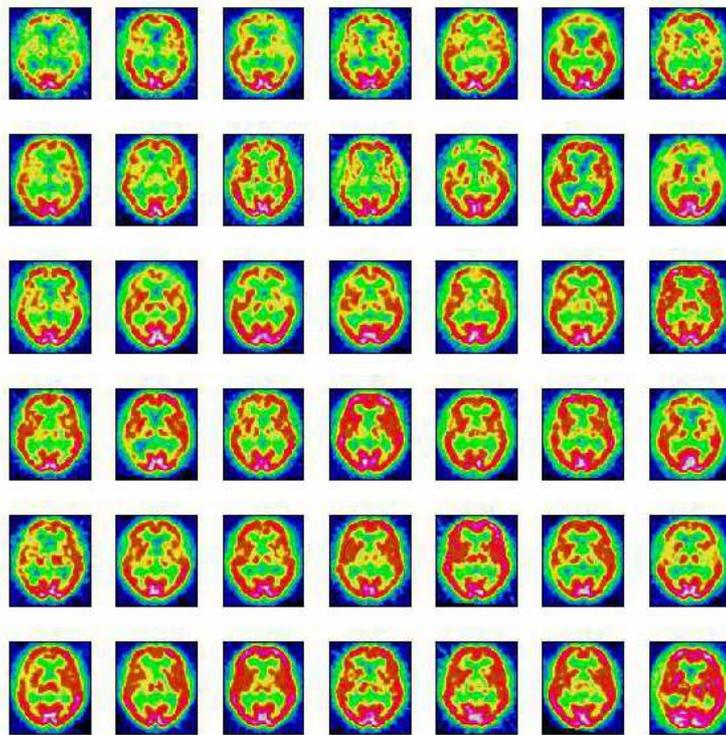


FIGURE 4.7 – Classement automatique par la méthode de la marche aléatoire d'images cérébrales obtenues par SPECT de 42 patients par hétérogénéité décroissante.

Bibliographie

- [1] Terrence M. Adams and Karl E. Petersen. Binomial-coefficient multiples of irrationals. *Monatsh. Math.*, 125(4) :269–278, 1998.
- [2] Arthur T. Benjamin, Christopher R. H. Hanusa, and Francis Edward Su. Linear recurrences through tilings and Markov chains. *Util. Math.*, 64 :3–17, 2003.
- [3] D. Berend, M. D. Boshernitzan, and G. Kolesnik. Irrational dilations of Pascal’s triangle. *Mathematika*, 48(1-2) :159–168 (2003), 2001.
- [4] F. Blanchard. β -expansions and symbolic dynamics. *Theoret. Comput. Sci.*, 65(2) :131–141, 1989.
- [5] Walter Borho. Kettenbrüche im Galoisfeld. *Abh. Math. Sem. Univ. Hamburg*, 39 :76–82, 1973.
- [6] Walter Borho and Gerhard Rosenberger. Eine Bemerkung zur Hecke-Gruppe $G(\lambda)$. *Abh. Math. Sem. Univ. Hamburg*, 39 :83–87, 1973.
- [7] Philippe Bougerol and Jean Lacroix. *Products of random matrices with applications to Schrödinger operators*, volume 8 of *Progress in Probability and Statistics*. Birkhäuser Boston Inc., Boston, MA, 1985.
- [8] J. Boyle. An application of fourier series to the most significant digit problem. *Amer. Math. Monthly*, 101 :879–886, 1994.
- [9] E. A. Carlen, M. C. Carvalho, and M. Loss. Determination of the spectral gap for Kac’s master equation and related stochastic evolution. *Acta Math.*, 191(1) :1–54, 2003.
- [10] Francis Comets, Tokuzo Shiga, and Nobuo Yoshida. Probabilistic analysis of directed polymers in a random environment : a review. In *Stochastic analysis on large scale interacting systems*, volume 39 of *Adv. Stud. Pure Math.*, pages 115–142. Math. Soc. Japan, Tokyo, 2004.
- [11] Karma Dajani and Martijn de Vries. Measures of maximal entropy for random β -expansions. *J. Eur. Math. Soc. (JEMS)*, 7(1) :51–68, 2005.
- [12] Karma Dajani and Martijn de Vries. Invariant densities for random β -expansions. *J. Eur. Math. Soc. (JEMS)*, 9(1) :157–176, 2007.
- [13] Alexandre I. Danilenko and Daniel J. Rudolph. Conditional entropy theory in infinite measure and a question of Krengel. *Israel J. Math.*, 172 :93–117, 2009.
- [14] Thierry de la Rue. 2-fold and 3-fold mixing : why 3-dot-type counterexamples are impossible in one dimension. *Bull. Braz. Math. Soc. (N.S.)*, 37(4) :503–521, 2006.
- [15] Thierry de la Rue. An introduction to joinings in ergodic theory. *Discrete Contin. Dyn. Syst.*, 15(1) :121–142, 2006.
- [16] Jean-Paul Delahaye and Nicolas Gauvrit. Pourquoi la loi de benford n’est pas mystérieuse. *Mathematics and Social Sciences*, 182 :7–15, 2008.
- [17] Arnaud Denjoy. Sur une fonction réelle de Minkowski. *J. Math. Pures Appl.*, 17 :105–151, 1938.
- [18] Persi Diaconis. *Weak and Strong Averages in Probability and the Theory of Numbers*. PhD thesis, 1974.

- [19] Persi Diaconis. Examples for the theory of infinite iteration of summability methods. *Can. J. Math.*, 29(3) :489–497, 1977.
- [20] Persi Diaconis and Laurent Saloff-Coste. Bounds for kac’s master equation. *Commun. Math. Phys.*, 209 :729–755, 2000.
- [21] Robert Luce Duncan. Note on the initial digit problem. *Fibonacci Quat.*, 7 :474–475, 1969.
- [22] Noam Elkies, Greg Kuperberg, Michael Larsen, and James Propp. Alternating-sign matrices and domino tilings. I. *J. Algebraic Combin.*, 1(2) :111–132, 1992.
- [23] Mark Embree and Lloyd N. Trefethen. Growth and decay of random Fibonacci sequences. *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.*, 455(1987) :2471–2485, 1999.
- [24] J. Feldman. New K -automorphisms and a problem of Kakutani. *Israel J. Math.*, 24(1) :16–38, 1976.
- [25] P. A. Ferrari and A. Galves. Coupling and regeneration for stochastic processes. Notes for a minicourse presented in XIII Escuela Venezolana de Matemáticas, 2000.
- [26] B J Flehinger. On the probability that a random integer has initial digit A . *Amer. Math. Monthly*, 73 :1056–1061, 1966.
- [27] Sarah Bailey Frick. Limited scope adic transformations. *Discrete Contin. Dyn. Syst. Ser. S*, 2(2) :269–285, 2009.
- [28] Harry Furstenberg. Noncommuting random products. *Trans. Amer. Math. Soc.*, 108 :377–428, 1963.
- [29] Harry Furstenberg. Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Math. Systems Theory*, 1 :1–49, 1967.
- [30] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete mathematics*. Addison-Wesley Publishing Company Advanced Book Program, Reading, MA, 1989. A foundation for computer science.
- [31] Leonard Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4) :1061–1083, 1975.
- [32] W. K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1) :97–109, 1970.
- [33] T. P. Hill. Base-invariance implies Benford’s law. *Proc. Amer. Math. Soc.*, 123(3) :887–895, 1995.
- [34] Bernard Host. Mixing of all orders and pairwise independent joinings of systems with singular spectrum. *Israel J. Math.*, 76(3) :289–298, 1991.
- [35] D. Ioffe and A. Loidor. Private communication, 2004.
- [36] Shunji Ito and Yōichirō Takahashi. Markov subshifts and realization of β -expansions. *J. Math. Soc. Japan*, 26 :33–55, 1974.
- [37] Dennis Jang, Jung Uk Kang, Alex Kruckman, Jun Kudo, and Steven J. Miller. Chains of distributions, hierarchical bayesian models and benford’s law. 1(1) :37–60, 2009.
- [38] É. Janvresse and R. Giuliano. A unifying probabilistic interpretation of benford’s law. *Unif. Distrib. Theory*, 5(2) :169–182, 2010.
- [39] É. Janvresse, B. Rittaud, and T. de la Rue. Almost-sure growth rate of generalized random fibonacci sequences. *Ann. IHP, Probab. Stat.*, 46(1) :135–158, 2010.
- [40] Élise Janvresse. Spectral gap for Kac’s model of Boltzmann equation. *Ann. Probab.*, 29(1) :288–304, 2001.
- [41] Élise Janvresse. Approach to equilibrium for Kac master equation. *Markov Process. Related Fields*, 8(2) :221–232, 2002. Inhomogeneous random systems (Cergy-Pontoise, 2001).
- [42] Élise Janvresse. Bounds on semigroups of random rotations on $SO(n)$. *Teor. Veroyatnost. i Primenen.*, 47(3) :606–612, 2002.
- [43] Élise Janvresse and Thierry de la Rue. From uniform distributions to Benford’s law. *J. Appl. Probab.*, 41(4) :1203–1210, 2004.

- [44] Élise Janvresse and Thierry de la Rue. The Pascal adic transformation is loosely Bernoulli. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(2) :133–139, 2004.
- [45] Élise Janvresse and Thierry de la Rue. A class of pairwise-independent joinings. *Ergodic Theory Dynam. Systems*, 28(5) :1545–1557, 2008.
- [46] Élise Janvresse and Thierry de la Rue. Zero krengel entropy does not kill poisson entropy. Prépublication, <http://hal.archives-ouvertes.fr/hal-00424090/fr/>, à paraître aux Annales de l’IHP, Probab. Stat., 2009.
- [47] Élise Janvresse and Thierry de la Rue. Averaging along uniform random integers. Prépublication, <http://hal.archives-ouvertes.fr/hal-00574623>, à paraître dans Uniform Distribution Theory, 2011.
- [48] Élise Janvresse, Thierry de la Rue, and Yvan Velenik. Pinning by a sparse potential. *Stochastic Process. Appl.*, 115(8) :1323–1331, 2005.
- [49] Élise Janvresse, Thierry de la Rue, and Yvan Velenik. Self-similar corrections to the ergodic theorem for the Pascal-adic transformation. *Stoch. Dyn.*, 5(1) :1–25, 2005.
- [50] Élise Janvresse, Thierry de la Rue, and Yvan Velenik. A note on domino shuffling. *Electron. J. Combin.*, 13(1) :Research Paper 30, 19 pp. (electronic), 2006.
- [51] Élise Janvresse, Tom Meyerovitch, Emmanuel Roy, and Thierry de la Rue. Poisson suspensions and entropy for infinite transformations. *Trans. Amer. Math. Soc.*, 362(6) :3069–3094, 2010.
- [52] Élise Janvresse, Benoît Rittaud, and Thierry de la Rue. How do random Fibonacci sequences grow? *Probab. Theory Related Fields*, 142(3-4) :619–648, 2008.
- [53] Élise Janvresse, Benoît Rittaud, and Thierry de la Rue. Growth rate for the expected value of a generalized random Fibonacci sequence. *J. Phys. A*, 42(8) :085005, 18, 2009.
- [54] Élise Janvresse, Benoît Rittaud, and Thierry de la Rue. Dynamics of λ -continued fractions and β -shifts, 2011.
- [55] M. Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III*, pages 171–197. University of California Press, Berkeley and Los Angeles, 1956.
- [56] Mark Kac. *Probability and related topics in physical sciences*, volume 1957 of *With special lectures by G. E. Uhlenbeck, A. R. Hibbs, and B. van der Pol. Lectures in Applied Mathematics. Proceedings of the Summer Seminar, Boulder, Colo.* Interscience Publishers, London-New York, 1959.
- [57] Steven Arthur Kalikow. Twofold mixing implies threefold mixing for rank one transformations. *Ergodic Theory Dynam. Systems*, 4(2) :237–259, 1984.
- [58] Donald E. Knuth. *The art of computer programming. Vol. 2.* Addison-Wesley Publishing Co., Reading, Mass., second edition, 1981. Seminumerical algorithms, Addison-Wesley Series in Computer Science and Information Processing.
- [59] Ulrich Krengel. Entropy of conservative transformations. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 7 :161–181, 1967.
- [60] Serge Lang. *Introduction to Transcendental Numbers.* Addison-Wesley Publishing Company, 1966.
- [61] François Ledrappier. Un champ markovien peut être d’entropie nulle et mélangeant. *C. R. Acad. Sci. Paris Sér. A-B*, 287(7) :A561–A563, 1978.
- [62] V. P. Leonov. The use of the characteristic functional and semi-invariants in the ergodic theory of stationary processes. *Soviet Math. Dokl.*, 1 :878–881, 1960.
- [63] Armin Leutbecher. Über die Heckeschen Gruppen $G(\lambda)$. *Abh. Math. Sem. Univ. Hamburg*, 31 :199–205, 1967.
- [64] Armin Leutbecher. Über die Heckeschen Gruppen $G(\lambda)$. II. *Math. Ann.*, 211 :63–86, 1974.

- [65] Sheng Lin Lu and Horng-Tzer Yau. Spectral gap and logarithmic Sobolev inequality for Kawasaki and Glauber dynamics. *Comm. Math. Phys.*, 156(2) :399–433, 1993.
- [66] Edouard Lucas. *Recherches sur plusieurs ouvrages de Léonard de Pise et sur diverses questions d'arithmétique supérieure*. Imprimerie des sciences mathématiques et physiques, 1877.
- [67] Eran Makover and Jeffrey McGowan. An elementary proof that random Fibonacci sequences grow exponentially. *J. Number Theory*, 121(1) :40–44, 2006.
- [68] David K. Maslen. The eigenvalues of Kac's master equation. *Math. Z.*, 243(2) :291–331, 2003.
- [69] Xavier Méla. *Dynamical Properties of the Pascal Adic and Related Systems*. PhD thesis, Chapel Hill, 2002.
- [70] Xavier Méla. A class of nonstationary adic transformations. *Ann. Inst. H. Poincaré Probab. Statist.*, 42(1) :103–123, 2006.
- [71] Xavier Méla and Karl Petersen. Dynamical properties of the Pascal adic transformation. *Ergodic Theory Dynam. Systems*, 25(1) :227–256, 2005.
- [72] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Communications and Control Engineering Series. Springer-Verlag London Ltd., London, 1993.
- [73] S.J. Miller and M.J. Nigrini. The modulo 1 central limit theorem and Benford's law for products. *Int. J. Algebra*, 2(1-4) :119–130, 2008.
- [74] R. Modzelewski and al. A new method (the random walk) for heterogeneity quantification of brain spect perfusion images. *EANM 04 - Annual Congress of the European Association of Nuclear Medicine (Helsinki)*, 2004.
- [75] R. Modzelewski, T. de la Rue, É. Janvresse, A. Hitzel, J.-F. Ménard, I. Gardin, E. Gerardin, A. Manrique, D. Hannequin, S. Lepretre, H. Tilly, and P. Vera. Development and validation of the random walk algorithm : Application to the classification of diffuse heterogeneity in brain spect perfusion images. *Journal of Computer Assisted Tomography*, 32(4) :651–659, 2008.
- [76] R. Modzelewski, É. Janvresse, T. de la Rue, and P. Vera. Brain perfusion heterogeneity measurement based on random walk algorithm : Choice and influence of inner parameters. *Comput. Med. Imaging Graph.*, 34(4) :289–297, 2010.
- [77] R. Modzelewski, É. Janvresse, T. de la Rue, and P. Vera. Heterogeneity quantification methods for brain spect perfusion images. Prépublication, 2011.
- [78] Simon Newcomb. Note on the Frequency of Use of the Different Digits in Natural Numbers. *Amer. J. Math.*, 4(1-4) :39–40, 1881.
- [79] Donald S. Ornstein, Daniel J. Rudolph, and Benjamin Weiss. Equivalence of measure preserving transformations. *Mem. Amer. Math. Soc.*, 37(262) :xii+116, 1982.
- [80] W. Parry. On the β -expansions of real numbers. *Acta Math. Acad. Sci. Hungar.*, 11 :401–416, 1960.
- [81] William Parry. *Entropy and generators in ergodic theory*. W. A. Benjamin, Inc., New York-Amsterdam, 1969.
- [82] Karl Petersen and Klaus Schmidt. Symmetric Gibbs measures. *Trans. Amer. Math. Soc.*, 349(7) :2775–2811, 1997.
- [83] Roger S. Pinkham. On the distribution of first significant digits. *Ann. Math. Statist.*, 32 :1223–1230, 1961.
- [84] James Propp. Generalized domino-shuffling. *Theoret. Comput. Sci.*, 303(2-3) :267–301, 2003. Tilings of the plane.
- [85] R. A. Raimi. The first digit problem. *Amer. Math. Monthly*, 83(7) :521–538, 1976.
- [86] A. Rényi. Representations for real numbers and their ergodic properties. *Acta Math. Acad. Sci. Hungar*, 8 :477–493, 1957.

- [87] Benoît Rittaud. On the average growth of random Fibonacci sequences. *J. Integer Seq.*, 10(2) :Article 07.2.4, 32 pp. (electronic), 2007.
- [88] Christian Robert. *Méthodes de Monte Carlo par chaînes de Markov*. Statistique Mathématique et Probabilité. [Mathematical Statistics and Probability]. Éditions Économica, Paris, 1996.
- [89] V. A. Rohlin. On endomorphisms of compact commutative groups. *Izvestiya Akad. Nauk SSSR. Ser. Mat.*, 13 :329–340, 1949.
- [90] David Rosen. A class of continued fractions associated with certain properly discontinuous groups. *Duke Math. J.*, 21 :549–563, 1954.
- [91] Jeffrey S. Rosenthal. Random rotations : characters and random walks on $SO(N)$. *Ann. Probab.*, 22(1) :398–423, 1994.
- [92] V. V. Ryzhikov. Joinings and multiple mixing of the actions of finite rank. *Funktsional. Anal. i Prilozhen.*, 27(2) :63–78, 96, 1993.
- [93] Clément Sire and Paul L. Krapivsky. Random fibonacci sequences. *J. Phys. A*, 34 :9065–9083, 2001.
- [94] T. Takagi. A simple example of the continuous function without derivative. *Proc. Phys. Math. Japan*, 1 :176–177, 1903.
- [95] Haruo Totoki. The mixing property of Gaussian flows. *Mem. Fac. Sci. Kyushu Univ. Ser. A*, 18 :136–139, 1964.
- [96] A. V. Ustinov. A short proof of Euler’s identity for continuants. *Mat. Zametki*, 79(1) :155–156, 2006.
- [97] A. M. Vershik and A. N. Livshits. Adic models of ergodic transformations, spectral theory, substitutions, and related topics. In *Representation theory and dynamical systems*, volume 9 of *Adv. Soviet Math.*, pages 185–204. Amer. Math. Soc., Providence, RI, 1992.
- [98] A.M. Vershik. A theorem on the Markov approximation in ergodic theory. *Journ. of Soviet Math.*, 28 :667–674, 1985.
- [99] Cédric Villani. *Contribution à l’étude mathématique des collisions en théorie cinétique*. PhD thesis, Univ. Paris-Dauphine, 2000.
- [100] Divakar Viswanath. Random Fibonacci sequences and the number $1.13198824\dots$. *Math. Comp.*, 69(231) :1131–1155, 2000.
- [101] Michel Waldschmidt. Séminaire sur les nombres transcendants 1972-73. *Publ. Math. Orsay*, 57, 1973.