

Chapitre 1

Suites récurrentes aléatoires

À l'exception de la première section, ce chapitre n'est pas essentiel à la suite du cours. Le but ici est surtout d'éveiller la curiosité du lecteur en lui présentant quelques exemples très simples à décrire mais dont la richesse des comportements est remarquable. Nous nous sommes inspirés à plusieurs reprises d'un bel article de Diaconis et Freedman [41] sur les "itérations aléatoires de fonctions" dont la lecture est très abordable.

1.1 Suites récurrentes aléatoires

Une *suite récurrente aléatoire* sur M est une suite de variables aléatoires (X_n) à valeurs dans M définie sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})^1$ solution d'une équation récurrente de la forme

$$X_{n+1} = f(\theta_{n+1}, X_n) \tag{1.1}$$

où :

- (a) $\theta_1, \theta_2, \dots$ sont des variables aléatoires indépendantes identiquement distribuées à valeurs dans un espace mesurable (Θ, \mathcal{A}) ;

¹Rappelons que Ω est un ensemble, \mathcal{F} une tribu sur Ω et $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ une probabilité (voir l'appendice).

(b)

$$\begin{aligned} f : \Theta \times M &\rightarrow M, \\ (\theta, x) &\mapsto f(\theta, x) = f_\theta(x) \end{aligned}$$

est une application mesurable ;

(c) X_0 (la condition initiale) est une variable aléatoire (éventuellement déterministe) indépendante de la suite (θ_n) .

Définition 1.1.1 Si m désigne la loi des variables (θ_n) , le couple (f, m) s'appelle un **système dynamique aléatoire** et la suite (X_n) solution de (1.1) une *trajectoire* du système (f, m) de condition initiale X_0 . La loi de X_0 est appelée la *loi initiale* de (f, m) .

Remarque 1.1.2 Dans le cas particulier où f ne dépend pas de θ , la récurrence (1.1) est déterministe. Son étude fait l'objet de la *théorie des systèmes dynamiques* (à temps discret) lorsque X_0 est déterministe, et de la *théorie ergodique* lorsque X_0 est aléatoire.

Commençons par quelques exemples.

1.2 Battre les cartes

Un joueur de poker qui bat un jeu standard de 52 cartes nous fournit un bel exemple de suite récurrente aléatoire sur le groupe symétrique $M = S_{52}$ des 52! permutations (les bijections de $\{1, \dots, 52\}$ dans lui-même). À l'instant initial, les cartes sont ordonnées et l'état initial du système est la permutation identité : $X_0 = e$. À l'étape $n + 1$, le joueur choisit une permutation θ_{n+1} au hasard dans un sous-ensemble $S \subset M$ et l'applique à X_n . La suite (X_n) est donc une suite récurrente aléatoire

$$X_{n+1} = \theta_{n+1} \circ X_n$$

où les variables θ_i sont i.i.d. à valeurs dans S .

L'ensemble S caractérise le type de "battage" et dépend de la dextérité du joueur. Typiquement une permutation de S est obtenue de la façon suivante : on choisit un

entier entre 1 et 52 (on coupe) puis on intercale les cartes $1, 2, \dots, k$ au milieu des cartes $k + 1, \dots, 52$.

Nous ne chercherons pas ici à décrire l'ensemble S avec précision mais nous supposerons que le battage est "mélangeant", c'est-à-dire

- (a) S engendre M : toute permutation $x \in M$ peut s'écrire comme un produit fini $x = s_1 \circ \dots \circ s_l$ d'éléments de S ;
- (b) la loi m des variables θ_n charge tous les éléments de S : $m(s) > 0$ pour tout $s \in S$.

Sous ces hypothèses, la suite (X_n) vérifie les propriétés (très intuitives) suivantes :

- (i) elle est *récurrente* : elle visite infiniment souvent tous les points de l'espace d'états ;
- (ii) elle satisfait une *loi des grands nombres* : la fréquence d'occupation d'un sous-ensemble $A \subset M$ est asymptotiquement égale au "volume" de $A = \text{card}(A)/52!$.

Plus précisément,

Proposition 1.2.1 *Pour tout $x \in M$,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mathbf{1}_{\{X_k=x\}} = \frac{1}{52!} = \frac{1}{\mathbb{E}(T_e)}$$

où

$$T_e = \inf\{n > 0 : X_n = e\}$$

désigne le premier instant de retour à e .

Cet énoncé est un cas particulier du *théorème ergodique presque sûr* pour les chaînes de Markov finies (*i.e.* sur un espace d'états fini) que nous démontrerons dans la section 2.2 du chapitre 2. Nous conseillons l'exercice suivant au lecteur trop impatient pour attendre le chapitre 2.

Exercice 1.2.2 Pour tout $i \geq 1$, notons Δ^i l'intervalle de temps écoulé entre la i^{e} et la $(i-1)^{\text{e}}$ visite à e , c'est-à-dire

$$\Delta^i = \inf\{k > 0 : \theta_{T^{i-1}+k} \circ \dots \circ \theta_{T^{i-1}+1} = e\},$$

où $T^0 = 0$ et $T^i = \Delta^i + \dots + \Delta^1$.

1) Montrer que les variables Δ^i sont indépendantes et de même loi. En déduire que

$$\lim_{i \rightarrow \infty} \frac{i}{T^i} = \frac{1}{E(T_e)}.$$

2) Montrer que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbf{1}_{\{X_k=e\}} = \lim_{i \rightarrow \infty} \frac{i}{T^i}.$$

3) Prouvez la proposition 1.2.1.

1.2.1 Oui, mais combien de fois ?

La proposition 1.2.1 est un résultat asymptotique, et de ce fait n'est pas entièrement satisfaisante pour le joueur de poker, qui préférerait savoir si le paquet est bien mélangé lorsqu'il cesse de battre les cartes. En d'autres termes (plus mathématiques), le joueur aimerait pouvoir comparer loi de X_n avec la loi uniforme sur M .

Introduisons la distance

$$d_n = \frac{1}{2} \sum_{x \in M} \left| P(X_n = x) - \frac{1}{52!} \right| \leq 1.$$

La proposition suivante donne un premier élément de réponse à la question du joueur.

Proposition 1.2.3 *Supposons que S soit tel que l'entier l de l'hypothèse (a) puisse être choisi indépendamment de x (il suffit par exemple de supposer que $e \in S$). Alors il existe des constantes $C > 0$ et $0 \leq \rho < 1$ (dépendantes de S) telles que*

$$d_n \leq C\rho^n.$$

Il s'agit là encore d'un résultat tout à fait typique des chaînes de Markov finies. Nous le démontrerons dans les chapitres 2 (théorème 2.2.14) et 3 (théorème 3.3.3) dans un cadre plus général. Malheureusement, les meilleures constantes C et ρ sont difficiles à calculer et les estimés faciles sont trop grossiers.

En 1986, Persis Diaconis² se rend célèbre en démontrant le résultat suivant :

²Persis Diaconis est professeur de mathématiques à Standford. Il a commencé sa carrière comme magicien professionnel avant de devenir mathématicien. Non seulement Diaconis démontre de jolis théorèmes sur le mélange des cartes mais il sait aussi les battre !

il suffit de battre 7 fois un jeu de 52 cartes.

Plus précisément, Diaconis démontre que le graphe de la fonction $n \rightarrow d_n$ est une “sigmoïde”. Pour $n < 7$, d_n est proche de 1 et pour $n \geq 8$, d_n est proche de 0. L’explication de ce phénomène (dit de “cut-off”) n’entre pas dans le cadre de ce cours mais le lecteur intéressé pourra consulter Diaconis [16], ou encore l’excellent cours de Saloff-Coste [28].

1.3 Marches aléatoires sur \mathbb{Z}^d

Soient e_1, \dots, e_d les vecteurs de la base canonique de \mathbb{R}^d . Un marcheur (sans doute ivre!) se déplace dans \mathbb{Z}^d . Sa position à l’instant n est la variable aléatoire $X_n \in \mathbb{Z}^d$ définie par la récurrence aléatoire : $X_0 = 0$, et

$$X_{n+1} = X_n + \theta_{n+1}$$

où $\theta_1, \theta_2, \dots$ sont des v.a. i.i.d uniformément distribuées dans $\{e_1, \dots, e_d, -e_1, \dots, -e_d\}$.

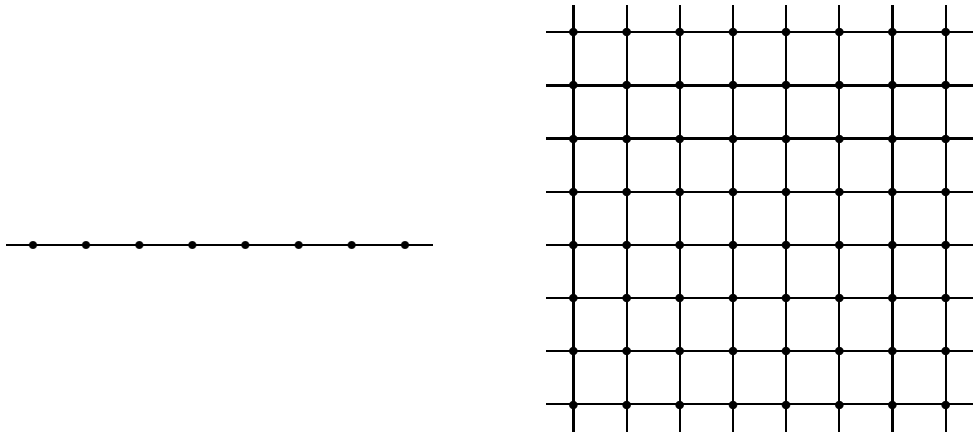


Figure 1.1 Les réseaux \mathbb{Z} et \mathbb{Z}^2 .

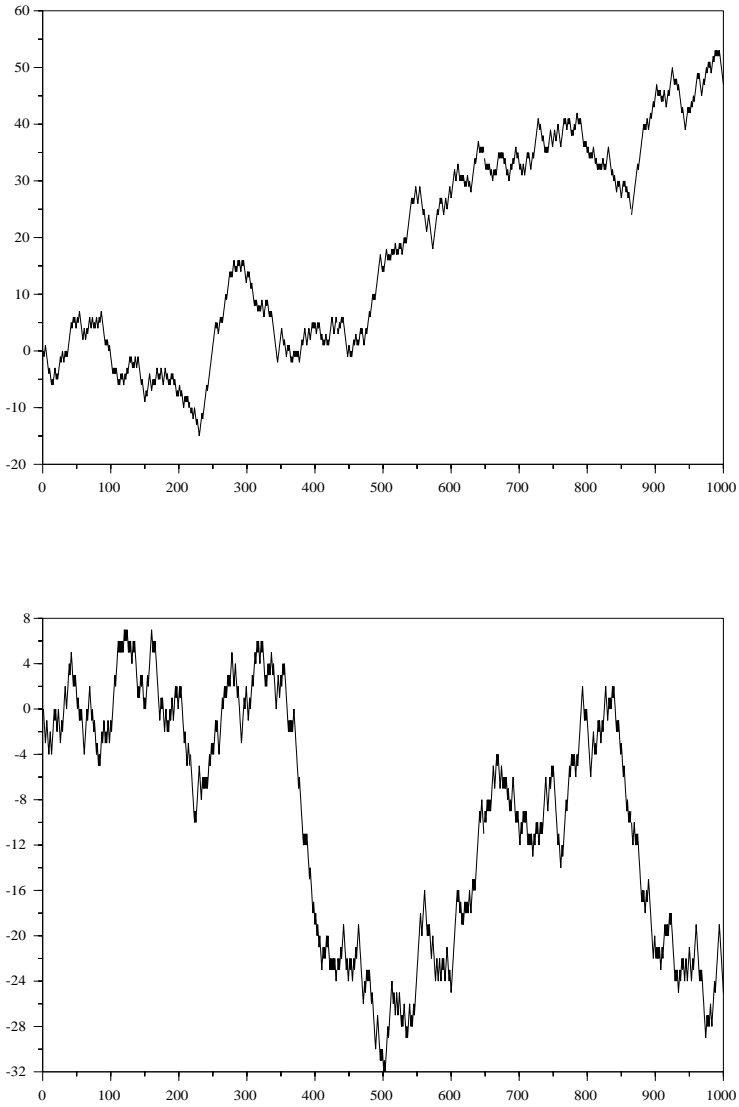


Figure 1.2 *Marches aléatoires sur \mathbb{Z} .*
Le temps \mathbb{N} est en abscisse et l'espace \mathbb{Z} en ordonnée.

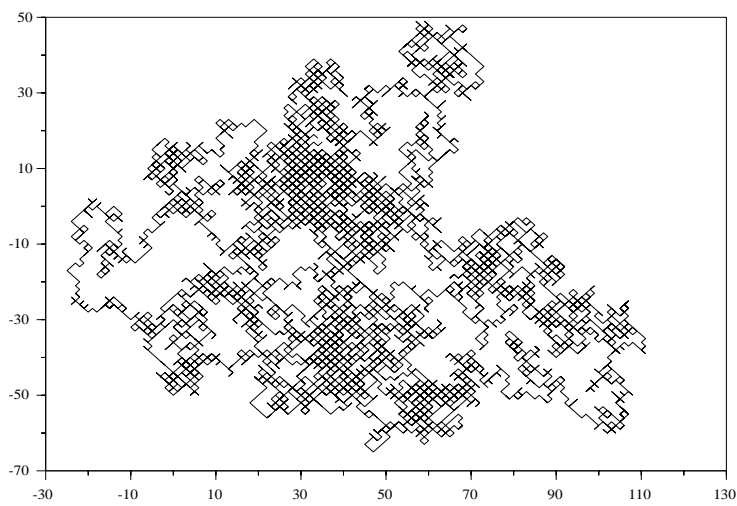
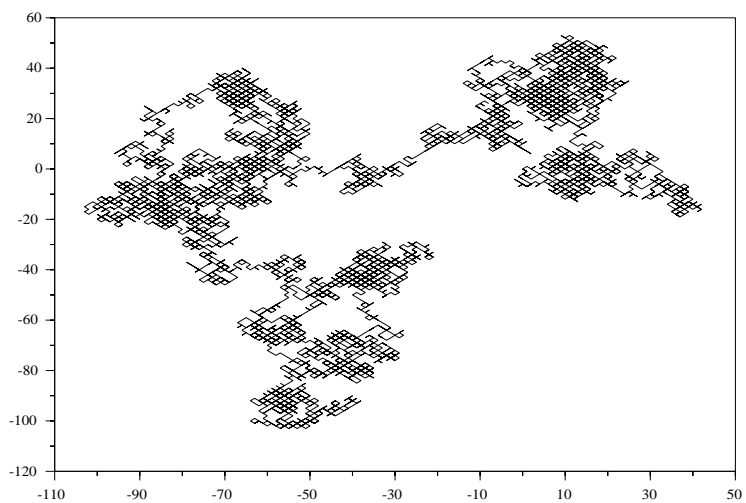


Figure 1.3 *Marches aléatoires sur \mathbb{Z}^2 .*

En 1921, Polya³ se pose le problème suivant : le marcheur est-il sûr de revenir à sa position de départ ? La réponse dépend crucialement de la dimension du réseau :

Théorème 1.3.1 (Polya, 1921). *Pour $d = 1$ ou 2 la marche aléatoire (X_n) est récurrente :*

$$P(\exists n > 0 : X_n = 0) = 1.$$

Pour $d \geq 3$, elle est transiente :

$$P(\exists n > 0 : X_n = 0) < 1.$$

Nous démontrerons ce résultat dans la section 2.3.2 du chapitre 2.

1.3.1 La constante de Polya

La *constante de Polya* est le nombre

$$p(d) = P(\exists n > 0 : X_n = 0).$$

D'après le théorème 1.3.1, $p(1) = p(2) = 1$ et $p(d) < 1$ pour $d \geq 3$. Des estimations numériques précises de $p(d)$ sont disponibles sur le site internet <http://www.mathsoft.com.alsove/constant/polya/polya.html> pour $d \in \{3, \dots, 64\}$. En voici quelques-unes

$$p(3) = 0,3405373295509991\dots$$

$$p(4) = 0,1932016732249839\dots$$

$$p(10) = 0,0561975359742678\dots$$

$$p(64) = 0,0079380451778596\dots$$

³Mathématicien hongrois né en 1887 et mort en 1985, George Polya a travaillé en théorie des nombres, probabilités, combinatoire et en analyse. Son résultat remarquable sur la marche aléatoire de \mathbb{Z}^d a inspiré de nombreux travaux, et est à l'origine de la théorie des marches aléatoires sur les graphes et les groupes infinis.

1.4 Convolution de Bernoulli

Voici un exemple très simple (mais très riche) de système dynamique aléatoire sur \mathbb{R} . Nous l'obtenons en modifiant légèrement la suite récurrente décrivant la marche de Polya unidimensionnelle.

Soit a un réel positif, $(\theta_n)_{n \geq 1}$ une suite de variables aléatoires i.i.d. de loi uniforme sur $\{-1, 1\}$ et (X_n) la suite récurrente aléatoire réelle définie par

$$X_{n+1} = aX_n + \theta_{n+1}.$$

Pour $a = 1$ nous retrouvons la marche de Polya unidimensionnelle étudiée dans la section précédente. Pour $a \neq 1$, la suite (X_n) s'obtient à partir de X_0 en composant aléatoirement les deux applications affines $f_1(x) = ax + 1$ et $f_{-1}(x) = ax - 1$.

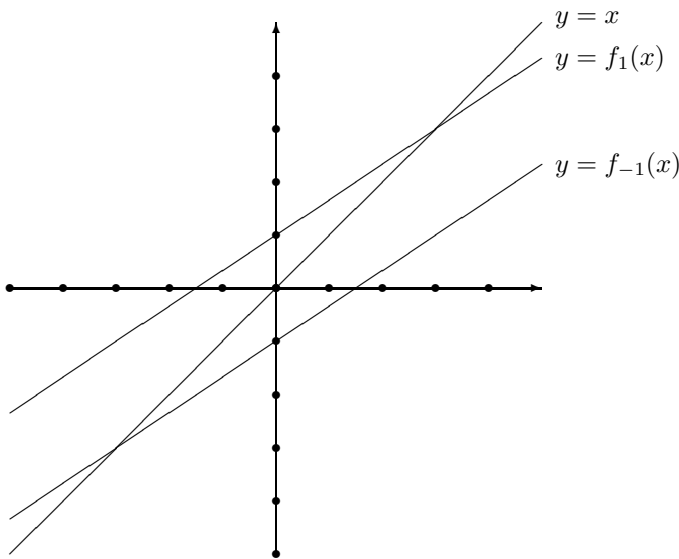


Figure 1.4 Les fonctions f_1 et f_{-1} pour $a = 2/3$.

Pour $a > 1$, le comportement asymptotique de (X_n) présente peu d'intérêt.

Exercice 1.4.1 Montrer que pour $a > 1$

$$P(\lim_{n \rightarrow \infty} X_n \in \{-\infty, +\infty\}) = 1.$$

Pour $0 < a < 1$, nous allons voir que le comportement asymptotique de la suite (X_n) est beaucoup plus riche que ne pourrait le laisser supposer la simplicité du modèle.

Pour tout $n \geq 0$,

$$X_n - a^n X_0 = a^{n-1} \theta_1 + \dots + \theta_n.$$

Soit

$$Y_n = \theta_1 + a\theta_2 + \dots + a^{n-1}\theta_n.$$

Puisque $|a| < 1$, la suite (Y_n) converge vers une variable $Y_\infty \in [-\frac{1}{1-a}, \frac{1}{1-a}]$.

Remarquons que les variables $X_n - a^n X_0$ et Y_n ont la même loi car les vecteurs $(\theta_1, \dots, \theta_n)$ et $(\theta_n, \dots, \theta_1)$ ont la même loi.

Pour toute fonction $\Phi : \mathbb{R} \rightarrow \mathbb{R}$ continue,

$$\lim_{n \rightarrow \infty} (\Phi(X_n) - \Phi(X_n - a^n X_0)) = 0,$$

car Φ restreinte à $[-\frac{1}{1-a} - |X_0|, \frac{1}{1-a} + |X_0|]$ est uniformément continue. Par conséquent,

$$\lim_{n \rightarrow \infty} \mathbb{E}(\Phi(X_n)) = \lim_{n \rightarrow \infty} \mathbb{E}(\Phi(Y_n)) = \mathbb{E}(\Phi(Y_\infty)), \quad (1.2)$$

par application du théorème de convergence dominée (voir l'appendice, théorème 8.2.3). Ce type de convergence s'appelle la *convergence en loi* (voir l'appendice) et nous venons donc de prouver l'énoncé suivant.

Proposition 1.4.2 *La suite (X_n) converge en loi vers la variable aléatoire Y_∞ .*

Soit ν_a la loi de Y_∞ et $F_a : \mathbb{R} \rightarrow [0, 1]$ sa *fonction de répartition* :

$$F_a(t) = \nu_a([-\infty, t]) = \mathbb{P}(Y_\infty \leq t).$$

Si F_a est continue au point t (c'est-à-dire si $\mathbb{P}(Y_\infty = t) = 0$), la suite de fonctions $(\mathbf{1}_{\{Y_n \leq t\}})_{n \geq 0}$ converge presque sûrement vers $\mathbf{1}_{\{Y_\infty \leq t\}}$ et donc, par convergence dominée,

$$F_a(t) = \lim_{n \rightarrow \infty} \mathbb{P}(Y_n \leq t) = \lim_{n \rightarrow \infty} \mathbb{P}(X_n \leq t). \quad (1.3)$$

Comme nous verrons bientôt que $F_a(t)$ est continue en tout point, cette propriété de convergence (1.3) est vraie sans restriction. De plus, nous démontrerons dans le chapitre 4 (corollaire 4.4.12) l'énoncé presque sûr suivant.

Proposition 1.4.3 *Pour tout intervalle $I = [\alpha, \beta] \subset \mathbb{R}$,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbf{1}_{\{X_k \in I\}} = \nu_a(I) = F_a(\beta) - F_a(\alpha)$$

presque sûrement.

Les propositions 1.4.2, 1.4.3 et l'équation (1.3) nous éclairent sur le comportement asymptotique de la suite (X_n) . Elles montrent que la suite (X_n) se comporte en première approximation, pour n grand, comme une suite de variables i.i.d. distribuées selon F_a .

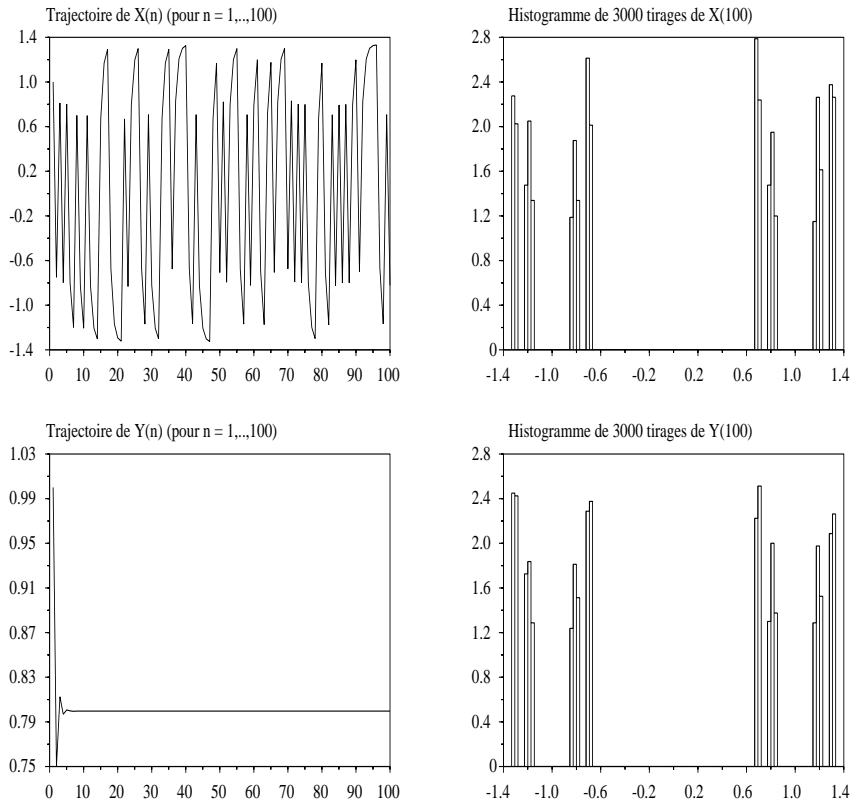


Figure 1.5 *Convolution de Bernoulli pour $a = 0,25$.*

Les résultats de simulations (figures 1.5 et 1.6) illustrent la différence de nature entre les suites (X_n) et (Y_n) . Alors que la suite (Y_n) converge presque sûrement (vers la v.a Y_∞) la suite (X_n) “oscille” aléatoirement. Par contre les histogrammes de (Y_n) et (X_n) “convergent” vers le même histogramme. Cela s’explique par la convergence en loi de (X_n) vers Y_∞ (proposition 1.4.2).

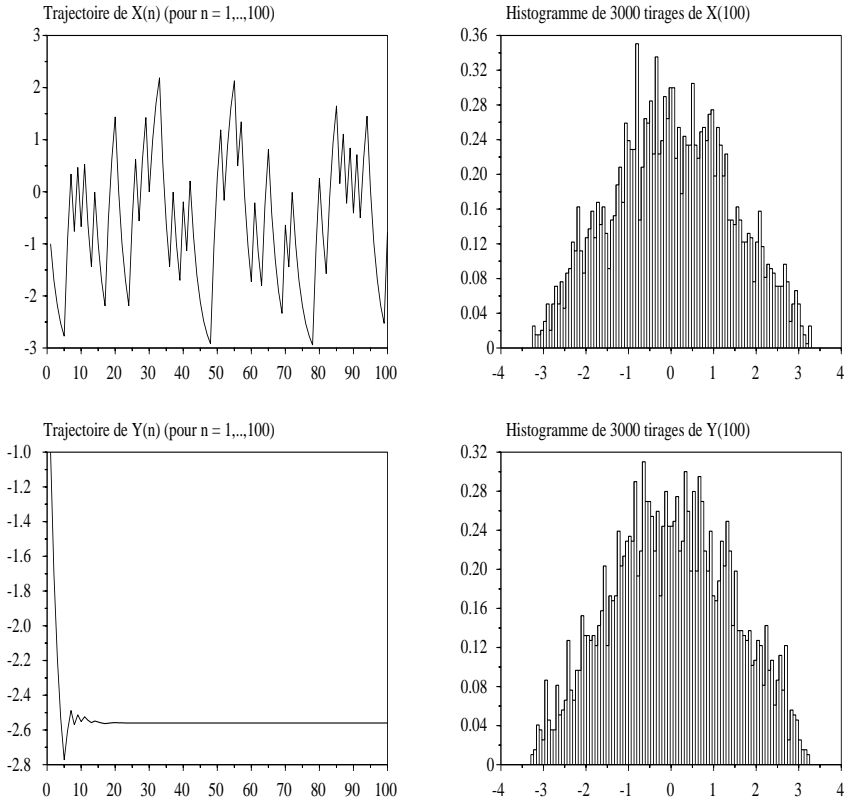


Figure 1.6 *Convolution de Bernoulli pour $a = 0,7$.*

Un problème fascinant, vieux de 70 ans environ, et toujours (partiellement) ouvert est de comprendre les propriétés fines de F_a (et ν_a). Nous allons discuter (et démontrer) quelques-unes de ces propriétés dans les paragraphes suivants.

1.4.1 Autosimilarité

Une première conséquence de (1.3) est la propriété *d'autosimilarité* ci-après.

Proposition 1.4.4 F_a est l'unique fonction de répartition solution de l'équation fonctionnelle

$$\forall t \in \mathbb{R}, G(t) = \frac{1}{2} \left[G\left(\frac{t-1}{a}\right) + G\left(\frac{t+1}{a}\right) \right]. \quad (1.4)$$

De plus, la fonction F_a est continue.

DÉMONSTRATION :

En effet, puisque la v.a. $Y_\infty^1 = \sum_{k \geq 2} a^{k-2} \theta_k$ est indépendante de θ_1 , et a même loi que Y_∞ , nous pouvons décomposer Y_∞ en

$$Y_\infty = aY_\infty^1 + \theta_1.$$

Par indépendance, nous voyons que

$$P(Y_\infty \leq t) = \frac{1}{2} P\left(Y_\infty \leq \frac{t+1}{a}\right) + \frac{1}{2} P\left(Y_\infty \leq \frac{t-1}{a}\right)$$

qui est en fait l'équation de convolution recherchée.

Réciproquement, si G est une fonction de répartition vérifiant (1.4), choisissons X_0 de loi G . Alors

$$P(X_{n+1} \leq t) = P(aX_n + \theta_{n+1} \leq t) = \frac{1}{2} \left[P\left(X_n \leq \frac{t-1}{a}\right) + P\left(X_n \leq \frac{t+1}{a}\right) \right]$$

et une récurrence immédiate sur n montre que X_n a pour loi G . L'égalité $G = F_a$ se déduit alors de la proposition 1.4.2 et du théorème d'unicité des probabilités (voir l'appendice, théorème 8.1.22).

Pour la continuité de F_a , voir Jessen et Wintner [49].

Remarque 1.4.5 Soit G_n la fonction de répartition de X_n . Alors d'après la preuve de la proposition 1.4.4,

$$G_n = T_a^n(G_0)$$

où T_a est la fonction (définie sur l'ensemble des fonctions de répartition) par la formule

$$T_a(G)(t) = \frac{1}{2} \left[G\left(\frac{t-1}{a}\right) + G\left(\frac{t+1}{a}\right) \right].$$

De même, si G_0 admet une densité g_0 , alors G_n admet pour densité

$$g_n = R_a^n(g_0)$$

où

$$R_a(g)(t) = \frac{1}{2a} \left[g\left(\frac{t-1}{a}\right) + g\left(\frac{t+1}{a}\right) \right].$$

Grâce à cette remarque, il est facile d'écrire un algorithme pour estimer numériquement (puis représenter graphiquement) les fonctions G_n et g_n .

Ci-dessous sont représentées g_n et G_n pour différentes valeurs de a et différentes valeurs de n lorsque X_0 suit une loi uniforme sur $[0, 1]$. Pour comprendre le caractère pour le moins singulier des figures 1.7 et 1.8, il faut lire la section suivante.

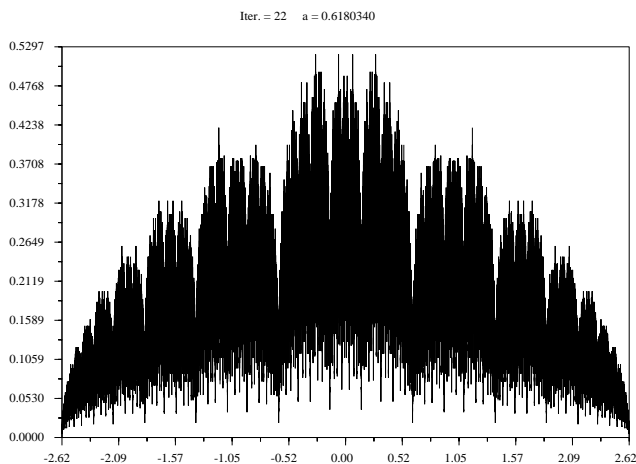


Figure 1.7 Graphe de g_{22} pour $a = 0,6180340$.

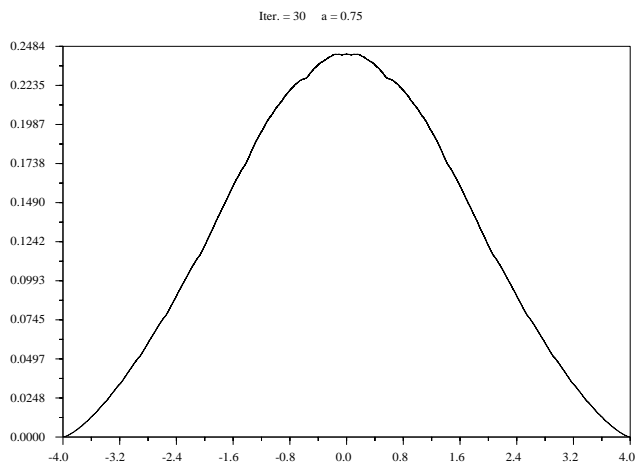


Figure 1.8 *Graphe de g_{30} pour $a = 0,75$.*

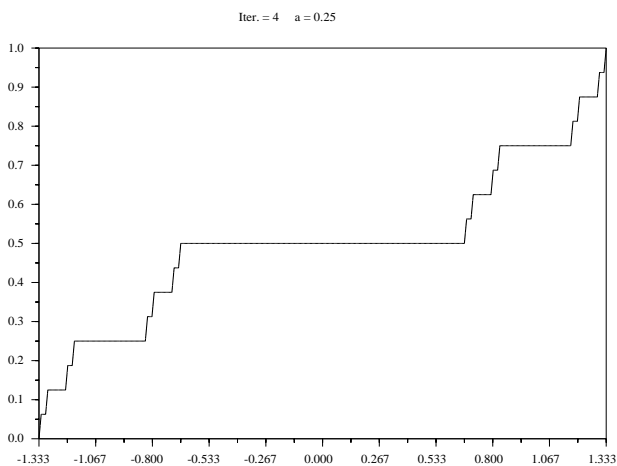


Figure 1.9 *Graphe de G_4 pour $a = 0,25$.*

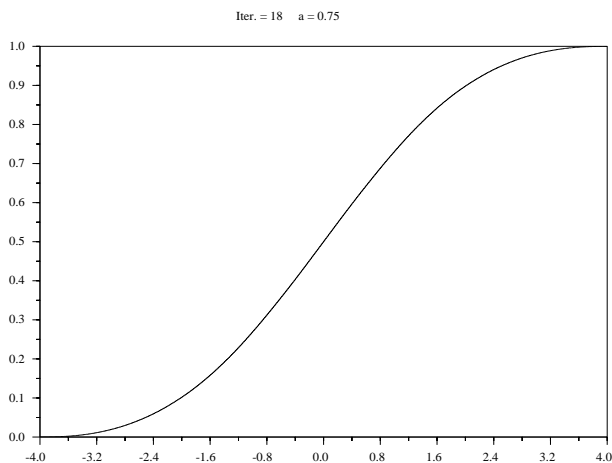


Figure 1.10 Graphe de G_{18} pour $a = 0,75$.

◇

... Pour en savoir plus ...

1.4.2 Propriétés fines

Le *support* de la loi ν_a est, par définition, l'ensemble $K_a \subset \mathbb{R}$ constitué des points $t \in \mathbb{R}$ tels que

$$\nu_a(J) > 0$$

pour tout intervalle ouvert J contenant t . De façon équivalente, K_a est l'ensemble des points $t \in \mathbb{R}$ où la fonction F_a est strictement croissante.

$$K_a = \{t \in \mathbb{R} : F_a(t_2) > F_a(t_1) \text{ pour tout } t_1 < t < t_2\}. \quad (1.5)$$

Il est facile de vérifier que K_a est un sous-ensemble fermé (c'est une conséquence immédiate de (1.5)) contenu dans

$$I_a = \left[-\frac{1}{1-a}, \frac{1}{1-a} \right],$$

car $Y_\infty \in I_a$ presque sûrement. De plus

$$K_a = f_1(K_a) \cup f_{-1}(K_a).$$

Cette propriété d'autosimilarité ensembliste est caractéristique des structures *fractales*. Elle se vérifie facilement, en utilisant (1.4).

Le lecteur de ce cours a déjà rencontré des lois discrètes et des lois à densité (voir Comets [3]) mais la typologie des lois sur \mathbb{R} est un peu plus compliquée. Une probabilité ν sur \mathbb{R} de fonction de répartition $F(t) = \nu(-\infty, t]$ est dite

- *continue*, si F est continue ;
- *absolument continue*, si

$$F(t) = \int_{-\infty}^t f(u) du$$

où f est une fonction intégrable, positive, d'intégrale 1 (appelée densité) ;

- *discrète*, si il existe un ensemble dénombrable t_1, t_2, \dots et des "poids" $p_i \geq 0$ tels que

$$\nu = \sum_i p_i \delta_{t_i}.$$

Dans ce cas F est en escaliers ;

- *singulière* si il existe un ensemble $N \subset \mathbb{R}$ de mesure de Lebesgue nulle et tel que $\nu(N) = 1$.

Remarquons qu'une loi discrète est singulière et qu'une loi absolument continue n'est jamais singulière. En général, et d'après un résultat de théorie de la mesure, une probabilité ν sur \mathbb{R} peut toujours s'écrire

$$\nu = \nu^{ac} + \nu^{cs} + \nu^d \tag{1.6}$$

où ν^{ac} est absolument continue, ν^{cs} est continue singulière (ou bien nulle) et ν^d est discrète (ou bien nulle). Une telle décomposition est unique.

Dans les années 30, le mathématicien Wintner et ses collaborateurs ont cherché à décrire le type de ν_a . Jessen et Wintner [49] démontrent en 1935 que ν_a est absolument continue **ou bien** continue singulière (voir proposition 1.4.6). Kershner et Wintner [50] remarquent que ν_a est singulière pour $a < \frac{1}{2}$ car K_a est un Cantor de mesure de Lebesgue nulle (voir proposition 1.4.8). Pour $a \geq \frac{1}{2}$, $K_a = I_a$ (lemme 1.4.11) et la fonction F_a est donc strictement croissante sur I_a . De plus, Wintner [56] remarque que $\nu_{1/2}$ est la loi uniforme sur $I_{1/2}$ (proposition 1.4.10) et prouve que F_a est dérivable de classe C^{k-1} pour $a = 2^{-1/k}$ et $k \geq 2$. Ces résultats donnent à penser que F_a est absolument continue pour $a \geq \frac{1}{2}$, mais Erdős [45] (théorème 1.4.13) en 1939 infirme cette conjecture en prouvant que ν_a est singulière quand a est l'inverse d'un *nombre de Pisot* ! Un nombre de Pisot est un *entier algébrique* (i.e. la racine d'un polynôme unitaire à coefficients entiers) dont les *conjugués* (i.e. les autres racines) sont de module < 1 . Par exemple, le nombre d'or $g = (\sqrt{5} + 1)/2$ est un nombre de Pisot car c'est la racine du polynôme $x^2 - x - 1$ dont le conjugué $(1 - \sqrt{5})/2$ a un module < 1 . De même les racines positives $\theta_1 \sim 1.32471$ de $x^3 - x - 1$ et $\theta_2 \sim 1.3802777$ de $x^4 - x^3 - 1$ sont des nombres de Pisot (ce sont d'ailleurs les seuls nombres de Pisot contenus dans $]1, \sqrt{2}[$. C'est ainsi que F_a est de classe C^1 pour $a = 2^{-1/2} \sim 0.70710$ mais n'est pas absolument continue pour $a = \frac{1}{\theta_2} \sim 0.72449$ ou $a = \frac{1}{\theta_1} \sim 0.754877$.

Proposition 1.4.6 (Loi des types purs, Jessen and Wintner, 1935 [49]).

Pour tout $0 < a < 1$, si ν_a n'est pas absolument continue elle est nécessairement singulière.

DÉMONSTRATION :

Soient $\nu_a = \nu_a^{ac} + \nu_a^{cs} + \nu_a^d$ la décomposition de ν_a (équation (1.6)) et $\tilde{\nu}_a = \tilde{\nu}_a^{ac} + \tilde{\nu}_a^{cs} + \tilde{\nu}_a^d$ la probabilité dont la fonction de répartition est

$$\tilde{F}_a(t) = \frac{1}{2} \left[F_a \left(\frac{t-1}{a} \right) + F_a \left(\frac{t+1}{a} \right) \right].$$

D'après (1.4), $F_a = \tilde{F}_a$, et donc $\nu_a^\alpha = \tilde{\nu}_a^\alpha$ pour tout $\alpha \in \{ac, cs, d\}$, par unicité de la décomposition (1.6). Comme de plus ν_a est continue, $\nu_a^d = \tilde{\nu}_a^d = 0$. Supposons ν_a non absolument continue. Alors ν_a^{cs} est non nulle et la fonction

$$F_a^{cs}(t) = \frac{1}{\nu_a^{cs}(\mathbb{R})} \nu_a^{cs}([-\infty, t])$$

est une fonction de répartition solution de l'équation fonctionnelle (1.4). Elle est donc nécessairement égale à F_a d'après la proposition 1.4.4. Ce qui prouve que $\nu_a = \nu_a^{cs}$.

Remarque 1.4.7 La preuve originelle de Jessen et Wintner de la proposition ci-dessus est différente car Jessen et Wintner voyaient ν_a comme la loi de Y_∞ et non pas comme la loi limite de la récurrence aléatoire (X_n) . Or, c'est cette dernière interprétation qui nous a permis de démontrer l'**unicité** de la solution de l'équation (1.4) (voir la proposition 1.4.4) sur laquelle repose entièrement la preuve donnée ici de la proposition 1.4.6.

Proposition 1.4.8 (Kershner and Wintner, 1935 [50]). *Pour $a < \frac{1}{2}$, ν_a est continue singulière. Plus précisément, l'ensemble K_a est un Cantor de mesure de Lebesgue nulle.*

DÉMONSTRATION :

Pour $a < \frac{1}{2}$, Y_n prend au plus 2^{n-1} valeurs distinctes et le reste $r_n = Y_\infty - Y_n$ a un module $\leq \frac{a^{n-1}}{a-1}$. La mesure ν_a est donc concentrée sur une réunion d'intervalles de longueur totale $l_n = 2^{n-1}a^{n-1}/(1-a)$. Comme $l_n \rightarrow 0$ quand $n \rightarrow \infty$, K_a est de mesure de Lebesgue nulle et ν_a est singulière.

Remarque 1.4.9 Pour $a < \frac{1}{2}$, la fonction F_a est un *escalier du diable* : une fonction continue croissante non constante mais presque partout constante ! Le graphe de la fonction de répartition de Y_n donne, pour n assez grand, une idée de l'allure d'une telle fonction.

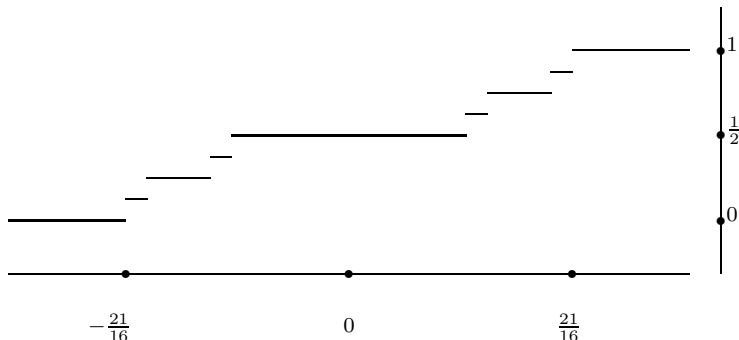


Figure 1.11 La fonction de répartition de Y_3 pour $a = \frac{1}{4}$.

Le cas $a = \frac{1}{2}$ est très particulier.

Proposition 1.4.10 (Wintner, 1935 [56]). $\nu_{1/2}$ est la probabilité uniforme sur $[-2, 2]$.

DÉMONSTRATION :

Soit $\Phi_t(Y_\infty) = \mathbf{E}(e^{itY_\infty})$ la fonction caractéristique de Y_∞ . Alors

$$\Phi_t(Y_\infty) = \prod_{k=0}^{\infty} \frac{1}{2} (e^{ita^k} + e^{-ita^k}) = \prod_{k=0}^{\infty} \cos(ta^k). \tag{1.7}$$

Pour $a = \frac{1}{2}$, il s'agit de la fonction caractéristique de la loi uniforme sur $[-2, 2]$. En effet

$$\begin{aligned} \frac{1}{4} \int_{-2}^2 e^{itx} dx &= \frac{\sin(2t)}{2t} = \frac{\sin(t) \cos(t)}{t} = \frac{\sin(2(t/2))}{2(t/2)} \cos(t) = \\ &\dots = \frac{2^n}{t} \sin(t/2^n) \prod_{k=0}^n \cos(t/2^k) = \prod_{k=0}^{\infty} \cos(t(\frac{1}{2})^k). \end{aligned}$$

Nous passons maintenant au cas $a > \frac{1}{2}$.

Lemme 1.4.11 Pour $a \geq \frac{1}{2}$, $K_a = I_a$. La fonction F_a est donc strictement croissante sur I_a .

DÉMONSTRATION :

Nous allons montrer que l'ensemble

$$S = \left\{ \sum_{k=0}^n a^k \xi_k, n \geq 0, \xi_0, \dots, \xi_n \in \{-1, 1\} \right\}$$

est dense dans $\left[-\frac{1}{1-a}, \frac{1}{1-a}\right]$. Par conséquent, pour tout intervalle

$$J =]\alpha, \beta[\subset [-(1-a)^{-1}, (1-a)^{-1}]$$

et $N \in \mathbb{N}$ tel que

$$\alpha + a^N/(1-a) < \beta - a^N/(1-a)$$

il existe $n \geq N$ tel que

$$P(Y_n \in]\alpha + a^N/(1-a), \beta - a^N/(1-a)[) \geq \frac{1}{2} > 0.$$

D'où

$$P(Y_\infty \in J) > 0.$$

Nous démontrons maintenant la densité de S . Posons $I_+ = [(1-2a)/(1-a), 1/(1-a)]$ et $I_- = [-1/(1-a), (2a-1)/(1-a)]$. Alors $I_a = I_+ \cup I_-$, $f_{-1}(I_a) = I_-$ et $f_1(I_a) = I_+$. Maintenant, étant donné $x \in I$, soit $(x_n, \xi_n) \in I \times \{-1, 1\}$ la suite définie par

(a) $x = x_0$;

(b) si $x_n \in I_+$ alors $\xi_n = +1$, sinon $\xi_n = -1$.

Alors $x_{n+1} = \frac{x_n - \xi_n}{a}$ (i.e. $f_{\xi_n}(x_{n+1}) = x_n$). D'où

$$a^{n+1}x_{n+1} - x = -\sum_{k=0}^{n+1} a^k \xi_k.$$

Et donc

$$\left| x - \sum_{k=0}^{n+1} a^k \xi_k \right| \leq a^n/(1-a).$$

Pour $a > \frac{1}{2}$, la fonction F_a semble donc plutôt régulière, mais comme nous l'avons déjà signalé, Erdős [45] en 1939 a prouvé que ν_a est singulière pour $a = \frac{1}{g}$ et g est nombre de Pisot, c'est-à-dire

(a) g est la racine d'un polynôme unitaire $P(X) = X^m + a_{m-1}X^{m-1} + \dots + a_0$ à coefficients dans \mathbb{Z} ;

(b) les autres racines g_2, \dots, g_m de P sont de module < 1 .

La preuve d'Erdős repose sur un lemme classique d'analyse que nous rappelons sans démonstration. Le lecteur pourra facilement prouver ce résultat en s'inspirant de l'exercice 3.2.12 de Bony [1].

Lemme 1.4.12 (Lemme de Riemann-Lebesgue). Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction intégrable. Alors

$$\lim_{t \rightarrow \infty} \int_{\mathbb{R}} f(x) e^{itx} dx = 0.$$

Théorème 1.4.13 (Erdős, 1939 [45]). Soit $1 < g < 2$ un nombre de Pisot et $a = 1/g$. Alors ν_a est singulière.

DÉMONSTRATION :

En vertu du lemme de Riemann-Lebesgue, il suffit de prouver que la fonction caractéristique de Y_∞ ne converge pas vers 0 quand $t \rightarrow \infty$.

L'expression $g^n + g_2^n + \dots + g_m^n$ est un polynôme symétrique des racines et peut donc s'écrire comme un polynôme à coefficients dans \mathbb{Z} des fonctions symétriques élémentaires $\sum g_i = -a_{m-1}$, $\sum_{i < j} g_i g_j = a_{m-2}, \dots, a_0 = \sum_{i_1 < \dots < i_m} a_{i_1} \dots a_{i_m}$. Ainsi $g^n + g_2^n + \dots + g_m^n \in \mathbb{Z}$ et

$$\text{dist}(g^n, \mathbb{Z}) \leq \theta^n$$

pour un certain $0 < \theta < 1$.

Soit $a = 1/g$ et $t(n) = \pi g^n$. D'après (1.7),

$$\Phi_{t(n)}(Y_\infty) = \prod_{k=0}^{\infty} \cos(g^{n-k}\pi) = \prod_{k=0}^n \cos(g^k\pi) C$$

où

$$C = \prod_{k=1}^{\infty} \cos(a^k\pi)$$

est non nul, car $\cos(a^k\pi) \neq 0$.

Pour tout entier $p \in \mathbb{Z}$,

$$||\cos(g^k\pi) - \cos(p\pi)|| \leq \pi|g^k - p|,$$

et donc, par définition de $\theta < 1$,

$$||\cos(g^k\pi) - 1| \leq \pi\theta^k.$$

D'où

$$\lim_{n \rightarrow \infty} |\Phi_{t(n)}(Y_\infty)| = |C| \prod_{k=0}^{\infty} |\cos(g^k\pi)| > 0.$$

1.4.3 Le problème d'Erdős

Soit $CS \subset]\frac{1}{2}, 1[$ l'ensemble des valeurs de a pour lesquelles ν_a est singulière. Le théorème d'Erdős pose le problème de la description de CS . En 1995, Solomyak [55] (voir aussi la très belle preuve de Solomyak et Peres [53]) a démontré que CS est de mesure de Lebesgue nulle, résolvant ainsi une conjecture datant des années 60. Mais, à ce jour, on ne sait toujours pas si CS contient des points qui ne sont pas des inverses de nombres de Pisot. Le lecteur intéressé par ces questions pourra consulter l'article de Peres, Schlag et Solomyak [52].

◇

1.5 Simulation de fractales

Une généralisation multidimensionnelle de la dynamique de la section 1.4 est la récurrence aléatoire définie sur \mathbb{R}^d par

$$X_{n+1} = A(\theta_{n+1})X_{n+1} + B(\theta_{n+1})$$

où les variables (θ_n) sont i.i.d. à valeurs dans $\{1, \dots, k\}$, $\{A(1), \dots, A(k)\}$ sont k matrices $d \times d$ à coefficients réels et $\{B(1), \dots, B(k)\}$ k vecteurs de \mathbb{R}^d .

Si A est une matrice $d \times d$ et $\|\cdot\|$ une norme sur \mathbb{R}^d , rappelons que la norme de A (notée $\|A\|$) est définie par

$$\|A\| = \sup\{\|Ax\| : x \in \mathbb{R}^m, \|x\| = 1\}.$$

Supposons $\|A(i)\| < 1$ pour tout $i = 1, \dots, k$. Alors la suite (Y_n) de terme général

$$Y_n = B(\theta_1) + A(\theta_1)B(\theta_2) + \dots + \left[\prod_{j=1}^n A(\theta_j) \right] B(\theta_{n+1})$$

converge presque sûrement vers une variable aléatoire $Y_\infty \in \mathbb{R}^d$.

Proposition 1.5.1 *Supposons que $\|A(i)\| < 1$ pour tout $i = 1, \dots, k$. Alors*

- (i) *la suite (X_n) converge en loi vers Y_∞ ;*
- (ii) *la loi π de Y_∞ est l'unique probabilité sur \mathbb{R}^d solution de l'équation*

$$\pi = \sum_{i=1}^k m(i)\pi \circ f_i^{-1}$$

où f_i désigne l'application affine $x \rightarrow A(i)x + B(i)$ et $m(i) = P(\theta_n = i)$.

Exercice 1.5.2 Démontrer la proposition précédente (on pourra s'inspirer des preuves des propositions 1.4.2 et 1.4.4).

Tout comme dans le cas des convolutions de Bernoulli, ce résultat de convergence en loi s'accompagne d'un théorème de convergence presque sûre.

Proposition 1.5.3 *Sous les hypothèses, et avec les notations, de la proposition 1.5.1,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \Phi(X_k) = \int_{\mathbb{R}^m} \Phi(x)\pi(dx)$$

presque sûrement, pour toute fonction $\Phi : \mathbb{R}^m \rightarrow \mathbb{R}$ continue et bornée.

Ce théorème ergodique (dont la preuve sera donnée dans le chapitre 4, corollaire 4.4.12) permet, pour $d = 1, 2$ ou 3 , de “visualiser” le support π à partir de la représentation graphique d'une trajectoire X_0, X_1, \dots particulière. Dans une image en noir

et blanc, les zones blanches (non visitées) n'appartiennent pas au support de π et les régions visitées sont d'autant plus foncées qu'elles sont chargées par π . Voici quelques images obtenues en dimension 2. La structure fractale est liée à la propriété d'autosimilarité de π (proposition 1.5.1 (ii)). Pour d'autres images et des développements on pourra consulter Barnsley [13].

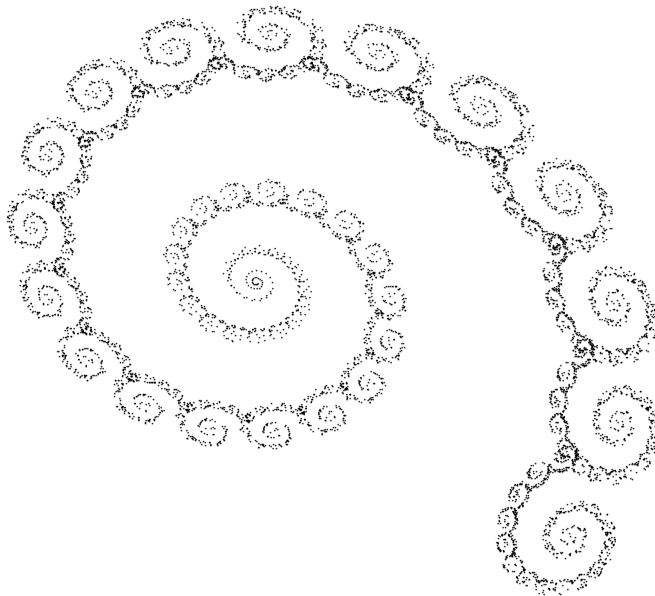
Une spirale

$$A(1) = \begin{bmatrix} 0.839 & -0.303 \\ 0.383 & 0.924 \end{bmatrix}, A(2) = \begin{bmatrix} -0.161 & -0.136 \\ 0.138 & -0.182 \end{bmatrix},$$

$$B(1) = \begin{bmatrix} 0.232 \\ -0.080 \end{bmatrix}, B(2) = \begin{bmatrix} 0.921 \\ 0.178 \end{bmatrix}$$

et $m(1) = m(2) = \frac{1}{2}$.

Spirale



Une Feuille

$$A(1) = \begin{bmatrix} 0 & 0 \\ 0 & 0.16 \end{bmatrix}, A(2) = \begin{bmatrix} 0.2 & -0.26 \\ 0.23 & 0.22 \end{bmatrix}, A(3) = \begin{bmatrix} -0.15 & 0.28 \\ 0.26 & 0.24 \end{bmatrix}, A(4) = \begin{bmatrix} 0.85 & 0.04 \\ -0.04 & 0.85 \end{bmatrix}$$

$$B(1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, B(2) = \begin{bmatrix} 0 \\ 1.6 \end{bmatrix}, B(3) = \begin{bmatrix} 0 \\ 0.44 \end{bmatrix}, B(4) = \begin{bmatrix} 0 \\ 1.6 \end{bmatrix}$$

et $m(1) = 0.01, m(2) = 0.07, m(3) = 0.07, m(4) = 0.85$

Feuille

