

Le phénomène de raréfaction dans la suite de Thue-Morse.

Alexandre Aksenov
sous la direction de Jean-Louis Verger-Gaugry

25 juin 2009

1.Introduction.

La suite de Thue-Morse est la suite

$$1 \bar{1} \bar{1}1 \bar{1}11\bar{1} \bar{1}11\bar{1}1\bar{1}\bar{1}1 \dots$$

définie par :

$$\tau(n) = (-1)^{\text{le nombre de chiffres un dans l'écriture binaire de } n}$$

(le symbole $\bar{1}$ signifie -1).

Les principaux objets de mon étude sont les sommes q -raréfiées de cette suite, c'est à dire les sommes du type

$$S_{q,i}(N) = \sum_{\substack{0 \leq n < N \\ n \equiv i \pmod{q}}} \tau(n) \quad (i = 0, 1 \dots q-1). \quad (1)$$

Voici un exemple d'énoncé :

Proposition 1 (Conjecture de Moser, Théorème de Newman ([14])). *Pour tout $N > 0$, on a : $S_{3,0}(N) > 0$.*

Pour le montrer, on sera amené à passer de la base 2 à la base 4 car dans cette base le nombre 3 a les mêmes propriétés que le chiffre 9 en base 10. Plus généralement, pour un nombre $q > 3$, les sommes q -raréfiées peuvent être étudiées en passant à la base 2^s où s est l'ordre de 2 dans le groupe $(\mathbb{Z}/q\mathbb{Z})^\times$.

Par exemple, dans le cas des nombres premiers on a le résultat suivant :

Théorème 1. *Soit p un nombre premier impair et $n \in \mathbb{N}$. On a alors :*

$$\mathbf{S}_p(n) = n^{\alpha_p} \tilde{\mathbf{G}} \left(\frac{\log n}{rs \log 2} \right) + \tilde{\mathbf{E}}(n) \quad (2)$$

où $\mathbf{S}_p(n) = (S_{p,i}(n))_{i=0,1,\dots,p-1}$;

$\tilde{\mathbf{G}}(x) = (\tilde{G}_i(x))_{i=0,1,\dots,p-1}$ est un vecteur colonne à p lignes dont chaque composante est une fonction continue nulle part dérivable de période 1 ;

$$\alpha_p = \frac{\log |\xi_m|}{rs \log 2}, \quad (3)$$

ξ_m étant une valeur propre de \mathbf{M} de module maximal et $r \in \{1, 2, 4\}$ le plus petit entier tel que $\xi_m^r \in \mathbb{R}_+^*$; le terme d'erreur $\tilde{\mathbf{E}}(n) = O(n^{\beta_p})$ avec

$$\beta_p = \beta_p(\lambda_2) = \begin{cases} \frac{\log \lambda_2}{s \log 2} & \text{si } \lambda_2 > 1, \\ 0 & \text{sinon,} \end{cases}$$

où λ_2 est le deuxième plus grand module de valeur propre de \mathbf{M} . Ici s désigne l'ordre de 2 dans le groupe $(\mathbb{Z}/p\mathbb{Z})^\times$ et \mathbf{M} est la matrice $p \times p$ suivante : $\mathbf{M} = (S_{p,i-j}(2^s))_{i,j}$.

La forme (3) de l'exposant fractal indique une subdivision de l'ensemble des nombres premiers impairs en classes : notons $\mathbb{P}_1 = \{3, 5, 11, 13, 19, 29, 37, 53, 59, 61, 83, 101, \dots\}$ l'ensemble des nombres premiers pour lesquels $s = p - 1$; pour ces nombres on a

$$\alpha_p = \frac{\log p}{(p-1) \log 2}.$$

Les classes suivantes sont alors : l'ensemble des nombres premiers pour lesquels $s = \frac{p-1}{2}$ et congrus à 1 modulo 4, noté $\mathbb{P}_{2,1} = \{17, 41, 97, 137, 193, \dots\}$, l'ensemble des nombres premiers pour lesquels $s = \frac{p-1}{2}$ et congrus à 3 modulo 4, noté $\mathbb{P}_{2,3} = \{7, 23, 47, 71, 79, 103, 167, \dots\}$ etc.

Cette subdivision rejoint les études menées sur la Conjecture d'Artin.

Conjecture (Artin). *Le sous-ensemble des nombres premiers $p \in \mathbb{P}$ tels que le groupe monogène $\langle 2 \rangle$ dans $(\mathbb{Z}/p\mathbb{Z})^\times$ soit d'ordre $p - 1$ est infini.*

D'après l'heuristique démontrée dans [11] sous l'hypothèse de Riemann généralisée, la fréquence de tels nombres premiers p vaut $\prod_q (1 - \frac{1}{q(q-1)}) \approx 0,37$ (l'indice q parcourant l'ensemble des nombres premiers). L'article [12] donne une discussion critique et approfondie de cette heuristique basée sur les résultats expérimentaux.

De même, on ne sait pas si les autres classes introduites plus haut sont finies ou infinies.

2. Le cas d'un nombre premier.

2.1. Le cas $p = 3$: la première méthode.

Cette partie et la suivante seront consacrées à l'étude de la 3-raréfaction. Nous développerons pour cela deux méthodes : d'abord celle de l'article [14], la plus directe et aussi celle qui a été la première sur le plan historique, puis celle de l'article [2] qui a l'avantage de s'adapter aux autres nombres que 3.

Introduisons¹, à côté des sommes $S_{3,i}(N)$, la notation suivante :

$$\tilde{S}_{3,i}(N) = \sum_{\substack{0 \leq n \leq N \\ n \equiv i \pmod{3}}} \tau(n) = S_{3,i}(N+1).$$

Calculons les sommes $S_{3,i}(N)$ où N est une puissance de deux. On obtient le théorème suivant :

¹C'est fait dans le souci d'uniformiser les notations des différents auteurs ; la fonction que D.J. Newman appelle $S(N)$ correspond dans nos notations à $\tilde{S}_{3,N}(N)$.

Théorème 2. Pour tout $k \in \mathbb{N}$, on a :

$$\begin{aligned} S_{3,0}(4^{k+1}) &= 2 \cdot 3^k, \\ S_{3,1}(4^{k+1}) &= -3^k, \\ S_{3,2}(4^{k+1}) &= -3^k, \\ S_{3,0}(2 \cdot 4^k) &= 3^k, \\ S_{3,1}(2 \cdot 4^k) &= -3^k, \\ S_{3,2}(2 \cdot 4^k) &= 0. \end{aligned}$$

Démonstration. Procédons par récurrence sur k . Le cas $k = 0$ se vérifie directement, et on a en général les relations suivantes :

$S_{3,0}(4^{k+1}) = 2S_{3,0}(2 \cdot 4^k)$ car les deux parties de la somme

$$S_{3,0}(4^{k+1}) = \sum_{\substack{0 \leq n < 4^{k+1} \\ n \equiv 0 \pmod{3}}} \tau(n) = \sum_{\substack{0 \leq n < 2 \cdot 4^k \\ n \equiv 0 \pmod{3}}} \tau(n) + \sum_{\substack{2 \cdot 4^k < n < 4^{k+1} \\ n \equiv 0 \pmod{3}}} \tau(n)$$

sont égales ; en effet, pour chaque $n < 4^k$, le nombre $4^{k+1} - 1 - n$ (multiple de 3 si n est un multiple de 3) contient des 1 à la place des 0 de l'écriture binaire de n et vice versa, d'où $\tau(n) = \tau(4^{k+1} - 1 - n)$;

$S_{3,1}(4^k + 1) = S_{3,2}(4^{k+1})$; en effet, dans ce cas le "complémentaire" $4^{k+1} - 1 - n$ d'un nombre n est congru à 1 modulo 3 si $n \equiv 2 \pmod{3}$ et vice versa ;

pour tout nombre pair N , on a :

$$S_{3,0}(N) + S_{3,1}(N) + S_{3,2}(N) = \sum_{0 \leq n < N} \tau(n) = 0 ;$$

$S_{3,2}(2 \cdot 4^k) = 0$ car, en définissant le complémentaire de $n < 2 \cdot 4^k$ comme étant le nombre $2 \cdot 4^k - 1 - n$, on obtient $\tau(2 \cdot 4^k - 1 - n) = -\tau(n)$ et $2 \cdot 4^k - 1 - n \equiv 2 \pmod{3}$ si $n \equiv 2 \pmod{3}$; par conséquent,

$$S_{3,2}(2 \cdot 4^k) = \sum_{\substack{0 \leq n < 2 \cdot 4^k \\ n \equiv 2 \pmod{3}}} \tau(n)$$

(qui est une somme de longueur paire) contient autant de termes +1 que de termes -1.

Ces relations montrent que les rapports des sommes $S_{3,i}(2 \cdot 4^k)$ et $S_{3,i}(4 \cdot 4^k)$ entre elles sont ceux prédits dans l'énoncé. Pour faire le pas de récurrence, montrons que pour tout $k \in \mathbb{N}$ on a : $S_{3,0}(4^{k+1}) = 3S_{3,0}(4^k)$. En découpant de la première somme en quatre parties, on obtient :

$$\begin{aligned} S_{3,0}(4^{k+1}) &= \sum_{n < 4^k} \tau(n) + \sum_{4^k \leq n < 2 \cdot 4^{k+1}} \mathbf{1}_{n \equiv 0 \pmod{3}} \tau(n) \\ &+ \sum_{2 \cdot 4^{k+1} \leq n < 3 \cdot 4^{k+1}} \mathbf{1}_{n \equiv 0 \pmod{3}} \tau(n) + \sum_{3 \cdot 4^{k+1} \leq n < 4^{k+1}} \mathbf{1}_{n \equiv 0 \pmod{3}} \tau(n) \\ &= 2S_{3,0}(4^k) - S_{3,2}(4^k) - S_{3,1}(4^k) = 3S_{3,0}(4^k). \end{aligned}$$

□

Soit maintenant un entier N quelconque d'écriture binaire

$$N = \sum_{k=0}^l c_k 2^k$$

où $l = \lfloor \log_2 N \rfloor$. Alors, $\tilde{S}_{3,i}(N)$ peut être exprimé de façon suivante à l'aide de ces formules :

$$\tilde{S}_{3,i}(N) = \sum_{\substack{0 \leq n < 2^l \\ n \equiv i \pmod{3}}} \tau(n) - \sum_{\substack{0 \leq n \leq N-2^l \\ n' \equiv i-2^l \pmod{3}}} \tau(n) = S_{3,i}(2^l) - \tilde{S}_{3,i-2^l}(N-2^l) \quad (4)$$

où le premier terme du membre de droite est calculé à partir du Théorème 1, et l'argument de la fonction \tilde{S} dans le deuxième terme a un chiffre un de moins que N . Si on reutilise la formule (4) jusqu'à arriver à un cas trivial, on obtient la formule suivante :

$$\tilde{S}_{3,i}(N) = S_{3,\rho_1}(2^{k_1}) - S_{3,\rho_2}(2^{k_2}) + S_{3,\rho_3}(2^{k_3}) + \dots + (-1)^w S(2^{k_w}) + (-1)^w \quad (5)$$

où $l = k_1 > k_2 > \dots > k_w \geq 0$ sont les numéros des positions telles que $c_{k_v} = 1$, w est le nombre de chiffres un dans l'écriture binaire de N et $\rho_1, \rho_2, \dots, \rho_w$ sont les restes modulo 3 définis par la relation de récurrence

$$\begin{cases} \rho_1 \equiv i \\ \rho_{v+1} \equiv \rho_v - 2^{k_v}. \end{cases}$$

Maintenant, on peut borner tous les termes de cette somme pour obtenir :

$$\begin{aligned} \tilde{S}_{3,0}(N) &\geq 3^{\frac{l-1}{2}} - \frac{2}{3} 3^{\frac{k_3}{2}} - \frac{2}{3} 3^{\frac{k_4}{2}} - \dots - \frac{2}{3} 3^{\frac{k_w}{2}} - \frac{1}{3} \\ &\geq 3^{\frac{l-1}{2}} - \sum_{a=2}^{+\infty} \frac{2}{3} 3^{\frac{l-a}{2}} = 3^{\frac{l-1}{2}} \left(1 - \frac{2}{3\sqrt{3}-3} \right) > 0 \end{aligned}$$

ce qui montre la Conjecture de Moser et la minoration $\tilde{S}_{3,0}(3N)/N^{\alpha_3} \geq 3^{\alpha_3-1} \left(1 - \frac{2}{3\sqrt{3}-3} \right)$ où $\alpha_3 = \log_4 3$.

L'article [14] obtient aussi par un argument similaire la majoration $S_{3,0}(3N) < \frac{2}{3-\sqrt{3}} 3^{\frac{l}{2}}$, d'où

$$S_{3,0}(3N)/N^{\alpha_3} < \frac{2 \cdot 3^{\alpha_3}}{3 - \sqrt{3}}, \quad (6)$$

mais cette majoration peut être améliorée (cf le Théorème 5) . Dans la Section 5.1 nous reviendrons sur cette majoration dans le but de l'améliorer et d'étudier les deux autres restes modulo 3.

2.2. Le cas $p = 3$: la deuxième méthode.

Dans cette partie on montrera la Conjecture de Moser en reprenant l'étude des sommes $S_{3,0}(N)$ entreprise dans l'article [2]. Pour chaque $n \in \mathbb{N}$ on a :

$$3 \cdot \mathbf{1}_{3|n} = 1 + j^n + j^{2n}$$

où $j = e^{2i\pi/3}$ ce qui permet de mettre la somme (1) sous la forme suivante :

$$S_{3,0}(N) = \frac{1}{3} \left(\sum_{n=0}^{N-1} \tau(n)(1 + j^n + j^{2n}) \right) = \frac{\eta(N)}{3} + \frac{2}{3} \Re \sum_{n=0}^{N-1} \tau(n) j^n \quad (7)$$

avec $\eta(N) = \sum_{n=0}^{N-1} \tau(n) \in \{-1, 0, 1\}$. Par la suite, nous considérerons $\eta(N)$ comme un “terme d’erreur” et nous nous intéresserons aux sommes $\psi(N) = \sum_{n=0}^{N-1} \tau(n)j^n$.

La suite $\tau(n)j^n$ a la propriété de 4-multiplicativité, c’est à dire elle se décompose en produit

$$\tau(n)j^n = \prod_{k=0}^L \tau(d_k)j^{d_k}$$

où les d_k sont les chiffres de n en base 4 et $L = \lceil \log_4 n \rceil$. Cette propriété permet d’obtenir pour tout $k \in \mathbb{N}$:

$$\begin{aligned} \psi(4^k) &= 3^k, \\ \psi(2 \cdot 4^k) &= (1-j)3^k \text{ et} \\ \psi(3 \cdot 4^k) &= 2 \cdot 3^k. \end{aligned}$$

Notons, d’après l’article [2],

$$c(0) = 0, \quad c(1) = 1, \quad c(2) = 1-j, \quad c(3) = 2, \quad c(4) = 3.$$

On a alors l’expression compacte suivante pour $\psi(N)$. Soit

$$N = \sum_{k=0}^L d_k 4^k$$

l’écriture de N en base 4. On a alors :

$$\psi(N) = \sum_{k=0}^L \left(\prod_{k'=k+1}^L \tau(d_{k'})j^{d_{k'}} \right) c(d_k)3^k.$$

Cette série permet de définir $\psi(x)$ pour n’importe quel $x \in \mathbb{R}_+$. Montrons que sa somme ne dépend pas du choix d’écriture si x admet deux écritures en base 4.

Supposons que le nombre x s’écrit

$$x = \sum_{k=k_1+1}^L d_k 4^k + (e+1)4^{k_1}$$

où $e \in \{0, 1, 2\}$ et $k_1 \in \mathbb{Z}$. La valeur obtenue à partir de cette écriture est

$$\psi_1(x) = \sum_{k=k_1+1}^L \left(\prod_{k'=k+1}^L \tau(d_{k'})j^{d_{k'}} \right) c(d_k)3^k + \left(\prod_{k'=k_1+1}^L \tau(d_{k'})j^{d_{k'}} \right) c(e+1)3^{k_1}.$$

L’autre écriture de x est alors

$$x = \sum_{k=k_1+1}^L d_k 4^k + e \cdot 4^{k_1} + \sum_{a=1}^{+\infty} 3 \cdot 4^{k_1-a}$$

d’où la valeur de la série correspondante est

$$\psi_2(x) = \sum_{k=k_1+1}^L \left(\prod_{k'=k+1}^L \tau(d_{k'})j^{d_{k'}} \right) c(d_k)3^k + \left(\prod_{k'=k_1+1}^L \tau(d_{k'})j^{d_{k'}} \right) \left(c(e)3^{k_1} + \sum_{a=1}^{+\infty} \tau(e)j^e \cdot c(3)3^{k_1-a} \right).$$

La dernière formule utilise le fait que $\tau(3)j^3 = 1$. Les termes qui correspondent aux chiffres qui précèdent k_1 sont identiques dans les deux formules, et la preuve de l'égalité $\psi_1(x) = \psi_2(x)$ se ramène donc à l'égalité suivante :

$$c(e+1)3^{k_1} = c(e)3^{k_1} + \sum_{a=1}^{+\infty} \tau(e)j^e \cdot c(3)3^{k_1-a},$$

qui, à son tour, devient claire quand on remplace $c(3)$ par 2, la somme des 3^{-a} par $\frac{1}{2}$, et $c(e+1)$ par $c(e) + \tau(e)j^e$.

La proposition suivante concerne les propriétés analytiques de la fonction ainsi définie.

Théorème 3. *La fonction $\psi(t)$ est continue et dérivable nulle part.*

Démonstration. Montrons d'abord que la fonction $\psi : \mathbb{R}_+ \rightarrow \mathbb{C}$ est continue. Soit une suite de réels $x_1, x_2, \dots, x_n, \dots$ qui tend vers x , et montrons que la suite $\psi(x_1), \psi(x_2), \dots, \psi(x_n), \dots$ tend vers $\psi(x)$. Si x n'admet pas d'écriture finie en base 4, alors pour chaque k_1 , il existe un rang n tel qu'à partir du rang n tous les chiffres des x_i jusqu'au k_1 -ième après la virgule sont les mêmes que ceux de x , d'où

$$|\psi(x_i) - \psi(x)| \leq \sum_{a=1}^{+\infty} 4 \cdot 3^{-k_1-a} = 2 \cdot 3^{-k_1}$$

ce qui devient arbitrairement petit quand k_1 tend vers l'infini, d'où la convergence. Si le nombre x admet une écriture de longueur finie et une écriture de longueur infinie, le même raisonnement s'applique en comparant l'écriture des x_i soit à la première écriture de x si $x_i > x$ soit à la seconde si $x_i < x$.

Par contre, elle n'est dérivable en aucun point. Soit, en effet, un nombre $x \in \mathbb{R}_+$ et considérons son écriture

$$x = \sum_{k=-L}^{+\infty} d_k 4^{-k}$$

qui ne se termine pas par une infinité de chiffres 3. Soit k_1, k_2, \dots une suite strictement croissante d'indices tels que $d_{k_i} \in \{0, 1, 2\}$. On a alors :

$$\frac{\psi(\sum_{k=-k_i}^L d_k 4^k) - \psi(\sum_{k=-k_i}^L d_k 4^k + 4^{-k_i})}{4^{-k_i}} = \frac{3^{-k_i}}{4^{-k_i}} \left(\prod_{k'=k_i+1}^L \tau(d_{k'}) j^{d_{k'}} \right) (c(d_{k_i} + 1) - c(d_{k_i})),$$

donc

$$\left| \frac{\psi(\sum_{k=-k_i}^L d_k 4^k) - \psi(\sum_{k=-k_i}^L d_k 4^k + 4^{-k_i})}{4^{-k_i}} \right| = \left(\frac{4}{3} \right)^{k_i}.$$

La dérivabilité de la fonction ψ est donc impossible. □

En termes de représentation graphique, la fonction $\psi(x)$ se construit à partir du bloc qu'on voit sur la Figure 1 en itérant l'opération de remplacement de chaque segment par une copie du bloc tout entier. On obtient donc à la limite la ligne appelée le flocon de Koch (cf [13]) représentée sur la Figure 2.

Mettons maintenant l'équation (7) sous la forme

$$S_{3,0}(3N) = \frac{\eta(3N)}{3} + N^\alpha F(\log_4 N) \tag{8}$$

où $\alpha = \alpha_3 = \log_4 3$ et F est la fonction continue 1-périodique définie par $F(\log_4 x) = 2\Re\psi(x)/(x^\alpha)$. Avec les notations du théorème 1 on a : $\tilde{G}_0(x) = F(x)$. L'article [14] donne l'encadrement suivant pour F :

$$0,0711 \approx 3^{\alpha-1} \left(1 - \frac{2}{3\sqrt{3}-3} \right) < F(t) < \frac{2 \cdot 3^\alpha}{3 - \sqrt{3}} \approx 3,77 \tag{9}$$

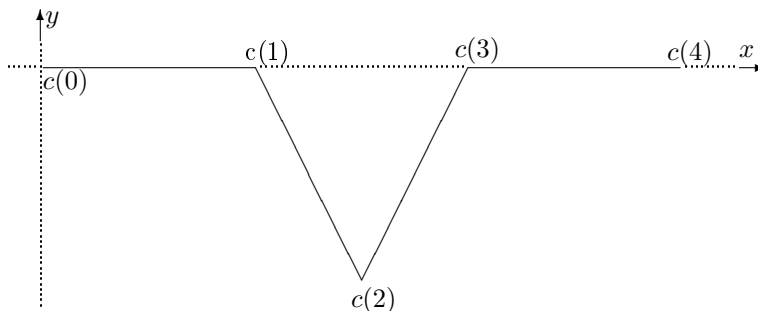


FIG. 1 – Le bloc du flocon de Koch.

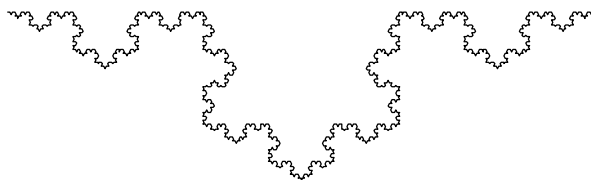


FIG. 2 – Le flocon de Koch.

et l'article [2] l'améliore pour obtenir

$$1,1547 \approx \frac{2\sqrt{3}}{3} = \inf F < \sup F = \frac{55}{3} \left(\frac{3}{65} \right)^\alpha \approx 1,6020. \quad (10)$$

Terminons cette partie avec des commentaires sur les représentations graphiques de la fonction F et des fonctions correspondantes obtenues à partir de $S_{3,1}(N)$ et $S_{3,2}(N)$ (les Figures 3,4,5). À chaque fois, il s'agit d'un nuage de points qui fait apparaître la courbe en tant qu'objet limite quand N devient grand. La présence des points en-dehors de la courbe est une conséquence du terme d'erreur.

La première courbe a été étudiée dans [2] avec plus de détails. Son minimum se trouve au point d'abscisse 0,5 et on voit deux maxima locaux (de hauteurs proches l'une de l'autre) aux points d'abscisse $\{\log_4 \frac{17}{3}\}$ et $\{\log_4 \frac{65}{3}\}$ comme prédit. Les points "erronés" sont au-dessous de la courbe dans leur majorité, ce qui correspond au fait que le terme d'erreur vaut $\tau(3N - 1)$ (quand il n'est pas nul, c'est à dire quand N est impair), et la plupart de ces nombres sont négatifs au début.

En ce qui concerne les deux autres courbes, on peut obtenir les résultats essentiels sur elles à partir d'un cadre plus général, qui est l'objet de la partie suivante.

2.3. Le cas d'un nombre premier $p > 3$.

La construction de la fonction fractale décrite dans la partie précédente était essentiellement automatique à partir du moment où on avait $\tau(0)j^0 + \tau(1)j^1 + \tau(2)j^2 + \tau(3)j^3 = 3 > 1$. Dans cette partie nous allons travailler de façon similaire avec un nombre p premier (impair) quelconque en suivant l'article [8]. Rappelons que s désigne le plus petit entier naturel non nul tel que $2^s \equiv 1 \pmod{p}$.

La somme citée plus haut est remplacée par les sommes suivantes :

$$\sum_{n=0}^{2^s-1} \tau(n) \zeta^{an} = \sum_{j \in \mathbb{F}_p} S_{p,j}(2^s) \zeta^{aj} = \xi_a \quad (11)$$

avec $a \in \mathbb{F}_p^\times$ et $\zeta = e^{\frac{2i\pi}{p}}$. Ces nombres sont des valeurs propres de la matrice

$$\mathbf{M} = \sum_{j \in \mathbb{F}_p} S_{p,j}(2^s) \mathbf{T}^j \quad (12)$$

où \mathbf{T} est la matrice $p \times p$ suivante :

$$\mathbf{T} = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & & & 0 \\ & 1 & 0 & & 0 \\ & (0) & \backslash & \ddots & \vdots \\ & & & & 1 & 0 \end{pmatrix}.$$

qui vérifie $\mathbf{T}^j = 1$. En effet, les vecteurs propres de \mathbf{T} sont les vecteurs $(\zeta^{-aj})_{j=0 \dots p-1}$ (pour $a \in \mathbb{F}_p$) avec la valeur propre ζ^a . Les vecteurs propres de la matrice \mathbf{M} sont les mêmes, de valeur propre donnée par (11).

Une manière plus pratique d'étudier ces valeurs propres est fournie par la factorisation de l'expression (12). On a en effet :

$$\mathbf{M} = \prod_{m=0}^{s-1} (1 - \mathbf{T}^{2^m}) \quad (13)$$

d'après la définition de la suite de Thue-Morse. Cela implique que la valeur propre ξ_a ne change pas si on remplace a par $2^m a$. La matrice \mathbf{M} a donc une valeur propre simple $\xi_0 = 0$ et une valeur propre non nulle par classe de conjugaison modulo le groupe homogène engendré par 2. Le produit de ces dernières vaut :

$$\prod_{a \in \mathbb{F}_p^\times / \langle 2 \rangle} \xi_a = \prod_{j \in \mathbb{F}_p^\times} (1 - \zeta^j) = p.$$

En particulier, si $s = p - 1$, la seule valeur propre non nulle vaut p et on a : $\mathbf{M} = p - 1 - \mathbf{T} - \mathbf{T}^2 - \dots - \mathbf{T}^{p-1}$ ce qui signifie que la p -raréfaction peut être étudiée exactement comme avant. À titre d'illustration, donnons le "graphique" (Figure 17) de la courbe fractale créée pour $p = 5$ de la même façon que ce qu'on avait pour $p = 3$.

Si ce n'est pas le cas, on sait tout de même que les valeurs propres sont toutes réelles ou toutes imaginaires pures. En effet, on a :

$$\begin{aligned} \xi_a &= \prod_{j' \in \langle 2 \rangle a} (1 - \zeta^{2j'}) = \left(\prod_{j' \in \langle 2 \rangle a} \zeta^{j'} \right) \left(\prod_{j' \in \langle 2 \rangle a} (\zeta^{-j'} - \zeta^{j'}) \right) \\ &= \zeta^{(\sum_{j' \in \langle 2 \rangle a} j')} \prod_{j' \in \langle 2 \rangle a} \left(-2i \sin \frac{2\pi j'}{p} \right) \\ &= (-2i)^s \prod_{j' \in a \langle 2 \rangle} \sin \frac{2\pi j'}{p}. \end{aligned} \quad (14)$$

Donnons maintenant l'idée de démonstration du théorème 1. Étudions le vecteur $\mathbf{S}_p(N) = (S_{p,i}(N))_{i=0,\dots,p-1}$ composante par composante dans la base des $(\zeta^{-aj})_j$. On a la relation de récurrence $\mathbf{S}_p(2^s N) = \mathbf{M}\mathbf{S}_p(N)$ valable pour tout $N \in \mathbb{N}$. Notons \mathbf{P}_u la projection de \mathbb{C}^n sur la somme directe V_u des sous-espaces propres qui correspondent aux valeurs propres de module plus grand que 1 (notons qu'aucune valeur propre ne peut être de module 1 car toutes les valeurs propres sont réelles ou imaginaires pures, et d'autre part, les valeurs $\{1, i, -1, -i\}$ sont exclues par la formule du produit de toutes les valeurs propres). Alors, la fonction $\mathbf{P}_u \mathbf{S}_p(N)$ peut être prolongée en une fonction continue

$$\mathbf{F} : \mathbb{R}_+ \rightarrow V_u$$

par un procédé qui généralise la construction de la fonction ψ dans la partie précédente. Le terme d'erreur $(\mathbf{I} - \mathbf{P}_u)\mathbf{S}_p(N)$ est borné en général.

Étudions maintenant ces valeurs propres dans le cas où $s = \frac{p-1}{2}$. Deux cas se distinguent ici : le cas où $p \equiv 1 \pmod{4}$ et celui où $p \equiv 3 \pmod{4}$.

Si $p \equiv 3 \pmod{4}$, les deux valeurs propres sont imaginaires pures conjuguées et leur produit vaut p . Ce sont donc $i\sqrt{p}$ et $-i\sqrt{p}$. Or, $i\sqrt{p}$ est aussi la valeur de la somme de Gauss quadratique $\sum_{1 \leq j < p/2} \left(\frac{j}{p}\right)\zeta^j$, ce qui indique, au signe près, la forme des coefficients de la matrice \mathbf{M} : ses coefficients diagonaux sont nuls et les autres sont égaux à 1 ou -1 suivant s'il s'agit d'un résidu quadratique modulo p . Le signe peut être précisé à partir de la formule 14 de façon suivante : $\xi_1 = +i\sqrt{p}$ si et seulement si un nombre pair d'éléments de $\langle 2 \rangle$ (qui sont aussi les résidus quadratiques) se trouve dans l'intervalle $[\frac{p}{2}, p]$. D'après le Théorème 4 du §4.2 de [1], on a aussi :

$$\xi_1 = (-1)^{\frac{h+1}{2}} i\sqrt{p}$$

où h est le nombre de classes d'idéaux du corps $\mathbb{Q}(-\sqrt{p})$.

Passons au cas où $p \equiv 1 \pmod{4}$, d'où, d'après la loi de réciprocité quadratique, $p \equiv 1 \pmod{8}$. Les premiers nombres premiers de ce type sont 17, 41, 97, 137, 193, ... D'après la formule (14), les deux valeurs propres de \mathbf{M} sont alors positives.

Introduisons les notations suivantes : $\langle 2 \rangle' = \mathbb{F}_p^\times \setminus \mathbb{F}_p^{\times 2}$;

$$S_{p,j}(2^s) = \begin{cases} C & \text{si } j = 0 \\ A & \text{si } j \in \langle 2 \rangle \\ B & \text{si } j \in \langle 2 \rangle' \end{cases}$$

$$\xi = \xi_1 = \sum_{j \in \mathbb{F}_p} S_{p,j}(2^s) \zeta^j = C + A \sum_{j \in \langle 2 \rangle} \zeta^j + B \sum_{j \in \langle 2 \rangle'} \zeta^j ;$$

$$\xi' = \xi_{j'} = \sum_{j \in \mathbb{F}_p} S_{p,j}(2^s) \zeta^{j'j} = C + B \sum_{j \in \langle 2 \rangle} \zeta^j + A \sum_{j \in \langle 2 \rangle'} \zeta^j$$

où j' est un élément de $\langle 2 \rangle'$.

On a alors : $\xi\xi' = p$ et $\xi + \xi' = 2C - A - B$ d'où ξ et ξ' sont des éléments conjugués d'une extension quadratique de \mathbb{Q} . Comme

$$\xi - \xi' = (A - B) \left(\sum_{j \in \langle 2 \rangle} \zeta^j - \sum_{j \in \langle 2 \rangle'} \zeta^j \right) = (A - B)\sqrt{p}$$

et $A \neq B$ (car sinon, on aurait $\xi = \xi'$ et ces deux nombres seraient entiers et égaux à une racine carrée de p), cette extension est $\mathbb{Q}(\sqrt{p})$.

On peut donc écrire

$$\begin{aligned}\xi &= a\sqrt{p} + bp & \text{et} \\ \xi' &= -a\sqrt{p} + bp\end{aligned}$$

où a et b sont des rationnels reliés aux entiers A, B et C par les formules suivantes :

$$\begin{cases} C = b(p-1), \\ A = a-b, \\ B = -a-b. \end{cases}$$

Par conséquent, $2a$ et $2b$ sont des entiers, et même a et b sont des entiers car A et B ont la même parité en tant que sommes de même longueur $(\frac{2^s-1}{p})$ de termes de la suite de Thue-Morse.

Dans cette situation les valeurs propres ξ et ξ' s'expriment à l'aide du nombre de classes h du corps $\mathbb{Q}(\sqrt{p})$ et de l'unité fondamentale ϵ de ce corps. En effet, d'après le Théorème 2 du §4.2 de [1], on a :

$$\begin{aligned}\xi &= \epsilon^{-h}\sqrt{p} & \text{et} \\ \xi' &= \epsilon^h\sqrt{p} & .\end{aligned}\tag{15}$$

Remarquons que l'étude des valeurs propres peut être menée indépendamment de l'hypothèse $s = \frac{p-1}{2}$ si on remplace la matrice \mathbf{M} par le produit qui correspond à (13) avec le sous-groupe des carrés au lieu du sous-groupe engendré par 2.

2.4. Retour à la 3-raréfaction.

Précisons les résultats obtenus pour $q = 3$. Nous allons reprendre ici les méthodes et les notations de la partie 2.1 et affiner la majoration pour $S_{3,0}(N)$, et traiter les deux autres résidus. Le résultat le plus important sera l'inégalité $S_{3,2}(N) \leq 0$. Précisons que tous les résultats de cette partie sont originaux.

Théorème 4. *Soit l un entier naturel, $N \leq 2^{2l} - 1$ et i un reste modulo 3. Alors,*

$$\tilde{S}_{3,0}(2^{2l} - 1) \geq \tilde{S}_{3,i}(N).$$

Démonstration. Par récurrence sur l . L'énoncé est évident pour $l = 0$ et $l = 1$, et pour l plus grand on a :

$$\tilde{S}_{3,0}(2^{2l+2} - 1) = S_{3,0}(2^{2l+1}) - S_{3,1}(2^{2l}) + \tilde{S}_{3,0}(2^{2l} - 1) = 3^l + 3^{l-1} + \tilde{S}_{3,0}(2^{2l} - 1)$$

alors que tout autre couple de valeurs pour N et i donnerait :

$$\tilde{S}_{3,i}(N) = \mathbf{1}_{c_{2l+1}=1} S_{3,i}(2^{2l+1}) + \varepsilon_1 S_{3,i_2}(2^{2l+1}) + \varepsilon_2 \tilde{S}_{3,i_3}(N_2)$$

où c_{2l+1} est le chiffre de poids $2l+1$ dans l'écriture binaire de N , i_2, i_3 sont des restes modulo 3, les coefficients ε_1 et ε_2 sont dans $\{0, 1, -1\}$ et N_2 est le nombre obtenu en enlevant à N (écrit en $2l+2$ chiffres) les deux chiffres de poids le plus fort. D'après les formules du Théorème 2, la somme des deux premiers termes est toujours inférieure à $3^l + 3^{l-1}$, ce qui permet de conclure par l'hypothèse de récurrence. \square

Passons maintenant au reste 2 modulo 3. La courbe 5 suggère l'inégalité suivante :

Théorème 5. $\tilde{S}_{3,2}(N) \leq 0$ pour tout N .

Démonstration. Soit N un entier naturel. Si N n'est pas une puissance de deux (auquel cas, l'énoncé est une conséquence directe du théorème 2), nous pouvons écrire en reprenant les notations de la partie 2.1 :

$$\tilde{S}_{3,2}(N) = S_{3,2}(2^{k_1}) - S_{3,\rho_1}(2^{k_2}) + \tilde{S}_{3,\rho_2}(N_2)$$

où k_1 et k_2 sont les numéros des deux chiffres un de poids le plus fort dans l'écriture binaire de N , et on a : $N = 2^{k_1} + 2^{k_2} + N_2$. On a quatre cas suivant la parité de k_1 et de k_2 :

1. Si k_1 est impair et k_2 est pair, alors

$$\tilde{S}_{3,2}(N) = -2 \cdot 3^{\frac{k_2}{2}-1} + \tilde{S}_{3,2}(N_2) < 0$$

car d'après l'hypothèse de récurrence (N_2 étant strictement plus petit que N), on a $\tilde{S}_{3,2}(N_2) \leq 0$.

2. Si k_1 et k_2 sont impairs, alors

$$\tilde{S}_{3,2}(N) = -3^{\frac{k_2-1}{2}} + \tilde{S}_{3,1}(N_2).$$

La courbe de la partie 2.2 suggère qu'on a toujours $\tilde{S}_{3,1}(N_2) < 0$ (si $N_2 \geq 1$), et on prouvera en effet cette intuition plus loin dans cette partie.

3. Si k_1 et k_2 sont pairs, supposons d'abord que $k_2 > 0$. On a alors :

$$\tilde{S}_{3,2}(N) = -3^{\frac{k_1}{2}-1} + 3^{\frac{k_2}{2}-1} + \tilde{S}_{3,0}(N_2). \quad (16)$$

D'après le théorème 4 on a dans ce cas :

$$\tilde{S}_{3,2}(N) \leq -3^{\frac{k_1}{2}-1} + 3^{\frac{k_2}{2}-1} + \tilde{S}_{3,0}(2^{k_2} - 1) \leq -3^{\frac{k_2}{2}+1} + 3^{\frac{k_2}{2}-1} + 2 \cdot 3^{\frac{k_2}{2}-1} = 0.$$

Si $k_2 = 0$, l'expression précédente devient :

$$\tilde{S}_{3,2}(N) = -3^{\frac{k_1}{2}-1} - \tilde{S}_{3,1}(1) = -3^{\frac{k_1}{2}-1} + 1 \leq 0$$

car $\frac{k_1}{2} - 1 \geq 1$.

4. Si k_1 est pair et k_2 est impair, on a :

$$\tilde{S}_{3,2}(N) = -3^{\frac{k_1}{2}-1} + 3^{\frac{k_2-1}{2}} + \tilde{S}_{3,2}(N_2) \leq 0$$

d'après l'hypothèse de récurrence.

Montrons maintenant qu'on a en effet $\tilde{S}_{3,1}(N) < 0$ pour tout N ; cela se fait par une distinction des cas du même type que précédemment mais aucun des cas ne pose un problème. On a

$$S_{3,1}(N) = S_{3,1}(2^{k_1}) - S_{3,\rho_1}(2^{k_2}) + \tilde{S}_{3,\rho_2}(N_2)$$

et tous les trois termes sont négatifs sauf dans le cas où k_1 et k_2 sont impairs. Dans ce cas on obtient :

$$S_{3,1}(N) = -3^{\frac{k_1-1}{2}} + S_{3,0}(N_2)$$

et le chiffre du plus fort poids dans N_2 se trouve à la position $k_1 - 3$ au maximum. On conclut en utilisant la majoration (6). Ceci termine l'étude du dernier cas. □

Remarquons que ce résultat met à faux l'inégalité énoncée dans l'article [5] sous la forme suivante : $C_{2,3} = \limsup\left(\frac{S_{3,2}(N)}{N^{\alpha_3}}\right) > 0$.

3. Sur la raréfaction par rapport à un nombre composé.

3.1. Le cas d'une puissance d'un nombre premier.

Les cas des puissances de 3 et des puissances de 5 ont été étudiés par P. Grabner dans l'article [9]. Voici sa formule pour la 3^r -raréfaction :

$$\begin{aligned} S_{3^r,0}(3^r N) &= \frac{1}{3^{r-1+\alpha_3}} (3^r N)^{\alpha_3} F(\log_4(3^{r-1} N)) \\ &+ 3^{\frac{r\alpha_3}{3}-r} N^{\frac{\alpha_3}{3}} F_1\left(\frac{1}{3} \log_4 3^r N\right) + \dots \\ &+ 3^{\frac{r\alpha_3}{3^{w-1}}-r} N^{\frac{\alpha_3}{3^{w-1}}} F_{r-1}\left(\frac{1}{3^{r-1}} \log_4 3^r N\right) + \frac{\eta(3^r N)}{3^r}. \end{aligned}$$

Le premier terme de cette somme est le terme dominant, et il ressemble exactement à la 3-raréfaction car la fonction F qui y apparaît est la même que dans la Section 2. Les figures 6 à 14 montrent les courbes fractales $S_{p,i}(N)/N^\alpha$ qui correspondent aux suites 9-raréfiées pour les différentes classes résiduelles. On observe à chaque fois le dessin de la courbe qui correspond au reste correspondant modulo 3 décalée horizontalement et multipliée par un coefficient. Les points séparés près de la courbe montrent que le terme d'erreur est plus important ici que pour la 3-raréfaction. Sur les Figures 8, 11 et 14 qui correspondent au reste 2 modulo 3 on voit apparaître un autre phénomène : le terme d'erreur modifie la courbe quand la valeur de celle-ci s'approche de 0. Les Figures 15 et 16 montrent un cas où les termes d'erreur disparaissent lentement, et les points avant $N \approx 10000$ se trouvent loin de la courbe limite mais assez près les uns des autres pour former ce qui apparaît comme une spirale qui monte en s'approchant de la courbe de la fonction F . Les fonctions fractales de la raréfaction sont données par les valeurs limites.

Des formules similaires sont valables dans tous les cas où 2 engendre le groupe $(\mathbb{Z}/p^r\mathbb{Z})^\times$, comme c'est le cas quand $p = 3$ ou 5. L'article [9] montre aussi (sur l'exemple de la raréfaction par rapport à un nombre de type $3^r 5^t$) qu'on peut procéder par analogie avec le cas quadratique dans le cas d'un nombre composé qui vérifie la condition : 2 engendre le sous-groupe des carrés modulo q .

3.2. Un cas improbable.

Étudions de plus près la question suivante : dans quels cas est-ce que 2 engendre le groupe $(\mathbb{Z}/p^r\mathbb{Z})^\times$? Un calcul montre que pour cela il suffit que 2 engendre le groupe $(\mathbb{Z}/p^2\mathbb{Z})^\times$, et, plus généralement, si $2^s \not\equiv 1 \pmod{p^2}$, les groupes quotients $(\mathbb{Z}/p^r\mathbb{Z})^\times / \langle 2 \rangle$ et $(\mathbb{Z}/p\mathbb{Z})^\times / \langle 2 \rangle$ sont isomorphes.

La recherche des nombres premiers tels que $2^s \equiv 1 \pmod{p^2}$ montre qu'il y en a exactement deux plus petits que 5×10^9 : 1093 et 3511. Pour $p = 1093$, on a $s = 364 = \frac{1092}{3}$ et pour $p = 3511$, on a $s = 1755 = \frac{3510}{2}$.

Une première conjecture sur la fréquence des nombres premiers "exceptionnels" tels que $2^s \equiv 1 \pmod{p^2}$ serait la suivante : le nombre de nombres premiers inférieurs à x croît comme $\int \frac{dx}{\log x}$, l'avant-dernier chiffre de 2^s (ou, ce qui est équivalent pour cette question, de 2^{p-1}) est choisi de façon indifférente parmi les x possibilités, d'où le nombre de nombres premiers "exceptionnels" inférieurs à x croirait comme $\int \frac{dx}{x \log x} = \log \log x$. L'ensemble de tels nombres serait donc infini.

4. Fonctions fractales de la raréfaction.

Chaque figure donnée ici représente un nuage de points d'ordonnée $S_{q,i}(qN)/N^\alpha$ et d'abscisse $\{\log_{2^s}(N)\}$ où q , i , s et α sont des paramètres de la figure (q sera une puissance d'un nombre premier p , i sera un reste modulo q , s sera l'ordre de 2 dans le groupe multiplicatif $(\mathbb{Z}/p\mathbb{Z})^\times$ et $\alpha = \alpha_p$ sera l'exposant fractal qui correspond au nombre p). Les bornes dans lesquelles varie N dépendent de la figure. Aucun trait de segment n'a été tracé entre les points, donc ce qui apparaît comme une courbe sur la figure n'est en fait rien d'autre que l'ensemble des points qui correspondent aux N assez grands.

Les figures 15 et 16 montrent que la fonction fractale $F_{27,0}(x)$ apparaît lorsque le nombre de points N dépasse 20000 pour atteindre au moins 65000. Atteindre une représentation graphique "stable" nécessite des calculs longs, notamment lorsque p est grand.

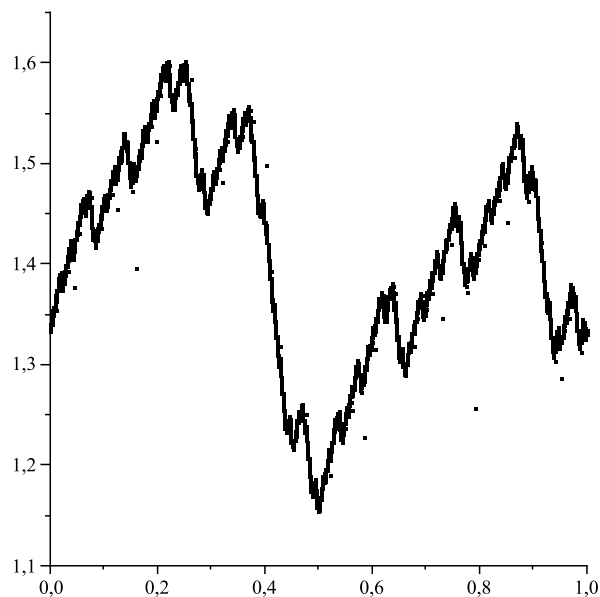


FIG. 3 – $S_{3,0}(3N)/N^{\alpha_3}$ en fonction de $\{\log_4(N)\}$ quand N varie de 2 à 10000.

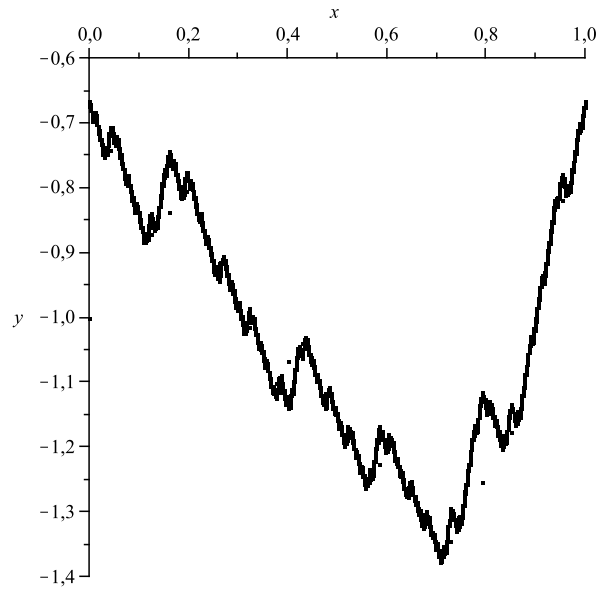


FIG. 4 – $S_{3,1}(3N)/N^{\alpha_3}$ en fonction de $\{\log_4(N)\}$ quand N varie de 1 à 10000.

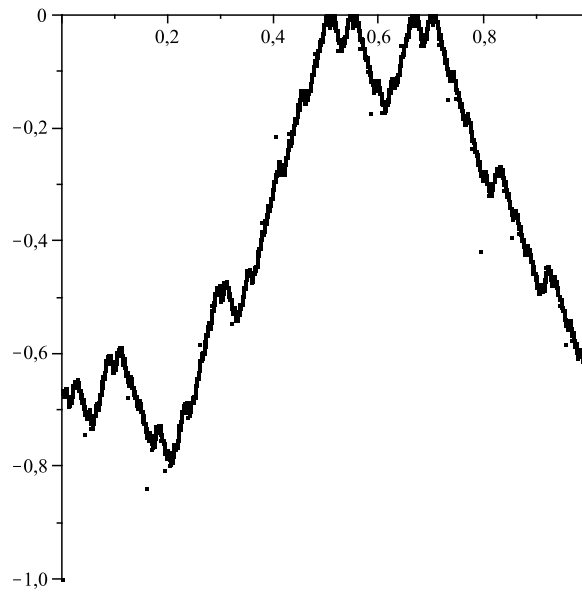


FIG. 5 – $S_{3,2}(3N)/N^{\alpha_3}$ en fonction de $\{\log_4(N)\}$ quand N varie de 1 à 10000.

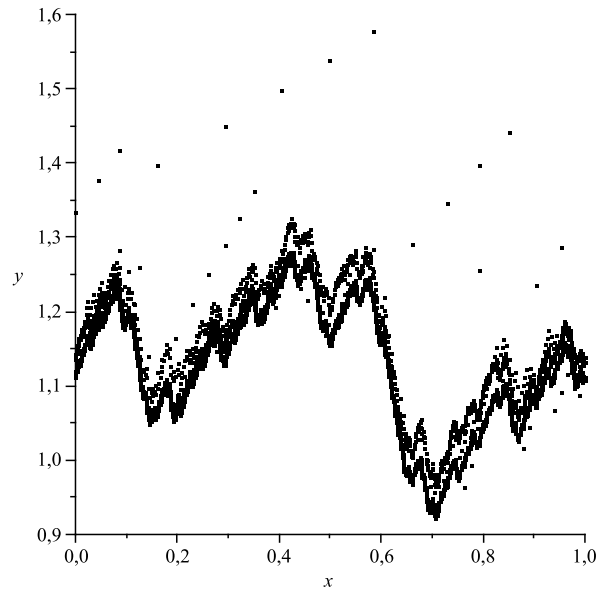


FIG. 6 – $S_{9,0}(9N)/N^{\alpha_3}$ en fonction de $\{\log_4(N)\}$ quand N varie de 2 à 10000.

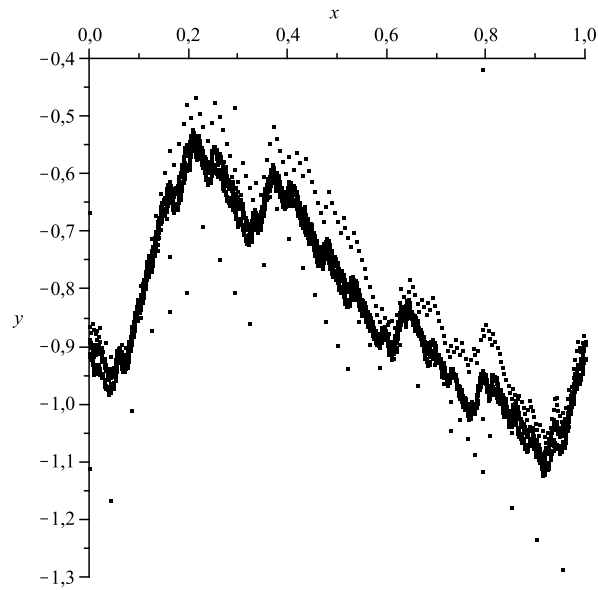


FIG. 7 – $S_{9,1}(9N)/N^{\alpha_3}$ en fonction de $\{\log_4(N)\}$ quand N varie de 2 à 10000.

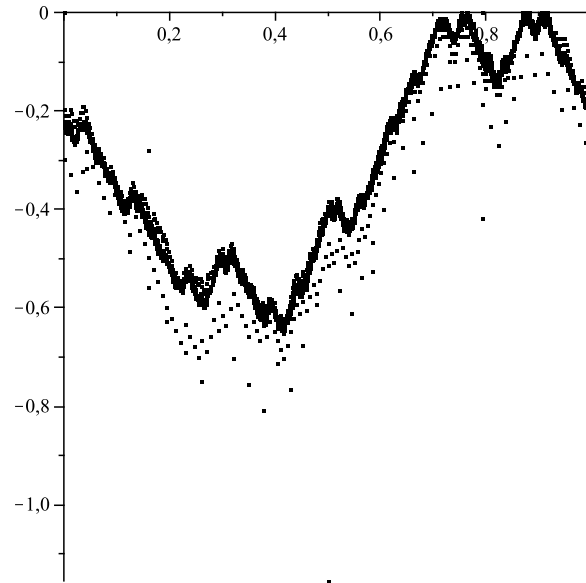


FIG. 8 – $S_{9,2}(9N)/N^{\alpha_3}$ en fonction de $\{\log_4(N)\}$ quand N varie de 2 à 10000.

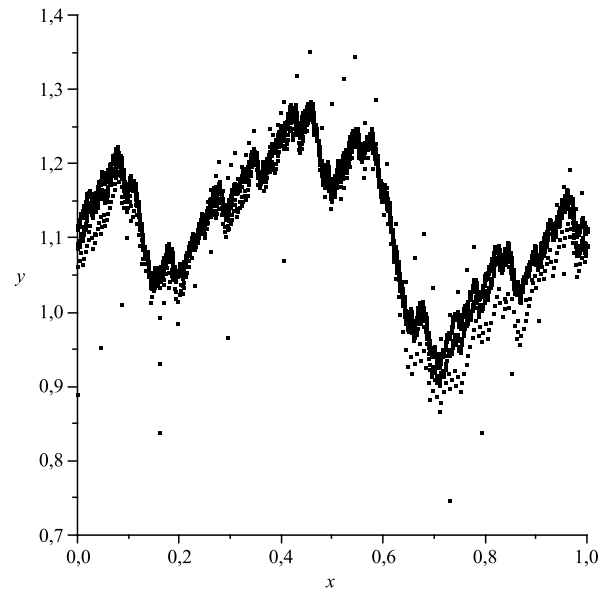


FIG. 9 – $S_{9,3}(9N)/N^{\alpha_3}$ en fonction de $\{\log_4(N)\}$ quand N varie de 2 à 10000.

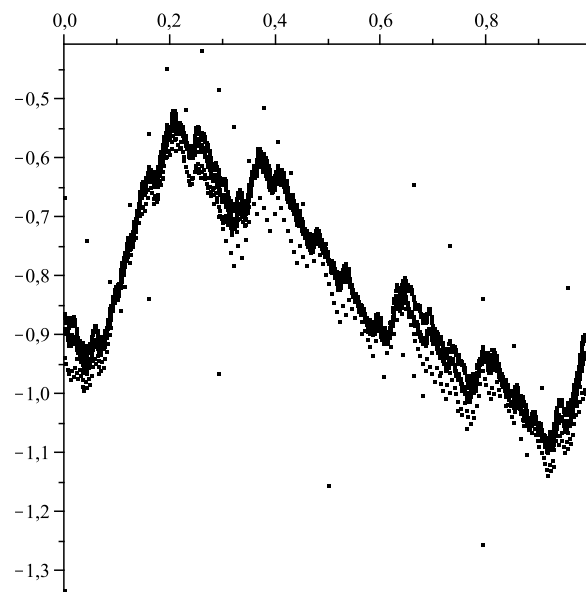


FIG. 10 – $S_{9,4}(9N)/N^{\alpha_3}$ en fonction de $\{\log_4(N)\}$ quand N varie de 2 à 10000.

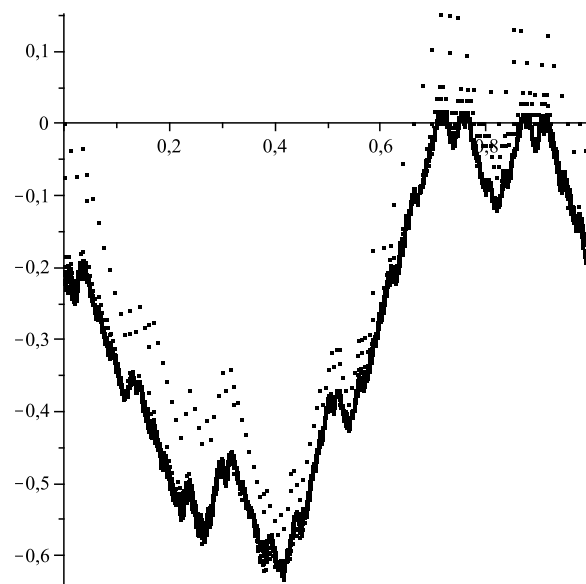


FIG. 11 – $S_{9,5}(9N)/N^{\alpha_3}$ en fonction de $\{\log_4(N)\}$ quand N varie de 4 à 10000.

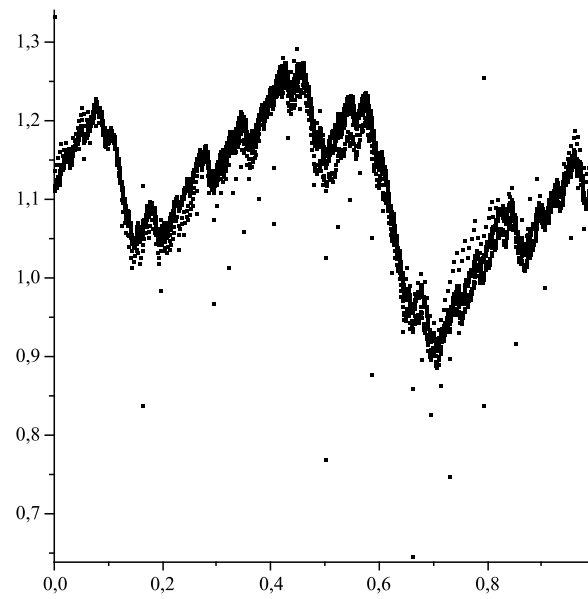


FIG. 12 – $S_{9,6}(9N)/N^{\alpha_3}$ en fonction de $\{\log_4(N)\}$ quand N varie de 2 à 10000.

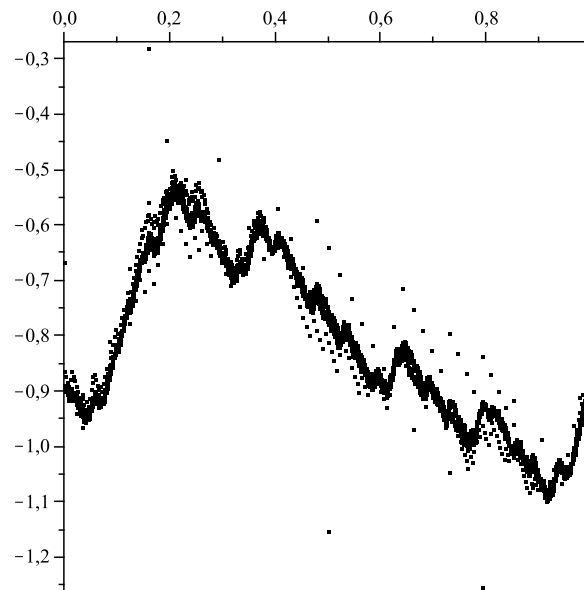


FIG. 13 – $S_{9,7}(9N)/N^{\alpha_3}$ en fonction de $\{\log_4(N)\}$ quand N varie de 2 à 10000.

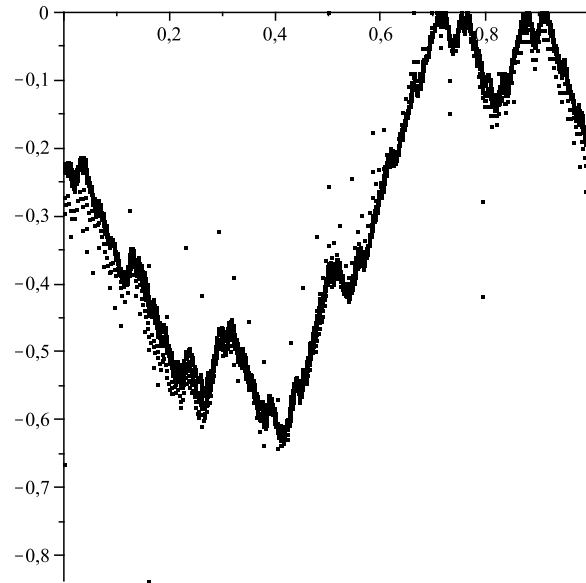


FIG. 14 – $S_{9,8}(9N)/N^{\alpha_3}$ en fonction de $\{\log_4(N)\}$ quand N varie de 2 à 10000.

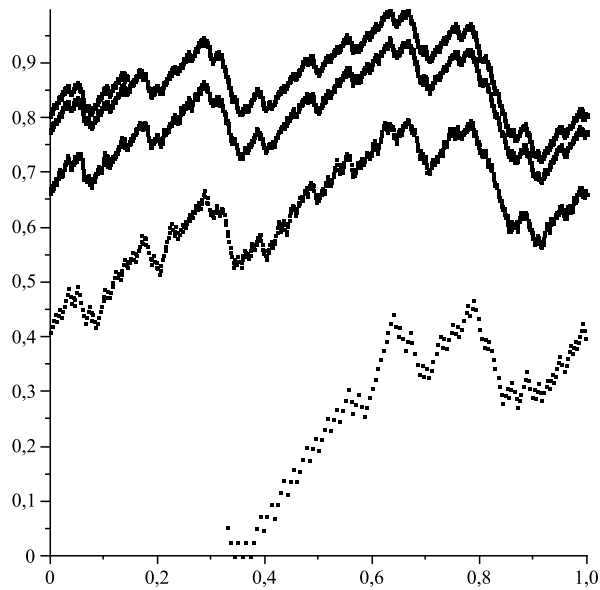


FIG. 15 – $S_{27,0}(27N)/N^{\alpha_3}$ en fonction de $\{\log_4(N)\}$ quand N varie de 100 à 20000.

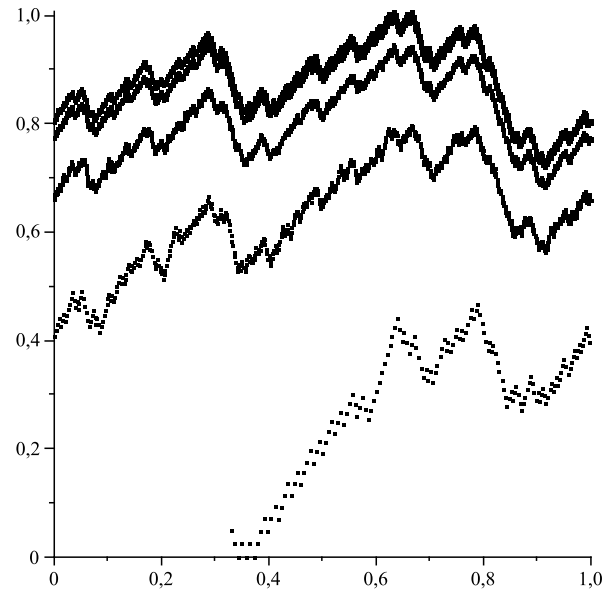


FIG. 16 – $S_{27,0}(27N)/N^{\alpha_3}$ en fonction de $\{\log_4(N)\}$ quand N varie de 100 à 65000.

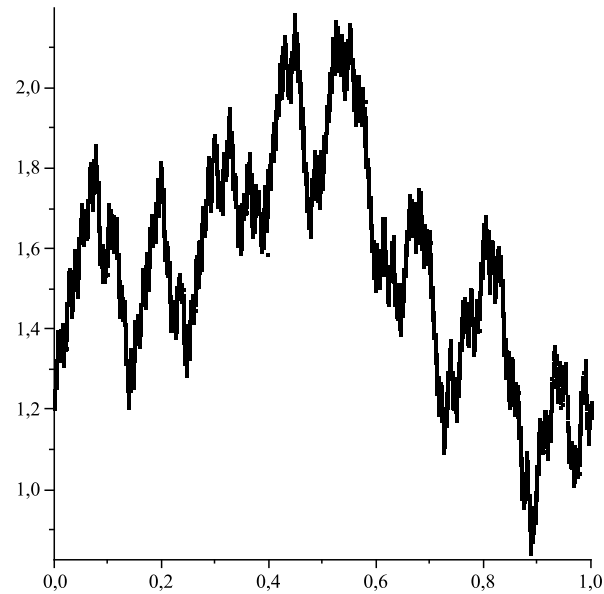


FIG. 17 – $S_{5,0}(5N)/N^{\alpha_5}$ en fonction de $\{\log_{16}(N)\}$ quand N varie de 2 à 50000.

Références

- [1] Z.I. Borevitch et I.R. Shafarevitch, *Théorie des nombres*, Gautier-Villars Paris (1967).
- [2] J. Coquet, A summation formula related to the binary digits, *Inv. Math.* **73** (1983), 123–137.
- [3] J. Coquet et M.Mendès-France, Suites à spectre vide et suites pseudo-aléatoires, *Acta Arithm.* **23** (1977), 99–106.
- [4] M. Drmota and M. Skalba, Rarified sums of the Thue-Morse sequence, *Trans. Amer. Math. Soc.* **352** (1999), 609–640.
- [5] J.M. Dumont, Discrépance des progressions arithmétiques dans la suite de Morse, *C. R. Acad. Sci. Paris Série I* **297** (1983), 145–148.
- [6] K.J. Falconer, The Geometry of fractal sets, Cambridge Tracts in Mathematics **85** Cambridge University Press (1986)
- [7] J.-P. Gazeau and J.-L. Verger-Gaugry, On the spectrum of the Thue-Morse quasicrystal and the rarefaction phenomenon, *J. Théor. Nombres Bordeaux* **20** (2008), 673–705.
- [8] S. Goldstein, K. Kelly and E. Speer, The Fractal Structure of Rarefied Sums of the Thue-Morse Sequence, *J. Number Theory* **42** (1992), 1–19.
- [9] P.J. Grabner, A note on the parity of the sum-of-digits function, *Actes 30ième Séminaire Lotharingien de Combinatoire (Gerolfingen,1993)*, 35–42.
- [10] P.G. Hartung, On the Pellian Equation, *J. Number Theory* **12** (1992), 110–112.
- [11] Ch. Hooley, On Artin’s conjecture, *J. Reine Angew. Math.* **225** (1967), 209–220.
- [12] D.H. Lehmer and E. Lehmer, Heuristics, anyone?, Studies in mathematical analysis and related topics, *Studies in matematics and statistics, IV*, Stanford University press (1962).
- [13] B.B. Mandelbrot, *The fractal geometry of nature*, W.H. Freeman and company New-York (1982).
- [14] D.J. Newman, On the number of binary digits in a mutiple of three, *Proc. Am. Math. Soc.* **21** (1969), 719–721.
- [15] J. Peyrière, Étude de quelques propriétés des produits de Riesz, *Ann. Inst. Fourier, Grenoble* **25**, 2 (1975), 127–169.