Bull. Soc. Math. France
147 (4), 2019, p. 661-704

BLOCS DE CHIFFRES DE TAILLE CROISSANTE DANS LES NOMBRES PREMIERS

PAR GAUTIER HANNA

RÉSUMÉ. — Dans cet article, nous démontrons un théorème à la Mauduit et Rivat (théorème des nombres premiers, principe d'aléa de Möbius) pour des fonctions qui comptent les blocs de chiffres dont la taille grandit lorsque la fonction tend à l'infini. Ces suites ne sont pas automatiques. Pour obtenir nos résultats, nous contrôlons des sommes de type I et II et utilisons une version améliorée de la propriété de propagation, ainsi que des méthodes classiques d'analyse harmonique.

ABSTRACT (*Blocks of digits of increasing size in primes*). — In this article, we prove a theorem à la Mauduit et Rivat (prime number theorem, Moebius randomness principle) for functions that count digital blocks whose length is a growing function tending to infinity. These sequences are not automatic. To obtain our results, we control sums of type I and II and use an adapted and refined version of the carry propagation property as well as standard methods from harmonic analysis.

1. Introduction

Dans [5], Gelfond fournit l'estimation asymptotique du nombre d'entiers en progression arithmétique dont la somme des chiffres est dans une classe de congruence fixée. Il termine son article par une série de questions, dont celle de l'estimation du nombre de nombres premiers inférieurs à x tels que

Texte reçu le 30 novembre 2016, modifié le 6 février 2019, accepté le 27 février 2019.

GAUTIER HANNA, CNRS, Institut Elie Cartan de Lorraine, UMR 7502, Vandoeuvre-lès-Nancy, F-54506, France • *E-mail :* gautier.hanna@univ-lorraine.fr

Classification mathématique par sujets (2010). — 11A63; 11B85, 11N05, 11L20.

Mots clefs. — Nombres premiers, Sommes d'exponentielles, Chiffres.

0037-9484/2019/661/$ 5.00
doi:10.24033/bsmf.2795

$s_q(p) \equiv a \bmod m$, où $s_q(\cdot)$ désigne la somme des chiffres en base q (q étant, et sera pour tout cet article, un entier supérieur ou égal à 2), et a et m deux entiers presque quelconques (une obstruction naturelle pour a et m, conséquence d'une généralisation de la preuve par 9, apparaît). Dans [14], Mauduit et Rivat arrivent à estimer cette quantité près de 40 ans après que cette question fut posée.

Peu de temps après la parution de [14], Kalai, dans [11], a posé une question qui peut s'apparenter à une extension de la question de Gelfond : il s'agit de montrer un théorème des nombres premiers pour une suite liée à une somme partielle des chiffres d'un entier. Green [7] répond partiellement à cette question, puis Bourgain [2] en s'inspirant des travaux de [14] complète la réponse de Green. Cette question possède une généralisation naturelle également proposée par Kalai [12]. Cette généralisation consiste à démontrer, si μ est la fonction de Möbius, $P \in \mathbb{F}_2[X_1, \dots, X_N]$ de degré inférieur à polylog (N) (le degré est ici le nombre maximal de variables entrant en compte dans un monôme, le polylog une puissance quelconque du logarithme), et si pour tout entier $n < 2^N$, $\epsilon_0(n), \dots, \epsilon_{N-1}(n)$ désignent les N premiers chiffres de n en base 2, ϵ_0 étant le chiffre des unités (les chiffres au-delà de N sont nécessairement nuls, il n'est pas utile de les marquer) que

$$\sum_{n<2^N} \mu(n)(-1)^{P(n)} = o(2^N), \tag{1}$$

avec ici

$$P(n) := P(\epsilon_0(n), \dots, \epsilon_{N-1}(n)). \tag{2}$$

Avec ces notations, puisque $s_2(n) = P(n)$ avec $P(X_1, \dots, X_N) = \sum_i X_i$, on voit bien que cette question peut être perçue comme une généralisation de celle de Gelfond. La première question de Kalai concernait les polynômes de degré au plus 1, la seconde s'intéresse aux polynômes de degré strictement plus grand que 1. Un exemple de polynôme de degré 2 est donné par $P(X_1, \dots, X_N) := \sum_i X_i X_{i+1}$ et $f(n) := (-1)^{P(n)}$ engendre alors la suite de Rudin-Shapiro. Dans ce cas, Tao [12] a esquissé une démonstration de (1) sans terme d'erreur explicite, et Mauduit et Rivat [15] en utilisant la méthode qu'ils ont développée dans [14] ont fourni un terme d'erreur explicite. De plus, Mauduit et Rivat ne se contentent pas de regarder la suite de Rudin-Shapiro, mais une classe de suite $(e(\alpha a(n)))_{n\geq 0}$ où $e(x) := \exp(2i\pi x), \alpha \in \mathbb{R}\backslash\mathbb{Z}$ et $(a(n))_{n\geq 0}$ est une suite liée aux comptages de blocs en base 2, ce que Mauduit et Rivat nomment suites de Rudin-Shapiro généralisée.

Dans [9], les résultats de [15] ont été étendus dans le cas où $(a(n))_{n\geq 0}$ est une suite β-récursive possédant certaines propriétés peu restrictives. Les suites β-récursives généralisent les suites bloc-additives (parfois nommées suites digitales) [1, Section 3.3], [3] dans le sens où cette fois-ci il est permis de compter les blocs $0 \cdots 0$. En particulier, dans [9], les résultats de [15] ont été étendus

à toutes les suites comptant les blocs, que ceux-ci soient connexes ou non (le bloc $1*1*00$ est non connexe) en base q quelconque. Les résultats de [15] et de [9] concernent des suites qui sont automatiques, et de ce fait sont impliqués par [16].

Une explicitation des calculs effectués dans [15] et [9] permet de montrer qu'il existe deux constantes absolues strictement positives C_1 et C_2 telles que

$$\sum_{n<2^N} \mu(n)(-1)^{P_N(n)} \ll N^{C_1} 2^{N - C_2 N^{\frac{1}{\beta 2^\beta}} + o\left(\frac{N}{\beta 2^\beta}\right)} \quad (N \to \infty), \tag{3}$$

avec $P_N(X_1, \dots, X_N) := \sum_{i \leq N-\beta+1} \tilde{P}(X_i, \dots X_{i+\beta-1})$ et $\tilde{P}$ un polynôme à β-variables vérifiant une non-trivialité proche de celle réclamée aux suite β-récursives, pour plus de précisions voir [8, Théorème 0.4.1]. En particulier le terme majorant de (3) reste $o(2^N)$ pour $\beta < \log N / \log 2$, degré tout à fait acceptable dans l'optique de la conjecture de Kalai. La conjecture de Kalai s'inscrit dans l'historique de la conjecture de Sarnak (voir [16] pour plus de références à ce sujet).

Une extension naturelle de ces résultats consiste à s'intéresser à la même question, mais pour des suites liées aux comptages de blocs de taille croissante. Si nous notons à présent $P : \mathbb{N} \to \mathbb{N}$ non plus un polynôme mais une application croissante, il convient de s'intéresser à

$$a_P(n) := \sum_{i \geq 0} \epsilon_{i+P(T_q(n))}(n) \cdots \epsilon_i(n) \tag{4}$$

où

$$T_q(\cdot) := \lfloor \log(\cdot) / \log q \rfloor \tag{5}$$

indique l'indice du dernier chiffre non nul dans l'écriture en base q.

En base 2, $a_P(n)$ compte les blocs composés de $P(T_2(n))$ '1' consécutifs ; si P est la fonction constante ceci correspond à la suite $(b_d(n))_{n \geq 0}$ introduite dans [15]. S'il était possible d'utiliser la méthode de Mauduit et Rivat pour $P(x) = x$; puisque celle-ci fournit un résultat d'équirépartition [15, Corollary 2], une conséquence serait que la moitié des nombres premiers seraient des nombres de Mersenne (voir Partie 11 pour plus de précision). Dans cet article, nous montrons que la vitesse $P(y) < \log y / \log q$ est admissible en démontrant le théorème suivant (ici et dans le reste de l'article, la constante implicite dépend de q) :

THÉORÈME 1.1. — *Soit P une fonction croissante, positive, et à valeurs entières pour laquelle il existe une constante strictement positive $c < 1/\log q$ telle que $P(y) \leq c \log y$ pour tout réel y assez grand. Alors uniformément en $\vartheta \in \mathbb{R}$:*

$$\left| \sum_{n \leq x} \Lambda(n) f_P(n) e(\vartheta n) \right| \ll c_1(q) (\log x)^{3 + \frac{\omega(q)}{4}} x q^{-\frac{1}{64} \gamma_P\left(\frac{1}{120} \lfloor \frac{\log x}{\log q} \rfloor, \lfloor \frac{\log x}{\log q} \rfloor\right)}, \tag{6}$$

avec

$$\gamma_P(l,k) = l\left(1 - \frac{\log\left(q^{P(k)} - 8\left(\sin\frac{\pi\|\alpha\|}{4}\right)^2\right)}{P(k)\log q}\right), \tag{7}$$

où Λ *est la fonction de von Mangoldt,* $f_P(n) = e(\alpha a_P(n))$ *et*

$$c_1(q) = q^{13/64}\max\left((\log q)^3, \tau(q)^{1/4}\right)(\log q)^{-3-\frac{\omega(q)}{4}}.$$

De plus ce théorème reste valable avec μ *en lieu et place de* Λ*.*

Remarque 1.2. — Si $P(x)$ vérifie les conditions du théorème, alors le majorant de l'équation (6) est $o(x)$ si $x \to \infty$. Pour plus de détail, nous laissons le lecteur se reporter à la remarque 9.1

Si $\alpha = 1/2$, f_P n'est pas automatique (voir Partie 10), de sorte que les résultats de cet article ne sont pas impliqués par [16].

2. Notations

Pour des raisons techniques nous ne regarderons pas la suite $(a_P(n))_{n\geq 0}$ mais une application $a_P : \mathbb{N}\times\mathbb{N} \to \mathbb{N}$ à deux variables définie par :

$$a_P(x,y) := \sum_{i\geq 0} \epsilon_{i+P(y)}(x)\cdots\epsilon_i(x). \tag{8}$$

Nous définissons alors pour ρ un entier, $a_P^{(\rho)}(x,y) := a_P(x \bmod q^\rho, y)$. Avec ces définitions, nous avons $a_P(n) = a_P(n, T_q(n))$ et $a_P^{(\rho)}(n) = a_P^{(\rho)}(n, T_q(n)) = a_P(n \bmod q^\rho, T_q(n))$, en rappelant que T_q est définie par (5).

Si α est un nombre réel, nous définissons $f_P(x,y) := e(\alpha a_P(x,y))$, $f_P(n) := f_P(n, T_q(n))$, *etc.* Nous définissons également pour $0 \leq \mu_1 \leq \mu_2$ des entiers

$$f_P^{(\mu_1,\mu_2)}(x,y) := f_P^{(\mu_2)}(x,y)\overline{f_P^{(\mu_1)}(x,y)}.$$

Il s'agit d'une fonction doublement tronquée qui jouera un rôle crucial dans les estimations des sommes de type II.

Si μ_0, μ_2 et n sont des entiers tels que $\mu_0 \leq \mu_2$, nous notons $r_{\mu_0,\mu_2}(n)$ l'entier u tel que $n = kq^{\mu_2} + uq^{\mu_0} + (n \bmod q^{\mu_0})$ et $0 \leq u < q^{\mu_2-\mu_0}$.

Remarque 2.1. — La fonction γ_P définie par (7) intervient dans le contrôle de la transformée de Fourier de f_P. Ainsi, tout comme dans [15, equation (25)] nous avons uniformément en $t \in \mathbb{R}$

$$1 = \sum_{0\leq h<q^\lambda}\left|\frac{1}{q^\lambda}\sum_{0\leq u<q^\lambda} f_P(uq^\kappa, k)e(-u(h+t))\right|^2 \leq q^{\lambda-2\gamma_P(\lambda,k)}$$

et donc $\gamma_P(\lambda,k) \leq \lambda/2$.

3. Panorama de la preuve

Depuis [15], une stratégie naturelle pour prouver un principe d'aléa de Möbius pour une fonction définie sur les chiffres est de vérifier les définitions de Mauduit et Rivat [15, Definition $1-2$]. S'il est possible dans notre cas, en ajustant les idées de [15], d'obtenir [15, Definition 2], en revanche, dès que P n'est plus constante, il est impossible d'obtenir l'autre condition.

En effet, si τ est tel que $P(\tau) \neq P(\tau-1)$, alors, pour n et ρ tels que $T_q(n) = \tau$ et $n \bmod q^\rho < n$, nous avons alors $P(T_q(n)) = P(\tau) \neq P(T_q(n \bmod q^\rho))$. Ceci veut dire que le nombre de chiffres considérés dans nos comptages de blocs pour n et $n \bmod q^\rho$ n'est pas le même et on ne peut pas espérer de simplification dans l'expression $a_P(n) - a_P(n \bmod q^\rho)$. Il n'y a donc *a priori* aucune raison de pouvoir contrôler le nombre de l tels que

$$a_P(lq^\kappa + k_1) - a_P(lq^\kappa + k_1 + k_2) \neq a_P((lq^\kappa + k_1) \bmod q^\rho) - a_P((lq^\kappa + k_1 + k_2) \bmod q^\rho)$$

si lq^κ est "beaucoup plus grand" que q^ρ. Il devient nécessaire de reprendre en profondeur les arguments de [15] pour espérer obtenir un résultat.

L'obtention du [15, Theorem 1] se fait à l'aide de l'étude des sommes de type I et de type II. L'adaptation de la technique utilisée dans [15] pour contrôler S_I, la somme de type I nécessite la définition suivante : pour n un entier en considérant l'écriture $n = u + vq^\kappa$ avec $0 \leq u < q^\kappa$, en posant $T_q(n) = l$, et $w = v \bmod q^\rho$, nous introduisons $n' = u + wq^\kappa + q^{\kappa+\rho}\lfloor q^{l-\rho} \rfloor$. Cette introduction permet de séparer l'information digitale de l'information multiplicative, ingrédient primordial dans la méthode de Mauduit et Rivat.

Pour S_{II}, l'introduction de la fonction doublement tronquée nécessite le Lemme 6.3, qui dit que si m et n sont des entiers, et m' et n' des petites perturbations de ces entiers, alors les produits respectifs auront le même nombre de chiffres. Il y a (au moins) deux manières différentes de démontrer ce lemme, l'une consistant en l'étude de la série de Fourier de la fonction $T_q(\cdot) := \lfloor \log(\cdot)/\log q \rfloor$, et elle nous a été suggérée par Olivier Robert, l'autre consiste à reprendre certaines idées évoquées dans [4, Lemma 3.5], et nous a été fournie par Thomas Stoll.

Ces introductions faites, plusieurs difficultés surgissent, mais leur résolution est plus de l'ordre de la technique et se fait à travers l'utilisation de résultats classiques (Kusmin-Landau, comportement en moyenne de la fonction τ, *etc.*)

Cet article est présenté comme suit. Les Parties 4 et 6 sont dédiées à la collecte de résultats utiles, respectivement, au traitement des sommes de type I et de type II. Ces résultats sont de différentes natures (analytiques, digitaux et harmoniques) et leur démonstration est assez différente du schéma global de la preuve principale : les inclure nuirait à la compréhension structurelle du traitement des sommes. Dans la Partie 5 nous exerçons un contrôle sur les sommes de type I, et dans la Partie 7, sur les sommes de type II. Dans la Partie 8 nous