Bull. Soc. Math. France
147 (2), 2019, p. 259-287

NOMBRES PREMIERS AVEC CONTRAINTES DIGITALES MULTIPLES

PAR BRUNO MARTIN, CHRISTIAN MAUDUIT & JOËL RIVAT

RÉSUMÉ. — Soit $q \geqslant 2$ un nombre entier. Pour tout nombre entier $n \geqslant 1$ et $k \in \{0, \ldots, q-1\}$ nous notons $|n|_k$ le nombre d'apparitions du chiffre k dans le développement de n en base q. Nous étudions la distribution jointe des fonctions $(n \mapsto |n|_k)_{1 \leqslant k < q}$ le long de la suite des nombres premiers. Nous obtenons notamment une formule asymptotique pour le cardinal de l'ensemble des nombres premiers p n'excédant pas x et équilibrés en base q (*i.e.* dont les chiffres en base q satisfont à la condition $||p|_i - |p|_j| \leqslant 1$ pour tout $(i, j) \in \{0, \ldots, q-1\}^2$).

Texte reçu le 28 septembre 2017, modifié le 20 juin 2018, accepté le 21 septembre 2018.

BRUNO MARTIN, Univ. Littoral Côte d'Opale, EA 2797 – LMPA – Laboratoire de Mathématiques pures et appliquées Joseph Liouville, 62228 Calais, France • *E-mail :* `Bruno.Martin@univ-littoral.fr`

CHRISTIAN MAUDUIT, Université d'Aix-Marseille et Institut Universitaire de France, Institut de Mathématiques de Marseille CNRS UMR 7373, 163 avenue de Luminy, Case 907, 13288 Marseille Cedex 9, France • *E-mail :* `mauduit@iml.univ-mrs.fr`

JOËL RIVAT, Université d'Aix-Marseille, Institut de Mathématiques de Marseille CNRS UMR 7373, 163 avenue de Luminy, Case 907, 13288 Marseille Cedex 9, France • *E-mail :* `rivat@iml.univ-mrs.fr`

Classification mathématique par sujets (2010). — 11A63, 11J71, 11L20, 11N05.

Mots clefs. — nombres premiers, sommes d'exponentielles, fonctions digitales, équirépartition modulo 1.

Ce travail a bénéficié de l'aide de l'Agence Nationale de la Recherche portant les références « ANR-14-CE34-0009 » MUDERA.

doi:10.24033/bsmf2781

ABSTRACT (*Prime numbers with multiple digital constraints*). — Let $q \geqslant 2$ be an integer. For every positive integer n and $k \in \{0, \dots, q-1\}$ we denote by $|n|_k$ the number of occurrences of the digit k in the representation of n in base q. We study the joint distribution of the functions $(n \mapsto |n|_k)_{1 \leqslant k < q}$ along the sequence of prime numbers. We obtain an asymptotic formula for the number of primes less than x which are balanced in base q (*i.e.* such that the digits in base q satisfy the condition $||p|_i - |p|_j| \leqslant 1$ for all $(i,j) \in \{0, \dots, q-1\}^2$).

1. Notations et rappels

Dans tout ce qui suit, q est un nombre entier supérieur ou égal à 2 et p désigne systématiquement un nombre premier. On note $\log_q$ la fonction logarithme en base q. Pour $x \in \mathbb{R}$, $\pi(x)$ désigne le nombre de nombres premiers n'excédant pas x, tandis que pour $a \in \mathbb{Z}$ et $b \in \mathbb{N}^*$, $\pi(x; a, b)$ désigne le nombre de nombres premiers p n'excédant pas x et tels que $p \equiv a \mod b$. Pour $t \in \mathbb{R}$, on note $\|t\|$ la distance de t au nombre entier le plus proche et l'on pose $\mathrm{e}(t) = e^{2i\pi t}$. Si $\boldsymbol{u} = (u_1, \dots, u_n)$ et $\boldsymbol{v} = (v_1, \dots, v_n)$ sont des éléments de $\mathbb{R}^n$, on note $\boldsymbol{u} \cdot \boldsymbol{v} = \sum_{j=1}^n u_j v_j$ leur produit scalaire usuel et $\|\boldsymbol{u}\|_\infty = \max_{1 \leqslant j \leqslant n} |u_j|$. La fonction indicatrice d'Euler est désignée par la lettre φ.

Rappelons que tout nombre entier strictement positif n admet un unique développement q-adique de la forme

$$n = \sum_{j=0}^{\nu} n_j q^j, \quad 0 \leqslant n_j \leqslant q-1, \quad n_\nu \geqslant 1. \tag{1}$$

Pour $n = 0$, on convient de poser $\nu = 0$ et $n_0 = 0$. Conformément à l'usage, on désigne pour tout $0 \leqslant k \leqslant q-1$ par $|\cdot|_k$ la fonction comptant le nombre d'apparitions du chiffre k dans le développement en base q, soit

$$|n|_k = \#\{0 \leqslant j \leqslant \nu \,:\, n_j = k\}$$

et on appelle fonction digitale (à valeurs entières) toute fonction $g : \mathbb{N} \to \mathbb{Z}$ de la forme

$$g(n) = \sum_{k=0}^{q-1} a_k |n|_k,$$

avec $(a_0, ..., a_{q-1}) \in \mathbb{Z}^q$. Un exemple classique de fonction digitale est donné par la fonction somme des chiffres en base q définie, pour tout nombre entier n, par

$$\mathrm{s}_q(n) = \sum_{j=0}^{\nu} n_j = \sum_{k=1}^{q-1} k|n|_k.$$

2. Présentation des résultats

L'objet de cet article est d'étudier la distribution jointe du $(q-1)$-uplet de fonctions $(n \mapsto |n|_k)_{1 \leqslant k \leqslant q-1}$ le long de la suite des nombres premiers.

2.1. Distribution dans les progressions arithmétiques. — C'est avec l'article [5] de Gelfond que s'ouvre le problème de déterminer la répartition dans les progressions arithmétiques des valeurs prises par une fonction digitale à valeurs entières le long d'une suite donnée. En particulier Gelfond pose la question de déterminer la répartition de $\mathrm{s}_q(p)$ dans les progressions arithmétiques lorsque p parcourt l'ensemble des nombres premiers. Ce problème a été résolu par Mauduit et Rivat qui obtiennent, étant donné un nombre entier $m \geqslant 2$, l'existence de $\sigma_{q,m} > 0$ tel que pour tous $a \in \mathbb{Z}$, $x \geqslant 2$, on a

$$\begin{aligned}\operatorname{card}\{p \leqslant x \,:\, \mathrm{s}_q(p) \equiv a \bmod m\} = \frac{\operatorname{pgcd}(m, q-1)}{m} & \pi(x; a, \operatorname{pgcd}(m, q-1)) \\ & + O\left(x^{1-\sigma_{q,m}}\right).\end{aligned}$$

Plusieurs auteurs ont étendu ce résultat à d'autres classes de fonctions digitales ou automatiques (cf. [10], [1], [12], [6], [14]). Nous obtenons dans cet article un premier exemple de résultat de ce type dans le cas multi-dimensionnel.

Théorème 2.1. — *Il existe $c = c(q) > 0$ tel que l'on a uniformément pour tous $x \geqslant 2$, $\boldsymbol{m} = (m_1, \dots, m_{q-1}) \in (\mathbb{Z} \setminus \{0\})^{q-1}$ et $\boldsymbol{\ell} = (\ell_1, \dots, \ell_{q-1}) \in \mathbb{Z}^{q-1}$,*

$$\begin{aligned}\operatorname{card}\{p \leqslant x \,:\, \forall k \in \{1, \dots, q-1\}, |p|_k \equiv \ell_k \bmod m_k\} = \\ \frac{d_{\boldsymbol{m}} \pi(x; v_{\boldsymbol{\ell}}, d_{\boldsymbol{m}})}{m_1 \dots m_{q-1}} + O\left((\log x)^3 x^{1-c/\|\boldsymbol{m}\|_\infty^4}\right),\end{aligned}$$

où l'on a posé

$$d_{\boldsymbol{m}} = \operatorname{pgcd}(m_1, 2m_2, \dots, km_k, \dots, (q-1)m_{q-1}, q-1) \quad \text{et} \quad v_{\boldsymbol{\ell}} = \sum_{k=1}^{q-1} k\ell_k. \tag{2}$$

La constante implicite ne dépend que de q.

Le théorème 2.1 n'a d'intérêt que si $(v_{\boldsymbol{\ell}}, d_{\boldsymbol{m}}) = 1$, puisque l'on a l'inclusion $\{p \leqslant x \,:\, \forall k \in \{1, \dots, q-1\}, |p|_k \equiv \ell_k \bmod m_k\} \subseteq \{p \leqslant x \,:\, p \equiv v_{\boldsymbol{\ell}} \bmod d_{\boldsymbol{m}}\}$.

2.2. Loi locale. — Étant donnée une fonction arithmétique $g : \mathbb{N}^* \to \mathbb{Z}$, un problème classique consiste à déterminer sa loi locale, c'est-à-dire obtenir une estimation asymptotique pour tout $k \in \mathbb{Z}$ de la quantité

$$\operatorname{card}\{n \leqslant x \,:\, g(n) = k\}$$

lorsque $x \to +\infty$. Les exemples classiques sont les fonctions Ω et ω (voir par exemple [15] chapitre II.6.1) qui dénombrent le nombre de facteurs premiers d'un nombre entier, comptés avec ou sans leur ordre de multiplicité. Plusieurs travaux récents traitent du cas où g est une fonction digitale. Le premier résultat de cet ordre a été obtenu par Mauduit et Sárközy dans [13] qui établissent une formule asymptotique pour $\mathrm{card}\{n < q^\nu \,:\, \mathrm{s}_q(n) = k\}$, uniforme pour $\nu \to +\infty$ et $0 \leqslant k \leqslant \frac{q-1}{2}\nu$. En s'appuyant sur ces résultats lorsque k est proche de la valeur moyenne $\frac{q-1}{2}\nu$, Fouvry et Mauduit étudient dans [4] l'ensemble

$$\mathcal{E} = \left\{ n \in \mathbb{N}^* \,:\, \mathrm{s}_q(n) = \tfrac{q-1}{2} \left\lfloor \log_q n \right\rfloor + B(\lfloor \log_q n \rfloor) \right\},$$

où $B : \mathbb{N} \to \mathbb{R}$ est une fonction telle que pour tout $j \in \mathbb{N}$, $\frac{q-1}{2}j + B(j) \in \mathbb{N}$ et telle que

(3) il existe $K > 0$ tel que pour tout $n \in \mathbb{N}^*, |B(n)| \leqslant K n^{1/4}$.

Par exemple, pour $g = \mathrm{s}_2$ et $B(j) = \lfloor j/2 \rfloor - j/2$, on a [1]

$$\mathcal{E} = \left\{ n \in \mathbb{N}^* \,:\, \mathrm{s}_2(n) = \left\lfloor \frac{1}{2} \log_2 n \right\rfloor \right\}.$$

Le théorème 1.1 de [4] montre que pour tout $x \geqslant 2$, on a

$$\mathrm{card}\{n \leqslant x \,:\, n \in \mathcal{E}\} = \sqrt{\frac{6}{\pi(q^2-1)}} \frac{x}{\sqrt{\log_q x}} + O\left(\frac{x}{\log_q x} \right), \tag{4}$$

où la constante implicite ne dépend que de K et q. Fouvry et Mauduit étudient alors la régularité statistique de l'ensemble $\mathcal{E}$ en établissant que pour tout $\beta \in \mathbb{R} \setminus \mathbb{Q}$, la suite $(n\beta)_{n \in \mathcal{E}}$ est équirépartie modulo 1 (théorème 1.2 de [4]).

Remarque 2.2. — La formule (4) n'est plus valable en toute généralité sous la condition [2] :

(5) il existe $K > 0$ tel que pour tout $n \in \mathbb{N}^*, |B(n)| \leqslant K n^{1/2}$,

mais le théorème 1.2 de [4] reste valide sous la condition (5). En effet, sous la condition (5), on peut assez rapidement vérifier qu'il existe des constantes $C_1, C_2 > 0$ dépendant de K et q telles que pour tout $x \geqslant x_0(K, q)$ on a

$$C_1 \frac{x}{\sqrt{\log x}} \leqslant \mathrm{card}\{n \leqslant x \,:\, n \in \mathcal{E}\} \leqslant C_2 \frac{x}{\sqrt{\log x}}$$

1. Rappelons que pour tout $y \in \mathbb{R}$, on a $\lfloor y/2 \rfloor = \lfloor \lfloor y \rfloor /2 \rfloor$.

2. Par exemple pour $q = 2$, et

$$B(j) = \begin{cases} \left\lfloor \sqrt{j} \right\rfloor & \text{si } j \text{ est pair,} \\ -\{j/2\} & \text{si } j \text{ est impair,} \end{cases}$$

on peut montrer, à l'aide de la formule de Stirling, que la suite

$$\left(\left(\frac{2^\nu}{\sqrt{\nu}} \right)^{-1} \mathrm{card}\{n \leqslant 2^\nu \,:\, n \in \mathcal{E}\} \right)_{\nu \in \mathbb{N}^*} = \left(\left(\frac{2^\nu}{\sqrt{\nu}} \right)^{-1} \binom{\nu}{\frac{\nu}{2} + B(\nu)} \right)_{\nu \in \mathbb{N}^*}$$

ne converge pas.

et ceci permet de constater, en examinant la démonstration du théorème 1.2 de [4], que la condition (5) reste suffisante pour garantir que pour tout $\beta \in \mathbb{R} \setminus \mathbb{Q}$ la suite $(\beta n)_{n \in \mathcal{E}}$ est équirépartie modulo 1.

Dans [11], Mauduit étudie les ensembles de nombres entiers reconnaissables par un q-automate infini déterministe associé à une marche aléatoire de moyenne nulle sur un réseau de dimension d, l'exemple le plus simple étant l'ensemble des nombres entiers $n \in \mathbb{N}^*$ satisfaisant à $|n|_0 = |n|_1 = \ldots = |n|_{q-1}$ (on a ici $d = q-1$). Plus généralement Drmota et Mauduit [2] considèrent

$$\mathcal{N} = \left\{ n \in \mathbb{N}^* \,:\, L_k(|n|_0, |n|_1, \ldots, |n|_{q-1}) = \lfloor \eta_k \log_q n \rfloor + v_k,\ k = 1, \ldots, m \right\},$$

où $\mathcal{L} = (L_1, \ldots, L_m)$ est une famille de formes linéaires définies sur $\mathbb{R}^q$, $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_m) \in \mathbb{R}^m$ et $\boldsymbol{v} = (v_1, \ldots, v_m) \in \mathbb{Z}^m$. Sous des hypothèses assez générales sur le triplet $(\mathcal{L}, \boldsymbol{\eta}, \boldsymbol{v})$ que nous ne détaillons pas ici, Drmota et Mauduit déterminent l'ordre de grandeur du cardinal de $\mathcal{N} \cap [1, x]$, et montrent que pour tout $\beta \in \mathbb{R} \setminus \mathbb{Q}$, la suite $(\beta n)_{n \in \mathcal{N}}$ est équirépartie modulo 1.

Dans [3], Drmota, Mauduit et Rivat étudient la loi locale de s_q au voisinage de sa valeur moyenne le long de la suite des nombres premiers : ils obtiennent que pour tous $\varepsilon \in]0, 1/2[$, $k \in \mathbb{N}^*$ tel que $(k, q) = 1$, on a

$$\begin{aligned} \operatorname{card}\{p \leqslant x \,:\, \mathrm{s}_q(p) = k\} = {} & \frac{q-1}{\varphi(q-1)} \frac{\pi(x)}{\sqrt{2\pi\sigma_q^2 \log_q x}} \\ & \cdot \left(e^{-\frac{(k - \mu_q \log_q x)^2}{2\sigma_q^2 \log_q x}} + O\left((\log x)^{-\frac{1}{2} + \varepsilon} \right) \right), \end{aligned} \tag{6}$$

où $\mu_q = \frac{q-1}{2}$ et $\sigma_q^2 = \frac{q^2-1}{12}$. Ce résultat est étendu dans [8] à toute fonction $g = \sum_{1 \leqslant k < q} a_k |\cdot|_k$ avec $(a_k)_{1 \leqslant k < q} \in \mathbb{Z}^{q-1}$. Nous obtenons dans cet article un résultat de ce type dans le cas multidimensionnel.

Théorème 2.3. — *Soit $\varepsilon \in]0, 1/2[$, $r \in \{1, \ldots, q-1\}$ et $\boldsymbol{k} = (k_1, \ldots, k_r) \in \mathbb{N}^r$ tel que $1 \leqslant k_1 < \ldots < k_r \leqslant q-1$. On a uniformément pour tous $x \geqslant 2$ et $\boldsymbol{b} = (b_1, \ldots, b_r) \in \mathbb{N}^r$,*

$$\begin{aligned} \operatorname{card}\{p \leqslant x \,:\, |p|_{k_1} = b_1, \ldots, |p|_{k_r} = b_r\} = {} & \delta_{\boldsymbol{k}} C_{r,q} \frac{\pi(x; \boldsymbol{b} \cdot \boldsymbol{k}, \delta_{\boldsymbol{k}})}{(\log_q x)^{r/2}} e^{-\frac{1}{2} R(u_{x,\boldsymbol{b}})} \\ & + O\left(\frac{\pi(x)}{(\log x)^{\frac{r+1}{2} - \varepsilon}} \right), \end{aligned} \tag{7}$$