Functiones et Approximatio
49.1 (2013), 7–27
doi: 10.7169/facm/2013.49.1.1

CONSTRUCTION OF NORMAL NUMBERS BY CLASSIFIED PRIME DIVISORS OF INTEGERS II

Jean-Marie De Koninck, Imre Kátai

Abstract: Given an integer $q \geqslant 2$, a *q-normal number* is an irrational number η such that any preassigned sequence of k digits occurs in the q-ary expansion of η at the expected frequency, namely $1/q^k$. In a series of recent papers, using the complexity of the multiplicative structure of integers along with a method we developed in 1995 regarding the distribution of subsets of primes in the prime factorization of integers, we initiated new methods allowing for the creation of large families of normal numbers. Here, we further expand on this initiative.

Keywords: normal numbers, primes, arithmetic function.

1. Introduction

Given an integer $q \geqslant 2$, a *q-normal number*, or simply a *normal number*, is an irrational number whose q-ary expansion is such that any preassigned sequence, of length $k \geqslant 1$, of base q digits from this expansion, occurs at the expected frequency, namely $1/q^k$.

Let $A_q = \{0, 1, \dots, q-1\}$ be the set of digits in base q. An expression of the form $i_1 i_2 \dots i_k$, where each $i_j \in A_q$, is said to be a word of length k. Given a word α, we shall write $\lambda(\alpha) = t$ to indicate that α is a word of length t. We shall also use the symbol Λ to denote the *empty word* and write $\lambda(\Lambda) = 0$. Also, we let A_q^k stand for the set of all words of length k and A_q^* stand for the set of all the words regardless of their length.

In 1995 (see [1]), we introduced the notion of a *disjoint classification of primes*, that is a collection of $q+1$ disjoint sets of primes $\mathcal{R}, \wp_0, \wp_1, \dots, \wp_{q-1}$, whose union is $\wp$, the set of all primes, where $\mathcal{R}$ is a finite set (perhaps empty) and where the other q sets are of positive densities $\delta_0, \delta_1, \dots, \delta_{q-1}$ (with clearly $\sum_{i=0}^{q-1} \delta_i = 1$). We then introduced the function $H : \mathbb{N} \to A_q^*$ defined by $H(n) = H(p_1^{a_1} \cdots p_r^{a_r}) = \ell_1 \dots \ell_r$,

The first author was partially supported by a grant from NSERC. The second author was partially supported by a grant from the European Union and the European Social Fund TÁMOP 4.2.1/B-09/1/KMR-2010-0003.

2010 Mathematics Subject Classification: primary: 11K16; secondary: 11N37, 11A41 where each ℓ_j is such that $p_j \in \wp_{\ell_j}$, and investigated the size of the set of positive integers $n \leqslant x$ for which $H(n) = \alpha$ for a given word $\alpha \in A_q^k$. More precisely, letting $\omega(n)$ stand for the number of distinct prime factors of n, and letting $P(n)$ and $p(n)$ stand respectively for the largest and smallest prime factor of n, writing $\pi(B)$ for the number of primes belonging to the set B and writing $x_1 = \log x$, $x_2 = \log x_1$ and so on, we proved the following result.

Theorem A. Let $\mathcal{R}, \wp_0, \wp_1, \dots, \wp_{q-1}$ be a disjoint classification of primes such that

$$\pi([u, u+v] \cap \wp_i) = \delta_i \pi([u, u+v]) + O\left(\frac{u}{\log^{c_1} u}\right) \tag{1.1}$$

holds uniformly for $2 \leqslant v \leqslant u$, $i = 0, 1, \dots, q-1$, where $c_1 \geqslant 5$ is a constant, $\delta_0, \delta_1, \dots, \delta_{q-1}$ are positive constants such that $\sum_{i=0}^{q-1} \delta_i = 1$. Let $\lim_{x\to\infty} w_x = +\infty$, $w_x = O(x_3)$, $\sqrt{x} \leqslant Y \leqslant x$ and $1 \leqslant k \leqslant c_2 x_2$, where c_2 is an arbitrary constant. Let A be a positive integer such that $A \leqslant x_2$ and $P(A) \leqslant w_x$. Then,

$$\begin{aligned}
&\#\{n = An_1 \leqslant Y : p(n_1) > w_x,\ \omega(n_1) = k,\ H(n_1) = i_1 \dots i_k\} \\
&\qquad = (1 + o(1))\delta_{i_1} \cdots \delta_{i_k} \frac{Y}{A \log Y} t_k(Y) \varphi_{w_x}\left(\frac{k-1}{x_2}\right) F\left(\frac{k-1}{x_2}\right),
\end{aligned}$$

where $t_k(x) = \dfrac{x_2^{k-1}}{(k-1)!}$,

$$\varphi_w(z) := \prod_{p \leqslant w} \left(1 + \frac{z}{p}\right)^{-1} \qquad \text{and} \qquad F(z) := \frac{1}{\Gamma(z+1)} \prod_p \left(1 + \frac{z}{p}\right)\left(1 - \frac{1}{p}\right)^z.$$

In 2011 (see [2]), we used Theorem A to construct large families of normal numbers, namely by establishing the following result.

Theorem B. Let $q \geqslant 2$ be an integer and let $\mathcal{R}, \wp_0, \wp_1, \dots, \wp_{q-1}$ be a disjoint classification of primes. Assume that, for a certain constant $c_1 \geqslant 5$,

$$\pi([u, u+v] \cap \wp_i) = \frac{1}{q} \pi([u, u+v]) + O\left(\frac{u}{\log^{c_1} u}\right) \tag{1.2}$$

uniformly for $2 \leqslant v \leqslant u$, $i = 0, 1, \dots, q-1$, as $u \to \infty$. Furthermore, let $H : \wp \to A_q^*$ be defined by

$$H(p) = \begin{cases} \Lambda & \text{if } p \in \mathcal{R}, \\ \ell & \text{if } p \in \wp_\ell \text{ for some } \ell \in A_q \end{cases} \tag{1.3}$$

and further let $T : \mathbb{N} \to A_q^*$ be defined by $T(1) = \Lambda$ and for $n \geqslant 2$ by

$$T(n) = T(p_1^{a_1} \cdots p_r^{a_r}) = H(p_1) \dots H(p_r). \tag{1.4}$$

Then, the number $0.T(1)T(2)T(3)T(4)\dots$ is a q-normal number.

As one will notice, Theorem B does not use the full power of Theorem A. Indeed, it is clear that condition (1.2) is much more restrictive than condition (1.1) since it does not allow for subsets of primes $\wp_j$ of distinct densities. In this paper, we will first weaken condition (1.2) to allow for the construction of even larger families of normal numbers. Then, we will extend our method in order to construct normal numbers using the sequence of shifted primes, and thereafter using the sequence n^2+1, $n=1,2,\ldots$

Finally, let us mention some notations we shall be using. As usual, φ will stand for the Euler function and $\mathrm{li}(x)$ for the logarithmic integral, that is $\mathrm{li}(x)=\displaystyle\int_0^x \frac{dt}{\log t}$. Throughout this text, the letters p, $p_1,p_2,\ldots$, $q_1,q_2,\ldots$, $\pi_0,\pi_1,\pi_2,\ldots$ will always denote primes, while the letter c will stand for a positive constant, but not necessarily the same at each occurrence.

2. Main results

Theorem 1. *Assume that $\mathcal{R},\wp_0,\ldots,\wp_{q-1}$ are disjoint sets of primes, whose union is $\wp$, and assume that there exists a positive number $\delta<1$ and a real number $c_1\geqslant 5$ such that*

$$\pi([u,u+v]\cap\wp_i)=\delta\pi([u,u+v])+O\left(\frac{u}{\log^{c_1}u}\right) \tag{2.1}$$

holds uniformly for $2\leqslant v\leqslant u$, $i=0,1,\ldots,q-1$, and similarly

$$\pi([u,u+v]\cap\mathcal{R})=(1-q\delta)\pi([u,u+v])+O\left(\frac{u}{\log^{c_1}u}\right).$$

Let H and T be defined as in (1.3) and (1.4). Then,

$$\xi=0.T(1)T(2)T(3)\ldots$$

is a q-normal number.

Examples

1. Let $\wp_0=\{p: p\equiv 1 \pmod 8\}$, $\wp_1=\{p: p\equiv 7 \pmod 8\}$ and $\mathcal{R}=\{2\}\cup\{p: p\equiv 3,5 \pmod 8\}$. With H, T and ξ as in the statement of Theorem 1, we may conclude that the number ξ is a binary normal number.
2. Let $P(x)=e_kx^k+\cdots+e_1x\in\mathbb{R}[x]$ be a polynomial with at least one irrational coefficient. Let I_0 and I_1 be two disjoint intervals in $[0,1)$ of equal length. Consider the set of primes $\wp_0=\{p:\{P(p)\}\in I_0\}$, $\wp_1=\{p:\{P(p)\}\in I_1\}$ and $\mathcal{R}=\wp\setminus(\wp_0\cup\wp_1)$. (Here, $\{P(p)\}$ stands for the fractional part of $P(p)$.) With H, T and ξ as in Theorem 1, we may conclude that ξ is a binary normal number.
3. It is well known that, given a prime $p\equiv 1 \pmod 4$, there exists a prime $\rho\in\mathbb{Z}[i]$ (the set of Gaussian integers) such that $\dfrac{\arg\rho}{\pi/2}\in[0,1)$ and $p=\rho\cdot\overline{\rho}$.

So, let the subsets of primes $\wp_0, \dots, \wp_{q-1}$ be defined in such a way that $p \in \wp_j$ if the corresponding Gaussian prime ρ satisfies

$$\frac{\arg \rho}{\pi/2} \in \left[\frac{j}{q}, \frac{j+1}{q}\right) \qquad (j = 0, 1, \dots, q-1)$$

and let $\mathcal{R} = \{2\} \cup \{p : p \equiv 3 \pmod 4\}$. Then, letting H, T and ξ be defined as in Theorem 1, we get that ξ is a normal number in base q.

Theorem 2. *Let $\mathcal{R}, \wp_0, \dots, \wp_{q-1}$, H and T be as in the statement of Theorem 1. Then the number*

$$\eta = 0.T(1)T(2)T(4)T(6)T(10)\dots T(p-1)\dots,$$

where p runs through the sequence of primes, is a q-normal number.

Theorem 3. *Let $f : \mathbb{N} \to \mathbb{N}$ be defined by $f(n) = n^2 + 1$. Consider the subset of primes $\widetilde{\wp} := \{p \in \wp : p \equiv 1 \pmod 4\}$. Assume that the sets $\wp_0, \wp_1, \dots, \wp_{q-1} \subseteq \widetilde{\wp}$ satisfy (2.1) and let*

$$\mathcal{R} = \wp \setminus \left(\bigcup_{j=0}^{q-1} \wp_j \right).$$

Let also H and T be defined as in (1.3) and (1.4). Then

$$\tau = 0.T(f(1))T(f(2))T(f(3))T(f(4))\dots$$

is a q-normal number.

3. Preliminary results

Lemma 1. *Let $f(n)$ be a real valued non negative arithmetic function. Let a_n, $n = 1, \dots, N$, be a sequence of integers. Let r be a positive real number, and let $p_1 < p_2 < \cdots < p_s \leqslant r$ be prime numbers. Set $Q = p_1 \cdots p_s$. If $d|Q$, then let*

$$\sum_{\substack{n=1 \\ a_n \equiv 0 \pmod d}}^{N} f(n) = \kappa(d)X + R(N, d),$$

where X and R are real numbers, $X \geqslant 0$, and $\kappa(d_1 d_2) = \kappa(d_1)\kappa(d_2)$ whenever d_1 and d_2 are co-prime divisors of Q.

Assume that for each prime p, $0 \leqslant \kappa(p) < 1$. Setting

$$I(N, Q) := \sum_{\substack{n=1 \\ (a_n, Q)=1}}^{N} f(n),$$

then the estimate

$$I(N,Q) = \{1+2\theta_1 H\} X \prod_{p|Q} (1-\kappa(p)) + 2\theta_2 \sum_{\substack{d|Q \\ d\leqslant z^3}} 3^{\omega(d)} |R(N,d)|$$

holds uniformly for $r \geqslant 2$, $\max(\log r, S) \leqslant \frac{1}{8}\log z$, *where* $|\theta_1| \leqslant 1$, $|\theta_2| \leqslant 1$, *and*

$$H = \exp\left(-\frac{\log z}{\log r}\left\{\log\left(\frac{\log z}{S}\right) - \log\log\left(\frac{\log z}{S}\right) - \frac{2S}{\log z}\right\}\right)$$

and

$$S = \sum_{p|Q} \frac{\kappa(p)}{1-\kappa(p)} \log p.$$

When these conditions are satisfied, there exists an absolute positive constant c *such that* $2H \leqslant c < 1$.

Proof. This result is Lemma 2.1 in the book of Elliott [3]. ■

Lemma 2 (Brun-Titchmarsh inequality). *There exists a positive constant* c *such that*

$$\pi(x; k, \ell) < c\frac{x}{\varphi(k)\log(x/k)} \qquad \text{for all } k < x.$$

Proof. For a proof, see the book of Halberstam and Richert [4]. ■

Lemma 3 (Bombieri-Vinogradov theorem). *Given any fixed number* $C > 0$, *there exists a number* $B = B(C) > 0$ *such that*

$$\sum_{k\leqslant\sqrt{x}/(\log^B x)} \max_{(k,\ell)=1} \max_{y\leqslant x} \left|\pi(y;k,\ell) - \frac{li(y)}{\varphi(k)}\right| = O\left(\frac{x}{\log^C x}\right).$$

Proof. For a proof, see the book of Iwaniec and Kowalski [5]. ■

For the statement of the next results, we shall need the following notations.

Let Z_x be a function tending to infinity but with the condition $\frac{\log Z_x}{\log x} \to 0$ as $x \to \infty$. Furthermore, let $K_x \to \infty$ as $x \to \infty$, but also satisfying $\dfrac{K_x \log Z_x}{\log x} \to 0$ as $x \to \infty$.

Let $Q = \prod_{p\leqslant Z_x} p$. Given an integer $m \geqslant 2$ such that $P(m) \leqslant Z_x$, we set

$$\mathcal{A}(x|m) = \#\{p \leqslant x : p \equiv 1 \pmod m,\ \gcd\left(\frac{p-1}{m}, Q\right) = 1\}.$$

Further set $\nu(Q) = \prod_{\substack{p|Q \\ p>2}} \left(1 - \frac{1}{p-1}\right)$.

We now introduce the strongly multiplicative function $\kappa(n)$ defined on primes p by

$$\kappa(p) = \begin{cases} 1 & \text{if } p = 2, \\ \frac{p-1}{p-2} & \text{if } p > 2. \end{cases} \tag{3.1}$$

Lemma 4. *Let Z_x and K_x be defined by $\log Z_x = (\log x)/x_2^2$ and $K_x = Bx_2$, where B is a large constant. Then, given any arbitrarily large constant C,*

$$\sum_{\substack{m \leqslant Z_x^{K_x} \\ P(m) \leqslant Z_x}} \left| \mathcal{A}(x|m) - \frac{\nu(Q)\kappa(m)}{m} li(x) \right| \ll \frac{x}{\log^C x}.$$

Proof. For now, we fix an integer $m \leqslant Z_x^{K_x}$ such that $P(m) \leqslant Z_x$. We plan to use Lemma 1. For this, we set $r = \pi(Z_x)$ and we let $q_1 < \cdots < q_T$ be the sequence of those primes $q_j \leqslant x$ satisfying $q_j - 1 \equiv 0 \pmod m$ for $j = 1, \dots, T$ (so that $T = \pi(x; m, 1)$); and also we let $a_n = (q_n - 1)/m$ for $n = 1, 2, \dots, T$ and set $f(n) = 1$. Now, define $R(m, d)$ implicitly by

$$\pi(x; dm, 1) = \sum_{\substack{p \leqslant x \\ \frac{p-1}{m} \equiv 0 \pmod d}} 1 = \eta(d)\pi(x; m, 1) + R(m, d), \tag{3.2}$$

where $\eta(d)$ is the strongly multiplicative function defined on primes p by

$$\eta(p) = \begin{cases} \frac{1}{p} & \text{if } p|m, \\ \frac{1}{p-1} & \text{if } (p, m) = 1. \end{cases}$$

Hence, as a consequence of Lemma 1, we obtain

$$\mathcal{A}(x|m) = \{1 + 2\theta_1 H\}\pi(x; m, 1) \prod_{p|Q} (1 - \eta(p)) + 2\theta_2 \sum_{\substack{d|Q \\ d \leqslant z^3}} 3^{\omega(d)} |R(m, d)|. \tag{3.3}$$

Now, since

$$S = \sum_{\substack{p|Q \\ p>2}} \frac{\log p}{p-2} = (1 + o(1)) \log Z_x \qquad (x \to \infty)$$

and

$$r = \pi(Z_x) \qquad \text{and} \qquad \log r = \log Z_x + O(\log\log x),$$

and since

$$\log z = K_x \log Z_x, \qquad \frac{\log z}{\log r} \sim K_x, \qquad \log\left(\frac{\log z}{S}\right) = \log K_x \qquad (x \to \infty),$$

we have, for x large,

$$H = \exp\{-K_x(\log K_x - \log\log K_x - z/K_x)\} \leqslant \exp\left\{-\frac{K_x}{2}\log K_x\right\}.$$

Hence, it follows from (3.3) that

$$\left|\mathcal{A}(x|m) - \pi(x;m,1)\frac{\varphi(m)}{m}\kappa(m)\nu(Q)\right| \leqslant 2H\pi(x;m,1)\nu(Q)\kappa(m) + 2\sum_{\substack{d|Q\\ d\leqslant z^3}} 3^{\omega(d)}|R(m,d)|, \quad (3.4)$$

where $R(m,d)$ satisfies, in light of (3.2),

$$|R(m,d)| \leqslant E(dm) + \frac{E(m)}{\varphi(d)}, \quad (3.5)$$

where

$$E(r) := \left|\pi(x;r,1) - \frac{\operatorname{li}(x)}{\varphi(r)}\right|.$$

Using (3.5), we have that

$$\begin{aligned}\sum_{\substack{d|Q\\ d\leqslant z^3}} 3^{\omega(d)}|R(m,d)| &\leqslant \sum_{\substack{d|Q\\ d\leqslant z^3}} 3^{\omega(d)}\left(E(dm) + \frac{E(m)}{\varphi(d)}\right)\\ &= \sum_{\substack{d|Q\\ d\leqslant z^3}} 3^{\omega(d)}E(dm) + \sum_{\substack{d|Q\\ d\leqslant z^3}} 3^{\omega(d)}\frac{E(m)}{\varphi(d)}\\ &= \sum\nolimits_1 + \sum\nolimits_2, \quad (3.6)\end{aligned}$$

say. Now, on the one hand,

$$\sum\nolimits_1 = \sum_{k\leqslant z^4} E(k)\prod_{p|k}(1+3) = \sum_{k\leqslant z^4} E(k)2^{2\omega(k)}. \quad (3.7)$$

On the other hand, we have

$$\sum\nolimits_2 \leqslant E(m)\sum_{d|Q}\frac{3^{\omega(d)}}{\varphi(d)} \leqslant E(m)\prod_{p|Q}\left(1+\frac{3}{p-1}\right) \leqslant cE(m)(\log Z_x)^3. \quad (3.8)$$

Thus, using (3.7) and (3.8) in (3.6), we obtain that

$$\begin{aligned}\sum_{\substack{d|Q\\ d\leqslant z^3}} 3^{\omega(d)}|R(m,d)| &\leqslant c(\log Z_x)^3E(m) + \sum_{k\leqslant z^4} E(k)2^{2\omega(k)}\\ &= T_1 + T_2, \quad (3.9)\end{aligned}$$

say. Now, because of Lemma 3, we have that, given any fixed constant C,

$$T_1 \ll \frac{x}{\log^C x}. \quad (3.10)$$

On the other hand, observe that since $a \leqslant b + \frac{1}{b}a^2$ for all $a, b \in \mathbb{R}^+$, we have

$$T_2 \leqslant 2^{2Bx_2} \sum_{k \leqslant z^4} E(k) + 2^{-2Bx_2} \sum_{k \leqslant z^4} E(k) 2^{4\omega(k)} = U_1 + U_2, \tag{3.11}$$

say. Using Lemmas 3 and 2 in order to estimate U_1 and U_2, respectively, it follows that (3.11) can be replaced by

$$T_2 \leqslant \frac{x}{(\log x)^{(A')(2B \log 2)}} + \frac{x}{\log x} (\log x)^{-2B \log 2} \sum_{k \leqslant z^4} \frac{2^{4\omega(k)}}{\varphi(k)}, \tag{3.12}$$

where B and A' are arbitrary positive constants. Hence, by an appropriate choice of B and A', it follows from (3.12) that

$$T_2 \ll \frac{x}{\log^C x}. \tag{3.13}$$

Then, using (3.10) and (3.13) in (3.9), placing the result in (3.4) and then summing the first term on the right hand side of (3.4) over m, we obtain from Lemma 2 that it is $\ll x/(\log^C x)$, thus completing the proof of Lemma 4. ■

Lemma 5. *Given positive integers k and A, set*

$$B_k(x, A) = \sum_{\substack{m_1 \leqslant Z_x^{K_x} \\ \omega(m_1) = k \\ p(m_1) > w_x,\ P(m_1) \leqslant Z_x}} \mathcal{A}(x|Am_1).$$

Let $\wp_0, \ldots, \wp_{q-1}$ be a disjoint classification of primes with corresponding densities $\delta_0, \ldots, \delta_{q-1}$. Then, given an arbitrary constant $C > 0$,

$$\sum_{\substack{A \leqslant w_x^{w_x} \\ P(A) \leqslant w_x}} \sum_{k \leqslant Bx_2} \sum_{i_1 \ldots i_k \in A_q^k} \left| \sum_{\substack{m_1 \leqslant Z_x^{K_x} \\ H(m_1) = i_1 \ldots i_k \\ p(m_1) > w_x,\ P(m_1) \leqslant Z_x}} \mathcal{A}(x|Am_1) - \delta_{i_1} \cdots \delta_{i_k} B_k(x, A) \right| \ll \frac{x}{\log^C x}.$$

Moreover,

$$\sum_{\substack{A \leqslant w_x^{w_x} \\ P(A) \leqslant w_x}} \sum_{k \leqslant Bx_2} \left| B_k(x, A) - \nu(Q)\, li(x) \frac{\kappa(A)}{A} \sum_{\substack{m_1 \leqslant Z_x^{K_x} \\ \omega(m_1) = k \\ p(m_1) > w_x,\ P(m_1) \leqslant Z_x}} \frac{\kappa(m_1)}{m_1} \right| \ll \frac{x}{\log^C x}.$$

Proof. The result is a direct consequence of Theorem A and Lemma 4. ∎

Lemma 6. *Given positive integers $h \geqslant 2k$,*

$$\sum_{\alpha \in A_q^h} \left(F_\beta(\alpha) - \frac{h}{q^k} \right)^2 \leqslant c \frac{hkq^h}{q^k}, \tag{3.14}$$

where c is some absolute constant.

Proof. On the one hand, we have

$$\sum\nolimits_1 := \sum_{\alpha \in A_q^h} F_\beta(\alpha) = \#\{(\gamma_1, \gamma_2) : \alpha = \gamma_1 \beta \gamma_2\} = \sum_{\ell=0}^{h-k} q^\ell q^{h-\ell-k} = q^{h-k}(h-k+1), \tag{3.15}$$

while on the other hand

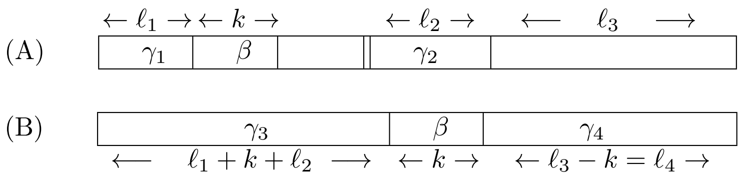
$$\sum\nolimits_2 := \sum_{\alpha \in A_q^h} F_\beta^2(\alpha) = \#\{(\gamma_1, \gamma_2, \gamma_3, \gamma_4) : \alpha = \gamma_1 \beta \gamma_2 = \gamma_3 \beta \gamma_4\}. \tag{3.16}$$

Now, write

$$\sum\nolimits_2 = \sum\nolimits_{2,0} + \sum\nolimits_{2,1} + \sum\nolimits_{2,2},$$

where in $\sum_{2,0}$, we impose the condition $\lambda(\gamma_1) = \lambda(\gamma_3)$, in $\sum_{2,1}$, we impose the condition $\lambda(\gamma_1) > \lambda(\gamma_3)$, and finally in $\sum_{2,2}$, we are restricted to $\lambda(\gamma_1) < \lambda(\gamma_3)$. In $\sum_{2,0}$, we have $\gamma_1 = \gamma_3$, so that $\sum_{2,0} = \sum_1$.

Let $\sum_{2,1,1}$ be the number of those γ_1, γ_3 for which $\lambda(\gamma_3) \leqslant \lambda(\gamma_1) + k$, and $\sum_{2,1,2}$ be the number of those γ_1, γ_3 for which $\lambda(\gamma_3) > \lambda(\gamma_1) + k$. Since γ_3 is a prefix of $\gamma_1\beta$, it follows that it has no more than k distinct values for a fixed γ_1, and therefore that $\sum_{2,1,1} \leqslant k \sum_1$. Assume now that $\lambda(\gamma_3) > \lambda(\gamma_1) + k$. Thus



Let us fix the position of β in (A) and in (B), that is the lengths ℓ_1 and ℓ_2. Then $\ell_1 + \ell_2 + \ell_4$ digits can be distributed freely, which yields $q^{\ell_1+\ell_2+\ell_4} = q^{h-2k}$ integers. Hence the number of those nonnegative integers ℓ_1, ℓ_2, ℓ_4 for which $\ell_1 + \ell_2 + \ell_4 = h - 2k$ is therefore equal to

$$\sum_{\ell_4=0}^{h-2k} (h - 2k - \ell_4 + 1) = \sum_{\nu=1}^{h-2k} \nu = \frac{(h-2k)(h-2k+1)}{2}.$$

Thus

$$\sum\nolimits_{2,1,2} = \frac{(h-2k)(h-2k+1)}{2q^{2k}} q^h = \frac{h^2 q^h}{2q^{2k}} + O\left(\frac{khq^h}{q^{2k}} \right),$$

so that (3.16) can be written as

$$\sum\nolimits_2 = \frac{h^2 q^h}{q^{2k}} + O\left(\frac{khq^h}{q^{2k}}\right), \tag{3.17}$$

Therefore, using (3.15) and (3.17), inequality (3.14) follows, thus completing the proof of Lemma 6. ■

4. Proof of Theorem 1

Let $\wp^* = \bigcup_{j=0}^{q-1} \wp_j$ and define

$$\omega_{\wp^*}(n) := \sum_{\substack{p|n \\ p\in\wp^*}} 1.$$

For each real number $u \geqslant 2$, let us set

$$\rho_u := T([u]+1)\dots T([2u]).$$

It is clear that

$$\lambda(\rho_u) = u \sum_{\substack{p\leqslant 2u \\ p\in\wp^*}} \frac{1}{p} + O(u) = q\delta u \log\log u + O(u). \tag{4.1}$$

Now let k be a fixed positive integer and consider the word $\beta = i_1 \dots i_k \in A_q^k$. We now let $F_\beta(\alpha)$ stand for the number of occurrences of the word β within the word α, that is we set

$$F_\beta(\alpha) := \#\{(\gamma_1, \gamma_2) : \alpha = \gamma_1\beta\gamma_2,\ \gamma_1, \gamma_2 \in A_q^*\}.$$

We shall prove that

$$\max_{\beta\in A_q^k} \left| F_\beta(\rho_u) - \frac{\lambda(\rho_u)}{q^k} \right| \leqslant \varepsilon(u)\lambda(\rho_u), \tag{4.2}$$

where $\varepsilon(u)$ tends to 0 monotonically as $u \to \infty$.

Once we will have proven (4.2), Theorem 1 will follow. Indeed, let ξ_N stand for the q-ary expansion of ξ up to the N-th digit. Now, given N, let u be a real number which satisfies the inequalities

$$N_1 := \sum_{j\leqslant 2u} \omega_{\wp^*}(j) \leqslant N < \sum_{j\leqslant 2u+1} \omega_{\wp^*}(j).$$

Let us further set $\xi_{N_1} := T(1)T(2)\dots T([2u])$. With this definition, we have that

$$0 \leqslant \lambda(\xi_N) - \lambda(\xi_{N_1}) = O(\log N). \tag{4.3}$$

Now, given an arbitrary positive integer ℓ satisfying $2^\ell < u$, let us write

$$\xi_{N_1} = \chi^{(\ell)}\ \rho_{u/2^\ell}\ \rho_{u/2^{\ell-1}}\ \ldots\ \rho_u,$$

where

$$\rho_v := T([v]+1)\ldots T([2v]).$$

It follows that

$$F_\beta(\xi_{N_1}) = F_\beta(\chi^{(\ell)}) + F_\beta(\rho_{u/2^\ell}) + \cdots + F_\beta(\rho_u) + O(\ell+1).$$

Hence, using (4.2) and (4.3), we obtain that

$$F_\beta(\xi_N) = F_\beta(\xi_{N_1}) + O(\log N) = \frac{\lambda(\xi_N)}{q^k} + O\left(\varepsilon(u/2^\ell)N + \lambda(\chi^{(\ell)})\right). \tag{4.4}$$

Now, choosing ℓ to be the unique integer satisfying $2^\ell \leqslant \sqrt{u} < 2^{\ell+1}$ and using the fact that $\lambda(\chi^{(\ell)})/N \to 0$ as $N \to \infty$, we then obtain from (4.4) that

$$\frac{F_\beta(\xi_N)}{N} \to \frac{1}{q^k} \qquad \text{as } N \to \infty, \tag{4.5}$$

thus proving that ξ is a q-normal number.

Thus, it remains to prove (4.2). Doing so, we will make repetitive use of (4.1). First we set $w_u = \log\log\log u$ and $Z_u = \exp\{(\log u)^{1-\varepsilon_u}\}$, where $\varepsilon_u \to 0$ as $u \to \infty$, and write each integer $n \geqslant 2$ as

$$n = \prod_{\substack{p^a \parallel n \\ p \leqslant w_u}} p^a \cdot \prod_{\substack{p^a \parallel n \\ w_u < p \leqslant Z_u}} p^a \cdot \prod_{\substack{p^a \parallel n \\ p > Z_u}} p^a = A(n)\cdot B(n)\cdot C(n),$$

say. Since

$$\sum_{u \leqslant n \leqslant 2u} \omega(A(n)) + \sum_{u \leqslant n \leqslant 2u} \omega(C(n)) = o(u\log\log u) \qquad (u \to \infty),$$

it follows that

$$F_\beta(\rho_u) = \sum_{u \leqslant n \leqslant 2u} F_\beta(T(B(n))) + o(u\log\log u) \qquad (u \to \infty). \tag{4.6}$$

Let $\mathcal{M}_u$ be the set of those positive integers m for which there exists at least one integer $n \in [u, 2u]$ such that $B(n) = m$, in which case we let

$$D(m) = \#\{n \in [u, 2u] : B(n) = m\}.$$

Then, from (4.6), we have

$$F_\beta(\rho_u) = \sum_{m \in \mathcal{M}_u} F_\beta(T(m))D(m) + o(u\log\log u) \qquad (u \to \infty). \tag{4.7}$$

Further define $\mathcal{M}_u^{(1)}$ as the set of those $m \in \mathcal{M}_u$ for which at least one of the following conditions holds:

(1) m is not squarefree,
(2) $m \geqslant Z_u^{K_u}$, $K_u = (\log u)^{\varepsilon_u/2}$,
(3) there exist $p_1|m$ and $p_2|m$ such that $p_1 < p_2 < 2p_1$,
(4) $|\omega(m) - \log\log u| > (\log\log u)^{3/4}$.

Let $\mathcal{M}_u^{(0)} = \mathcal{M}_u \setminus \mathcal{M}_u^{(1)}$. Observing that $F_\beta(T(m)) \leqslant \omega(m)$, we easily obtain that

$$\sum_{m \in \mathcal{M}_u^{(1)}} F_\beta(T(m))D(m) = o(u \log\log u) \qquad (u \to \infty). \tag{4.8}$$

By a standard sieve argument, we easily get that, as $u \to \infty$,

$$D(m) = (1 + o(1))\frac{u}{m} \prod_{w_u \leqslant p \leqslant Z_u} \left(1 - \frac{1}{p}\right) = (1 + o(1))\frac{u}{m}\frac{\log w_u}{\log Z_u} \qquad (m \in \mathcal{M}_u^{(0)}). \tag{4.9}$$

Thus, using (4.8) and (4.9) in (4.7), we obtain

$$F_\beta(\rho_u) = (1 + o(1))u\frac{\log w_u}{\log Z_u} \sum_{m \in \mathcal{M}_u^{(0)}} \frac{F_\beta(T(m))}{m} + o(u \log\log u) \qquad (u \to \infty).$$

Hence, it remains to prove that, given arbitrary words β_1 and β_2 belonging to A_q^k,

$$\sum_{m \in \mathcal{M}_u^{(0)}} \frac{F_{\beta_1}(T(m))}{m} = (1 + o(1)) \sum_{m \in \mathcal{M}_u^{(0)}} \frac{F_{\beta_2}(T(m))}{m} \qquad (u \to \infty). \tag{4.10}$$

We shall now use a technique we have already used to prove Theorem 1 of our 1995 paper [1]. We define the sequence $t_0 < t_1 < \cdots$ as follows:

$$t_0 = w_u, \qquad t_{j+1} = t_j + \frac{t_j}{(\log t_j)^5} \qquad \text{for} \quad j = 0, 1, \ldots.$$

Let r be defined implicitly by $t_r \leqslant Z_u < t_{r+1}$ and set $I_j = [t_j, t_{j+1})$ for each integer $j \geqslant 0$.

Let h be fixed, $|h - \log\log u| \leqslant (\log\log u)^{3/4}$, $0 \leqslant j_1 < j_2 < \cdots < j_h \leqslant r - 1$ with $j_{\ell+1} \geqslant 2j_\ell$. Further define $\mathcal{M}_u^{(0)}(j_1, \ldots, j_h)$ as the set of those $m = \pi_1\pi_2 \cdots \pi_h$ for which $\pi_j \in I_{\ell_j}$ for $j = 1, \ldots, h$.

Observe that any $m \in \mathcal{M}_u^{(0)}(j_1, \ldots, j_h)$ satisfies

$$\ell_{j_1+1} \cdot \ell_{j_2+1} \cdots \ell_{j_h+1} \geqslant m \geqslant \ell_{j_1} \cdot \ell_{j_2} \cdots \ell_{j_h}$$

and that

$$\begin{aligned}1 \leqslant \frac{\ell_{j_1+1}\cdot\ell_{j_2+1}\cdots\ell_{j_h+1}}{\ell_{j_1}\cdot\ell_{j_2}\cdots\ell_{j_h}} &\leqslant \prod_{j=1}^{h}\left(1+\frac{1}{(\log\ell_j)^5}\right)\\ &\leqslant \exp\left\{\sum_{j=1}^{h}\frac{1}{(\log\ell_j)^5}\right\} \leqslant \exp\left\{\sum_{j=0}^{h-1}\frac{1}{(\log w_u + j\log 2)^5}\right\}\\ &= 1+o(1) \qquad (u\to\infty).\end{aligned}$$

This means that instead of proving (4.10), we only need to prove

$$\sum_{m\in\mathcal{M}_u^{(0)}(\ell_{j_1},\dots,\ell_{j_h})}\frac{F_{\beta_1}(T(m))}{m} = (1+o(1))\sum_{m\in\mathcal{M}_u^{(0)}(\ell_{j_1},\dots,\ell_{j_h})}\frac{F_{\beta_2}(T(m))}{m} \qquad (u\to\infty). \tag{4.11}$$

Now let $\mathcal{M}_u^{(0)}(\ell_{j_1},\dots,\ell_{j_h}|\wp_{\nu_1},\dots,\wp_{\nu_h})$ be the set of those $m=\pi_1\pi_2\cdots\pi_h\in\mathcal{M}_u^{(0)}(\ell_{j_1},\dots,\ell_{j_h})$ for which $\pi_\ell\in\wp_{\nu_\ell}$.

Then, repeating the computation done in [1], we obtain that

$$\frac{\#\mathcal{M}_u^{(0)}(\ell_{j_1},\dots,\ell_{j_h}|\wp_{\nu_1},\dots,\wp_{\nu_h})}{\#\mathcal{M}_u^{(0)}(\ell_{j_1},\dots,\ell_{j_h})} = (1+o(1))\tau(\nu_1)\cdots\tau(\nu_h) \qquad (u\to\infty), \tag{4.12}$$

where $\tau(\nu)=\delta$ if $\nu\in\{0,1,\dots,q-1\}$ and $\tau(q)=1-q\delta$. Assume that among $\nu_1,\dots,\nu_h$, the value q occurs t_1 times. Then, on the right hand side of (4.12), $\tau(\nu_1)\cdots\tau(\nu_h)=(1-q\delta)^{t_1}\cdot\delta^{h-t_1}$, which does depend only on t_1. It is clear that $F_\beta(T(m))$ is constant in every $\mathcal{M}_u^{(0)}(\ell_{j_1},\dots,\ell_{j_h}|\wp_{\nu_1},\dots,\wp_{\nu_h})$. So, let $e_1<\cdots<e_{t_1}\leqslant h$ be arbitrary integers and consider those $\wp_{\nu_1},\dots,\wp_{\nu_h}$ for which $\nu_{e_j}=q$ for $j=1,\dots,t_1$ and $\nu_\ell\neq q$ if $\ell\neq e_j$. Let $v_0<v_1<\cdots<v_{h-t_1-1}$ be the sequence of integers defined by

$$\{v_0,\dots,v_{h-t_1-1}\}=\{1,\dots,h\}\setminus\{e_1,\dots,e_{t_1}\}.$$

Moreover, let $\nu_{v_j}\in\{0,1,\dots,q-1\}$ for $j=0,1,\dots,h-t_1-1$ be arbitrary digits. If $m\in\mathcal{M}_u^{(0)}(\ell_{j_1},\dots,\ell_{j_h}|\wp_{\nu_1},\dots,\wp_{\nu_h})$, then

$$F_\beta(T(m)) = F_\beta(\nu_{v_0}\nu_{v_1}\dots\nu_{v_{h-t_1-1}}). \tag{4.13}$$

Now, one can easily show that the number of those $n\in[u,2u]$ for which $h-t_1\leqslant k^2$ is $o(u)$. Hence, we may assume that $h-t_1>k^2$. Then, in light of (4.12), (4.13) and Lemma 6, we easily obtain (4.11) and thereby (4.2) and (4.5), thus completing the proof of Theorem 1.

5. Proof of Theorem 2

To prove Theorem 2, we use the notations of Theorem 1 and essentially the same kind of technique.

Let $u \geqslant 3$ and set

$$\kappa_u := T(p_1 - 1) \dots T(p_k - 1),$$

where $u \leqslant p_1 < \cdots < p_k < 2u$ are all the primes included in the interval $[u, 2u)$.

Now,

$$\lambda(\kappa_u) = \sum_{u \leqslant p_\ell < 2u} \omega_{\wp^*}(p_\ell - 1) = q\delta \cdot \frac{u}{\log u} \cdot \log\log u + o\left(\frac{u \log\log u}{\log u}\right), \tag{5.1}$$

an estimate that follows essentially from the fact that $p - 1$ has no more than 3 prime divisors $\pi > u^{1/3}$ if $u \leqslant p < 2u$. Since equation (5.1) provides the exact size of $\lambda(\kappa_u)$, it will be used repetitively below.

Then, choose the intervals I_j as we did in the proof of Theorem 1 and let us estimate

$$W(u) := \sum_{\substack{u \leqslant p_\ell \leqslant 2u \\ \pi_1 | p_\ell - 1, \pi_2 | p_\ell - 1 \\ \pi_1, \pi_2 \in I_j \text{ for some } j}} \omega(p_\ell - 1).$$

We have

$$W(u) = \sum_{w_u < \pi_1 < \pi_2 < 2\pi_1 < Z_u} \sum_{\substack{p \in (u, 2u) \\ p - 1 \equiv 0 \pmod{\pi_1 \pi_2}}} \omega(p - 1). \tag{5.2}$$

Since $\pi_1 \pi_2 | p - 1$, we have from (5.2) that

$$W(u) \leqslant \sum_{w_u < \pi_1 < \pi_2 < 2\pi_1 < Z_u} \sum_{\pi_0 < u^{1/4}} \pi(2u; \pi_0 \pi_1 \pi_2, 1) + \pi(2u; \pi_1 \pi_2, 1). \tag{5.3}$$

Then, using the Brun-Titchmarsh inequality (Lemma 2), it is clear that the right hand side of (5.3) is less than

$$c \frac{u}{\log u} \sum_{\pi_0 < u^{1/4}} \frac{1}{\pi_0} \sum_{w_u < \pi_1 < \pi_2 < 2\pi_1 < Z_u} \frac{1}{(\pi_1 - 1)(\pi_2 - 1)} \\ + c \frac{u}{\log u} \sum_{w_u < \pi_1 < \pi_2 < 2\pi_1 < Z_u} \frac{1}{(\pi_1 - 1)(\pi_2 - 1)}. \tag{5.4}$$

On the other hand, one can easily establish that

$$\sum_{\pi_1 < \pi_2 < 2\pi_1} \frac{1}{(\pi_1 - 1)(\pi_2 - 1)} \leqslant c \sum_p \frac{1}{(p - 1) \log p} < \infty, \tag{5.5}$$

while it is clear that

$$\sum_{\pi_0 < u^{1/4}} \frac{1}{\pi_0} \ll \log\log u. \tag{5.6}$$

Using estimates (5.5) and (5.6) in (5.4), it follows that (5.3) can be replaced by

$$W(u) = o\left(\frac{u \log\log u}{\log u}\right). \tag{5.7}$$

Now, let $\beta = i_1 \dots i_k \in A_q^k$. We will now estimate $F_\beta(\kappa_u)$. We have

$$F_\beta(\kappa_u) = \sum_{j=1}^{k} F_\beta(T(p_j-1)) + O(\pi([u,2u])). \tag{5.8}$$

Since

$$\sum_{u \leqslant p \leqslant 2u} \sum_{\substack{\pi_0 | p-1 \\ \pi_0 \notin (w_u, Z_u)}} 1 = o(\mathrm{li}(u) \log\log u) \qquad (u \to \infty),$$

relation (5.8) becomes

$$F_\beta(\kappa_u) = \sum_{j=1}^{k} F_\beta(T(A(p_j-1)B(p_j-1))) + o(\mathrm{li}(u)\log\log u) \qquad (u \to \infty). \tag{5.9}$$

Let $\mathcal{A}([u,2u]|m) = \mathcal{A}(2u|m) - \mathcal{A}(u|m)$. Using the notation of Lemma 5, it follows that $A(p_j-1)B(p_j-1) = m$ holds for $\mathcal{A}([u,2u]|m)$ numbers, so that

$$\sum_{\substack{m \leqslant Z_u^{K_u} \\ P(m) \leqslant Z_u}} \left| \mathcal{A}([u,2u]|m) - \frac{\nu(Q)\kappa(m)}{m}(\mathrm{li}(2u) - \mathrm{li}(u)) \right| \ll \frac{u}{\log^C u},$$

where C is an arbitrary positive number. Using this in (5.9), it follows that, with $\nu(Q) = \prod_{2<p\leqslant Z_u}\left(1 - \frac{1}{p-1}\right)$ and with the strongly multiplicative function κ defined in (3.1),

$$F_\beta(\kappa_u) = \sideset{}{^*}\sum_{\substack{m \leqslant Z_u^{K_u} \\ P(m) \leqslant Z_u}} \frac{\nu(Q)\kappa(m)}{m}(\mathrm{li}(2u) - \mathrm{li}(u))F_\beta(T(m)) + o\left(\mathrm{li}(u)\log\log u\right), \tag{5.10}$$

where the star $(*)$ on the sum indicates that, in light of (5.7), we omitted those m's for which there exist two primes π_1 and π_2 such that $\pi_1\pi_2|m$, $w_u < \pi_1 < \pi_2 \leqslant Z_u$ with $\pi_1, \pi_2 \in I_j$ for some j. Note that in this same sum, we dropped those integers $m > Z_u^{K_u}$ since

$$\sum_{\substack{m > Z_u^{K_u} \\ P(m) \leqslant Z_u}} \mathcal{A}([u,2u]|m)\omega(m) = o\left(\frac{u}{\log u}\log\log u\right).$$

Let us write

$$m = \prod_{\substack{p^a \| m \\ p \leqslant w_u}} p^a \cdot \prod_{\substack{p^a \| m \\ w_u < p \leqslant Z_u}} p^a = A m_1,$$

say. Since $\omega(A) \leqslant w_u = O(\log\log\log u)$, it follows that estimate (5.10) becomes

$$F_\beta(\kappa_u) = \nu(Q)(\mathrm{li}(2u) - \mathrm{li}(u)) \sideset{}{^+}\sum_{A} \frac{\kappa(A)}{A} \sideset{}{^*}\sum_{m_1} \frac{\kappa(m_1)}{m_1} F_\beta(T(m_1)) + o(\mathrm{li}(u)\ \log\log u), \tag{5.11}$$

where the (+) sign on the sum indicates that we dropped those integers A which are large, namely those for which $A \geqslant w_u^{w_u}$, say.

Now, since

$$\begin{aligned} 1 \leqslant \kappa(m_1) &= \prod_{p\mid m_1}\left(1+\frac{1}{p}\right)\prod_{p\mid m_1}\frac{p^2-p}{p^2-p-2} \\ &= \exp\left(\sum_{p\mid m_1}\log\left(1+\frac{2}{p^2-p-2}\right)\right)\prod_{p\mid m_1}\left(1+\frac{1}{p}\right) \\ &\leqslant \exp\left(3\sum_{p>w_u}\frac{1}{p^2}\right)\prod_{p\mid m_1}\left(1+\frac{1}{p}\right) \\ &\leqslant \left(1+\frac{1}{w_u}\right)\sum_{d\mid m_1}\frac{|\mu(d)|}{d}, \end{aligned}$$

we obtain that

$$0 \leqslant \kappa(m_1)-1 \leqslant \sum_{\substack{d\mid m_1\\ d>1}}\frac{|\mu(d)|}{d}+\frac{1}{w_u}\sum_{d\mid m_1}\frac{|\mu(d)|}{d}.$$

Hence, using this last estimate in (5.11), we obtain

$$F_\beta(\kappa_u) = \nu(Q)\frac{u}{\log u}\sum_A\frac{\kappa(A)}{A}\sum_{m_1}{}^{*}\frac{1}{m_1}F_\beta(T(m_1)) + o(\mathrm{li}(u)\,\log\log u).$$

Proceeding as we did in the proof of Theorem 1, we then obtain that

$$\sum_{m_1}{}^{*}\frac{1}{m_1}F_{\beta_1}(T(m_1)) = (1+o(1))\sum_{m_1}{}^{*}\frac{1}{m_1}F_{\beta_2}(T(m_1)) \qquad (u\to\infty),$$

and therefore that

$$F_{\beta_1}(\kappa_u) = F_{\beta_2}(\kappa_u) + o\left(\frac{u}{\log u}\log\log u\right) \qquad (u\to\infty),$$

for every words β_1 and β_2 belonging to A_q^k. This observation concludes the proof of Theorem 2.

6. The proof of Theorem 3

We first introduce the multiplicative function ρ defined on prime powers by

$$\rho(p^a) = \begin{cases} 1 & \text{if } p=2 \text{ and } a=1,\\ 0 & \text{if } p=2 \text{ and } a\geqslant 2,\\ 2 & \text{if } p\equiv 1 \pmod 4,\\ 0 & \text{if } p\equiv 3 \pmod 4. \end{cases}$$

First observe that, given an integer $D > 1$, the congruence $n^2 + 1 \equiv 0 \pmod{D}$ has $\rho(D)$ distinct solutions.

As in the proof of Theorem 1, let $Z_u = \exp\{(\log u)^{1-\varepsilon_u}\}$ and, for each integer $n \geqslant 2$, set

$$E(n|Z_u) = \prod_{\substack{p^a \| n \\ p \leqslant Z_u}} p^a.$$

Further define

$$\mathcal{A}([u,2u]|m) := \#\{n \in [u,2u] : E(n^2+1|Z_u) = m\}$$

and let K_u be a function tending to infinity with u and that we will determine later.

Our first goal will be to find an upper bound for

$$S := \sum_{\substack{m > Z_u^{K_u} \\ P(m) \leqslant Z_u}} \mathcal{A}([u,2u]|m)\omega(m).$$

Letting $\omega_Y(n) := \sum_{p|n,\ p \leqslant Y} 1$, we have that

$$\sum_{n \in [u,2u]} \omega_{Z_u}(n^2+1)^2 \ll u \sum_{p_1,p_2 \leqslant Z_u} \frac{\rho(p_1)}{p_1}\frac{\rho(p_2)}{p_2} + u \sum_{p \leqslant Z_u} \frac{\rho(p)}{p}$$
$$\ll u(\log\log Z_u)^2 \ll u(\log\log u)^2.$$

Hence, it follows from this that, setting

$$S_1 := \sum_{\substack{m > Z_u^{K_u} \\ P(m) \leqslant Z_u}} \mathcal{A}([u,2u]|m), \tag{6.1}$$

we have, letting $E(u)$ be a function which tends to infinity with u and that we will determine later,

$$\begin{aligned} S &= \sum_{\substack{m > Z_u^{K_u} \\ P(m) \leqslant Z_u \\ \omega_{Z_u}(m) < E(u)\log\log u}} \mathcal{A}([u,2u]|m)\omega(m) + \sum_{\substack{m > Z_u^{K_u} \\ P(m) \leqslant Z_u \\ \omega_{Z_u}(m) \geqslant E(u)\log\log u}} \mathcal{A}([u,2u]|m)\omega(m) \\ &< E(u)(\log\log u)S_1 + \sum_{\substack{n \in [u,2u] \\ \omega_{Z_u}(n^2+1) \geqslant E(u)\log\log u}} \omega_{Z_u}(n^2+1) \\ &\leqslant E(u)(\log\log u)S_1 + \frac{1}{E(u)\log\log u} \sum_{n \in [u,2u]} \omega_{Z_u}(n^2+1)^2 \\ &\ll E(u)(\log\log u)S_1 + \frac{1}{E(u)} u \log\log u. \end{aligned} \tag{6.2}$$

Let us now bound S_1. Clearly, if $n \in [u, 2u]$ is counted on the right hand side of (6.1), then the corresponding integer m for which $E(n^2+1|Z_u) = m > Z_u^{K_u}$ has a divisor m_1 satisfying $Z_u^{K_u-1} \leqslant m_1 \leqslant Z_u^{K_u}$.

Now consider the subintervals $J_\ell := [2^\ell Z_u^{K_u-1}, 2^{\ell+1} Z_u^{K_u-1}]$ with $\ell = 0, 1, \ldots, \ell_0$, where ℓ_0 is the unique positive integer such that $2^{\ell_0-1} \leqslant Z_u < 2^{\ell_0}$, so that

$$[Z_u^{K_u-1}, Z_u^{K_u}] \subset \bigcup_{\ell=0}^{\ell_0} J_\ell.$$

Now, for a given integer $n \in [u, 2u]$, if the corresponding integers m and m_1 mentioned above are such that m_1 is the minimal divisor of m and thus of n^2+1, with $m_1 \in J_\ell$ for some $\ell \geqslant 0$, then $p(n^2+1) > 2^\ell$. It follows from this that

$$S_1 \leqslant \sum_{\ell=0}^{\ell_0} T_\ell, \tag{6.3}$$

where

$$T_\ell \leqslant \sum_{\substack{m_1 \in J_\ell \\ P(m_1) \leqslant Z_u \\ p(m_1) > 2^\ell}} \#\{n \in [u, 2u] : m_1 | n^2+1,\ (n^2+1, \pi_0) = 1 \text{ for all primes } \pi_0 < 2^\ell\}.$$

Now, using the Selberg Sieve, it follows that

$$\begin{aligned} T_\ell &\leqslant u \sum_{\substack{m_1 \in J_\ell \\ P(m_1) \leqslant Z_u \\ p(m_1) > 2^\ell}} \frac{\rho(m_1)}{m_1} \prod_{\pi_0 < 2^\ell} \left(1 - \frac{\rho(\pi_0)}{\pi_0}\right) \ll \frac{u}{\ell} \sum_{\substack{m_1 \in J_\ell \\ P(m_1) \leqslant Z_u \\ p(m_1) > 2^\ell}} \frac{\rho(m_1)}{m_1} \\ &\leqslant \frac{u}{\ell} \frac{1}{2^\ell Z_u^{K_u-1}} \sum_{\substack{m_1 \in J_\ell \\ P(m_1) \leqslant Z_u \\ p(m_1) > 2^\ell}} \rho(m_1) = \frac{u}{\ell} \frac{1}{2^\ell Z_u^{K_u-1}} H_\ell, \end{aligned} \tag{6.4}$$

say.

Again, using a sieving technique, we get that

$$H_\ell \ll \prod_{p < 2^\ell} \left(1 - \frac{\rho(p)}{p}\right) \prod_{Z_u < p < Z_u^{K_u}} \left(1 - \frac{\rho(p)}{p}\right) \sum_{a \in J_\ell} \rho(a) \ll \frac{1}{\ell K_u} 2^\ell Z_u^{K_u-1}. \tag{6.5}$$

Using (6.5) in (6.4), we obtain that

$$T_\ell \ll \frac{u}{\ell^2 K_u}.$$

Substituting this back in (6.3), we obtain that

$$S_1 \ll \frac{u}{K_u}.$$

Using this in (6.2) and choosing $E(u)$ and K_u in such a manner that $E(u)/K_u \to 0$ as $u \to \infty$, we finally obtain that

$$S = o\,(u \log\log u) \qquad (u \to \infty). \tag{6.6}$$

Now assume that $m \leqslant Z_u^{K_u}$. Then, in light of (6.6) and using Lemma 1, we obtain that

$$\mathcal{A}([u, 2u]|m) = (1 + o(1))u \frac{\rho(m)}{m} \prod_{\pi_0 < Z_u} \left(1 - \frac{\rho(\pi_0)}{\pi_0}\right) \qquad (u \to \infty). \tag{6.7}$$

The rest of the proof of Theorem 3 then runs along the same lines as those of Theorems 1 and 2, in particular by using our 1995 results (see [1]).

Indeed, let

$$\tau_u := T(f([u] + 1)) \dots T(f([2u])).$$

By using (6.7) and our earlier estimates, we obtain that, in light of (6.6), as $u \to \infty$,

$$F_\beta(\tau_u) = (1 + o(1))u \prod_{\pi_0 < Z_u} \left(1 - \frac{\rho(\pi_0)}{\pi_0}\right) \sum_{\substack{m \leqslant Z_u^{K_u} \\ P(m) \leqslant Z_u}} \frac{\rho(m)}{m} F_\beta(T(m)) + o(u \log\log u). \tag{6.8}$$

Let $w_u = \log\log\log u$ and write m in the form $m = Am_1$, where $P(A) \leqslant w_u$, $p(m_1) > w_u$. As in our 1995 paper [1], we let ℓ_j, for $j = 0, 1, \dots$, be the sequence defined by

$$\ell_0 = w_u \qquad \text{and} \qquad \ell_{j+1} = \ell_j + \frac{\ell_j}{(\log \ell_j)^5}$$

and set $I_j := [\ell_j, \ell_{j+1}]$.

Now, on the right hand side of (6.8), drop those integers $m = Am_1$ for which one of the following five conditions holds:

1. m_1 is not squarefree,
2. $A > w_u^{w_u}$,
3. there exist two prime divisors q_1, q_2 of m_1 for which $q_1 < q_2 < 4q_1$,
4. $|\omega(m_1) - \log\log u| > \frac{1}{2} \log\log u$,
5. $\rho(m) = 0$.

It is easy to see that the whole contribution of the dropped elements located on the right hand side of (6.8) is $o(u \log\log u)$.

Let us now consider all the remaining m's. Let $m_1 = q_1 q_2 \cdots q_h$ and let $\ell_{i_1}, \dots, \ell_{i_h}$ be such indices for which $q_j \in I_{\ell_j}$ $(j = 1, \dots, h)$. Further let $\mathcal{K}(\ell_{i_1}, \dots, \ell_{i_h})$ be the set of those m_1's for which each factor $q_j | m_1$ belongs to I_{ℓ_j} for $j = 1, \dots, h$. We then have

$$\min m_1 \geqslant \prod_{j=1}^{h} \ell_{i_j} \qquad \text{and} \qquad \max m_1 \leqslant \left(\prod_{j=1}^{h} \ell_{i_j}\right) \exp\left\{\sum_{j=1}^{h} \frac{c}{(\log \ell_{i_j})^5}\right\}.$$

But since $\ell_{i_j} \geqslant 4^{j-1}\ell_0$, it follows that

$$\sum_{j=1}^{h} \frac{c}{(\log \ell_{i_j})^5} \leqslant \sum_{j=1}^{h} \frac{c}{(\log \ell_0 + (j-1)\log 4)^5} \leqslant \frac{c_1}{(\log \ell_0)^5},$$

and therefore $m_1 \in \mathcal{K}(\ell_{i_1}, \ldots, \ell_{i_h})$ implies that

$$\frac{\rho(m_1)}{m_1} = (1+o(1))2^h \prod_{j=1}^{h} \ell_{i_j} \qquad (u \to \infty),$$

implying that, for those $m_1 \in \mathcal{K}(\ell_{i_1}, \ldots, \ell_{i_h})$, we have that $\frac{\rho(m_1)}{m_1}$ tends to a constant (independent of m_1) as $u \to \infty$.

Now, the number of integers $m_1 \in \mathcal{K}(\ell_{i_1}, \ldots, \ell_{i_h})$ is equal to

$$D := \prod_{j=1}^{h} \pi(I_{\ell_{i_j}} \cap \widetilde{\wp}). \tag{6.9}$$

Let us set $\mathcal{R}_0 = \widetilde{\wp} \setminus \bigcup_{j=0}^{q-1} \wp_i$. On the one hand, it is obvious that, in light of (2.1),

$$\begin{aligned} \pi([u,v] \cap \mathcal{R}_0) &= \pi(u+v; 4, 1) - \pi(u; 4, 1) - \sum_{j=0}^{q-1} \pi([u,v] \cap \wp_j) \\ &= \left(\frac{1}{2} - q\delta\right) \pi([u,v]) + O\left(\frac{u}{\log^{c_1} u}\right), \end{aligned} \tag{6.10}$$

uniformly for $2 \leqslant v \leqslant u$.

We shall now subdivide the sequence $\{i_1, \ldots, i_h\}$ into two monotone subsequences $\{u_1, \ldots, u_t\}$ and $\{v_1, \ldots, v_{h-t}\}$ in the following way. Count those elements of $\mathcal{K}(\ell_{i_1}, \ldots, \ell_{i_h})$ in which $q_{u_j} \in \mathcal{R}_0$ and $q_{v_\nu} \in \wp_{e_\nu}$ for $e_\nu \in \{0, \ldots, q-1\}$ $(\nu = 1, \ldots, h-t)$, and denote this number by $\mathcal{D}(u_1, \ldots, u_t | e_1, \ldots, e_{h-t})$. Then, in light of the definition of D given in (6.9) and of estimate (6.10), it follows that, uniformly for $0 \leqslant t \leqslant h$,

$$\frac{\mathcal{D}(u_1, \ldots, u_t | e_1, \ldots, e_{h-t})}{D} = (1+o(1)) \left(\frac{1}{2} - q\delta\right)^t \delta^{h-t} \qquad (h \to \infty).$$

Hence, we obtain that for every class $\mathcal{K}(\ell_{i_1}, \ldots, \ell_{i_h})$, we have that

$$\begin{aligned} &\sum_{m \in \mathcal{K}(\ell_{i_1}, \ldots, \ell_{i_h})} \frac{\rho(m_1)}{m_1} F_{\beta_1}(T(m_1)) \\ &\qquad = (1+o(1)) \sum_{m \in \mathcal{K}(\ell_{i_1}, \ldots, \ell_{i_h})} \frac{\rho(m_1)}{m_1} F_{\beta_2}(T(m_1)) \qquad (h \to \infty), \end{aligned}$$

which in turn implies that

$$F_{\beta_1}(\tau_u) = F_{\beta_2}(\tau_u) + o(u \log\log u) \qquad (u \to \infty)$$

for every words β_1 and β_2 belonging to A_q^k, which completes the proof of Theorem 3.

References

[1] J.M. De Koninck and I. Kátai, *On the distribution of subsets of primes in the prime factorization of integers*, Acta Arithmetica **72** (1995), no.2, 169–200.
[2] J.M. De Koninck and I. Kátai, *Construction of normal numbers by classified prime divisors of integers*, Functiones et Approximatio **45.2** (2011), 231–253.
[3] P.D.T.A. Elliott, *Probabilistic Number Theory I, Mean Value Theorems*, Springer-Verlag, Berlin, 1979.
[4] H.H. Halberstam and H.E. Richert, *Sieve Methods*, Academic Press, London, 1974.
[5] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, AMS Colloquium Publications, Vol. 53, 2004.

Addresses: Jean-Marie De Koninck: Département de mathématiques et de statistique, Université Laval, Québec, Québec G1V 0A6, Canada;
Imre Kátai: Computer Algebra Department, Eötvös Loránd University, 1117 Budapest, Pázmány Péter Sétány I/C, Hungary.

E-mail: jmdk@mat.ulaval.ca, katai@compalg.inf.elte.hu

Received: 9 January 2012