Arch. Math. 88 (2007), 524–536

0003/889X/060524-13, *published online* 2007-05-14
DOI 10.1007/s00013-007-1928-6

Archiv der Mathematik

On the index of composition of integers from various sets

J. M. De Koninck[1], I. Kátai[2] and M. V. Subbarao[3]

Abstract. Given an integer $n \geq 2$, let $\lambda(n) := (\log n)/(\log \gamma(n))$, where $\gamma(n) = \prod_{p|n} p$, stand for the index of composition of n, with $\lambda(1) = 1$. We study the distribution function of $(\lambda(n) - 1) \log n$ as n runs through particular sets of integers, such as the shifted primes, the values of a given irreducible cubic polynomial and the shifted powerful numbers.

Mathematics Subject Classification (2000). 11A25, 11N25, 11N37.

Keywords. Arithmetic function, distribution function, shifted primes, powerful numbers.

1. Introduction. Recently, De Koninck and Doyon [2] studied the global and local behaviour of the index of composition of an integer, namely the function $\lambda(n) := \dfrac{\log n}{\log \gamma(n)}$, where $\gamma(n)$ stands for the product of the distinct primes dividing n (for convenience, $\lambda(1) = \gamma(1) = 1$). In a sense, $\lambda(n)$ measures the level of compositeness of n. More recently, De Koninck and Kátai [3] extended the study of this function by establishing estimates for $\sum_{x \leq n \leq x+\sqrt{x}} \lambda(n)$, $\sum_{x \leq n \leq x+\sqrt{x}} 1/\lambda(n)$ and $\sum_{x \leq p \leq x+x^{2/3}} \lambda(p+1)$.

In this paper, we study the distribution function of $\eta(n) := (\lambda(n) - 1) \log n$ as n runs through particular sets of integers, such as the shifted primes, the values of a given irreducible cubic polynomial with positive leading coefficient and the shifted powerful numbers.

---

[1]Research supported in part by a grant from NSERC.
[2]Research supported by the Applied Number Theory Research Group of the Hungarian Academy of Science and by a grant from OTKA.
[3]† Professor M.V. Subbarao passed away on February 15, 2006.

2. Notations and preliminary results. Let $\mathbb{N}$ and $\mathbb{Z}$ stand respectively for the set of positive integers and the set of all integers. In what follows, the letters p and q (with or without subscript) always stand for prime numbers, while c and C stand for absolute positive constants, not necessarily the same at each occurrence. Moreover, given any integer $n \geq 2$, let $P(n)$ stand for the largest prime factor of n. As usual, ϕ stands for Euler's function and μ for the Möbius function. We let λ_* stand for the Liouville function defined by $\lambda_*(n) = (-1)^{\Omega(n)}$, where $\Omega(n)$ is the number of prime factors of n counted with their multiplicities. We shall also write $\omega(n)$ for the number of distinct prime factors of n, and the logarithmic integral of x as $\mathrm{li}(x) := \displaystyle\int_2^x \frac{dt}{\log t}$.

A positive integer n is said to be powerful if $p^2 | n$ whenever the prime number p divides n. Let $\mathcal{B}$ be the set of powerful numbers. For each $K \in \mathcal{B}$, we set

$$\mathcal{A}_K := \{n = K \cdot m : \gcd(K, m) = 1,\ \mu^2(m) = 1\} \quad \text{and}$$
$$\mathcal{A}_K(x) := \#\{n \leq x : n \in \mathcal{A}_K\}.$$

One can prove that, given any $K \in \mathcal{B}$,

$$\mathcal{A}_K(x) = \alpha(K)x + O\left(\sqrt{\frac{x}{K}}\rho(K)\right) \qquad (x \to \infty), \tag{1}$$

where

$$\alpha(K) = \frac{6}{\pi^2 K} \prod_{p|K} \left(1 + \frac{1}{p}\right)^{-1} \qquad \text{and} \qquad \rho(K) = \prod_{p|K} \left(1 + \frac{1}{\sqrt{p}}\right)$$

and where the constant implicit in $O(\dots)$ is absolute, that is does not depend on K.

Indeed, in order to show (1), one may proceed as follows. First, it is well known that

$$\mathcal{A}_1(x) = \sum_{n \leq x} |\mu(n)| = \frac{6}{\pi^2}x + O(\sqrt{x}). \tag{2}$$

On the other hand, since

$$\sum_{\substack{n=1 \\ (n,K)=1}}^{\infty} \frac{|\mu(n)|}{n^s} = \frac{\zeta(s)}{\zeta(2s)} \prod_{p|K} \frac{1}{1 + 1/p^s},$$

it follows that

$$\mathcal{A}_K(x) = \sum_{v \in D_K} \lambda_*(v) \mathcal{A}_1\left(\frac{x}{Kv}\right), \tag{3}$$

where D_K is the set of all those positive integers all of whose prime factors divide K. Now, using (2) in (3), we obtain that

$$\begin{aligned}\mathcal{A}_K(x) &= \sum_{\substack{v\le x/K\\ v\in D_K}} \lambda_*(v)\cdot\frac{6}{\pi^2}\frac{x}{Kv} + O\left(\sqrt{\frac{x}{K}}\sum_{v\in D_K}\frac{1}{\sqrt{v}}\right)\\ &= \frac{6}{\pi^2}\frac{x}{K}\sum_{v\in D_K}\frac{\lambda_*(v)}{v} + O\left(\frac{x}{K}\sum_{v\ge x/K}\frac{1}{v}\right) + O\left(\sqrt{\frac{x}{K}}\sum_{v\in D_K}\frac{1}{\sqrt{v}}\right)\\ &= \alpha(K)x + O\left(\sqrt{\frac{x}{K}}\rho(K)\right),\end{aligned}$$

thereby establishing (1).

Now, for each powerful number $K>1$, set

$$\kappa(K) = \log\left(\frac{K}{\gamma(K)}\right),$$

letting $\kappa(1)=0$. Then, given any $n\in\mathcal{A}_K$, it is clear that

$$\eta(n) = (\lambda(n)-1)\log n = \frac{\kappa(K)}{1-\frac{\kappa(K)}{\log n}}. \tag{4}$$

Lemma 1. *If $K, L\in\mathcal{B}$ and $\kappa(K)=\kappa(L)$, then $K=L$.*

Proof. By hypothesis, we have $K/\gamma(K) = L/\gamma(L)$. Hence, given a prime power $p^\beta \| K$ (with $\beta\ge 2$), we have $p^{\beta-1}\| K/\gamma(K)$, so that $p^{\beta-1}\| L/\gamma(L)$, which means that $p^\beta\| L$. Since this is true for any prime power, it follows that $K=L$.

Let us now reorder the elements of $\mathcal{B}$. We enumerate them as $K_1, K_2, \dots$ in such a way that $\kappa(K_1) < \kappa(K_2) < \dots$. In this manner, we clearly have that $\kappa(K_i)\to\infty$ as $i\to\infty$.

Let ξ be the random variable taking the values $\kappa(K_i)$ with probability $\alpha(K_i)$, that is $P(\xi = \kappa(K_i)) = \alpha(K_i)$. Let $F(y)$ be the distribution function of ξ. Then it is clear that $F(u)=F(v)$ if $\kappa(K_i) < u < v < \kappa(K_{i+1})$ and also that F is continuous for all real $y\notin\{\kappa(K_1),\kappa(K_2),\dots\}$. $\square$

3. The distribution function of $\eta(n)$ as n runs through the set positive integers.

Theorem 1. *For each point of continuity y of F,*

$$\lim_{x\to\infty}\frac{1}{x}\#\{n\le x : \eta(n) < y\} = F(y).$$

Proof. Let i be a fixed positive integer and let $y\in(\kappa(K_i),\kappa(K_{i+1}))$. Our goal is to estimate the size of the set of positive integers $n\le x$ such that $\eta(n)<y$. It follows from (4) that there exists an absolute constant c such that, if $n\in\mathcal{A}_K$,

$$\kappa(K) < \eta(n) < \kappa(K) + c\frac{\kappa^2(K)}{\log n}.$$

Hence, if $\eta(n) < y$, its powerful part K must satisfy $\kappa(K) \in \{\kappa(K_1), \ldots, \kappa(K_i)\}$, so that we may write that

$$\{n : \eta(n) < y\} \subseteq \bigcup_{j=1}^{i} \mathcal{A}_{K_j}.$$

Now, for a fixed K_j, let us consider the integers $n \in \mathcal{A}_{K_j}$. Since $\kappa(K_j) < y$ and $\dfrac{\kappa^2(K_j)}{\log n} \to 0$ provided $n \in \mathcal{A}_{K_j}$ and $n \to \infty$, it follows that for every $n \in \mathcal{A}_{K_j}$ with the exception of at most finitely many of them, we have that $\eta(n) < y$, thereby completing the proof of Theorem 1. □

4. The case of shifted primes. We now consider the case of the shifted primes $p - 1$.

Before we go any further, let us mention two important results concerning the counting function for the number of primes in an arithmetic progression, namely the function

$$\pi(x; D, \ell) := \#\{p \le x : p \equiv \ell \pmod{D}\}.$$

First, we shall be using the fact that it follows from the Siegel-Walfisz Theorem (see Walfisz [8]) that, for some positive constant c_1,

$$\pi(x; K\delta^2, 1) = \frac{\mathrm{li}(x)}{\phi(K\delta^2)} \left(1 + O\left(e^{-c_1\sqrt{\log x}}\right)\right) \tag{5}$$

uniformly for all δ for which $\delta^2 K \le (\log x)^c$, where the constant implicit in the error term is absolute. Also, we will be using the Brun-Titchmarsh inequality (see for instance Crandall and Pomerance [1, Theorem 1.4.7]), that is

$$\pi(x; D, \ell) \le C\frac{\mathrm{li}(x)}{\phi(D)}. \tag{6}$$

Now, for a fixed $K \in \mathcal{B}$, let

$$\mathcal{S}_K := \{p : p - 1 = Km,\ \gcd(K, m) = 1,\ \mu^2(m) = 1\} \quad \text{and}$$
$$\mathcal{S}_K(x) := \#\{p \le x : p \in \mathcal{S}_K\}.$$

We shall now estimate the size of $\mathcal{S}_K(x)$. To do so, we first observe that

$$\#\{p \le x : \text{there exists a prime } q > (\log x)^{1/3} \text{ such that } q^2 | p - 1\} \ll \frac{\mathrm{li}(x)}{(\log x)^{1/3}}. \tag{7}$$

To see that (7) holds, observe that, using (6), we obtain that

$$\begin{aligned}\sum_{\substack{p\le x\\ p\equiv 1 \pmod{q^2}\\ q>(\log x)^{1/3}}} 1 &\le \sum_{q>(\log x)^{1/3}} \pi(x;q^2,1) \ll \sum_{q>(\log x)^{1/3}} \frac{1}{\phi(q^2)}\mathrm{li}(x)\\ &= \mathrm{li}(x) \sum_{q>(\log x)^{1/3}} \frac{1}{q(q-1)} \ll \mathrm{li}(x) \sum_{q>(\log x)^{1/3}} \frac{1}{q^2}\\ &< \mathrm{li}(x)\int_{(\log x)^{1/3}}^{\infty} \frac{dt}{t^2} \ll \frac{\mathrm{li}(x)}{(\log x)^{1/3}},\end{aligned}$$

which clearly establishes (7).

Let $P_y := \prod_{p<y} p$. Choose $y = \frac{\log x}{3}$ and consider $K \in \mathcal{B}$ with $P(K) < y$ and $K < (\log x)^c$ for some constant $c > 0$. Then, in view of (7), we have

$$\mathcal{S}_K(x) = \sum_{\delta | P_y} \mu(\delta)\pi(x;K\delta^2,1) + O\left(\frac{\mathrm{li}(x)}{(\log x)^{1/3}}\right). \tag{8}$$

But for $\delta|P_y$ and since by Chebychev's inequality, $P_y \le e^{1.05y} = x^{0.35}$, it follows that $K\delta^2 < KP_y^2 < Cx^{0.75}$ for some constant $C > 0$. Therefore, combining estimates (8), (5) and (6), we get that

$$\begin{aligned}(9)\ \mathcal{S}_K(x) &= \mathrm{li}(x)\sum_{\delta|P_y} \frac{\mu(\delta)}{\phi(K\delta^2)} + O\left(\mathrm{li}(x)e^{-c_1\sqrt{\log x}} \sum_{\delta|P_y}\frac{1}{\phi(K\delta^2)}\right)\\ &\qquad + O\left(\mathrm{li}(x) \sum_{\substack{\delta|P_y\\ \delta>(\log x)^c}} \frac{1}{\phi(K\delta^2)}\right)\\ &= \mathrm{li}(x)E_y(K) + O\left(\mathrm{li}(x)e^{-c_1\sqrt{\log x}}U_K(x)\right) + O\left(\mathrm{li}(x)V_K(x)\right),\end{aligned}$$

say. Now, since

$$\frac{1}{\phi(K\delta^2)} \le \frac{1}{\phi(K)}\cdot\frac{1}{\phi(\delta^2)} \qquad \text{and} \qquad \sum_{\delta}\frac{1}{\phi(\delta^2)} = O(1),$$

it follows that

$$U_K(x) \ll \frac{1}{\phi(K)}. \tag{10}$$

Moreover, since

$$\sum_{\delta>(\log x)^c} \frac{1}{\phi(\delta^2)} = \sum_{\delta>(\log x)^c} \frac{1}{\delta\phi(\delta)} \ll \sum_{\delta>(\log x)^c} \frac{\log\log\delta}{\delta^2} \ll \frac{\log\log\log x}{(\log x)^c},$$

we have

$$V_K(x) = O\left(\frac{\log\log\log x}{(\log x)^c}\right). \tag{11}$$

On the other hand,

$$E_y(K) = \frac{1}{\phi(K)} \prod_{p|K}\left(1 - \frac{1}{p^2}\right) \prod_{\substack{(p,K)=1 \\ p<(\log x)^{1/3}}} \left(1 - \frac{1}{p(p-1)}\right),$$

so that setting

$$E(K) := \frac{1}{\phi(K)} \prod_{p|K}\left(1 - \frac{1}{p^2}\right) \prod_{(p,K)=1} \left(1 - \frac{1}{p(p-1)}\right),$$

it is clear that

$$E_y(K) = (1 + o(1))E(K) \quad \text{as } y = y(x) \to \infty. \tag{12}$$

Gathering estimates (10), (11) and (12) in (9), we obtain

$$\frac{\mathcal{S}_K(x)}{\text{li}(x)} = E(K) + o(1) \qquad (x \to \infty).$$

We may thus conclude that $(\lambda(p-1)-1)\log(p-1)$ has a distribution function. Hence, letting $G(y)$ be the distribution function of the random variable ξ defined by $P(\xi = \kappa(K)) = E(K)$, we have thus established

Theorem 2. *For each point of continuity y of G,*

$$\lim_{x\to\infty} \frac{1}{li(x)} \#\{p \le x : (\lambda(p-1) - 1)\log(p-1) < y\} = G(y).$$

Remark. Repeating the argument used for F, one easily obtains that G is continuous at each real number y if and only if $y \notin \{\kappa(K) : K \in \mathcal{B}\}$.

5. The case of an irreducible cubic polynomial. Let $f(n)$ be an irreducible cubic polynomial with coefficients in $\mathbb{Z}$ with a positive leading coefficient. C. Hooley [5, Chapter 4] proved that

$$\#\{n \le x : \text{ there exists a prime } q > \frac{1}{6}\log x \text{ with } q^2|f(n)\} \ll \frac{x}{\sqrt{\log x}}. \tag{13}$$

Let $\rho(\ell)$ be the number of (incongruent) roots of the congruence $f(\nu) \equiv 0 \pmod{\ell}$. Now, given an arbitrary constant $c > 0$, let

$$\xi_1 = \frac{1}{6}\log x, \qquad K \in \mathcal{B}, \qquad K < \xi_1^c, \qquad P(K) < \xi_1.$$

Moreover, let

$$V_K(x) := \#\{n \le x : f(n) = Km, \ \gcd(K,m) = 1, \ \mu^2(m) = 1\}.$$

By using the Eratosthenian sieve (see for instance Halberstam and Richert [4, Chapter 1]), we have

$$V_K(x) = \sum_{\substack{\delta_1 \mid K \\ \delta_2 \mid \frac{P_{\xi_1}}{\gamma(K)}}} \mu(\delta_1)\mu(\delta_2)\#\{n \le x : f(n) \equiv 0 \pmod{K\delta_1\delta_2^2}\}.$$

Therefore,

$$V_K(x) = \frac{x}{K} \prod_{\substack{p<\xi_1 \\ (p,K)=1}} \left(1 - \frac{\rho(p^2)}{p^2}\right) \cdot$$

$$\prod_{p^\alpha \| K} \left(1 - \frac{\rho(p^{\alpha+1})}{p^{\alpha+1}}\right) + O\left(\frac{x}{\log^2 x}\right) + O(T_K(x)), \tag{14}$$

say, where in view of (13),

$$\sum_{K<(\log x)^c} T_K(x) \ll \frac{x}{\sqrt{\log x}}. \tag{15}$$

In what follows, we shall use a classical result of T. Nagell [7, Chapter III], from which it follows that $\rho(\ell) \le c \cdot 2^{\omega(\ell)}$. In particular, this allows us to write that

$$1 \ge \prod_{p \ge \xi_1} \left(1 - \frac{\rho(p^2)}{p^2}\right) \ge \exp\left\{-2\sum_{p\ge\xi_1} \frac{\rho(p^2)}{p^2}\right\} \ge \exp\{-\frac{c}{\xi_1}\} = 1 - O\left(\frac{1}{\xi_1}\right).$$

Thus, setting $D = \prod_p \left(1 - \frac{\rho(p^2)}{p^2}\right)$, we have that

$$\prod_{\substack{p<\xi_1 \\ (p,K)=1}} \left(1 - \frac{\rho(p^2)}{p^2}\right) \prod_{p^\alpha\|K} \left(1 - \frac{\rho(p^{\alpha+1})}{p^{\alpha+1}}\right) = D\left(1 + O\left(\frac{1}{\xi_1}\right)\right) \prod_{p^\alpha\|K} \frac{\left(1 - \frac{\rho(p^{\alpha+1})}{p^{\alpha+1}}\right)}{\left(1 - \frac{\rho(p^2)}{p^2}\right)},$$

so that in light of (15), (14) may be written as

$$V_K(x) = x\xi(K) + O\left(\frac{x}{\sqrt{\log x}}\right),$$

where

$$\xi(K) := \frac{D}{K} \prod_{p^\alpha\|K} \frac{\left(1 - \frac{\rho(p^{\alpha+1})}{p^{\alpha+1}}\right)}{\left(1 - \frac{\rho(p^2)}{p^2}\right)}.$$

Letting $F_f(y)$ be the distribution function of the random variable ξ_f defined by $P(\xi_f = \kappa(K)) = \xi(K)$, we can deduce, using the same approach as in the earlier two theorems, the following result.

Theorem 3. *Given an irreducible cubic polynomial f with positive leading coefficient, then, provided $y \neq \kappa(K)$ for all $K \in \mathcal{B}$,*

$$\lim_{x\to\infty} \frac{1}{x}\#\{n \le x : (\lambda(f(n)) - 1)\log f(n) < y\} = F_f(y).$$

Remark. Letting f be as in the statement of Theorem 3, one can show that

$$\begin{aligned}\sum_{n\le x} \lambda(f(n)) &= x + \sum_{K \text{ powerful}} \sum_{n\le x} \frac{\kappa(K)}{\log f(n)} + O\left(\sum_{K \text{ powerful}} \sum_{n\le x} \frac{\kappa^2(K)}{\log^2 f(n)}\right) \\ &= x + c_2 \frac{x}{\log x} + O\left(\frac{x}{\log^{3/2} x}\right),\end{aligned}$$

for some computable constant c_2.

6. The case of shifted powerful numbers. Given a powerful number K, consider the set

$$\mathcal{T}_K := \{n \in \mathcal{B} : n+1 = K\nu,\ \gcd(K,\nu) = 1,\ \mu^2(\nu) = 1\}.$$

It is well known that each powerful number n can be written uniquely in the form $n = r^3 m^2$, where $m \in \mathbb{N}$ and r is a squarefree number. Setting $\mathcal{B}^{(r)} := \{n \in \mathbb{N} : n = r^3 m^2,\ m = 1, 2, 3, \ldots\}$, we have

$$\mathcal{B} = \bigcup_{\substack{r=1\\ \mu^2(r)=1}}^{\infty} \mathcal{B}^{(r)}.$$

Now, let $\mathcal{T}_K^{(r)} := \mathcal{T}_K \cap \mathcal{B}^{(r)}$, so that

$$\mathcal{T}_K^{(r)} = \{n \in \mathbb{N} : n = r^3 m^2 \text{ and } n+1 = K\nu,\ \gcd(K,\nu) = 1,\ \mu^2(\nu) = 1\}.$$

Furthermore, we introduce the counting functions

$$\mathcal{T}_K(x) := \{n \le x : n \in \mathcal{T}_K\} \qquad \text{and} \qquad \mathcal{T}_K^{(r)}(x) := \{n \le x : n \in \mathcal{T}_K^{(r)}\}.$$

With good estimates of $\mathcal{T}_K(x)$ and $\mathcal{T}_K^{(r)}(x)$, at least in the range $K < (\log x)^c$ (for an arbitrary constant $c > 0$), $P(K) < \sqrt{\log x}$, $r < (\log x)^{1/4}$, we shall be able to establish that

$$\lim_{x\to\infty} \frac{1}{\mathcal{B}(x)}\#\{n \le x : n \in \mathcal{B}, (\lambda(n+1) - 1)\log n < y\}$$

exists, where $\mathcal{B}(x) := \#\{n \le x : n \in \mathcal{B}\}$.

To do so, let $1 \le a \le (\log x)^c$, where $c > 0$ is a given constant. Set $f_a(m) := am^2 + 1$ and let $\rho_a(\nu)$ be the number of solutions of $f_a(m) \equiv 0 \pmod{\nu}$. It is known that ρ_a is a multiplicative function and that, for each $\alpha \in \mathbb{N}$,

$$\rho_a(p^\alpha) = \begin{cases} 1 + \left(\frac{-a}{p}\right) & \text{if } (p, 2a) = 1, \\ 0 & \text{if } p|a, \end{cases}$$

and that if a is odd,

(1) $\rho_a(2) = 1$,

(2) $\rho_a(2^2) = \begin{cases} 2 & \text{if } 4 | a+1, \\ 0 & \text{if } 4 \nmid a+1, \end{cases}$

(3) if $\alpha \geq 3$, then $\rho_a(2^\alpha) = 4 \cdot \epsilon_a(\alpha)$, where

$$\epsilon_a(\alpha) = \begin{cases} 1 & \text{if } y^2 \equiv -a \pmod{2^\alpha} \text{ is solvable}, \\ 0 & \text{otherwise}. \end{cases}$$

If $r_0 = r_0(x)$ a function slowly tending to $+\infty$ with x, then

$$\lim_{x\to\infty} \frac{1}{\mathcal{B}(x)} \#\{n = m^2 r^3 \leq x : r \geq r_0\} = 0.$$

Let $K \in \mathcal{B}$ be fixed. Let $1 \leq a \leq r_0^3$, $x^{1/4} \leq z \leq x^4$ and $S_{K,a}(z)$ be the number of positive integers $m \leq z$ for which $f_a(m) = K\nu$, where $(K, \nu) = 1$ and where either ν is squarefree or if $p^2 | \nu$, then $p > \log x$. By using the Eratosthenian sieve (see Halberstam and Richert [4, Chapter 1]), one can obtain that

$$S_{K,a}(z) = \sum_{\delta_1, \delta_2} \mu(\delta_1)\mu(\delta_2) \#\{m \leq z : f_a(m) \equiv 0 \pmod{K\delta_1\delta_2^2}\},$$

where, in this sum, δ_1 runs over the squarefree divisors of K, while δ_2 runs over those squarefree numbers which are coprime to Ka, and for which the inequality $P(\delta_2) \leq \log x$ holds. Therefore

$$S_{K,a}(z) = z \sum_{K\delta_1\delta_2^2 \leq z} \frac{\rho_a(K\delta_1\delta_2^2)\mu(\delta_1)\mu(\delta_2)}{K\delta_1\delta_2^2} + O\left(\sum_{K\delta_1\delta_2^2 \leq az^2} \rho_a(K\delta_1\delta_2^2) \right). \tag{16}$$

The error term in (16) can easily be seen to be $O(\sqrt{z})$, due essentially to the fact that $P(\delta_2) \leq \log x$. On the other hand, the summation in the main term on the right hand side of (16) may also be taken to run over those $K\delta_1\delta_2^2 > z$ since

$$\left| \sum_{K\delta_1\delta_2^2 > z} \frac{\rho_a(K\delta_1\delta_2^2)\mu(\delta_1)\mu(\delta_2)}{K\delta_1\delta_2^2} \right| \leq \sum_{K\delta_1\delta_2^2 > z} \frac{\rho_a(K\delta_1)\rho_a(\delta_2^2)}{K\delta_1\delta_2^2} \ll \frac{1}{z^{1/4}},$$

say. Hence, writing $K = 2^\delta K_1$ with K_1 odd, we have

$$\frac{S_{K,a}(z)}{z} = \frac{1+o(1)}{K_1} \prod_{\substack{p^\alpha \| K_1 \\ (p,a)=1}} \left(1 - \frac{\rho_a(p)}{p}\right)$$

$$\prod_{(p, 2K_1 a)=1} \left(1 - \frac{\rho_a(p)}{p^2}\right) \cdot \delta(K, a) \qquad (z \to \infty), \tag{17}$$

where

$$\delta(K,a) = \begin{cases} 0 & \text{if } \gcd(K,a) > 1, \\ 1 & \text{if } a \text{ is even and } K \text{ is odd}, \\ \frac{1}{2^\delta}\left(\rho_a(2^\delta) - \frac{1}{2}\rho_a(2^{\delta+1})\right) & \text{if } a \text{ is odd and } K \text{ is even}, \\ 1 - \frac{\rho_a(2^2)}{2^2} & \text{if } a \text{ is odd and } K \text{ is odd}. \end{cases}$$

It is clear that

$$\mathcal{T}_K(x) = \sum_{\substack{r \le r_0 \\ (K,r)=1}} \mu^2(r)\mathcal{T}_K^{(r)}(x) + o(\mathcal{B}(x)) \tag{18}$$

and furthermore that

$$\mathcal{T}_K^{(r)}(x) = S_{K,r^3}\left(\frac{\sqrt{x}}{r^{3/2}}\right) + O(\Delta_r), \tag{19}$$

where Δ_r stands for the number of positive integers $n \le x$, $n \in \mathcal{T}_K^{(r)}$ such that $p^2|n+1$ for some prime $p > \log x$. The main difficulty in obtaining a "closed formula" for $\mathcal{T}_K(x)$ is to estimate the size of $\sum_{r \le r_0} O(\Delta_r)$.

If we set

$$D(x,Y) := \#\{n \le x : n \in \mathcal{K} \text{ for which there exists } p^2|n+1,\ p > Y\},$$

it is clear that

$$\sum_{r \le r_0} O(\Delta_r) \ll D(x, \log x). \tag{20}$$

We shall actually prove that, for some constant $C > 0$,

$$D(x, \sqrt{\log x}) \le C\frac{\sqrt{x}}{(\log x)^{1/3}} = o(\mathcal{B}(x)), \tag{21}$$

which together with (18), (19) and (20) will clearly be enough to show that

$$\mathcal{T}_K(x) = \sum_{\substack{r \le r_0 \\ (K,r)=1}} \mu^2(r)S_{K,r^3}\left(\frac{\sqrt{x}}{r^{3/2}}\right) + o\left(\mathcal{B}(x)\right), \tag{22}$$

in light of the well known estimate

$$\mathcal{B}(x) = d\sqrt{x} + O(x^{1/3}) \qquad \text{with } d = \frac{\zeta(3/2)}{\zeta(3)} \approx 2.17 \tag{23}$$

(see for instance Ivić and Shiu [6]).

To prove (21), for a given constant $c_3 > 0$, we let $r < (\log x)^{c_3}$ and count those positive integers $n = r^3m^2 \le x$ for which $p^2|r^3m^2+1$. If $p^2 \le x/r^3$, then no more than $2\frac{\sqrt{x}}{r^{3/2}p^2}$ such m's exist. So, assume that m_0 is the smallest positive integer m for which $p|r^3m^2+1$. Then, all the other m's can be written as $m = m_0+tp$ for some positive integer t. Let us search for those integers t for which $r^3(m_0+tp)^2+1 \equiv 0 \pmod{p^2}$. In this case, we have

$$\begin{aligned}(r^3m_0^2+1)+2m_0tr^3p+t^2p^2r^3 &\equiv 0 \pmod{p^2},\\ \frac{r^3m_0^2+1}{p}+2m_0tr^3 &\equiv 0 \pmod{p}.\end{aligned}$$

But since $(m_0,p)=1$ and $(2r^3,p)=1$, one obtains the value of $t \pmod{p}$.

Since $m_0+tp < \frac{\sqrt{x}}{r^{3/2}}$, no more than one such t may occur. Consequently, for $r < (\log x)^{c_3}$ and $p > \frac{\sqrt{x}}{r^{3/2}}$, at most one such m exists under the conditions $r^3m^2+1 \le x$ and $p^2 | r^3m^2+1$.

Therefore, given a large number W,

$$\begin{aligned} D(x,\sqrt{\log x}) &\le \sum_{\substack{r^3m^2\le x\\ r>W}} 1 + \sum_{r\le W}\pi(\sqrt{x}) + 2\sum_{r<W}\frac{\sqrt{x}}{r^{3/2}}\sum_{\sqrt{\log x}\le p\le \frac{\sqrt{x}}{r^{3/2}}}\frac{1}{p^2}\\ &\le \sqrt{x}\sum_{f>W}\frac{1}{r^{3/2}} + W\pi(\sqrt{x}) + 2\frac{\sqrt{x}}{\sqrt{\log x}}\sum_{r<W}\frac{1}{r^{3/2}}\\ &\le c_4\frac{\sqrt{x}}{\sqrt{W}} + c_5\frac{W\sqrt{x}}{\log x} + 2\frac{\sqrt{x}}{\sqrt{\log x}}\\ &\le c_6\frac{\sqrt{x}}{(\log x)^{1/3}}, \end{aligned}$$

by choosing $W=(\log x)^{2/3}$, which completes the proof of (21).

Now, set

$$\sigma(K,r) := \frac{1}{K_1}\prod_{p|K_1}\left(1-\frac{\rho_r(p)}{p}\right)\prod_{(p,2K_1r)=1}\left(1-\frac{\rho_r(p)}{p^2}\right)\cdot\delta(K,r^3).$$

Using (17), we obtain that

$$S_{K,r^3}\left(\frac{\sqrt{x}}{r^{3/2}}\right) = \frac{\sqrt{x}}{r^{3/2}}\sigma(K,r)+o(\sqrt{x}). \tag{24}$$

It follows from (23) and (24) that

$$\frac{1}{\mathcal{B}(x)}S_{K,r^3}\left(\frac{\sqrt{x}}{r^{3/2}}\right) = \frac{1}{dr^{3/2}}\sigma(K,r)+o(1) \qquad (x\to\infty). \tag{25}$$

Therefore, setting

$$\Delta(K) := \sum_{(K,r)=1}\mu^2(r)\frac{\sigma(K,r)}{r^{3/2}},$$

it follows from (22) and (25) that

$$\mathcal{T}_K(x) = (1+o(1))\Delta(K)\sqrt{x} \qquad (x\to\infty).$$

Hence, letting $H(y)$ be the distribution function of the random variable ξ defined by $P(\xi = \kappa(K)) = \Delta(K)/d$, we can deduce using the same approach as in the earlier theorems the following result.

Theorem 4. *For each point of continuity y of H,*

$$\lim_{x\to\infty} \frac{1}{d\sqrt{x}} \#\{n \le x : n \in \mathcal{B},\ \eta(n+1) < y\} = H(y).$$

7. Further remarks.

Lemma 2. *Let $F \in \mathbb{Z}[x]$, $F(x) = f_1(x)\cdots f_r(x)$ be a product of irreducible polynomials such that $gcd(f_i(x), f_j(x)) = 1$ for every $i \ne j$.*

(a) *If $deg\, f_j \le 3$ $(j = 1, \ldots, r)$, then for every $\xi_1 > 0$ there exists $\xi_2 > 0$ such that*

$$\{n \le x : q^2|F(n) \text{ for some prime } q > (\log x)^{\xi_1}\} \le \frac{cx}{(\log x)^{\xi_2}}.$$

(b) *If $deg\, f_j \le 2$ $(j = 1, \ldots, r)$, then*

$$\#\{p \le x : q^2|F(p) \text{ for some prime } q > (\log x)^{\xi_1}\} \le \frac{c\, li(x)}{(\log x)^{\xi_2}}.$$

Proof. In the case $r = 1$, this can be proved by essentially repeating the argument found in Chapter 4 of Hooley [5]. Indeed, let $D_{i,j}$ be the resultant of $(f_i(x), f_j(x))$. It is known that $q^2|f_i(n)f_j(n)$ with $q \nmid D_{i,j}$ implies that either $q^2|f_i(n)$ or $q^2|f_j(n)$. By this observation, Lemma 2 is proved.

Given a function F satisfying the conditions of Lemma 2, then using Lemma 2 and a routine application of the Eratosthenian sieve (see Halberstam and Richert [4, Chapter 1]), one can show that

$$\begin{aligned} \lim_{x\to\infty} \frac{1}{x} \#\{n \le x : F(n) \in \mathcal{A}_K\} &= A_K, \\ \lim_{x\to\infty} \frac{1}{\pi(x)} \#\{p \le x : F(p) \in \mathcal{A}_K\} &= B_K, \end{aligned}$$

for some constants A_K and B_K such that

$$\sum_{K\in\mathcal{B}} A_K = 1 \qquad \text{and} \qquad \sum_{K\in\mathcal{B}} B_K = 1.$$

Letting θ and ψ be random variables such that

$$P(\theta = \kappa(K)) = A_K, \qquad P(\psi = \kappa(K)) = B_K$$

and letting H_θ and H_ψ be their corresponding distribution functions, then we have the following result.

Theorem 5. *Under the conditions of Lemma 2,*

$$\lim_{x\to\infty}\frac{1}{x}\#\{n\le x:\eta(F(n))<y\} = H_\theta(y),$$
$$\lim_{x\to\infty}\frac{1}{\pi(x)}\#\{p\le x:\eta(F(p))<y\} = H_\psi(y),$$

provided $y\neq\kappa(K)$ *for all* $K\in\mathcal{B}$.

References

[1] R. Crandall and C. Pomerance, Prime Numbers. 2nd edition, Springer, New York, 2005.

[2] J. M. De Koninck and N. Doyon, À propos de l'indice de composition des nombres. Monatshefte für Mathematik **139**, 151–167 (2003).

[3] J. M. De Koninck and I. Kátai, On the mean value of the index of composition. Monatshefte für Mathematik **145**, 131–144 (2005).

[4] H. Halberstam and H. E. Richert, Sieve Methods. Academic Press, New York, 1974.

[5] C. Hooley, Applications of Sieve Methods to the Theory of Numbers. Cambridge Tracts in Mathematics, Vol. **70**, Cambridge University Press, 1976.

[6] A. Ivić and P. Shiu, The distribution of powerful numbers. Illinois J. Math. **26**, 576–590 (1982).

[7] T. Nagell, Introduction to Number Theory. Wiley, New York, 1951.

[8] A. Walfisz, Zur additiven Zahlentheorie, II. Math. Zeitschr. **40**, 592–607 (1936).

J. M. De Koninck, Dép. de mathématiques, Université Laval, Québec, Québec G1K 7P4, Canada
e-mail: jmdk@mat.ulaval.ca

I. Kátai, Computer Algebra Department, Eötvös Loránd University, 1117 Budapest, Pázmány Péter Sétány I/C, Hungary
e-mail: katai@compalg.inf.elte.hu

M. V. Subbarao, Department of Mathematics, University of Alberta, Edmonton, Alberta 76G 2G1, Canada

Received: 3 March 2006

Revised: 28 October 2006