Monatsh. Math. 145, 131–144 (2005)
DOI 10.1007/s00605-004-0288-6

Monatshefte für
Mathematik
Printed in Austria

On the Mean Value of the Index of Composition of an Integer

By

J. M. De Koninck[1,*] and **I. Kátai**[2,†]

[1] Université Laval, Québec, Canada
[2] Eötvös Loránd University, Budapest, Hungary

Communicated by J. Schoissengeier

Received January 15, 2004; accepted in revised form July 26, 2004
Published online March 25, 2005

Abstract. For each integer $n \geqslant 2$, let $\lambda(n) = \frac{\log n}{\log \gamma(n)}$ be the *index of composition* of n, where $\gamma(n) = \prod_{p|n} p$. For convenience, we write $\lambda(1) = \gamma(1) = 1$. We obtain sharp estimates for $\sum_{x \leqslant n \leqslant x+\sqrt{x}} \lambda(n)$ and $\sum_{n \leqslant x} \lambda(n)$, as well as for $\sum_{x \leqslant n \leqslant x+\sqrt{x}} 1/\lambda(n)$ and $\sum_{n \leqslant x} 1/\lambda(n)$. Finally we study the sum of λ running over shifted primes.

2000 Mathematics Subject Classification: 11A25, 11N25, 11N37
Key words: Arithmetic function, prime factorisation

1. Introduction

For each integer $n \geqslant 2$, let $\lambda(n) = \frac{\log n}{\log \gamma(n)}$ be the *index of composition* of n, where $\gamma(n) = \prod_{p|n} p$. For convenience, we write $\lambda(1) = \gamma(1) = 1$. The index of composition of an integer measures essentially the multiplicity of its prime factors. It was shown by De Koninck and Doyon [1] that the average order of $\lambda(n)$ is 1 and more precisely that

$$\sum_{n \leqslant x} \lambda(n) = x + O\left(\frac{x}{\log x}\right). \tag{1}$$

A similar result was obtained for $\sum_{n \leqslant x} \frac{1}{\lambda(n)}$.

Here, we first prove a short interval version of (1) using a result of Filaseta and Trifonov [2]. Then, we provide estimates for $\sum_{x \leqslant n \leqslant x+\sqrt{x}} \lambda(n)$ and $\sum_{x \leqslant n \leqslant x+\sqrt{x}} 1/\lambda(n)$ with an error term $O\left(\frac{x}{\log^{r+1} x}\right)$, where r is any given positive integer. From these, we deduce a sharpening of the estimate (1) and similarly for $\sum_{n \leqslant x} \frac{1}{\lambda(n)}$. Finally we study the sum of λ running over shifted primes.

* Research supported in part by a grant from NSERC.
† Research supported by the Applied Number Theory Research Group of the Hungarian Academy of Science and by a grant from OTKA.

2. Main Results

Theorem 1. *Let $h(x) = x^{1/5}\log^3 x$. Then*

$$\sum_{x\leqslant n\leqslant x+h(x)} (\lambda(n)-1) = O\left(\frac{h(x)}{\log x}\right) \tag{2}$$

and

$$\sum_{x\leqslant n\leqslant x+h(x)} \left(1-\frac{1}{\lambda(n)}\right) = O\left(\frac{h(x)}{\log x}\right). \tag{3}$$

Remark. Observe that since $\lambda(n)\geqslant 1$, the summands on the left hand side of (2) and (3) are non negative.

Theorem 2. *Given any positive integer r, there exist computable constants $c_1, c_2, \dots, c_r, c'_2, \dots, c'_r$ such that*

$$\frac{1}{\sqrt{x}}\sum_{x\leqslant n\leqslant x+\sqrt{x}} \lambda(n) = 1 + c_1\frac{1}{\log x} + c_2\frac{1}{\log^2 x} + \cdots + c_r\frac{1}{\log^r x} + O\left(\frac{1}{\log^{r+1} x}\right) \tag{4}$$

and

$$\sum_{n\leqslant x}\lambda(n) = x + c_1\frac{x}{\log x} + c'_2\frac{x}{\log^2 x} + \cdots + c'_r\frac{x}{\log^r x} + O\left(\frac{x}{\log^{r+1} x}\right). \tag{5}$$

Theorem 3. *Given any positive integer r, there exist computable constants $d_1, d_2, \dots, d_r, d'_2, \dots, d'_r$ such that*

$$\frac{1}{\sqrt{x}}\sum_{x\leqslant n\leqslant x+\sqrt{x}} \frac{1}{\lambda(n)} = 1 + d_1\frac{1}{\log x} + d_2\frac{1}{\log^2 x} + \cdots + d_r\frac{1}{\log^r x} + O\left(\frac{1}{\log^{r+1} x}\right) \tag{6}$$

and

$$\sum_{n\leqslant x}\frac{1}{\lambda(n)} = x + d_1\frac{x}{\log x} + d'_2\frac{x}{\log^2 x} + \cdots + d'_r\frac{x}{\log^r x} + O\left(\frac{x}{\log^{r+1} x}\right). \tag{7}$$

Theorem 4. *Given any positive integer r, there exist computable constants $\alpha_2, \alpha_3, \dots, \alpha_r$ such that*

$$\frac{1}{\pi(x+x^{2/3})-\pi(x)}\sum_{x<p\leqslant x+x^{2/3}} \lambda(p+1) = 1 + \alpha_2\frac{1}{\log^2 x} + \alpha_3\frac{1}{\log^3 x} + \cdots + \alpha_r\frac{1}{\log^r x} + O\left(\frac{1}{\log^{r+1} x}\right) \tag{8}$$

and

$$\sum_{p \leqslant x} \lambda(p+1) = \frac{x}{\log x} + \alpha_2 \frac{x}{\log^2 x} + \alpha_3 \frac{x}{\log^3 x} + \cdots + \alpha_r \frac{x}{\log^r x} + O\left(\frac{x}{\log^{r+1} x}\right). \tag{9}$$

3. Preliminary Results

Let $\mathbf{N}$ stand for the set of positive integers. A number $n \in \mathbf{N}$ is said to be squarefull (or powerful) if $n = 1$ or if $p|n \Longrightarrow p^2|n$.

Lemma 1. *Let β be an arbitrary non negative integer. As $y \to \infty$,*

(i)
$$\sum_{\substack{k \geqslant y \\ k \text{ squarefull}}} \frac{\log^\beta k}{k} \ll \frac{\log^\beta y}{\sqrt{y}}.$$

Moreover, given any real number $\alpha > \frac{1}{2}$,

(ii)
$$\sum_{k \text{ squarefull}} \frac{1}{k^\alpha} < +\infty.$$

Proof. It is well known that

$$N(t) := \sum_{\substack{n \leqslant t \\ n \text{ squarefull}}} 1 = c\sqrt{t} + O(t^{1/3}), \qquad c = \frac{\zeta(3/2)}{\zeta(3)} \approx 2.1732,$$

where ζ stands for the Riemann Zeta Function (see for instance Ivić and Shiu [4]). Therefore, integration by parts yields

$$\begin{aligned}\sum_{\substack{k \geqslant y \\ k \text{ squarefull}}} \frac{\log^\beta k}{k} &= \int_y^\infty \frac{\log^\beta t}{t}\, dN(t) \\ &= \frac{\log^\beta t}{t} N(t)\Big|_y^\infty + \int_y^\infty \frac{\log^\beta t}{t^2} N(t)\, dt - \beta \int_y^\infty \frac{\log^{\beta-1} t}{t^2} N(t)\, dt \\ &= -c\frac{\log^\beta y}{\sqrt{y}} + O\left(\frac{\log^\beta y}{y^{2/3}}\right) + O\left(\frac{\log^\beta y}{\sqrt{y}}\right) \ll \frac{\log^\beta y}{\sqrt{y}},\end{aligned}$$

which completes the proof of (i). For a proof of (ii), see De Koninck and Doyon [1].

For each positive integer k and each real number y, let

$$M(y|k) := \sum_{\substack{n \leqslant y \\ (n,k)=1}} \mu^2(n) \quad \text{and} \quad M(y) = M(y|1) = \sum_{n \leqslant y} \mu^2(n), \tag{10}$$

where μ stands for the Moebius function. For each positive integer k, let $\mathscr{B}_k$ be the set of all positive integers all of whose prime factors divide k, and let $1 \in \mathscr{B}_k$. Moreover, let $\Omega(n) := \sum_{p^\alpha \| n} \alpha$ for each integer $n \geqslant 2$ and set $\Omega(1) = 0$.

Walfisz [6] proved that there exists a positive constant A such that

$$M(x) = \frac{6}{\pi^2}x + T(x), \qquad T(x) \ll \sqrt{x}\,\exp\{-A(\log x)^{3/5}(\log\log x)^{-1/5}\}. \quad (11)$$

We shall use this result in the proof of the next lemma. But first, for each positive integer k and real number $\delta > 0$, let us set

$$\rho(k) = \prod_{p|k}\left(1+\frac{1}{p}\right)^{-1} = \sum_{v \in \mathscr{B}_k} \frac{(-1)^{\Omega(v)}}{v}, \qquad \psi_\delta(k) = \prod_{p|k}\left(1-\frac{1}{p^\delta}\right)^{-1}. \quad (12)$$

Lemma 2. *Let ν be a fixed positive integer. Then for each positive integer k,*

(i) $M(y|k) = \sum_{v \in \mathscr{B}_k} (-1)^{\Omega(v)} M\left(\frac{y}{v}\right)$,

(ii) $M(y|k) = \frac{6}{\pi^2}\rho(k)y + O(\sqrt{y}\,\exp\{-\eta(\log y)^{3/5}(\log\log y)^{-1/5}\})$ *uniformly for $k \leqslant (\log y)^\nu$ as $y \to \infty$, for some positive constant η.*

Proof. (i) follows from the identity

$$\sum_{\substack{n=1\\(n,k)=1}}^{\infty} \frac{\mu^2(n)}{n^s} = \frac{\prod_p\left(1+\frac{1}{p^s}\right)}{\prod_{p|k}\left(1+\frac{1}{p^s}\right)} = \sum_{n=1}^{\infty}\frac{\mu^2(n)}{n^s}\prod_{p|k}\left(1-\frac{1}{p^s}+\frac{1}{p^{2s}}-\cdots\right)$$

which is valid for all $\Re s > 1$, and the uniqueness of Dirichlet series representation.

In order to prove (ii), we first observe that

$$\tau(k) < \exp\left\{2\frac{\log k}{\log\log k}\right\} \quad (k \geqslant 3),$$

where $\tau(k)$ stands for the number of positive divisors of k (see for instance Nicolas and Robin [5]). It follows from this estimate that, given any $\delta > 0$, $\nu \in \mathbf{N}$, and $k \leqslant (\log y)^\nu$, as y becomes large, we have

$$\psi_\delta(k) \leqslant \tau(k) \leqslant \exp\left\{2\frac{\log k}{\log\log k}\right\} \leqslant \exp\left\{2\nu\frac{\log\log y}{\log\log\log y}\right\}. \quad (13)$$

Now from (i) and (12), we have

$$\begin{aligned} M(y|k) &= \frac{6}{\pi^2}\rho(k)y + O\left(y\sum_{\substack{v>y\\ v\in\mathscr{B}_k}}\frac{1}{v}\right) + O\left(\sum_{\substack{v\leqslant y\\ v\in\mathscr{B}_k}}\left|T\left(\frac{y}{v}\right)\right|\right)\\ &= \frac{6}{\pi^2}\rho(k)y + O(E_1) + O(E_2), \end{aligned} \quad (14)$$

say. First observe that

$$\sum_{\substack{v>y\\ v\in\mathscr{B}_k}}\frac{1}{v} = \sum_{\substack{v>y\\ v\in\mathscr{B}_k}}\frac{1}{v^{1-\delta}}\frac{1}{v^\delta} < \frac{1}{y^{1-\delta}}\sum_{\substack{v>y\\ v\in\mathscr{B}_k}}\frac{1}{v^\delta} \leqslant \frac{\psi_\delta(k)}{y^{1-\delta}},$$

so that using (13), we get

$$E_1 \ll \psi_\delta(k) y^\delta \ll y^{1/4}, \tag{15}$$

assuming that $\delta < 1/5$, say.

To estimate E_2, we proceed as follows. If $v \leqslant y^{1-\delta}$ for some fixed real number $\delta > 0$, then, in view of Walfisz's result (11),

$$\left| T\left(\frac{y}{v}\right) \right| \leqslant \sqrt{\frac{y}{v}} \exp\{-A_\delta (\log y)^{3/5} (\log \log y)^{-1/5}\} \tag{16}$$

for some positive constant A_δ, while the left hand side of (16) is $\leqslant D\sqrt{\frac{y}{v}}$ if $v \leqslant y$, for some constant $D > 0$, so that we may write

$$E_2 \leqslant \begin{cases} \sqrt{\frac{y}{v}} \exp\{-A_\delta (\log y)^{3/5} (\log \log y)^{-1/5}\} & \text{if } v \leqslant y^{1-\delta}, \\ D\sqrt{\frac{y}{v}} & \text{if } v \leqslant y. \end{cases} \tag{17}$$

Now, using the definition of ψ_δ and then (13), we have

$$\sqrt{y} \sum_{\substack{v \leqslant y^{1-\delta} \\ v \in \mathscr{B}_k}} \frac{1}{\sqrt{v}} \leqslant \sqrt{y}\ \psi_{1/2}(k) \leqslant \sqrt{y}\ \exp\left\{2\nu \frac{\log \log y}{\log \log \log y}\right\}. \tag{18}$$

On the other hand, if $v > y^{1-\delta}$, then

$$v = v^{1-\delta} v^\delta > y^{(1-\delta)^2} v^\delta,$$

in which case, again using (13), we have

$$\begin{aligned} \sqrt{y} \sum_{\substack{v > y^{1-\delta} \\ v \in \mathscr{B}_k}} \frac{1}{\sqrt{v}} &< \frac{\sqrt{y}}{y^{(1-\delta)^2/2}} \sum_{v \in \mathscr{B}_k} \frac{1}{v^{\delta/2}} \\ &\leqslant y^{\frac{1}{2}(1-(1-\delta)^2)} \psi_{\delta/2}(k) \\ &\leqslant y^{\frac{1}{2}(1-(1-\delta)^2)}\ \exp\left\{2\nu \frac{\log \log y}{\log \log \log y}\right\}. \end{aligned} \tag{19}$$

Choosing δ small and combining (18) with (19), we get from (17) that there exists a positive constant η such that

$$E_2 \ll \sqrt{y}\ \exp\{-\eta (\log y)^{3/5} (\log \log y)^{-1/5}\}. \tag{20}$$

Gathering (15) and (20), (ii) follows, and Lemma 2 is proved.

For each positive integer k and each real number y, let

$$\Pi(y|k) := \sum_{\substack{p \leqslant y \\ p \equiv -1 \pmod{k} \\ \left(\frac{p+1}{k}, k\right) = 1}} \mu^2\left(\frac{p+1}{k}\right)$$

and further let $\pi(x; D, \ell)$ be the number of primes $p \leqslant x$ such that $p \equiv \ell \pmod{D}$.

Lemma 3. Let k *be a squarefull integer. Then*

$$\Pi(y|k) = \sum_{\delta|k} \sum_{(\xi,k)=1} \mu(\delta)\mu(\xi)\pi(y; k\delta\xi^2, -1).$$

Proof. Since $\sum_{d|n} \mu(d)$ is always 0, unless $n = 1$ and since

$$\sum_{d^2|n} \mu(d) = \begin{cases} 1 & \text{if } n \text{ is squarefree,} \\ 0 & \text{otherwise,} \end{cases}$$

it follows that

$$\begin{aligned} \Pi(y|k) &= \sum_{\substack{p \leqslant y \\ p \equiv -1 \pmod{k}}} \sum_{\delta \left| \left(\frac{p+1}{k}, k\right)\right.} \mu(\delta) \sum_{\xi^2 | \frac{p+1}{k}} \mu(\xi) \\ &= \sum_{\delta|k} \sum_{(\xi,k)=1} \mu(\delta)\mu(\xi) \sum_{\substack{p \leqslant y \\ p \equiv -1 \pmod{k\delta\xi^2}}} 1, \end{aligned}$$

which completes the proof of Lemma 3.

Let ϕ stand for Euler's function and let $\mathrm{li}(y) := \int_2^y \frac{dt}{\log t}$.

Lemma 4. *Let* $\varepsilon > 0$ *be a small number. Let* ν *and* r *be fixed positive integers and let* $y^{\frac{7}{12}+\varepsilon} \leqslant H \leqslant y$. *Then, as* $y \to \infty$,

$$\Pi(y+H|k) - \Pi(y|k) = \frac{li(y+H) - li(y)}{k} \prod_{q \nmid k} \left(1 - \frac{1}{q(q-1)}\right) + O\left(\frac{H}{k} \frac{1}{\log^{r+2} y}\right)$$

uniformly for $k \leqslant (\log y)^{\nu}$.

Proof. From Lemma 3 and Huxley's Theorem for primes in arithmetic progressions (see Huxley [3]), we have that, given any number $c_0 > 0$,

$$\begin{aligned} \Pi(y+H|k) - \Pi(y|k) &= \sum_{\delta|k} \sum_{(\xi,k)=1} \mu(\delta)\mu(\xi)(\pi(y+H; k\delta\xi^2, -1) - \pi(y; k\delta\xi^2, -1)) \\ &= (\mathrm{li}(y+H) - \mathrm{li}(y)) \sum_{\delta|k} \sum_{(\xi,k)=1} \frac{\mu(\delta)\mu(\xi)}{\phi(k\delta\xi^2)} \\ &\quad + O\left(\sum_{\delta|k} \sum_{\substack{(\xi,k)=1 \\ \xi \leqslant (\log y)^{c_0}}} \frac{H|\mu(\delta)|}{\phi(k\delta\xi^2)} \exp(-\sqrt{\log y}) \right) \\ &\quad + O\left(\sum_{\delta|k} \sum_{\substack{(\xi,k)=1 \\ \xi > (\log y)^{c_0}}} \frac{2H}{k\delta\xi^2} \right). \end{aligned}$$

Clearly this last error term is

$$\ll \frac{H}{k}\frac{1}{(\log y)^{c_0}}\sum_{\delta|k}\frac{1}{\delta} \ll \frac{H}{k}\frac{1}{\log^{r+1} y},$$

for $k \leqslant (\log y)^{\nu}$, provided c_0 is chosen sufficiently large, whereas the first error term is

$$\ll \frac{H\exp(-\sqrt{\log y})}{\phi(k)}\sum_{\delta|k}\frac{|\mu(\delta)|}{\phi(\delta)} \ll \frac{H}{k}\frac{1}{\log^{r+1} y}.$$

Gathering these estimates and observing that

$$\sum_{\delta|k}\sum_{(\xi,k)=1}\frac{\mu(\delta)\mu(\xi)}{\phi(k\delta\xi^2)} = \frac{1}{\phi(k)}\sum_{\delta|k}\frac{\mu(\delta)}{\delta}\sum_{(\xi,k)=1}\frac{\mu(\xi)}{\phi(\xi^2)} = \prod_{q\nmid k}\left(1-\frac{1}{q(q-1)}\right),$$

the proof of Lemma 4 is complete.

Lemma 5. *Let $h(x) = x^{1/5}\log^3 x$ and let $E(x)$ be the number of integers $n \in [x, x+h(x)]$ which can be written as $n = p^2\nu$ for some prime number $p \geqslant \sqrt{h(x)}$. Then*

$$E(x) \ll \frac{h(x)}{\log x}.$$

Proof. This estimate is an immediate consequence of Theorem A, which is an unpublished result of Michael Filaseta who kindly communicated it to the first author.

Theorem A (Filaseta). *Let k be an integer $\geqslant 2$. Let $g(x)$ be a function satisfying $1 \leqslant g(x) \leqslant \log x$ for x sufficiently large, and set*

$$h = x^{1/(2k+1)}g(x)^3.$$

Then the number of k-free numbers in the interval $(x, x+h]$ is

$$\frac{h}{\zeta(k)} + O\left(\frac{h(\log x)}{g(x)^3}\right) + O\left(\frac{h}{g(x)}\right).$$

Proof. We modify the argument given in Section 4 of Filaseta and Trifonov [2]. Take $z = \varepsilon\log x$, where $\varepsilon = \varepsilon(k) > 0$ is sufficiently small. A sieve-of-Eratosthenes argument gives that the number of integers in $(x, x+h]$ free from divisors of the form p^k with $p \leqslant z$ is

$$\begin{aligned} h\prod_{p\leqslant z}\left(1-\frac{1}{p^k}\right) + O(2^z) &= h\prod_{p}\left(1-\frac{1}{p^k}\right) + O(h/z) + O(2^z) \\ &= h\prod_{p}\left(1-\frac{1}{p^k}\right) + O\left(\frac{h}{\log x}\right) + O(x^{\varepsilon\log 2}) \\ &= \frac{h}{\zeta(k)} + O\left(\frac{h}{g(x)}\right). \end{aligned}$$

We now obtain an upper bound for the number of these integers that are divisible by the k^{th} power of a prime $> z$. If $p > 2x^{1/k}$, then $p^k > 2^k x > x+h$ and, hence, p^k

does not divide integers in the interval $(x, x+h]$. Let c be a sufficiently large constant (for the purpose of having a condition in Theorem 7 of [2] hold later in the argument). We break up our consideration of primes into two parts, those primes p in $I = (z, ch]$ and those primes p in $J = (ch, 2x^{1/k}]$. If M_u denotes the number of integers in $(x, x+h]$ divisible by u^k, then

$$\sum_{p \in I} M_p \leqslant \sum_{p \in I} \left(\left[\frac{h}{p^k} \right] + 1 \right) \leqslant \sum_{p \geqslant z} \frac{h}{p^k} + \pi(ch)$$
$$\leqslant \frac{h}{z} + O(h/\log x) = O(h/\log x) = O(h/g(x)).$$

Thus, it remains to show that

$$\sum_{p \in J} M_p = O\left(\frac{h(\log x)}{g(x)^3} \right) + O\left(\frac{h}{g(x)} \right).$$

It suffices to obtain the same bound for $\sum_{u \in J} M_u$ (where u runs through the integers in J). Since such u exceed h, we deduce that $M_u \in \{0, 1\}$ and $M_u = 1$ precisely when there is a multiple of u^k in $(x, x+h]$. We use disjoint intervals J_i, with $1 \leqslant i \leqslant r$, of the form $(N, 2N]$ satisfying

$$J \subseteq \bigcup_{i=1}^{r} J_i \subseteq (ch/2, 2x^{1/k}].$$

In particular, $r \ll \log x$. Fix $J_i = (N, 2N]$. Note that

$$\frac{ch}{2} \leqslant N \leqslant x^{1/k}.$$

As in [2], we obtain that

$$\sum_{u \in J_i} M_u \ll |\{u \in (N, 2N] : \|f(u)\| < \delta\}|,$$

where $f(u) = x/u^k$, $\delta = hN^{-k}$, and $\|f(u)\|$ denotes the distance from $f(u)$ to the nearest integer. We appeal to Theorem 7 from [2] (with $s = k$ and $X = x$). Observe that the condition there on δ holds as c is sufficiently large. We obtain that

$$\sum_{u \in J_i} M_u \ll x^{1/(2k+1)} + \delta x^{1/(6k+3)} N^{(6k^2+k-1)/(6k+3)}$$
$$\ll x^{1/(2k+1)} + h x^{1/(6k+3)} N^{-1/3}.$$

Summing over i, we deduce

$$\sum_{u \in J} M_u \ll x^{1/(2k+1)} \log x + h x^{1/(6k+3)} h^{-1/3}$$
$$\ll x^{1/(2k+1)} \log x + x^{1/(2k+1)} g(x)^2.$$

The desired upper bound now follows, thus completing the proof of Theorem A.

4. Proof of Theorem 1

Each positive integer n can be written uniquely as

$$n = km, \qquad (m,k) = 1, \qquad \mu^2(m) = 1, \quad k \text{ squarefull},$$

so that

$$\gamma(n) = \gamma(k)m = \frac{\gamma(k)}{k} mk = \frac{\gamma(k)}{k} n$$

and therefore

$$\lambda(n) = \frac{\log n}{\log n + \log \frac{\gamma(k)}{k}} = \frac{1}{1 - \frac{\log (k/\gamma(k))}{\log n}}. \tag{21}$$

From this it follows in particular that there exists a positive constant B_1 such that

$$1 \leqslant \lambda(n) \leqslant 1 + B_1 \frac{\log (k/\gamma(k))}{\log n}$$

if $k < x^{1/3}$, say. Thus, writing $k(n)$ for the squarefull part of n, we have

$$\begin{aligned}\sum_{x \leqslant n \leqslant x+h(x)} (\lambda(n) - 1) &\leqslant \frac{B_1}{\log x} \sum_{k \leqslant h(x)} (\log (k/\gamma(k))) \cdot \#\{n \in [x, x+h(x)] : k(n) = k\} \\ &\quad + \sum_{\substack{x \leqslant n \leqslant x+h(x) \\ k(n) > h(x)}} (\lambda(n) - 1) \\ &= \frac{B_1}{\log x} S_1 + S_2, \end{aligned} \tag{22}$$

say. Using Lemma 1, we have that

$$S_1 \leqslant 2 \sum_{\substack{k \leqslant h(x) \\ k \text{ squarefull}}} (\log (k/\gamma(k))) \frac{h(x)}{k} \leqslant B_2 h(x) \tag{23}$$

for some absolute constant $B_2 > 0$, since the series over k converges by (ii) of Lemma 2.

It remains to estimate S_2. Let $U = \{n \in [x, x + h(x)] : k(n) > h(x)\}$ and let $U_1 \subseteq U$ be the subset of those $n \in U$ for which there is some squarefull divisor $t | k(n)$ such that $\log^4 x < t < h(x)$. It is clear that, using (i) of Lemma 1 with $\beta = 0$ and since $\lambda(n) \ll \log x$ for $n \leqslant 2x$,

$$\sum_{n \in U_1} (\lambda(n) - 1) \ll \log x \cdot h(x) \cdot \sum_{\substack{t \text{ squarefull} \\ t > \log^4 x}} \frac{1}{t} \ll \frac{\log x}{\log^2 x} h(x) = \frac{h(x)}{\log x}. \tag{24}$$

Let $U_2 = U \setminus U_1$. Given $n \in U_2$, let $t = \pi_1^{\beta_1} \cdots \pi_\ell^{\beta_\ell}$, with $\pi_1 < \cdots < \pi_\ell$ primes, be the smallest squarefull divisor of $k(n)$ which is larger than $h(x)$. Since $k(n)$ is such a candidate, then such a divisor t must exist. Furthermore, for every $j \in [1, \ell]$, either $\frac{t}{\pi_j}$ or $\frac{t}{\pi_j^2}$ is squarefull, and therefore $\frac{t}{\pi_j^2} \leqslant \log^4 x$, so that $\pi_j^2 \geqslant \frac{h(x)}{\log^4 x}$. Since $\pi_j^2 | k(n)$ and $n \notin U_1$, it follows that $\pi_j^2 > h(x)$, so that the conditions of Lemma 5 are satisfied. Observe furthermore that if $n \in U_2$, then

$$\log (k/\gamma(k)) < \log \frac{2x}{\sqrt{h(x)}} < \left(1 - \frac{1}{10}\right) \log x + O(\log \log x)$$

and therefore, using (21),

$$\lambda(n) < \frac{1}{1 - \frac{\frac{9}{10}\log x + O(\log\log x)}{\log x}} < \frac{1}{\frac{1}{10} + O\left(\frac{\log\log x}{\log x}\right)},$$

and so by Lemma 5, we obtain that

$$\sum_{n \in U_2} (\lambda(n) - 1) \ll \frac{h(x)}{\log x}. \tag{25}$$

Combining (22), (23), (24) and (25), we obtain (2). Since the proof of (3) is similar, we shall omit it.

5. Proofs of Theorems 2 and 3

We start with a general remark concerning the main idea used in the upcoming proofs. As we shall see, the contribution of $k = k(n)$, the squarefull part of n, in the estimates of Theorems 2 and 3 (or of $k = k(p+1)$ in Theorem 4) may be restricted to $k \leqslant \log^\nu x$ where ν is a large generic positive constant. This follows by trivial estimation and (i) of Lemma 1. Therefore, in the representation (21) for $\lambda(n)$, one can replace $\log n$ in the denominator by $\log x$ with negligible error. Hence this procedure permits, in view of the unique representation $n = km = k(n)m(n)$, $(k,m) = 1$, to split all sums in the following fashion: inner sum over m is performed using (ii) of Lemma 2, noting that partial summation permits one to evaluate (in terms of asymptotic expansion in decreasing powers of the logarithm) the quantity $\sum_{m \leqslant x, (m,k)=1} \log^{-j} m$ for $j = 0, 1, 2, \ldots$, where as always m denotes a generic squarefree number.

5.1. Regarding Theorems 2 and 3, it is enough to prove the short interval version, that is (4) and (6). Indeed, in each case, the "long interval" version, namely (5) or (7), follows by integrating with respect to x, as we will show in Section 5.2. Moreover, we shall only prove (4), the proof of (6) being very similar.

We start as in the proof of Theorem 1 by writing

$$\lambda(n) = \frac{1}{1 - \frac{\log (k/\gamma(k))}{\log n}},$$

from which it follows that

$$\lambda(n) = \frac{1}{1 - \frac{\log (k/\gamma(k))}{\log x}} + O\left(\frac{1}{x^{1/3}}\right),$$

whenever $n \in [x, x + \sqrt{x}\,]$ and $k(n) < (\log x)^\nu$, ν being an arbitrary positive integer.

First, we shall obtain an upper bound for the number of integers $n \in \mathscr{M}$, namely those integers which can be written in the form $n = p^2 a \in [x, x+\sqrt{x}\,]$ for some prime number $p > \sqrt{x}$. Observe that if $p^2 a \in [x, x+\sqrt{x}\,]$, then $p^2(a+1) > x+\sqrt{x}$, so that for each prime number p, there exists at most one positive integer a with this property.

We first count those integers $n = p^2 a$ for which $p < \frac{\sqrt{x}}{\log^R x}$. Their contribution is at most $C\frac{\sqrt{x}}{\log^{R+1} x}$ for some constant $C > 0$. On the other hand, if $p > \frac{\sqrt{x}}{\log^R x}$, then

$$a < \frac{(x+\sqrt{x})(\log^{2R} x)}{x} < 2\log^{2R} x.$$

Since $au^2 \in [x, x+\sqrt{x}\,]$ implies that $u \in \left[\sqrt{\frac{x}{a}}, \sqrt{\frac{x}{a}}\sqrt{1+\frac{1}{\sqrt{x}}}\right]$, and since the length of this interval is bounded, no more than $C\log^{2R} x$ such integers exist.

Gathering these estimates, we obtain that

$$\#\mathscr{M} \leqslant \frac{C\sqrt{x}}{\log^{2R} x} + C\log^{2R} x.$$

Now let $\mathscr{M}_1$ be the set of those integers $n = p^2 a \in [x, x+\sqrt{x}\,]$ for which $p^2 | k(n)$, where $(\log x)^\nu < p^2 < \sqrt{x}$. Then clearly

$$\#\mathscr{M}_1 \leqslant \sum_{(\log x)^\nu < p^2 < \sqrt{x}} \frac{\sqrt{x}}{p^2} \ll \frac{\sqrt{x}}{(\log x)^{\nu/2}}.$$

We now estimate

$$S(x) := \#\{n \in [x, x+\sqrt{x}] : k(n) > (\log x)^{\nu_1}\} = S_1(x) + S_2(x),$$

where in $S_1(x)$ we count those n for which there is a squarefull divisor $t | k(n)$ belonging to $[(\log x)^{\nu_1}, \sqrt{x}\,]$, the others being counted by $S_2(x)$. Repeating the argument used in the proof of Theorem 1 and using Lemma 1, we obtain that

$$S(x) \leqslant \frac{\sqrt{x}}{\log^{R+2} x},$$

provided ν_1 is taken large enough.

It remains to consider the sum

$$S_0(x) := \sum_{\substack{k \text{ squarefull} \\ k < (\log x)^{\nu_1}}} \frac{1}{1 - \frac{\log(k/\gamma(k))}{\log x}} \cdot \left(M\left(\frac{x+\sqrt{x}}{k}\middle| k\right) - M\left(\frac{x}{k}\middle| k\right)\right).$$

By (ii) of Lemma 2, we have

$$S_0(x) = \sum_{\substack{k \text{ squarefull} \\ k < (\log x)^{\nu_1}}} \frac{1}{1 - \frac{\log(k/\gamma(k))}{\log x}} \cdot \frac{6}{\pi^2}\frac{\rho(k)}{k}\sqrt{x} + O\left(\frac{\sqrt{x}}{\log^R x}(\log x)^{\nu_1+1}\right). \qquad (26)$$

The error term can be bounded by $\frac{\sqrt{x}}{\log^{r+1} x}$ provided we choose $R \geqslant (\nu_1+1)+r+1$. On the other hand, the main sum in (26) can be rewritten as

$$\sqrt{x}\left\{\vartheta_0(x) + \frac{\vartheta_1(x)}{\log x} + \cdots + \frac{\vartheta_r(x)}{\log^r x} + O\left(\frac{\vartheta_{r+1}(x)}{\log^{r+1} x}\right)\right\},$$

where

$$\vartheta_j(x) = \frac{6}{\pi^2} \sum_{\substack{k \text{ squarefull} \\ k < (\log x)^{\nu_1}}} (\log(k/\gamma(k)))^j \frac{\rho(k)}{k} \quad (j = 0, 1, 2, \ldots, r).$$

Set

$$\vartheta_j := \lim_{x\to\infty} \vartheta_j(x),$$

observing that indeed the limit exists and in fact that, using Lemma 1, we have

$$\vartheta_j(x) - \vartheta_j \ll \frac{(\log\log x)^j}{(\log x)^{\nu_1/2}}.$$

It is easy to see that $\vartheta_1 = 1$. Hence, we obtain that (4) holds with $c_i = \vartheta_i$ for $i = 2, 3, \ldots, r$. The proofs of Theorems 2 and 3 are thus complete.

Observe incidentally that

$$c_1 = \frac{6}{\pi^2} \sum_{k \text{ squarefull}} \frac{\rho(k)\log(k/\gamma(k))}{k} = \sum_p \frac{\log p}{p(p-1)} \approx 0.75536,$$

as in De Koninck and Doyon [1], and that one can easily show that $d_1 = -c_1$.

5.2. We now show how (5) follows from (4). For this purpose, let $E(x) := \sum_{n \leqslant x} \lambda(n)$, so that

$$\frac{E(x+\sqrt{x}) - E(x)}{\sqrt{x}} = \frac{1}{\sqrt{x}} \sum_{x < n \leqslant x+\sqrt{x}} \lambda(n) = \sum_{i=0}^{r} \frac{c_i}{\log^i x} + O\left(\frac{1}{\log^{r+1} x}\right) \tag{27}$$

with $c_0 = 1$, and therefore

$$\int_2^X \frac{E(x+\sqrt{x}) - E(x)}{\sqrt{x}}\, dx = \sum_{i=0}^{r} c_i \int_2^X \frac{dx}{\log^i x} + O\left(\frac{X}{\log^{r+1} X}\right). \tag{28}$$

On the other hand, since for $x \leqslant n \leqslant x + \sqrt{x}$, we have $\sqrt{n} = \sqrt{x}(1 + O(1/\sqrt{x}))$, so that

$$\frac{1}{\sqrt{x}} \sum_{x < n \leqslant x+\sqrt{x}} \lambda(n) = \sum_{x \leqslant n \leqslant x+\sqrt{x}} \frac{\lambda(n)}{\sqrt{n}} + O(x^{-1/2}\log x). \tag{29}$$

Now, again since $\sqrt{x} = \sqrt{n} + O(1)$ for $x \leqslant n \leqslant x + \sqrt{x}$, we have

$$\begin{aligned}
\int_2^X \sum_{x \leqslant n \leqslant x+\sqrt{x}} \frac{\lambda(n)}{\sqrt{n}}\, dx &= \sum_{2 \leqslant n \leqslant X+\sqrt{X}} \frac{\lambda(n)}{\sqrt{n}} \int_{n-\sqrt{n}+O(1)}^{n} dx \\
&= \sum_{2 \leqslant n \leqslant X+\sqrt{X}} \frac{\lambda(n)}{\sqrt{n}} (\sqrt{n} + O(1)) \\
&= \sum_{2 \leqslant n \leqslant X} \lambda(n) + O(\sqrt{X}\log X).
\end{aligned} \tag{30}$$

Finally, by partial integration, one easily obtains that, for each integer $i\in[1,r]$,

$$\int_e^X \frac{dx}{\log^i x} = \frac{X}{\log^i X} + i\frac{X}{\log^{i+1} X} + i(i+1)\frac{X}{\log^{i+2} X} + \cdots + O\left(\frac{X}{\log^{r+1} X}\right). \quad (31)$$

Gathering all estimates (27) through (31), estimate (5) follows, thus completing the proof of Theorem 2.

6. Proof of Theorem 4

Assume that $p\in[x, x+x^{2/3}]$ and write

$$p+1 = km, \quad (m,k)=1, \qquad \mu^2(m)=1, \quad k \text{ squarefull}.$$

Then, using the same approach as in the proof of Theorem 1, we can write

$$\lambda(p+1) = \frac{1}{1-\frac{\log(k/\gamma(k))}{\log(p+1)}}.$$

Now let ξ be a large constant. We shall first bound the set of those shifted primes $p+1\in[x, x+x^{2/3}]$ which have a squarefull divisor $t>(\log x)^{\xi}$. The contribution of those primes p with a corresponding $t<x^{2/3}$ can be estimated using a sieve approach. Indeed, letting $P(x)$ stand for the number of these primes p, we have

$$\begin{aligned} P(x) &\ll \sum_{(\log x)^{\xi}<t\leqslant\sqrt{x}} (\pi(x+x^{2/3};t,-1)-\pi(x;t,-1)) \\ &\quad + \sum_{\sqrt{x}<t<x^{2/3}} \left(\left[\frac{x+x^{2/3}}{t}\right] - \left[\frac{x}{t}\right]\right) \\ &\ll \frac{x^{2/3}}{\log x}\cdot\frac{1}{(\log x)^{\xi/2}} + O(x^{2/3-1/4}). \end{aligned}$$

It remains to estimate those shifted primes $p+1\in[x, x+x^{2/3}]$ for which there is a prime number q such that

$$p+1 = aq^2 \quad \text{with } q^2 > x^{2/3}. \quad (32)$$

To each prime q, there can only correspond one integer a. Moreover, it follows from (32) that $q<\sqrt{2x}$ and therefore that there can be at most $\sqrt{x}$ such shifted primes $p+1$.

We now consider the set $\mathscr{F}$ consisting of the shifted primes $p+1\in[x, x+x^{2/3}]$ for which the corresponding squarefull number k satisfies $k\leqslant(\log x)^{\xi}$. Given $p\in\mathscr{F}$, we have

$$\lambda(p+1) = \frac{1}{1-\frac{\log(k/\gamma(k))}{\log x}} + O\left(\frac{1}{x^{1/4}}\right),$$

say, and therefore

$$F := \sum_{p \in \mathscr{F}} \lambda(p+1) = \sum_{\substack{k \text{ squarefull} \\ k \leqslant (\log x)^{\xi}}} \frac{1}{1 - \frac{\log(k/\gamma(k))}{\log x}} (\Pi(x + x^{2/3}|k) - \Pi(x|k))$$
$$+ O(x^{2/3-1/4}).$$

Hence, by using Lemma 4, we get that

$$F = (\mathrm{li}(x + x^{2/3}) - \mathrm{li}(x)) \sum_{\substack{k \text{ squarefull} \\ k \leqslant (\log x)^{\xi}}} \frac{1}{k} \prod_{q \nmid k} \left(1 - \frac{1}{q(q-1)}\right) \frac{1}{1 - \frac{\log(k/\gamma(k))}{\log x}}$$
$$+ O\left(x^{2/3} \frac{\log\log x}{(\log x)^{r+2}}\right).$$

From here on, the proof is similar to that of Theorem 2 and we shall therefore omit it.

Incidentally, observe that α_2 can easily be computed and in fact that

$$\alpha_2 = \sum_{k \text{ squarefull}} \frac{\log(k/\gamma(k))}{k^2} \prod_{p \nmid k} \left(1 - \frac{1}{p(p-1)}\right) \approx 0.069.$$

7. Final Remarks

A key element in the proofs of Theorems 2 and 3 is the estimate (ii) of Lemma 2. But our results used only the fact that the error term in this estimate is $O\left(\frac{x}{\log^R x}\right)$ for any fixed $R > 0$. However, it should be mentioned that using the full force of the error term and at the cost of some additional computations, one can show that estimates (4) and (6) still hold when the indicated sums run over shorter intervals, such as for instance the interval $[x, x + \sqrt{x}\,\log^{-r} x]$.

Acknowledgements. The authors are indebted to Professor Michael Filaseta for allowing them to include here an important unpublished result (Theorem A). The authors are also grateful to the referees for valuable suggestions which helped to improve the final version of this paper.

References

[1] De Koninck JM, Doyon N (2003) À propos de l'indice de composition des nombres. Monatsh Math **139**: 151–167.
[2] Filaseta M, Trifonov O (1996) The distribution of fractional parts with applications to gap results in Number Theory. Proc London Math Soc **73**: 241–278
[3] Huxley M (1972) On the difference between consecutive primes. Invent Math **15**: 164–170
[4] Ivić A, Shiu P (1982) The distribution of powerful integers. Illinois J Math **26**: 576–590
[5] Nicolas JL, Robin G (1983) Majorations explicites pour le nombre de diviseurs de *n*. Bull Can Math **26**: 485–492
[6] Walfisz A (1963) Weylsche Exponentialsummen in der neueren Zahlentheorie. Berlin: Springer

Authors' addresses: Jean-Marie De Koninck, Département de mathématiques, Université Laval, Québec G1K 7P4, Canada, e-mail: jmdk@mat.ulaval.ca; Imre Kátai, Computer Algebra Department, Eötvös Loránd University, 1117 Budapest, Pázmány Péter Sétány I/C, Hungary, e-mail: katai@compalg.inf.elte.hu