Uniform Distribution Theory 4 (2009), no.2, 115–132

uniform distribution theory

UNIFORM DISTRIBUTION MODULO 1 OF THE HARMONIC PRIME FACTOR OF AN INTEGER

Imre Kátai — Florian Luca

ABSTRACT. Define the *harmonic mean prime factor of* n as the harmonic mean of the distinct prime factors of n. In this note, we show that the harmonic mean prime factor of n is uniformly distributed modulo 1.

Communicated by Michael Drmota

1. Introduction

Let n be a positive integer. Let $\Omega(n)$ and $\omega(n)$ count the number of prime divisors of n counted with and without multiplicity, respectively. The arithmetic properties of various averages of the prime factors of n has been studied in various recent papers. For example, it was shown in [1] that if one puts

$$p_a(n) = \frac{1}{\omega(n)} \sum_{p|n} p$$

for the arithmetic mean of the distinct prime factors of n, then $p_a(n)$ is uniformly distributed modulo 1. Furthermore, it was also shown that for large x the number of $n \le x$ such that $p_a(n)$ is an integer of order of magnitude $O(x/\log\log x)$. This was made more precise by the first author in [4], who showed that the counting function of the above n is asymptotically equal to $(c+o(1))x/\log\log x$ as x tends to infinity with some positive constant c which he did not compute. In [6], it was shown that the counting function of the composite $n \le x$ fulfilling the stronger condition that $p_a(n)$ is a prime factor of n is of size $x\exp(-c_x\sqrt{\log x \log\log x})$, where $1/\sqrt{2} + o(1) \le c_x \le 2 + o(1)$ as x tends to infinity. We expect similar

2000 Mathematics Subject Classification: 11K65, 11N67.
Keywords: Uniform distribution, values of arithmetic functions.

results to hold if one works with the function

$$p_A(n)=\frac{1}{\Omega(n)}\sum_{\substack{p^a|n\\ p^a>1}} p,$$

although we have not seen them anywhere in the literature. In [7], it was shown that both geometric mean prime factors $p_g(n)=(\prod_{p|n}p)^{1/\omega(n)}$ and $p_G(n)=n^{1/\Omega(n)}$ of n are uniformly distributed modulo 1. Observe that by unique factorization these numbers are integers if and only if n is a prime power. In this paper, we look at the functions

$$p_h(n)=\frac{\omega(n)}{\sum_{p|n}\frac{1}{p}} \qquad \text{and} \qquad p_H(n)=\frac{\Omega(n)}{\sum_{\substack{p^a|n\\ p^a>1}}\frac{1}{p}}$$

and show that these functions are uniformly distributed modulo 1, too. Observe also that none of these numbers is an integer except when n is a prime power.

Throughout we write $p(n)$ and $P(n)$ for the smallest and largest prime factors of n. We also put $\phi(n)$ and $\sigma(n)$ for the Euler function of n and sum of divisors function of n, respectively. We use the Landau symbols O and o and the Vinogradov symbols $\gg$ and $\ll$ with their usual meanings. We write $\log x$ for the natural logarithm of x and for a positive integer k we let $\log_k x$ to be the recursively defined function given by $\log_1 x=\max\{1,\log x\}$ and $\log_k x=\max\{1,\log\log_{k-1}x\}$ for $k>1$.

2. Main results

In this paper, we prove the following theorem.

Theorem 1. *Let $g(n)$ be an additive function such that $g(p)<c_1/p$ and $0<g(p^a)<c_2$ for all primes p and positive integers a with some positive constants c_1 and c_2. Let*

$$\nu(n)=\frac{\omega(n)}{g(n)} \qquad \text{and} \qquad \rho(n)=\frac{\omega(n+1)}{g(n)}.$$

Then

(i) *$\nu(n)$ is uniformly distributed modulo 1;*

(ii) *$\rho(n)$ is uniformly distributed modulo 1.*

The same holds when $\omega(n)$ is replaced by $\Omega(n)$.

Our theorem applies to the functions $g(n) = \sum_{p|n} 1/p$, $g(n) = \sum_{\substack{p^a|n \\ p>1}} 1/p$, $g(n) = \log(n/\phi(n))$ and $g(n) = \log(\sigma(n)/n)$, from where we deduce, in particular, the uniform distribution modulo 1 of the harmonic mean prime factors.

We close this section by describing the ideas behind the proof of Theorem 1. In [1] and [7] the uniform distributions modulo 1 of $p_a(n)$ and $p_g(n)$ were proved in the following way. Most integers n have a large $P(n)$ which appears with exponent 1 in their prime factorization. Thus, we can write $n = Pm$, where $P > P(m)$. Fix m. Then both

$$p_a(n) = \frac{\omega(m)p_a(m) + P}{\omega(m)+1} \qquad \text{and} \qquad p_g(n) = p_g(m)^{\omega(m)/(\omega(m)+1)} P^{1/(\omega(m)+1)}$$

are very sensitive to variations in P since P varies in a large interval. Fixing m, it was then argued that $p_a(n)$ and $p_g(n)$ were uniformly distributed modulo 1 via H. Weyl's criterion using exponential sums in [1] for the function $p_a(n)$ and in a direct way from the definition of uniform distribution for $p_g(n)$ in [7]. In fact, that method is strong enough to give nontrivial upper bounds on the discrepancy of these sequences.

The current problem is different since

$$p_h(m) = \frac{\omega(m)+1}{\omega(m)/p_h(m) + 1/P}$$

is not sensitive to variations in P. In fact, the difference $p_h(n) - p_h(m)$ approaches zero as n tends to infinity in a set of asymptotic density 1, which explains why the methods of [1] and [7] do not apply to the current problem. In fact, quite oppositely, the value of $p_h(n)$ is controlled by the small primes in n not by the large primes in n. So, in the next section, we shall prove that the values of $\nu(n)$ and $\rho(n)$ appearing in Theorem 1 are controlled by the prime factors of n less than $\log_2 n$ on a set of n of asymptotic density 1. Along the way, we are also led naturally to consider the problem of estimating the number of $n \le x$, where x is a large real number, which have a fixed number k of distinct prime factors in the normal interval for the function $\omega(n)$ and which additionally are free of small primes, which is what we do next.

3. Integers free of small primes with a fixed number of prime factors

We let x be a large real number. We let $2 \le y < x$ be some parameter which tends to infinity with x and which will be specified later. For a positive integer

n we put

$$A_y(n) = \prod_{\substack{p^a \| n \\ p \le y}} p^a \qquad \text{and} \qquad B_y(n) = \prod_{\substack{p^a \| n \\ p > y}} p^a.$$

For a positive integer k we write

$$N_k(x) = \#\{n \le x : \Omega(n) = k\} \qquad \text{and} \qquad \Pi_k(x) = \#\{n \le x : \omega(n) = k\}.$$

We put

$$\begin{aligned} N_k(x|y) &= \#\{n \le x : \Omega(n) = k \text{ and } p(n) > y\}, \\ \Pi_k(x|y) &= \#\{n \le x : \omega(n) = k \text{ and } p(n) > y\}. \end{aligned}$$

Now let $y = O(\log_2 x)$. Put $P_y = \prod_{p \le y} p$. Since for any complex number z we have

$$\sum_{\substack{n \le x \\ \gcd(n, P_y) = 1}} z^{\Omega(n)} = \sum_{n \le x} z^{\Omega(n)} \sum_{d | \gcd(n, P_y)} \mu(d) = \sum_{d | P_y} \mu(d) \sum_{dm \le x} z^{\Omega(dm)},$$

we obtain, by identifying coefficients and observing that $\omega(d) = \Omega(d)$ for all $d \mid P_y$ since P_y is square-free, the identity

$$N_k(x|y) = \sum_{d | P_y} \mu(d) N_{k - \omega(d)}(x/d). \tag{1}$$

In the above formula, we understand that $N_i(x)$ is 1 or 0 whenever i is zero or negative, respectively. The largest value of $\omega(d)$ and d occurring in the right hand side in formula (1) above is

$$\pi(y) \le \frac{2y}{\log y} = O\left(\frac{\log_2 x}{\log_3 x}\right) \qquad \text{and} \qquad P_y \le 4^y < e^{2y} = (\log x)^{O(1)},$$

respectively.

For suitable values of k, the summands $N_{k-\omega(d)}(x/d)$ will be estimated by using the formula of Sathe-Selberg (see [8]), namely that

$$\begin{aligned} \Pi_k(x) &= \frac{x}{\log x} \frac{(\log_2 x)^{k-1}}{(k-1)!} \left(1 + O\left(\frac{1}{\log_2 x}\right)\right) \\ N_k(x) &= \frac{x}{\log x} \frac{(\log_2 x)^{k-1}}{(k-1)!} \left(1 + O\left(\frac{1}{\log_2 x}\right)\right) \end{aligned} \tag{2}$$

provided that $1 \le k \le (2 - \delta) \log_2 x$, where $\delta > 0$ can be arbitrarily small but otherwise fixed. The constants implied in the above O's depend on δ.

Now, let $z = O(\log_4 x)$ and let k be such that $|k - \log_2 x| < z\sqrt{\log_2 x}$. Then for large x the condition $k < (2-\delta)\log_2 x$ is satisfied with $\delta = 0.5$. Now observe that for $d \mid P_y$, we have

$$\begin{aligned}\log_2(x/d) &= \log(\log x - \log d) = \log_2 x + O\left(\frac{\log d}{\log x}\right)\\ &= \log_2 x + O\left(\frac{\log_2 x}{\log x}\right) = \log_2 x + o(1) \quad \text{as} \quad x \to \infty,\end{aligned}$$

so the condition

$$k - \omega(d) < (2-\delta)\log_2(x/d)$$

holds for all the numbers k that we are considering and for all $d \mid P_y$ with $\delta = 0.4$ once x is sufficiently large. Observe that for such k and d we also have that

$$\begin{aligned}k - \omega(d) - 1 &= \log_2 x + O(z\sqrt{\log_2 x} + \pi(y)) = \log_2 x + O\left(\frac{\log_2 x}{\log_3 x}\right)\\ &= (1 + o(1))\log_2 x \quad \text{as} \quad x \to \infty, \end{aligned} \tag{3}$$

so, in particular, all the numbers $k - \omega(d) - 1$ are positive. Thus, according to (2), we have

$$N_{k-\omega(d)}(x/d) = \frac{x}{d\log(x/d)} \frac{(\log\log(x/d))^{k-\omega(d)-1}}{(k-\omega(d)-1)!}\left(1 + O\left(\frac{1}{\log_2 x}\right)\right). \tag{4}$$

Observe that

$$\frac{1}{\log(x/d)} = \frac{1}{\log x}\left(1 + O\left(\frac{\log d}{\log x}\right)\right) = \frac{1}{\log x}\left(1 + O\left(\frac{\log_2 x}{\log x}\right)\right), \tag{5}$$

while

$$\begin{aligned}(\log_2(x/d))^{k-\omega(d)-1} &= (\log_2 x)^{k-\omega(d)-1}\left(1 + O\left(\frac{\log d}{\log x}\right)\right)^{O(\log_2 x)}\\ &= (\log_2 x)^{k-\omega(d)-1}\left(1 + O\left(\frac{(\log_2 x)^2}{\log x}\right)\right).\end{aligned} \tag{6}$$

Thus, using the estimates (2), (4), (5) and (6) and observing that the error term in (2) dominates the error terms in estimates (5) and (6), we get

$$N_{k-\omega(d)}(x/d) = \frac{x}{d\log x}\frac{(\log_2 x)^{k-\omega(d)-1}}{(k-\omega(d)-1)!}\left(1+O\left(\frac{1}{\log_2 x}\right)\right). \tag{7}$$

Now we put

$$Y_1 = \frac{5\sqrt{\log_2 x}}{(\log_4 x)^2} \qquad \text{and} \qquad Y_2 = \frac{10\log_3 x}{\log_4 x},$$

and split the set of divisors of d of P_y into three subsets as follows:

$$\begin{aligned} \mathcal{D}_1 &= \{d \mid P_y : \omega(d) > Y_1\}, \\ \mathcal{D}_2 &= \{d \mid P_y : Y_2 < \omega(d) \le Y_1\}, \\ \mathcal{D}_3 &= \{d \mid P_y : \omega(d) \le Y_2\}. \end{aligned} \tag{8}$$

Observe that by the multinomial formula we have

$$S_i = \sum_{d\in\mathcal{D}_i}\frac{1}{d} < \sum_{\ell>Y_i}\frac{1}{\ell!}\left(\sum_{p\le y}\frac{1}{p}\right)^{\ell} \tag{9}$$

for all $i = 1$, 2 and 3, where we put $Y_3 = 0$. By Mertens' formula, we have

$$\sum_{p\le y}\frac{1}{p} = \log_2 y + O(1),$$

therefore in the sum appearing in the right hand side of (9), the ratio between the term corresponding to $\ell+1$ and the one corresponding to ℓ is

$$\frac{1}{\ell+1}\left(\sum_{p<y}\frac{1}{p}\right) \le \frac{\log_4 x + O(1)}{Y_i} < \frac{1}{10}$$

for large x and both $i = 1$ and $i = 2$, which shows that the first term dominates the entire sum. Thus, using also the elementary estimate $L! \ge (L/e)^L$ with $L = L_i = \lfloor Y_i \rfloor + 1$ and $i = 1$ and 2, we get

$$S_i \ll \frac{1}{L_i!}\left(\sum_{p\le y}\frac{1}{p}\right)^{L_i} \le \left(\frac{e\log_4 x + O(1)}{L_i}\right)^{L_i} < \exp\left(-0.5Y_i\log Y_i\right)$$

for x sufficiently large and for $i = 1$ and 2 since for large values of x we have

$$\frac{e\log_4 x + O(1)}{L_i} < \frac{1}{L_i^{1/2}} \qquad \text{for both } i = 1,\ 2.$$

In particular, we get

$$\begin{aligned} S_1 &\le \exp\left(-\frac{\sqrt{\log_2 x}\,\log_3 x}{(\log_4 x)^2}\right), \\ S_2 &\le \frac{1}{\log_2 x}, \end{aligned} \tag{10}$$

uniformly in our range for k. For S_3, we just use the trivial estimate

$$S_3 \le \sum_{d|P_y} \frac{1}{d} = \prod_{p\le y}\left(1+\frac{1}{p}\right) \ll \log y.$$

Now we use the above sums to estimate $N_{k-\omega(d)}(x/d)$ depending on which of the three subsets $\mathcal{D}_i$ for $i = 1,2,3$ contains d. When $d \in \mathcal{D}_1$, we use the trivial fact that $N_{k-\omega(d)}(x/d) \le x/d$ to get

$$\sum_{d\in\mathcal{D}_1} N_{k-\omega(d)}(x/d) \le x\sum_{d\in\mathcal{D}_1}\frac{1}{d} = xS_1 \le x\exp\left(-\frac{\sqrt{\log_2 x}\,\log_3 x}{\log_4 x}\right). \tag{11}$$

We show that in our range the above right hand side is smaller than

$$N_k(x)/(\log_2 x)^2.$$

Indeed, observe that by the Stirling formula for approximating the factorial, the estimates

$$\begin{aligned} \frac{N_k(x)}{(\log_2 x)^2} &\gg \frac{x}{\log x(\log_2 x)^{5/2}} e^k \left(\frac{\log_2 x}{k-1}\right)^{k-1} \\ &= \frac{x}{(\log_2 x)^{5/2}} e^{O(|k-\log_2 x|)}\left(1+O\left(\frac{|k-\log_2 x|}{\log_2 x}\right)\right)^{O(\log_2 x)} \\ &= \frac{x}{(\log_2 x)^{5/2}} e^{O(\sqrt{\log_2 x}\,\log_4 x)} > x\exp\left(-c_1\sqrt{\log_2 x}\,\log_4 x\right) \end{aligned} \tag{12}$$

hold for large x with some appropriate positive constant c_1. Comparing estimates (11) and (12), we see that estimate

$$\sum_{d\in\mathcal{D}_1} N_{k-\omega(d)}(x/d) = O\left(\frac{N_k(x)}{(\log_2 x)^2}\right) \tag{13}$$

holds. We now study the values $d \in \mathcal{D}_2$. Observe that for such values of d we have

$$\begin{aligned}
\frac{(\log_2 x)^{k-\omega(d)-1}}{(k-\omega(d)-1)!} &= \frac{(\log_2 x)^{k-1}}{(k-1)!} \times \left(\frac{k}{\log_2 x}\right)^{O(\omega(d))} \\
&= \frac{(\log_2 x)^{k-1}}{(k-1)!} \left(1 + O\left(\frac{|k-\log_2 x|}{\log_2 x}\right)\right)^{O(Y_1)} \\
&= \frac{(\log_2 x)^{k-1}}{(k-1)!} \exp\left(O\left(\frac{z\sqrt{\log_2 x}Y_1}{\log_2 x}\right)\right) \\
&= \frac{(\log_2 x)^{k-1}}{(k-1)!} \exp\left(O\left(\frac{1}{(\log_4 x)^2}\right)\right) \\
&= (1+o(1))\frac{(\log_2 x)^{k-1}}{(k-1)!} \qquad \text{as} \qquad x \to \infty. \qquad (14)
\end{aligned}$$

The above calculation together with estimate (4) shows that

$$N_{k-\omega(d)}(x/d) \ll \frac{x}{d \log x} \frac{(\log_2 x)^{k-1}}{(k-1)!} \ll \frac{N_k(x)}{d}.$$

Thus, using estimate (10) we get

$$\sum_{d \in \mathcal{D}_2} N_{k-\omega(d)}(x/d) \ll N_k(x) \sum_{d \in \mathcal{D}_2} \frac{1}{d} \ll N_k(x) S_2 \ll \frac{N_k(x)}{\log_2 x}. \qquad (15)$$

Now we deal with $d \in \mathcal{D}_3$. We repeat calculation (14), but for $d \in \mathcal{D}_3$ we get

$$\begin{aligned}
\frac{(\log_2 x)^{k-\omega(d)-1}}{(k-\omega(d)-1)!} &= \frac{(\log_2 x)^{k-1}}{(k-1)!} \left(1 + O\left(\frac{|k-\log_2 x|}{\log_2 x}\right)\right)^{O(Y_2)} \\
&= \frac{(\log_2 x)^{k-1}}{(k-1)!} \exp\left(O\frac{z\sqrt{\log_2 x}Y_2}{\log_2 x}\right) \\
&= \frac{(\log_2 x)^{k-1}}{(k-1)!} \left(1 + O\left(\frac{\log_3 x}{(\log_2 x)^{1/2}}\right)\right),
\end{aligned}$$

which via estimate (10) leads to the conclusion that

$$N_{k-\omega(d)}(x/d) = \frac{N_k(x)}{d} \left(1 + O\left(\frac{\log_3 x}{(\log_2 x)^{1/2}}\right)\right) \qquad \text{for} \qquad d \in \mathcal{D}_3. \qquad (16)$$

Going back to formula (1), and using estimates (13), (15) and (16), we get that

$$\begin{aligned} N_k(x|y) &= N_k(x)\sum_{d\in\mathcal{D}_3}\frac{\mu(d)}{d}+O\left(\frac{N_k(x)}{\log_2 x}+\left(\sum_{d\in\mathcal{D}_3}\frac{1}{d}\right)\frac{N_k(x)\log_3 x}{(\log_2 x)^{1/2}}\right)\\ &= N_k(x)\sum_{d|P_y}\frac{\mu(d)}{d}+O\left(N_k(x)\left(\frac{1}{\log_2 x}+\frac{\log_3 x\log y}{(\log_2 x)^{1/2}}+S_2+S_3\right)\right)\\ &= N_k(x)\left(\prod_{p\le y}\left(1-\frac{1}{p}\right)+O\left(\frac{(\log_3 x)^2}{(\log_2 x)^{1/2}}\right)\right). \qquad (17)\end{aligned}$$

Observe that since

$$\prod_{p\le y}\left(1-\frac{1}{p}\right)\gg\frac{1}{\log y}\gg\frac{1}{\log_3 x},$$

the above estimate (17) is in fact an asymptotic.

We now turn our attention to $\Pi_k(x|y)$. If n is counted by $\Pi_k(x|y)$ but not by $N_k(x|y)$, then there is a prime $p>y$ such that $p^2\mid n$ and n/p^2 is counted by one of $\Pi_\ell(x/p^2)$ for some $\ell\in\{k-1,k\}$. In particular,

$$\#\{n\le x:\omega(n)=k,\ p(n)>y \text{ and } \Omega(n)>k\}\le\sum_{\ell=k-1}^{k}\sum_{y<p\le\sqrt{x}}\Pi_\ell(x/p^2).$$

In the range $p<\log x$, we have that the inequality $\ell<(2-\delta)\log_2(x/p^2)$ holds for all our numbers $\ell\in\{k-1,k\}$ and primes p with $\delta=0.3$ and the condition $y=O(\log_2(x/p^2))$ still holds once x is large. Applying now estimate (2), we get that

$$\Pi_\ell(x/p^2)=\frac{x}{p^2\log(x/p^2)}\frac{(\log_2(x/p^2))^{\ell-1}}{(\ell-1)!}.$$

However, $\log(x/p^2)=\log x-2\log p=\log x+O(\log_2 x)=(1+o(1))\log_2 x$ as $x\to\infty$, while

$$\begin{aligned}(\log_2(x/p^2))^{\ell-1} &= (\log(\log x+O(\log_2 x)))^{\ell-1}\\ &= \left(\log_2 x+O\left(\frac{\log_2 x}{\log x}\right)\right)^{\ell-1}\\ &= (\log_2 x)^{\ell-1}\left(1+O\left(\frac{1}{\log x}\right)\right)^{O(\ell)}\\ &= (\log_2 x)^{\ell-1}\left(1+O\left(\frac{k}{\log x}\right)\right)\\ &= (1+o(1))(\log_2 x)^{\ell-1} \quad \text{as} \quad x\to\infty.\end{aligned}$$

Thus,

$$\Pi_\ell(x/p^2) \ll \frac{x}{p^2 \log x} \frac{(\log_2 x)^{\ell-1}}{(\ell-1)!} \ll \frac{N_k(x)}{p^2}.$$

In the range $p \geq \log x$ we use the trivial estimate $N_k(x/p^2) \leq x/p^2$. Thus, we get

$$\#\{n \leq x : \ \omega(n) = k, \ p(n) > y \ \text{ and } \mu(n) = 0\} \ll N_k(x) \sum_{y<p\leq \log x} \frac{1}{p^2}$$
$$+ x \sum_{\log x \leq p \leq x^{1/2}} \frac{1}{p^2} \ll \frac{N_k(x)}{y \log y} + \frac{x}{\log x \log_2 x}.$$

The term $x/(\log x \log_2 x)$ above is of a much smaller order of magnitude than $N_k(x)/(y(\log y))$ (compare with (12), for example), showing, via the asymptotic estimate (17), that in fact

$$|N_k(x|y) - \Pi_k(x|y)| \ll \frac{N_k(x|y)}{y \log y}.$$

Thus we get

$$\Pi_k(x|y) = N_k(x|y)\left(1 + O\left(\frac{1}{y \log y}\right)\right).$$

Therefore we have proved the following statement.

Lemma 1. *Let $x \geq 2$, $y = O(\log_2 x)$, $z = O(\log_4 x)$ and $|k - \log_2 x| < z\sqrt{\log_2 x}$. Then*

$$N_k(x|y) = N_k(x)\left(\prod_{p\leq y}\left(1 - \frac{1}{p}\right) + O\left(\frac{(\log_3 x)^2}{(\log_2 x)^{1/2}}\right)\right),$$
$$\Pi_k(x|y) = N_k(x|y)\left(1 + O\left(\frac{1}{y(\log y)}\right)\right).$$

4. The proof of Theorem 1

We let x large and put $y = \log_2 x$ and $z = \log_4 x$. We start by discarding some positive integers $n \leq x$. We discard the set $\mathcal{N}_1$ of positive integers $n \leq x$ such that $p^2 \mid n$ for some prime $p > y$. For a fixed prime, the number of such positive integers n is at most $\lfloor x/p^2 \rfloor < x/p^2$, so that

$$\#\mathcal{N}_1 \leq \sum_{y<p\leq\sqrt{x}} \frac{x}{p^2} \ll \frac{x}{y(\log y)} = o(x) \tag{18}$$

as $x \to \infty$. Next we discard the set $\mathcal{N}_2$ of positive integers $n \le x$ for which we have $|\omega(n) - \log_2 x| > z\sqrt{\log_2 x}$. By the Turán-Kubilius estimate

$$\sum_{n \le x} (\omega(n) - \log_2 x)^2 = O(x \log_2 x),$$

it follows easily that

$$\#\mathcal{N}_2 \ll \frac{x}{z^2} = o(x) \tag{19}$$

as $x \to \infty$. Next we let $\mathcal{N}_3$ be the set of $n \le x$ not in $\mathcal{N}_1 \cup \mathcal{N}_2$ such that $\omega(A_y(n)) > 10 \log_2 y$. For such n, it follows that if we write $\tau(m)$ for the number of divisors of m, then we have $\tau(A_y(n)) > 2^{10 \log_2 y} > (\log y)^6$ because $10 \log 2 > 6$. However, observe that

$$\begin{aligned}\sum_{n \le x} \tau(A_y(n)) &= \sum_{n \le x} \sum_{\substack{d \mid n \\ P(d) \le y}} 1 = \sum_{\substack{d \le x \\ P(d) \le y}} \left\lfloor \frac{x}{d} \right\rfloor \le x \sum_{P(d) \le y} \frac{1}{d} = x \prod_{p \le y} \left(1 - \frac{1}{p}\right)^{-1} \\ &\ll x \log y,\end{aligned}$$

from where we immediately deduce that

$$\#\mathcal{N}_3 \ll \frac{x}{(\log y)^5} = o(x) \tag{20}$$

as $x \to \infty$. Observe that when $n \notin \mathcal{N}_3$, we have $A_y(n) < y^{10 \log_2 y}$. Moreover, if also $n \notin \mathcal{N}_1$, then $n = A_y(n)m$, where m is square-free and free of primes $\le y$. Furthermore, $\omega(m) = \omega(n) - \omega(A_y(n)) > \omega(n) - 10 \log_4 x$, while $m \le x/A_y(n)$ and

$$\begin{aligned}\log_2(x/A_y(n)) &= \log_2 x + O\left(\frac{\log(A_y(n))}{\log x}\right) = \log_2 x + O\left(\frac{\log_2 y \log y}{\log x}\right) \\ &= \log_2 x + o(1)\end{aligned}$$

as $x \to \infty$. From the above estimates we deduce that if additionally $n \in \mathcal{N}_2$, then the inequality

$$|\omega(m) - \log_2(x/A_y(n))| < 2z\sqrt{\log_2(x/A_y(n))} \tag{21}$$

holds. Next we let $g_y(n)$ to be the additive function given by $g_y(p^a) = g(p^a)$ if $p \le y$ and be zero otherwise. We put $h_y(n) = g(n) - g_y(n)$. Observe that

$$\begin{aligned}\sum_{n\le x} h_y(n) &= \sum_{n\le x}\sum_{\substack{p^a\| n\\ p>y}} g(p^a) \le \sum_{\substack{p^a\le x\\ p>y}} g(p^a)\sum_{\substack{n\le x\\ p^a|n}} 1 \le \sum_{\substack{p^a\le x\\ p>y}} g(p^a)\frac{x}{p^a}\\ &\le x\sum_{y<p}\frac{g(p)}{p} + x\sum_{\substack{y<p\\ a\ge 2}}\frac{g(p^a)}{p^a} \ll \frac{x}{y\log y} = \frac{x}{\log_2 x\log_3 x},\end{aligned}$$

which shows that if we put $\mathcal{N}_4$ for the set of $n \le x$ such that

$$h_y(n) > \frac{1}{\log_2 x\sqrt{\log_3 x}},$$

then

$$\#\mathcal{N}_4 \ll \frac{x}{\sqrt{\log_3 x}} = o(x) \tag{22}$$

as $x \to \infty$. From now on, we assume that $n \notin \bigcup_{i=1}^4 \mathcal{N}_i$. Thus, if we put

$$\nu_y(n) = \frac{\omega(n)}{g_y(n)},$$

then

$$\nu_y(n) - \nu(n) = \frac{\omega(n)}{g_y(n)} - \frac{\omega(n)}{g(n)} = \frac{h_y(n)\omega(n)}{g(n)g_y(n)} \ll \frac{1}{g(n)g_y(n)\sqrt{\log_3 x}}.$$

Since most positive integers n have a "small prime power", where by "small" we mean smaller than any function tending to infinity arbitrarily slowly, and $g(p^a)$ is positive for all prime powers p^a, it follows easily that if we put $\mathcal{N}_5$ for the set of positive integers $n \le x$ such that

$$g(n)g_y(n) < \frac{1}{(\log_3 x)^{1/4}},$$

then $\#\mathcal{N}_5 = o(x)$ as $x \to \infty$. Thus, if $n \notin \mathcal{N}_5$, then

$$0 \le \nu_y(n) - \nu(n) \le \frac{1}{(\log_3 x)^{1/4}} = o(1) \tag{23}$$

as $x \to \infty$.

We are now ready to prove part (i) of Theorem 1. We fix any positive integer t. In order to show that $\nu(n)$ is uniformly distributed modulo 1 it suffices, via H. Weyl's criterion, to show that

$$T = \sum_{n\le x} \mathbf{e}(t\nu(n)) = o(x) \qquad \text{as} \qquad x \to \infty,$$

where, as usual, we put $\mathbf{e}(\alpha) = \exp(2\pi i\alpha)$. Observe that by the above remarks we have that

$$T = \sum_{\substack{n \le x \\ n \notin \bigcup_{i=1}^4 \mathcal{N}_i}} \mathbf{e}(t\nu_y(n)) + o(x) \qquad \text{as} \qquad x \to \infty,$$

that is, in the above exponential sum T we may replace $\nu(n)$ by $\nu_y(n)$ and we may discard sets of $n \le x$ of cardinality $o(x)$ as $x \to \infty$. By the Erdős-Wintner theorem (see [3]), we know that the function $g(n)$ has a limiting distribution; that is,

$$F(s) = \frac{1}{x}\#\{n \le x : g(n) < s\}$$

exists for any real number s. By Lévy's theorem (see Lévy's original paper [5] or Lemma 1.22 on page 46 in [2]), we know that F is continuos and $F(0) = 0$. Define the function

$$H_t(s) = \lim_{x\to\infty} \#\{n \le x : t/(g(n) < s\} = 1 - F(t/s).$$

Since $F(0) = 0$, F is continuous and strictly monotonic over the nonnegative reals, therefore for each $\varepsilon > 0$ there exist $\delta > 0$ and $K < \infty$ such that

$$\limsup_{x\to\infty} \frac{1}{x}\#\{n \le x : g(n) \notin [\delta, K]\} \le \varepsilon.$$

Let δ_1 be so small such that

$$\sum_{m=\lfloor t/K \rfloor}^{\lfloor t/\delta \rfloor} (H_t(m + \delta_1) - H_t(m - \delta_1)) < \varepsilon.$$

We now get, by denoting with $\|s\|$ the distance from s to the nearest integer, that

$$\limsup_{x\to\infty} \frac{1}{x}\#\{n \le x : \|t/g(n)\| < \delta_1 \text{ and } g(n) \in [\delta, K]\} < 2\varepsilon.$$

Using also estimate (23), it follows that

$$\limsup_{x\to\infty} \frac{1}{x}\#\{n \le x : \|t/g(A_y(n))\| < \delta_1/2 \text{ and } g(n) \in [\delta, K]\} < 3\varepsilon$$

once x is sufficiently large (depending on ε). Thus, for x large, putting $\mathcal{N}_6$ for the set of $n \le x$ such that either $g(n) \notin [\delta, K]$ or $\|t/g(A_y(n))\| < \delta_1/2$, then the cardinality of $\bigcup_{i=1}^6 \mathcal{N}_i$ is $< 4\varepsilon x$. Put

$$\mathcal{S}_y = \{A_y(n) : n \le x \text{ and } n \notin \bigcup_{i=1}^6 \mathcal{N}_i\}.$$

For each number $A \in \mathcal{S}_y$, we put $\eta_A = t/g(A)$, and

$$\mathcal{B}(A) = \{n \le x : n = Am \text{ where } p(m) > y \text{ and } \mu(m)^2 = 1\}.$$

Observe that for $n \in \mathcal{B}(A)$ we have

$$\nu_y(n) = \frac{\omega(n)}{g_y(n)} = \frac{\omega(A) + \omega(m)}{g_y(A)}.$$

Thus,

$$T = \sum_{A \in \mathcal{S}_y} \sum_{m \in \mathcal{B}(A)} \mathbf{e}(t\nu_y(n)) = \sum_{A \in \mathcal{S}_y} \mathbf{e}(\eta_A \omega(A)) \sum_{n \in \mathcal{B}(A)} \mathbf{e}(\eta_A \omega(m)) + O(\varepsilon x).$$

Since $\varepsilon > 0$ is arbitrary, it certainly suffices to show that the inner sums satisfy

$$\sum_{n \in \mathcal{B}(A)} \mathbf{e}(\eta_A \omega(m)) = o(\#\mathcal{B}(A)) \tag{24}$$

as $x \to \infty$ uniformly in $A \in \mathcal{S}_y$. To this end, we split the inner sum according to the number of distinct prime divisors of m. We get

$$\begin{aligned}
\sum_{n \in \mathcal{B}(A)} \mathbf{e}(\eta_A \omega(m)) &= \sum_{|k - \log_2 x| < z\sqrt{\log_2 x}} \mathbf{e}(k\eta_A)\Pi_k(x/A|y) + o(\#\mathcal{B}(A)) \\
&\ll \sum_{|k - \log_2(x/A)| < 2z\sqrt{\log_2(x/A)}} \mathbf{e}(k\eta_A) N_k(x/A|y) + o(\#\mathcal{B}(A)) \quad \text{as} \quad x \to \infty
\end{aligned}$$

(see Lemma 1 and formula (21)). The crucial point which will take care of the problem is to observe that the function $N_k(x|y)$ does not vary much "locally" in the parameter k in our range. Indeed, if $\ell \le z$, then, by Lemma 1 we have

$$\begin{aligned}
N_{k+\ell}(x|y) &= \frac{(1 + o(1))}{\log y} N_k(x) = \frac{x(1 + o(1))}{\log x \log y} \frac{(\log_2 x)^{k+\ell-1}}{(k + \ell - 1)!} \\
&= (1 + o(1)) N_k(x|y) \left(1 + O\left(\frac{|k + \ell - \log_2 x|}{\log_2 x}\right)\right)^{\ell} \\
&= (1 + o(1)) N_k(x|y) \left(1 + O\left(\frac{z^2}{(\log_2 x)^{1/2}}\right)\right) \\
&= (1 + o(1)) N_k(x|y) \quad \text{as} \quad x \to \infty,
\end{aligned}$$

and the same estimates hold with x replaced by x/A for all values of $A \in \mathcal{S}_y$. We split the interval $|k - \log_2(x/A)| < 2z\sqrt{\log_2(x/A)}$ into subintervals $[k_i, k_{i+1}]$ of length $L = \lfloor z \rfloor$ for $i = 1, 2, \ldots, i_A$, except that the last one might be shorter.

Clearly, $i_A \ll \sqrt{\log_2 x}$. Then we get

$$\begin{aligned} T &\ll \sum_{i=1}^{i_A} N_{k_i}(x|y) \sum_{k_i \le k < k_i+1} \mathbf{e}(k\eta_A) + o(\#\mathcal{B}(A)) \\ &= \sum_{i=1}^{i_A} N_{k_i}(x|y)\mathbf{e}(k_i\eta_A)\left(\frac{\mathbf{e}(k_{i+1}-k_i)\eta_A)-1}{\mathbf{e}(\eta_A)-1}\right) + o(\#\mathcal{B}(A)) \\ &\ll \delta_1^{-1}\sum_{i=1}^{i_A} N_{k_i}(x|y) + o(\#\mathcal{B}(A)) \\ &= \frac{1}{\delta_1 L}\sum_{|k-\log_2(x/A)|<2z\sqrt{\log_2(x/A)}} N_k(x|y) + o(\#\mathcal{B}(A)) \\ &= \frac{1}{\delta_1 z}\sum_{|k-\log_2(x/A)|<2z\sqrt{\log_2(x/A)}} \Pi_k(x|y) + o(\#\mathcal{B}(A)) \\ &= \frac{1}{\delta_1 z}\#\mathcal{B}(A) + o(\#\mathcal{B}(A)) = o(\#\mathcal{B}(A)) \quad \text{as} \quad x \to \infty, \end{aligned}$$

which completes the proof of the estimate (24) and, in particular, of the fact that $\nu(n)$ is uniformly distributed modulo 1. It is also clear that the argument works if instead of $\omega(n)$ in the definition of $\nu(n)$ we take $\Omega(n)$. In fact, the result remains also true if we replace $\omega(n)$ by $\log \tau_r(n)$ for some integer $r \ge 2$, where

$$\tau_r(n) = \#\{n = x_1 \cdots x_r : x_i \text{ positive integers for all } i = 1, \ldots, r\}.$$

Now we prove part (ii) of Theorem 1. We work with the same subsets $\mathcal{N}_i$ for $i = 1, \ldots, 7$ and with $n \le x$ not in these subsets. We also ask that $n+1$ does not belong to $\bigcup_{i=1}^{3} \mathcal{N}_i$. Furthermore, instead of the function $\nu_y(n)$ we work with the function

$$\rho_y(n) = \frac{\omega(n+1)}{g_y(n)}.$$

In particular, $\mathcal{N}_5$ will now consist of the positive integers $n \le x$ such that the condition (23) holds with the function $\nu(n)$ replaced by the function $\rho(n)$. We write

$$n = Am \qquad \text{and} \qquad n+1 = A_1 m_1 \tag{25}$$

for some coprime $A,\ A_1 \in \mathcal{S}_y$ one of which is even and some $m \le x/A$, $m_1 \le x/A_1$ with $p(mm_1) > y$. We let $\mathcal{B}(A, A_1)$ be the set of $m \le x/A$ coprime to P_y such that with $n = Am$ we have $n+1 = A_1 m_1$ for some m_1 coprime to P_y. Following the arguments for the proof of the part (i), it follows easily that it suffices to show that the estimate

$$\sum_{m \in \mathcal{B}(A,A_1)} \mathbf{e}\left(\frac{t\omega(Am+1)}{g(A)}\right) = o(\#\mathcal{B}(A,A_1))$$

holds uniformly for coprime $A, A_1 \in \mathcal{S}_y$ one of which is even as $x \to \infty$. Now observe that for fixed A and A_1, condition (25) puts m into a certain arithmetic progression $C \pmod{A_1}$. Since m is also coprime to P_y, this progression can be extended to $C_1, \ldots, C_s$ residue classes modulo $\mathrm{lcm}[A_1, P_y^2]$. The residue classes $C_1, \ldots, C_s$ are such that the numbers

$$m_1 = \frac{Am+1}{A_1} = \frac{A\mathrm{lcm}[A_1, P_y^2]}{A_1}\ell + \left(\frac{AC_i+1}{A_1}\right)$$

are also coprime to P_y. We also fix the class $C = C_i$ and write $\mathcal{B}(A, A_1, C)$ for the set of such n. It clearly suffices to show that the estimates

$$\sum_{m \in \mathcal{B}(A,A_1,C)} \mathbf{e}\left(\frac{t\omega(Am+1)}{g(A)}\right) = o(\#\mathcal{B}(A,A_1,C))$$

hold uniformly in our range for A, A_1, C as $x \to \infty$. In order to prove the above estimate, we slice again the interval $|k - \log_2 x| < z\sqrt{\log_2 x}$ into subintervals of length $L = \lfloor z \rfloor$ as in proof of part (i) of the theorem. Assume that $\omega(n) = k$. Then $\ell = \omega(m_1) = k - \omega(A_1) = k + O(z)$. The ingredient that we use is a result of Wolke and Zhan (see [9]), who showed that for all $\varepsilon > 0$ and constant $K > 0$ there exists a positive constant $\eta = \eta(\varepsilon, K)$ depending on ε and K such that the estimate

$$\sum_{d \le x^{1/2-\varepsilon}} \max_{(a,d)=1} \max_{w \le x} \left| \sum_{\substack{n \le w \\ \omega(n)=k \\ n \equiv a \pmod{d}}} 1 - \frac{1}{\phi(d)} \sum_{\substack{n \le w \\ (n,d)=1 \\ \omega(n)=k}} 1 \right| \ll \frac{\Pi_k(x)}{(\log x)^K} \tag{26}$$

holds uniformly in $1 \le k \le \eta \log x/(\log_2 x)^2$. We take $\varepsilon = 1/2$ and $K = 10$. We also take $d = \mathrm{lcm}[A_1, P_y^2]/A_1 \le P_y^2 \le e^{4y} = (\log x)^4$, and observe that the condition $(x/A_1)^{1/4} > d$ holds for large x because $A_1 < y^{10\log_2 y} < \log x$ for large x. Observe also that $\ell = \omega(m_1)$ satisfies the inequality

$$\ell \le \eta \log(x/A_1)/(\log_2(x/A_1))^2$$

for large values of x with $\eta = \eta(1/2, 10)$. We deduce from estimate (26) that

$$\left| \sum_{\substack{m \in \mathcal{B}(A, A_1, C) \\ \omega(m) = \ell}} 1 - \frac{1}{\phi(d)} M_\ell(x/A_1|y) \right| \ll \frac{\Pi_\ell(x)}{(\log x)^{10}}. \tag{27}$$

Since

$$\frac{M_\ell(x/A_1|y)}{\phi(d)} \gg \frac{\Pi_\ell(x/A_1)}{(\log x)^4 \log y} \gg \frac{\Pi_\ell(x)}{A_1 (\log x)^4 \log y} \gg \frac{\Pi_\ell(x)}{(\log x)^6},$$

it follows easily from estimate (27) that

$$\sum_{\substack{m \in \mathcal{B}(A, A_1, C) \\ \omega(m) = \ell}} 1 = (1 + o(1)) \frac{M_\ell(x/A_1|y)}{\phi(d)}$$

uniformly in our range for k as $x \to \infty$. Now we proceed as in the last part of the proof of the part (i) and conclude the proof of (ii). The theorem is therefore completely proved.

Acknowledgement. We thank the anonymous referee for a careful reading of a preliminary version of this manuscript and for suggestions which improved the quality of the paper. This work started when both authors attended the Colloquium on Uniform Distribution in Luminy in January of 2008. They thank the organizers for the opportunity of attending this conference. Research of F. L. was supported in part by Grant SEP-CONACyT 79685 and PAPIIT 100508.

REFERENCES

[1] BANKS, W.D. – GARAEV, M.Z. – LUCA, F. – SHPARLINSKI, I.E.: *Uniform distribution of the fractional part of the average prime factor*, Forum Mathematicum **17** (2005), 885–903.

[2] ELLIOTT, P.D.T.A.: *Probabilistic Number Theory I: Mean-Value Theorems*, Springer-Verlag, New York, 1979.

[3] ERDŐS, P. – WINTNER, A.: *Additive arithmetical functions and statistical independence*, Amer. J. Math. **61** (1939), 713–721.

[4] KÁTAI, I.: *On the average prime divisors*, Ann. Univ. Sci. Budapest, Sect. Comp. **27** (2007), 137–144.

[5] LÉVY, P.: *Sur les séries dont les termes sont des variables éventuelles indépendants*, Studia Math. **3** (1931), 119–155.

[6] LUCA, F. – PAPPALARDI, F.: , *Composite positive integers with an average prime factor*, Acta Arith. **129** (2007), 197–201.

[7] LUCA, F. – SHPARLINSKI, I.E.: *On the distribution modulo* 1 *of the geometric mean prime divisor*, Bol. Soc. Mat. Mexicana **12** (2006), 155–164.
[8] SELBERG, A.: *Note on a paper by L. G. Sathe*, J. Indian Math. Soc. **18** (1954), 83–87.
[9] WOLKE, D. – ZHAN, T.: *On the distribution of integers with a fixed number of prime factors*, Math. Z. **213** (1993), 133–144.

Received January 20, 2009
Accepted January 8, 2010

Imre Kátai
Department of Computer Algebra
Eőtvős Loránd University
Pázmány Péter sétány 1/C
H-1117 Budapest
HUNGARY
E-mail: katai@compalg.inf.elte.hu

Florian Luca
Instituto de Matemáticas
Universidad Nacional Autonoma de México
Ap. Postal 61-3 (Xangari)
C.P. 58089 Morelia, Michoacán
MÉXICO
E-mail: fluca@matmor.unam.mx