COLLOQUIUM MATHEMATICUM

Online First version

CONSECUTIVE NEIGHBOUR SPACINGS BETWEEN THE PRIME DIVISORS OF AN INTEGER

BY

JEAN-MARIE DE KONINCK (Québec) and IMRE KÁTAI (Budapest)

Dedicated to the memory of Professor Aleksandar Ivić

Abstract. Writing $p_1(n) < \cdots < p_r(n)$ for the distinct prime divisors of a given integer $n \geq 2$ and letting, for a fixed $\lambda \in (0,1]$, $U_\lambda(n) := \#\{j \in \{1,\ldots,r-1\} : \log p_j(n)/\log p_{j+1}(n) < \lambda\}$, we recently proved that $U_\lambda(n)/r \sim \lambda$ for almost all integers $n \geq 2$. Now, given $\lambda \in (0,1)$ and $p \in \wp$, the set of prime numbers, let $\mathcal{B}_\lambda(p) := \{q \in \wp : \lambda < \frac{\log q}{\log p} < 1/\lambda\}$ and consider the arithmetic function $u_\lambda(n) := \#\{p \mid n : (n/p, \mathcal{B}_\lambda(p)) = 1\}$. Here, we prove that $\sum_{n\leq x}(u_\lambda(n) - \lambda^2 \log\log n)^2 = (C + o(1))x \log\log x$ as $x \to \infty$, where C is a positive constant which depends only on λ, and thereafter we consider the case of shifted primes. Finally, we study a new function $V(n)$ which counts the number of divisors of n with large neighbour spacings and establish the mean value of $V(n)$ and of $V^2(n)$.

1. Introduction. Given an integer $n \geq 2$, let $\omega(n)$ stand for the number of distinct prime divisors of $n \geq 2$ (setting $\omega(1) = 0$) and let

$$p_1(n) < \cdots < p_{\omega(n)}(n) \quad \text{or briefly} \quad p_1 < \cdots < p_{\omega(n)} \tag{1.1}$$

be those prime divisors. Many have shown interest for the relative size of these prime factors $p_j(n)$.

Let $\xi(n) \to \infty$ as $n \to \infty$. In 1946, P. Erdős [4] proved that given any small number $\varepsilon > 0$,

$$e^{e^{k(1-\varepsilon)}} < p_k(n) < e^{e^{k(1+\varepsilon)}} \quad (\xi(n) \leq k \leq \omega(n)) \quad \text{for almost all } n.$$

In 1976, J. Galambos [5] strengthened this result by showing that, given any small $\varepsilon > 0$ and a function $k = k(x)$ which tends to infinity with x in such a manner that $k(x) = o(\log\log x)$ as $x \to \infty$, we have

$$\lim_{x\to\infty} \frac{1}{x} \#\{n \leq x : e^{-(1+\varepsilon)k} \log x < \log p_{\omega(n)-k}(n) < e^{-(1-\varepsilon)k} \log x\} = 1.$$

Galambos also established that, given any small $\varepsilon > 0$ and a function

2020 *Mathematics Subject Classification*: Primary 11N37; Secondary 11A05.
Key words and phrases: prime divisors.
Received 12 December 2021.
Published online *.

DOI: 10.4064/cm8745-2-2022

$j = j(x)$ which tends to infinity with x in such a manner that

$$j(x) \le (1-\varepsilon)\log\log x,$$

for any fixed real number $z > 1$, we have

$$\lim_{x\to\infty} \frac{1}{x}\#\left\{n \le x : \frac{\log p_{j+1}(n)}{\log p_j(n)} < z\right\} = 1 - \frac{1}{z}.$$

In 1987, De Koninck and Galambos [2] showed that for almost all n (with the prime factors of n written as in (1.1)) and for any fixed positive integer k, the corresponding

$$(\log\log p_{j+1} - \log\log p_j, \log\log p_{j+2} - \log\log p_{j+1}, \\ \dots, \log\log p_{j+k} - \log\log p_{j+k-1})$$

is distributed as a k-tuple of independent exponential random variables with parameter 1.

Finally, in 2007, Granville [6], [7] proved that if $S_k(x)$ stands for the number of positive integers $n \le x$ such that $\omega(n) = k$, then for all but $o(S_\ell(x))$ of the integers $n \in S_\ell(x)$, the sets

$$\left\{\frac{\log\log p}{\frac{1}{\ell}\log\log(n^{1/\ell})} : p \,|\, n,\ p \le n^{1/\ell}\right\}$$

are Poisson distributed.

Recently, in [3], we further expanded on the above results by examining the distribution of the consecutive neighbour spacings between the prime divisors of a typical integer n. Here, we examine other functions which provide more information on the spacings between the prime divisors of an integer.

2. Setting the table. Let us write the distinct prime factors of an integer n as in (1.1) and introduce the functions

$$\gamma_j(n) := \frac{\log p_j(n)}{\log p_{j+1}(n)} \qquad (j = 1, \dots, r-1).$$

From here on, the letters p and q (and at times the letter π) will denote primes, whereas the letter $\wp$ will stand for the set of all prime numbers. Given positive real numbers $u < v$, we set

$$Q(u,v) := \prod_{\substack{u<p<v \\ p\in\wp}} p.$$

Also, given an integer $n \ge 2$ and a prime divisor p of n which is smaller than $P(n)$, the largest prime divisor of n, we set

$$\nu_p = \nu_p(n) := \min\{q \,|\, n : q > p\}.$$

In light of the above notation, we have

$$\mathrm{GCD}\left(\frac{n}{p\nu_p}, Q(p,\nu_p)\right) = 1.$$

When the context is clear, we shall write (a,b) instead of $\mathrm{GCD}(a,b)$. Given a real number $\lambda \in (0,1]$, we introduce the function

$$U_\lambda(n) := \sum_{\substack{p|n \\ \frac{\log p}{\log \nu_p(n)} < \lambda}} 1,$$

or equivalently

$$U_\lambda(n) := \sum_{\substack{p|n \\ \log\log \nu_p(n) - \log\log p > \log \frac{1}{\lambda}}} 1.$$

In a recent paper [3], we proved the following.

Theorem A (De Koninck–Kátai). *For an arbitrary real number $\lambda \in (0,1]$,*

$$\sum_{n\le x} U_\lambda(n) = (1+o(1))\lambda x \log\log x \qquad (x\to\infty).$$

Moreover, for every $\varepsilon > 0$,

$$\lim_{x\to\infty} \frac{1}{x}\#\left\{n \le x : \left|\frac{U_\lambda(n)}{\omega(n)} - \lambda\right| > \varepsilon\right\} = 0.$$

Now, given $\lambda \in (0,1)$ and $p \in \wp$, consider the set

$$\mathcal{B}_\lambda(p) := \left\{q \in \wp : \lambda < \frac{\log q}{\log p} < 1/\lambda\right\}.$$

We will say that a positive integer m is *coprime to a set A of primes* and write $(m,A) = 1$ if $(m,p) = 1$ for every $p \in A$.

In the current paper, we study the arithmetic function

$$u_\lambda(n) := \#\{p \,|\, n : (n/p, \mathcal{B}_\lambda(p)) = 1\}.$$

Observe that it follows from Theorem A that

$$\sum_{n\le x} u_\lambda(n) = (1+o(1))\lambda^2 x \log\log x \qquad (x\to\infty), \tag{2.1}$$

$$\sum_{n\le x} u_\lambda(n)^2 = \lambda^4 x(\log\log x)^2 + O(x\log\log x). \tag{2.2}$$

Our first goal in this paper is to fine tune the above estimates so that we can prove that $\sum_{n\le x}(u_\lambda(n) - \lambda^2\log\log x)^2 = (C+o(1))x\log\log x$ as $x\to\infty$, where C is a positive constant which depends only on λ, and thereafter to consider the case of shifted primes. Finally, we study a new function $V(n)$

which counts the number of divisors of n with large neighbour spacings and establish the mean value of $V(n)$ and of $V^2(n)$.

3. Main results. The following is an improvement of estimates (2.1) and (2.2).

Theorem 3.1. *As* $x \to \infty$,

$$\frac{1}{x}\sum_{n\le x}\big(u_\lambda(n) - \lambda^2 \log\log x\big)^2 = (1+o(1))\psi(\lambda)\log\log x, \tag{3.1}$$

where

$$\psi(\lambda) = \lambda^2 + 2\lambda^2(1-\lambda^2) - 4\lambda^4 \log\frac{1}{\lambda}. \tag{3.2}$$

We then have the following analogue result for shifted primes.

Theorem 3.2. *As* $x \to \infty$,

$$\frac{1}{\pi(x)}\sum_{p\le x} u_\lambda(p+1) = (1+o(1))\lambda^2\log\log x,$$

$$\frac{1}{\pi(x)}\sum_{p\le x}\big(u_\lambda(p+1) - \lambda^2\log\log x\big)^2 = (1+o(1))\psi(\lambda)\log\log x,$$

where, as usual, $\pi(x)$ *stands for the number of primes not exceeding* x.

Now, given real numbers $1 < \xi_1 < \xi_2$, let

$$\Lambda(\xi_1,\xi_2) := \frac{\log \xi_1}{\log \xi_2}.$$

Let $k \ge 2$ be an integer and consider the intervals $\mathcal{I}_j := (u_j, v_j) \subseteq (0,1)$ for $j = 1,\dots,k-1$, and let $\lambda \in (0,1)$. Then, given a large number x, we let $\mathcal{K}$ be the collection of all prime k-tuples $(p_1,\dots,p_k)$ such that $p_1 < \cdots < p_k \le x$ and

$$\Lambda(p_j, p_{j+1}) \in \mathcal{I}_j \qquad (j = 1,\dots,k-1).$$

We then let $V(n)$ be the number of those divisors d of n of the form $d = p_1\cdots p_k$, where $(p_1,\dots,p_k)\in\mathcal{K}$, and for which we also have

$$\left(\frac{n}{p_1\cdots p_k}, Q(p_1^\lambda, p_k^{1/\lambda})\right) = 1.$$

Further, let $\widetilde{V}(n)$ be the number of those divisors counted by $V(n)$ but which satisfy the additional condition $Y_1 < p_1 < p_k < Y_2$, where $Y_1 = Y_1(x)$ and $Y_2 = Y_2(x)$ are defined as follows. Given a real number $x \ge e^{e^e}$, we set

$$Y_1 := Y_1(x) = \exp\{(\log x)^{\varepsilon(x)}\}, \quad Y_2 := Y_2(x) = \exp\{(\log x)^{1-\varepsilon(x)}\}, \tag{3.3}$$

where

$$\varepsilon(x) := \frac{1}{2}\,\frac{\log\log\log x}{\log\log x}.$$

Finally, set

$$\Delta_1 = \Delta_1(\mathcal{I}_1,\dots,\mathcal{I}_{k-1}) := (v_1-u_1)\cdots(v_{k-1}-u_{k-1}),$$
$$\Delta_2 = \Delta_2(\mathcal{I}_1,\dots,\mathcal{I}_{k-1}) := \left(\frac{1}{u_1}-\frac{1}{v_1}\right)\cdots\left(\frac{1}{u_{k-1}}-\frac{1}{v_{k-1}}\right).$$

Then we have the following.

THEOREM 3.3. *Let k, λ, $\mathcal{I}_j$ $(j=1,\dots,k-1)$ and $\mathcal{K}$ be as above. Then*

(i) $\frac{1}{x}\sum_{n\le x}\widetilde{V}(n) = \lambda^2\Delta_1\log\log x + O(1)$,
(ii) $\frac{1}{x}\sum_{n\le x}V(n) = \lambda^2\Delta_1\log\log x + O(\log\log\log x)$.

Finally, we will prove the following.

THEOREM 3.4. *We have*

$$\sum_{n\le x}\widetilde{V}^2(n) = \lambda^4\Delta_1\Delta_2 xS^2 + MxS + o(xS), \tag{3.4}$$

where

$$M := 2\lambda^2\Delta_1\Delta_2(\lambda-\lambda^2-2\lambda^2\log 1/\lambda)+\lambda^2\Delta_1,$$
$$S := \sum_{Y_1\le p\le Y_2}\frac{1}{p} = \log\log x + O(\log\log\log x).$$

Moreover,

$$\sum_{n\le x}V^2(n) = \lambda^4\Delta_1\Delta_2 x(\log\log x)^2 + O(x\log\log x\cdot\log\log\log x). \tag{3.5}$$

4. Preliminary results. Here, we state some classical results from prime number theory. We start with Mertens' theorem, which in fact can be formulated in three equivalent forms.

THEOREM B (Mertens). *For large x, we have*

(i) $\displaystyle\sum_{p\le x}\frac{\log p}{p} = \log x + O(1)$,

(ii) $\displaystyle\sum_{p\le x}\frac{1}{p} = \log\log x + B + O\left(\frac{1}{\log x}\right)$ *for some constant B*,

(iii) $\displaystyle\prod_{p\le x}\left(1-\frac{1}{p}\right) = \frac{e^D}{\log x}\left(1+O\left(\frac{1}{\log x}\right)\right)$ *for some constant D.*

Proof. This is Theorem 10.1 in the book of De Koninck and Doyon [1], where a detailed proof is also given. ∎

LEMMA 4.1. *Given real numbers $y > x \geq e$, we have*

(a) $\displaystyle\sum_{x<p\leq y} \frac{\log p}{p} = \log y - \log x + O\left(\frac{1}{\log x}\right)$,

(b) $\displaystyle\sum_{x<p\leq y} \frac{1}{p} = \log\left(\frac{\log y}{\log x}\right) + O\left(\frac{1}{\log x}\right)$.

Proof. To prove part (a), one will not succeed by simply using Theorem B(i). A stronger estimate is required. In 1962, Rosser and Schoenfeld [9] proved that there exist constants $E < 0$ and $a > 0$ such that

$$\sum_{p\leq x} \frac{\log p}{p} = \log x + E + O\left(\frac{1}{e^{a\sqrt{\log x}}}\right). \tag{4.1}$$

It is then clear that (a) is an easy consequence of (4.1).

Part (b) is an immediate consequence of Theorem B(ii). ∎

With the notation introduced right before the statement of Theorem 3.3 along with the conditions attached to the related variables, we have the following.

LEMMA 4.2. *For each $j = 1, \dots, k-1$,*

$$\sum_{u_j \leq \Lambda(p_j,p_{j+1}) < v_j} \frac{\log p_j}{p_j} = (v_j - u_j)\log p_j + O\left(\frac{1}{\log p_j}\right).$$

Proof. This is a direct consequence of Lemma 4.1(a). ∎

LEMMA 4.3. *For each $j = 1, \dots, k-1$,*

$$\sum_{p_j^{1/v_j} < p_{j+1} < p_j^{1/u_j}} \frac{1}{p_{j+1}\log p_{j+1}} = \left(\frac{1}{u_j} - \frac{1}{v_j}\right)\frac{1}{\log p_j} + O\left(\frac{1}{\log^2 p_j}\right).$$

Proof. Using the prime number theorem in the form

$$\pi(x) = \frac{x}{\log x} + O\left(\frac{x}{\log^2 x}\right),$$

we get

$$\begin{aligned}\sum_{p_j^{1/v_j} < p_{j+1} < p_j^{1/u_j}} \frac{1}{p_{j+1}\log p_{j+1}} &= \int_{p_j^{1/v_j}}^{p_j^{1/u_j}} \frac{1}{t\log t}\, d\pi(t) \\ &= \int_{p_j^{1/v_j}}^{p_j^{1/u_j}} \frac{1}{t\log^2 t}\left(1 + O\left(\frac{1}{\log t}\right)\right) dt\end{aligned}$$

$$= -\frac{1}{\log t}\left(1+O\left(\frac{1}{\log t}\right)\right)\Bigg|_{p_j^{1/v_j}}^{p_j^{1/u_j}}$$
$$= \left(\frac{1}{u_j}-\frac{1}{v_j}\right)\frac{1}{\log p_j}+O\left(\frac{1}{\log^2 p_j}\right),$$

thus proving our claim. ∎

LEMMA 4.4. *We have*

$$\sum_{\{p_1,\dots,p_k\}\in\mathcal{K}} \frac{\log p_1}{p_1\cdots p_k \log p_k} = \varDelta_1 \log\log x + O(1).$$

Proof. Making repetitive use of Lemma 4.2, we obtain

$$\sum_{\{p_1,\dots,p_k\}\in\mathcal{K}} \frac{\log p_1}{p_1\cdots p_k \log p_k}$$
$$= (v_1-u_1)\sum_{p_2,\dots,p_k}\frac{\log p_2}{p_2\cdots p_k\log p_k}\left(1+O\left(\frac{1}{\log^2 p_2}\right)\right)$$
$$\vdots$$
$$= (v_1-u_1)\cdots(v_{k-1}-u_{k-1})\sum_{p_k}\frac{\log p_k}{p_k\log p_k}\left(1+O\left(\frac{1}{\log^2 p_2}\right)\right)$$
$$= \varDelta_1\sum_{p_k}\frac{1}{p_k}\left(1+O\left(\frac{1}{\log^2 p_2}\right)\right)$$
$$= \varDelta_1\log\log x+O(1). \blacksquare$$

LEMMA 4.5. *For all $x \ge e^{e^e}$,*

$$S := \sum_{Y_1\le p\le Y_2}\frac{1}{p} = \log\log x - \log\log\log x + O\left(\frac{1}{\sqrt{\log\log x}}\right). \tag{4.2}$$

Proof. Using estimate (ii) of Theorem B, we have

$$S = \log\log Y_2 - \log\log Y_1 + O\left(\frac{1}{\log Y_1}\right)$$
$$= (1-\varepsilon(x))\log\log x - \varepsilon(x)\log\log x + O\left(\frac{1}{e^{\varepsilon(x)\log\log x}}\right)$$
$$= \log\log x - 2\varepsilon(x)\log\log x + O\left(\frac{1}{e^{\log\log\log x/2}}\right),$$

from which (4.2) follows immediately. ∎

Finally, we recall Lemmas 5 and 6 of our recent paper [3] which we rename here as Lemmas A and B.

Lemma A. *Let* $x \geq e^{e^{100}}$ *and let* $Y_1 = Y_1(x)$ *and* $Y_2 = Y_2(x)$ *be the functions defined in* (3.3). *Let* $\pi_1 < \cdots < \pi_s$ *be* s *primes located in the interval* (Y_1, Y_2). *Write their product as* $B = \pi_1 \cdots \pi_s$. *Further, set*

$$\eta := \sum_{i=1}^{s} \frac{1}{\pi_i} \quad \textit{and} \quad S_B(x) := \sum_{\substack{n\leq x\\(n,B)=1}} 1.$$

Assume that $\eta \leq K$, *where* K *is an arbitrary number, and let* h *be a positive integer satisfying* $h \geq 3e^2K$. *Then, letting* ϕ *stand for the Euler totient function, we have*

$$\left| S_B(x) - \frac{\phi(B)}{B}x \right| \leq x(3e)^{-h} + 2Y_2^h,$$

so that in particular, if we choose $h = \lfloor \log\log\log x \rfloor$, *there exists a positive constant* c *such that*

$$\left| S_B(x) - \frac{\phi(B)}{B}x \right| \leq \frac{cx}{(\log\log x)^2}.$$

Lemma B. *Let* x, B, h, η *and* K *be as in Lemma* A. *Let* D *be a positive integer* $\leq Y_2^{c_0}$, *where* c_0 *is an arbitrary positive constant, and assume that* $(B, D) = 1$. *Consider the sum*

$$\overline{S}_{B,D}(x) := \sum_{\substack{p\leq x\\ p+1\equiv 0 \,(\mathrm{mod}\, D)\\ (p+1,B)=1}} 1.$$

Then, setting

$$\widetilde{\phi}_0(n) = \frac{\phi(n)}{n} \prod_{p|n} \left(1 - \frac{1}{(p-1)^2}\right),$$

for some positive constant c_1 *we have*

$$\left| \overline{S}_{B,D}(x) - \widetilde{\phi_0}(B)\frac{\mathrm{li}(x)}{\phi(D)} \right| \leq c_1 \frac{\mathrm{li}(x)\,(3e)^{-h}}{\phi(D)} + Y_2^h,$$

where $\mathrm{li}(x) := \int_2^x \frac{dt}{\log t}$, *so that in particular, by choosing* $h = \lfloor \log\log\log x \rfloor$, *for some positive constant* c_2 *we have*

$$\left| \overline{S}_{B,D}(x) - \widetilde{\phi_0}(B)\frac{\mathrm{li}(x)}{\phi(D)} \right| \leq c_2 \frac{\mathrm{li}(x)}{\phi(D)\,(\log\log x)^2}.$$

5. The proofs of Theorems 3.1 and 3.2. Our main goal in this section is to prove relation (3.1).

Given an integer $n \geq 2$, we start by subdividing the primes p dividing n into three subsets: $p < Y_1$, $p \in [Y_1, Y_2]$, and $p > Y_2$. Correspondingly, we write $u_\lambda(n)$ as the sum of three new functions:

$$u_\lambda(n) = u_\lambda^{(1)}(n) + \widetilde{u_\lambda}(n) + u_\lambda^{(2)}(n). \tag{5.1}$$

First observe that

$$\frac{1}{x}\sum_{n\le x} u_\lambda^{(1)}(n) \le \sum_{p<Y_1}\frac{1}{p} \ll \log\log Y_1 = \varepsilon(x)\log\log x = \frac{1}{2}\log\log\log x \tag{5.2}$$

and

$$\begin{aligned} \frac{1}{x}\sum_{n\le x} u_\lambda^{(2)}(n) &\le \sum_{Y_2<p\le x}\frac{1}{p} = \log\log x - \log\log Y_2 + O(1) \\ &= \log\log x - (1-\varepsilon(x))\log\log x + O(1) \\ &= \varepsilon(x)\log\log x + O(1) = \frac{1}{2}\log\log\log x + O(1). \end{aligned} \tag{5.3}$$

Now, on the other hand, using Lemma A and Theorem B(iii), we have

$$\begin{aligned} \sum_{n\le x}\widetilde{u_\lambda}(n) &= \sum_{p\in[Y_1,Y_2]} \#\{m \le x/p : (m, \mathcal{B}_\lambda(p)) = 1\} \\ &= \sum_{p\in[Y_1,Y_2]} \frac{x}{p} \prod_{q\in\mathcal{B}_\lambda(p)}\left(1-\frac{1}{q}\right)\left(1+O\left(\frac{1}{\log p}\right)\right) \\ &= x\lambda^2 \sum_{p\in[Y_1,Y_2]}\frac{1}{p} + O\left(x\sum_{p>Y_1}\frac{1}{p\log p}\right) = x\lambda^2 S + o(x). \end{aligned} \tag{5.4}$$

We also have

$$\begin{aligned} &\sum_{n\le x}\widetilde{u_\lambda}^2(n) \\ &\quad = 2\sum_{Y_1<p_1<p_2<Y_2} \#\left\{m \le \frac{x}{p_1p_2} : q \mid m \Rightarrow q \notin \mathcal{B}_\lambda(p_1)\cup\mathcal{B}_\lambda(p_2)\right\} \\ &\qquad + \sum_{\substack{p_1=p_2\\ Y_1<p<Y_2}} \#\{m \le x/p : (m, \mathcal{B}_\lambda(p)) = 1\} \\ &\quad =: 2S_1(x) + S_2(x). \end{aligned} \tag{5.5}$$

It is easy to see that $S_2(x) = \sum_{n\le x}\widetilde{u_\lambda}(n)$. Thus, in light of estimate (5.4), we may write

$$S_2(x) = x\lambda^2 S + o(x). \tag{5.6}$$

In order to estimate $S_1(x)$, we will again make use of Lemma A. To do so, we need to examine the possible overlap of the sets $\mathcal{B}_\lambda(p_1)$ and $\mathcal{B}_\lambda(p_2)$. First observe that

$$\mathcal{B}_\lambda(p_1)\cap\mathcal{B}_\lambda(p_2) = \emptyset \quad \text{if} \quad \frac{\log p_1}{\log p_2} < \lambda^2,$$

in which case we have, again using Theorem B(iii),

$$\prod_{q\in\mathcal{B}_\lambda(p_1)\cup\mathcal{B}_\lambda(p_2)}\left(1-\frac{1}{q}\right)=\prod_{q\in\mathcal{B}_\lambda(p_1)}\left(1-\frac{1}{q}\right)\cdot\prod_{q\in\mathcal{B}_\lambda(p_2)}\left(1-\frac{1}{q}\right)=\lambda^4\left(1+O\left(\frac{1}{\log p_1}\right)\right).\tag{5.7}$$

Now, in the other situation, that is, if

$$\lambda^2<\frac{\log p_1}{\log p_2}<\lambda,$$

then

$$\prod_{q\in\mathcal{B}_\lambda(p_1)\cup\mathcal{B}_\lambda(p_2)}\left(1-\frac{1}{q}\right)=\prod_{p_1^\lambda<q<p_2^{1/\lambda}}\left(1-\frac{1}{q}\right)=\lambda^2\frac{\log p_1}{\log p_2}\left(1+O\left(\frac{1}{\log p_1}\right)\right).\tag{5.8}$$

Thus, combining (5.7) and (5.8), and recalling the implicit definition of $S_1(x)$ given in (5.5), we obtain

$$\begin{aligned}S_1(x)&=\lambda^4x\sum_{p_1<p_2^{\lambda^2}}\frac{1}{p_1p_2}+\lambda^2x\sum_{\lambda^2<\frac{\log p_1}{\log p_2}<\lambda}\frac{\log p_1}{p_1p_2\log p_2}+o(x)\\&=\lambda^4x\left(\frac{S^2}{2}-\sum_{p_2}\frac{1}{p_2}\left(\sum_{p_2^{\lambda^2}<p_1<p_2}\frac{1}{p_1}\right)\right)\\&\quad+\lambda^2x\sum_{p_2}\frac{1}{p_2\log p_2}\sum_{p_2^{\lambda^2}<p_1<p_2}\frac{\log p_1}{p_1}+o(x)\\&=\lambda^4x\frac{S^2}{2}-\lambda^4x\sum_{p_2}\frac{1}{p_2}\left(\sum_{p_2^{\lambda^2}<p_1<p_2}\frac{1}{p_1}\right)+\lambda^2x(1-\lambda^2)S+o(x).\end{aligned}\tag{5.9}$$

Using Theorem B(ii), we obtain

$$\sum_{p_2^{\lambda^2}<p_1<p_2}\frac{1}{p_1}=\log\left(\frac{\log p_2}{\lambda^2\log p_2}\right)+O\left(\frac{1}{\log p_2}\right)=2\log\frac{1}{\lambda}+O\left(\frac{1}{\log p_2}\right).$$

Inserting this last estimate in (5.9) gives

$$S_1(x)=\lambda^4\frac{S^2}{2}x-2\cdot\log\frac{1}{\lambda}\cdot\lambda^4Sx+\lambda^2(1-\lambda^2)Sx+o(x).\tag{5.10}$$

Gathering (5.6) and (5.10) in (5.5) gives

$$\begin{aligned}(5.11)\quad \frac{1}{x}\sum_{n\le x}\widetilde{u_\lambda}^2(n) &= \lambda^4S^2+\lambda^2S+\left(2\lambda^2(1-\lambda^2)-4\left(\log\frac{1}{\lambda}\right)\lambda^4\right)S+o(1)\\ &= \lambda^4S^2+\psi(\lambda)S+o(1),\end{aligned}$$

where $\psi(\lambda)$ is the function defined in (3.2).

Using (5.11) and (5.4), we thus have

$$\begin{aligned}(5.12)\quad \frac{1}{x}\sum_{n\le x}\left(\widetilde{u_\lambda}(n)-\lambda^2S\right)^2 &= \frac{1}{x}\sum_{n\le x}\widetilde{u_\lambda}^2(n)-2\lambda^2S\frac{1}{x}\sum_{n\le x}\widetilde{u_\lambda}(n)+\lambda^4S^2\\ &= \psi(\lambda)S+o(1).\end{aligned}$$

Using the estimate of Lemma 4.5 in (5.12), we may rewrite (5.12) as

$$(5.13)\quad \frac{1}{x}\sum_{n\le x}\left(\widetilde{u_\lambda}(n)-\lambda^2\log\log x\right)^2=\psi(\lambda)\log\log x+O(\log\log\log x).$$

Inserting (5.2), (5.3) and (5.13) in (5.1), we can prove that

$$(5.14)\quad \frac{1}{x}\sum_{n\le x}\left(u_\lambda(n)-\lambda^2\log\log x\right)^2=\psi(\lambda)\log\log x+o(\log\log x).$$

Indeed, first, for each integer $n\ge 1$, let us set

$$\begin{aligned}\delta_n &:= \widetilde{u_\lambda}(n)-\lambda^2S,\\ \kappa_n &:= (u_\lambda(n)-\widetilde{u_\lambda}(n))-\lambda^2(\log\log x-S).\end{aligned}$$

The numbers δ_n and κ_n are tied by the relation

$$\kappa_n=(u_\lambda(n)-\lambda^2\log\log x)-\delta_n,$$

so that

$$(5.15)\quad (u_\lambda(n)-\lambda^2\log\log x)^2=\delta_n^2+\kappa_n^2+2\delta_n\kappa_n.$$

Clearly, in light of (5.13) and of Lemma 4.5,

$$(5.16)\quad \frac{1}{x}\sum_{n\le x}\delta_n^2=\psi(\lambda)\log\log x+O(\log\log\log x).$$

Now, setting $\rho_n:=\#\{p\,|\,n : p\notin(Y_1,Y_2)\}$, we have

$$(5.17)\quad |\kappa_n|\le\omega(n)+\lambda^2(\log\log x-S)\le\rho_n+O(\log\log\log x).$$

On the other hand, it is clear that

$$(5.18)\quad \sum_{n\le x}\rho_n\ll x\log\log\log x\quad\text{and}\quad\sum_{n\le x}\rho_n^2\ll x(\log\log\log x)^2.$$

Using (5.18) in (5.17), we obtain

$$(5.19)\quad \sum_{n\le x}\kappa_n\ll x\log\log\log x\quad\text{and}\quad\sum_{n\le x}\kappa_n^2\ll x(\log\log\log x)^2.$$

Hence, by using the Cauchy–Schwarz inequality, it follows from (5.16) and (5.19) that

$$(5.20)\quad \Big|\sum_{n\le x}\delta_n\kappa_n\Big| \le \sum_{n\le x}|\delta_n\kappa_n| \le \Big(\sum_{n\le x}\delta_n^2\Big)^{1/2}\cdot\Big(\sum_{n\le x}\kappa_n^2\Big)^{1/2}$$
$$\ll \sqrt{x\log\log x}\cdot\sqrt{x(\log\log\log x)^2} = x\sqrt{\log\log x}\cdot\log\log\log x.$$

Using estimates (5.16), (5.19) and (5.20) in identity (5.15) completes the proof of (5.14) and thus at the same time the proof of Theorem 3.1.

The proof of Theorem 3.2 goes along the same lines as the proof of Theorem 3.1, except that it uses the Bombieri–Vinogradov theorem (see Theorem 17.1 in the book of Iwaniec and Kowalski [8]) and the above Lemma B.

6. Proof of Theorem 3.3. It is clear that

$$(6.1)\quad \sum_{n\le x}\widetilde{V}(n) = \sum_{\substack{Y_1<p_1<p_k<Y_2\\ \{p_1,\dots,p_k\}\in\mathcal{K}}} \#\Big\{m\le \frac{x}{p_1\cdots p_k} : \Big(m, \frac{Q(p_1^\lambda, p_k^{1/\lambda})}{p_1\cdots p_k}\Big)=1\Big\}$$
$$=: \sum_{\substack{Y_1<p_1<p_k<Y_2\\ \{p_1,\dots,p_k\}\in\mathcal{K}}} H(x\,|\,p_1,\dots,p_k).$$

It follows from Lemma A and Theorem B(iii) that

$$(6.2)\quad H(x\,|\,p_1,\dots,p_k) = \frac{x}{p_1\cdots p_k}\prod_{\substack{p_1^\lambda<q<p_k^{1/\lambda}\\ q\notin\{p_1,\dots,p_k\}}}\Big(1-\frac{1}{q}\Big)\Big(1+O\Big(\frac{1}{\log p_1}\Big)\Big)$$
$$= \frac{x}{p_1\cdots p_k}\cdot\frac{\lambda\log p_1}{(1/\lambda)\log p_k}\Big(1+O\Big(\frac{1}{\log p_1}\Big)\Big).$$

Using (6.2) in (6.1) and the estimate of Lemma 4.4, we obtain

$$(6.3)\quad \sum_{n\le x}\widetilde{V}(n) = \lambda^2 x\sum_{\{p_1,\dots,p_k\}\in\mathcal{K}}\frac{\log p_1}{p_1\cdots p_k\log p_k} = \lambda^2\varDelta_1 x\log\log x + O(x),$$

thus proving Theorem 3.3(i). It remains to prove (ii). To do so, it is clearly sufficient to prove that

$$(6.4)\quad \sum_{n\le x}(V(n)-\widetilde{V}(n)) \ll x\log\log\log x.$$

To evaluate the above sum, we need to estimate the contribution of those $\{p_1,\dots,p_k\}\in\mathcal{K}$ for which either $p_1<Y_1$ or $p_k>Y_2$.

Now, given $\{p_1,\dots,p_k\}\in\mathcal{K}$, we have

$$u_1\cdots u_k\le \Lambda(p_1,p_k)=\prod_{j=1}^{k-1}\Lambda(p_j,p_{j+1})\le v_1\cdots v_k.$$

Therefore, if $p_1<Y_1$, then $p_k<Y_1^{C_1}$, where $C_1=\frac{1}{u_1}\cdots\frac{1}{u_k}$, and if $p_k>Y_2$, then $p_1>Y_2^{C_2}$, where $C_2=1/C_1$.

Gathering the above, we find that

$$\sum_{n\le x}(V(n)-\widetilde{V}(n))\ll x\biggl(\sum_{p<Y_1^{C_1}}\frac{1}{p}+\sum_{Y_2^{C_2}<p\le x}\frac{1}{p}\biggr)\ll x\log\log\log x,$$

thus proving (6.4) and thereby completing the proof of Theorem 3.3(ii).

7. Proof of Theorem 3.4. We first estimate

$$S(x):=\sum_{n\le x}\widetilde{V}^2(n).$$

In each n counted in $\widetilde{V}^2(n)$, we can separate its divisors $\{p_1,\dots,p_k\}\in\mathcal{K}$, $\{q_1,\dots,q_k\}\in\mathcal{K}$ into three categories:

(a) those such that $\{p_1,\dots,p_k\}=\{q_1,\dots,q_k\}$,
(b) those for which $p_k<q_1$,
(c) those for which $q_k<p_1$.

Since the contribution of those in category (b) is the same as that of those in category (c), we may write

$$S(x)=\sum_{n\le x}\widetilde{V}(n)+2S_b(x), \tag{7.1}$$

where

$$S_b(x)=\sum_{\substack{\{p_1,\dots,p_k\}\in\mathcal{K}\\ \{q_1,\dots,q_k\}\in\mathcal{K}\\ p_k<q_1}}\#\biggl\{m\le\frac{x}{p_1\cdots p_kq_1\cdots q_k}:$$
$$\biggl(m,\frac{Q(p_1^\lambda,p_k^{1/\lambda})}{p_1\cdots p_k}\biggr)=1 \text{ and } \biggl(m,\frac{Q(q_1^\lambda,q_k^{1/\lambda})}{q_1\cdots q_k}\biggr)=1\biggr\}.$$

Observe that

$$q_1^{\lambda^2}<q_1^\lambda<q_1.$$

So, there are three possibilities for the location of p_k in the above chain of inequalities:

(i) $q_1^\lambda<p_k<q_1$,
(ii) $q_1^{\lambda^2}<p_k<q_1^\lambda$,
(iii) $p_k<q_1^{\lambda^2}$.

The first possibility can be ignored, since in this case we have $\log q_1 \in \left[\lambda \log p_1, \frac{1}{\lambda}\log p_k\right]$, which is not counted in our case. Therefore, we are left with situations (ii) and (iii), which we rename as scenarios (1) and (2), as follows:

(1) $q_1^{\lambda^2} < p_k < q_1^{\lambda}$ (or equivalently $\lambda^2 < \Lambda(p_k, q_1) < \lambda$),
(2) $p_k < q_1^{\lambda^2}$ (or equivalently $\Lambda(p_k, q_1) < \lambda^2$).

We separate the sum $S_b(x)$ into two subsums, depending on the two scenarios, by writing

$$S_b(x) = S_b^{(1)}(x) + S_b^{(2)}(x). \tag{7.2}$$

We first consider $S_b^{(2)}(x)$. In that sum, $(Q(p_1^{\lambda}, p_k^{1/\lambda}), Q(q_1^{\lambda}, q_k^{1/\lambda})) = 1$ and therefore

$$\begin{aligned}\prod_{\pi \mid Q(p_1^{\lambda}, p_k^{1/\lambda})}\left(1 - \frac{1}{\pi}\right) \prod_{\pi \mid Q(q_1^{\lambda}, q_k^{1/\lambda})}\left(1 - \frac{1}{\pi}\right) \\ = \lambda^4 \frac{\log p_1}{\log p_k}\, \frac{\log q_1}{\log q_k}\left(1 + O\left(\frac{1}{\log p_1}\right)\right).\end{aligned} \tag{7.3}$$

On the other hand, in $S_b^{(1)}(x)$, we have

$$\prod_{\pi \mid Q(p_1^{\lambda}, q_k^{1/\lambda})}\left(1 - \frac{1}{\pi}\right) = \lambda^2 \frac{\log p_1}{\log q_k}\left(1 + O\left(\frac{1}{\log p_1}\right)\right). \tag{7.4}$$

Using (7.3) and (7.4), and making repeated use of Lemma 4.2, we obtain

$$\begin{aligned} S_b^{(2)}(x) &= x\lambda^4 \sum_{\substack{\{p_1,\ldots,p_k\} \in \mathcal{K} \\ \{q_1,\ldots,q_k\} \in \mathcal{K} \\ p_k < q_1^{\lambda^2}}} \frac{\log p_1}{p_1 \cdots p_k \log p_k}\, \frac{\log q_1}{q_1 \cdots q_k \log q_k}\left(1 + O\left(\frac{1}{\log p_1}\right)\right) \\ &= x\lambda^4 \Delta_1 \sum_{p_k < q_1^{\lambda^2}} \frac{\log q_1}{p_k q_1 \cdots q_k \log q_k}\left(1 + O\left(\frac{1}{\log p_1}\right)\right). \end{aligned} \tag{7.5}$$

Recall that from Lemma 4.3 we have

$$\sum_{u_j \le \Lambda(q_j, q_{j+1}) < v_j} \frac{1}{q_{j+1} \log q_{j+1}} = \left(\frac{1}{u_j} - \frac{1}{v_j}\right)\frac{1}{\log q_j} + O\left(\frac{1}{\log^2 q_j}\right).$$

Applying this relation for each $j = k-1, k-2, \ldots, 1$, we obtain

$$\sum_{p_k < q_1^{\lambda^2}} \frac{\log q_1}{p_k q_1 \cdots q_k \log q_k} = \Delta_2 \sum_{p_k < q_1^{\lambda^2}} \frac{1}{p_k q_1} =: \Delta_2 S_3. \tag{7.6}$$

To estimate S_3, we proceed as follows. We would like to run the summation over $Y_1 < p_k < q_1 < Y_2$, but if we do so, we must remove those terms for which $q_1^{\lambda^2} < p_k$. This is why we write (recalling the definition of S given in (4.2))

$$S_3 = \sum_{Y_1<p_k<q_1<Y_2} \frac{1}{p_k q_1} - \sum_{\substack{Y_1<p_k<q_1<Y_2 \\ q_1<p_k^{1/\lambda^2}}} \frac{1}{p_k q_1} \tag{7.7}$$

$$= \frac{S^2}{2} - \sum_{Y_1<p_k<Y_2} \frac{1}{p_k} \sum_{p_k<q_1<p_k^{1/\lambda^2}} \frac{1}{q_1}$$

$$= \frac{S^2}{2} - \sum_{Y_1<p_k<Y_2} \frac{1}{p_k}\left(\log\frac{1}{\lambda^2} + O\left(\frac{1}{\log p_k}\right)\right)$$

$$= \frac{S^2}{2} - \log\frac{1}{\lambda^2}\cdot S + O(1).$$

Inserting (7.7) in (7.6), we see that estimate (7.5) can be replaced by

$$S_b^{(2)}(x) = \lambda^4 \varDelta_1 \varDelta_2 x \left(\frac{S^2}{2} - \log\frac{1}{\lambda^2}\cdot S\right) + o(x). \tag{7.8}$$

On the other hand, we have

$$S_b^{(1)}(x) = x\lambda^2 \sum_{\substack{\{p_1,\dots,p_k\}\in\mathcal{K} \\ \{q_1,\dots,q_k\}\in\mathcal{K} \\ \lambda^2<\varLambda(p_k,q_1)<\lambda \\ Y_1<p_1,\, q_k<Y_2}} \frac{\log p_1}{p_1\cdots p_k q_1\cdots q_k \log q_k}. \tag{7.9}$$

Making use of Lemmas 4.2 and 4.3, relation (7.9) becomes

$$S_b^{(1)}(x) = x\varDelta_1\varDelta_2\lambda^2 \sum_{\lambda^2<\varLambda(p_k,q_1)<\lambda} \frac{\log p_1}{p_k q_1 \log q_1} + o(x\log\log x), \tag{7.10}$$

$$= x\varDelta_1\varDelta_2\lambda^2 \sum_{Y_1<q_1<Y_2} (\lambda-\lambda^2)\frac{\log q_1}{q_1\log q_1} + o(x\log\log x)$$

$$= x\varDelta_1\varDelta_2\lambda^2(\lambda-\lambda^2)S + o(x\log\log x).$$

Inserting estimates (7.8) and (7.10) in (7.2), and recalling (6.3), we can replace (7.1) by

$$S(x) = \lambda^4\varDelta_1\varDelta_2 x S^2 + MxS + o(xS), \tag{7.11}$$

where

$$M = 2\lambda^2\varDelta_1\varDelta_2(\lambda - \lambda^2 - 2\lambda^2\log 1/\lambda) + \lambda^2\varDelta_1.$$

Recalling the estimate of S given in Lemma 4.5, we find that the first estimate of Theorem 3.4, namely (3.4), is proved.

To establish that (3.5) follows from (3.4), we will use an approach somewhat similar to the one used to complete the proof of Theorem 3.1. It goes as follows. We first set

$$\sigma_n := V(n) - \widetilde{V}(n). \tag{7.12}$$

Now, assume that $p_1 \cdots p_k \,|\, n$ for some $\{p_1, \dots, p_k\} \in \mathcal{K}$. It is clear that knowing p_1 and n determines the values of $p_2, \dots, p_k$. Similarly, knowing the values of p_k and n will reveal the values of $p_1, \dots, p_{k-1}$. It follows from this observation that

$$\sigma_n \le \sum_{\substack{p|n \\ p<Y_1}} 1 + \sum_{\substack{q|n \\ q>Y_2}} 1 \quad \text{and} \quad \sigma_n^2 \le \sum_{\substack{pq|n \\ p,q\notin(Y_1,Y_2)}} 1.$$

From these two inequalities, it follows that there exist absolute positive constants c_1 and c_2 such that

$$\sum_{n\le x} \sigma_n \le c_1 x \log\log\log x \quad \text{and} \quad \sum_{n\le x} \sigma_n^2 \le c_2 x (\log\log\log x)^2. \tag{7.13}$$

Now, observe that it follows from (7.12) that

$$V^2(n) - \widetilde{V}^2(n) = \sigma_n(\sigma_n + 2\widetilde{V}(n)). \tag{7.14}$$

Observe also that

$$(\sigma_n + 2\widetilde{V}(n))^2 \le 4\sigma_n^2 + 8\widetilde{V}^2(n),$$

implying that, in light of the second inequality in (7.13) and of Theorem 3.3(i), we have

$$\sum_{n\le x} (\sigma_n + 2\widetilde{V}(n))^2 \ll x(\log\log\log x)^2 + x(\log\log x)^2 \ll x(\log\log x)^2. \tag{7.15}$$

Applying the Cauchy–Schwarz inequality to expression (7.14) and using once more the second inequality in (7.13) as well as the upper bound (7.15), we obtain

$$\begin{aligned} \sum_{n\le x} \big(V^2(n) - \widetilde{V}^2(n)\big) &\le \Big(\sum_{n\le x} \sigma_n^2\Big)^{1/2} \cdot \Big(\sum_{n\le x} (\sigma_n + 2\widetilde{V}(n))^2\Big)^{1/2} \\ &\ll \sqrt{x}\, \log\log\log x \cdot \sqrt{x}\, \log\log x \\ &= x \log\log x \cdot \log\log\log x. \end{aligned} \tag{7.16}$$

Combining (7.16) with (3.4) and (7.11) proves (3.5), thus completing the proof of Theorem 3.4.

8. Final remarks. In conclusion, we present an additional result which we state as our fifth theorem.

Theorem 8.1. *We have*

$$\frac{1}{\pi(x)}\sum_{p\le x} V(p+1)=\lambda^2\Delta_1\log\log x+O(\log\log\log x),$$

$$\frac{1}{\pi(x)}\sum_{p\le x} V^2(p+1)=\lambda^4\Delta_1\Delta_2(\log\log x)^2+O(\log\log x\cdot\log\log\log x).$$

We omit the proof, but let us mention that one can use the Bombieri–Vinogradov theorem (already mentioned above in the proof of Theorem 3.2) and then proceed essentially as in the proofs of Theorems 3.3 and 3.4.

Acknowledgements. The work of the first author was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

[1] J.-M. De Koninck and N. Doyon, *The Life of Primes in 37 Episodes*, Amer. Math. Soc., Providence, RI, 2021.

[2] J.-M. De Koninck and J. Galambos, *The intermediate prime divisors of integers*, Proc. Amer. Math. Soc. 101 (1987), 213–216.

[3] J.-M. De Koninck and I. Kátai, *On the consecutive prime divisors of an integer*, Funct. Approx. Comment. Math. 66 (2022), 7–33.

[4] P. Erdős, *On the distribution function of additive functions*, Ann. of Math. 47 (1946), 1–20.

[5] J. Galambos, *The sequences of prime divisors of integers*, Acta Arith. 31 (1976), 213–218.

[6] A. Granville, *Prime divisors are Poisson distributed*, Int. J. Number Theory 3 (2007), 1–18.

[7] A. Granville, *Erratum: "Prime divisors are Poisson distributed" [Int. J. Number Theory 3 (2007), 1–18]*, Int. J. Number Theory 3 (2007), 649–651.

[8] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, Amer. Math. Soc. Colloq. Publ. 53, Amer. Math. Soc., Providence, RI, 2004.

[9] J. B. Rosser and L. Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illinois J. Math. 6 (1962), 64–94.

Jean-Marie De Koninck
Département de mathématiques et de statistique
Université Laval
Québec
Québec G1V 0A6, Canada
E-mail: jmdk@mat.ulaval.ca

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