COLLOQUIUM MATHEMATICUM

VOL. 88 2001 NO. 2

ON THE UNIMODAL CHARACTER OF THE FREQUENCY FUNCTION OF THE LARGEST PRIME FACTOR

BY

JEAN-MARIE DE KONINCK and JASON PIERRE SWEENEY (Québec, PQ)

Abstract. The main objective of this paper is to analyze the unimodal character of the frequency function of the largest prime factor. To do that, let $P(n)$ stand for the largest prime factor of n. Then define $f(x,p) := \#\{n \le x \mid P(n) = p\}$. If $f(x,p)$ is considered as a function of p, for $2 \le p \le x$, the primes in the interval $[2,x]$ belong to three intervals $I_1(x) = [2, v(x)]$, $I_2(x) =]v(x), w(x)[$ and $I_3(x) = [w(x), x]$, with $v(x) < w(x)$, such that $f(x,p)$ increases for $p \in I_1(x)$, reaches its maximum value in $I_2(x)$, in which interval it oscillates, and finally decreases for $p \in I_3(x)$. In fact, we show that $v(x) \ge \sqrt{\log x}$ and $w(x) \le \sqrt{x}$. We also provide several conditions on primes $p \le q$ so that $f(x,p) \ge f(x,q)$.

1. Introduction. For each integer $n \ge 2$, let $P(n)$ stand for its largest prime factor. Given a fixed large number x, for each prime number $p \le x$, let $f(x,p)$ stand for the number of integers $n \in [2,x]$ such that $P(n) = p$,

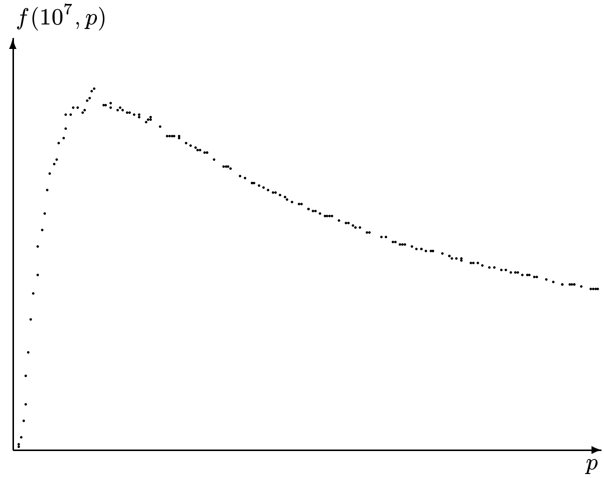

Fig. 1

2000 *Mathematics Subject Classification*: Primary 11N25.

This work was supported in part by the Natural Science and Engineering Research Council of Canada and by the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche du Québec.

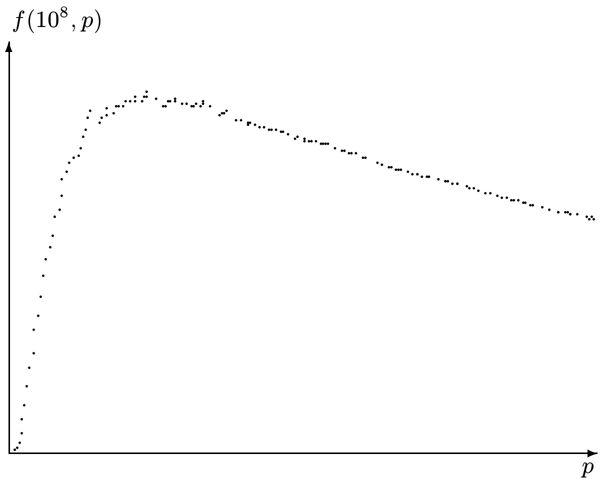

Fig. 2

that is, $f(x,p) := \#\{n \le x \mid P(n) = p\}$. The expression $f(x,p)$ considered as a function of p, for $2 \le p \le x$, has a somewhat smooth behavior, as can be seen in Figures 1 and 2 in the cases $x = 10^7$ and $x = 10^8$, respectively. In fact, for small values of x, it is unimodal in p; such is the case for $f(10^3, p)$, which is increasing for $2 \le p \le 7$, reaches its maximum value 55 at $p = 7$ and decreases thereafter, that is, for $7 < p \le 997$. The function $f(10^4, p)$ is not unimodal because although it reaches its maximum value 224 at $p = 19$, we note that $f(10^4, 23) = 216$, $f(10^4, 29) = 196$ and $f(10^4, 31) = 197$. Even though, as x becomes larger, the function $f(x,p)$ becomes more complicated, it does maintain a unimodal character in the sense that the primes in the interval $[2, x]$ belong to three intervals $I_1(x) = [2, v(x)]$, $I_2(x) =]v(x), w(x)[$ and $I_3(x) = [w(x), x]$, with $v(x) < w(x)$, such that $f(x,p)$ increases for $p \in I_1(x)$, reaches its maximum value in $I_2(x)$, in which interval it oscillates, and finally decreases for $p \in I_3(x)$. For instance, one can establish that the function $f(10^7, p)$ is increasing for $2 \le p \le 89$, attains its maximum at $p = 113$ (with $f(10^7, 113) = 19101$) and decreases for $523 < p \le 9\,999\,991$.

Note that, as was shown in De Koninck [2], the maximum value of $f(x,p)$, for x large and fixed, is attained at some prime $p \in I_2(x)$ satisfying

$$p = e^{(1+o(1))\sqrt{(1/2)\log x \log\log x}}.$$

Our main goal in this paper is to determine bounds on the functions $v(x)$ and $w(x)$. However, we also prove other results pertaining to whether $f(x,p)$ is decreasing or not.

2. Notations. Throughout this paper, p and q stand for prime numbers, while p_ν stands for the νth prime number.

As we shall see in Section 4,

$$p_{\nu+1} - p_\nu < p_{\nu+1}^{\xi} \quad (\nu \geq 1) \tag{2.1}$$

holds for some real number $\xi \leq 0.535$. Now set $\theta = 1/(3-\xi)$, so that $0.333 \leq \theta \leq 0.406$.

Note that if $x > 169$, the interval $]x^\theta, \sqrt{x}[$ contains at least two prime numbers. One can check this using a computer for $169 < x \leq 5\,000\,000$, while for $x > 5\,000\,000$ it follows by Bertrand's postulate since then $x^\theta \leq x^{0.406} < \sqrt{x}/4$.

Given a real number $x > 169$ and a positive integer ν such that

$$x^\theta < p_\nu < p_{\nu+1} < \sqrt{x}, \tag{2.2}$$

define the corresponding values D, R and α as follows:

$$D = D(x,\nu) = \left[\frac{x}{p_\nu}\right] - \left[\frac{x}{p_{\nu+1}}\right], \qquad R = R(x,\nu) = \left[\frac{x}{p_\nu p_{\nu+1}}\right], \tag{2.3}$$

$$\alpha = \alpha(D,R) = \frac{1}{2} - \left(\log\left(\frac{\log D}{\log(D/R)}\right) + \frac{1}{\log D}\right). \tag{2.4}$$

It is a simple matter of algebra to deduce that

$$\frac{x}{p_\nu} - \frac{x}{p_{\nu+1}} - 1 < D < \frac{x}{p_\nu} - \frac{x}{p_{\nu+1}} + 1, \tag{2.5}$$

and from there that $D \geq 2$. We will also use the quantity $d := D/R$, mainly because, as R becomes large, d corresponds approximately to the difference between p_ν and $p_{\nu+1}$:

$$(p_{\nu+1} - p_\nu) - \frac{1}{R} < d < \left(1 + \frac{1}{R}\right)(p_{\nu+1} - p_\nu) + \frac{1}{R}. \tag{2.6}$$

From equation (2.6), we obtain $D > 2R - 1$, and since both numbers are integers, $D \geq 2R$, which is $d \geq 2$.

Finally, given an integer $n \geq 2$, let $\lambda(n)$ stand for the maximum number of prime numbers which can be included in an interval of the form $]z, z+n]$, $z \geq 1$, that is,

$$\lambda(n) = \max_{z \geq 1}(\pi(z+n) - \pi(z)). \tag{2.7}$$

3. Main results. We first examine $f(x,p)$ on the interval $I_1(x) = [2, v(x)]$. Given a fixed large number x, we determine a lower bound for $v(x)$. We shall use known estimates of the associated function $\Psi(x,y) := \#\{n \leq x \mid P(n) \leq y\}$ in our proof of the following result.

THEOREM 1. *For large x and all positive integers ν such that $p_{\nu+1} \le \sqrt{\log x}$,*

$$f(x, p_\nu) < f(x, p_{\nu+1}).$$

So we have $v(x) \ge \sqrt{\log x}$ for large x.

We then examine $f(x, p)$ on the interval $I_2(x) = [v(x), w(x)]$. The behavior of the studied function on this intermediate interval is quite difficult to characterize for two consecutive primes. However, we can establish the following result, which provides conditions on primes $p < q$ that ensure $f(x, p) \ge f(x, q)$. This shows that the maximum of $f(x, p)$ is asymptotically smaller than any power of x, a fact which is confirmed by De Koninck's estimation mentioned in the introduction.

THEOREM 2. *Let a and d be fixed real numbers satisfying $0 < a < d < 1/2$. Let $\xi_1(x)$ and $\xi_2(x)$ be two functions satisfying*

$$x^a < \xi_1(x) < \xi_2(x) < x^d \quad \textit{and} \quad \xi_1(x) = o(\xi_2(x))$$

for all $x > x'$ for a certain real number x'. Then there exists a real number x^* such that for each pair of prime numbers (p, q) such that $x^a < p < \xi_1(x)$ and $\xi_2(x) < q < x^d$,

$$f(x, p) \ge f(x, q) \qquad (x \ge x^*).$$

We now examine $f(x, p)$ on the interval $I_3(x) = [w(x), x]$. We will strive to establish results for two consecutive primes, much more specific than the preceding result. It seems trivial from the definition of $f(x, p)$ that $w(x) \le \sqrt{x}$, but this is not the case. We must work a bit to obtain the following result, mainly due to the $p_\nu < \sqrt{x} \le p_{\nu+1}$ case.

THEOREM 3. *Let x be a positive real number and let ν be a positive integer such that $\sqrt{x} \le p_{\nu+1}$. Then*

$$f(x, p_\nu) \ge f(x, p_{\nu+1}).$$

So we have $w(x) \le \sqrt{x}$ for large x.

The next theorem, which is the main result of the paper, while not establishing a bound on $w(x)$, does provide a precise condition for identifying consecutive primes for which $f(x, p)$ is decreasing. This result also reveals much about the general behavior of $f(x, p)$, as we will see in Theorem 5.

THEOREM 4. *For a real number $x > 169$ and a positive integer ν such that $x^\theta < p_\nu < p_{\nu+1} < \sqrt{x}$ (where θ is defined in "Notations"), if D, R and α are the corresponding values defined by* (2.3) *and* (2.4), *the following hold*:

(i) *with* $\lambda(n)$ *as in* (2.7),

$$f(x,p_\nu)-f(x,p_{\nu+1})\ge D-R-\sum_{i=1}^{R}\lambda\Big(\frac{D}{i}+1\Big);$$

(ii) *if*

$$\alpha\ge\frac{R}{D}, \tag{3.1}$$

then

$$f(x,p_\nu)\ge f(x,p_{\nu+1}).$$

The final theorem is based directly on Theorem 4. It shows that, in a given interval that depends on x, we can prove that $f(x,p)$ is decreasing for consecutive primes simply on the basis of the distance between them.

THEOREM 5. *Let* $x>169$ *and let* ν_0 *be a positive integer satisfying* $x^\theta<p_{\nu_0}<p_{\nu_0+1}<\sqrt{x}$ *and such that the associated values* $D_0=D(x,\nu_0)$, $R_0=R(x,\nu_0)$ *and* $\alpha_0=\alpha(D_0,R_0)$ *defined by* (2.3) *and* (2.4) *satisfy* (3.1). *Assume that a positive integer* ν_1 *satisfies*

$$p_{\nu_0+1}\le p_{\nu_1}<p_{\nu_1+1}<\sqrt{x}\quad\text{and}\quad p_{\nu_1+1}-p_{\nu_1}\ge\frac{D_0}{R_0}+1. \tag{3.2}$$

Then $f(x,p_{\nu_1})\ge f(x,p_{\nu_1+1})$.

4. Preliminary results. We shall be making use of several well known results, the first of which was obtained be Ennola in 1969, the second by Hildebrand in 1985.

THEOREM A (Tenenbaum [7], p. 367). *For* $k\ge1$, $z\ge0$, $a_i>0$,

$$\frac{z^k}{k!}\prod_{i=1}^{k}\frac{1}{a_i}<N_k(z)\le\frac{(z+\sum_{i=1}^{k}a_i)^k}{k!}\prod_{i=1}^{k}\frac{1}{a_i},$$

where

$$N_k(z):=\#\Big\{(v_1,\dots,v_k)\,\Big|\,v_1\ge0,\dots,v_k\ge0,\ \sum_{i=1}^{k}v_ia_i\le z\Big\}.$$

THEOREM B (Hildebrand [4]). *Let* $\varepsilon>0$ *and* $u:=\log x/\log y$. *Then*

$$\Psi(x,y)=x\varrho(u)\Big(1+O_\varepsilon\Big(\frac{\log(u+1)}{\log y}\Big)\Big)$$

uniformly for

$$x\ge3,\qquad 1\le u\le\frac{\log x}{(\log\log x)^{5/3+\varepsilon}}.$$

Here $\varrho(u)$ is the well known Dickman function defined by $\varrho(u) = 1$ $(0 \le u \le 1)$ and

$$\varrho(u) = \varrho(k) - \int_k^u \varrho(\nu - 1)\,\frac{d\nu}{\nu} \qquad (k < u \le k+1).$$

Hence, Dickman's function is a solution of $u\varrho'(u) + \varrho(u-1) = 0$ $(u > 1)$.

Theorem C (Baker & Harman [1]). *There exists a real number $\xi \le 0.535$ such that for all integers $\nu \ge 1$,*

$$p_{\nu+1} - p_\nu \le p_\nu^\xi < p_{\nu+1}^\xi. \tag{4.1}$$

Theorem D (Montgomery & Vaughan [6]). *For all real numbers $y \ge 2$,*

$$\lambda(y) \le \frac{2y}{\log y}.$$

5. Proof of main results

5.1. *Proof of Theorem 1.* We begin with the simple identity

$$f(x, p_\nu) = \Psi(x, p_\nu) - \Psi(x, p_{\nu-1}),$$

which yields

$$f(x, p_{\nu+1}) - f(x, p_\nu) = \Psi(x, p_{\nu+1}) + \Psi(x, p_{\nu-1}) - 2\Psi(x, p_\nu).$$

So, to establish the growth of $f(x, p)$, it will be sufficient to prove that

$$\Psi(x, p_{\nu+1}) > 2\Psi(x, p_\nu). \tag{5.1}$$

Theorem A provides us with the following bounds for $\Psi(x, p_\nu)$:

$$\frac{(\log x)^\nu}{\nu!}\prod_{i=1}^{\nu}\frac{1}{\log p_i} < \Psi(x, p_\nu) \le \frac{(\log x + \sum_{i=1}^{\nu}\log p_i)^\nu}{\nu!}\prod_{i=1}^{\nu}\frac{1}{\log p_i}.$$

From this, to establish (5.1), we need only prove that

$$2\frac{(\log x + \sum_{i=1}^{\nu}\log p_i)^\nu}{\nu!}\prod_{i=1}^{\nu}\frac{1}{\log p_i} < \frac{(\log x)^{\nu+1}}{(\nu+1)!}\prod_{i=1}^{\nu+1}\frac{1}{\log p_i}$$

or, equivalently,

$$2(\nu+1)\left(1 + \frac{\sum_{i=1}^{\nu}\log p_i}{\log x}\right)^\nu < \frac{\log x}{\log p_{\nu+1}}. \tag{5.2}$$

For x large, since $p_{\nu+1} \le \sqrt{\log x}$, we have $\log p_{\nu+1} \le (\log\log x)/2$ and $\nu < \nu + 1 < p_\nu < \sqrt{\log x}$. Hence, using the estimate $\prod_{i=1}^{r} p_i \le 3^{p_r}$, $r = 1, 2, \ldots$ (see Hanson [3]), we deduce that

$$
\begin{aligned}
2(\nu+1)&\left(1+\frac{\sum_{i=1}^{\nu}\log p_i}{\log x}\right)^{\nu}\\
&=2(\nu+1)\left(1+\frac{\log\prod_{i=1}^{\nu}p_i}{\log x}\right)^{\nu}\le 2(\nu+1)\left(1+\frac{p_\nu\log 3}{\log x}\right)^{\nu}\\
&=2(\nu+1)\exp\left(\nu\log\left(1+\frac{p_\nu\log 3}{\log x}\right)\right)\le 2(\nu+1)\exp\left(\frac{\nu p_\nu\log 3}{\log x}\right)\\
&<2\sqrt{\log x}\exp(\log 3)=6\sqrt{\log x}\\
&<\frac{2\sqrt{\log x}}{\log\log x}\cdot\sqrt{\log x}=\frac{2\log x}{\log\log x}\le\frac{\log x}{\log p_{\nu+1}},
\end{aligned}
$$

which establishes (5.2) and the result is proven.

5.2. *Proof of Theorem 2.* The proof is based on the estimate of the $\Psi(x,y)$ function given by Theorem B. The admissibility region of this theorem is equivalent to

$$(\log\log x)^{5/3+\varepsilon}\le\log y\le\log x.$$

We will examine, for large x, the region $x^\delta\le y\le x$ with $\delta>0$. For a fixed value of $\varepsilon>0$, there exists $x_0=x_0(\delta)$ such that

$$(\log\log x)^{5/3+\varepsilon}<\delta\log x\le\log y\qquad(x\ge x_0).$$

The estimate of Theorem B is then valid uniformly in the region $x^\delta\le y\le x$ with $x\ge x_0$.

We now notice that, in this last region,

$$
\begin{aligned}
\frac{\log(u+1)}{\log y}&=\frac{\log\left(\dfrac{\log x}{\log y}+1\right)}{\log y}\le\frac{\log\left(\dfrac{\log x}{\log x^\delta}+1\right)}{\log x^\delta}\\
&=\frac{\log(1/\delta+1)}{\delta\log x}=O_\delta\left(\frac{1}{\log x}\right).
\end{aligned}
$$

We can now state a modified version of Hildebrand's Theorem.

THEOREM B′. *For all $\delta>0$, there exists $x_0=x_0(\delta)$ such that*

$$\Psi(x,y)=x\varrho(u)\left(1+O_\delta\left(\frac{1}{\log x}\right)\right)$$

uniformly for

$$x\ge x_0,\qquad x^\delta\le y\le x\qquad(\textit{or equivalently }1\le u\le 1/\delta).$$

Converting this approximation to the function $f(x,p)$, we obtain

$$f(x,p)=\Psi\left(\frac{x}{p},p\right)=\frac{x}{p}\varrho\left(\frac{\log x}{\log p}-1\right)\left(1+O_\delta\left(\frac{1}{\log x}\right)\right),$$

provided $x^\delta \le p \le x^\gamma$ $(0 < \delta < \gamma < 1/2)$ and $x \ge x_0$. Here, with $\delta' = \delta/(1+\delta)$ we have $(x/p)^{\delta'} \le p$, and since $p^2 \le x^{2\gamma} < x$, we have $p \le x/p$.

Let δ and γ be two fixed real numbers such that $0 < \delta < \gamma < 1/2$. There then exists a fixed number x_0 that renders the approximation of Theorem B$'$ valid. Let a and d be fixed numbers such that $\delta < a < d < \gamma$. Suppose that p and q are two prime numbers such that

$$x^a < p < \xi_1(x) \quad \text{and} \quad \xi_2(x) < q < x^d,$$

where

$$x^a < \xi_1(x) < \xi_2(x) < x^d \quad \text{and} \quad \xi_1(x) = o(\xi_2(x))$$

for all $x > x'$ for a certain fixed real number x'. Then, provided that $x \ge x_1$, where $x_1 = \max\{x_0, x'\}$, we have

$$\begin{aligned}\Psi\left(\frac{x}{p}, p\right) &= \frac{x}{p}\varrho\left(\frac{\log x}{\log p} - 1\right)\left(1 + O_\delta\left(\frac{1}{\log x}\right)\right)\\ &\ge \frac{x}{\xi_1(x)}\varrho\left(\frac{1}{a} - 1\right)\left(1 - \frac{M}{\log x}\right)\end{aligned}$$

and

$$\begin{aligned}\Psi\left(\frac{x}{q}, q\right) &= \frac{x}{q}\varrho\left(\frac{\log x}{\log q} - 1\right)\left(1 + O_\delta\left(\frac{1}{\log x}\right)\right)\\ &\le \frac{x}{\xi_2(x)}\varrho\left(\frac{1}{d} - 1\right)\left(1 + \frac{M}{\log x}\right)\end{aligned}$$

where M is a fixed constant depending only on δ and γ. Hence there exists $x_1 > x_0$ such that

$$\Psi\left(\frac{x}{p}, p\right) \ge \frac{x}{\xi_1(x)}\varrho\left(\frac{1}{a} - 1\right) \cdot 0.99$$

and

$$\Psi\left(\frac{x}{q}, q\right) \le \frac{x}{\xi_2(x)}\varrho\left(\frac{1}{d} - 1\right) \cdot 1.01$$

when $x \ge x_1$. Since $\xi_1(x) = o(\xi_2(x))$, there exists $x_2 > x_1$ such that

$$f(x,p) \ge f(x,q) \qquad (x \ge x_2).$$

Choosing $\delta = a/2$ and $\gamma = (1/2+d)/2$, we obtain the assertion of Theorem 2.

5.3. *Proof of Theorem 3.* We first determine a representation for $f(x,p)$ in the region $x^{1/2} \le p \le x$. By definition,

$$f(x,p) = \Psi\left(\frac{x}{p}, p\right), \tag{5.3}$$

and

$$\Psi(x,y) = [x] \qquad (y \ge x). \tag{5.4}$$

Combining (5.3) and (5.4), we easily obtain

$$f(x,p_\nu)=\left[\frac{x}{p_\nu}\right]\qquad (p_\nu\ge\sqrt{x}). \tag{5.5}$$

Therefore if $x^{1/2}\le p_\nu\le p_{\nu+1}\le x$, then

$$f(x,p_\nu)\ge f(x,p_{\nu+1}).$$

Hence in the region $\sqrt{x}\le p\le x$, $f(x,p)$ is a decreasing function of p.

To complete the proof of the theorem, we must examine the particular case where $p_\nu<\sqrt{x}\le p_{\nu+1}$. It is then necessary to extend the covered interval to the left of $\sqrt{x}$. We will do that by using Buchstab's identity

$$\Psi(x,y)=\Psi(x,z)-\sum_{y<p\le z}\Psi\left(\frac{x}{p},p\right)\qquad (x\ge 1,\ z\ge y\ge 1). \tag{5.6}$$

Notice that this identity is a direct consequence of (5.3). We now consider the region $x^{1/3}<p_\nu<x^{1/2}$. Replacing x by x/p_ν, y by p and z by x/p_ν in (5.6) we obtain

$$\begin{aligned}\Psi\left(\frac{x}{p_\nu},p_\nu\right)&=\Psi\left(\frac{x}{p_\nu},\frac{x}{p_\nu}\right)-\sum_{p_\nu<p\le x/p_\nu}\Psi\left(\frac{x}{p_\nu p},p\right)\\&=\left[\frac{x}{p_\nu}\right]-\sum_{p_\nu<p\le x/p_\nu}\left[\frac{x}{p_\nu p}\right],\end{aligned} \tag{5.7}$$

where we used the fact that $x/p^2<p_\nu$.

We now return to the non-trivial case of $p_\nu<\sqrt{x}\le p_{\nu+1}$. In this situation, by (5.5) and (5.7),

$$f(x,p_{\nu+1})=\left[\frac{x}{p_{\nu+1}}\right]\quad\text{and}\quad f(x,p_\nu)=\left[\frac{x}{p_\nu}\right]-\sum_{p_\nu<p\le x/p_\nu}\left[\frac{x}{p_\nu p}\right].$$

To justify the use of (5.7), we observe that

$$p_\nu\ge\frac{p_{\nu+1}}{2}\ge\frac{\sqrt{x}}{2}=\frac{x^{1/3}x^{1/6}}{2}\ge x^{1/3}\qquad (x\ge 64), \tag{5.8}$$

where the first inequality is true by Bertrand's postulate.

We now consider the terms of the summation on the right side of (5.7). Since $p_{\nu+1}\le p\le x/p_\nu$, and $p_{\nu+1}<2p_\nu$, once again by Bertrand's postulate, we obtain

$$1\le\frac{x}{p_\nu p}\le\frac{x}{p_\nu p_{\nu+1}}<\frac{2x}{p_{\nu+1}^2}\le 2,$$

from which it follows that

$$\left[\frac{x}{p_\nu p}\right]=1.$$

There are now three distinct cases to consider, namely $p_{\nu+1} < [x/p_\nu]$, $p_{\nu+1} > [x/p_\nu]$ and $p_{\nu+1} = [x/p_\nu]$.

In the first case,

$$\sum_{p_\nu < p \le x/p_\nu} \left[\frac{x}{p_\nu p}\right] = \sum_{p_{\nu+1} \le p \le [x/p_\nu]} 1 \le \left[\frac{x}{p_\nu}\right] - p_{\nu+1},$$

the last inequality being valid since there are $[x/p_\nu] - p_{\nu+1} + 1$ admissible numbers in the second summation, at least one of which is not prime. Hence

$$f(x,p_\nu) - f(x,p_{\nu+1}) \ge \left[\frac{x}{p_\nu}\right] - \left(\left[\frac{x}{p_\nu}\right] - p_{\nu+1}\right) - \left[\frac{x}{p_{\nu+1}}\right] = p_{\nu+1} - \left[\frac{x}{p_{\nu+1}}\right] \ge 0$$

since $p_{\nu+1} \ge \sqrt{x}$.

In the second case,

$$\sum_{p_\nu < p \le x/p_\nu} \left[\frac{x}{p_\nu p}\right] = 0,$$

because the sum is empty. We then easily conclude that

$$f(x,p_\nu) - f(x,p_{\nu+1}) = \left[\frac{x}{p_\nu}\right] - \left[\frac{x}{p_{\nu+1}}\right] \ge 0.$$

In the last case,

$$\sum_{p_\nu < p \le x/p_\nu} \left[\frac{x}{p_\nu p}\right] = \sum_{p_{\nu+1} \le p \le [x/p_\nu]} 1 = 1,$$

which implies that

$$f(x,p_\nu) - f(x,p_{\nu+1}) = \left[\frac{x}{p_\nu}\right] - 1 - \left[\frac{x}{p_{\nu+1}}\right] = \left(p_{\nu+1} - \left[\frac{x}{p_{\nu+1}}\right]\right) - 1 \ge 0,$$

the last inequality being valid if and only if $p_{\nu+1} > \sqrt{x}$, because we then have $[x/p_{\nu+1}] < p_{\nu+1}$. So we must eliminate the possibility that, in this third case, $p_{\nu+1} = \sqrt{x}$. To do that, let us consider the corresponding value of p_ν. Supposing that $p_{\nu+1} = \sqrt{x}$, we have $p_\nu \le \sqrt{x} - 2$, which implies that

$$\frac{x}{p_\nu} \ge \frac{x}{\sqrt{x} - 2} > \sqrt{x} + 1 = p_{\nu+1} + 1.$$

As this is contrary to the initial hypothesis $p_{\nu+1} = [x/p_\nu]$, we may conclude that, in this third case, $p_{\nu+1}$ must be larger than $\sqrt{x}$.

In view of (5.8), this completes the proof of Theorem 3 for $x \ge 64$. The property follows for all $x \ge 1$ by inspection.

5.4. *Proof of Theorem 4(i).* In this section, we will build sufficient conditions to ensure the decrease of the function $f(x,p)$ in a subregion of

$$x^{1/3} < p_\nu < p_{\nu+1} < x^{1/2}.$$

We will use the representation given by (5.7), which is valid in this interval, to calculate the following difference:

$$
\begin{aligned}
&f(x,p_\nu) - f(x,p_{\nu+1}) \\
&\quad = \left(\left[\frac{x}{p_\nu}\right] - \sum_{p_\nu<p\le x/p_\nu} \left[\frac{x}{p_\nu p}\right]\right) - \left(\left[\frac{x}{p_{\nu+1}}\right] - \sum_{p_{\nu+1}<p\le x/p_{\nu+1}} \left[\frac{x}{p_{\nu+1} p}\right]\right) \\
&\quad = \left(\left[\frac{x}{p_\nu}\right] - \left[\frac{x}{p_\nu p_{\nu+1}}\right] - \sum_{p_{\nu+1}<p\le x/p_{\nu+1}} \left[\frac{x}{p_\nu p}\right] - \sum_{x/p_{\nu+1}<p\le x/p_\nu} \left[\frac{x}{p_\nu p}\right]\right) \\
&\qquad - \left(\left[\frac{x}{p_{\nu+1}}\right] - \sum_{p_{\nu+1}<p\le x/p_{\nu+1}} \left[\frac{x}{p_{\nu+1} p}\right]\right) \\
&\quad = \left(\left[\frac{x}{p_\nu}\right] - \left[\frac{x}{p_{\nu+1}}\right] - \left[\frac{x}{p_\nu p_{\nu+1}}\right]\right) \\
&\qquad - \sum_{p_{\nu+1}<p\le x/p_{\nu+1}} \left(\left[\frac{x}{p_\nu p}\right] - \left[\frac{x}{p_{\nu+1} p}\right]\right) - \sum_{x/p_{\nu+1}<p\le x/p_\nu} \left[\frac{x}{p_\nu p}\right] \\
&\quad = A - B - C,
\end{aligned}
$$

say. Our main objective is to establish a lower bound for $A-B-C$. We first have $A = D - R$, where D and R are defined in (2.3).

To estimate B, we split up the interval $]p_{\nu+1}, x/p_{\nu+1}]$ using the following partition:

$$\left\{\frac{x}{(N_0+1)p_\nu}, \frac{x}{N_0 p_{\nu+1}}, \frac{x}{N_0 p_\nu}, \frac{x}{(N_0-1)p_{\nu+1}}, \dots, \frac{x}{3p_\nu}, \frac{x}{2p_{\nu+1}}, \frac{x}{2p_\nu}, \frac{x}{p_{\nu+1}}\right\}.$$

We can determine the exact value of N_0 by noticing that since

$$R \le \frac{x}{p_\nu p_{\nu+1}} < R+1,$$

we have

$$\frac{x}{(R+1)\,p_\nu} < p_{\nu+1} \le \frac{x}{Rp_\nu},$$

so that $N_0 = R$.

We must also determine a condition that will guarantee that all subintervals will be distinct. It turns out that condition (2.2) is sufficient since, as $\theta = 1/(3-\xi)$,

$$R \le \frac{x}{p_\nu p_{\nu+1}} < x^{1-2\theta} < p_{\nu+1}^{(1-2\theta)/\theta} = p_{\nu+1}^{1-\xi} \le \frac{p_{\nu+1}}{p_{\nu+1} - p_\nu},$$

the last inequality being true because of Theorem C. Hence

$$R(p_{\nu+1} - p_\nu) < p_{\nu+1} \iff (R-1)p_{\nu+1} < Rp_\nu \iff \frac{x}{Rp_\nu} < \frac{x}{(R-1)p_{\nu+1}}.$$

Now that we have confirmed the validity of our partition, let us set

$$I = \bigcup_{i=1}^{R} \left] \frac{x}{(i+1)p_\nu}, \frac{x}{ip_{\nu+1}} \right] \quad \text{and} \quad I' = \bigcup_{i=2}^{R} \left] \frac{x}{ip_{\nu+1}}, \frac{x}{ip_\nu} \right],$$

so that

$$\left] p_{\nu+1}, \frac{x}{p_{\nu+1}} \right] \subset I \cup I'. \tag{5.9}$$

Let $p \in I$, say

$$p \in \left] \frac{x}{(k+1)p_\nu}, \frac{x}{kp_{\nu+1}} \right]$$

for a certain k, $1 \le k \le R$. We then have $[x/(p_\nu p)] = k$ and $[x/(p_{\nu+1}p)] = k$, and so

$$\left[\frac{x}{p_\nu p} \right] - \left[\frac{x}{p_{\nu+1} p} \right] = 0. \tag{5.10}$$

Now, let $p \in I'$, say

$$p \in \left] \frac{x}{lp_{\nu+1}}, \frac{x}{lp_\nu} \right]$$

for a certain l, $2 \le l \le R$. In this case, $[x/(p_\nu p)] = l$ and $[x/(p_{\nu+1}p)] = l-1$, so that

$$\left[\frac{x}{p_\nu p} \right] - \left[\frac{x}{p_{\nu+1} p} \right] = 1. \tag{5.11}$$

By (5.9)–(5.11), an upper estimate for B is obtained in the following way:

$$\begin{aligned} B &\le \sum_{p \in I} \left(\left[\frac{x}{p_\nu p} \right] - \left[\frac{x}{p_{\nu+1} p} \right] \right) + \sum_{p \in I'} \left(\left[\frac{x}{p_\nu p} \right] - \left[\frac{x}{p_{\nu+1} p} \right] \right) \\ &= \sum_{p \in I} 0 + \sum_{p \in I'} 1 = \sum_{i=2}^{R} \sum_{p \in]x/(ip_{\nu+1}), x/(ip_\nu)]} 1 \le \sum_{i=2}^{R} \lambda\left(\frac{x}{ip_\nu} - \frac{x}{ip_{\nu+1}} \right), \end{aligned} \tag{5.12}$$

where $\lambda(n)$ is defined in (2.7).

To estimate C, observe that for any prime $p \in]x/p_{\nu+1}, x/p_\nu]$, the inequalities

$$\left[\frac{x}{p_\nu p} \right] \ge \left[\frac{x}{p_\nu} \cdot \frac{p_\nu}{x} \right] = 1 \quad \text{and} \quad \left[\frac{x}{p_\nu p} \right] \le \left[\frac{x}{p_\nu} \cdot \frac{p_{\nu+1}}{x} \right] = \left[\frac{p_{\nu+1}}{p_\nu} \right] < 2$$

hold, the second one being true by Bertrand's postulate. It follows that each term in C is equal to 1, that is,

$$C = \sum_{x/p_{\nu+1} < p \le x/p_\nu} \left[\frac{x}{p_\nu p} \right] = \sum_{p \in]x/p_{\nu+1}, x/p_\nu]} 1 \le \lambda\left(\frac{x}{p_\nu} - \frac{x}{p_{\nu+1}} \right). \tag{5.13}$$

Gathering (2.3), (5.12) and (5.13), we obtain

$$\begin{aligned} f(x,p_\nu) - f(x,p_{\nu+1}) &= A - B - C \\ &\geq D - R - \sum_{i=1}^{R} \lambda\left(\frac{x}{ip_\nu} - \frac{x}{ip_{\nu+1}}\right). \end{aligned} \tag{5.14}$$

To complete the proof of Theorem 4(i), we shall transform the terms of the type $x/(ip_\nu) - x/(ip_{\nu+1})$ inherent in D. Using (2.5), we have

$$\left|\frac{D}{i} - \left(\frac{x}{ip_\nu} - \frac{x}{ip_{\nu+1}}\right)\right| < 1 \qquad (1 \leq i \leq R).$$

for which we obtain the important inequality

$$\lambda\left(\frac{x}{ip_\nu} - \frac{x}{ip_{\nu+1}}\right) \leq \lambda\left(\frac{D}{i} + 1\right) \qquad (1 \leq i \leq R). \tag{5.15}$$

Thus, in view of (5.14), the proof of Theorem 4(i) is complete.

5.5. *Proof of Theorem 4(ii).* The main objective is to establish $f(x,p_\nu) - f(x,p_{\nu+1}) \geq 0$. By part (i) of the theorem, it is sufficient for this to have

$$D - 2R - \sum_{i=1}^{R} \lambda\left(\frac{D}{i}\right) \geq 0. \tag{5.16}$$

where we used the trivial inequality $\lambda(n+1) \leq \lambda(n) + 1$. The following reasoning is true for all values of D and R. However, the end result is only significant when D and R are large. When these quantities are small, say $D < 2e^4$, then inequality (5.16) is best handled computationally, since the required values of $\lambda(n)$ can easily be calculated individually.

To ensure that (5.16) holds, first observe that, by Theorem D,

$$\begin{aligned} D - 2R - \sum_{i=1}^{R} \lambda\left(\frac{D}{i}\right) &\geq D - 2R - 2D\sum_{i=1}^{R} \frac{1}{i\log(D/i)} \\ &= D - 2R - 2D\Sigma_R, \end{aligned} \tag{5.17}$$

say. Notice that Σ_R exists because $D \geq 2$ (see "Notations"). To simplify Σ_R, let

$$s(x) = \frac{1}{x\log(D/x)}.$$

Since

$$s'(x) = \frac{-1}{\{x\log(D/x)\}^2}\left\{\log\left(\frac{D}{x}\right) + 1\right\} < 0 \qquad (x \in [1,R]),$$

it follows that

$$\begin{aligned} \Sigma_R = \sum_{i=1}^{R} s(i) &\leq \int_1^R s(t)\,dt + s(1) = -\log\log\left(\frac{D}{t}\right)\Big|_1^R + \frac{1}{\log D} \\ &= \log\log D - \log\log\left(\frac{D}{R}\right) + \frac{1}{\log D} = \log\left(\frac{\log D}{\log(D/R)}\right) + \frac{1}{\log D}. \end{aligned}$$

Combining this result with (5.17), we obtain

$$D-2R-\sum_{i=1}^{R}\lambda\left(\frac{D}{i}\right)\geq D-2R-2D\left[\log\left(\frac{\log D}{\log(D/R)}\right)+\frac{1}{\log D}\right].$$

We have thus established that (5.16) holds if

$$\frac{1}{2}-\left(\log\left(\frac{\log D}{\log(D/R)}\right)+\frac{1}{\log D}\right)\geq\frac{R}{D}.$$

This completes the proof of Theorem 4(ii).

5.6. *Proof of Theorem 5.* We will make use of the following lemma.

LEMMA. *Let* $x>169$ *and let* ν_0 *be a positive integer satisfying* $x^\theta<p_{\nu_0}<p_{\nu_0+1}<\sqrt{x}$ *and such that the associated values* $D_0=D(x,\nu_0)$, $R_0=R(x,\nu_0)$ *and* $\alpha_0=\alpha(D_0,R_0)$ *defined by* (2.3) *and* (2.4) *satisfy* (3.1). *If* ν_1 *is a positive integer satisfying* $x^\theta<p_{\nu_1}<p_{\nu_1+1}<\sqrt{x}$, *then* (3.1) *also holds for the corresponding values* $D_1=D(x,\nu_1)$, $R_1=R(x,\nu_1)$ *and* $\alpha_1=\alpha(D_1,R_1)$ *provided* $R_1\leq R_0$ *and* $d_1\geq\max(d_0,3)$, *where* $d_0=D_0/R_0$ *and* $d_1=D_1/R_1$.

Proof. By hypothesis,

$$\alpha_0\geq\frac{R_0}{D_0},\qquad \alpha_0(D_0,R_0)=\frac{1}{2}-\log\left(\frac{\log D_0}{\log(D_0/R_0)}\right)-\frac{1}{\log D_0}.$$

By definition, $\alpha=\alpha(dR,R)=1/2-g(d,R)$, where

$$g(d,R):=\log\left(\frac{\log R}{\log d}+1\right)+\frac{1}{\log d+\log R}. \tag{5.18}$$

It is a simple matter, using basic calculus, to deduce that:

- For a fixed value of d ($d\geq3$), $g_d(R):=g(d,R)$ is strictly increasing for $R\geq1$.
- For a fixed value of R ($R\geq1$), $g_R(d):=g(d,R)$ is strictly decreasing for $d\geq2$.

Since, by hypothesis, $d_1\geq\max(d_0,3)$ and $R_1\leq R_0$, we have $g(d_1,R_1)\leq g(d_1,R_0)\leq g(d_0,R_0)$, which implies that $\alpha_1\geq\alpha_0$. Condition (3.1) is then met for the pair (D_1,R_1) because

$$\frac{D_1}{R_1}=d_1\geq d_0=\frac{D_0}{R_0}\geq\frac{1}{\alpha_0}\geq\frac{1}{\alpha_1},$$

which completes the proof of the lemma.

We now establish Theorem 5. First, $R_1\leq R_0$ since $p_{\nu_0}p_{\nu_0+1}\leq p_{\nu_1}p_{\nu_1+1}\leq x/2$. Then the hypothesis concerning $p_{\nu_1+1}-p_{\nu_1}$, combined with (2.6), yields $d_0+1\leq p_{\nu_1+1}-p_{\nu_1}<d_1+1/R_1$. As $R_1\geq1$ by construction, we obtain $d_0<d_1$. Since $d_0\geq2$ (see “Notations”), we have $p_{\nu_1+1}-p_{\nu_1}\geq4$ and thus $d_1\geq3$. All the requirements of Theorem 5 are now met. Hence, as (D_1,R_1) respects (3.1), it follows from Theorem 4(ii) that $f(x,p_{\nu_1})\geq f(x,p_{\nu_1+1})$.

6. Final remarks

6.1. *Numerical remarks.* We begin this section with a table of numerical examples. We show here the explicit values of $v(x)$ and $w(x)$ for small powers of 10, along with the corresponding number of oscillations in the intermediate interval $I_2(x) = [v(x), w(x)]$.

x	$v(x)$	# osc. in $I_2(x)$	$w(x)$
10 000	19	0	19
100 000	23	0	23
1 000 000	47	7	73
10 000 000	89	17	199
100 000 000	113	29	463

It would seem reasonable to assume that, for $x > 10^8$, we always have $v(x) < w(x)$, meaning that $f(x,p)$ does oscillate. However, this is certainly very difficult to prove.

6.2. *Illustration of Theorems 4 and 5.* We will illustrate an application of Theorems 4 and 5 in the case of $x = 1\,000\,000$. For this value of x, the interval in which Theorem 4 applies is $[x^\theta, \sqrt{x}] = [273, 1000]$. In this interval, setting $\nu = 128$, we obtain $p_\nu = 719$, $p_{\nu+1} = 727$, $D = 15$, $R = 1$, $R/D = 0.07$, $\alpha = 0.13$ and so condition (3.1) is respected and we must have $f(1\,000\,000, 719) \geq f(1\,000\,000, 727)$. An explicit calculation of these values yields $f(1\,000\,000, 719) = 1297$ and $f(1\,000\,000, 727) = 1284$.

We now have determined a pair of primes for which condition (3.1) holds. According to Theorem 5, it is now sufficient to find consecutive primes larger than 727 with a difference larger than $D/R+1 = 15/1+1 = 16$ to ensure that $f(x,p)$ is decreasing. The pair $(887, 907)$ satisfies these conditions; and we do indeed have $f(1\,000\,000, 887) \geq f(1\,000\,000, 907)$, that is, $1093 > 1073$.

We note that it is possible to establish a relationship between the quantities D and R when applying Theorem 4(ii). Clearly, α must be positive to respect (3.1), which implies that

$$\log\left(\frac{\log D}{\log(D/R)}\right) < \frac{1}{2}.$$

We then obtain the following inequality connecting D and R:

$$D > R^{2.54},$$

which is more revealing when written as $d > R^{1.54}$. This shows again, as in Theorem 5, that the difference between two consecutive primes is closely tied with the behavior of $f(x,p)$. As the difference between two consecutive primes grows, so does the probability that condition (3.1) will hold. In fact, a conjecture that holds numerically for all values of x shown in the above table is the following:

Conjecture. *If* $v(x) \le p_\nu, p_{\nu+1} \le w(x)$ *and* $p_{\nu+1} - p_\nu = 2$, *then* $f(x, p_\nu) \le f(x, p_{\nu+1})$.

Another way of stating this conjecture is that in the interval $I_2(x)$, the function $f(x,p)$ will increase for twin primes.

6.3. *Further remarks assuming other strong conjectures.* Notice that our main results are limited in scope by the first inequality given in Theorem C. For example, if the Riemann Hypothesis is true, then (4.1) holds with $\xi = 1/2 + \varepsilon_0$ for each $\varepsilon_0 > 0$ (see Ivić [5], p. 321), in which case Theorems 4 and 5 hold with $\theta = 4/10 + \varepsilon$, $\varepsilon > 0$. Moreover, if Cramer's conjecture is true, namely if $p_{\nu+1} - p_\nu \ll \log^2 p_\nu$ (see Ivić [5], p. 299), then one can take $\xi = \varepsilon_0$, say, in which case

$$\theta = \frac{1}{3-\xi} = \frac{1}{3-\varepsilon_0} = \frac{1}{3} + \varepsilon, \qquad \varepsilon > 0.$$

This last conjecture provides us with the optimal interval for the setting of Theorem 4, namely $[x^{1/3+\varepsilon}, x]$, with $\varepsilon > 0$.

REFERENCES

[1] R. C. Baker and G. Harman, *The difference between consecutive primes*, Proc. London Math. Soc. (3) 72 (1996), 261–280.

[2] J. M. De Koninck, *On the largest prime divisors of an integer*, in: Extreme Value Theory and Applications, J. Galambos *et al.* (eds.), Kluwer, 1994, 447–462.

[3] D. Hanson, *On the product of primes*, Canad. Math. Bull. 15 (1972), 33–37.

[4] A. Hildebrand, *On the number of positive integers* $\le x$ *and free of prime factors* $> y$, J. Number Theory 22 (1986), 289–307.

[5] A. Ivić, *The Riemann Zeta-Function*, Wiley, 1985.

[6] H. L. Montgomery and R. C. Vaughan, *The large sieve*, Mathematika 20 (1973), 119–134.

[7] G. Tenenbaum, *Introduction à la théorie analytique et probabiliste des nombres*, Cours Spéc. 1, Soc. Math. France, 1995.

Département de Mathématiques et de Statistique
Université Laval
Québec G1K 7P4, Canada
E-mail: jmdk@mat.ulaval.ca

Received 18 May 1999;
revised 20 June 2000 (3761)