ACTA ARITHMETICA
146.3 (2011)

Exponential sums involving the largest prime factor function

by

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

1. Introduction. Let $P(n)$ stand for the largest prime factor of the integer $n \geq 2$ and set $P(1) = 1$. A well known result of I. M. Vinogradov [7] asserts that, given any irrational number α, the sequence αp_n, $n = 1, 2, \ldots$, where p_n stands for the nth prime, is uniformly distributed in $[0, 1]$. In 2005, Banks, Harman and Shparlinski [1] proved that for every irrational number α, the sequence $\alpha P(n)$, $n = 1, 2, \ldots$, is uniformly distributed mod 1. They did so by using the well known Weyl criteria (see the book of Kuipers and Niederreiter [5]) and thus by establishing that

$$\lim_{x\to\infty} \frac{1}{x} \sum_{n\leq x} e(\alpha P(n)) = 0, \tag{1.1}$$

where $e(z) := \exp\{2\pi i z\}$.

Let $\mathcal{M}$ stand for the set of all complex-valued multiplicative functions and let $\widetilde{\mathcal{M}}$ be the subset of those functions $f \in \mathcal{M}$ such that $|f(n)| \leq 1$ for positive integers n. Daboussi (see Daboussi and Delange [2]) proved that given $f \in \widetilde{\mathcal{M}}$ and any irrational number α,

$$\lim_{x\to\infty} \sup_{f\in\widetilde{\mathcal{M}}} \frac{1}{x} \sum_{n\leq x} f(n) e(n\alpha) = 0.$$

Let $\mathcal{M}_1$ stand for the subset of those functions $f \in \mathcal{M}$ such that $|f(n)| = 1$ for all positive integers n. In this paper, we first generalize (1.1) by showing that for any irrational number α and any function $f \in \mathcal{M}_1$, we have

$$\sum_{n\leq x} f(n) e(\alpha P(n)) = o(x) \quad (x \to \infty). \tag{1.2}$$

We also show that this general result further holds if one replaces $e(\alpha P(n))$ by $T(P(n))$, where T is any function defined on primes satisfying $|T(p)| = 1$

2010 *Mathematics Subject Classification*: 11N13, 11L07.
Key words and phrases: largest prime factor, exponential sums, uniform distribution.

for all primes p and such that $\sum_{p\le x} T(p) = o(\pi(x))$, where $\pi(x)$ stands for the number of primes $\le x$.

We then move our interest to shifted primes by establishing that (1.2) holds if one replaces $P(n)$ by $P(n-1)$, provided $f \in \mathcal{M}_1$ satisfies an additional condition.

Finally, we examine the counting function

$$E(x,q,a) := \#\{p \le x : P(p-1) \equiv a \pmod q\}.$$

In [1], Banks, Harman and Shparlinski proved that

$$E(x,q,a) \ll \frac{\mathrm{li}(x)}{\phi(q)} \qquad (\log q \le (\log x)^{1/3}),$$

where the constant implicit in $\ll$ is absolute,

$$\mathrm{li}(x) := \int_2^x \frac{dt}{\log t}$$

and ϕ stands for the Euler function. They also mentioned that the matching lower bound $E(x,q,a) \gg \mathrm{li}(x)/\phi(q)$ should most likely hold as well, but could not prove it. Here we prove their guess to be true.

In what follows, $c, c_1, c_2, \dots$ always denote absolute real constants.

2. Main results

Theorem 1. *Given an irrational number α and a function $f \in \mathcal{M}_1$,*

$$\lim_{x\to\infty} \frac{1}{x} \sum_{n\le x} f(n) e(\alpha P(n)) = 0,$$

where $e(z) := \exp\{2\pi i z\}$.

Theorem 2. *Let $f \in \mathcal{M}_1$ and let $\wp$ stand for the set of all prime numbers. Let $T : \wp \to \mathbb{C}$ be such that $|T(p)| = 1$ for each $p \in \wp$ and such that $\sum_{p\le x} T(p) = o(\pi(x))$, where $\pi(x)$ stands for the number of primes not exceeding x. Then*

$$\lim_{x\to\infty} \frac{1}{x} \sum_{n\le x} f(n) T(P(n)) = 0.$$

Note that one can show that Theorems 1 and 2 remain valid when replacing $P(n)$ by $P_k(n)$, the kth largest prime factor of n.

Theorem 3. *Given an arbitrary fixed number $A > 0$, there exists an absolute constant $c > 0$ such that, for all $x \ge 2$,*

$$E(x,q,a) \ge c\frac{\mathrm{li}(x)}{\phi(q)} \qquad ((a,q) = 1,\ q \le (\log x)^A).$$

THEOREM 4. *Let $f \in \mathcal{M}_1$ and assume that*

$$\mathsf{S}(t) := \sum_p \frac{1 - \Re(f(p)p^{-it})}{p}$$

converges for some $t \in \mathbb{R}$. Then, given any irrational number α,

$$\lim_{x\to\infty} \sum_{n\le x} f(n)\, e(\alpha P(n-1)) = 0.$$

3. Preliminary results. The following two lemmas are essentially due to Halász [4]. We state them as follows.

LEMMA 1. *Let $f \in \mathcal{M}$ with $|f(n)| \le 1$ for all $n \in \mathbb{N}$. Assume that the series $\mathsf{S}(a_0)$ is convergent for some real number a_0. Then there exists a constant $C_0 \in \mathbb{C}$ and a slowly oscillating function $L_0(u)$, with $|L_0(u)| = 1$, such that*

$$\sum_{n\le x} f(n) = C_0\, L_0(\log x)\, x^{1+ia_0} + o(x).$$

REMARK. Observe that the constant C_0 is nonzero if there exists at least one integer $r \ge 0$ for which $f(2^r) \ne -1$.

LEMMA 2. *Let $f \in \mathcal{M}$ with $|f(n)| \le 1$ for all $n \in \mathbb{N}$. Then*

$$\sum_{n\le x} f(n) = o(x)$$

if $\mathsf{S}(b)$ diverges for every real number b or if $f(2^r) = -1$ for $r = 1, 2, \ldots$.

The next lemma, which may be of independent interest, plays a crucial role in what follows.

LEMMA 3. *Let $(a(n))_{n\ge1}$ be a sequence of complex numbers of modulus 1 and set $A(x) := \sum_{n\le x} a(n)$. Also let $\tau \in \mathbb{R}$ and set $A_\tau(x) := \sum_{n\le x} a(n)n^{i\tau}$. If $A(x) = o(x)$, then $A_\tau(x) = o(x)$.*

REMARK. As a consequence of Lemma 3, it follows that if $A_{\tau_1}(x) = o(x)$ for some real number τ_1, then $A_\tau(x) = o(x)$ for every real number τ.

Proof of Lemma 3. Since $A(x) = o(x)$, there exist decreasing functions $\varepsilon(x)$ and $\delta(x)$, both tending to 0 as $x \to \infty$, such that

$$|A(x+y) - A(x)| \le \delta(x)y, \tag{3.1}$$

uniformly for $\varepsilon(x)x \le y \le x$.

Now observe that

$$A_\tau(x+y) - A_\tau(x) = x^{i\tau} \sum_{x<n\le x+y} a(n) e^{i\tau \log(n/x)}$$
$$= x^{i\tau}(A(x+y) - A(x)) + O\Big(|\tau| \sum_{x<n\le x+y} \log \frac{n}{x}\Big).$$

Therefore,

$$|A_\tau(x+y) - A_\tau(x)| \le |A(x+y) - A(x)| + c_1|\tau|\frac{y^2}{x}. \tag{3.2}$$

We shall now prove that

$$\limsup_{X\to\infty} \frac{|A_\tau(X)|}{X} = 0. \tag{3.3}$$

To do so, we first let $M > 0$ be an arbitrarily large integer and choose X large enough so that we have both $\delta(X/M) < 1/M^2$ and $\varepsilon(X/M) < 1/M^2$. Finally let $x = X/M$. Since

$$A_\tau(Mx) = A_\tau(x) + \sum_{j=2}^{M} (A_\tau(jx) - A_\tau((j-1)x)),$$

it follows, in light of (3.1) and (3.2), that

$$|A_\tau(Mx)| \le |A_\tau(x)| + \sum_{j=2}^{M} |A_\tau(jx) - A_\tau((j-1)x)|$$
$$\le x + \sum_{j=1}^{M-1} x\delta(jx) + c_1|\tau| x \sum_{j=1}^{M-1} \frac{1}{j} \le x + xM\delta(x) + c_2 x|\tau| \log M,$$

from which it follows that

$$\frac{|A_\tau(Mx)|}{Mx} \le \frac{1}{M} + \delta(x) + c_2|\tau|\frac{\log M}{M},$$

which in turn implies

$$\limsup_{X\to\infty} \frac{|A_\tau(X)|}{X} \le c_3|\tau|\frac{\log M}{M}.$$

Since M can be taken arbitrarily large, (3.3) follows, thus completing the proof of Lemma 3.

4. The proofs of Theorems 1 and 2. Let $f \in \mathcal{M}_1$, α an irrational number and $S(x) := \sum_{n\le x} f(n)$. Assume for now that f is completely multiplicative. We shall consider separately the two cases

$$\text{(i)} \ \lim_{x\to\infty} \frac{S(x)}{x} = 0, \qquad \text{(ii)} \ \frac{S(x)}{x} \not\to 0 \text{ as } x \to \infty.$$

It is well known (see Tenenbaum [6]) that

$$\psi(x,y) := \#\{n \le x : P(n) \le y\} = (1+o(1))x\rho(u) \qquad (x \to \infty), \tag{4.1}$$

where $\rho(u)$ stands for the Dickman function and $u := (\log x)/(\log y)$ is fixed.

Therefore, it is clear that, using (4.1) for a fixed positive $\delta < 1/2$,

$$\begin{aligned}\lim_{x\to\infty} \frac{1}{x} &(\#\{n \le x : P(n) \le x^\delta\} + \#\{n \le x : P(n) > x^{1-\delta}\}) \\ &= \lim_{x\to\infty} \frac{1}{x} (\psi(x,x^\delta) + x - \psi(x, x^{1-\delta})) = \rho(1/\delta) + 1 - \rho(1/(1-\delta)) \ll \delta.\end{aligned} \tag{4.2}$$

So, let $0 < \delta < 1/2$ be fixed. For some prime q, $x^\delta < q < x^{1-\delta}$, define

$$S_q(x) := \sum_{\substack{n \le x \\ P(n) < q}} f(n) \quad \text{and} \quad D_q = \prod_{q \le p \le x} p.$$

Observe that for any $n \le x$, one has $P(n) < q$ if and only if $\gcd(n, D_q) = 1$. Using the fact that f is completely multiplicative, we deduce that

$$S_q(x) = \sum_{d|D_q} \mu(d) f(d) S(x/d). \tag{4.3}$$

Now consider the sum

$$\Sigma_1 = \Sigma_1(x) := \sum_{x^\delta < q < x^{1-\delta}} f(q) e(\alpha q) S_q(x/q).$$

It follows from (4.2) that

$$\Big| \sum_{n \le x} f(n) e(\alpha P(n)) - \Sigma_1 \Big| \le c_4 \delta x.$$

This last estimate implies that Theorem 1 will be proved (in this case) if we can show that $\Sigma_1 = \Sigma_1(x)$ tends to 0 as $x \to \infty$.

Now since $S(x) = o(x)$, there exists a function $\varepsilon_1(x)$ which tends to 0 as $x \to \infty$ and is such that $|S(x)| \le \varepsilon_1(x) x$.

From (4.3) and the definition of Σ_1, we have

$$\begin{aligned}|\Sigma_1| &\le x \sum_{x^\delta < q < x^{1-\delta}} \frac{1}{q} \sum_{\substack{d|D_q \\ dq < x^{1-\delta^2}}} \frac{\varepsilon_1(x^{\delta^2})}{d} + x \sum_{\substack{d|D_q \\ x^{1-\delta^2} \le qd < x}} \frac{1}{qd} \\ &= x\Sigma_A + \Sigma_B,\end{aligned} \tag{4.4}$$

say. Clearly,

$$\Sigma_A \le \varepsilon_1(x^{\delta^2}) \sum_{x^\delta<q<x^{1-\delta}} \frac{1}{q} \prod_{q\le p<x} \left(1+\frac{1}{p}\right) \tag{4.5}$$
$$\le c_5\varepsilon_1(x^{\delta^2}) \sum_{x^\delta<q<x^{1-\delta}} \frac{\log x}{q\log q}$$
$$\le c_6\varepsilon_1(x^{\delta^2})\frac{1}{\delta}.$$

In order to estimate Σ_B, we proceed as follows. For a fixed prime q, each divisor d in the sum lies in $[z, x^{\delta^2}z]$, where $z = x^{1-\delta^2}/q$. Splitting this interval into dyadic subintervals of the form $[2^j z, 2^{j+1}z]$, we observe that

$$\sum_{\substack{d|D_q \\ d\in]2^j z, 2^{j+1}z[}} \frac{1}{d} \le c_7 \prod_{p<q}\left(1-\frac{1}{p}\right) \le \frac{c_8}{\log q}.$$

Since the maximum value of j in the above expression is $c_9\delta^2\log x$, it follows that

$$\Sigma_B \le c_{10}\delta^2 \sum_{x^\delta<q<x^{1-\delta}} \frac{\log x}{q\log q} \le c_{11}\delta^2\frac{\log x}{\delta\log x} = c_{11}\delta. \tag{4.6}$$

Using (4.5) and (4.6) in (4.4), we obtain

$$\left|\frac{\Sigma_1}{x}\right| \le c_{11}\delta + c_6\frac{\varepsilon_1(x^{\delta^2})}{\delta},$$

which implies that

$$\limsup_{x\to\infty} \frac{|\Sigma_1(x)|}{x} \ll \delta.$$

Since δ can be chosen arbitrarily small, it follows that $|\Sigma_1(x)|/x \to 0$ as $x\to\infty$, which completes the proof of Theorem 1 in case (i) when f is assumed to be completely multiplicative, a fact that we only used to deduce (4.3).

To drop this last condition, we proceed as follows. We define $f_1 = f_{1,x}\in\mathcal{M}$ by $f_1(p^\alpha) = f(p^\alpha)$ if $p\notin[x^\delta, x^{1-\delta}]$ and $f_1(p^\alpha) = f(p)^\alpha$ otherwise. Set

$$S^{(1)}(x) := \sum_{n\le x} f_1(n),$$

and, for $x^\delta < q < x^{1-\delta}$, let

$$S_q^{(1)}(x) := \sum_{d|D_q} \mu(d)f(d)S^{(1)}(x/d).$$

In light of these definitions, it is easy to see that

$$|S(x) - S^{(1)}(x)| \le x \sum_{x^\delta < q < x^{1-\delta}} \frac{1}{q^2} \ll x^{1-\delta}$$

and

$$\Big| \sum_{n \le x} (f(n) - f_1(n)) e(\alpha P(n)) \Big| \ll \delta x + x^{1-\delta},$$

so that the theorem is proved in case (i) without the restriction that f is completely multiplicative.

It remains to consider case (ii). In this case, it follows from Lemma 2 that there exists a real number τ for which $\mathsf{S}(\tau)$ converges. From Lemma 3 we have, as $x \to \infty$,

$$\frac{1}{x} \sum_{n \le x} f(n) e(\alpha P(n)) \to 0 \quad \text{and} \quad \frac{1}{x} \sum_{n \le x} f(n) n^{-i\tau} e(\alpha P(n)) \to 0.$$

In light of these observations, it is sufficient to consider the case $\tau = 0$, that is

$$\mathsf{S}(0) = \sum_p \frac{1 - \Re(f(p))}{p} \text{ is convergent.} \tag{4.7}$$

Let $f(p^r) = e(F(p^r))$ with $-1/2 \le F(p^r) \le 1/2$. It is clear that (4.7) holds if and only if

$$\sum_p \frac{F^2(p)}{p} < \infty. \tag{4.8}$$

Let Y be a fixed large number and set

$$A_{X,Y} := \sum_{Y < p < X} \frac{F(p)}{p}.$$

Further define the multiplicative functions $f_Y(n)$ and $g_Y(n)$ by

$$f_Y(p^r) := \begin{cases} f(p^r) & \text{if } p \le Y, \\ 1 & \text{if } p > Y, \end{cases} \qquad g_Y(p^r) := \begin{cases} f(p^r) & \text{if } p > Y, \\ 1 & \text{if } p \le Y. \end{cases}$$

It is clear that $f(n) = f_Y(n) \cdot g_Y(n)$.

Further let

$$G_Y(n) := \sum_{\substack{p^r \| n \\ p > Y}} F(p^r).$$

It follows from the Turán–Kubilius inequality that

$$\sum_{n \le x} |G_Y(n) - A_{X,Y}|^2 \le c_{12} x \sum_{\substack{p \ge Y \\ r \ge 1}} \frac{F^2(p^r)}{p^r} = c_{12} x B_Y^2, \tag{4.9}$$

say. From (4.8), it follows that $B_Y \to 0$ as $Y \to \infty$. On the other hand, since $g_Y(n) = e(G_Y(n))$, it is clear, in light of (4.9), that

$$\sum_{n\le x} |g_Y(n) - e(A_{X,Y})|^2 \le c_{13} x B_Y^2.$$

Therefore,

$$\Big|\sum_{n\le x} f(n)e(\alpha P(n)) - e(-A_{X,Y}) \sum_{n\le x} f_Y(n)e(\alpha P(n))\Big| \le c_{14} x B_Y. \tag{4.10}$$

We shall now establish that

$$\frac{1}{x}\sum_{n\le x} f_Y(n)e(\alpha P(n)) \to 0 \quad (x\to\infty). \tag{4.11}$$

We further define the multiplicative function $\widetilde{f_Y}(n)$ by

$$\widetilde{f_Y}(p^r) := \begin{cases} 1 & \text{if } p > Y^{1/r}, \\ f_Y(p^r) & \text{otherwise.} \end{cases}$$

First observe that

$$\Big|\sum_{n\le x} f_Y(n)e(\alpha P(n)) - \sum_{n\le x} \widetilde{f_Y}(n)e(\alpha P(n))\Big| \le \sum_{\substack{p^r\ge Y\\ p\le Y}} \frac{x}{p^r} \le \varepsilon_1(Y)x, \tag{4.12}$$

where $\varepsilon_1(Y)\to 0$ as $Y\to\infty$.

Let $h_Y(n)$ be the function defined implicitly by

$$\widetilde{f_Y}(n) = \sum_{d|n} h_Y(d).$$

It is easy to see that

$$h_Y(p) = \begin{cases} \widetilde{f_Y}(p) - 1 & \text{if } p \le Y, \\ 0 & \text{if } p > Y, \end{cases}$$

and that similarly $h_Y(p^r) = 0$ if $p > Y$.

On the other hand, since $h_Y(p^r) = \widetilde{f_Y}(p^r) - \widetilde{f_Y}(p^{r-1})$, it follows that $h_Y(p^r) = 0$ if $p^{r-1} > Y$.

From the definition of h_Y, it is clear that

$$\sum_{n\le x} \widetilde{f_Y}(n)e(\alpha P(n)) = \sum_{d\le x} h_Y(d) \sum_{dm\le x} e(\alpha P(dm)). \tag{4.13}$$

If $h_Y(d) \ne 0$, then $p^r \,\|\, d$ implies that $p < Y$ and $p^{r-1} \le Y$, so that $p^r \le Y^2$. Consequently, $d \le Y^{2\pi(Y)} \le Y^{2Y}$. Furthermore, $h_Y(d) \le 2^{\pi(Y)}$.

For a fixed positive integer d, we have

$$\sum_{m\le x/d} e(\alpha P(dm)) = \sum_{m\le x/d} e(dP(m)) + O\Big(\sum_{\substack{m\le x/d\\ P(m)\le P(d)}} 1\Big). \tag{4.14}$$

Using the main result of Banks, Harman and Shparlinski [1], namely that for any fixed irrational number α,

$$\lim_{x\to\infty}\frac{1}{x}\sum_{n\le x} e(\alpha P(n)) = 0,$$

we deduce, using (4.14) in (4.13), that

$$\lim_{x\to\infty}\frac{1}{x}\sum_{n\le x}\widetilde{f_Y}(n)e(\alpha P(n)) = 0. \tag{4.15}$$

Hence, it follows from estimate (4.15), taking into account (4.12), that (4.11) is proved. Finally, gathering (4.10) and (4.11), Theorem 1 is proved.

Theorem 2 can be established along the lines of the proof of Theorem 1 and its proof will therefore be omitted.

5. The proof of Theorem 3. Let $0<\eta_1<\eta_2<1/2$. It is clear that

$$\begin{aligned} E(x,Q,a) &\ge \sum_{\substack{x^{\eta_1}<Q<x^{\eta_2}\\ Q\equiv a\,(\mathrm{mod}\, q)}} \pi(x;Q,1) - \sum_{\substack{Q<Q'\\ x^{\eta_1}<Q<x^{\eta_2}\\ Q\equiv a\,(\mathrm{mod}\, q)}} \pi(x;QQ',1) \\ &= \Sigma_1 - \Sigma_2, \end{aligned} \tag{5.1}$$

say, where as usual $\pi(x;b,a) := \#\{p\le x : p\equiv a \pmod b\}$. It follows from the Bombieri–Vinogradov theorem that

$$\Sigma_1 = \mathrm{li}(x)\sum_{\substack{x^{\eta_1}<Q<x^{\eta_2}\\ Q\equiv a\,(\mathrm{mod}\, q)}}\frac{1}{Q-1} + O\biggl(\frac{x}{(\log x)^A}\biggr), \tag{5.2}$$

assuming that $x^{\eta_2}\le \sqrt{x}/(\log x)^{2A+5}$, a condition which is equivalent to

$$\frac{1}{2}-\eta_2 \ge (2A+5)\frac{\log\log x}{\log x}. \tag{5.3}$$

Summing over Q allows us to write (5.2) as

$$\Sigma_1 = \biggl(\log\frac{\eta_2}{\eta_1}\biggr)\frac{\mathrm{li}(x)}{\phi(q)} + O\biggl(\frac{x}{(q\log x)^D}\biggr) \tag{5.4}$$

uniformly for $q\le(\log x)^c$, where D is any preassigned value.

In order to estimate Σ_2, we use standard sieve techniques. Actually Σ_2 represents the number of solutions of $p-1=bQQ'\le x$, where b,Q,Q' vary as follows:

$$Q\equiv a \pmod q,\qquad Q\in[x^{\eta_1},x^{\eta_2}],\qquad Q<Q',\qquad b=1,2,\ldots.$$

We first fix b and Q, and we assume that there is at least one pair of numbers p,Q' which is a solution of $p-1=bQQ'\le x$, in which case we have $b < x^{1-2\eta_1}$ and $bQ < x^{1-\eta_1}$. Let η_1 be close to $1/2$. Then

$$E_{b,Q} := \#\{p, Q' : p-1 = bQQ' \le x,\ Q \equiv a \pmod q\} \le c_{15}\frac{x}{\log^2 x\ \phi(bQ)}. \tag{5.5}$$

Using the well known estimate $\sum_{b\le y} 1/\phi(b) \le c_{16}\log y$, it follows from (5.5) that

$$\Sigma_2 = \sum_{b,Q} E_{b,Q} \le c_{15}\frac{x}{\log^2 x}c_{16}\sum_{x^{\eta_1}<Q<x^{\eta_2}}\frac{\log(x/Q^2)}{Q-1} \le c_{17}\frac{x}{\log x}\,\frac{1}{\phi(q)}(1-2\eta_1)\log\frac{\eta_2}{\eta_1}. \tag{5.6}$$

Choosing η_2 so that it satisfies (5.3) and η_1 so that $c_{17}(1-2\eta_1) < 1/2$, and then gathering (5.4) and (5.6) in (5.1), we obtain

$$E(x,q,a) \ge \frac{1}{2}\left(\log\frac{\eta_2}{\eta_1}\right)\frac{\operatorname{li}(x)}{\phi(q)},$$

thus completing the proof of Theorem 3.

6. The proof of Theorem 4. Again using the analogue of Lemma 3, namely in the form

$$\lim_{x\to\infty}\frac{1}{x}\sum_{n\le x} f(n)n^{i\tau}e(\alpha P(n-1)) = 0 \ \Leftrightarrow\ \lim_{x\to\infty}\frac{1}{x}\sum_{n\le x} f(n)e(\alpha P(n-1)) = 0,$$

we may assume that $\tau = 0$, that is,

$$\mathsf{S}(0) = \sum_p \frac{1-\Re(f(p))}{p} < \infty.$$

Arguing as in the proof of case (ii) of Theorem 1, we reduce the problem to proving that the expression

$$\sum_{n\le x}\widetilde{f_Y}(n)e(\alpha P(n-1)) = \sum_{d\le x} h_Y(d)\sum_{m\le x/d} e(\alpha P(dm-1)) \tag{6.1}$$

is $o(x)$ as $x\to\infty$.

First let us define

$$\psi(x,y;a,q) := \#\{n \le x : P(n) \le y,\ n \equiv a \pmod q\}.$$

Since, in the first sum on the right hand side of (6.1), d runs over a finite set of integers which does not change as $x\to\infty$, it is enough to prove that

$$\lim_{X\to\infty}\frac{1}{X}\sum_{m\le X} e(\alpha P(dm-1)) = 0. \tag{6.2}$$

We have $P(dm-1)=q$ if $dm-1=q\nu$, $P(\nu)\le q$, that is, $q\nu+1\equiv 0 \pmod d$, $\nu\equiv \ell_q \pmod d$, $P(\nu)\le q$, $\nu\le x/q$. This quantity is precisely $\psi(xd/q,q;\ell_q,d)$.

It follows that

$$\sum_{m\le X} e(\alpha P(dm-1)) = \sum_{q<xd} e(\alpha q)\psi\Big(\frac{xd}{q},q;\ell_q,d\Big).$$

Let $\varepsilon>0$ be an arbitrary real number. It follows from (4.2) that

$$\sum_{m\le x} e(\alpha P(dm-1)) = \sum_{x^{\varepsilon}<q<x^{1-\varepsilon}} e(\alpha q)\psi\Big(\frac{xd}{q},q;\ell_q,d\Big)+R_x, \tag{6.3}$$

where $|R_x|\le \varepsilon x$. It has been established by Granville [3] that, if $\gcd(a,d)=1$ and $d^{1+\varepsilon}\le y\le x$, then

$$\psi(x,y;a,d)\sim \frac{1}{d}\psi(x,y) \qquad (x\to\infty). \tag{6.4}$$

Observing that

$$\psi\Big(\frac{xd}{q},q\Big)=(1+o(1))\rho\Big(\frac{\log xd}{\log q}-1\Big)\frac{xd}{q} \qquad (x\to\infty),$$

we deduce, by (6.4), that the right hand side of (6.3) is, as $x\to\infty$, equal to

$$\begin{aligned} xd\sum_{x^{\varepsilon}<q<x^{1-\varepsilon}} \rho\Big(\frac{\log xd}{\log q}-1\Big)\frac{e(\alpha q)}{q} + o(1)xd\sum_{x^{\varepsilon}<q<x^{1-\varepsilon}} \rho\Big(\frac{\log xd}{\log q}-1\Big)\frac{1}{q}+R_x \\ = S_1(x)+S_2(x)+R_x.\end{aligned}$$

In order to prove (6.2), it remains to show that

$$S_1(x)=o(x) \quad \text{and} \quad S_2(x)=o(x). \tag{6.5}$$

First we set

$$J_x:=\Big[\frac{1}{1-\varepsilon}-1+\frac{\log d}{\log x},\frac{1}{\varepsilon}-1+\frac{\log d}{\log x}\Big].$$

If $q\in[x^{\varepsilon},x^{1-\varepsilon}]$, then $(\log xd)/(\log q)-1\in J_x$. On the other hand, note that $J_x\subseteq[1/(1-\varepsilon),1/\varepsilon]$, and that in this interval, ρ is bounded, and therefore,

$$S_2(x)\ll o(1)xd\sum_{x^{\varepsilon}<q<x^{1-\varepsilon}}\frac{1}{q}\ll o(1)x\log\frac{1}{\varepsilon}=o(x) \qquad (x\to\infty), \tag{6.6}$$

which proves the second estimate in (6.5).

To estimate $S_1(x)$, we proceed as follows. First set

$$B(y):=\sum_{x^{\varepsilon}\le q<y}\frac{e(\alpha q)}{q}.$$

By using the theorem of I. M. Vinogradov according to which

$$\max_{2x^\varepsilon \le y \le x} \frac{1}{\pi(y)} \Big| \sum_{x^\varepsilon \le q < y} e(\alpha q) \Big| = \delta(x) \to 0 \qquad \text{as } x \to \infty,$$

we find immediately that

$$\max_{2x^\varepsilon \le y \le x} |B(y)| = \delta_1(x) \to 0 \qquad \text{as } x \to \infty.$$

On the other hand, since

$$\sum_{x^\varepsilon \le q < y} \frac{1}{q} \le \log\left(\frac{\log y}{\varepsilon \log x}\right) \qquad \text{for } x^\varepsilon < y \le 2x^\varepsilon,$$

it follows that

$$\max_{x^\varepsilon \le y \le x} |B(y)| = \delta_2(x) \to 0 \qquad \text{as } x \to \infty.$$

From the definitions of $S_1(x)$ and $B(y)$, we have

$$\begin{aligned} S_1(x) &= xd \int_{x^\varepsilon}^{x^{1-\varepsilon}} \rho\left(\frac{\log xd}{\log u} - 1\right) dB(u) \\ &= xd\, \rho\left(\frac{\log xd}{\log u} - 1\right) B(u) \Big|_{x^\varepsilon}^{x^{1-\varepsilon}} \\ &\quad + xd \int_{x^\varepsilon}^{x^{1-\varepsilon}} B(u) \rho'\left(\frac{\log xd}{\log u} - 1\right) \frac{\log xd}{u(\log u)^2}\, du. \end{aligned} \tag{6.7}$$

Since both $\rho(u)$ and $\rho'(u)$ are bounded in J_x, it follows from (6.7) and the above bounds on $B(u)$ that

$$\left| \frac{1}{x} S_1(x) \right| \le do(1) + do(1) \int_{x^\varepsilon}^{x^{1-\varepsilon}} \frac{\log xd}{u(\log u)^2}\, du \qquad (x \to \infty). \tag{6.8}$$

On the other hand,
(6.9)

$$\int_{x^\varepsilon}^{x^{1-\varepsilon}} \frac{1}{u(\log u)^2}\, du = \int_{\varepsilon \log x}^{(1-\varepsilon)\log x} \frac{dv}{v^2} = -\frac{1}{v}\Big|_{\varepsilon \log x}^{(1-\varepsilon)\log x} = \left(\frac{1}{\varepsilon} - \frac{1}{1-\varepsilon}\right) \frac{1}{\log x}.$$

Gathering (6.6), (6.8) and (6.9) completes the proof of (6.5), as required. Since $\varepsilon > 0$ is arbitrary, it follows from (6.3) that

$$\frac{1}{x} \sum_{m \le x} e(\alpha P(dm-1)) \to 0 \qquad (x \to \infty)$$

for every d, thus proving (6.2) and thereby (6.1), which completes the proof of Theorem 4.

Acknowledgements. The authors would like to thank the referee for some very relevant comments which helped improve the quality of this paper.

References

[1] W. D. Banks, G. Harman and I. E. Shparlinski, *Distributional properties of the largest prime factor*, Michigan Math. J. 53 (2005), 665–681.

[2] H. Daboussi et H. Delange, *Quelques propriétés des fonctions multiplicatives de module au plus égal à* 1, C. R. Acad. Sci. Paris Ser. A 278 (1974), 657–660.

[3] A. Granville, *Smooth numbers: computational number theory and beyond*, in: Algorithmic Number Theory, Math. Sci. Res. Inst. Publ. 44, Cambridge Univ. Press, Cambridge, 2008, 267–323.

[4] G. Halász, *Über die Mittelwerte multiplikativer zahlentheoretischer Funktionen*, Acta Math. Acad. Sci. Hungar. 19 (1968), 365–403.

[5] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, New York, 1974.

[6] G. Tenenbaum, *Introduction à la théorie analytique et probabiliste des nombres*, 3ème éd., Collection Échelles, Belin, Paris, 2008.

[7] I. M. Vinogradov, *The Method of Trigonometrical Sums in the Theory of Numbers*, translated, revised and annotated by K. F. Roth and A. Davenport, Interscience, New York, 1954.

Jean-Marie De Koninck
Département de mathématiques et de statistique
Université Laval
Québec, QC, 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@compalg.inf.elte.hu

Received on 23.10.2009
and in revised form on 11.7.2010 (6187)