Ann. Sci. Math. Québec **29** (2005), no. 2, 131–145.

ON CERTAIN SUMS RELATED TO MULTIPLE DIVISIBILITY BY THE LARGEST PRIME FACTOR

WILLIAM D. BANKS, FLORIAN LUCA AND IGOR E. SHPARLINSKI

RÉSUMÉ. Pour un nombre entier positif n, $P(n)$ désigne le plus grand facteur premier de n. Nous montrons que pour deux nombres réels fixés ϑ, μ et un entier $k \geq 2$, avec $\vartheta \in (-1, 0]$ et $\mu < k-1$, l'estimation

$$\sum_{\substack{P(n)^k \mid n \\ n \leq x}} P(n)^\mu n^\vartheta = x^{1+\vartheta} \exp\left(-(1+o(1))\sqrt{2(k-1-\mu)\log x \log_2 x}\right)$$

a lieu lorsque $x \to \infty$, ou $\log_2 x = \log\log x$. Avec $\vartheta = \mu = 0$ et $k = 2$, nous retrouvons le résultat connu à l'effet que, lorsque $x \to \infty$, le nombre d'entiers positifs $n \leq x$ pour lesquels $P(n)^2 \mid n$ est $x \exp\left(-(1+o(1))\sqrt{2\log x \log_2 x}\right)$. Nous obtenons aussi des bornes supérieure et inférieure pour le nombre d'entiers positifs $n \leq x$ pour lesquels $P(\varphi(n))^k \mid \varphi(n)$, où $\varphi(n)$ est la fonction d'Euler.

ABSTRACT. For a positive integer n, let $P(n)$ denote the largest prime factor of n. We show that for two fixed real numbers ϑ, μ and an integer $k \geq 2$, with $\vartheta \in (-1, 0]$ and $\mu < k-1$, the estimate

$$\sum_{\substack{P(n)^k \mid n \\ n \leq x}} P(n)^\mu n^\vartheta = x^{1+\vartheta} \exp\left(-(1+o(1))\sqrt{2(k-1-\mu)\log x \log_2 x}\right)$$

holds as $x \to \infty$, where $\log_2 x = \log\log x$. With $\vartheta = \mu = 0$ and $k = 2$, we recover the known result that, as $x \to \infty$, the number of positive integers $n \leq x$ for which $P(n)^2 \mid n$ is $x \exp\left(-(1+o(1))\sqrt{2\log x \log_2 x}\right)$. We also obtain upper and lower bounds for the number of positive integers $n \leq x$ for which $P(\varphi(n))^k \mid \varphi(n)$, where $\varphi(n)$ is the Euler function.

1. Introduction. For an integer $n \geq 2$, let $P(n)$ denote the largest prime factor of n, and put $P(1) = 1$. In many situations, especially when counting integers subject to various arithmetical constraints, it can be important to understand the set of positive integers n for which $P(n)$ is *small* relative to n, or the set of positive integers n for which $P(n)^2 \mid n$. Such sets often contribute to the error term in a given estimate, hence it is useful to have tight bounds for the number of elements in these sets. For example,

Reçu le 16 septembre 2004 et, sous forme définitive, le 13 février 2005.

such sets of integers have played an important role in estimating exponential sums with the *Euler function* $\varphi(n)$ (see [4, 5]), in studying the "average prime divisor" of an integer (see [3]), or in evaluating various means and moments of the *Smarandache function* of n (see [6, 9]).

Positive integers with the first property (that is, $P(n)$ is small) are called *smooth*, and the distribution of smooth numbers has been extensively studied.

Although there is no special name for integers with the second property (that is, $P(n)^2 \mid n$), they have also been studied in the literature. For example, in [10] it has been shown that for a fixed real number $r \geq 0$, the following estimate holds as $x \to \infty$:

$$\sum_{\substack{P(n)^2 \mid n \\ n \leq x}} \frac{1}{P(n)^r} = x \exp\left(-(1 + o(1))\sqrt{(2r+2)\log x \log\log x}\right). \tag{1}$$

The error term $o(1)$ in the above estimate can be taken to be of the form

$$O_r(\log\log\log x / \log\log x).$$

More general sums taken over integers $n \leq x$ with $P(n)$ in a fixed arithmetic progression have been estimated in [8]. In Theorem 1 below, we complement the result (1) by showing that if ϑ, μ are fixed real numbers and $k \geq 2$ is an integer such that $\vartheta \in (-1, 0]$ and $\mu < k - 1$, then the following estimate holds as $x \to \infty$:

$$\begin{aligned} S(k, \vartheta, \mu, x) &= \sum_{\substack{P(n)^k \mid n \\ n \leq x}} P(n)^{\mu} n^{\vartheta} \\ &= x^{1+\vartheta} \exp\left(-(1 + o(1))\sqrt{2(k - 1 - \mu)\log x \log\log x}\right). \end{aligned} \tag{2}$$

Our method closely resembles that of [8]. In particular, taking $\vartheta = \mu = 0$ in (2), we obtain the estimate

$$\#\mathcal{N}_k(x) = x \exp\left(-(1 + o(1))\sqrt{2(k-1)\log x \log\log x}\right) \tag{3}$$

for the cardinality of the set

$$\mathcal{N}_k(x) = \{n \leq x : P(n)^k \mid n\}.$$

When $k = 2$, the bound (3) can be recovered from (1) by taking $r = 0$, and it appears also in [6].

We also consider a similar question for values of the Euler function $\varphi(n)$. In Theorem 2, we give upper and lower bounds for the cardinality of the sets

$$\mathcal{F}_k(x) = \{n \leq x : P(\varphi(n))^k \mid \varphi(n)\}, \qquad k = 2, 3, \ldots$$

Finally, for an odd integer n, we define $t(n)$ to be the multiplicative order of 2 modulo n, and we give upper bounds on the cardinality of the sets

$$\mathcal{T}_k(x) = \{n \leq x : n \text{ odd},\ P\left(t(n)\right)^k \mid t(n)\}, \qquad k = 2, 3, \ldots$$

Throughout the paper, x denotes a large positive real number. We use the Landau symbols 'O' and 'o' as well as the Vinogradov symbols '$\ll$', '$\gg$' and '$\asymp$' with their usual meanings. The constants or convergence implied by these symbols depend only on our parameters ϑ, μ and k. For a positive integer ℓ, we write $\log_\ell x$ for the function defined inductively by $\log_1 x = \max\{\log x, 1\}$ and $\log_\ell x = \log_1(\log_{\ell-1} x)$ for $\ell \ge 2$, where log denotes the natural logarithm function. In the case $\ell = 1$, we omit the subscript to simplify the notation; however, it should be understood that all the values of all logarithms that appear are at least 1.

2. Preliminary results. For the results in this section, we refer the reader to Section III.5.4 in the book of Tenenbaum [11].

Let $\rho : \mathbb{R}_{\ge 0} \to \mathbb{R}$ denote the *Dickman function*. We recall that $\rho(u) = 1$ for $0 \le u \le 1$, ρ is continuous at $u = 1$ and differentiable for all $u > 1$, and ρ satisfies the difference-differential equation

$$u\rho'(u) + \rho(u-1) = 0 \qquad (u > 1).$$

Thus,

$$\rho(u) = \frac{1}{u}\int_{u-1}^{u} \rho(v)\,dv \qquad (u \ge 1).$$

For $u > 0$, $u \ne 1$, let $\xi(u)$ denote the (unique) real nonzero root of the equation $e^\xi = 1 + u\xi$. By convention, we also put $\xi(1) = 0$.

Lemma 1. *The following estimate holds:*

$$\xi(u) = \log(u \log u) + O\left(\frac{\log_2 u}{\log u}\right) \qquad (u > 1).$$

Lemma 2. *For any fixed real number $u_0 > 1$, the following estimate holds:*

$$\rho'(u) = -\xi(u)\rho(u)\left(1 + O(1/u)\right) \qquad (u \ge u_0).$$

Lemma 3. *The following estimate holds:*

$$\rho(u) = u^{-(1+o(1))u} \qquad (u \to \infty).$$

As usual, we say that a positive integer n is *y-smooth* if $P(n) \le y$. For all real numbers $x \ge y \ge 2$, we denote by $\Psi(x,y)$ the number of y-smooth integers $n \le x$:

$$\Psi(x,y) = \#\{n \le x : P(n) \le y\}.$$

Lemma 4. *For any fixed $\varepsilon > 0$ and*

$$\exp\left((\log_2 x)^{5/3+\varepsilon}\right) \le y \le x,$$

we have

$$\Psi(x,y) = x\rho(u)\left(1 + O\left(\frac{\log u}{\log y}\right)\right),$$

when $u = (\log x)/(\log y) \to \infty$.

Lemma 5. *For all $x \geq y \geq 2$, we have*

$$\Psi(x, y) \ll xe^{-u/2}.$$

We also denote by $\Phi(x, y)$ the number of integers $n \leq x$ for which $\varphi(n)$ is y-smooth:

$$\Phi(x, y) = \#\{n \leq x : P(\varphi(n)) \leq y\}.$$

By Theorem 3.1 of [2], the following bound holds.

Lemma 6. *For any fixed $\varepsilon > 0$ and*

$$(\log_2 x)^{1+\varepsilon} \leq y \leq x,$$

we have

$$\Phi(x, y) \leq x \exp(-(1 + o(1))\, u \log_2 u)$$

when $u = (\log x)/(\log y) \to \infty$.

Finally, let $\Theta(x, y)$ be the number of integers $n \leq x$ for which $t(n)$ is y-smooth:

$$\Theta(x, y) = \#\{n \leq x : P(t(n)) \leq y\}.$$

By Theorem 5.1 of [2], the following bound holds.

Lemma 7. *For $\exp\left(\sqrt{\log x \log_2 x}\right) \leq y \leq x$ we have*

$$\Theta(x, y) \leq x \exp(-(\tfrac{1}{2} + o(1))\, u \log_2 u)$$

when $u = (\log x)/(\log y) \to \infty$.

3. Main Results. It is easy to see that a positive integer n satisfies $P(n)^k \mid n$ if and only if $n = mP(m)^{k-1}$, where $m = n/P(n)^{k-1}$; therefore, the sum in (2) can be rewritten as

$$S(k, \vartheta, \mu, x) = \sum_{\substack{P(n)^k \mid n \\ n \leq x}} P(n)^{\mu} n^{\vartheta} = \sum_{mP(m)^{k-1} \leq x} P(m)^{\mu + k\vartheta} m^{\vartheta}.$$

Theorem 1. *Let $\vartheta \in (-1, 0]$ and $\mu < k - 1$ be fixed. Then the following estimate holds as $x \to \infty$:*

$$S(k, \vartheta, \mu, x) = x^{1+\vartheta} \exp\left(-(1 + o(1))\sqrt{2(k - 1 - \mu) \log x \log_2 x}\right).$$

Proof. Given real numbers $x \geq y \geq 2$, we put

$$\mathcal{S}(x, y) = \{n \leq x : P(n) \leq y\};$$

thus, $\Psi(x, y) = \#\mathcal{S}(x, y)$. For any integer m such that $mP(m)^{k-1} \leq x$, setting $p = P(m)$ we see that $m = pn$, where $n \leq x/p^k$ and $P(n) \leq p$; in other words,

$n \in \mathcal{S}(x/p^k, p)$. Conversely, for a prime p and an integer $n \in \mathcal{S}(x/p^k, p)$, the number $m = pn$ clearly satisfies $mP(m)^{k-1} \le x$. We therefore have the basic identity:

$$S(k, \vartheta, \mu, x) = \sum_{mP(m)^{k-1} \le x} P(m)^{\mu + k\vartheta} m^{\vartheta} = \sum_{p \le x} p^{\mu + k\vartheta} \sum_{n \in \mathcal{S}(x/p^k, p)} n^{\vartheta}. \tag{4}$$

We begin by considering the contribution to (4) coming from "small" primes p. Let

$$y_1 = \exp\left(\sqrt{\frac{(1+\vartheta)^2}{16(k-1-\mu)} \log x \log_2 x} \right).$$

Then

$$\begin{aligned} \sum_{p \le y_1} p^{\mu + k\vartheta} \sum_{n \in \mathcal{S}(x/p^k, p)} n^{\vartheta} &\le \sum_{p \le y_1} p^{\mu + k\vartheta} \sum_{n \le \Psi(x/p^k, p)} n^{\vartheta} \\ &\ll \sum_{p \le y_1} p^{\mu + k\vartheta} \int_1^{\Psi(x/p^k, p)} t^{\vartheta} dt \ll \sum_{p \le y_1} p^{\mu + k\vartheta} \Psi(x/p^k, p)^{1+\vartheta} \end{aligned} \tag{5}$$

since $\vartheta \in (-1, 0]$.

To estimate the last sum in (5), we define

$$y_1^* = \exp\left(\tfrac{1}{3} \sqrt{\log x} / \log_2 x \right)$$

and consider separately the contributions coming from primes $p \le y_1^*$ and from primes $p > y_1^*$. For a real number $z > 0$, let

$$u_z = \frac{\log(x/z^k)}{\log z} = \frac{\log x}{\log z} - k.$$

If $p \le y_1^*$, the inequality

$$u_p \ge \frac{\log x}{\log y_1^*} - k \ge 3\sqrt{\log x}\, \log_2 x - k \ge 2\sqrt{\log x}\, \log_2 x$$

holds provided that x is large enough; hence, by Lemma 5, we have

$$\begin{aligned} \sum_{p \le y_1^*} p^{\mu + k\vartheta} \Psi(x/p^k, p)^{1+\vartheta} &\ll \sum_{p \le y_1^*} p^{\mu + k\vartheta} \left(\frac{x}{p^k} \right)^{1+\vartheta} \exp\left(-(1+\vartheta)\sqrt{\log x}\, \log_2 x \right) \\ &= x^{1+\vartheta} \exp\left(-(1+\vartheta)\sqrt{\log x}\, \log_2 x \right) \sum_{p \le y_1^*} p^{\mu - k} \\ &\ll x^{1+\vartheta} \exp\left(-(1+\vartheta)\sqrt{\log x}\, \log_2 x \right) \end{aligned} \tag{6}$$

since $\mu - k < -1$.

For primes p in the range $y_1^* < p \le y_1$, we remark that

$$\frac{\log u_p}{\log p} \le \frac{\log_2 x}{\log p} < \frac{\log_2 x}{\log y_1^*} \ll \frac{(\log_2 x)^2}{\sqrt{\log x}} = o(1);$$

hence, by Lemma 4:

$$\sum_{y_1^*<p\le y_1} p^{\mu+k\vartheta}\Psi(x/p^k,p)^{1+\vartheta} = \sum_{y_1^*<p\le y_1} p^{\mu+k\vartheta}\left(\frac{x}{p^k}\,\rho(u_p)\left(1+O\left(\frac{\log u_p}{\log p}\right)\right)\right)^{1+\vartheta}$$
$$\ll x^{1+\vartheta}\sum_{y_1^*<p\le y_1} p^{\mu-k}\rho(u_p)^{1+\vartheta}.$$

For each $p \le y_1$, we also have the lower bound

$$u_p = \frac{\log x}{\log p} - k \ge \frac{\log x}{\log y_1} - k = (1+o(1))u_*,$$

where

$$u_* = \left(\frac{4(k-1-\mu)\log x}{(1+\vartheta)^2\log_2 x}\right)^{1/2}.$$

Using Lemma 3 and the fact that ρ is decreasing for large u, we deduce that

$$\rho(u_p)^{1+\vartheta} \le \exp\left(-(1+o(1))\sqrt{4(k-1-\mu)\log x\log_2 x}\right),$$

which implies the bound

$$\sum_{y_1^*<p\le y_1} p^{\mu+k\vartheta}\Psi(x/p^k,p)^{1+\vartheta} \le x^{1+\vartheta}\exp\left(-(1+o(1))\sqrt{4(k-1-\mu)\log x\log_2 x}\right). \quad (7)$$

Combining the estimates (6) and (7) and substituting into (5), it follows that

$$\sum_{p\le y_1} p^{\mu+k\vartheta}\sum_{n\in\mathcal{S}(x/p^k,\,p)} n^\vartheta \le x^{1+\vartheta}\exp\left(-(1+o(1))\sqrt{4(k-1-\mu)\log x\log_2 x}\right).$$

Therefore, we see that the contribution to (4) from “small” primes $p \le y_1$ is negligible compared to the claimed bound on $S(k,\vartheta,\mu,x)$.

Next, we estimate the contribution to (4) coming from “large” primes $p > y_2$, where

$$y_2 = \exp\left(\sqrt{4(k-1-\mu)^{-1}\log x\log_2 x}\right).$$

We have

$$\sum_{p>y_2} p^{\mu+k\vartheta}\sum_{n\in\mathcal{S}(x/p^k,\,p)} n^\vartheta \le \sum_{p>y_2} p^{\mu+k\vartheta}\sum_{n\le x/p^k} n^\vartheta \ll \sum_{p>y_2} p^{\mu+k\vartheta}\int_1^{x/p^k} t^\vartheta dt$$
$$\ll \sum_{p>y_2} p^{\mu+k\vartheta}\left(\frac{x}{p^k}\right)^{1+\vartheta} = x^{1+\vartheta}\sum_{p>y_2} p^{\mu-k} \ll x^{1+\vartheta}y_2{}^{\mu-k+1}$$
$$= x^{1+\vartheta}\exp\left(-(1+o(1))\sqrt{4(k-1-\mu)\log x\log_2 x}\right).$$

Hence, the contribution to (4) from “large” primes $p > y_2$ is also negligible.

Thus, to complete the proof of Theorem 1, it suffices to show that “medium” primes $y_1 < p \le y_2$ make the appropriate contribution to (4); that is, we need to show that

$$\sum_{y_1 < p \le y_2} p^{\mu + k\vartheta} \sum_{n \in \mathcal{S}(x/p^k,\, p)} n^{\vartheta} = x^{1+\vartheta} \exp\left(-(1+o(1))\sqrt{2(k-1-\mu)\log x \log_2 x}\right). \quad (8)$$

To do this, let us first fix a prime $p \in (y_1, y_2]$. Using Lemma 4, we obtain the lower bound

$$\sum_{n \in \mathcal{S}(x/p^k,\, p)} n^{\vartheta} \ \ge\ (x/p^k)^{\vartheta} \Psi(x/p^k, p) \ \gg\ \rho(u_p)\, \frac{x^{1+\vartheta}}{p^{k+k\vartheta}}, \quad (9)$$

where

$$u_p = \frac{\log(x/p^k)}{\log p} = \frac{\log x}{\log p} - k$$

as before. For the upper bound, we have by partial summation:

$$\begin{aligned} \sum_{n \in \mathcal{S}(x/p^k,\, p)} n^{\vartheta} &= \int_{1^-}^{x/p^k} t^{\vartheta} d\left(\Psi(t,p)\right) \\ &= t^{\vartheta}\Psi(t,p)\Big|_{t=1^-}^{t=x/p^k} + |\vartheta| \int_{1^-}^{x/p^k} t^{\vartheta-1}\Psi(t,p)\, dt \\ &\ll \rho(u_p)\, p^{\mu-k} x^{1+\vartheta} + \int_{1^-}^{x/p^k} t^{\vartheta-1}\Psi(t,p)\, dt. \end{aligned}$$

To bound the last term, we split the integral at $w = \rho(u_p)^{1/(1+\vartheta)} x/p^k$. For the lower range, we have

$$\int_{1^-}^{w} t^{\vartheta-1}\Psi(t,p)\, dt \le \int_{1^-}^{w} t^{\vartheta} dt \ll w^{1+\vartheta} = \rho(u_p)\, \frac{x^{1+\vartheta}}{p^{k+k\vartheta}}. \quad (10)$$

For the upper range, writing $u_{t,p} = (\log t)/(\log p)$, we have by Lemma 4:

$$\begin{aligned} \int_{w}^{x/p^k} t^{\vartheta-1}\Psi(t,p)\, dt &= \int_{w}^{x/p^k} \rho(u_{t,p}) t^{\vartheta} \left(1 + O\left(\frac{\log u_{t,p}}{\log p}\right)\right) dt \\ &\le\ \rho(u_{w,p}) \left(1 + O\left(\frac{\log_2 x}{\log p}\right)\right) \int_{w}^{x/p^k} t^{\vartheta} dt \ \ll\ \rho(u_{w,p})\, \frac{x^{\vartheta+1}}{p^{k+k\vartheta}}. \end{aligned} \quad (11)$$

Since

$$\frac{\log x}{\log y_2} \le u_p < \frac{\log x}{\log y_1}$$

holds for $y_1 < p \le y_2$, using Lemma 3 we derive that

$$\log \rho(u_p) \asymp \sqrt{\log x \log_2 x} \asymp \log p.$$

Thus,

$$u_{w,p} = \frac{\log w}{\log p} = \frac{\log(x/p^k)}{\log p} + \frac{\log \rho(u_p)}{(1-\vartheta)\log p} \geq u_p + O(1),$$

Using Lemma 3 again, it follows that

$$\rho(u_{w,p}) = \rho(u_p)\exp\left(O(u_p)\right) = \rho(u_p)\exp\left(o\left(\sqrt{\log x \log_2 x}\right)\right).$$

Substituting this estimate into (11) and combining this with (10), we obtain the upper bound

$$\sum_{n\in\mathcal{S}(x/p^k,\,p)} n^{\vartheta} \leq \rho(u_p)\,\frac{x^{1+\vartheta}}{p^{k+k\vartheta}}\,\exp\left(o\left(\sqrt{\log x \log_2 x}\right)\right).$$

Taking into account the lower bound (9), we obtain that

$$\sum_{n\in\mathcal{S}(x/p^k,\,p)} n^{\vartheta} \asymp \rho(u_p)\,\frac{x^{1+\vartheta}}{p^{k+k\vartheta}}\,\exp\left(o\left(\sqrt{\log x \log_2 x}\right)\right). \tag{12}$$

Now let $\delta = (\log x \log_2 x)^{-1/4}$. Put $z_0 = y_1$, and let $z_\ell = z_0(1+\delta)^\ell$ for each $\ell \geq 1$. Let L denote the smallest positive integer for which $z_L > y_2$.

For each $\ell = 1, \ldots, L$, let $\mathcal{J}_\ell$ denote the half-open interval $\mathcal{J}_\ell = (z_{\ell-1}, z_\ell]$; clearly,

$$\sum_{\ell=1}^{L-1}\sum_{p\in\mathcal{J}_\ell} \rho(u_p)p^{\mu-k} \leq \sum_{y_1<p\leq y_2} \rho(u_p)p^{\mu-k} \leq \sum_{\ell=1}^{L}\sum_{p\in\mathcal{J}_\ell} \rho(u_p)p^{\mu-k}. \tag{13}$$

Now let

$$v_\ell = u_{z_\ell} \qquad (1 \leq \ell \leq L).$$

For every prime $p \in \mathcal{J}_\ell$, we have

$$\begin{aligned} u_p = \frac{\log x}{\log p} - k &= \frac{\log x}{\log z_\ell + O(\delta)} - k \\ &= \frac{\log x}{\log z_\ell} - k + O\left(\frac{\delta \log x}{(\log z_\ell)^2}\right) = v_\ell + O\left(\frac{\delta}{\log_2 x}\right), \end{aligned}$$

since $\log z_\ell \geq \log y_1 \gg \sqrt{\log x \log_2 x}$. Thus, using Lemmas 1 and 2, we obtain that

$$\rho(u_p) = \rho(v_\ell) + O\left(|u_p - v_\ell| \max_{t\in\mathcal{J}_\ell} |\rho'(u_t)|\right) = \rho(v_\ell) + O\left(\delta \max_{t\in\mathcal{J}_\ell} \rho(u_t)\right).$$

Consequently,

$$\sum_{p\in\mathcal{J}_\ell} \rho(u_p)p^{\mu-k} = \rho(v_\ell)\left(\sum_{p\in\mathcal{J}_\ell} p^{\mu-k}\right)\left(1 + O(\delta)\right). \tag{14}$$

Let $\pi(t)$ denote, as usual, the number of prime numbers $p \leq t$. As the asymptotic relation $\log t \asymp \sqrt{\log x \log_2 x} = \delta^{-2}$ holds uniformly for all $t \in (y_1, y_2]$, we have

$$\pi(t) = \frac{t}{\log t} + O\left(\frac{t}{\log^2 t}\right) = \frac{t}{\log t}\left(1 + O(\delta^2)\right). \tag{15}$$

In particular,

$$\#\mathcal{J}_\ell = \pi(z_\ell(1+\delta)) - \pi(z_\ell) = \frac{\delta z_\ell}{\log z_\ell}\left(1+O(\delta)\right).$$

Noting that $p^{\mu-k} \asymp z_\ell^{\mu-k}$ for every $p \in \mathcal{J}_\ell$, we obtain that

$$\sum_{p\in\mathcal{J}_\ell} p^{\mu-k} \asymp \#\mathcal{J}_\ell \cdot z_\ell^{\mu-k} = \delta z_\ell^{\mu-k+1}\frac{1}{\log z_\ell}\left(1+O(\delta)\right). \tag{16}$$

Combining (16) with (14) using the resulting estimate in (13), we have therefore shown that

$$\left(\delta+O(\delta^2)\right)\sum_{\ell=1}^{L-1} H(\ell) \ll \sum_{y_1<p\le y_2} \rho(u_p)p^{\mu-k} \ll \left(\delta+O(\delta^2)\right)\sum_{\ell=1}^{L} H(\ell), \tag{17}$$

where

$$H(\ell) = \rho(v_\ell) z_\ell^{\mu-k+1}\frac{1}{\log z_\ell} \qquad (1\le \ell\le L).$$

Now let z_* be chosen such that

$$\rho(u_{z_*})z_*^{\mu-k+1}\frac{1}{\log z_*} = \max_{y_1\le z\le y_2} \rho(u_z)z^{\mu-k+1}\frac{1}{\log z}.$$

By Lemma 3, we have for all $z\in[y_1,y_2]$:

$$\rho(u_z)z^{\mu-k+1}\frac{1}{\log z} = \exp\left(-(1+o(1))\left(\frac{\log x\log_2 x}{2\log z} + (k-1-\mu)\log z\right)\right);$$

therefore,

$$\log z_* = (1+o(1))\sqrt{\frac{\log x\log_2 x}{2(k-1-\mu)}},$$

and

$$M = \rho(u_{z_*})z_*^{\mu-k+1}\frac{1}{\log z_*} = \exp\left(-(1+o(1))\sqrt{2(k-1-\mu)\log x\log_2 x}\right).$$

Let $\kappa\in\{1,\ldots,L\}$ be chosen such that

$$H(\kappa) = \max_{1\le\ell\le L} H(\ell).$$

We remark that for each $\ell\in\{1,\ldots,L\}$ and every number $z\in\mathcal{J}_\ell$, we have

$$\rho(u_z) = \rho(v_\ell)\left(1+O(\delta)\right) \asymp \rho(v_\ell)$$

as in our proof of (14). Thus, for $z\in\mathcal{J}_\ell$,

$$\rho(u_z)z^{\mu-k+1}\frac{1}{\log z} \asymp \rho(u_{z_\ell})z_\ell^{\mu-k}\frac{1}{\log z_\ell}.$$

In particular,

$$\rho(u_{z_L}) z_L^{\mu-k+1} \frac{1}{\log z_L} \asymp \rho(u_{y_2}) y_2^{\mu-k+1} \frac{1}{\log y_2} = o(M),$$

which shows that $z_* \notin \mathcal{J}_L$. Therefore $y_1 < z_* \le y_2$, which implies that

$$\left(\delta + O(\delta^2)\right) H(\kappa) \ll \sum_{y_1 < p \le y_2} \rho(u_p) p^{\mu-k} \ll \left(\delta + O(\delta^2)\right) L\, H(\kappa).$$

We also have that if $z_* \in \mathcal{J}_\ell$, then

$$M = \rho(u_{z_*}) z_*^{\mu-k+1} \frac{1}{\log z_*} \asymp \rho(u_{z_\ell}) z_\ell^{\mu-k+1} \frac{1}{\log z_\ell} = H(\ell) \le H(\kappa) \le M.$$

We observe that (17) immediately implies that

$$\left(\delta + O(\delta^2)\right) H(\kappa) \ll \sum_{y_1 < p \le y_2} \rho(u_p) p^{\mu-k} \ll \left(\delta + O(\delta^2)\right) L\, H(\kappa),$$

and since $L \asymp \delta^{-1} \sqrt{\log x \log_2 x} \asymp (\log x \log_2 x)^{3/4}$, it follows that

$$\sum_{y_1 < p \le y_2} \rho(u_p) p^{\mu-k} \asymp M^{1+o(1)},$$

which together with (12) implies (8). This completes the proof. $\square$

Theorem 2. *The bounds*

$$\#\mathcal{F}_k(x) \le x \exp\left(-(1+o(1)) \sqrt{(k-1) \log x \log_3 x} \right)$$

and

$$\#\mathcal{F}_k(x) \ge x \exp\left(-(1+o(1)) \sqrt{2(k+1) \log x \log_2 x} \right)$$

hold as $x \to \infty$.

Proof. Let $\mathcal{S}_1$ denote the set of positive integers $n \le x$ for which $\varphi(n)$ is y-smooth.

Let $\mathcal{S}_2$ be the set of positive integers $n \le x$ such that $P(\varphi(n))^k \mid \varphi(n)$ and $n \notin \mathcal{S}_1$. For each $n \in \mathcal{S}_2$, there exists a prime $\ell > y$ such that $\ell^k \mid \varphi(n)$. As in the proof of Theorem 5 of [7], one can show that there are at most $xy^{-k+1+o(1)}$ such positive integers $n \le x$, that is,

$$\#\mathcal{S}_2 \le xy^{-k+1+o(1)}.$$

Taking

$$y = \exp\left(\sqrt{\frac{1}{(k-1)} \log x \log_3 x} \right),$$

we see that Lemma 6 applies, and using the fact that

$$\#\mathcal{F}_k(x) \le \#\mathcal{S}_1 + \#\mathcal{S}_2,$$

we obtain the stated upper bound.

For the lower bound, we remark that $P(\varphi(n))^k \mid \varphi(n)$ whenever n has the form $n = mP(\varphi(m))^k$ for some integer m. If $m \le x/y^k$ is positive and y-smooth, it follows that

$$n = mP(\varphi(m))^k \le mP(m)^k \le x.$$

Moreover, since $P(\varphi(m)) \le P(m) \le y$, we see that for each n, there are at most y distinct representations of the form $n = mP(\varphi(m))^k$. We now choose

$$y = \exp\left(\sqrt{\frac{1}{2(k+1)}\log x \log_2 x}\right).$$

Taking into account that

$$v = \frac{\log(x/y^k)}{\log y} = \sqrt{\frac{2(k+1)\log x}{\log_2 x}} - k,$$

using Lemmas 3 and 4 we derive that

$$\begin{aligned}\#\mathcal{F}_k(x) &\ge \Psi(x/y^k, y)y^{-1} = v^{-(1+o(1))v} x y^{-k-1}\\ &= x\exp\left(-(1+o(1))\sqrt{2(k+1)\log x \log_2 x}\right).\end{aligned}$$

This completes the proof. □

Theorem 3. *The bound*

$$\#\mathcal{T}_k(x) \le x\exp\left(-(1+o(1))\sqrt{\frac{k-1}{2}\log x \log_3 x}\right)$$

holds as $x \to \infty$.

Proof. Let $\mathcal{S}_1$ denote the set of positive integers $n \le x$ for which $t(n)$ is y-smooth.

Let $\mathcal{S}_2$ be the set of positive integers $n \le x$ such that $P(t(n))^k \mid t(n)$ and $n \notin \mathcal{S}_1$. For each $n \in \mathcal{S}_2$, there exists a prime $\ell > y$ such that $\ell^k \mid t(n) \mid \varphi(n)$. As in the proof of Theorem 5 of [7], one can show that there are at most $xy^{-k+1+o(1)}$ such positive integers $n \le x$, that is,

$$\#\mathcal{S}_2 \le xy^{-k+1+o(1)}.$$

Taking

$$y = \exp\left(\sqrt{\frac{1}{(k-1)}\log x \log_3 x}\right)$$

we see that Lemma 7 applies, and using the fact that

$$\#\mathcal{T}_k(x) \le \#\mathcal{S}_1 + \#\mathcal{S}_2,$$

we obtain the stated upper bound. □

4. Remarks. A close analysis of the arguments used in the proof of Theorem 1 shows that if $\varepsilon > 0$ is fixed, then estimate of Theorem 1 holds uniformly for all $\vartheta \in [-1+\varepsilon, -\varepsilon)$, $\mu < k-1-\varepsilon$ and $k = o((\log x/\log_2 x)^{1/3})$. Furthermore, using the *Siegel-Walfisz* theorem together with (15), one can easily see that the same asymptotic formula as (2) holds for

$$S(k, \vartheta, \mu, a, b, x) = \sum_{\substack{n \le x \\ P(n)^k | n \\ P(n) \equiv a \pmod{b}}} P(n)^\mu n^\vartheta$$

for all positive coprime integers a and b with $b \le (\log x)^A$, where A is an arbitrary constant. Moreover, since k is arbitrary in (2), it follows that if we replace the range of summation $\mathcal{N}_k(x)$ by

$$\mathcal{N}_k(x) \backslash \mathcal{N}_{k+1}(x) = \{n \le x : P(n)^k \| n\},$$

then the same asymptotic formula as (2) holds for this “smaller” sum. We do not give further details in this direction.

One might consider a problem related to the estimation of $\mathcal{N}_k(x)$, namely that of estimating the number of positive integers $n \le x$ such that $p_j^2 \nmid n$ for $j = 1, \ldots, k$, where $p_1 > \ldots > p_k$ are the k largest prime factors of n.

Certainly, a more careful study of $\#\mathcal{F}_k(x)$ (perhaps in the style of our proof of Theorem 1) should lead to stronger lower and upper bounds in Theorem 2. On the other hand, it is unlikely that one can establish unconditionally the precise rate of growth of the counting function $\#\mathcal{F}_k(x)$ due to the current lack of matching lower and upper bounds for $\Phi(x, y)$. It is reasonable to expect that the upper bound of Lemma 6 is tight; in fact, similar lower bounds for $\Phi(x, y)$ may be derived from certain widely accepted conjectures about the distribution of smooth shifted primes (see the discussion in Section 8 of [2]). Nevertheless, the strongest unconditional results on smooth shifted primes (see [1], for example) are not nearly sufficient to establish a lower bound of the same order of magnitude as the upper bound of Lemma 6.

There is no reason to doubt that $\#\mathcal{T}_k(x) = x^{1+o(1)}$; however, at present we do not even see how to prove that $\#\mathcal{T}_k(x) \ge x^\alpha$ for some fixed $\alpha > 0$. Examining the integers $n = 3^k$ one easily concludes that $\#\mathcal{T}_k(x) \gg \log x$. This bound can be improved by considering powers of other small primes, but it still leaves open the question about the true order of $\#\mathcal{T}_k(x)$.

Finally, it would also be interesting to prove that the set

$$\mathcal{Q}(x) = \{p \le x : p \text{ prime}, \ P(p-1)^2 \,|p-1\}$$

is infinite. This could be considered as an weak form of the conjecture that $p = m^2 + 1$ holds infinitely often.

Acknowledgements. Most of this work was done during a visit by the first two authors to Macquarie University; the hospitality and support of this institution are gratefully acknowledged. During the preparation of this paper, W. B. was supported in part by NSF grant DMS-0070628, F. L. was supported in part by grants SEP-CONACYT 37259-E and 37260-E, and I. S. was supported in part by ARC grant DP0211459.

Résumé substantiel en français. Pour chaque entier $n \geq 2$, $P(n)$ désigne le plus grand facteur premier de n. Dans plusiers situations, en particulier celles où l'on compte le nombre d'entiers sujet à divers contraintes arithmétiques, il est souvent important de comprendre l'ensemble des entiers positifs n pour lesquels $P(n)$ est *petit* par rapport à n, ou l'ensemble des entiers positifs n pour lesquels $P(n)^2 \mid n$. De tels ensembles souvent contribuent au terme d'erreur dans une estimation. Ainsi, il est utile d'avoir de très bonnes bornes sur le nombre d'éléments dans ces ensembles. Par exemple, de tels ensembles d'entiers jouent un rôle important dans l'estimation des sommes exponentielles avec la *fonction d'Euler* $\varphi(n)$ (voir [4, 5]), dans l'étude du «diviseur premier moyen» d'un entier (voir [3]), ou dans l'évaluation de moyennes diverses et de moments de la *fonction de Smarandache* de n (voir [6, 9]).

Les entiers positifs ayant la première propriété (à savoir, $P(n)$ est petit) sont dits *lisses*, et la distribution d'entiers lisses a été considérablement étudiée. Malgré le fait qu'aucun nom ne leur a été attribué, les entiers satisfaisant la seconde propriété (à savoir, $P(n)^2 \mid n$) ont été aussi étudiés dans le corpus. Par exemple, dans [10] on a montré que pour un nombre réel fixé $r \geq 0$, l'estimation suivante a lieu lorsque $x \to \infty$:

$$\sum_{\substack{P(n)^2 \mid n \\ n \leq x}} \frac{1}{P(n)^r} = x \exp\left(-(1+o(1))\sqrt{(2r+2)\log x \log_2 x}\right). \qquad (1)$$

Ici, $\log_2 x = \log\log x$. Le terme d'erreur $o(1)$ dans l'estimation ci-dessus peut être pris sous la forme $O_r(\log_3 x/\log_2 x)$, où $\log_3 x = \log\log\log x$. Plus généralement, les sommes prises sur des entiers $n \leq x$ où $P(n)$ est dans une progression arithmétique fixée ont été estimées dans [8]. Dans ce papier, nous complétons l'estimation (1) en établissant le résultat suivant :

Théorème 1. *Soient $\vartheta \in (-1, 0]$ et $\mu < k - 1$ deux nombres fixés. L'estimation suivante a lieu :*

$$\sum_{\substack{P(n)^k \mid n \\ n \leq x}} P(n)^\mu n^\vartheta = x^{1+\vartheta} \exp\left(-(1+o(1))\sqrt{2(k-1-\mu)\log x \log_2 x}\right). \qquad (2)$$

Notre méthode ressemble à celle dans [8]. En particulier, en prenant $\vartheta = \mu = 0$ dans (2), nous obtenons l'estimation

$$\#\mathcal{N}_k(x) = x \exp\left(-(1+o(1))\sqrt{2(k-1)\log x \log_2 x}\right) \qquad (3)$$

pour le cardinal de l'ensemble

$$\mathcal{N}_k(x) = \{n \leq x : P(n)^k \mid n\}.$$

Lorsque $k = 2$, la borne (3) peut être récupérée de (1) en prenant $r = 0$, et ceci apparait aussi dans [6].

Nous considérons également une question similaire pour les valeurs de la fonction d'Euler $\varphi(n)$, et nous donnons des bornes inférieures et supérieures pour le cardinal des ensembles

$$\mathcal{F}_k(x) = \{n \leq x : P(\varphi(n))^k \mid \varphi(n)\}, \qquad k = 2, 3, 4, \ldots.$$

Notre résultat est le suivant :

Théorème 2. *Les bornes*

$$\#\mathcal{F}_k(x) \le x \exp\left(-(1+o(1))\sqrt{(k-1)\log x \log_3 x}\right)$$

et

$$\#\mathcal{F}_k(x) \ge x \exp\left(-(1+o(1))\sqrt{2(k+1)\log x \log_2 x}\right)$$

ont lieu lorsque $x \to \infty$*.*

Finalement, pour un entier impair n, nous définissons $t(n)$ comme étant l'ordre multiplicatif de 2 modulo n, et nous donnons des bornes supérieures sur le cardinal des ensembles

$$\mathcal{T}_k(x) = \{n \le x : n \text{ odd},\ P\left(t(n)\right)^k \mid t(n)\}, \qquad k = 2, 3, 4, \dots .$$

Notre résultat est le suivant :

Théorème 3. *La borne*

$$\#\mathcal{T}_k(x) \le x \exp\left(-(1+o(1))\sqrt{\frac{k-1}{2}\log x \log_3 x}\right)$$

a lieu lorsque $x \to \infty$*.*

References

1. R. C. Baker and G. Harman, *Shifted primes without large prime factors*, Acta Arith. **83** (1998), 331–361.
2. W. D. Banks, J. B. Friedlander, C. Pomerance and I. E. Shparlinski, *Multiplicative structure of values of the Euler function,*, High Primes and Misdemeanours: Lectures in Honour of the 60th Birthday of Hugh Cowie Williams, 29–47; Fields Inst. Commun., vol. 41, Amer. Math. Soc., Providence, RI, 2004.
3. W. Banks, M. Garaev, F. Luca and I. Shparlinski, *Uniform distribution of the fractional part of the average prime divisor*, Forum Math. **17**, 885–901.
4. W. D. Banks and I. E. Shparlinski, *Congruences and exponential sums with the Euler function,*, High Primes and Misdemeanours: Lectures in Honour of the 60th Birthday of Hugh Cowie Williams, 49–59; Fields Inst. Commun., vol. 41, Amer. Math. Soc., Providence, RI, 2004.
5. W. Banks and I. E. Shparlinski, *Congruences and rational exponential sums with the Euler function*, Rocky Mountain J. Math. **36** (2006), 1415–1426.
6. K. Ford, *The normal behavior of the Smarandache function*, Smarandache Notions J. **10** (1999), 81–86.
7. J. B. Friedlander, C. Pomerance and I. E. Shparlinski, *Period of the power generator and small values of Carmichael's function*, Math. Comp. **70** (2001), 1591–1605.
8. A. Ivić, *On sums involving reciprocals of the largest prime factor of an integer, II*, Acta Arith. **71** (1995), 229–251.
9. A. Ivić, *On a problem of Erdös involving the largest prime factor of* n, Monatsh. Math. **145** (2005), no. 1, 35-46.

10. A. Ivić and C. Pomerance, *Estimates for certain sums involving the largest prime factor of an integer*, Topics in Analytic Number Theory, vol. I-II, pp. 769–789 (Budapest, 1981); Colloq. Math. Soc. János Bolyai, vol. 34, North-Holland, Amsterdam, 1984.
11. G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge Studies in Advanced Mathematics, vol. 46, Cambridge University Press, Cambridge, 1995.

W. D. BANKS
DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MISSOURI
COLUMBIA, MO 65211 USA
E-MAIL: `bbanks@math.missouri.edu`

F. LUCA
INSTITUTO DE MATEMÁTICAS
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
C.P. 58089, MORELIA
MICHOACÁN, MÉXICO
E-MAIL: `fluca@matmor.unam.mx`

I. E. SHPARLINSKI
DEPARTMENT OF COMPUTING, MACQUARIE UNIVERSITY
SYDNEY, NSW 2109, AUSTRALIA
E-MAIL: `igor@ics.mq.edu.au`