Ricerche di Matematica
Vol. XXXIX, fasc. 1°, (1990), 131-140

On the average prime factor of an integer and some related problems.

JEAN-MARIE DE KONINCK (*) - ALEKSANDAR IVIĆ (**)

RIASSUNTO. – Si investigano formule asintotiche che riguardano funzioni additive $f(n) = \sum_{p|n} p^\rho L(p)$ e $F(n) = \sum_{p^\alpha \| n} \alpha p^\rho L(p)$, includendo il loro comportamento negli intervalli corti. Si suppone che $\rho < 0$ e $L(x)$ sia una funzione a oscillazione lenta. Le due funzioni connesse con f e F sono $P_*(n) = (\sum_{p|n} p)/\omega(n)$ e $P^*(n) = (\sum_{p^\alpha \| n} \alpha p)/\Omega(n)$, che rappresentano il divisore primo medio di n. Anche il loro comportamento negli intervalli corti è discusso.

KEY WORDS: *Arithmetic functions, Prime factors, Slowly oscillating functions.*

1. Introduction.

For an integer $n \geq 2$, let $P(n)$ denote the largest prime factor of n, and let $\omega(n)$ and $\Omega(n)$ denote the number of distinct prime factors and the number of all prime factors of n, respectively. In [3] De Koninck - Ivić showed that

$$(1.1) \qquad \sum_{2 \leq n \leq x} P(n) = x^2 \left(\frac{d_1}{\log x} + \frac{d_2}{\log^2 x} + \ldots + \frac{d_n}{\log^n x} + O\left(\frac{1}{\log^{M+1} x} \right) \right)$$

for any fixed integer $M \geq 1$, where $d_1 = \pi^2/12$, and all the constants d_j

(*) Jean-Marie De Koninck - Département de mathématiques et statistique - Université laval - Québec GIK 7P4 - Canada.
(**) Aleksandar Ivić - Katedra Matematike - RGF-a Universiteta u Beogradu - Djušina 7, 11000 Beograd - Yugoslavia.

are effectively computable. Moreover, (1.1) remains valid (with different constants d_j) if $P(n)$ is replaced by any of the two functions

$$P_*(n) = \frac{\beta(n)}{\omega(n)}, \quad P^*(n) = \frac{B(n)}{\Omega(n)}, \tag{1.2}$$

where

$$\beta(n) = \sum_{p|n} p, \quad B(n) = \sum_{p^\alpha \| n} \alpha p. \tag{1.3}$$

Here and later p denotes primes, and $p^\alpha \| n$ means that p^α exactly divides n. Various sums involving the additive functions $\beta(n)$ and $B(n)$ were investigated in [2], [7], [8], [11] and [13]. The functions $P_*(n)$ and $P^*(n)$ may be called the average prime factor (divisor) of n, where in the first case the averaging is over distinct prime factors of n, and in the second it is over all prime factors.

The functions $\beta(n)$ and $B(n)$ are special cases of the more general additive functions

$$f(n) = \sum_{p|n} p^\rho L(p), \quad F(n) = \sum_{p^\alpha \| n} \alpha p^\rho L(p), \tag{1.4}$$

where ρ is a real number, and $L(x)$ is a slowly oscillating (or slowly varying) function. Slowly oscillating functions arise naturally in many branches of number theory, and for a comprehensive account the reader is referred to the monographs of Bingham et al. [1] and E. Seneta [15]. They are continuous, positive functions for $x \ge x_0 > 0$ and satisfy $\lim_{x\to\infty} L(cx)/L(x) = 1$ for any $c > 0$. By a fundamental result of J. Karamata [14], who founded the theory of slowly oscillating functions, the limit is uniform for $0 < a \le c \le b < \infty$ and any $0 < a < b$. This result is used to show that any slowly oscillating function is necessarily of the form

$$L(x) = A(x) \exp\left(\int_{x_0}^{x} \eta(t)\frac{dt}{t}\right), \lim_{x\to\infty} A(x) = A > 0, \lim_{x\to\infty} \eta(x) = 0. \tag{1.5}$$

Sums involving $P(n)$ and the additive functions defined by (1.4) are investigated by De Koninck - Mercier [6]. Therein it is shown that under suitable conditions

$$\sum_{2\le n\le x} f(P(n)) = (1 + o(1)) \sum_{2\le n\le x} f(n) \qquad (x \to \infty), \tag{1.6}$$

and in fact both sums are evaluated asympotically. Sums of the form

$$\sum_{2\le n\le x} \frac{f(n)}{\omega(n)}, \tag{1.7}$$

are extensively treated by De Koninck - Ivić [5], who discovered their probabilistic importance. Their methods can be also used to deal with the sums

$$\sum_{2\le n\le x} \frac{F(n)}{\Omega(n)}, \tag{1.8}$$

where both f and F are defined by (1.4). Obviously (1.7) and (1.8), for $\rho = 1$, $L(x) = 1$, reduce to the summatory functions of $P_*(n)$ and $P^*(n)$, respectively.

In the investigations carried out in [4] and [5] it was supposed that $L(x)$ was of the special form

$$L(x) = A \exp\left(\int_{x_0}^{x} \eta(t)\frac{dt}{t}\right), \lim_{x\to\infty} A(x) = A > 0, \lim_{x\to\infty} \eta(x) = 0. \tag{1.9}$$

Since every slowly oscillating function is by (1.5) asymptotic to a function of the form (1.9), this is not a severe restriction. Nevertheless, it seems of interest to investigate these problems, which include the asymptotic evaluation of $\sum_{2\le n\le x} f(n)$ and the sums (1.7), (1.8), when $L(x)$ is of a most general form. When $\rho > 0$, such an analysis will be carried out in Section 2 for $\sum_{2\le n\le x} f(n)$. We shall also establish (1.6) when $\rho > 0$ in the general case, and discuss some related problems.

In Section 3, we turn to the behavior of $f(n)$ and $F(n)$ is short intervals. It was proved by De Koninck - Ivić [4] that, if $f(n)$, $F(n)$ are defined by (1.4) with $\rho > 0$ and $L(x)$ of the form (1.9), $h = o(x)$ and $h \ge x^{7/12}\log^{22} x$ as $x \to \infty$, then

$$\sum_{x<n\le x+h} f(n) = (\zeta(1+\rho) + o(1))\frac{hx^\rho L(x)}{\log x}, \tag{1.10}$$

$$\sum_{x<n\le x+h} F(n) = (\zeta(1+\rho)+o(1))\frac{hx^\rho L(x)}{\log x}\,. \tag{1.11}$$

In [4] we have also shown the following results on the behavior of the average prime factor of an integer in short intervals: if $h = o(x)$ and $h \ge x^{7/12}\log^{22} x$ as $x \to \infty$, then

$$\sum_{x<n\le x+h} P_*(n) = (E+o(1))\frac{hx}{\log x}\,,\qquad E = \sum_{m=1}^{\infty}\frac{1}{m^2(\omega(m)+1)} = 1.30642\ldots\,, \tag{1.12}$$

$$\sum_{x<n\le x+h} P^*(n) = (F+o(1))\frac{hx}{\log x}\,,\qquad F = \sum_{m=1}^{\infty}\frac{1}{m^2(\Omega(m)+1)} = 1.28518\ldots\,, \tag{1.13}$$

We shall indicate how (1.12) and (1.13) may be obtained without the restriction that $L(x)$ is of the special form (1.9). In fact this follows as the special case from the corresponding results on the behavior of $f(n)/\omega(n)$ and $F(n)/\Omega(n)$ is short intervals.

2. Evaluation of the summatory functions.

We shall consider now the summatory functions of $f(n)$ and $F(n)$ as defined by (1.4), where we assume that $\rho > 0$ and $L(x)$ is an arbitrary slowly oscillating function. Note that, for any given $\delta > 0$ and $x \ge x_1 = x_1(\delta)$ we have from (1.5)

$$(A-\delta)L^*(x) \le L(x) \le (A+\delta)L^*(x)\,, \tag{2.1}$$

where

$$L^*(x) = \exp\left(\int_{x_0}^{x}\eta(t)\frac{dt}{t}\right)\,,\ \lim_{x\to\infty}\eta(x) = 0\,, \tag{2.2}$$

so that $L^*(x)$ is differentiable. If p denotes primes and m, n denote natural numbers, then for any given $\delta > 0$

$$\begin{aligned}\sum_{n\le x} f(n) &= \sum_{n\le x}\sum_{p\mid n} p^\rho L(p) = \sum_{pm\le x} p^\rho L(p) = \sum_{m\le\frac{x}{2}}\sum_{p\le\frac{x}{m}} p^\rho L(p)\\ &= \sum_{m\le\frac{x}{2}}\sum_{x_1<p\le\frac{x}{m}} p^\rho L(p) + \sum_{m\le\frac{x}{2}}\sum_{p\le x_1} p^\rho L(p)\\ &= (A+O(\delta))\sum_{m\le\frac{x}{2}}\sum_{x_1<p\le\frac{x}{m}} p^\rho L^*(p) + O_\delta(x)\\ &= (A+O(\delta))\sum_{m\le\frac{x}{2}}\int_{x_1-0}^{x/m} t^\rho L^*(t)d\pi(t) + O_\delta(x)\\ &= (A+O(\delta))\sum_{m\le\frac{x}{2}}\int_{x_1}^{x/m}\frac{t^\rho L^*(t)}{\log t}dt +\\ &+ (A+O(\delta))\sum_{m\le\frac{x}{2}}\int_{x_1-0}^{x/m} t^\rho L^*(t)dR(t) + O_\delta(x)\\ &= S_1 + S_2 + O_\delta(x)\,,\end{aligned}$$

say. Here we used the prime number theorem in the standard form

$$\pi(x) = \sum_{p\le x} 1 = \mathrm{li}\,x + R(x)\,,\ \mathrm{li}\,x = \int_2^x\frac{dt}{\log t}\,,\ R(x) = O\left(xe^{-\sqrt{\log x}}\right)\,.$$

Change of variable $t = x/u$ gives

$$\begin{aligned}S_1 &= (A+O(\delta))\sum_{m\le\frac{x}{2}}\int_{x_1}^{x/m}\frac{t^\rho L^*(t)}{\log t}dt\\ &= (A+O(\delta))\sum_{m\le x/x_1}\int_m^{x/x_1} x^{1+\rho}u^{-\rho-2}\frac{L^*(x/u)}{\log(x/u)}du \qquad (2.3)\\ &= (A+O(\delta))x^{1+\rho}\int_1^{x/x_1}[u]u^{-\rho-2}\frac{L^*(x/u)}{\log(x/u)}du\,.\end{aligned}$$

The major contribution to $\sum_{n\le x} f(n)$ comes from S_1, since we shall presently show that

$$S_2 = (A+O(\delta))\sum_{m\le\frac{x}{2}}\int_{x_1-0}^{x/m} t^\rho L^*(t)dR(t) \ll x^{1+\rho}L(x)e^{-\sqrt{\log x}}\,. \tag{2.4}$$

Since $(L^*(x))' = x^{-1}\eta(x)L^*(x)$, integration by parts gives

$$S_2 \ll \sum_{m\le \frac{x}{2}} (|R(x/m)|(x/m)^\rho L^*(x/m)+1) +$$

$$+ \sum_{m\le\frac{x}{2}} \int_{x_1}^{x/m} |R(t)|t^{\rho-1}(\rho+|\eta(t)|)L^*(t)dt\,.$$

But $R(x) \ll x\mathrm{e}^{-\sqrt{\log x}}$ and $\eta(x)$ is bounded, so that

$$S_2 \ll x + x^{\rho+1}\sum_{m\le\frac{x}{2}} \mathrm{e}^{-\sqrt{\log(x/m)}}L^*(x/m)m^{-\rho-1}+$$

$$+ \sum_{m\le\frac{x}{2}}\int_{x_1}^{x/m} t^\rho \mathrm{e}^{-\sqrt{\log t}}L^*(t)dt$$

$$\ll x^{\rho+1}\sum_{m\le\frac{x}{2}} \mathrm{e}^{-\sqrt{\log(x/m)}}L^*(x/m)m^{-\rho-1}\,,$$

since for any $\rho > -1$

$$\int_{x_0}^{x} t^\rho L(t)dt \sim \frac{x^{\rho+1}}{\rho+1}L(x) \quad (x\to\infty)$$

if and only if $L(x)$ is slowly oscillating (see [1]). The portion of the last sum for which $\sqrt{x} < m \le x/2$ is clearly negligible. Setting $L_1(x) = \mathrm{e}^{-\sqrt{\log x}}L^*(x)$, it is seen that $L_1(x)$ is of the form (1.9). Therefore it is differentiable and $(L_1(x)x^{\rho/2})' > 0$, whence $L_1(x)x^{\rho/2}$ is increasing for sufficiently large x. Thus we have

$$\sum_{m\le\sqrt{x}} L_1(x/m)m^{-\rho-1} = x^{-\rho/2}\sum_{m\le\sqrt{x}} L_1(x/m)(x/m)^{\rho/2}m^{-\rho/2-1}$$

$$\le x^{-\rho/2}L_1(x/1)(x/1)^{\rho/2}\sum_{m\le\sqrt{x}} m^{-\rho/2-1}$$

$$\ll L_1(x) \ll \mathrm{e}^{-\sqrt{\log x}}L(x)\,,$$

since $L^*(x) \ll L(x)$. This establishes (2.4).

To evaluate S_1, note that if $0 < h(u) \ll u^{-c}$ for some $c > 1$ and h is Riemann integrable, then for any slowly oscillating function $L(x)$, we have

$$\int_1^{x/x_1} h(u)L(x/u)du = (1+o(1))L(x)\int_1^\infty h(u)du \tag{2.5}$$

if x_1 is a sufficiently large constant. This follows e.g. from Lemma 1 of J.P. Tull [16]. Noting that $L^*(x)/\log x$ is slowly oscillating and $\delta > 0$ can be arbitrarily small, we see that (2.3) yields, as $x\to\infty$.

$$S_1 \sim (A+O(\delta))x^{1+\rho}\frac{L^*(x)}{\log x}\int_1^\infty [u]u^{-\rho-2}du$$

$$\sim x^{1+\rho}\frac{L(x)}{\log x}\int_{1-0}^\infty [u]u^{-\rho-2}du$$

$$= \frac{x^{1+\rho}L(x)}{(1+\rho)\log x}\int_{1-0}^\infty u^{-\rho-1}d[u] = \frac{x^{1+\rho}L(x)}{(1+\rho)\log x}\zeta(1+\rho)\,.$$

Collecting the above estimates, we obtain

THEOREM 1. – *If $f(n)$ is given by (1.4) where $\rho > 0$ and $L(x)$ is slowly oscillating, then as $x\to\infty$*

$$\sum_{2\le n\le x} f(n) = \left(\frac{\zeta(1+\rho)}{1+\rho}+o(1)\right)\frac{x^{1+\rho}L(x)}{\log x}\,. \tag{2.6}$$

By the same method, only with some more technical complications, we could obtain the analogous result if $f(n)$ is replaced by $F(n)$. The very generality of $L(x)$ hinders a result sharper than (2.6). In cases of special $L(x)$, we can obtain, by the preceding method, a much more precise result. Thus for $f(n) = \beta(n)(\rho = 1, L(x) = 1)$ we actually obtain

$$\sum_{2\le n\le x}\beta(n) = x^2\int_1^{x/2}\frac{[u]}{u^3\log(\frac{x}{u})}du + O\left(x^2\mathrm{e}^{-\sqrt{\log x}}\right)\,, \tag{2.7}$$

and the error term could be slightly sharpened by using the strongest form of the prime number theorem (see Ch. 12 of [12]). If we use now in (2.7)

$$\frac{1}{\log(\frac{x}{u})} = \frac{1}{\log x}\left\{\sum_{j=0}^{M}\left(\frac{\log u}{\log x}\right)^j + O\left(\left(\frac{\log u}{\log x}\right)^{M+1}\right)\right\}$$

and simplify, we obtain the analogue of (1.1) for $\beta(n)$, which easily yields then (1.1) itself by the method of De Koninck - Ivić [3]. In fact, (2.7) shows that

$$\sum_{2\le n\le x}\beta(n) = x^2L(x) + O\left(x^2\mathrm{e}^{-\sqrt{\log x}}\right)\,,$$

where

$$L(x)=\int_1^{x/2}\frac{[u]}{u^3\log(\frac{x}{u})}du$$

is slowly oscillating function whose asymptotic expansion is of the form given by the right-hand side of (1.1). From Theorem 1, we easily deduce

THEOREM 2. – *If $f(n)$ is given by (1.4), where $\rho>0$ and $L(x)$ is slowly oscillating, then as $x\to\infty$*

$$\sum_{2\le n\le x} f(P(n))=\left(\frac{\zeta(1+\rho)}{1+\rho}+o(1)\right)\frac{xf(x)}{\log x}. \tag{2.8}$$

This result was proved by De Koninck - Mercier [6], but under a rather severe restriction on $L(x)$, whereas the above result holds for most general $L(x)$. To obtain (2.8), we use Theorem 1 by observing that the difference between the left sides of (2.6) and (2.8) is

$$\sum_{2\le n\le x}\sum_{p|n,p<P(n)}p^\rho L(p)=\sum_{pm\le x,p<P(m)}p^\rho L(p)=\sum_{pm\le x,p<\sqrt{x},p<P(m)}p^\rho L(p)$$
$$\ll x\sum_{p\le\sqrt{x}}p^{\rho-1}L(p)$$
$$\ll x^{1+\frac{\rho}{2}}L^*(\sqrt{x})\log\log x\ll x^{1+\frac{3\rho}{4}},$$

since $\sqrt{x}<p\le x$ is impossible. Here we used the fact that $L(x)\ll x^\epsilon$ and $p^\rho L(p)\ll p^\rho L^*(p)$, the last function being increasing for large p. But for $\rho>0$

$$x^{1+\frac{3\rho}{4}}=o\left(x^{1+\rho}\frac{L(x)}{\log x}\right)\quad(x\to\infty),$$

hence (2.8) follows.

3. Sums over short intervals.

We pass now to sums over short intervals. The result that generalizes (1.12) and (1.13), which concern the average prime divisor in short intervals, is contained in

THEOREM 3. – *Let $f(n)$ and $F(n)$ be defined by (1.4), where $\rho>0$ and $L(x)$ is a slowly oscillating function. If $h=o(x)$ as $x\to\infty$ and $h\ge x^{7/12}\log^{22}x$, then*

$$\sum_{x<n\le x+h} f(n)=(\zeta(1+\rho)+o(1))\frac{hx^\rho L(x)}{\log x} \tag{3.1}$$

and

$$\sum_{x<n\le x+h} F(n)=(\zeta(1+\rho)+o(1))\frac{hx^\rho L(x)}{\log x}. \tag{3.2}$$

This result, with $L(x)$ of the form (1.9), was proved by De Koninck - Ivić [4]. The general case follows, as in Section 2, by using the function $L^*(x)$ (defined by (2.2)) and the method of [4]. The details are therefore omitted. We remark, however, that the range $h\ge x^{7/12}\log^{22}x$ comes from the result of the second author [10], who established the asymptotic formula

$$\pi(x+h)-\pi(x)=(1+o(1))\frac{h}{\log x}\quad(x\to\infty) \tag{3.3}$$

for $h\ge x^{7/12}\log^{22-\delta}x$ with some $\delta>0$. Recently D.R. Heath-Brown [9] gave an extensive treatment of primes in short intervals, and showed in particular that (3.3) is true for $h\ge x^{7/12}$. Thus in this range, (3.1) and (3.2) also hold.

We could also prove

THEOREM 4. – *If $f(n)$ is defined by (1.4), where $\rho>0$ and $L(x)$ is slowly oscillating, $h=o(x)$ as $x\to\infty$ and $h\ge x^{7/12}\log^{22}x$, then*

$$\sum_{x<n\le x+h}\frac{f(n)}{\omega(n)}=(c_\rho+o(1))\frac{hf(x)}{\log x},\ c_\rho=\sum_{m=1}^\infty\frac{1}{m^{\rho+1}(\omega(m)+1)}. \tag{3.4}$$

We omit the proof, since it is based on the methods of [4] and [5], and the use of the function $L^*(x)$ similarly as in the proof of Theorem 1.

Finally we remark that in all the preceding discussions we assumed $\rho>0$ in (1.4). It is possible to treat also the cases $\rho=0$ and $\rho<0$, as was done by De Koninck -Mercier [6] and De Koninck - Ivić [5]. However in these cases, especially the former, additional hypotheses on $L(x)$ seem to be necessary in order to obtain satisfactory asymptotics.

REFERENCES

[1] N.H. Bingham, C.M. Goldie and J.L. Teugels: *Regular variation*, Cambridge University Press, Cambridge, 1987.
[2] J.M. De Koninck and A. Ivić: *Topics in Arithmetical Functions*, Notas de Matemática **72**, North-Holland, Amsterdam, 1980.
[3] J.M. De Koninck and A. Ivić: *The distribution of the average prime divisor of an integer*, Archiv der Math. **43** (1984), 37-43.
[4] J.M. De Koninck and A. Ivić: *The average prime divisor of an integer in short intervals*, Archiv der Math. **52** (1989), 440-448.
[5] J.M. De Koninck and A. Ivić: *Random sums related to prime divisors of an integer*, Publs. Inst. Mathématique (Belgrade), to appear.
[6] J.M. De Koninck and A. Mercier: *Les fonctions arithmétiques et le plus grand facteur premier*, Acta Arith. **52** (1989), 25-48.
[7] P. Erdös and A. Ivić: *Estimates for sums involving the largest prime factor of an integer and certain related additive functions*, Studia Scien. Math. Hung. **15** (1980), 183-199.
[8] P. Erdös, A. Ivić and C. Pomerance: *On sums involving reciprocals of the largest prime factor of an integer*, Glasnik Matematički **21** (1986), 27-44.
[9] D.R. Heath-Brown: *The number of primes in a short interval*, J. Reine Angew. Math. **389** (1988), 22-63.
[10] A. Ivić: *On sums of large differences between consecutive primes*, Math. Annalen **241** (1979), 1-9.
[11] A. Ivić: *Sums of reciprocals of the largest prime factor of an integer*, Archiv der Math. **36** (1981), 57-61.
[12] A. Ivić: *The Riemann zeta-function*, John Wiley & Sons, New York, 1985.
[13] A. Ivić and C. Pomerance: *Estimates of certain sums involving the largest prime factor of an integer*, in:, Coll. Math. Soc. J. Bolyai **34**, Topics in classical number theory, North-Holland, Amsterdam 1984, 769-789.
[14] J. Karamata: *Sur un mode de croissance régulière des fonctions*, Mathematica (Cluj) **4** (1930), 38-53.
[15] E. Seneta: *Regularly varying functions*, LNM 508, Berlin-Heidelberg- New York, 1976.
[16] J.P. Tull: *Convolutions of slowly oscillating functions*, Proc. Amer. Math. Soc. **13** (1962), 225-228.

Pervenuto alla Redazione il 24 ottobre 1989