JOURNAL OF NUMBER THEORY 33, 293–310 (1989)

Additive Functions and the Largest Prime Factor of Integers

JEAN-MARIE DEKONINCK*

Département de Mathématiques, Université Laval, Québec, Canada G1K 7P4

IMRE KÁTAI†

Eötvös Loránd University, Computer Centre, 1117 Budapest, Bogdánfy u. 10/B, Hungary

AND

ARMEL MERCIER‡

Département de Mathématiques, Université du Québec, Chicoutimi, Canada G7H 2B1

Communicated by P. T. Bateman

Received March 31, 1988

For an integer $n \geqslant 2$, let $P(n)$ be the largest prime factor of n. For an arbitrary strongly additive function f, let $V_f(x) = \sum_{a_n \leqslant x} f(a_n)$ and $U_f(x) = \sum_{a_n \leqslant x} f(P(a_n))$, where $a_1 < a_2 < \cdots$ is an infinite sequence of natural numbers. Assume that $f(p) = R(p)$, for each prime p, where $R(x) = x^\rho L(x)$, with $\rho \geqslant 0$ and L a slowly oscillating function; further assume that, if $\rho > 0$, $L(x)$ grows fast enough with x. We show that, if there exist positive constants ε, c, x_0 such that $\#\{a_n \leqslant x \mid P(a_n) > x^{1/2+\varepsilon}\} > c\#\{a_n \mid n \leqslant x\}$ holds for each $x > x_0$, then $U_f(x) \sim V_f(x)$. We further prove that

$$\sum_{n \leqslant x} \min(f(n), f(n+1), f(n+2)) \sim \sum_{n \leqslant x} \min(f(P(n)), f(P(n+1)), f(P(n+2))).$$

We also compare $\sum_{j=0}^{k-1} f(n+j)$ with $\sum_{j=0}^{k-1} f(P(n+j))$ when $n \to \infty$ and $k = k_n$ increases with n. Finally we show that $\lim_{n \to \infty} x^{-1} \sum_{n \leqslant x} f(n)/f(P(n)) = 1$.

* Supported by Grant CRSNG A-8729.

† Supported by the Research Fund of Hungary 907.

‡ Supported by Grant CRSNG A-3508.

0022-314X/89 $3.00

1. Introduction

For an integer $n \geqslant 2$, let $P(n)$ be the largest prime factor of n. For an arbitrary strongly additive function f, let

$$A_f(x) = \sum_{2 \leqslant n \leqslant x} f(n), \qquad B_f(x) = \sum_{2 \leqslant n \leqslant x} f(P(n)).$$

Alladi and Erdös [1] proved that if $f(n) = \beta(n) = \sum_{p \mid n} p$, then

$$A_\beta(x) \sim B_\beta(x) \sim \frac{\pi^2}{12} \frac{x^2}{\log x}.$$

The asymptotic expansion of $A_\beta(x)$ has been given by DeKoninck and Ivić [3].

Similar questions were considered for a quite large class of functions by DeKoninck and Mercier [5]. Namely, they proved that if f is a strongly additive function and $f(p) = R(p)$, for each prime number p, where $R(x)$ is a function of the form $R(x) = x^\rho L(x)$, with slowly oscillating L, then

(i) in the case $\rho > 0$

$$A_f(x) \sim B_f(x) \sim \frac{x^{\rho+1}\zeta(\rho+1)}{\rho+1} \frac{L(x)}{\log x}, \tag{1}$$

(ii) while in the case $\rho = 0$, $(L'/L)(x) x \log x \to \infty$,

$$A_f(x) \sim B_f(x) \sim x \int_2^x \frac{L(u)}{u \log u}\, du. \tag{2}$$

They treated further cases as well and pointed out that in some of them the relation $A_f(x) \sim B_f(x)$ is no longer satisfied.

In this paper we shall show that the summatory functions of $f(n)$ and of $f(P(n))$ are asymptotically equal under the conditions (i) and (ii), even if summation is taken over various subsets of integers, or integers contained in a short interval. In these general cases we are unable to give a simple asymptotic expression (like the second relation in (1), (2) for the corresponding sums).

Throughout this paper, except in the remark following Theorem 5, we assume that $R(x) = x^\rho L(x)$, R: $[1, \infty) \to [1, \infty)$, $\rho \geqslant 0$, and L is a slowly oscillating function, i.e., such that

$$\lim_{\substack{x_1 < x_2 < 2x_1 \\ x_1 \to \infty}} \frac{L(x_2)}{L(x_1)} \to 1, \tag{3}$$

furthermore that

Case #1: $\rho > 0$ or

Case #2: $\rho = 0$ and L is differentiable and $(L'/L)(x)\, x \log x \to \infty$ is satisfied.

Assume furthermore that f is a strongly additive function defined on the set of primes p by $f(p) = R(p)$. We denote by $\mathscr{F}$ the set of functions f satisfying the above requirements.

2. Summing on Certain Sequences of Integers

Let $\mathscr{A} = \{a_1 < a_2 < \cdots\}$ be an infinite sequence of natural numbers. Let

$$V_f(x) := \sum_{a_n \leqslant x} f(a_n), \qquad U_f(x) := \sum_{a_n \leqslant x} f(P(a_n))$$

and $A(x) = \#\{a_n \leqslant x\}$.

Theorem 1. *Let $f \in \mathscr{F}$. Assume that there exist positive constants ε, c, x_0 such that*

$$\#\{a_n \leqslant x \mid P(a_n) > x^{1/2+\varepsilon}\} > cA(x) \tag{4}$$

holds for each $x > x_0$. Then

$$U_f(x) \sim V_f(x). \tag{5}$$

First we prove the following

Lemma 1. *Let $f \in \mathscr{F}$ and assume furthermore that case #2 holds. Then there is a constant c_1 that may depend on L, such that, for $x^{1/4} \leqslant y \leqslant x$,*

$$\max_{\substack{n \leqslant x \\ P(n) \leqslant y}} f(n) \leqslant c_1 L(y). \tag{6}$$

Proof of Lemma 1. For an arbitrary $\kappa > 0$ let $u_0(\kappa)$ be a constant such that $\lambda(u) \geqslant \kappa$ whenever $u \geqslant u_0(\kappa)$. The existence of $u_0(\kappa)$ follows from the fact that $\lim_{x \to \infty} \lambda(x) = +\infty$. Then, for $u_0(\kappa) < u < v$,

$$\log \frac{L(v)}{L(u)} = \int_u^v \frac{\lambda(t)}{t \log t}\, dt \geqslant \kappa \int_u^v \frac{dt}{t \log t} = \kappa \log \frac{\log v}{\log u},$$

and so

$$\frac{L(v)}{L(u)} \geqslant \left(\frac{\log v}{\log u}\right)^{\kappa}. \tag{7}$$

In (7), set $\kappa = 2$, $u = y^{4/k}$, $v = y$, then, for each $k \in \mathbf{N}$ such that $y^{4/k} > u_0(2)$, we have

$$L(y^{4/k}) \leqslant \frac{16}{k^2} L(y). \tag{8}$$

Now take any positive integer n such that $P(n) \leqslant y$; then clearly we can write $n = n_1 \cdot n_2$, where $P(n_1) < u_0(2)$ and each prime factor of n_2 is not less than $u_0(2)$. Using (8) and the fact that for each $4 < k \in \mathbf{N}$ one has $P_k(n) \leqslant x^{1/k} \leqslant y^{4/k}$,

$$\begin{aligned} f(n_2) &= \sum_{\substack{p \mid n \\ P(n_1) < p \leqslant y}} L(p) \leqslant \sum_{4 \leqslant k < 4 \log y / \log u_0(2)} L(y^{4/k}) \\ &\leqslant L(y) \sum_{4 \leqslant k < 4 \log y / \log u_0(2)} \frac{16}{k^2} < \frac{8\pi^2}{3} L(y). \end{aligned}$$

Since $\max_{P(m) < u_0(2)} f(m)$ is bounded, it follows that (6) is true. ∎

Proof of Theorem 1. Assume first that Case #2 holds. Let ε be a small positive constant, and

$$\begin{aligned} \mathscr{B}_1 &= \{a_n \leqslant x \mid P(a_n) \leqslant x^{1/2 + \varepsilon/2}\}, \\ \mathscr{B}_2 &= \{a_n \leqslant x \mid P(a_n) > x^{1/2 + \varepsilon/2}\}. \end{aligned} \tag{9}$$

From Lemma 1 we have

$$\sum_{\mathscr{B}_1} f(a_n) \ll A(x) L(x^{1/2 + \varepsilon/2}).$$

If $a_n \in \mathscr{B}_2$, then $a_n = P(a_n) b_n$, $b_n \leqslant x^{1/2 - \varepsilon/2}$, consequently, $f(a_n) - f(P(a_n)) = f(b_n) \ll L(x^{1/2 - \varepsilon/2})$. So, we have

$$V_f(x) - U_f(x) \ll A(x) L(x^{1/2 + \varepsilon/2}).$$

From (4) it is clear that

$$U_f(x) \gg A(x)\, L(x^{1/2 + \varepsilon}).$$

Since $L(x^{1/2 + \varepsilon/2}) = o(1)\, L(x^{1/2 + \varepsilon})$ holds for $x \to \infty$, therefore (5) is true.

Assume now that Case #1 holds. Since $\log R(y)/\log y \to \rho$, then

$$y^{\rho-\delta_y} < R(y) < y^{\rho+\delta_y},$$

with a suitable function δ_y, $\delta_y \to 0$. Since the number $d(n)$ of the divisors of n is $\ll n^\delta$ for each positive constant δ, then for an integer $n = p_1 \cdots p_r$ $(p_1 < \cdots < p_r)$ we have

$$f(n) = R(p_1) + \cdots + R(p_r) \ll R(p_r)\, n^\delta.$$

Splitting $\mathscr{A}$ into two parts, $\mathscr{B}_1$ and $\mathscr{B}_2$, as above, and proceeding as in Case #2, we can deduce Theorem 1 in this case as well. ∎

Remarks. (1)

$$\#\{n \leqslant x \mid P(n^2+1) > x^{1+\varepsilon}\} \gg x \tag{10}$$

holds for a suitable $\varepsilon > 0$ as $x \to \infty$.

(2) Let $l \neq 0$ be an integer, p run over the whole set of primes. Then

$$\#\{p \leqslant x \mid P(p+l) > x^{1/2+\varepsilon}\} \gg \pi(x) \tag{11}$$

holds for a suitable $\varepsilon > 0$ as $x \to \infty$.

For the proof of (1)–(2) see Hooley [6]. Since the inequalities (10), (11) guarantee the fulfilment of (4) in Theorem 1, the following relations hold:

$$\sum_{p \leqslant x} f(p+l) \sim \sum_{p \leqslant x} f(P(p+l)), \tag{12}$$

$$\sum_{n \leqslant x} f(n^2+1) \sim \sum_{n \leqslant x} f(P(n^2+1)). \tag{13}$$

3. On $\min(f(n), f(n+1), f(n+2))$

THEOREM 2. *For each f in $\mathscr{F}$, one has*

$$\sum_{n \leqslant x} \min(f(n), f(n+1), f(n+2))$$
$$\sim \sum_{n \leqslant x} \min(f(P(n)), f(P(n+1)), f(P(n+2))). \tag{14}$$

LEMMA 2. *Let $\alpha \geqslant \frac{1}{2}$,*

$$d_\alpha(x) = \frac{1}{x}\,\#\{n \leqslant x \mid P(n) \geqslant x^\alpha\}. \tag{15}$$

Then, as $x \to \infty$,

$$d_\alpha(x) \to \log \frac{1}{\alpha}. \tag{16}$$

Proof. n is counted on the right-hand side of (15) if it is of the form $n = m \cdot p$, $p \geqslant x^\alpha$, $m \in \mathbf{N}$, $n \leqslant x$. Therefore

$$x d_\alpha(x) = \sum_{x^\alpha \leqslant p \leqslant x} \left[\frac{x}{p}\right] = x \sum_{x^\alpha < p \leqslant x} \frac{1}{p} + O(x/\log x),$$

which implies (16) immediately. ∎

Proof of Theorem 2. First observe that $\log \frac{1}{2} > \frac{2}{3}$. Therefore there exist positive constants ε, δ such that

$$d_{1/2+\varepsilon}(x) > \tfrac{2}{3} + \delta \tag{17}$$

for each large x.

Let

$$\mathscr{D} = \{n \leqslant x \mid P(n+j) > x^{1/2+\varepsilon}, j = 0, 1, 2\}.$$

For an $n \leqslant x$, let $e(n)$ be the number of those integers among $P(n)$, $P(n+1)$, $P(n+2)$ which are greater than $x^{1/2+\varepsilon}$. It is clear that

$$\sum_{n \leqslant x} e(n) = 3x\, d_{1/2+\varepsilon}(x) + O(1),$$

which by (17) implies that

$$\sum_{n \leqslant x} e(n) > (2 + 3\delta)x + O(1). \tag{18}$$

Consequently

$$\# \{n \leqslant x \mid e(n) = 3\} > \delta x + O(1). \tag{19}$$

In other words

$$\# \mathscr{D} > \frac{\delta}{2} x \tag{20}$$

holds for each large x. From here on, the proof is essentially the same as that of Theorem 1; hence, we omit the details. ∎

It seems plausible that the set of integers n with

$$\min_{j=0,\ldots,k} P(n+j) > n^{1/2+\varepsilon}$$

has a positive lower density. This would imply that

$$\sum_{n \leqslant x} \min_{j=0,\ldots,k} f(n+j) \sim \sum_{n \leqslant x} \min_{j=0,\ldots,k} f(P(n+j)).$$

We are unable to prove this even for $k=3$.

4. The Behaviour of $f(n)$ on Sequences of k Consecutive Integers

Theorem 3. *Assume that $f \in \mathscr{F}$. Let $k_1 \leqslant k_2 \leqslant \cdots$ be a sequence of positive integers tending to infinity for $n \to \infty$ and let furthermore $k_{2n}/k_n = O(1)$. Define*

$$V(n,k) = \sum_{j=0}^{k-1} f(n+j), \qquad U(n,k) = \sum_{j=0}^{k-1} f(P(n+j)). \tag{21}$$

Then, there exists a sequence $\mathscr{B}$ of natural numbers having zero density such that, for all $n \in \overline{\mathscr{B}} = \mathbf{N} \backslash \mathscr{B}$, one has

$$V(n,k_n) = (1+o(1))\, U(n,k_n). \tag{22}$$

We shall deduce Theorem 3 from Theorem 4, which is essentially due to J. B. Friedlander [5]. Since we shall need a modification of it and our proof is much simpler, we give its proof.

Theorem 4. *Let $a(n,k) = n(n+1)\cdots(n+k-1)$. Let δ_x be positive tending to zero arbitrarily slowly, K_x be a function taking on positive integer values $K_x \to \infty$ as $x \to \infty$ and $K_x < x$. Let*

$$F(n,k) = F(n,k,x) = \sum_{\substack{p \mid a(n,k) \\ p > x^{1/2+\delta_x}}} \log p. \tag{23}$$

Then

$$|F(n,K_x) - \tfrac{1}{2} K_x \log x| = o(K_x \log x) \tag{24}$$

for all but at most $o(x)$ integers $n \in [x/2, x]$.

Proof. Let $1 < k < \log x$, $\delta = \delta_x$ and consider the set $\mathscr{A}$ of integers $n \in [x/2, x]$. Let $A(n, d)$ be the number of multiples of d in the interval $]n, n+k-1]$, that is

$$A(n, d) = \left[\frac{n+k-1}{d}\right] - \left[\frac{n}{d}\right] = \frac{k}{d} + O(1). \tag{25}$$

Let furthermore Λ be von Mangoldt's function. It is clear that

$$\log a(n, k) = k \log x + O(k) \tag{26}$$

and furthermore that

$$\log a(n, k) = \sum_{d < 2x} \Lambda(d) A(n, d). \tag{27}$$

Let the right side of (27) be split into five parts: $E_1(n) + E_2(n) + E_3(n) + E_4(n) + F(n, k)$. In E_1 we sum over the powers of primes with exponent greater than 1. In E_2, E_3, E_4, F we sum over primes p from the range $p \leqslant k$, $k < p \leqslant \sqrt{x}$, $\sqrt{x} < p \leqslant x^{1/2+\delta}$, $p > x^{1/2+\delta}$, respectively.

Using (25) we obtain successively

$$\Sigma_1 := \sum_{n \in \mathscr{A}} E_1(n) \ll xk \sum_{\substack{p, a \\ a \geqslant 2}} \frac{\log p}{p^a} \ll xk, \tag{28}$$

$$E_2(n) = k \sum_{p \leqslant k} \frac{\log p}{p} + O(k) = k \log k + O(k), \tag{29}$$

and

$$\Sigma_2 := \sum_{n \in \mathscr{A}} E_4(n) = \sum_{\sqrt{x} < p \leqslant x^{1/2+\delta}} \log p \sum_{\substack{p \mid a(n,k) \\ n \in \mathscr{A}}} 1 \tag{30}$$

$$\ll kx \sum_{\sqrt{x} < p \leqslant x^{1/2+\delta}} \frac{\log p}{p} \ll \delta kx \log x.$$

Now we estimate $E_3(n)$ by the Turán–Kubilius inequality. It is clear that $A(n, p) = 0$ or 1 if $p > k$. Let us consider the sums

$$S_1(k) := \sum_{n \in \mathscr{A}} E_3(n), \qquad S_2(k) := \sum_{n \in \mathscr{A}} E_3^2(n).$$

We shall prove that

$$\mathscr{D}(k) := \sum_{n \in \mathscr{A}} (E_3(n) - k\alpha(k))^2 \ll k(\log x)^2 x, \tag{31}$$

where $\alpha(k)=\log(\sqrt{x}/k)$. Let $B(p_1, p_2)$ be the number of those integers $n\in\mathscr{A}$ for which $p_1, p_2 \mid a(n,k)$, $k\leqslant p_i\leqslant\sqrt{x}$. Then

$$B(p_1,p_2)=\begin{cases}\frac{x}{2}\frac{k}{p}+O(k) & \text{if}\quad p_1=p_2=p,\\ \frac{x}{2}\frac{k^2}{p_1p_2}+O(k^2) & \text{if}\quad p_1\neq p_2\end{cases}$$

Furthermore,

$$S_1(k)=\sum_{k<p\leqslant\sqrt{x}}(\log p)\sum_{\substack{n\in\mathscr{A}\\ p\mid a(n,k)}}1=\frac{xk}{2}\alpha(k)+O(k),$$

and

$$\begin{aligned}S_2(k)&=\sum_{p_i\in(k,\sqrt{x}]}(\log p_1)(\log p_2)B(p_1,p_2)\\ &=\frac{kx}{2}\sum_{k<p\leqslant\sqrt{x}}\frac{(\log p)^2}{p}\\ &\quad+\frac{xk^2}{2}\sum_{p_1\neq p_2}\frac{(\log p_1)(\log p_2)}{p_1p_2}+O(k^2x/(\log x)^2). \qquad (32)\end{aligned}$$

Since the last sum on the right-hand side of (32) equals $\alpha^2(k)+O(1)$, we have

$$S_2(k)=(\alpha(k)k)^2\frac{x}{2}+\frac{kx}{2}\sum\frac{(\log p)^2}{p}+O(k^2x).$$

On the other hand,

$$\mathscr{D}(k)=S_2(k)-2k\alpha(k)S_1(k)+(k\alpha(k))^2\frac{x}{2}+O(k\alpha(k)),$$

thus implying (31).

Since $F(n,k)=\log a(n,k)-E_1(n)-E_2(n)-E_3(n)-E_4(n)$, we obtain by (26),

$$\begin{aligned}&\left|F(n,k)-\frac{k}{2}\log x\right|\\ &\qquad\leqslant E_1(n)+E_4(n)+\left|E_2(n)+E_3(n)-\frac{k}{2}\log x\right|+O(k). \qquad (33)\end{aligned}$$

From (29),

$$\left| E_2(n) + E_3(n) - \frac{k}{2} \log x \right| \leqslant |E_3(n) - k\alpha(k)| + O(k). \tag{34}$$

Using the Cauchy–Schwarz inequality and (31), we have

$$\sum_{n \in \mathscr{A}} |E_3(n) - k\alpha(k)| \ll x \sqrt{k} \log x, \tag{35}$$

and so by (28) and (30) we have

$$\sum_{n \in A} \left| F(n, k) - \frac{k}{2} \log x \right| \ll x \sqrt{k} \log x + \delta k x \log x + xk. \tag{36}$$

We now apply inequality (36) with $k = [\log K_x]$, say. Since $F(n, k) = O(k \log x)$, then

$$F(n, K_x) = \sum_{j=0}^{H-1} F(n + jk, k) + O(k \log x), \qquad H = \left[\frac{Kx}{k} \right],$$

and so from (36) we obtain that

$$\begin{aligned} &\sum_{n \in A} \left| F(n, K_x) - \frac{K_x}{2} \log x \right| \\ &\quad \ll \frac{K_x}{k} \sum_{x/2 < n < 2x} \left| F(n, k) - \frac{k}{2} \log x \right| + O(kx \log x) \\ &\quad \ll \delta K_x x \log x + \frac{x K_x \log x}{\sqrt{\log K_x}}. \end{aligned} \tag{37}$$

From (37), putting $\delta = \delta_x \to 0$, we infer

$$\sum_{n \in \mathscr{A}} \left| F(n, K_x) - \frac{K_x}{2} \log x \right| = o(x K_x \log x) \qquad (x \to \infty) \tag{38}$$

which implies (24) immediately. ∎

Proof of Theorem 3. Let us once more consider the integers

$$n \in \mathscr{A} = \{ n \mid n \in [x/2, x] \}.$$

We shall prove the theorem assuming that Case #2 holds. Case #1 can be considered similarly. Let

$$K_x = k_{[x/2]}. \tag{39}$$

Then by (24), for all but $o(x)$ integers n in $\mathscr{A}$,

$$F(n, K_x) > \tfrac{1}{3}K_x \log x,$$

and because of (39) we have

$$F(n, k_n) > ck_n \log x,$$

for some positive constant c. For a non-exceptional n there exist at least ck_n integers in $0 \leqslant j < k_n$ such that $P(n+j) > x^{1/2+\delta_x}$. Consequently

$$U(n, k_n) \gg k_n L(x^{1/2+\delta_x}). \tag{40}$$

Furthermore, for each $m \in [x/2, x]$, $P(m/P(m)) < x^{1/2}$, and so, by Lemma 1

$$f(m) - f(P(m)) \ll L(x^{1/2}),$$

and thus

$$V(n, k_n) - U(n, k_n) \ll k_n L(x^{1/2}). \tag{41}$$

If δ_x is chosen so that $L(x^{1/2}) = o(1) L(x^{1/2+\delta_x})$, then (40) and (41) imply that

$$V(n, k_n) = (1 + o(1))\, U(n, k_n)$$

for all non-exceptional n. ∎

5. On the Quotient $f(n)/f(P(n))$

In their paper [2], Alladi and Erdös proved that

$$\sum_{n \leqslant x} \frac{\beta(n)}{P(n)} \sim x, \tag{42}$$

where $\beta(n) = \sum_{p \mid n} p$, which by $\beta(n) \geqslant P(n)$ implies that $\beta(n)/P(n) \to 1$ on a set of integers having density 1. Here we consider the expression

$$Q(x) = Q_f(x) \overset{\text{def}}{=} \frac{1}{x} \sum_{2 \leqslant n \leqslant x} \frac{f(n)}{f(P(n))}$$

and prove:

Theorem 5. *Let $f \in \mathscr{F}$, then $Q(x) \to 1$ as $x \to \infty$.*

Proof. We concentrate our attention on Case #2 (that is, when $\rho = 0$ and $\lambda_L(x) \to +\infty$). The treatment of Case #1 is somewhat simpler and runs on similar lines. Let $\Psi(x, y)$ be the number of integers n up to x for which $P(n) \leqslant y$. We shall use the known inequality

$$\Psi(x, y) \leqslant c_1 x \exp\left(-c \frac{\log x}{\log y}\right), \tag{43}$$

valid uniformly in $x, y \geqslant 2$ with absolute positive constants c, c_1.

Starting from the formula

$$\log \frac{L(w_2)}{L(w_1)} = \int_{w_1}^{w_2} \frac{\lambda(u)}{u \log u}\, du, \qquad w_1 < w_2,$$

and by using the monotonity of λ, we obtain

$$\lambda(w_1) \log \frac{\log w_2}{\log w_1} \leqslant \log \frac{L(w_2)}{L(w_1)} \leqslant \lambda(w_2) \log \frac{\log w_2}{\log w_1},$$

that is

$$\left(\frac{\log w_2}{\log w_1}\right)^{\lambda(w_1)} \leqslant \frac{L(w_2)}{L(w_1)} \leqslant \left(\frac{\log w_2}{\log w_1}\right)^{\lambda(w_2)}. \tag{44}$$

In the sequel p and q run over the set of primes. It is clear that

$$\sum_{2 \leqslant n \leqslant x} \frac{f(n)}{f(P(n))} = x + \sum_{q < p} \frac{f(q)}{f(p)} \Psi\left(\frac{x}{pq}, p\right) + O(1). \tag{45}$$

Let $\sum$ denote the sum $\sum_{q<p}$ on the right-hand side. We need to prove that $\sum = o(x)$.

From elementary estimations for the distribution of primes we obtain immediately that

$$\sum_{p > H} \frac{(\log p)^{-s}}{p} \ll \frac{1}{s(\log H)^s} \tag{46}$$

and furthermore that

$$\sum_{q < H} \frac{(\log q)^s}{q} \ll \frac{(\log H)^s}{s}, \tag{47}$$

uniformly in $s \geqslant 1$, $H \geqslant 2$.

From (44) it is clear that

$$\frac{f(q)}{f(p)} \Psi\left(\frac{x}{pq}, q\right) \leqslant \left(\frac{\log q}{\log p}\right)^{\lambda(q)} \frac{x}{pq}. \tag{48}$$

Let ε be an arbitrary small positive constant and consider the contribution of couples p, q with $q > x^{\varepsilon}$ to the sum $\sum$. By (46) we have

$$\sum_{q > x^{\varepsilon}} \frac{(\log q)^{\lambda(q)}}{q} \sum_{p > q} \frac{1}{p(\log p)^{\lambda(q)}} \ll \sum_{q > x^{\varepsilon}} \frac{1}{q\lambda(q)},$$

and the sum on the right-hand side tends to zero, since $\lambda(q) \geqslant \lambda(x^{\varepsilon})$, $\lambda(x^{\varepsilon}) \to \infty$.

Let now z be in the interval $2 \leqslant z \leqslant x$ and let

$$S(z) := \sum \frac{f(q)}{pqf(p)},$$

where the summation is extended over those couples p, q for which $q \leqslant z < p$ and $p \leqslant x$. Since the terms on the right-hand side can be bounded by

$$\frac{L(q)}{L(z)} \frac{L(z)}{L(p)} \frac{1}{pq},$$

we have

$$S(z) \ll \left(\sum_{q \leqslant z} \frac{L(q)}{qL(z)}\right)\left(\sum_{p \geqslant z} \frac{L(z)}{pL(p)}\right).$$

By (44) and (46) we obtain that the second sum is less than $1/\lambda(z)$. Since $\lambda(q) \geqslant 1$ for each large q, the first sum is less than

$$O(1) + \sum_{q < z} \frac{1}{q}\left(\frac{\log q}{\log z}\right) = O(1).$$

We have

$$S(z) \ll \frac{1}{\lambda(z)},$$

uniformly in z. Hence we have

$$\sum_{q \leqslant z < p} \frac{f(q)}{f(p)} \frac{x}{pq} \ll \frac{x}{\lambda(z)}. \tag{49}$$

Let us now choose first $z = x^{\varepsilon}$. We have

$$\sum = \sum_1 + o(x),$$

where in $\sum_1$ we sum over q, p with $q < p \le x^{\varepsilon}$. We split $\sum_1$ into two parts, $\sum_1 = \sum_2 + \sum_3$ according to $q \leqslant \sqrt{p}$, or $\sqrt{p} < q < p$, respectively.

Let

$$\mathscr{F}_s := [x^{\varepsilon/2^{s+1}}, x^{\varepsilon/2^s}] \qquad (s = 0, 1, 2, ...).$$

To estimate $\sum_2$ and $\sum_3$ we shall use inequality (43). For $q < x^{\varepsilon/2^{s+1}}$, $p \in \mathscr{F}_s$ we have

$$\begin{aligned} \Psi\left(\frac{x}{pq}, p\right) &\ll \frac{x}{pq} \exp\left(-c \frac{\log x}{\log p}\right) \\ &\ll \frac{x}{pq} \exp\left(-c\left(\frac{\log x}{(\varepsilon/2^s)\log x}\right)\right) \\ &= \frac{x}{pq} \exp\left(-c \frac{2^s}{\varepsilon}\right). \end{aligned} \tag{50}$$

Then

$$\sum_2 \leqslant \sum_{s=0}^{\infty} \sum^{(s)},$$

where

$$\sum^{(s)} = \sum_{\substack{q < x^{\varepsilon/2^{s+1}} \\ p \in \mathscr{F}_s}} \frac{f(q)}{f(p)} \Psi\left(\frac{x}{pq}, p\right).$$

By (49) and (50) we have

$$\sum^{(s)} \ll \frac{x}{\lambda(x^{\varepsilon/2^{s+1}})} \exp\left(-c \frac{2^s}{\varepsilon}\right).$$

So

$$\sum_2 \ll x \sum_{s=0}^{\infty} \frac{\exp(-c(2^s/\varepsilon))}{\lambda(x^{\varepsilon/2^{s+1}})}. \tag{51}$$

Observing that the sum on the right-hand side of (51) tends to zero as $x \to \infty$, we have

$$\sum_2 = o(x).$$

To estimate $\sum_3$ we observe that $\log q \geqslant \frac{1}{2}\log p$; consequently,

$$\Psi\left(\frac{x}{pq}, p\right) \ll \frac{x}{pq} \exp\left(-c' \frac{\log x}{\log q}\right),$$

with $c' = c/2$, say,

$$\sum_3 \ll x \sum_{\substack{q,\, p \\ q < p < x^{\varepsilon}}} \left(\frac{\log q}{\log p}\right)^{\lambda(q)} \times \frac{1}{pq} \exp\left(-c' \frac{\log x}{\log q}\right) = x \sum_A, \tag{52}$$

where

$$\sum_A = \sum_q \frac{(\log q)^{\lambda(q)}}{q} \exp\left(-c' \frac{\log x}{\log q}\right) \cdot \sum_q, \tag{53}$$

with

$$\sum_q = \sum_{q < p < x^{\varepsilon}} \frac{1}{p(\log p)^{\lambda(q)}}.$$

It is clear that

$$\sum_q \ll \int_q^{x^{\varepsilon}} \frac{1}{u(\log u)^{\lambda(q)+1}}\, du \ll (\log q)^{-\lambda(q)}. \tag{54}$$

Thus, using (54) in (53), we obtain

$$\begin{aligned} \sum_A &\ll \sum_{q < x^{\varepsilon}} q^{-1} \exp(-c' \log x/\log q) \\ &\ll \int_2^{x^{\varepsilon}} \frac{du}{u \log u \exp(c' \log x/\log u)} \\ &= \int_{\log 2}^{\varepsilon \log x} \frac{dv}{v \exp(c' \log x/v)}. \end{aligned}$$

Since the function $v \exp(c' \log x/v)$ is easily seen to be decreasing on the interval $[\log 2, \varepsilon \log x]$, it follows that this last integral is no larger than

$$(\varepsilon \log x - \log 2) \cdot \frac{1}{\varepsilon \log x \exp(c' \log x/\varepsilon \log x)} \ll \frac{1}{\exp(c'/\varepsilon)},$$

which tends to zero as ε tends to zero. Thus $\sum_4 = o(1)$. Combining this with (52), we obtain $\sum_3 = o(x)$. This settles Case #2 and hence completes the proof of the theorem. ∎

Remark. It is interesting to consider the expression $Q_f(x)$ for other strongly additive functions f which do not belong to $\mathscr{F}$, namely, those f for which $f(p) = L(p)$, where L is increasing, slowly oscillating and for which the corresponding function $\lambda(x) = \lambda_L(x) = x \log x L'(x)/L(x)$ satisfies $\lim_{x \to \infty} \lambda(x) = 0$. In this case, we show that

$$Q(x) = (1 + o(1)) \frac{L_1(x)}{L(x)}, \tag{55}$$

where

$$L_1(x) \stackrel{\text{def}}{=} \int_2^x \frac{L(t)}{t \log t} \, dt.$$

The proof goes as follows. Let

$$A(x) = \sum_{p \leqslant x} \frac{L(p)}{p}, \qquad L_1(x) = \int_2^x \frac{L(u)}{u \log u} \, du,$$

$$\log L(x) = \int_2^x \frac{\lambda(u)}{u \log u} \, du, \quad \lambda(u) > 0,\ \lambda(u) \to 0, \quad \Delta(x) = \frac{1}{L_1^2(x)} \int_2^x \frac{L^2(u)}{u \log u} \, du.$$

Then

$$\Delta(x) \leqslant \frac{L(x)}{L_1^2(x)} \int_2^x \frac{L(u)}{u \log u} \, du = \frac{L(x)}{L_1(x)},$$

since L is monotonic.

We shall prove that $L(x)/L_1(x) \to 0$, i.e., $\Delta(x) \to 0$. Let $\delta > 0$ be an arbitrary constant. Then for large x, $\lambda(u) < \varepsilon$ if $x^\delta \leqslant u \leqslant x$, and so

$$\log \frac{L(x)}{L(x^\delta)} \leqslant \varepsilon \int_{x^\delta}^x \frac{1}{u \log u} \, du = \varepsilon \log \frac{\log x}{\delta \log x} = \varepsilon \log(1/\delta).$$

Then we have $L(x)/L(x^\delta) \leqslant (1/\delta)^\varepsilon$ which is equivalent to $L(x^\delta) \geqslant \delta^\varepsilon L(x)$. Hence

$$L_1(x) \geqslant \int_{x^\delta}^x \frac{L(x^\delta)}{u \log u} \, du \geqslant \delta^\varepsilon L(x) \int_{x^\delta}^x \frac{1}{u \log u} \, du = \delta^\varepsilon L(x) \log(1/\delta),$$

and so

$$\frac{L(x)}{L_1(x)} \leqslant \left(\frac{1}{\delta}\right)^\varepsilon \frac{1}{\log(1/\delta)}.$$

Since $\varepsilon > 0$ can be chosen arbitrarily small if $x > x_0(\varepsilon)$, we have

$$\limsup_{x \to \infty} \Delta(x) \leqslant \frac{1}{\log(1/\delta)}.$$

Since $\delta > 0$ is arbitrary, letting $\delta(x) \to 0$, we have $\Delta(x) \to 0$.

From the Turán–Kubilius inequality,

$$\sum_{n \leqslant x} (f(n) - A(x))^2 \ll x \sum_{p \leqslant x} \frac{L^2(p)}{p},$$

and since

$$\sum_{p \leqslant x} \frac{L^2(p)}{p} \sim \int_2^x \frac{L^2(u)}{u \log u}\, du, \qquad A(x) \sim \sum_{p \leqslant x} \frac{L(p)}{p} \sim L_1(x),$$

$$\sum_{n \leqslant x} \frac{1}{L(P(n))} \sim \frac{x}{L(x)}, \qquad \sum_{p \leqslant x} \frac{1}{L^2(P(n))} \sim \frac{x}{L^2(x)},$$

then

$$\frac{1}{x} \sum_{n \leqslant x} \left| \frac{f(n) - A(x)}{f(P(n))} \right| \leqslant \left(\frac{1}{x} \sum_{n \leqslant x} |f(n) - A(x)|^2 \right)^{1/2} \times \left(\frac{1}{x} \sum_{n \leqslant x} \frac{1}{f^2(P(n))} \right)^{1/2} = \sqrt{\xi}\, \sqrt{\eta}.$$

Then $\eta \sim 1/L^2(x)$, $\xi \asymp \int_2^x (L^2(u)/u \log u)\, du$ and consequently

$$\frac{1}{x} \sum_{n \leqslant x} \left| \frac{f(n) - A(x)}{f(P(n))} \right| \ll \left(\frac{1}{L^2(x)} \int_2^x \frac{L^2(u)}{u \log u}\, du \right)^{1/2} \leqslant \left(\frac{L_1(x)}{L(x)} \right)^{1/2}.$$

So we proved that

$$\frac{1}{x} \sum_{n \leqslant x} \left| \frac{f(n) - A(x)}{f(P(n))} \right| = o(1) \frac{L_1(x)}{L(x)}.$$

Hence we get that

$$\sum_{n \leqslant x} \frac{f(n)}{f(P(n))} - A(x) \sum_{n \leqslant x} \frac{1}{f(P(n))} = o(x) \frac{L_1(x)}{L(x)}.$$

Since $A(x) \sim L_1(x)$ and $\sum_{n \leqslant x} 1/f(P(n)) \sim x/L(x)$, then

$$\sum_{n \leqslant x} \frac{f(n)}{f(P(n))} = (1 + o(1)) \frac{L_1(x)}{L(x)}.$$

This proves (55).

References

1. K. Alladi and P. Erdős, On an additive arithmetic function, *Pacific J. Math.* **71** (1977), 275–294.
2. K. Alladi and P. Erdős, On the asymptotic behaviour of large prime factors of integers, *Pacific J. Math.* **82** (1979), 295–315.
3. J. M. DeKoninck and A. Ivić, The distribution of the average prime divisor of an integer, *Arch. Math.* **43** (1984), 37–43.
4. J. M. DeKoninck and A. Mercier, Les fonctions arithmétiques et le plus grand facteur premier, *Acta Arith.*, in press.
5. J. B. Friedlander, Large prime factors in small intervals, *in* "Topics in Classical Number Theory," Vol. 34, pp. 511–518, Colloquia Math. Soc. J. Bolyai, North-Holland, Amsterdam, 1984.
6. C. Hooley, "Applications of Sieve Methods to the Theory of Numbers, Cambridge Univ. Press, London/New York.