111. On Certain Averages of $\omega(n)$

By Leo MURATA
Department of Mathematics, Faculty of Science,
Tokyo Metropolitan University

(Communicated by Shokichi IYANAGA, M. J. A., Dec. 12, 1980)

Let $\omega(n)$ be the number of distinct prime factors of the natural number n, and $\Omega(n)$ be the total number of prime factors of n. We observe their averages for $M \subseteq N$ and x:

$$V(M, x)=\frac{\sum_{n\in M(x)}\omega(n)}{|M(x)|},\qquad W(M, x)=\frac{\sum_{n\in M(x)}\{\Omega(n)-\omega(n)\}}{|M(x)|},$$

where $M(x)=\{n;\ n\in M,\ n\leqslant x\}$ and $|\ |$ designates the cardinal. For $M=N$, which can be regarded as the "standard" case, it is well-known that (cf. [1, Theorem 430]):

$$V(N, x)=\log\log x+A+O\Big(\frac{1}{\log x}\Big),$$

$$W(N, x)=\sum_p \frac{1}{p(p-1)}+O(x^{-1/2}),$$

where $\sum_p$ denotes the sum over all primes, $A=\gamma+\sum_p\{\log(1-1/p)+1/p\}$, and γ is Euler's constant. A few results are known as to the value of $V(M, x)$ or that of $W(M, x)$ for specially chosen M. For example, H. Halberstam ([2]) proved that if $f(x)$ is an irreducible polynomial with integral coefficients and

$$M^*=\{f(p);\ p: \text{rational prime}\},$$

then

$$V(M^*, x)\sim\log\log x,$$

but no estimate is obtained for error terms for this M^*.

Our aim is to observe $V(M, x)$ and $W(M, x)$ for other types of M. First we take up the case

$$N_d=\{n;\ 1\leqslant\|n\|\leqslant d\},$$

where d is a fixed positive integer and $\|n\|=\min_{p:\text{ prime}}(|n-p|)$, the distance from n to its nearest prime.

Theorem 1.

$$(1)\qquad V(N_d, x)=\log\log x+\Big\{A+\sum_p\frac{1}{p(p-1)}-\log 2+\beta_d(x)\Big\}+d^2\cdot D_d\cdot O\Big(\frac{\log\log x}{\log x}\Big),$$

where D_d is a computable constant depending only on d, given more precisely later on, and $\beta_d(x)$ is a function satisfying

$$\frac{1}{2}+O\Big(\frac{\log\log x}{\log x}\Big)\leqslant\beta_d(x)\leqslant 1+O\Big(\frac{\log\log x}{\log x}\Big),$$

where constants implied by O-symbols are absolute.

Through a numerical calculation we obtain

$$V(N_d,\ x)-V(N,\ x)\geqslant\sum_p\frac{1}{p(p-1)}-\log 2+\frac{1}{2}+O\Big(\frac{\log\log x}{\log x}\Big)>0.5799,$$

for sufficiently large x. Thus we can say that, if we restrict the domain of average to those composite integers in d-neighbourhoods of prime numbers, the corresponding average of $\omega(n)$ will be definitely larger than the "standard" average.

Concerning the function $W(M,\ x)$, we obtain

Theorem 2.

$$W(N_d,\ x)=\sum_p\frac{1}{(p-1)^2}+d^2\cdot D_d\cdot O\Big(\frac{1}{\log x}\Big),$$

where D_d is same to that of Theorem 1 and the constant implied by O-symbol is absolute.

This shows

$$W(N_d,\ x)-W(N,\ x)\sim\sum_p\frac{1}{p(p-1)^2}>0.6019.$$

We can generalize these theorems through regarding the d to be an increasing function $f(x)$ of x:

Theorem 3. *Let ε be an arbitrary positive constant, R be the minimum natural integer satisfying $\varepsilon>1/R$, $f(x)$ be a positive valued increasing function such that $f(x)=O((\log x)^{1-\varepsilon})$ as $x\to\infty$, and $M_f(x)$ be a set*

$$M_f(x)=\{n;\ 1\leqslant\|n\|\leqslant f(x)\}.$$

Then

$$V(M_f(x),\ x)=\log\log x+\Big\{A+\sum_p\frac{1}{p(p-1)}-\log 2+\gamma_f(x)\Big\}$$
$$+O((\log x)^{1-R\varepsilon}(\log\log x)(\log\log\log x)^{R-1}),$$

where $\gamma_f(x)$ is a function satisfying

$$\frac{1}{2}+O\Big(\frac{\log\log x}{\log x}\Big)\leqslant\gamma_f(x)\leqslant 1+O\Big(\frac{\log\log x}{\log x}\Big),$$

and the constants implied by O-symbols depend at most on ε.

Theorem 4. *Under the same assumptions of Theorem 1, we have*

$$W(M_f(x),\ x)=\sum_p\frac{1}{(p-1)^2}+O((\log x)^{1-R\varepsilon}(\log\log\log x)^{R-1}),$$

where the constant implied by O-symbol depends only on ε.

In this paper we give a sketch of the proof of Theorem 1 only.

Let $\boldsymbol{P}$ be the set of all primes, d be a fixed positive integer, and we define for integar i, j or b:

$P_i(x)=\{n;\ n=p+i,\ p\in\boldsymbol{P},\ n\leqslant x\}$, i.e. a sequence of shifted primes,

$I_j(x)=\{n;\ n$ is contained in *just* j sequences among $P_{-d}(x),\ \cdots,$ $P_{-1}(x),\ P_1(x),\ \cdots,\ P_d(x)\}$,

$$Q_j(x)=I_j(x)\cap \boldsymbol{P},$$
$$\boldsymbol{P}_b(x)=\{p\,;\,p\in\boldsymbol{P},\ p+b\in\boldsymbol{P},\ p\leq x\}.$$

Then we have

Lemma 1.

$$\sum_{n\in P_i(x)}\omega(n)=\Big\{\log\log x+A+\sum_p\frac{1}{p(p-1)}-\log 2+\alpha_i(x)\Big\}\cdot \mathrm{Li}(x),$$

where the function $\alpha_i(x)$ *satisfies*

$$\frac{1}{2}+O\Big(\frac{\log\log x}{\log x}\Big)\leqslant\alpha_i(x)\leqslant 1+O\Big(\frac{\log\log x}{\log x}\Big),$$

and the constants implied by O-symbols are absolute.

Lemma 2.

$$\sum_{\substack{q\leqslant x\\ q\in \boldsymbol{P}}}\sum_{\substack{p\in \boldsymbol{P}_b(x)\\ p\equiv a(q)}}1=D_d\cdot O\Big(\frac{x\log\log x}{\log^2 x}\Big)$$

where

$$D_d=\max_{\substack{2d\geqslant b\geqslant 1\\ b:\ \mathrm{even}}}\Big\{\prod_{p\mid b}\Big(1-\frac{1}{p}\Big)^{-1}\Big\},$$

and the constant implied by O-symbol is absolute.

In order to prove Lemma 1, we have to utilize Bombieri's theorem ([3]), Brun-Titchmarsh's theorem ([5, Theorem 3.8]), and M. Goldfeld's result ([4]). Lemma 2 can be deduced from the following two estimates ([5, Corollary 2.4.1] and [5, Theorem 3.11]):

$$(2)\qquad \sum_{\substack{p\in \boldsymbol{P}_b(x)\\ p\equiv l\,(\mathrm{mod}\,k)}}1=\Big(\prod_{p\mid kb}\frac{p}{p-1}\Big)\cdot O\Big(\frac{x}{\varphi(k)\log^2(x/k)}\Big),$$

$$(3)\qquad |\boldsymbol{P}_b(x)|=8\Big(\prod_{p>2}\frac{p(p-2)}{(p-1)^2}\Big)\Big(\prod_{\substack{p\mid b\\ p\neq 2}}\frac{p-1}{p-2}\Big)\cdot O\Big(\frac{x}{\log^2 x}\Big).$$

Now return to the proof of Theorem 1. It is easy to see that

$$(4)\qquad |N_d(x)|=\sum_{|i|=1}^{d}|P_i(x)|-\sum_{j=2}^{2d}(j-1)|I_j(x)|-\sum_{j=1}^{2d}|Q_i(x)|,$$

$$(5)\qquad \sum_{n\in N_d(x)}\omega(n)=\sum_{|i|=1}^{d}\sum_{n\in P_i(x)}\omega(n)-\sum_{j=2}^{2d}(j-1)\sum_{n\in I_j(x)}\omega(n)-\sum_{j=1}^{2d}|Q_j(x)|.$$

Concerning (4), we have obviously,

$$\sum_{|i|=1}^{d}|P_i(x)|=2d\{\pi(x)+O(1)\},$$

and from the definitions of $I_j(x)$ and $Q_j(x)$, we obtain by (2) and (3) that

$$\sum_{j=2}^{2d}(j-1)|I_j(x)|=2d\cdot d^2\cdot D_d\cdot O\Big(\frac{x}{\log^2 x}\Big),$$
$$\sum_{j=1}^{2d}|Q_j(x)|=2d\cdot D_d\cdot O\Big(\frac{x}{\log^2 x}\Big).$$

Lemma 1 gives an estimate of the first term of (5):

$$\sum_{|i|=1}^{d}\sum_{n\in P_i(x)}\omega(n)=2d\Big\{\log\log x+A+\sum_p\frac{1}{p(p-1)}-\log 2+\sum_{|i|=1}^{d}\alpha_i(x)\Big\}\cdot\mathrm{Li}(x).$$

Furthermore,

$$\sum_{j=2}^{2d}(j-1)\sum_{n\in I_j(x)}\omega(n)\leqslant 2d\sum_{j=2}^{2d}\sum_{n\in I_j(x)}\omega(n)=2d\sum_{b,c}\sum_{n\in P_b(x)\cap P_c(n)}\omega(n),$$

where $\sum_{b,c}$ denotes the sum over all such pairings of integers b and c that satisfy $-d\leqslant b<c\leqslant d$, $b\cdot c\neq 0$. Since

$$\sum_{n\in P_b(x)\cap P_c(x)}\omega(n)=\sum_{\substack{q\leqslant x\\ q\in P}}\sum_{\substack{p\in P_{c-b}(x)\\ p\equiv -b\,(q)}}1+O(1),$$

we obtain by Lemma 2 that

$$\sum_{j=2}^{2d}(j-1)\sum_{n\in I_j(x)}\omega(n)=2d\cdot d^2\cdot D_d\cdot O\Big(\frac{x\log\log x}{\log^2 x}\Big).$$

Consequently, from (4) and (5), we obtain

$$|N_d(x)|=2d\Big\{1+d^2\cdot D_d\cdot O\Big(\frac{1}{\log x}\Big)\Big\}\frac{x}{\log x},$$

and

$$\sum_{n\in N_d(x)}\omega(n)=2d\Big\{\log\log x+A+\sum_p\frac{1}{p(p-1)}-\log 2+\beta_d(x)$$
$$+d^2\cdot D_d\cdot O\Big(\frac{\log\log x}{\log x}\Big)\Big\}\frac{x}{\log x},$$

where

$$\beta_d(x)=\frac{1}{2d}\sum_{|i|=1}^{d}\alpha_i(x).$$

Theorem 1 can be proved immediately.

The proofs of Theorems 2–4 are accomplished in a similar way as the above, but for Theorems 3 and 4, besides (2) and (3), more precise estimates are required; let g be a positive integer, $b_i(1\leqslant i\leqslant g)$ be integers satisfying $1\leqslant b_1<b_2\cdots<b_g\leqslant 2d$, and $\boldsymbol{P}_{b_1,\dots,b_g}(x)=\{p\,;\,p\leqslant x,$ $p\in\boldsymbol{P},\ p+b_i\in\boldsymbol{P},\ i=1,2,\cdots,g\}$, then we have

$$|\boldsymbol{P}_{b_1,\dots,b_g}(x)|=D_d^g\cdot O\Big(\frac{x}{\log^{g+1}x}\Big),$$

$$\sum_{\substack{p\in \boldsymbol{P}b_1,\dots,b_g(x)\\ p\equiv l\,(\mathrm{mod}\,k)}}1=D_d^g\Big(\frac{k}{\varphi(k)}\Big)^g\Big(\prod_{p|k}\frac{p}{p-1}\Big)\cdot O\Big(\frac{x/k}{\log^{g+1}(x/k)}\Big),$$

where D_d is as above and constants implied by O-symbols are absolute (both of these two are deduced from [5, Theorem 2.4]).

Our Theorem 1 gives only a range of values of $\beta_d(x)$. More precise estimate of $\beta_d(x)$ would be obtained if an asymptotic formula of the following form could be proved:

$$|S_d(x)|\sim C_d\pi(x),$$

where $S_d(x)=\{p\,;\,p\in\boldsymbol{P},\ p\leqslant x,\ p+d$ has a prime factor greater than $\sqrt{x}\}$, and C_d is a constant depending only on d. We have a conjecture that

$$|S_d(x)|\sim(\log 2)\pi(x),$$

which would imply

$$V(N_d,x)-V(N,x)\sim W(N,x).$$

Remark. I should like to notice that the following asymptotic formula is easy to prove:

$$|T_d(x)| \sim (\log 2)x,$$

where $T_d(x)=\{n;\, n\leqslant x,\, n+d$ has a prime factor greater than $\sqrt{x}\}$.

I am thankful to Prof. M. Tanaka, who gave me an advice on an original version of this paper owing to which I could improve my results.

References

[1] G. H. Hardy and E. M. Wright: The Theory of Numbers. Forth edition, Oxford (1959).
[2] H. Halberstam: On the distribution of additive number-theoretic functions III. J. London Math. Soc., **31**, 15–27 (1965).
[3] E. Bombieri: On the large sieve. Mathematika, **12**, 201–225 (1965).
[4] M. Goldfeld: On the number of primes p for which $p+a$ has a large prime factor. Ibid., **16**, 23–27 (1969).
[5] H. Halberstam and H. E. Richert: Sieve Methods. Academic Press (1974).