ON THE NUMBER OF RESTRICTED PRIME FACTORS OF AN INTEGER. III

by Karl K. NORTON

§1. INTRODUCTION

Let P be the set of all (positive rational) prime numbers, and let E be an arbitrary nonempty subset of P. Throughout this paper, let p denote a general member of P, and for non-negative integers a, write $p^a \parallel n$ if $p^a \mid n$ and $p^{a+1} \nmid n$. For each positive integer n, define

$$\omega(n; E) = \sum_{p \mid n,\ p \in E} 1, \qquad \Omega(n; E) = \sum_{p^a \parallel n,\ p \in E} a .$$

We usually write $\omega(n; P) = \omega(n), \Omega(n; P) = \Omega(n)$. In this paper, we shall estimate the functions

$$\begin{aligned} S(x, y; E, \omega) &= \operatorname{card}\{n \leqslant x : \omega(n; E) > y\}, \\ S(x, y; E, \Omega) &= \operatorname{card}\{n \leqslant x : \Omega(n; E) > y\} \end{aligned} \tag{1.1}$$

when y is appreciably larger than the normal order of $\omega(n; E)$ and $\Omega(n; E)$; y may even be as large as the maximum order of $\omega(n; E)$ or $\Omega(n; E)$, respectively. (Here and throughout, card B means the number of members of the set B, and if $Q(n)$ is a statement about the integer n, we often write $\{n \leqslant x : Q(n)\}$ instead of $\{n : 1 \leqslant n \leqslant x \text{ and } Q(n)\}$.)

Define

$$E(x) = \sum_{p \leqslant x,\ p \in E} p^{-1} \qquad (x \text{ real}) . \tag{1.2}$$

In [13], it was observed that if $E(x) \to +\infty$ as $x \to +\infty$, then both the average order and the normal order of $\omega(n; E)$ are equal to $E(n)$, and the same statement holds for $\Omega(n; E)$. In [13], we obtained sharp inequalities for the functions (1.1) when $0 < y < 2E(x)$, roughly. In [14], we gave asymptotic formulas for the same functions when $E(x) \to +\infty$ and $y = E(x) + o(E(x))$ as $x \to +\infty$. It is well-known, however, that

$$E(x) \leqslant \log \log x + O(1) \qquad \text{for} \quad x \geqslant 2 ,$$

1980 Mathematics Subject Classification. Primary 10H15, 10H25. Secondary 10A20, 10A21.

whereas if x is large, $\omega(n; E)$ and $\Omega(n; E)$ may be much larger than log log x for some values of $n \leqslant x$. For example, the method of [6, pp. 262-263, 359] shows that

$$\limsup_{n \to +\infty} \frac{\omega(n) \log \log n}{\log n} = 1, \tag{1.3}$$

and a more precise version of (1.3) was obtained in [12, pp. 96-100]. (See also the remarks at the beginning of §3 below.) Before stating estimates for the functions (1.1) when y is large, it seems worthwhile to generalize results like (1.3) to $\omega(n; E)$. First define

$$\pi(x; E) = \sum_{p \leqslant x,\, p \in E} 1 \qquad (x \text{ real}), \tag{1.4}$$

and write

$$\log_2 x = \log \log x, \qquad \log_r x = \log(\log_{r-1} x) \quad \text{for} \quad r = 3, 4, \ldots . \tag{1.5}$$

THEOREM 1.6. *Suppose that there exists a real number* $\gamma(E) > 0$ *such that*

$$\pi(x; E) = \gamma(E)(x/\log x)\{1 + O_E(1/\log x)\} \quad \text{for all} \quad x \geqslant 2. \tag{1.7}$$

Then for each $n \geqslant 3$, *we have*

$$\omega(n; E) \leqslant \frac{\log n}{\log_2 n} + \frac{\{1 + \log \gamma(E)\} \log n}{(\log_2 n)^2} + O_E\left(\frac{\log n}{(\log_2 n)^3}\right), \tag{1.8}$$

with equality for infinitely many n.

Here and throughout, the notation $O_{\delta, \varepsilon, \ldots}$ implies a constant depending at most on $\delta, \varepsilon, \ldots$, while O without subscripts implies an *absolute* constant. Likewise, for $i = 1, 2, \ldots$, we shall write $c_i(\delta, \varepsilon, \ldots)$ for a positive number depending at most on $\delta, \varepsilon, \ldots$, while c_i will mean a positive absolute constant.

It is interesting to observe that a much weaker hypothesis than (1.7) still implies that the maximum order of $\omega(n; E)$ is approximately $(\log n)(\log_2 n)^{-1}$. See the remarks after the proof of Theorem 1.6 in §3.

After (1.3) and Theorem 1.6, it is natural to ask how often $\omega(n; E)$ and $\Omega(n; E)$ assume values appreciably larger than their normal order $E(n)$. It appears that rather little was known about this problem until very recently. The earliest contribution was by Hardy and Ramanujan [5] (reprinted in [15, pp. 262-275]), whose estimate for card $\{n \leqslant x : \omega(n) = m\}$ leads easily to a good upper bound for $S(x, y; P, \omega)$ (essentially the same as the bound given in Theorem 1.14 below). However, they did not state explicitly a result of the latter type. For arbitrary E, much weaker upper bounds for $S(x, y; E, \omega)$ and $S(x, y; E, \Omega)$ can be derived from a general theorem of Turán [19] on the distribution of values of additive functions. (See also Turán [18] or Hardy and Wright [6, pp. 356-358] for the case $E = P$, and see [13, §§1, 3] and [14, pp. 18-19] for remarks on all of this early work.) For the particular functions $\omega(n; E)$ and $\Omega(n; E)$, Turán's bounds were improved considerably in the author's paper [13; (5.16), (5.15), (1.11)], where it was observed that for any set E,

$$S(x, \alpha E(x); E, \omega) \leqslant x \exp\{(\alpha - 1 - \alpha \log \alpha) E(x)\} \tag{1.9}$$

for real $x \geqslant 1$, $\alpha \geqslant 1$, where $E(x)$ is defined by (1.2). A similar (slightly less precise) result was stated for $\Omega(n; E)$ when $1 \leqslant \alpha < p_1$, where p_1 is the smallest member of E. No lower bound was obtained in either case for $\alpha \geqslant 2$, so that the precision of (1.9) for large α was not clear. In a later paper [2], Erdös and Nicolas obtained a rather good estimate in the special case $E = P$. They showed that for any fixed α with $0 < \alpha < 1$,

$$\operatorname{card}\{n \leqslant x : \omega(n) > \alpha(\log x)(\log_2 x)^{-1}\} = x^{1-\alpha+o(1)} \tag{1.10}$$

as $x \to +\infty$. (In fact, they obtained a somewhat more precise result resembling Theorem 4.13 below.) However, they did not get an analogous result for $\Omega(n)$, nor did they generalize to $\omega(n; E)$ or $\Omega(n; E)$. Furthermore, their method did not give good upper estimates for $S(x, y; P, \omega)$ when y is appreciably smaller than $(\log x)(\log_2 x)^{-1}$. We propose to remedy all of these drawbacks to some extent. First, we obtain the following lower bound by a refinement of the Erdös-Nicolas method:

THEOREM 1.11. *Suppose that there exists a real number* $\gamma(E) > 0$ *such that* (1.7) *holds. Let* $\varepsilon > 0$, *and suppose that* $x \geqslant c_1(E, \varepsilon)$ *and*

$$c_2(E) \leqslant y \leqslant (\log x)(\log_2 x)^{-1} + \{1 + \log \gamma(E) - \varepsilon\}(\log x)(\log_2 x)^{-2}. \tag{1.12}$$

Then

$$S(x, y; E, \omega) \geqslant x \exp\{-y(\log y + \log_2 y - \log \gamma(E) - 1) + O_E(y(\log_2 y)/\log y)\}. \tag{1.13}$$

(1.8) shows that only a very small weakening of the hypothesis (1.12) would be of any interest. In Theorem 3.20, we assume much less than (1.7) and derive a result similar to Theorem 1.11 (but somewhat weaker).

Concerning upper bounds for $S(x, y; E, \omega)$, we have obtained only a modest improvement of (1.9); see Theorem 4.8 and Corollary 4.12. It should be emphasized that (1.9) and Theorem 4.8 hold for an arbitrary set E (without the assumption (1.7)). Using the same methods, we deduce

THEOREM 1.14. *Suppose that there exists a real number* $\gamma(E) > 0$ *such that* (1.7) *holds. If* $x \geqslant 3$ *and* $y \geqslant \gamma(E) \log_2 x$, *then*

$$S(x, y; E, \omega) \leqslant x \exp\{-y(\log y - \log_3 x - \log \gamma(E) - 1) - \gamma(E) \log_2 x + O_E(y/\log_2 x)\}. \tag{1.15}$$

Although there is a considerable gap between (1.13) and (1.15), the results are more general and somewhat sharper than those of Erdös and Nicolas [2]. In particular, we get a generalization of (1.10) (see Theorem 4.13). Theorems 1.11 and 1.14 also yield immediately the following result which could not be obtained by the Erdös-Nicolas method:

COROLLARY 1.16. *Suppose that there exists a real number* $\gamma(E) > 0$ *such that* (1.7) *holds. If* $0 < \alpha < 1$ *and* $x \geqslant c_3(E, \alpha)$, *then*

$$S(x, (\log x)^\alpha; E, \omega) = x \exp\{-\alpha(\log x)^\alpha \log_2 x + O((\log x)^\alpha \log_3 x)\}.$$

It should be mentioned that when $E = P$ (the set of all primes) and $y/\log_2 x$ is bounded and not too close to 1, Theorems 1.11 and 1.14 can be replaced by a striking asymptotic formula which was recently obtained by H. Delange (for the proof, see [2]):

THEOREM 1.17 (Delange). *Let* x, α, r_1, r_2 *be real with* $x \geqslant 3$, $1 < r_1 \leqslant \alpha \leqslant r_2$. *Then*

$$S(x, \alpha \log_2 x; P, \omega) = \frac{F(\alpha)\, \alpha^{1/2 + \alpha \log_2 x - [\alpha \log_2 x]}}{(2\pi)^{1/2}(\alpha - 1)} \cdot \frac{x}{(\log x)^{1 - \alpha + \alpha \log \alpha}(\log_2 x)^{1/2}} \left\{1 + O_{r_1, r_2}\left(\frac{1}{\log_2 x}\right)\right\},$$

where $[z]$ *means the largest integer* $\leqslant z$ *and*

$$F(\alpha) = \frac{1}{\Gamma(\alpha+1)} \prod_p \left(1 + \frac{\alpha}{p-1}\right)\left(1 - \frac{1}{p}\right)^{\alpha}.$$

Delange obtained a similar result for card $\{n \leqslant x : \omega(n) \leqslant \alpha \log_2 x\}$ when $x \geqslant 3$, $(\log_2 x)^{-1} \leqslant \alpha \leqslant r_3 < 1$ (see [2]). In this connection, it is interesting to note the estimate

$$F(\alpha) = \exp\{-\alpha \log \alpha - \alpha \log_2 \alpha + (1-\gamma)\alpha + O(\alpha/\log \alpha)\}$$

for real $\alpha \geqslant 2$, where γ is Euler's constant. (Some effort is required to show this, and we omit the proof.)

For values of α near 1, Kubilius [8, Theorem 9.2] proved a result on the distribution of $\omega(n)$ which is similar to Theorem 1.17. His theorem was later extended by himself [9] and Laurinčikas [10] to somewhat more general additive functions, and it was generalized to $\omega(n; E)$ and $\Omega(n; E)$ by Norton [14]. The estimates for $S(x, y; E, \omega)$ derived in the present paper are not as precise as Theorem 1.17 or the earlier work cited, but they are more general with respect to E (except for [14]), and they hold for much larger values of y.

We now consider the function $\Omega(n; E)$. Here we assume that E is any nonempty set of primes; in particular, nothing like (1.7) is assumed. For completeness, we begin by stating the following easy result:

THEOREM 1.18. *Let* p_1 *be the smallest member of* E. *Then*

$$\Omega(n; E) \leqslant (\log n)(\log p_1)^{-1} \qquad \text{for all} \qquad n \geqslant 1, \tag{1.19}$$

with equality if and only if $n = p_1^a$ *for some integer* $a \geqslant 0$.

This follows from

$$n \geqslant \prod_{p^a \| n,\ p \in E} p^a \geqslant \prod_{p^a \| n,\ p \in E} p_1^a = p_1^{\Omega(n;E)}.$$

We now proceed to estimate $S(x, y; E, \Omega)$ (defined by (1.1)). For $y \geqslant 2E(x)$, rather little previous work has been done on this problem, and all of it was restricted to the special case $E = P$ (the set of all primes). Selberg [17, p. 87] stated without detailed proof the following asymptotic formula:

$$\text{card}\,\{n \leqslant x : \Omega(n) = m\} \sim A 2^{-m} x \log x$$

for integers m satisfying $(2+\varepsilon)\log_2 x \leqslant m \leqslant B \log_2 x$. (Here $\varepsilon > 0$ is arbitrarily small, while A and B are positive absolute constants; it is not clear from [17] how large B could be.) Selberg also gave an asymptotic formula for card $\{n \leqslant x : \omega(n) = m\}$ when $m/\log_2 x$ is bounded. His work was recently extended to considerably larger values of m (roughly $m < (\log x)^{3/5}$) by Kolesnik and Straus [7], whose theorems are quite complicated. These results, together with the formula

$$S(x, y; P, \Omega) = \sum_{y < m \leqslant Y} \text{card}\,\{n \leqslant x : \Omega(n) = m\} + S(x, Y; P, \Omega)$$

and different tools for estimating $S(x, Y; P, \Omega)$ from above, would yield some information about $S(x, y; P, \Omega)$. However, it appears that neither [17] nor [7] would thus lead to an estimate for $S(x, y; P, \Omega)$ which is both simple and reasonably precise when $y/\log_2 x$ is unbounded. To the best of our knowledge, the only previous result of the latter type is due to Erdös and Sárközy [3], who recently proved that

$$S(x, y; P, \Omega) \leqslant c_4\, y^4\, 2^{-y}\, x \log x \qquad \text{for} \qquad x \geqslant 3,\, y \geqslant 1. \tag{1.20}$$

We shall generalize their work to $S(x, y; E, \Omega)$ and get a sharper upper bound. Although the result could be phrased in terms of the function $E(x)$ (defined by (1.2)), it is more convenient to state it in terms of a real number v which in practice is taken to be an approximation to $E(x)$. (For example, if $E = P$, we could take $v = \log_2 x$.)

THEOREM 1.21. *Let x, v, y be real with $x \geqslant 1, v \geqslant 1$, and $y \geqslant 0$. Let p_1 be the smallest member of E, and define*

$$\Lambda = \Lambda(x, v; E) = \max\{2, |E(x) - v|\}. \tag{1.22}$$

Then

$$S(x, y; E, \Omega) \leqslant c_5(p_1)\, p_1^{-y}\, x v^{1/2}\, e^{(p_1 - 1)v + p_1 \Lambda}. \tag{1.23}$$

We remark that (1.23) is our best upper bound when $y > p_1 v - v^{1/2}$, but it can be improved for smaller values of y (see Lemma 5.3).

Concerning the problem of estimating $S(x, y; E, \Omega)$ from below, we shall state only the following simple result:

THEOREM 1.24. *Let p_1 be the smallest member of E. If $x \geqslant p_1$ and $0 \leqslant y \leqslant (\log x)(\log p_1)^{-1} - 1$, then*

$$S(x, y; E, \Omega) \geqslant (1/2)\, p_1^{-y-1}\, x.$$

To prove this, let $k = [y] + 1$ (so k is the smallest integer greater than y), and observe that the multiples n of p_1^k have the property that $\Omega(n; E) \geqslant k > y$.

There are just $[xp_1^{-k}]$ of these $n \leqslant x$, and since $[z] \geqslant z/2$ for $z \geqslant 1$, we get the result.

It is clear that Theorem 1.24 is essentially best possible in certain extreme cases (for example, if $E = \{p_1\}$, or if $x = p_1^a$ and $y = a - 1$).

When $E = P$ (the set of all primes), we can take $v = \log_2 x$. Then $\Lambda = O(1)$, and we have the following corollary of Theorems 1.21 and 1.24:

COROLLARY 1.25. *If* $x \geqslant e^e$ *and* $0 \leqslant y \leqslant (\log x)(\log 2)^{-1} - 1$, *then*

$$2^{-y-2}\,x \leqslant S(x, y; P, \Omega) \leqslant c_6\, 2^{-y}\, x (\log x)(\log_2 x)^{1/2}.$$

Corollary 1.25 should be compared with the Erdös-Sárközy result (1.20) and with the asymptotic formula of Selberg mentioned after Theorem 1.18. When $y < 2\log_2 x$ (roughly), more precise estimates for $S(x, y; P, \Omega)$ can be obtained from [13] and [14].

In a later paper, we shall show that if p_1 is the smallest member of E and $\varepsilon > 0$ is fixed, then the precise order of magnitude of $S(x, y; E, \Omega)$ is

$$p_1^{-y}\, x \exp\{(p_1 - 1) E(x)\}$$

when $E(x)$ is sufficiently large and

$$p_1 E(x) \leqslant y \leqslant (1-\varepsilon)(\log x)(\log p_1)^{-1}.$$

This theorem is much more difficult to prove than Theorem 1.21. Its proof depends on Theorem 1.21 and on an extension of Halász's work [4] concerning the local distribution of $\Omega(n; E)$. Theorem 1.21 remains our best upper bound when y is close to $(\log x)(\log p_1)^{-1}$ (cf. Theorem 1.18), and it seems to be the most we can achieve by a fairly simple method.

§2. NOTATION

The symbols a, m, n always represent integers with $a \geqslant 0, m \geqslant 0, n > 0$. The letter p always denotes a prime, while $v, w, x, y, z, \alpha, \beta, \delta, \varepsilon, \sigma$ are real numbers. $[x]$ means the largest integer $\leqslant x$. The notation $\log_r x$ is defined by (1.5), and the notations $O, O_{\delta,\varepsilon,\ldots}, c_i, c_i(\delta, \varepsilon, \ldots)$ are explained after Theorem 1.6. If a condition such as "$x \geqslant c_i(\delta, \varepsilon, \ldots)$" is used as a hypothesis, it is to be understood that $c_i(\delta, \varepsilon, \ldots)$ is sufficiently large. We shall occasionally use the notations $\ll$, $\gg$ to imply constants which are *absolute*. (Thus $A = O(B)$ is equivalent to $A \ll B$.)