ON SOME PROPERTIES OF PRIME
FACTORS OF INTEGERS

P. ERDÖS

Dedirated to the memory of TADASI NAKAYAMA

Let $n = \prod_{i=1}^{\nu(n)} p_i^{\alpha_i}$. A well known theorem of Hardy and Ramanujan states: $\nu(n) = (1 + o(1))\log\log n$ holds for all n if we neglect a sequence of density 0 [5]. Define for $2 \leq j \leq \nu(n)$

$$\prod_{i=1}^{j-1} p_i^{\alpha_i} = p_j^{\gamma_j(n)}.$$

I had often occasion to use the fact that $\gamma_j(n)$ is "usually" $O(1)$ [2]. Put

$$\max_{2 \leq j \leq \nu(n)} \gamma_j(n) = P(n).$$

In the present note we shall prove the following:

THEOREM 1. *For almost all integers* n

(1) $$P(n) = (1 + o(1)) \log_3 n/\log_4 n.$$

The phrase "almost all integers" means that (1) holds for all n if we neglect a sequence of integers of density 0, $\log_k n$ denotes the k-fold iterated logarithm.

We will also outline the proof of the following further results:

THEOREM 2. *There is a continuous strictly increasing function* $\varphi(c)$, $\varphi(0) = 0$, $\varphi(\infty) = 1$, *so that for almost all integers* n

$$\frac{1}{\log_2 n} \sum_{\gamma_j(n) \leq c} 1 \to \varphi(c).$$

In other words there are $(\varphi(c) + o(1))\log_2 n$ values of j for which $\gamma_j(n) \leq c$.

THEOREM 3. *The density of integers for which*

$$\min \gamma_j(n) < c/\log_2 n$$

617

is $\psi(c)$ where $\psi(0) = 0$, $\psi(\infty) = 1$ and $\psi(c)$ is continuous and strictly increasing.

Put $p_j^{\eta_j} = p_{j+1}$. In a previous paper [3] I proved then the density of integers n for which

$$(2) \qquad \min_{1 \le j \le \nu(n)-1} \eta_j < 1 + c/\log_2 n$$

equals $1 - \exp(-c)$ $(\exp z = e^z)$, also the density of integers n for which

$$(3) \qquad \max_{1 \le j \le \nu(n)-1} \eta_j > c \log_2 n$$

is $1 - \exp(-1/c)$. I further proved [4] that for almost all n

$$(4) \qquad \sum_{j=1}^{\nu(n)-1} \eta_j = (1 + o(1)) \log_2 n \log_3 n.$$

It would be easy to deduce from Theorem 2 and from the result of de Bruijn [1] that for almost all n

$$(5) \qquad \sum_{j=2}^{\nu(n)-1} r_j(n) = (1 + o(1)) \log_2 n \int_0^\infty \varphi(c). \quad \left(\int_0^\infty \varphi(c) < \infty \right)$$

Now we prove Theorem 1. To prove our Theorem we have to show that for every $\varepsilon > 0$ the density of integers for which

$$(6) \qquad P(n) > (1 + \varepsilon) \log_3 n / \log_4 n$$

is 0 and the density of integers for which

$$(7) \qquad P(n) < (1 - \varepsilon) \log_3 n / \log_4 n$$

is also 0. First we prove (6). Because of the slow growth of $\log_3 n / \log_4 n$ it clearly will suffice to show that the number of integers $n \le x$ for which

$$(8) \qquad P(n) > \left(1 + \frac{\varepsilon}{2}\right) \log_3 x / \log_4 x$$

is $o(x)$.

First of all we observe that the number of integers $n \le x$ which are divisible by a square $k^2 > c$ is less than

$$(9) \qquad \sum_{k^2 > c} \left[\frac{x}{k^2} \right] < \frac{2x}{c^{1/2}}.$$

Hence by (9) we obtain by a simple argument that (8) will follow if we show that for every $\varepsilon > 0$ for all but $o(x)$ integers $n \le x$ we have for every $2 \le j \le \nu(n)$

(10) $$\prod_{i<j} p_i < \exp\left(\left(1 + \frac{\varepsilon}{4}\right)(\log_3 x/\log_4 x)\log p_j\right).$$

Put $T = \left(1 + \frac{\varepsilon}{8}\right)\log_3 x/\log_4 x$. (10) will easily follow from

LEMMA 1. *Let $k > k_0(\varepsilon)$ be sufficiently large. Then the number of integers $n \leq x$ which for some $l \geq 0$ have more than T prime factors p satisfying*

(11) $$2^{k^l} < p < 2^{k^{l+1}}$$

is $o(x)$.

Assume that the Lemma has already been proved. Let n be an integer which has for every $l \geq 0$ not more than T distinct prime factors satisfying (11). For these integers we clearly have for every $2 \leq j \leq \nu(n)$

$$\prod_{i<j} p_i < \prod_{l=0}^{\infty}(p_j^{1/k^l})^T = p_j^{(1+1/k-1)T} < \exp\left(\left(1 + \frac{\varepsilon}{4}\right)\log p_j \log_3 x/\log_4 x\right)$$

for $k > k_0(\varepsilon)$, hence (10) is proved.

Thus to prove (10) we only have to prove Lemma 1. By the well known theorem of Mertens $\sum_{p<y} 1/p = \log\log y + c + 0(1/\log y)$ we have (in $\sum_l 2^{k^l} < p < 2^{k^{l+1}}$)

(12) $$\sum_l 1/p < c_1 \log k.$$

The number of integers $n \leq x$ which for a given l have more than T distinct prime factors satisfying (11) is by (12) clearly less than

(13) $$x\left(\sum_l 1/p\right)^T/T! < x(c_1 \log k)^T/T! <$$
$$x\left(\frac{ec_1 \log k}{T}\right)^T < \frac{x}{(\log_2 x)^{1+\varepsilon/10}}.$$

for $x > x_0(\varepsilon)$. Since $2^{k^l} \leq x$ we have at most $\log_2 x$ choices for l, thus the number of integers $n \leq x$ which for some l have more than T distinct prime factors satisfying (11) is by (13) less than $x/(\log_2 x)^{\varepsilon/10} = o(x)$, which proves Lemma 1 and hence (10), (8) and (6) are proved.

By the same method we can prove that for every $\varepsilon > 0$ and $\eta > 0$ there is an $l = l(\varepsilon, \eta)$ so that the density of integers n for which

$$r_j(n) > (1 + \varepsilon)\log_3 p_j/\log_4 p_j$$

holds for more than l values of j is less than η.

The proof can be easily deduced from the fact that for every k the series
$(A(p) = [(1 + \varepsilon) \log_3 p / \log_4 p])$

$$(14) \qquad \sum_p \frac{1}{p} \left(\left(\sum_{p^{1/k} < q < p} 1/q \right)^{A(p)} / A(p)! \right)$$

converges. The proof of (14) is similar to that of (13).

To complete the proof of Theorem 1 we now prove (7). Instead of (7) we will prove that for all but $o(x)$ integers $n < x$

$$(15) \qquad\qquad P(n) > (1 - \varepsilon) \log_3 x / \log_4 x.$$

Let r run through the integers of the interval $\left(\frac{1}{2} \log_2 x, \log_2 x \right)$ and denote by I_r the interval

$$\exp\left(\left(1 + \frac{1}{\log_4 x} \right)^r \right), \quad \exp\left(\left(1 + \frac{1}{\log_4 x} \right)^{r+1} \right).$$

To prove (15) it will suffice to prove the following:

LEMMA 2. *For all but $o(x)$ integers $n \leq x$ every n has at least*

$$\left[\left(1 - \frac{\varepsilon}{2} \right) \log_3 x / \log_4 x \right] = T_1$$

distinct prime factors in some I_r.

(15) immediately follows from Lemma 2. Let n have at least T_1 distinct prime factors in I_r and let p_j the greatest p/n in I_r. Then clearly for sufficiently large x

$$\prod_{i < j} p_i > p_j^{(T_1 - 1)(1 - 1/\log_4 x)} > p_j^{(1 - \varepsilon) \log_3 x / \log_4 x}$$

which proves (15).

Thus to prove (15) (and hence to complete the proof of Theorem 1) we only have to prove Lemma 2. First we need three further Lemmas. Denote by $a_i^{(r)}$, $i = 1, \ldots$ the integers which are the product of T_1 distinct prime factors of I_r. We have

LEMMA 3. *Put* $(\log(1 + 1/\log_4 x))^{T_1} / T_1! = F(x)$. *We then have*

$$(16) \qquad\qquad \sum_i 1/a_i^{(r)} = (1 + o(1)) F(x).$$

By the theorem of Mertens we have

(17) $$\sum_{p \text{ in } I_r} 1/p = \log\left(1 + 1/\log_4 x\right) + o\left(\frac{1}{\log_2 x}\right).$$

Denote by $b_j^{(s)}$, $j = 1, \ldots$ the integers composed of s distinct primes in I_r (for $s = T_1$ the $b_j^{(s)}$ are the $a_i^{(r)}$). Clearly

(18) $$\sum_j \frac{1}{b_j^{(s+1)}} = \frac{1}{s+1} \sum_j \frac{1}{b_j^{(s)}} \sum{}' \frac{1}{p}$$

where the dash indicates that the summation is extended over the p in I_r for which $p \nmid b_j^{(s)}$. Clearly by (17) we have

(19) $$\sum{}' \frac{1}{p} = \log(1 + 1/\log_4 x) + o(1/\log_2 x).$$

Lemma 3 follows from (18) and (19) by a simple computation.

LEMMA 4. *Denote by $A_r(x)$ the number of integers which are divisible by at least one $a_i^{(r)}$. We have*

$$A_r(x) = (1 + o(1))xF(x).$$

Clearly by Lemma 3

(20) $$A_r(x) \leq \sum_i \left[\frac{x}{a_i^{(r)}}\right] \leq x \sum_i \frac{1}{a_i^{(r)}} = (1 + o(1))xF(x).$$

Denote on the other hand by $B_i(x)$ the number of integers $n \leq x$ which are multiples of $a_i^{(r)}$ but of no other $a_j^{(r)}$. Clearly by (17)

(21) $$B_i(x) \geq \left[\frac{x}{a_i^{(r)}}\right] - \sum_{p \text{ in } I_r} \left[\frac{x}{p a_i^{(r)}}\right] = (1 + o(1)) \frac{x}{a_i^r}$$

Further clearly

(22) $$A_r(x) \geq \sum_i B_i(x).$$

From (21), (22) and Lemma 3 we have

(23) $$A_r(x) \geq (1 + o(1))x \sum_i \frac{1}{a_i^{(r)}} = (1 + o(1))xF(x)$$

(20) and (23) proves Lemma 4.

Denote by $A_{r_1, r_2}(x)$ $(r_1 \neq r_2)$ the number of integers $n \leq x$ which are divisible by at least one $a_i^{(r_1)}$ and at least one $a_i^{(r_2)}$

LEMMA 5.

$$A_{r_1, r_2}(x) = (1 + o(1)) \, xF(x)^2.$$

The proof of Lemma 5 is the same as Lemma 4, (we use $(a_i^{(r_1)}, a_j^{(r_2)}) = 1$ and $a_i^{(r_1)} a_j^{(r_2)} = o(x)$) and can be left to the reader.

Now we are ready to prove Lemma 2. Denote by $f(n)$ the number of r's for which n is divisible by an $a_i^{(r)}$. We have to show that for all but $o(x)$ integers $n \leq x$, $f(n) > 0$. In fact we shall prove more. We shall show that for all but $o(x)$ integers $n \leq x \left(\text{put } \frac{1}{2} (\log \log x). \, F(x) = \mathfrak{G}(x) \right)$

$$(24) \qquad f(n) = (1 + o(1)) \frac{1}{2} (\log \log x). \, F(x) = (1 + o(1)) \mathfrak{G}(x)$$

(24) implies by a simple computation that for almost all n, $f(n) \to \infty$. We prove (24) by Turán's method [6].

We evidently have

$$(25) \qquad \sum_{n=1}^{x} (f(n) - \mathfrak{G}(x))^2 = \sum_{n=1}^{x} f(n)^2 - 2 \, \mathfrak{G}(x) \sum_{n=1}^{x} f(n) + x \mathfrak{G}^2(x).$$

Now clearly by Lemma 4 and the definition of $\mathfrak{G}(x)$

$$(26) \qquad \sum_{n=1}^{x} f(n) = \sum_{r} A_r(x) = (1 + o(1)) x \mathfrak{G}(x).$$

Further by a simple argument we have from Lemma 5

$$(27) \qquad \sum_{n=1}^{x} f^2(n) = 2 \sum_{r_1 < r_2} A_{r_1, r_2}(x) + \sum_{r} A_r(x) = (1 + o(1)) x \mathfrak{G}^2(x).$$

Thus from (25), (26) and (27)

$$(28) \qquad \sum_{n=1}^{x} (f(n) - \mathfrak{G}(x))^2 = o(x(\mathfrak{G}(x))^2),$$

(28) immediately implies (24) (using Tchebicheff's inequality). This completes the proof of Lemma 2 and Theorem 1.

By somewhat more trouble we could prove the following sharpening of Lemma 2: Let $C_1, \ldots, C_s$, $s = o(x)$ be classes of integers. Assume that if a and b belong to different classes then $(a, b) = 1$. Denote by $g_i(x)$ the number of integers $n \leq x$ which are divisible by at least one integer of C_i and assume that

$$\lim_{x = \infty} \frac{1}{x} \sum_{i=1}^{s} g_i(x) = \infty.$$

Then all but $o(x)$ integers $n \leq x$ have a divisor from at least one C_i, $1 \leq i \leq s$.

The proof is more difficult than that of Lemma 2, and I have to use Brun's method. The difficulty is that the analog of Lemma 5 breaks down.

We only outline the proof of Theorem 2.

LEMMA 6. *There is a continuous strictly increasing function $\varphi(c)$, $\varphi(0) = 0$, $\varphi(\infty) = 1$ so that to every $\varepsilon > 0$ there is a j_0 for which for every fixed $j > j_0$ the density of integers n with $r_j(n) > c$ differs from $\varphi(c)$ by ε.*

The proof of Lemma 6 can easily be deduced from the results of N. G. de Bruijn [1] and is not difficult.

Theorem 2 follows from Lemma 6 by the methods of probabilistic number theory but the proof is not quite simple, we have first to show that if $j' - j$ is large then the values of $r_j(n)$ and $r_{j'}(n)$ are nearly independent and then Turán's method [6] can be applied without much difficulty.

Theorem 3 can be proved similarly as (3) but the proof is more complicated.

By using the results of de Bruijn one could sharpen Theorem 1 and one could perhaps obtain an asymptotic expansion for $P(n)$ valid for almost all integers, but I have not even determined the second term of this hypothetical asymptotic expansion.

REFERENCES

[1] N. G. de Bruijn, On the number of positive integers$\leq x$ and free of prime factors$>y$. Indigationes Math., **13** (1945), 50-60.

[2] See, E. G. P. Erdös, On a problem of Chowla and some related problems, Proc. Cambridge Phil. Soc., **32** (1936), p. 532.

[3] P. Erdös, Some remarks on prime factors of integers, Canadian J. Math., **11** (1959), 161-167.

[4] P. Erdös, Some remarks about additive and multiplicative functions, Bull. Amer. Math. Soc., **52** (1946), p. 535 (Theorem 10).

[5] G. H. Hardy and S. Ramanujan, The normal number of prime factors of n, Quart. J. Math., **48** (1917), 76-92.

[6] P. Turán, On a theorem of Hardy and Ramanujan, J. London Math. Soc., **9** (1934), 274-276.

University College, London