Reprint from
aequationes mathematicae
Vol. 2, fasc. 2/3, 1969 BIRKHÄUSER VERLAG BASEL pages 177–183

On the Distribution of Prime Divisors

P. ERDŐS (Budapest, Hungary)
Dedicated to the 75th birthday of Professor A. OSTROWSKI

Denote by $V(n)$ the number of distinct prime factors of n. A well known theorem of HARDY and RAMANUJAN [8] states that for almost all n $V(n)=(1+o(1))\log_2 n$ and a special case of a result of KAC and myself [3] states that the density of integers n satisfying

$$V(n) > \log_2 n + c(\log_2 n)^{1/2}$$

is $1/\sqrt{2\pi}\int_c^\infty c^{-x^2/2}dx$ (almost all n means for all neglecting a sequence of density 0, $\log_k n$ denotes the k-fold iterated logarithm).

Denote by $v(n; u, v)$ the number of prime factors p of n satisfying $u<p<v$. Let $u=u(x)$, $v=v(x)$, and assume that $\log_2 v-\log_2 u\to\infty$, TURÁN [11] proved that for all but $o(x)$ integers $n<x$ $v(n; u, v)=(1+o(1))(\log_2 v-\log_2 u)$. We now investigate the case when u and v depend on n. If the dependence is regular Turán's method carries through without too much difficulty. In the general case, somewhat unexpectedly, $\log_2 v-\log_2 u\to\infty$ is not sufficient for $v(n; u, v)=(1+o(1))(\log_2 v-\log_2 u)$, but in fact we show that this holds uniformly in u and v under rather mild conditions.

In fact we prove

THEOREM 1. *Assume* $(\log_2 v-\log_2 u)/\log_3 n\to\infty$. *Then we have for almost all n uniformly in u and v*

$$V(n; u, v) = (1 + o(1))(\log_2 v - \log_2 u).$$

Theorem 1 is best possible. In fact we have

THEOREM 2. *There are two continuous functions $f_1(c)$ and $f_2(c)$, $f_1(0)=\infty$, $f_1(\infty)=1$, $f_2(c)$ is strictly decreasing for $0<c<\infty$; $f_2(c)=0$ for $0\leqslant c\leqslant 1$, $f_2(\infty)=1$, $f_2(c)$ is strictly increasing in $1<c<\infty$, satisfying for almost all n and for every $c>0$*

$$\max V(n; u, v) = (1 + o(1))f_1(c)(\log_2 v - \log_2 u)$$

and

$$\min V(n; u, v) = (1 + o(1))f_2(c)(\log_2 v - \log_2 u)$$

where the max *and* min *is taken with n fixed over the values $1\leqslant u<v\leqslant n$ satisfying*

$$\log_2 v - \log_2 u > c\log_3 n$$

We will prove Theorem 1 in full detail, the proof of Theorem 2 is similar but more complicated and will be omitted (see [4]).

Theorems 1 and 2 can be generalised for a large class of additive functions but

Received December 1, 1967.

we will not discuss this here. I only formulate the probabilistic theorem which corresponds to theorems 1 and 2 in case of Rademacher functions $r_k(x)$.

THEOREM 3. *For all x neglecting a set of measure* 0 *we have*

$$\lim \frac{1}{v-u} \sum_{k=u}^{v} r_k(x) = 0$$

uniformly in u and v if $(v-u)/\log u \to \infty$.

Further

$$\lim \frac{1}{v-u} \sum_{k=u}^{v} r_k(x) = f_1(c)$$

where $u \to \infty$ and $v > u + c \log u$. $f(c) = 1$ for $0 \leqslant c \leqslant 1$, $\mathrm{f}(\infty) = 0$.

$f(c)$ is strictly decreasing in $1 < c < \infty$.

The proof of Theorem 3 follows from simple independence arguments and will not be given here. Theorem 3 could be generalised for other independent functions but I have not investigated how far this generalisation will go.

Let $p_1 < \cdots < p_{v(n)}$ be the distinct prime factors of n. In a previous paper [5] I stated (it can be proved by the methods of probabilistic number theory [5]), that roughly speaking the order of magnitude of the i-th prime factor of n is $\exp \exp i\,(\exp z = c^z)$. More precisely for every $\varepsilon > 0$ and $\eta > 0$ there is an i_0 so that for all but εx integers $n \leqslant x$ we have for all $i_0 < i \leqslant v(n)$

$$(1-\eta)\, i < \log\log p_i < (1+\eta)\, i\,.$$

In fact in [5] a sharper result is stated.

By the same method I can prove that for every $\varepsilon > 0$ and $\eta > 0$ there is an i_0 so that for all but εx integers $n \leqslant x$ we have for all $i_0 < i \leqslant v(n)$

$$(1-\eta)\, i < \log\log n - \log\log p_i < (1+\eta)\, i\,.$$

I now state without proof a few results about prime factors of integers which can be obtained by standard methods of probabilistic number theory (see [9]).

We have for almost all integers n

$$\sum_{\substack{p_i \mid n \\ p_i > \exp\exp i}} \frac{1}{i} = \left(\tfrac{1}{2} + o(1)\right) \log_3 n\,. \tag{1}$$

In fact more generally we have for almost all n

$$\sum{}' \frac{1}{i} = (1 + o(1)) \int_c^\infty e^{-x^2/2}\, dx \log_3 n \tag{2}$$

where the dash indicates that the summation is extended over the $p_i|n, p_i>$ exp exp $(i+c\, i^{1/2})$.

(1) and (2) follows from a result of CHUNG and myself [2] (which is a generalisation of a result of PAUL LEVY [10]) and BRUNS method (similarly as in [3]).

On the other hand it is not true that for almost all n

$$\sum_{\substack{p_i|n \\ p_i > \exp\exp i}} 1 = \left(\tfrac{1}{2} + o(1)\right) \log_2 n\,. \tag{3}$$

Instead of (3) the following result holds: Let $j_1 < \cdots$ be a sequence of integers. Put $\sum_{j_r<y} 1 = A(y)$ and assume $A(y) \to \infty$, $A(2y)/A(y) \to 1$ as $y \to \infty$. Then for almost all integers

$$\sum_{\substack{p_{j_r}|n \\ p_{j_r} > \exp\exp j_r}} 1 = \left(\tfrac{1}{2} + o(1)\right) A(\log_2 n)$$

From the arc sine law [6] and BRUNS method we obtain the following result. The density of integers n for which

$$\frac{1}{\log_2 n} \sum_{\substack{p_i|n \\ p_i < \exp\exp i}} 1 < \alpha$$

holds, equals $2/\pi$ arc sin $\alpha^{1/2}$. This shows that (3) is not true.

Denote by $f(\alpha)$ the density of integers the largest prime factor of which is $< n^\alpha$. It is easy to see that $f(\alpha)$ exists and is a continuous strictly increasing function of α, $f(0)=0, f(1)=1$. DICKMAN, DE BRUIJN, BUCHSTAB [1] and others obtained more or less explicit formulas for $f(\alpha)$. Denote by $f_i(\alpha)$ the density of the integers n for which $p_{v(n)-i} < n^\alpha$. It is easy to see by methods similar to those used for $f(\alpha)=f_0(\alpha)$ that $f_i(\alpha)$ is a strictly increasing continuous function of α, $f_0(\alpha)\, f_i(0)=0, f_i(1/i+1)=1$. As far as I know $f_i(\alpha)$ has never been computed explicitely. It follows by the methods of probabilistic number theory [3] that the density of integers n for which

$$p_i > \exp\exp(i + c\, i^{1/2})$$

and

$$p_{V(n)-j} > \exp \log n / e^{j - c\, j^{1/2}}$$

approaches as $i \to \infty, j \to \infty$ $1/\sqrt{2\pi} \int_c^\infty e^{-x^2/2} dx$.

Denote by $\alpha(i, k)$ the density of the integers the i-th prime factor of which is p_k. Clearly $\alpha(i, k)$ exists for every i and k (because the sequence of numbers the i-th prime factor of which equals p_k can be obtained by set theoretic operations from a finite number of arithmetic progressions), and is positive for $k \geqslant i$. It might be of interest to determine max $\alpha(i, k)$, or at least to obtain an asymptotic formula for it. I only succeeded in obtaining here some rather crude results.

Now we prove Theorem 1. Because of the slow growth of the iterated logarithms it clearly will suffice to prove the following.

THEOREM 1'. *To every $\varepsilon>0$ there is an A so that for every $x>x_1(\varepsilon, A)$ the number of integers $n<x$ for which for every* $(u<v\leqslant x)$

$$\log_2 v - \log_2 u > A \log_3 x \tag{4}$$

we have

$$(1-\varepsilon)(\log_2 v - \log_2 u) < V(n; u, v) < (1+\varepsilon)(\log_2 v - \log_2 u)$$

is $x+o(x)$.

To prove Theorem 1' we only have to show that for $x>x_0(\varepsilon, A)$ the number of integers $n<x$ for which there are values $u<v\leqslant x$ satisfying (4) and for which

$$V(n; u, v) < (1-\varepsilon)(\log_2 v - \log_2 u) \tag{5}$$

or

$$V(n; u, v) > (1+\varepsilon)(\log_2 v - \log_2 u) \tag{6}$$

is $o(x)$.

Put $w_i = \exp\exp i$. First we prove

LEMMA 1. *The number of integers $n\leqslant x$ for which*

$$\max_i V(n; w_i, w_{i+1}) > \log_3 x \tag{7}$$

is $o(x)$.

The number of integers $n\leqslant x$ for which $V(n; w_i, w_{i+1})>\log_3 x$ holds is clearly at most $\sum' [x/a_j]$ where in $\sum'$ the summation is extended over the integers a_j which are the product of $t=[\log_3 x]$ distinct primes p, $w_i<p<w_{i+1}$. Now clearly by the well known theorem of Mertens $\sum_{p<y} 1/p = \log_2 y + c + o(1)$ we have uniformly in i

$$\sum{}' \left[\frac{x}{a_j}\right] < x\Bigg(\sum_{w_i<p<w_{i+1}} \frac{1}{p}\Bigg)^t \Bigg/ t = x\,\frac{1+o(1)}{t!} = o\left(\frac{x}{\log_2 x}\right). \tag{8}$$

In (7) there are only $\log_2 x$ choices of i, thus Lemma 1 follows immediately from (8).

From Lemma 1 we easily deduce that to prove Theorem 1' it suffices to consider in (5) and (6) only those u's and v's for which

$$u = w_i, \quad v = w_j, \quad 1 \leqslant i < j \leqslant \log_2 x, \quad j - i > A \log_3 x, \tag{9}$$

LEMMA 2. *For all but $o(x)$ integers $n\leqslant x$*

$$V(n; x^{1/\log_2 x}, x) = (1+o(1)) \log_3 x.$$

Lemma 2 follows immediately by the method of TURÁN [11].

From Lemma 2 it easily follows that in (9) we can assume $j \leqslant \log_2 x - \log_3 x$ or $w_j \leqslant \exp(\log x/\log_2 x)$.

Denote ($p^\alpha \| n$ means that $p^\alpha | n$ but $p^{\alpha+1} \nmid n$)

$$A(n) = \prod{}' p^\alpha$$

where the dash indicates $p^\alpha \| n$, $p \leqslant \exp(\log x/\log_2 x)$.

LEMMA 3. *For all but $o(x)$ integers $n \leqslant x$ we have $A(n) < x^{1/2}$.*

Lemma 3 is well known, but for the sake of completeness we give the simple proof. We evidently have by the theorem of MERTENS ($L = \exp(\log x/\log_2 x)$)

$$\prod_{n=1}^{x} A(n) < \prod_{p<L} p^{\sum_{k=1}^{\infty} x/p^k} = \exp\left(x \sum_{p<L} \frac{\log p}{p-1}\right) < \exp(2x \log x/\log_2 x). \tag{10}$$

From (10) we obtain that the number of integers $n \leqslant x$ for which $A(n) \geqslant x^{1/2}$ is less than $4x/\log_2 x$, which proves the Lemma.

LEMMA 4. *Let i and j satisfy* (9). *Then the number of integers $S(i, j)$ satisfying $A(n) < x^{1/2}$ which satisfy* (5) *or* (6) *is $o(x/(\log_2 x)^2)$ uniformly in i and j.*

Before we prove Lemma 4 we deduce Theorem 1′ from our four Lemmas. We already deduced from Lemmas 1 and 2 that to prove theorem 1′ it suffices to consider the values i and j satisfying (9) and from Lemma 3 we obtain that we only have to consider the $n \leqslant x$ with $A(n) < x^{1/2}$. There are fewer than $(\log_2 x)^2$ values of i and j satisfying (9), hence from Lemma 4 the number of integers which satisfy (5) and (6) for some i and j satisfying (9) is $o(x)$ as stated.

Thus to complete the proof of Theorem 1, we only have to prove Lemma 4. Let $i < j$ satisfy (9) and denote

$$A_{i,j}(n) = \prod p^\alpha, \quad \text{where} \quad p^\alpha \| n, \quad w_i < p < w_j.$$

LEMMA 5. *Let $t \leqslant x^{1/2}$. The number of integers $f_{i,j}(x, t)$ for which $n \leqslant x$ and $A_{i,j}(n) = t$ is less than $c_1 x/t \exp(j-i)$.*

$f_{i,j}(x, t)$ clearly equals the number of integers $m \leqslant x/t$ which have no prime factor p, satisfying $w_i < p < w_j$. It immediately follows from BRUNS method [7] (here we use $A(n) < x^{1/2}$)) and the well known theorem of MERTENS $\prod_{w_i<p<w_j} (1-1/p) < c_2/\exp(j-i)$ that $f_{i,j}(x, t)$ is less than

$$c_3 \frac{x}{t} \prod_{w_i<p<w_j} \left(1 - \frac{1}{p}\right) < c_1 x/t \exp(j-i),$$

which proves Lemma 5.

Denote now by $a_1 < \cdots$ the integers not exceeding $x^{1/2}$ which are composed entirely from the primes $w_i < p < w_j$ and for which

$$V(a_r) < (1-\varepsilon)(j-i) \tag{11}$$

or

$$V(a_r) > (1+\varepsilon)(j-i). \tag{12}$$

We evidently have

$$\sum \frac{1}{a_r} < \sum_{u<(1-\varepsilon)(j-1)} \left(\sum \frac{1}{P}\right)^u \Big/ u! + \sum_{u>(1+\varepsilon)(j-1)} \left(\sum \frac{1}{P}\right)^u \Big/ u! \tag{13}$$

where P runs through all the powers of the primes p, satisfying $w_i < p < w_j$. By the theorem of Mertens we obtain by a simple calculation (using $j-i > A \log_3 x$)

$$\left.\begin{aligned} \sum \frac{1}{a_r} &< \sum_{u<(1-\varepsilon)(j-i)} \frac{(j-i+c_4)^u}{u!} + \sum_{u>(1+\varepsilon)(j-i)} \frac{(j-i+c_4)^u}{u!} \\ &< \exp(j-i+c_4)(1-c_5\varepsilon)^{-c_6\varepsilon(j-i)} = o\left(\exp(j-i)/(\log_2 x)^2\right) \end{aligned}\right\} \tag{14}$$

for sufficiently large $A = A_0(\varepsilon)$. (14) states a well known property of the exponential series, namely in the expansion $e^z = \sum_{k=0}^{\infty} z^k/k!$, the main contribution comes from the terms $k = (1+o(1))z$.

We evidently have from Lemma 5 and (14)

$$S(i,j) = \sum_r f_{i,j}(x, a_r) < \sum_r \frac{c_1 x}{a_r \exp(j-i)} = o\left(\frac{x}{(\log_2 x)^2}\right),$$

which proves Lemma 4 and thus completes the proof of Theorems 1′ and 1.

REFERENCES

[1] DE BRUIJN, N.G., *On the number of positive integers $\leqslant x$ and free of prime factors and $\geqslant y$*, *Nederl. Akad. Wetensch. Indag. Math. 13*, 50–60 (1951).

[2] CHUNG, K.L., and ERDŐS, P., *Probability Limit Theorems Assuming Only the First Moment*, Mem. Amer. Math. Soc. No. 6, *1951*.

[3] ERDŐS, P. and KAC, M., *The Gaussian Law of Errors in the Theory of Additive Number Theoretic Functions*, Amer. J. Math. *62*, 738–742 (1940).

[4] ERDŐS, P., *Some Remarks on Prime Factors of Integers*, Canad. J. Math. *11*, 161–167 (1959), and *On some Properties of Prime Factors of Integers*, Nagoya Math. J. *27*, 617–623 (1966).

[5] ERDŐS, P., *On the distribution function of additive functions*, Ann. of Math. *47*, 1–20 (1946).

[6] See e.g. ERDŐS, P., KAC, M., *On the number of positive sums of independent random variables*, *Bull. Amer. Math. Soc. 53*, 1011–1020 (1947).

[7] See e.g. HALBERSTAM, H., ROTH, K.F., *Sequences*, Vol. I (Clarendon Press, Oxford 1966), Chapt. IV.

[8] HARDY, G.H. and RAMANUJAN, S., *The normal number of prime factors of n*, Quart. J. Math. *48*, 76–92 (1917).

[9] KUBYLYUS, I. P., *Probabilistic methods in number theory*, Amer. Math. Soc. Transl. [2] *19*, 47–85 (1962).
[10] LEVY, P., *Théorie de l'addition des variables aléatoires* (Gauthier-Villars, Paris 1937).
[11] TURÁN, P., *On a theorem of Hardy and Ramanujan*, J. London Math. Soc. *9*, 274–276 (1934).

Mathematical Research Institute
Hungarian Academy of Sciences